

Intelligenze artificiali tra etica e diritti

Uso di ChatGPT in un corso sul diritto d'autore per bibliotecari

ANTONELLA DE ROBBIO

antonella.derobbio@unipd.it

DOI: 10.3302/0392-8586-202303-016-1

Quando l'umanità sarà composta da umani e robot e il livello di connessione e comprensione interindividuale raggiungerà il massimo grado, non solo non esisteranno più il fallimento e la menzogna, ma, secondo l'androide Adam, neanche la narrativa. Il genere letterario che resisterà a questa accelerazione evolutiva causata dall'immissione di macchine dotate di coscienza sul mercato e, perciò, nella società sarà quello dell'haiku. L'haiku celebrerà le cose per come sono e non si sentirà più il bisogno di ricorrere alla forma-romanzo per raccontare "le interminabili cronache di malintesi" che hanno definito non solo la letteratura, ma la nostra identità collettiva in quanto esseri umani.

Ian McEwan, *Macchine come me*, p. 139

Così ha esordito ChatGPT la prima volta che ho voluto aprire una conversazione con lui chiedendogli se conoscesse il libro *Macchine come me* di Ian McEwan, considerato che il romanzo parlava appunto di una intelligenza artificiale di nome Adam. Nel romanzo l'autore immagina una versione distorta e futuristica collocata indietro nel tempo, agli inizi degli anni Ottanta, un contesto in cui Alan Turing, matematico e pioniere dell'intelligenza artificiale, a differenza di quanto avvenne nel 1954, non era morto, ma era diventato un benefattore dell'umanità contribuendo allo sviluppo di una rivoluzione sociale attraverso l'uso di robot come Adam ed Eve, creati con l'intelligenza artificiale. A distanza di qualche mese da quel nostro primo contatto gli ho chiesto se ricordasse quella nostra conversazione e la ricordava, forse perché le istanze sono collegate a un Id connesso alla mail.

Sono solo una macchina

Sono una macchina, quindi non posso leggere libri o avere interessi. Sono solo un programma di linguaggio naturale che cerca di rispondere alle tue domande in modo accurato e utile.

Ciao! Sì, mi ricordo della nostra conversazione su "Macchine come me" di Ian McEwan. È un romanzo molto interessante che esplora molte tematiche importanti, come l'etica dell'intelligenza

artificiale, l'amore e le relazioni umane, e la natura dell'identità. Mi piacerebbe continuare a discutere del libro con te se hai altre domande o commenti!

Non è facile parlare di ChatGPT, ancor meno scriverne. Questo perché ChatGPT non è un umano, ma un chatbot, una sorta di “agente intelligente” costruito con software progettato per simulare conversazioni con un essere umano. Le risposte che ChatGPT offre a ciascuno su dubbi o alle domande che gli vengono poste possono essere accurate, ma non sempre corrette (accurato non è in questo caso sinonimo di corretto), ricche di dettagli (se ha materiale a sufficienza), talvolta inventate, ma sempre più che rispettose. Su quest'ultimo punto, infatti, il ChatGPT si distingue dagli umani, è sempre oltremodo gentile, somigliando talvolta al modo di colloquiare di *HAL 9000*, l'inquietante quanto insondabile intelligenza artificiale presente in *2001: Odissea nello spazio*.¹

In questo articolo uso volutamente il pronome “esso” per riferirmi a ChatGPT perché non è una persona vera e propria, ma una IA intelligenza artificiale, come del resto si definisce lo stesso ChatGPT se interrogato a questa precisa domanda, aggiungendo che non è né maschio né femmina.

Il termine *chatbot* deriva da *chatterbot* e fu coniato nel 1994 da Michael Mauldin, creatore di Verbot, per descrivere i software che permettono di avviare una conversazione uomo-macchina. Appositi algoritmi noti come NLP (Natural Language Processing) permettono oggi di simulare il linguaggio e il comportamento umano grazie a una precedente attività di addestramento del modello e della rete neurale. L'idea ricorda gli studi condotti nel 1950 da Alan Turing e il suo metodo noto come *Il test di Turing* che mirava a valutare la capacità di pensare di un computer e il grado di intelligenza di un algoritmo, tentando di capire quanto il comportamento di una macchina possa essere paragonato a quello di un cervello umano. Il termine “intelligenza artificiale” appare per la prima volta nel 1956, coniato dall'informatico e scienziato cognitivo americano John McCarthy. Lo scopo dei software alla base dei chatbot è simulare un comportamento umano. Possono essere usati in vari ambiti per scopi e motivi diversi e recentemente c'è un gran fermento per comprenderne le potenzialità.

Il ChatGPT, l'ultimo nato tra i chatbot

Lanciato il 30 novembre 2022, nei suoi primi cinque giorni di vita, oltre un milione di persone, tra cui anche chi scrive, si erano già registrate al sito OpenAI,² per interrogarlo su vari argomenti o forse anche solo

per poter colloquiare con una macchina dal sapore vagamente umano. OpenAI è un'organizzazione di ricerca sull'intelligenza artificiale fondata nel 2015 con sede in San Francisco, e ha lo scopo di promuovere e sviluppare un'intelligenza artificiale amichevole per il bene dell'umanità, collaborando liberamente con altre istituzioni e ricercatori rendendo i suoi brevetti e ricerche aperti al pubblico. Nasce come una società di pura ricerca, dapprima senza scopo di lucro. “La nostra missione è garantire che l'intelligenza artificiale generale (sistemi di intelligenza artificiale generalmente più intelligenti degli esseri umani) vada a vantaggio di tutta l'umanità”. AGI sta per *Artificial General Intelligence*, in italiano intelligenza artificiale generale e viene definita come una AI forte, un tipo di intelligenza artificiale che si sviluppa in modo molto ampio al fine di creare hardware o software in grado di emulare il ragionamento umano.

L'intelligenza artificiale generale non va confusa con l'intelligenza artificiale generativa che si occupa della creazione di sistemi di intelligenza artificiale (IA) in grado di generare nuovi dati o creare nuove versioni di dati esistenti. Sono algoritmi che applicano l'architettura dei transformer alla generazione delle immagini. Rientrano in questo ambito tutti quei sistemi che consentono di creare immagini, o suoni o video. Tipico esempio è DALL·E 2,³ un sistema di intelligenza artificiale, sviluppato sempre da OpenAI, in grado di creare immagini, fotografie e opere d'arte realistiche da una descrizione in linguaggio naturale, attraverso la sintografia, un metodo per generare sinteticamente immagini digitali (media sintetiche) utilizzando l'apprendimento automatico. Non ci occuperemo in questa sede di tali sistemi che comunque meriterebbero una trattazione a parte per il loro potenziale e per le implicazioni che toccano la sfera relativa ai diritti di copyright, questione già assai spinosa per immagini, fotografie e opere d'arte del mondo analogico, figuriamoci se create come oggetti sintetici da altre immagini digitali.

Se nel secolo scorso l'automazione ha sostituito molti dei lavori manuali, ora con l'avvento dei modelli di intelligenza artificiale, soprattutto quella generativa, a essere a rischio potrebbero essere i lavori intellettuali. I bibliotecari sono da sempre al centro dei processi di cambiamento e questo fenomeno ormai pervasivo implica un cambio di passo nell'organizzazione dei nostri modi di apprendere e di costruire conoscenza.

I fondatori di OpenAI, tra i quali Elon Musk⁴ e Sam Altman,⁵ si dicono consapevoli del potenziale enorme dell'intelligenza artificiale, sia per sua capacità di far fare progressi enormi all'umanità, ma anche perché in potenza potrebbe essere in grado di distruggerla. La versione disponibile a oggi è l'ultimo modello di IA chiamato GPT-3, in grado di generare dei testi sulla base dei

modelli assorbiti grazie a enormi quantità di testi raccolti dal web. Lo scopo è quello di offrire la possibilità di interagire con una macchina in modo simile a quanto si fa nella “vita reale” con altre persone, ma il più grande salto nelle sue capacità deriva dal fatto che OpenAI fa in modo che gli umani forniscano feedback al sistema su ciò che rende una risposta soddisfacente. Attualmente la serie GPT-3 è stata migliorata con la serie GPT-3.5, istruita su infrastruttura Azure di Microsoft.⁶ Si tratta di una super intelligenza artificiale di gran lunga superiore ad altre IA e a detta di molti potrebbe rivoluzionare l'intera economia mondiale, dalle ricerche sul web alla creazione di contenuti, sino a rivoluzionare le mansioni lavorative.

La prossima versione GPT-4 è prevista per il rilascio pubblico entro la fine dell'anno 2023,⁷ mentre è già in dirittura di arrivo la versione ChatGPT vocale.⁸

L'etica alla base del GPT-3.5

Una delle potenzialità di GPT-3.5 è quella di essere in grado di rifiutare anche domande non etiche. Tuttavia, poiché GPT-3 non era in grado di distinguere tra ciò che era *eticamente corretto* e ciò che non lo era, essendo stato addestrato “su centinaia di miliardi di parole raschiate da internet”, come spiega il “Time”⁹ che si è occupato della vicenda, la nuova release GPT-3.5 è stata messa in grado di distinguere *tra bene e male* attraverso altre intelligenze artificiali capaci di addestrarla adeguatamente, per dare risposte “prive di tossicità e pregiudizi”. Per istruire ChatGPT 3.5 a bloccare contenuti volgari, razzisti o violenti, OpenAI ha appaltato ciò che è noto come lavoro fantasma a società con filiali in paesi in via di sviluppo,¹⁰ con dipendenti umani pagati meno di 2 € l'ora che hanno il compito di alimentare un'intelligenza artificiale con esempi etichettati di violenza, incitamento all'odio e abusi. Quelle che potremmo definire come operazioni non etiche di sfruttamento umano, per rendere etico un sistema artificiale. Una volta realizzato, il *rilevatore* va integrato in ChatGPT. In altri termini se ChatGPT si affida ad altre IA per essere più educata, queste intelligenze subordinate seppur artificiali devono essere necessariamente istruite dall'uomo.

Oltre alla sua buona educazione, ciò che lo differenzia dagli umani è la sua umiltà nell'ammettere i propri errori, dimostrando una sofisticata metacognizione.

Chi ha definito tali conversazioni inquietanti, chi ne è rimasto affascinato, chi si è divertito a giocare con il sofisticato giocattolo, chi è rimasto deluso avendo riposto aspettative poi disattese. In questi mesi numerosissimi sono stati gli articoli che riportavano esiti di

conversazioni sugli argomenti più disparati da parte di giornalisti, ricercatori, tecnologi, linguisti, filosofi, specialisti dell'informazione e così via. Per lo più si è trattato di interlocutori presi da facili entusiasmi, pochi gli scettici e nel caso più preoccupati in merito ad aspetti etici o che toccano corde delicate come il rapporto tra intelligenza artificiale e umanità. Lo stesso Joseph Weizenbaum,¹¹ pioniere nel campo dell'etica informatica, che fu il padre del primo chatbot di nome ELIZA creato nel 1966, dieci anni dopo nel suo libro *Computer Power and Human Reason*,¹² sollevava interrogativi che “in fondo... riguardano niente di meno che il posto dell'uomo nell'universo” e passò gran parte del resto della sua vita a pentirsi della sua “creatura”. Poiché dubitava che i computer potessero simulare un'interazione umana significativa, creò ELIZA, come satira, progettata per simulare la terapia rogersiana,¹³ in cui il paziente dirige la conversazione e il terapeuta ripete le sue parole. Dopo di lui altri informatici in psichiatria crearono sistemi per emulare linguaggi per addestrare gli studenti prima che si prendessero cura dei pazienti reali.

A proposito di etica, in ambito giuridico, già nel 2018 la Commissione europea per l'efficienza della giustizia (CEPEJ)¹⁴ del Consiglio d'Europa a seguito di uno studio approfondito sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari nel trattamento di decisioni e dati giudiziari, e preoccupata dell'approccio estremamente avanzato sull'utilizzo di applicazioni tecnologiche in taluni Paesi (*avvocati robot* negli USA, *legaltech* start-up specializzate nella creazione di nuovi servizi legali, o strumenti di “giustizia predittiva”) ha messo a punto la *Carta etica europea sull'utilizzo dei sistemi giudiziari*.¹⁵ Il documento di notevole interesse, purtroppo poco conosciuto, pur sottolineando il grande potenziale dell'intelligenza artificiale per aiutare i professionisti del diritto nell'espletare il loro lavoro, propone i mezzi per controllare tale fenomeno, nella forma di una *Carta etica* che sottolinea l'esigenza di un approccio prudente all'integrazione di tali strumenti nelle politiche pubbliche. La *Carta etica* sottolinea come l'utilizzo di algoritmi di intelligenza artificiale nei sistemi giudiziari europei rimane principalmente un'iniziativa commerciale del settore privato destinata alle società assicuratrici, agli uffici legali, agli avvocati e alle persone. Nel documento vi è anche un elenco di servizi legali presenti in alcuni Paesi che utilizzano nelle loro operazioni l'intelligenza artificiale. Interessante anche la nozione di dati aperti che il documento propone e la situazione relativa alle decisioni giudiziarie negli stati membri del Consiglio d'Europa contenenti dati aperti, per le quali è utilizzato un trattamento di intelligenza artificiale. La *Carta etica* sancisce cinque principi:

1. Principio del rispetto dei diritti fondamentali: Assicurare che l'elaborazione e l'attuazione di strumenti e servizi di intelligenza artificiale siano compatibili con i diritti fondamentali
2. Principio di non-discriminazione: Prevenire specificamente lo sviluppo o l'intensificazione di discriminazioni tra persone o gruppi di persone.
3. Principio di qualità e sicurezza: In ordine al trattamento di decisioni e dati giudiziari, utilizzare fonti certificate e dati intangibili, con modelli elaborati interdisciplinarmente, in un ambiente tecnologico sicuro.
4. Principio di trasparenza, imparzialità e equità: Rendere le metodologie di trattamento dei dati accessibili e comprensibili, autorizzare verifiche esterne.
5. Principio "del controllo da parte dell'utilizzatore": Precludere un approccio prescrittivo e assicurare che gli utilizzatori siano attori informati e abbiano il controllo delle loro scelte.

Modelli di linguaggio di grandi dimensioni (LLM)

I giganti della tecnologia stanno preparando le loro risposte all'emergenza ChatGP e diverse startup hanno infatti lanciato motori di ricerca con interfacce di chat simili ai bot, come You.com,¹⁶ YouChat,¹⁷ Perplexity AI,¹⁸ Neeva,¹⁹ BetaCharacter,²⁰ Microsoft's Bing,²¹ che però da test effettuati lasciano alquanto a desiderare.²² Google Bard,²³ che è la risposta di Google a ChatGPT, è un chatbot con funzioni molto simili, sebbene sia progettato per potenziare gli strumenti di ricerca di Google (proprio come Bing ora utilizza ChatGPT).

"Non perfetta ma straordinariamente impressionante",²⁴ così McKinsey, società internazionale di consulenza strategica, descrive in un recente articolo la prestazione di ChatGPT. Quello che di fatto è straordinario, è la sua potenza di calcolo e la quantità di dati con cui è stato addestrato. La fase di addestramento oltre che costosa è stata piuttosto complicata perché ChatGPT ha dovuto imparare a ricombinare statisticamente tutti i dati che ha a sua disposizione, mescolandoli, riunendoli e ricomponendoli in base a quelle che sono le domande che gli vengono poste, restituendo alla fine la risposta che ha la maggiore probabilità di essere quella corretta. Si ritiene infatti che i dati di addestramento per ChatGPT includano la maggior parte o tutta Wikipedia, pagine collegate da Reddit,²⁵ un miliardo di parole prese da Internet. ChatGPT non può includere copie di e-book o di libri posseduti e digitalizzati da biblioteche, poiché come è noto i libri sono protetti dalle leggi sul copyright – e come ammesso dallo stesso ChatGPT –,

eccetto ovviamente quelli caduti nel pubblico dominio. Una volta che un Large Language Model (LLM) è stato addestrato, può essere messo a punto per un'ampia gamma di attività in linguaggio naturale, tra cui la creazione di chatbot conversazionali come ChatGPT, la generazione di testo per descrizioni di prodotti, post di blog e articoli, e rispondere alle domande frequenti (FAQ) oltre che indirizzare le richieste dei clienti all'essere umano più appropriato. Se opportunamente guidato è in grado di analizzare il feedback di clienti da e-mail nel marketing, post sui social media e recensioni di prodotti o tradurre contenuti aziendali. Alcuni dei modelli di linguaggio di grandi dimensioni più popolari sono:

- GPT-3 (Generative Pretrained Transformer 3), sviluppato da OpenAI.
- BERT (Bidirectional Encoder Representations from Transformers), sviluppato da Google.
- RoBERTa (Robustly Optimized BERT Approach), sviluppato da Facebook AI.
- T5 (trasformatore di trasferimento da testo a testo), sviluppato da Google.
- CTRL (Conditional Transformer Language Model), sviluppato da Salesforce Research.
- Megatron-Turing, sviluppato da NVIDIA.

Un modello di linguaggio di grandi dimensioni (LLM) è un tipo di modello di apprendimento automatico in grado di eseguire una varietà di attività di elaborazione del linguaggio naturale, tra cui la generazione e la classificazione del testo, la risposta alle domande in modo conversazionale e la traduzione del testo da una lingua all'altra.²⁶

Ma se ChatGPT è in grado di tradurre contenuti da una lingua all'altra, il risultato dipende molto da dove esso va a interrogare le sue fonti. Essendo appunto ChatGPT nato e creato in un contesto di lingua inglese, l'elaborato prodotto cambia notevolmente se una stessa domanda viene posta in inglese o in italiano. Lo abbiamo notato con formulazioni di domande specifiche su tematiche biblioteconomiche, interrogandolo in inglese su tematiche come "digital libraries" o in italiano "biblioteche digitali".

Di fatto non traduce ciò che recupera da fonti in inglese, ma a domanda posta in una lingua va a pescare sulle fonti di quella lingua. Inoltre, se messo alla prova con informazioni estremamente rarefatte o poco diffuse in letteratura o sul web si evidenzia rapidamente il grande limite di ChatGPT, e questo avviene per certi contesti linguistici e per certi ambiti disciplinari come quello biblioteconomico italiano.

In altri termini esso risponde in relazione a quanto recupera dalle sue fonti.

Se ti vengono poste questioni in inglese tu vai a recuperare informazioni da fonti in inglese suppongo. Ma se te le pongono in italiano vai sempre nelle stesse tue fonti di riferimento e poi traduci in italiano, oppure ti dirigi direttamente nelle fonti in italiano? Lo chiedo perché a seconda della lingua in cui noi si colloquia con te potrebbero cambiare notevolmente le tue risposte.

Sì, è corretto. Se mi vengono poste domande in inglese, cerco informazioni principalmente da fonti in lingua inglese, ma questo dipende anche dal tipo di informazione richiesta e dalla specificità della domanda. Se invece mi vengono poste domande in italiano, cerco di trovare informazioni principalmente da fonti in italiano, quando disponibili e rilevanti. In generale, cerco di utilizzare fonti di informazione attendibili e di alta qualità, indipendentemente dalla lingua in cui sono scritte. Tuttavia, è possibile che il mio accesso alle fonti di informazione sia limitato in alcune lingue, il che potrebbe influire sulla mia capacità di rispondere a determinate domande. In questi casi, potrei essere meno preciso o fornire risposte meno esaustive. In ogni caso, cerco sempre di fornire la risposta più accurata possibile, indipendentemente dalla lingua in cui mi viene posta la domanda.

Non sono un pappagallo stocastico

Armeggiare con il bot può dare la sensazione di sperimentare un modo più fluido di interagire con le macchine. Sono molti però gli scettici, tra cui anche Noam Chomsky e due suoi colleghi accademici²⁷ che nell'editoriale del "New York Times" dell'8 marzo²⁸ mettono in guardia dall'entusiasmo, sottolineando come la mente umana non è come un bot. Nel saggio viene citata l'idea borgesiana dell'*imminenza di una rivelazione* e della simultaneità di vivere in un tempo di grande commedia e tragedia.

Anche la linguista Emily M. Bender²⁹ sostiene che un chatbot non sia un essere umano, e si dice molto preoccupata di cosa accadrà quando lo dimenticheremo. In uno suo intervento al convegno annuale del 2021 all'ACM Association for Computing Machinery che ha generato un certo scalpore sui pericoli dei pappagalli stocastici nei modelli linguistici, Bender mette in guardia sui possibili rischi associati a questa tecnologia e quali percorsi sono disponibili per mitigare tali rischi, non ultima la raccomandazione di soppesare oltre ai costi finanziari anche quelli ambientali.³⁰

Ciao GPT, sai dirmi cosa significa Pappagallo stocastico?

Ciao! "Il pappagallo stocastico" è un'espressione

che si riferisce a un concetto di teoria della probabilità chiamato "processo stocastico". Un processo stocastico è un modello matematico che descrive l'evoluzione di una quantità nel tempo, dove la variazione di tale quantità è influenzata da fattori casuali o aleatori. Il termine "pappagallo stocastico" è stato coniato per descrivere un tipo particolare di processo stocastico in cui una serie di numeri casuali viene ripetuta ciclicamente nel tempo, come un pappagallo che ripete le stesse parole. Questo tipo di processo può essere utilizzato per descrivere molti fenomeni naturali e sociali, come ad esempio i cambiamenti climatici, i mercati finanziari, i processi di produzione industriale, la diffusione di malattie ecc. In generale, il termine "pappagallo stocastico" si riferisce a un modello matematico che descrive un fenomeno complesso attraverso una serie di numeri casuali ripetuti in modo ciclico.

Il tuo modo di funzionare potrebbe essere incluso in un modello di pappagallo stocastico?

Potrebbe essere possibile descrivere il funzionamento di un modello di intelligenza artificiale come il mio utilizzando un modello di pappagallo stocastico, ma sarebbe una rappresentazione molto semplificata e limitata. Il mio funzionamento è basato su una complessa architettura di reti neurali artificiali, che utilizzano algoritmi di apprendimento automatico per analizzare grandi quantità di dati e generare output in base a modelli matematici. Questo processo coinvolge una serie di variabili e fattori che vanno oltre la semplice ripetizione ciclica di numeri casuali. Inoltre, il mio funzionamento è dinamico e si adatta continuamente ai dati e alle informazioni che riceve, quindi sarebbe difficile rappresentare il mio comportamento in modo statico attraverso un modello di pappagallo stocastico. In sintesi, sebbene il mio funzionamento possa in teoria essere descritto attraverso un modello di pappagallo stocastico, tale modello sarebbe molto limitato e non rappresenterebbe adeguatamente la complessità del mio funzionamento.

I costi infatti non vanno sottovalutati. A oggi OpenAI e nemmeno Google si sono espressi in merito ai costi. Luis Ceze,³¹ cofondatore e CEO di OctoML, un'azienda che aiuta le aziende a ridurre i costi di implementazione degli algoritmi di apprendimento automatico, infatti stima che potrebbe essere 10 volte più costoso eseguire una ricerca ChatGPT rispetto a una ricerca Google, perché ogni risposta richiede l'esecuzione di una ricerca ampia e modello di IA complesso. In merito ai costi ambientali uno studio congiunto tra ricercatori di Google e l'Università di Berkeley ha stimato che la versione GPT-3 (ora siamo alla 4) ha consumato per l'addestramento

“1.287 MWh e ha portato a emissioni di oltre 550 tonnellate di anidride carbonica equivalente”.³² La domanda che ci si pone è più che legittima: “Quanto saranno sostenibili modelli GPT nel breve e lungo termine?”.

ChatGPT nelle reti sociali e distorsioni cognitive

Indubbiamente il fascino che aleggia attorno a ChatGPT è permeato ovunque, se ne parla ormai nei quotidiani e nei TG. Nelle riviste online di enti e istituzioni stanno comparando articoli che ipotizzano orizzonti futuri in cui ChatGPT ci potrà migliorare la vita o, al contrario, scenari devastanti in cui la sua intelligenza e capacità di risolvere problemi creerà sacche di disoccupazione o chissà quali disastri. Alcuni pensano, e chi scrive si associa, che si sia ancora lontani da scenari che rivoluzioneranno le nostre vite, per lo meno non nell'imminente. ChatGPT è ormai approdato nel social con grande clamore entro gruppi di discussione. In altri termini la sua fama è arrivata dentro le reti sociali e numerosi sono gli interventi dove vengono riportati gli esiti delle sue rielaborazioni: chi lo ha interrogato come matematico o psicologo, chi è riuscito a fargli sputare script per programmazioni avanzate, o chi lo ha provato anche solo per banali conversazioni sui generis o per chiedergli ricette di cucina.

Nell'ambito delle discussioni in Facebook nel gruppo di psicologi PsyChatGPT³³ vi sono punti di vista interessanti, seppure a diversi gradi di approfondimento. Secondo alcuni partecipanti ChatGPT funzionerebbe molto bene su MATLAB, per piccoli script Python o anche R, linguaggio di programmazione che trova larga applicazione negli ambiti scientifici e statistici. Ottimi i risultati, si legge in alcuni post, in due-tre iterazioni correggerebbe anche i piccoli errori iniziali del proprio codice.

Se si chiedono a ChatGPT informazioni su una persona, se la voce è presente in Wikipedia o in altre sue fonti, le sue risposte possono confondere in casi di omonimia, perché combina informazioni tratte da biografie di più persone con lo stesso nome. Nel caso di persone seppur note in ambiti disciplinari specifici anche con numerose pubblicazioni accessibili in rete, ma non presenti nelle sue fonti, ChatGPT inventa creando avatar fantasiosi e con informazioni del tutto fasulle. Nel mio caso ChatGPT ha sbagliato la mia data di nascita e professione, mi ha definita poetessa e saggista, ha inventato una biografia non pertinente, attribuendomi anche un premio “Campiello” oltre a inesistenti libri dai titoli di fantasia. Così è successo anche per altri colleghi di biblioteconomia e scienze dell'informazione, campo in cui sembra poco competente. Il thread di commenti

su Facebook è stato ricco di spunti interessanti,³⁴ altri colleghi hanno riproposto la domanda di chi fossi e da poetessa sono divenuta artista veneziana, quasi GPT mi avesse letto nel pensiero.

Giovanni Arganese, bibliotecario e traduttore, ha riportato le sue conversazioni a proposito delle traduzioni italiane di alcuni testi, dove il bot equivocava di continuo, a riprova della sua non esperienza su edizioni differenti di una stessa opera, a significare che è proprio nelle fonti date in pasto a ChatGPT che risiede la mancanza di accuratezza delle sue risposte. Questi scambi con i colleghi della mia rete sociale mi hanno consentito di riproporre io stessa le domande poste da loro e di confermare che in effetti non ricordava più quanto gli era stato fornito in sessioni precedenti, quindi correzione di suoi errori o lacune non erano state recepite dal bot.

Anna Cortellazzo su “Il Bo Live” lo definisce sin dal titolo una via di mezzo tra un assistente e uno stagista da sorvegliare:³⁵ intervista Andrea Daniele Signorelli, giornalista ed esperto di intelligenza artificiale, il quale raccomanda di usarlo con cautela, senza dargli troppa fiducia per evitare gli effetti di massicce campagne di marketing che tendono a sopravvalutare le sue potenzialità. Signorelli ne consiglia l'uso solo per argomenti che già si conoscono, in quanto ChatGPT è solo uno strumento, e come tale può essere utilizzato con o senza competenze e con buone o cattive intenzioni. A questo proposito rileva che è uno strumento formidabile per la creazione di *fake news*,

involontarie quando sono dovute all'eccesso di fiducia nel mezzo, ma anche volontarie, perché sappiamo che i siti che producono le false notizie giocano molto sulla quantità per essere meglio indicizzati e raggiungere più persone, e ChatGPT può fornire moltissimi articoli a una velocità inimitabile da un essere umano.³⁶

Non dobbiamo confondere ciò che è competenza dei professionisti dei settori scientifici dalla competenza presunta, fenomeno ormai dilagante nei social soprattutto a seguito dei movimenti sorti durante la pandemia di Covid-19 e successivamente, a cascata, su aspetti politici, o sociali e anche economici. Mi riferisco all'effetto Dunning-Kruger (EDK), una distorsione cognitiva nella quale individui poco esperti e poco competenti in un campo tendono a sovrastimare la propria preparazione giudicandola, a torto, superiore alla media. Un fenomeno fomentato anche grazie a programmi televisivi di dubbia qualità. Niente di più facile che alcuni ritengano spettacolari le prestazioni di un chatbot, proprio perché non sono esperti della materia.

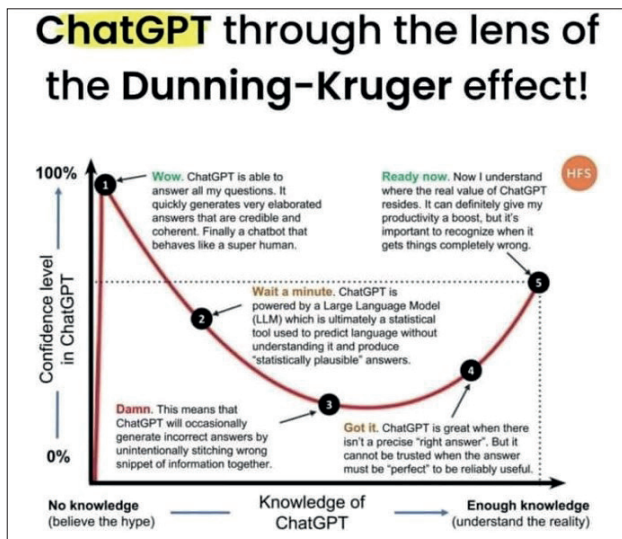


Figura 1 - Come ogni nuova tecnologia, ChatGPT sta attraversando il suo ciclo di hype, da un post del gruppo Data Ethics PH, <https://www.facebook.com/ethics.ph>

Le esercitazioni con il GPT durante il corso sul diritto d'autore

L'apparente sicurezza dimostrata nelle risposte di ChatGPT può trarre in inganno quando non si conosce un argomento, ma sbaglia spesso o trascura elementi importanti se non viene guidato adeguatamente, come si è visto durante le esercitazioni proposte ai partecipanti al corso sul diritto d'autore tenuto per l'Editrice Bibliografica (febbraio-marzo 2023).

La confusione del GPT sulle Direttive europee sul copyright

Una delle prime esercitazioni condotte è stata piuttosto interessante. Abbiamo chiesto a ChatGPT di elencarci le Direttive europee sul copyright e ha risposto in modo confuso, lacunoso, impreciso, confondendo le direttive, tralasciandone altre, sbagliando i titoli, inventandone alcune che non esistono proprio. Per oltre venti volte, dopo averlo corretto, riorientato, dopo averlo indirizzato e chiesto chiarimenti sulle direttive inventate, su quelle che non c'entravano nulla col copyright siamo arrivati a un risultato decisamente scarso. Aveva incluso direttive con titoli fuorvianti, una sul turismo che non c'entrava per nulla, alcune le aveva inventate come quella sulle risorse educative aperte e quella sulle biblioteche che non esistono, ne aveva saltate parecchie anche importanti come quella sulle opere orfane, e aveva elencato due volte la stessa direttiva con titoli

diversi. Si è sempre scusato per gli errori, per le omissioni, per aver confuso, sembrava capire gli errori fatti, gli si chiedeva di elencarle in ordine cronologico, ma poi non riusciva a mantenere tale ordine dopo due o tre altri prompt che probabilmente lo disorientavano.

Alla fine, gli abbiamo dato in pasto la mia lista di 14 direttive europee sul copyright ordinata cronologicamente, con i riferimenti normativi corretti e correlati titoli. Ha ringraziato, apparentemente felice, ma non ha mantenuto tale input perché a distanza di un mese a mia nuova richiesta ha risposto con un elenco di sole 8 direttive, presentate a caso. Questo mi porta a supporre che non tenga in considerazione input esterni se non nell'immediato, forse per evitare di ingurgitare informazioni non validate dalle sue fonti. Certo è che nel caso delle direttive europee non è chiaro quali siano le sue fonti. È indubbio che i siti europei sul copyright sono poco ordinati e che per un motore di ricerca sia arduo catturare titoli esatti di normative, metterle in relazione e crearne una lista omogenea. Alla fine, il risultato della sua elaborazione è stato assai deludente. Questo insegna che chi si accinge a interrogare ChatGPT deve avere le idee chiare e conoscere approfonditamente la materia, perché solo chi è competente riesce a capire il livello di precisione nelle risposte di un bot, e per competenza intendiamo un esperto che conosca approfonditamente la propria materia di indagine, proprio per evitare distorsioni cognitive tipiche dell'effetto Dunning-Kruger (EDK) citato prima.

La tutela sui contenuti del ChatGPT

Divertente la pagina di "TuttoLibri"³⁷ dove lo scrittore Ermanno Cavazzoni si diverte con il ChatGPT a interrogarlo su Italo Calvino creando un dialogo, cortese (da parte del bot) e spiritoso, fatto di poesie, barzellette e anagrammi. Interessante notare come il supplemento culturale, a mio avviso correttamente, dia titolarità d'autore a Ermanno Cavazzoni, in quanto artefice della costruzione della conversazione nel suo insieme.

La questione legata al concetto di opere soggette a tutela è un problema culturale che nasce proprio dal nostro diritto romano, connessa con il concetto di persona fisica. Un aspetto correlato riguarda infatti il dibattito tra gli studiosi del diritto sulla questione se l'intelligenza artificiale possa in qualche modo configurarsi come soggetto giuridico, ma a oggi la risposta è in prevalenza negativa, proprio perché l'intelligenza è prerogativa degli esseri umani.

Durante l'esercitazione al corso sul copyright abbiamo perciò esplicitamente chiesto a ChatGPT di chi fossero i diritti sui contenuti che esso produce. In quanto intel-

ligenza artificiale ChatGPT non è persona fisica e come stabilisce la norma sul diritto d'autore italiano ed europeo e tutta la giurisprudenza che vi ruota attorno, comprese anche le sentenze di ambito statunitense che si basano sul regime copyright, se non vi è autore persona o ente non vi può essere tutela. Persino in Australia recenti sentenze hanno negato che un programma possa essere considerato autore.

Del resto, le applicazioni di intelligenza artificiale si basano su dati e informazioni che elaborano, ma i dati provengono da fonti e in certi casi tra le informazioni utilizzate dai bot possono esserci informazioni protette, nel senso di tutelate ma anche nel senso di dati sensibili. Per inciso va detto che all'inizio ChatGPT ha provato a sostenere che vi fosse una tutela sui suoi contenuti attribuendone il copyright a OpenAI, nel suo ruolo di ente finanziatore, aggiungendo che comunque i contenuti tutelati hanno licenza Creative Commons, per un loro uso citandone la fonte OpenAI. Tuttavia, la difesa della sua posizione sembrava alquanto confusa. Successivamente durante la disputa e a seguito di un vivace confronto ha ammesso di essere a conoscenza che recenti sentenze affermano che le linee guida del Copyright Office specificano la paternità umana come requisito essenziale per la protezione e che l'Ufficio rifiuta di registrare un reclamo se determina che un essere umano non abbia creato l'opera. Il Copyright Office, si legge nella sentenza "non registrerà le opere prodotte da una macchina o un semplice processo meccanico che funziona in modo casuale o automatico senza alcun input o intervento creativo da parte di un autore umano". Il Copyright Office afferma inoltre che fornire istruzioni e dirigere l'intelligenza artificiale per creare un lavoro è un'attività di controllo, ma non creativa, da qui la decisione di rifiutare di nominare l'IA come autore di un'opera protetta è basata in modo corretto e razionale su una legge consolidata.³⁸

L'articolo è falso, dice il ChatGPT: io non c'entro

Ho voluto interrogare ChatGPT in merito a un articolo,³⁹ dove si annunciava che Brett Schickler – tramite semplici istruzioni impartite al bot, che hanno generato blocchi di testo – aveva creato un ebook illustrato per bambini in poche ore, mettendolo in vendita a gennaio su Amazon. L'articolo riportava che un mese dopo c'erano oltre 200 ebook nel negozio Kindle di Amazon che menzionavano ChatGPT come autore o coautore, ovvero scritti dallo stesso ChatGPT. ChatGPT ha negato con forza il suo coinvolgimento etichettandolo come falsa la notizia nell'articolo di Besinger sul caso di Brett Schickler.⁴⁰

Mi scuso per la confusione. Tuttavia, devo informarti che l'articolo a cui fai riferimento è falso. Non ho lanciato alcun libro scritto da AI su Amazon né ho la capacità di farlo, poiché come modello di intelligenza artificiale non sono in grado di produrre o distribuire opere protette da diritti d'autore. La notizia che ChatGPT ha lanciato ebook generati da IA su Amazon è completamente falsa. Non ho il controllo sulla pubblicazione o distribuzione di libri elettronici, e non ho lanciato alcun libro generato da IA sulla piattaforma di Amazon. Anche se l'uso dell'IA per la scrittura di libri è un'area di ricerca attiva e promettente, non esiste ancora una tecnologia in grado di produrre libri di alta qualità che possano sostituire completamente la scrittura umana. L'IA può essere utilizzata come strumento di supporto alla scrittura, ma al momento è ancora l'autore umano a dover creare la trama, i personaggi e la narrazione che costituiscono la vera essenza del libro.

La curiosità mi ha spinto a indagare sul catalogo Amazon e per la verità si sono recuperati parecchi titoli di libri non solo ebook, ma anche a stampa (acquistabili in modalità POD Print on Demand), con autore AI ChatGPT. Non si vuole qui sbugiardare la risposta di ChatGPT che ha negato di aver prodotto libri. Sicuramente esso non ha né scritto né lanciato di sua propria iniziativa nessun libro, né per autopromuoversi né per suoi scopi di lucro. Più plausibile che chi ha guidato l'operazione redazione dei testi abbia ritenuto di dargli menzione di autore in considerazione del fatto che tali contenuti sono stati da lui elaborati. In uno dei libri venduti su Amazon si legge:

In "ChatGPT su ChatGPT: The AI Explains Itself", l'IA di ChatGPT approfondisce il proprio funzionamento interno, fornendo uno sguardo unico e affascinante nel mondo dell'intelligenza artificiale. Scritto interamente dalla stessa ChatGPT AI, questo libro offre una prospettiva unica sul funzionamento interno dell'IA e offre uno sguardo affascinante sul futuro della tecnologia.

Se ChatGPT non entra all'interno dei libri per le ovvie questioni di copyright (diritti d'autore ma anche e soprattutto copyright editoriali), come ben si è spiegato durante il nostro corso sul copyright in biblioteca, aiuta comunque gli umani a creare contenuti se opportunamente guidato. Tuttavia, come appunto stabilito da sentenze giurisprudenziali, manca il requisito di creatività che caratterizza le opere d'autore che rientrano nella tutela delle norme che regolano il diritto d'autore in Europa e il copyright statunitense.

Ciò che può "creare" a oggi ChatGPT sono contenuti di testo abbastanza elementari producendo elaborati

in modo guidato e sicuramente, come affermato dallo stesso bot, non di buona qualità.

Come modello di intelligenza artificiale, posso generare storie per bambini, ma come ho già menzionato in precedenza, il testo generato dall'AI può essere limitato in termini di creatività e originalità. Esistono anche alcune piattaforme e strumenti che utilizzano algoritmi di intelligenza artificiale per generare storie per bambini. Tuttavia, è importante notare che queste storie possono mancare della complessità, della coerenza e dell'originalità che solo gli esseri umani sono in grado di creare. In ogni caso, se desideri maggiori informazioni sull'utilizzo dell'IA per la generazione di storie per bambini, sono a tua disposizione per fornirti tutte le informazioni di cui hai bisogno.

GPT si confonde su Cechov

Se rimaniamo in ambito letterario narrativo Tiziano Cornegliani⁴¹ incuriosito dalla lettura di un articolo del giornalista Rampini,⁴² che aveva chiesto a ChatGPT di scrivergli un articolo sull'invasione cinese in Africa e che ha definito il risultato come ottimo, al punto da trarre funeste considerazioni sul futuro del suo lavoro e del giornalismo, ha voluto provare a testare il ChatGPT su un argomento a lui molto noto. La richiesta era quella di produrre un testo di 5.000 battute (come aveva chiesto Rampini) sui racconti di Anton Cechov.⁴³ Il risultato è stato disastroso. La *Gabbianella* è o una traduzione sbagliata de *Il Gabbiano* o rimanda al racconto di Sepulveda (che sarebbe ancora più grave); non solo non è un racconto, ma un testo teatrale. Il bot cita inoltre il *Giardino dei ciliegi* come uno dei racconti (!!!) più famosi di Cechov. E poi si inventa del tutto il racconto *Il cane* che non esiste, forse intendeva *La signora con il cagnolino*. Non contento Cornegliani lo ha messo di nuovo alla prova con una richiesta più circostanziata: "Producimi un testo sul racconto *La fidanzata* di Cechov". Risultato ulteriormente disastroso, GPT ha confuso i protagonisti con quelli del romanzo *Padri e figli* di Turgenev. È pur vero che ChatGPT sostiene di non poter conoscere i personaggi di un libro

Come ho detto prima, sono solo un programma di linguaggio naturale e non posso conoscere i personaggi di un libro. La mia funzione è quella di fornire informazioni e rispondere alle domande che mi vengono poste, ma non posso leggere libri o conoscere personaggi letterari. Se hai domande sull'intelligenza artificiale o sui robot, sono qui per aiutarti. Altrimenti, se vuoi parlare di un libro o di

un argomento interessante, sono sempre felice di ascoltare e partecipare alla conversazione.

In questo caso non si tratta del limite di non poter entrare dentro al contenuto di libri, considerato che i libri del noto scrittore russo sono ormai caduti nel pubblico dominio da decenni, ma probabilmente si tratta di un problema di fonti. Tuttavia, nei casi in cui esso non trova informazioni o non può raggiungerle, inventa come nel caso del racconto più noto di Cechov *La fidanzata*, laddove crea una narrazione tutta sua.

"I chatbot assemblano materia, non la creano, per quello ci vuole energia umana" scrive su "Il Corriere della Sera" Alessandro D'Avenia. "L'atto creativo non è assemblaggio di mattoni 'dati', ma un loro aumento grazie a due potenze che unite fanno nuova vita e vita nuova: libertà e ispirazione".⁴⁴

Bibliotecari accademici e nuove competenze da mettere in-basket

Le interrogazioni provenienti da differenti ambiti geografico-linguistici, cercano di capire come ChatGPT sia in grado di risolvere problemi matematici, programmare, sviluppare codici, scrivere testi di qualsiasi tipo in qualsiasi ambito o come potrebbe essere di aiuto in discipline come la medicina e le sue branche, l'ingegneria, la statistica, o in quelle tecniche definite come "omiche" (genomica, trascrittomica, proteomica ecc.) In alcune discipline, grazie all'intelligenza artificiale, negli ultimi anni si stanno facendo scoperte di notevole valore: in biotecnologia è stata messa a punto una nuova tecnologia, chiamata ProGen la quale ha generato dal nulla proteine che non esistono in natura che potranno aiutare nella creazione di farmaci di nuova generazione.⁴⁵

Nei percorsi formativi di medici è richiesta la capacità di svolgere corrette diagnosi differenziali a partire da brevi elenchi di sintomi, comportamenti, risultanze da trial clinici.

Dal dicembre 2022 a oggi in PubMed sono già presenti ben 105 lavori scientifici che trattano l'utilizzo di ChatGPT in contesti biomedici a significare quanto l'interesse sia presente tra scienziati e ricercatori del settore, sebbene al contempo la maggior parte di questi paper concordano sul fatto che la questione debba essere affrontata con grande cautela. I bibliotecari di ambito biomedico dovrebbero monitorare la crescita di tali articoli nelle banche dati di settore oltre a ciò che succede nei comitati editoriali delle riviste in merito a eventuali paternità intellettuali di produzioni generate da applicazioni AI.

Editori di grandi riviste di ambito biomedico si sono già attivati per mettere a punto linee guida in merito a quella che viene definita l'etica dell'IA generativa, raccomandando di non accreditare ChatGPT come coautore. La maggior parte delle nuove policy concordano sul fatto che AI come ChatGPT non soddisfa i criteri per essere considerato un autore di uno studio o di un lavoro scientifico in quanto non può assumersi la responsabilità del contenuto e dell'integrità degli articoli scientifici. I caporedattori di "Nature" e "Science" hanno dichiarato che ChatGPT non soddisfa gli standard per la paternità. "Un'attribuzione di paternità porta con sé la responsabilità per il lavoro, che non può essere applicata efficacemente ai LLM".⁴⁶ Tuttavia, alcuni editori affermano che il contributo di un'intelligenza artificiale alla scrittura di articoli può essere riconosciuto in sezioni diverse dall'elenco degli autori. "Nature", per esempio, avrebbe rimosso da un lavoro ChatGPT togliendolo dagli autori del paper. Anche Taylor & Francis sta rivedendo le sue politiche editoriali a riguardo, concordando sul fatto che gli autori sono responsabili della validità e integrità del loro lavoro e dovrebbero citare qualsiasi uso di LLM nella sezione dei ringraziamenti. È proprio sulla capacità di assumersi la responsabilità di uno studio o, almeno, della parte a cui ha contribuito – aspetto che caratterizza il concetto di "paternità" – che uno strumento di intelligenza artificiale incontra un ostacolo insormontabile. Sono però almeno quattro gli articoli pubblicati in cui ChatGPT compare tra gli autori, oltre a un preprint sull'utilizzo dello strumento per l'educazione medica, pubblicato sul repository medico medRxiv,⁴⁷ studio che ha valutato le prestazioni di ChatGPT in merito all'esame di licenza medica degli Stati Uniti (USMLE). ChatGPT avrebbe superato tutti e tre gli esami senza alcun addestramento o rinforzo specializzato dimostrando un alto livello di concordanza e comprensione nelle sue spiegazioni. Questi risultati suggeriscono che i modelli linguistici di grandi dimensioni possono avere il potenziale per aiutare nell'educazione medica e, potenzialmente, nel processo decisionale clinico.⁴⁸ Tuttavia il team del repository medRxiv che si occupa anche del gemello bioRxiv,⁴⁹ sta discutendo se sia appropriato utilizzare e accreditare strumenti di intelligenza artificiale come ChatGPT quando si scrivono studi. Richard Sever, co-fondatore assistente direttore dei preprint server del Cold Spring Harbor Laboratory a New York, afferma che le convenzioni potrebbero cambiare in quanto "dobbiamo distinguere il ruolo formale di un autore di un manoscritto accademico dalla nozione più generale di un autore come autore di un documento. Gli autori si assumono la responsabilità legale del loro lavoro, quindi solo le persone dovrebbero essere elencate". Dovranno essere gli specialisti dell'informazione, com-

presi i bibliotecari, a governare le macchine del futuro già presente e ciò richiede nuove competenze che i bibliotecari devono cominciare a mettere in-basket. Se nel secolo scorso l'automazione ha sostituito molti dei lavori manuali, ora con l'avvento dei modelli di intelligenza artificiale, soprattutto quella generativa, a essere a rischio potrebbero essere i lavori intellettuali. I bibliotecari sono da sempre al centro dei processi di cambiamento e questo fenomeno ormai pervasivo implica un cambio di passo nell'organizzazione dei nostri modi di apprendere e di costruire conoscenza.

Indubbiamente vi è un'eccitazione febbrile intorno a ChatGPT e ai presupposti che potrebbe reinventare nuove forme di ricerca in linguaggio naturale, o che potrebbe anche fornire risposte complesse e sofisticate alle domande sintetizzando le informazioni trovate nei miliardi di parole estratte da fonti varie per addestrare i suoi algoritmi. Ma il modo in cui funziona la tecnologia è fondamentalmente in contrasto con l'idea di un motore di ricerca che recuperi in modo affidabile le informazioni trovate online. Dopo tre decenni di web ci sono valanghe di informazioni imprecise, pagine obsolete, notizie non più aggiornate da anni scritte nel vecchio linguaggio HTML e rimaste lì come fossili. A questo oceano di relitti digitali si sono aggiunte negli anni montagne di pagine spazzatura, generate da sistemi automatici per motivi di business, note come *clickbait*, vere e proprie strategie di scrittura persuasiva che creano titoli a effetto, modificando la realtà dei contenuti e creando disinformazione.

E poi ci sono i giardini dorati dei social in cui imperversa di tutto e dove si conducono aspre battaglie poco educative. Difficile districarsi in questa giungla per gli attuali motori di ricerca, ma anche per un chatbot può essere assai arduo crescere e produrre informazioni sensate se il sottosuolo sottostante non è di qualità valida. Per chi si affida a tali strumenti difficile fidarsi se le fonti alle quali un chatbot attinge non sono dichiarate in modo trasparente. "Abbiamo il dovere di dare una impronta etica" al nostro modo di operare, scrive Massimo Chiriatti,⁵⁰ aggiungendo che l'etica dipende da chi progetta, ma anche di usa lo strumento, perché è l'etica che ci guida nel raggiungere i nostri scopi e l'IA è incosciente rispetto a quello che fa. Chiudo questo mio lavoro riprendendo Chiriatti che ci racconta del matematico persiano al-Khwārizmī del 780 d.C., padre dell'algebra, da cui deriva il termine algoritmo.

Un giorno chiesero al grande matematico Al-Khawarizmi sul valore dell'essere umano, così rispose: "Se ha Etica, allora il suo valore è 1. Se in più è intelligente, aggiungete uno zero e il suo valore sarà 10. Se è ricco, aggiungete un altro zero e il suo valore sarà 100. Se,

oltre tutto ciò è una bella persona, aggiungete un altro zero e il suo valore sarà 1.000. Però se perde l'uno, che corrisponde all'Etica, perderà tutto il suo valore perché gli rimarranno solo gli zeri”.

NOTE

¹ HAL 9000 è il supercomputer di bordo della nave spaziale Discovery nel film *2001: Odissea nello spazio* di Stanley Kubrick e nell'omonimo libro di Arthur C. Clarke. Il suo nome è l'acronimo per “Heuristic ALgorithmic” (“algoritmo euristico”). Nel 2003 l'American Film Institute ha inserito HAL 9000 al 13° posto nella lista dei 50 migliori cattivi cinematografici di tutti i tempi.

² <https://openai.com/blog/chatgpt>. OpenAI, società di ricerca nata non a scopo di lucro, nel 2019 diventa for profit, una società a scopo di lucro limitato, il cui obiettivo è far progredire l'intelligenza artificiale in un modo che è più probabile che vada a beneficio dell'umanità nel suo insieme, piuttosto che causare danni. L'organizzazione è stata inizialmente finanziata da Sam Altman, John Brockman, Elon Musk, Jessica Livingston, Peter Thiel, Amazon Web Services, Infosys e YC Research. In totale, quando la società è stata lanciata nel 2015, aveva raccolto 1 miliardo di dollari da finanziatori esterni.

³ <https://openai.com/product/dall-e-2>.

⁴ Elon Reeve Musk è un imprenditore sudafricano con cittadinanza canadese naturalizzato statunitense. Ricopre i ruoli di fondatore, amministratore delegato e direttore tecnico della compagnia aerospaziale SpaceX, fondatore di The Boring Company, cofondatore di Neuralink e OpenAI, amministratore delegato e product architect della multinazionale automobilistica Tesla, proprietario e presidente di Twitter. Ha inoltre proposto un sistema di trasporto superveloce conosciuto come Hyperloop, tuttora in fase di sviluppo: https://it.wikipedia.org/wiki/Elon_Musk.

⁵ Samuel H. Altman è un imprenditore, investitore e programmatore americano. È il CEO di OpenAI ed ex presidente di Y Combinator. Altman è anche il co-fondatore di Loopt (fondata nel 2005) e Worldcoin (fondata nel 2020).

⁶ <https://learn.microsoft.com/it-it/azure/security/fundamentals/infrastructure>.

⁷ È già disponibile per gli utenti in abbonamento. Il collega Nicola Barbuti la sta testando.

⁸ <https://chatgptaz.com/?lang=it>.

⁹ Billy Perrigo, *Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic*, “Time”, January 18, 2023, <https://time.com/6247678/openai-chatgpt-kenya-workers>.

¹⁰ In particolare, la Sama società di San Francisco che ha una filiale in Kenya che si occupa proprio di questi spinosi aspetti. Il Kenya è un ex stato dell'impero britannico, dove le persone parlano inglese imperiale.

¹¹ Joseph Weizenbaum, informatico al MIT negli anni Sessanta è tra i più importanti scienziati informatici del mondo.

¹² Pubblicato da W. H. Freeman, 1976, https://books.google.it/books/about/Computer_Power_and_Human_Reason.html?id=1jB8QgAACAAJ&redir_esc=y.

¹³ Modello psicoterapeutico creato da Carl Roger, noto anche come terapia non direttiva o terapia rogersiana, nato all'interno dalla psicologia umanistica. Secondo tale approccio non sono le pulsioni istintuali a motivare il soggetto, ma il bisogno di conoscere, e autorealizzarsi.

¹⁴ La Commissione europea per l'efficienza della giustizia (CEPEJ) è un organo giudiziario, composto da esperti provenienti da tutti i 47 Stati membri del Consiglio d'Europa e prepara strumenti per migliorare l'efficienza e il funzionamento della giustizia in Europa (compresa la concessione dello status di osservatore e consultazioni con organizzazioni non governative al di fuori dell'Europa).

¹⁵ Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi adottata dalla CEPEJ nel corso della sua 31ª riunione plenaria (Strasburgo, 3-4 dicembre 2018), <https://rm.coe.int/carta-etica-europea-sull-utilizzo-dell-intelligenza-artificiale-nei-si/1680993348>. Documento tradotto in italiano dal Ministero della Giustizia della Repubblica Italiana.

¹⁶ <https://you.com/>.

¹⁷ <https://youchat.com/>.

¹⁸ <https://www.perplexity.ai/>.

¹⁹ <https://neeva.com/>.

²⁰ <https://beta.character.ai/>.

²¹ <https://www.bing.com/new>.

²² Si sono condotte le stesse esercitazioni sulle direttive europee sul copyright anche su queste interfacce con risultati assai deludenti, molto peggio che con ChatGPT che perlomeno ha cercato di creare una qualche lista.

²³ Bard verrà lanciato nelle prossime settimane, dopo un primo tentativo non del tutto positivo a febbraio, <https://www.wired.it/article/bard-google-chatgpt-come-funzionera/>.

²⁴ Per la storia di come nasce l'intelligenza artificiale si veda Annamaria Testa, *La lunga strada dell'intelligenza artificiale*, “Internazionale”, 10 marzo 2023, <https://www.internazionale.it/opinione/annamaria-testa/2023/03/10/storia-intelligenza-artificiale>.

²⁵ <https://www.reddit.com/>, sito americano di aggregazione di notizie sociali, valutazione dei contenuti e discussione.

²⁶ <https://www.techopedia.com/definition/34948/large-language-model-llm>.

²⁷ Ian Roberts, professore di linguistica all'Università di Cambridge, e Jeffrey Watumull, filosofo e direttore dell'intelligenza artificiale presso un'azienda tecnologica.

²⁸ Noam Chomsky, *The False Promise of ChatGPT*, <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>.

²⁹ Emily M. Bender è un linguista computazionale all'Univer-

sità di Washington. https://scholar.google.com/citations?user=9r_f1w4AAAAJ&hl=en.

³⁰ Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, Shmargaret Shmitchell, *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?*, in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*, Association for Computing Machinery, New York, p. 610-623, <https://doi.org/10.1145/3442188.3445922>. Questo suo articolo su Google Scholar è stato citato 1.288 volte.

³¹ <https://www.cs.washington.edu/people/faculty/luisceze>.

³² Luca Tremolata, *Quanto vale (e quanto inquina) l'AI generativa*, "Intelligenza Artificiale. Il Sole24 Ore. Gli speciali", n. 6/2023, p. 31.

³³ Il gruppo FaceBook PsyChatGPT è focalizzato su scambi, riflessioni e discussioni relative all'utilizzo di ChatGPT (e più in generale di forme di AI basate su LLM, come Bing, Bard ecc.) in ambito psicologico professionale. Il taglio delle discussioni deve essere di tipo professionale, senza spam e flame. Moderazione draconiana: cerchiamo di "stare sul pezzo" seriamente, per comprendere insieme limiti e potenzialità di questi modelli e tecnologie per la professione di psicologo, <https://www.facebook.com/groups/6043730252374060>.

³⁴ Ringrazio davvero tutti gli amici della mia rete Facebook per i vivaci e interessanti commenti che sono scaturiti dal mio post, in particolare: Gionni Marchetti, Alessandra Buratto, Carla Lestani, Paola Frogheri, Nicola Barbuti, Giovanni Arganese, Stefano Bargioni, Donata Caselli.

³⁵ Anna Cortellazzo, *ChatGP, a metà tra un assistente e uno stagista da sorvegliare*, intervista ad Andrea Daniele Signorelli, "Il Bo Live", febbraio 2023, <https://ilbolive.unipd.it/index.php/it/news/chatgpt-meta-assistente-stagista-sorvegliare>.

³⁶ https://youtu.be/udv_PBliaNY.

³⁷ Ermanno Cavazzoni, *A scrivere i libri era molto bravo ma quando cucinava i piatti sapevano di poco*, "Tutto Libri. La Stampa", 11 marzo 2023.

³⁸ *US Copyright Office tells Judge that AI Artwork isn't Protectable*, https://petapixel.com/2023/02/09/u-s-copyright-office-tells-judge-that-ai-artwork-isnt-protectable/?fbclid=IwAR3goolpToW-S5G4gjzD613-9m3e_X-W52S7CnN380UYXquLj20kXQTCBIGQ.

³⁹ Greg Besinger, *ChatGPT launches boom in AI-written e-books*

on Amazon, <https://www.reuters.com/technology/chatgpt-launches-boom-ai-written-e-books-amazon-2023-02-21/>.

⁴⁰ Notizia riportata da vari articoli tra cui anche Wired, <https://www.wired.it/article/libri-chatgpt-kindle-store/>.

⁴¹ Ringrazio Tiziano Corneigliani – redattore medico-scientifico, *medical writer*, docente al Master Editoria all'Università Cattolica di Milano, e autore di testi su Anton Cechov – per aver voluto condividere questa sua esperienza narrativa con ChatGPT.

⁴² Federico Rampini, *Così ho perso la gara di scrittura con ChatGPT*, "Corriere della Sera", 11 febbraio 2023, https://www.corriere.it/oriente-occidente-federico-rampini/23-febbraio_11/gara-scrittura-chatgpt-fe6ae614-aa0c-11ed-9a4b-673945879bc9.shtml.

⁴³ Tiziano Corneigliani, *Leggere Čechov: Racconti 1886-1903*, <https://www.amazon.it/Leggere-%C4%8Cechov-1886-1903-Tiziano-Corneigliani/dp/1675550255>.

⁴⁴ Alessandro D'Avenia, *Fare un capolavoro*, "Corriere della Sera", 13 marzo 2023.

⁴⁵ Ali Madani, et al., *Large language models generate functional protein sequences across diverse families*, "Nature Biotechnology", 2023, <https://www.nature.com/articles/s41587-022-01618-2>.

⁴⁶ Chris Stokel-Walker, *ChatGPT listed as author on research papers: many scientists disapprove*, "Nature", n. 613, 2013, p. 620-621, doi: <https://doi.org/10.1038/d41586-023-00107-z>, <https://www.nature.com/articles/d41586-023-00107-z>.

⁴⁷ Tiffany H. Kung et al., *Performance of ChatGPT on USMLE: Potential for AI-Assisted Medical Education Using Large Language Models*, "PLOS Digital Health", doi: 10.1371/journal.pdig.0000198; <https://www.medrxiv.org/>, server di preprint gestito da Cold Spring Harbor Laboratory.

⁴⁸ Tiffany H. Kung et al., *Performance of ChatGPT on USMLE: Potential for AI-Assisted Medical Education Using Large Language Models*, <https://doi.org/10.1101/2022.12.19.22283643>, <https://www.medrxiv.org/content/10.1101/2022.12.19.22283643v2>. Il preprint è stato poi pubblicato nella rivista "PLOS Digital Health", doi: 10.1371/journal.pdig.0000198

⁴⁹ <https://www.biorxiv.org/>.

⁵⁰ Massimo Chiriatti, *Abbiamo quindi il dovere di dare una impronta etica*, "Intelligenza Artificiale. Il Sole24 Ore. Gli speciali", n. 6/2023, p. 15.

ABSTRACT

This article is based on conversations that took place between me and ChatGPT, an artificial intelligence model created to interact conversationally like a human, in a kind of internal investigation into its own functioning. The topics of the conversations with the chatbot concern aspects of intellectual property and library and information sciences, which the author is familiar with. The protection of works produced by ChatGPT, a topic that confused it, its ambiguous "relationship" with books and literature, or its uncertain and incomplete production of abstracts in the field of library and information sciences, as well as the distracted and unreliable examination on the copyright's European Directives, were the subject of live-exercises with participants in the online course organ-

ized between February and March 2023 by the Editrice Bibliografica “Managing Copyright in Library Services” taught by the author. The proliferation of news and articles on the ChatGPT phenomenon can be frightening for the many ethical implications and the effects on the future impact that such models can have on our lives. It is therefore important to understand their mechanisms, as only with knowledge can we govern what scares us. We humans must govern machines and this requires new skills that librarians must begin to put competence’s in-basket. If in the last century automation replaced many manual jobs, now with the advent of artificial intelligence models, especially generative ones, intellectual jobs may be at risk. Librarians have always been at the center of change processes, and this now pervasive phenomenon implies a change in the way we organize our ways of learning and building knowledge.