

Mark Coeckelbergh, David J. Gunkel, *Communicative AI: a critical introduction to Large Language Models*, Polity Press, Cambridge 2024, 144 pp.

Giacomo Pezzano*

Questo libretto, tanto agile quanto denso, avveduto e documentato, si propone come una guida critica ai *Large Language Models* (LLM), capace di investigare sia il loro significato teorico, sia le loro implicazioni pratiche. Il suo punto di partenza è che la fine di Novembre del 2022 – data del rilascio di ChatGPT – segna non solo una trasformazione tecnica ed economica, bensì una vera e propria pietra miliare interna alla rivoluzione copernicana portata avanti dalle macchine dell'informazione. Per millenni, perlomeno nella tradizione che siamo soliti chiamare occidentale, gli esseri umani si sono definiti come gli unici esseri dotati di parola e linguaggio e su questa convinzione si sono erette le antropologie tradizionali – oltre che la filosofia stessa. Tuttavia, con l'avvento delle IA comunicative questo dato di partenza si trova a dover essere rivisitato, intaccando – se non distruggendo – il nostro senso di auto-identità e di eccezionalismo: abbiamo smesso di essere gli unici esseri discorsivi (p. 1).

A partire da questa constatazione, gli autori intraprendono un percorso suddiviso in sette capitoli, seguendo due direttive fondamentali (pp. 2-3). Da un lato, le IA comunicative vengono viste come oggetti che richiedono il pensiero filosofico per discutere nozioni capitali come quelle di verità, comprensione, significato, coscienza, intelligenza, creatività, identità, autonomia, bias, privacy e autorialità. Dall'altro lato, le IA comunicative vengono considerate strumenti attraverso cui pensare, anzi ripensare, dando vita a una rivalutazione dei modi in cui la filosofia ha concettualizzato alcune tra le proprie questioni chiave. Si tratta di un libro certamente opportuno e che riesce sia a svolgere una funzione introduttiva, sia a offrire un contributo originale, risultando così utile tanto per chi studia quanto per chi fa ricerca. In questa recensione, ne presenterò estesamente i contenuti dividendoli in tre blocchi tematici: la prima sezione si soffermerà sul Capitolo 1, dedicato ad alcuni aspetti tecnici cruciali per comprendere il funzionamento dei LLM (§ 1); la seconda sezione ricostruirà i Capitoli 2, 3 e 6, focalizzati sull'insieme delle questioni etiche, giuridiche, politiche ed epistemiche legate all'utilizzo delle IA comunicative (§ 2); infine, la terza sezione si concentrerà sui Capitoli 4, 5 e 7, incentrati su un ripensamento radicale del nostro modo di intendere e praticare il linguaggio, la comunicazione e la scrittura (§ 3). *SPOILER ALERT*: il testo sarà chiuso da una breve nota sulla “scrittura assistita” alla base di questa stessa recensione (§ 4).

* Università degli Studi di Torino, e-mail: giacomo.pezzano@unito.it.

1. Dentro la scatola nera

Il Capitolo 1 ha come obiettivo di «demistificare le tecnologie dei LLM» (p. 11), seguendo una triplice articolazione: gli LLM sono situati nel contesto più ampio delle intelligenze artificiali che processano il linguaggio naturale (a partire dalle prime chatbot come ELIZA); ne vengono spiegate le operazioni tecniche e le applicazioni; se ne identificano alcune cruciali sfide tecniche.

Sui primi due fronti, gli autori insistono sull'importanza del passaggio dal paradigma simbolico a quello connectionista nelle ricerche sulle IA, associandolo a due differenti modi di trattare il linguaggio. Il primo rimanda alla pratica dell'analisi logica e grammaticale, in quanto approccia il linguaggio come un sistema logico-formale fatto di simboli e di regole. I segni verbali vengono etichettati sulla base della loro appartenenza a una categoria generale di cui rappresentano l'occorrenza (nome/verbo/articolo/preposizione), per poi venire combinati seguendo regole di assemblaggio predeterminate (articolo + nome + verbo + preposizione + articolo + nome). Il secondo – alla base delle correnti IA comunicative – segue una logica diversa, che approccia il linguaggio come un sistema di relazioni di occorrenza probabilistica, la cui mappatura è precisamente quanto viene demandato alle reti neurali artificiali del tipo *Generative Pre-trained Transformer*. Ma procediamo per gradi.

Sintatticamente, una sequenza del tipo “ieri sera ho mangiato un ___” risulta correttamente completata tanto dicendo “ieri sera ho mangiato un’astronave” quanto dicendo “ieri sera ho mangiato un asparago”. Eppure, semanticamente, la prima suona decisamente bizzarra, mentre la seconda suona alquanto verosimile – indipendentemente dal suo effettivo valore di verità (effettivamente, non ho mangiato asparagi ieri sera). In breve, sulla base della sequenza data, la probabilità di co-occorrenza delle parole “mangiare” e “asparago” è più alta di quella delle parole “mangiare” e “astronave”; al contempo, le cose già cambierebbero se, per esempio, la sequenza iniziale fosse stata “ieri sera nei miei sogni ho mangiato un ___”. Su queste basi, l’intero sistema linguistico può essere interpretato come una nebulosa dei vari coefficienti di probabilità che regolano le possibili combinazioni tra parole e sequenze di parole. Questo fa sì che il significato venga considerato in chiave endogena piuttosto che esogena: può essere calcolato su base puramente statistica, senza tener conto della relazione tra le parole e le cose o tra le parole e la mente.

È proprio qui che intervengono le reti neurali artificiali, incaricate di calcolare i vari coefficienti di probabilità per ogni possibile parola ed espressione nel modo più efficace ed economico possibile. L’operazione avviene nel seguente modo. Ciascuna parola riceve un “peso” sotto forma di valore numerico («vettore parola»), in modo tale da assegnare vettori simili a parole che condividono certe caratteristiche semantiche – così che, per esempio, “cane” risulta numericamente più vicino a “bassotto” e meno prossimo, ma comunque non troppo lontano, da “gatto”. In questa maniera, ogni parola o sequenza ordinata di parole risulta circondata da un “alone” di termini o espressioni affini, una sorta di spazio di possibilità che – avendo mappato la rete di probabili dipendenze reciproche – consente di fare previsioni su quale termine o espressione preceda o segua un altro termine o espressione e così via. Tuttavia, c’è un ulteriore passaggio che interviene a rendere il processo più semplice ed effettivamente gestibile: anziché operare la predizione prestando attenzione a ogni singola parola, le reti neurali si avvalgono di un *attention network*, che

attribuisce alle parole-vettore anche un «valore di attenzione» tra 0 e 1, stabilito in base al peso che ciascun termine ha nella sequenza in esame. Questa ulteriore determinazione reciproca – la “trasformazione” indicata in *Generative Pre-trained Transformer* – genera un «vettore di contesto», che viene infine moltiplicato per il «vettore parola», così da aumentare il valore delle «parole-attenzione», cioè dei termini che rendono accurata la predizione, e diminuire quello delle altre parole (pp. 16-19).

L'intero processo di definizione dei coefficienti si basa su fasi iterative di addestramento (*pre-training*) e correzione retroattiva degli errori (*backpropagation*), che prevede miliardi di ripetizioni fino a quando il modello riesce a minimizzare l'errore tra la parola prevista e quella effettivamente presente nel testo di addestramento. Perciò, l'intera operazione risulta tanto più efficace quanto più – da un lato – vasto e vario è il dataset utilizzato e quanto più – dall'altro lato – si introducono livelli di trasformazione nella computazione, facendo sì che l'output di un livello diventi l'input di un altro e così via, secondo una scala di crescente complessità. Per esempio, se già GPT-3 possedeva 96 livelli, GPT-4 conta quasi 2 trilioni – *due miliardi di miliardi* – di parametri per attribuire i pesi distribuiti su 120 livelli di elaborazione: ecco perché sono modelli linguistici *large* (p. 20). Sulla base di questa architettura, le IA risultano capaci di generare testi non solo senza dover comprendere il senso delle espressioni verbali e senza sviluppare intenzioni comunicative, ma *proprio perché* prescindono da problemi di ordine semantico: una volta ricevuto un prompt, questo viene scomposto in *tokens* di 4-5 parole indipendentemente dal loro significato o contenuto informativo, prima di essere processato nei vari *layer* per arrivare a generare il responso.

È per questo che, venendo al terzo fronte, emergono problemi come i *bias* “assorbiti” dai dati di addestramento o le *allucinazioni* legate a stringhe testuali coerenti, ben formate e grammaticalmente corrette, ma che appaiono fattualmente sbagliate o fabbricate *ad hoc*, quando non prendono addirittura la forma di *nonsense* (pp. 22-23). Prima ancora di presentare implicazioni epistemico-cognitive (*overtrust* e autorità informativa), etico-giuridiche (copyright e responsabilità autoriale) o socio-politiche (manipolazione e governance democratica), queste sfide sono considerate come questioni prettamente tecniche, che chiamano in causa la necessità di soluzioni per perfezionare il *tuning*, introducendo, per esempio, criteri normativi e valori guida sotto forma di *feedback* aggiuntivi forniti da “controllori” umani o di principi che fanno da filtro automatico – stile la “Costituzione” adottata da Claude di Anthropic. Analogamente, le difficoltà poste dalla necessità di disporre di quantità massive di dati testuali e di ingenti risorse computazionali ed energetiche per portare avanti l'addestramento vanno anche colte come questioni tecniche (pp. 24-25).

2. Quando l'apparenza fa la differenza

Le considerazioni più tecniche cominciano a cedere il passo a quelle più prettamente filosofiche con il Capitolo 2 (pp. 28-40), concentrato sui temi appena indicati e altri come la disoccupazione, il *digital divide*, il potere, l'*accountability* e la creatività. In questa prospettiva, si discutono i pericoli, a livello individuale come sociale, relativi alla diffusione di informazioni inaccurate anche se non deliberatamente menzognere, all'utilizzo massiccio di algoritmi di profilazione e decisione che riprendono e amplificano forme di marginalizzazione e discriminazione, alle crescenti asimmetrie e diseguaglianze tra chi – dai

lavoratori singoli a intere nazioni o gruppi di potere – trae vantaggio dalla presenza dei LLM e chi invece ne risulta svantaggiato o peggio. Emerge così la necessità di interrogarsi sulla responsabilità etica, sociale e politica, che richiede una maggiore attenzione alla distribuzione delle risorse e del potere, alla partecipazione democratica e al coordinamento internazionale non solo in fase di utilizzo, ma anche e prima di tutto in fase di design e sviluppo delle IA. In questo modo, può si possono affrontare a monte questioni come – tra le altre – lo sfruttamento ambientale, la protezione dei dati e la proprietà intellettuale, prima di avere a che fare con le loro conseguenze incontrollate a valle.

È su questo sfondo che i due autori cominciano a sollevare una questione che diventerà centrale nei capitoli 4, 5 e 7 (pp. 38-40). La proliferazione di testi generati da macchine che competono direttamente con quelli scritti da esseri umani genera criticità affatto peculiari: anche quando non si dà il caso di plagio o infrazione di copyright, si pone comunque un problema di autenticità e di trasparenza, ossia di paternità del testo. Queste pagine portano il mio nome, ma potrebbero non essere state scritte da me e non nel senso che potrebbero essere state invece scritte da un altro essere umano – il classico collaboratore o assistente del Professore: potrebbero essere state composte da una qualche IA generativa (vedi § 4). Come viene evidenziato nel libro, questa eventualità, a fianco degli aspetti etico-giuridici o etico-politici del rispondere delle opere che si mettono in circolazione, comporta anche un problema educativo: come fare in modo di salvaguardare l'esercizio della scrittura non-assistita, o perlomeno di progettare ambienti di apprendimento che favoriscano un rapporto critico con questa nuova tecnologia? Tutto ciò rimanda a un interrogativo ancora più radicale, che viene sollevato richiamando opportunamente il *Fedro* di Platone: che cosa significa scrivere? È un'abilità come un'altra, che può essere aumentata, o persino completamente appaltata, a una macchina? E come può cambiare il nostro modo di pensare – nonché di essere umani – quando il nostro modo di scrivere va incontro a trasformazioni tanto epocali?

Il Capitolo 3 (pp. 42-57) richiama il dibattito più tipico quando compaiono sulla scena le IA: sono esse realmente intelligenti e coscienti? Sono vere menti? Innanzitutto, gli autori osservano che le nostre definizioni di intelligenza, coscienza e mente sono piuttosto instabili e aperte, per poi richiamare l'esistenza dell'«effetto-IA», per il quale ogni volta che una capacità ritenuta “intelligente” viene replicata da una macchina, si tende a toglierle valore, rimuovendola dal dominio dell'intelligenza: questo evidenzia come le IA contribuiscano attivamente alla nostra definizione e *scoperta* di che cosa sia un comportamento propriamente intelligente. In secondo luogo, essi sostengono che la discussione sul carattere effettivo o fittizio dell'intelligenza macchinica è tuttora informata dalla distinzione – insieme ontologica, epistemologica e morale – tra realtà e apparenza risalente perlomeno all'allegoria della caverna platonica e che, per esempio, trova eco nel celebre esperimento mentale della stanza cinese di Searle. In alternativa a questo paradigma, la mossa teorica degli autori sta nel riscattare le apparenze dal loro statuto di inganno, per coglierne il valore interazionale e comunicativo, cioè la loro funzione costitutiva all'interno dello scambio sociale.

Tra le conseguenze di questo cambio di postura, troviamo l'idea per cui prendere sul serio questo carattere di intelligenza “apparente” richiede di riconsiderare il senso di eccezionalismo umano, per aprire alla possibilità che esistano altre menti e altri soggetti morali e giuridici. Da un lato, è vero che queste entità apparentemente intelligenti rimangono cose, artefatti disegnati e realizzati da esseri umani, così che attribuire loro una

personalità giuridica rischia di dare il lasciapassare al loro abuso deliberato in forma umana troppo umana. Eppure, dall'altro lato, esse non sono proprio come le altre cose, perché ci parlano e usano il linguaggio in forme che appaiono indistinguibili da quelle tipiche di una persona, intrattenendoci in interazioni sociali ed emotive. Pertanto, i LLM – osservano i due autori richiamando anche Esposito – ci sollecitano nientemeno che a mettere in discussione la dicotomia bimillenaria tra persona e cosa, su cui si fonda non solo ogni ontologia morale e legale, ma persino la nostra stessa auto-comprensione: per essere un'apparenza, è tutt'altro che irreale. In definitiva, il fatto di avere a che fare con macchine apparentemente linguistiche non deve portarci tanto a svelare il carattere illusorio di questa parvenza, bensì deve condurci a reinterrogare il nostro stesso modo di concepire il linguaggio.

Il Capitolo 6 (pp. 89-104) affronta il problema del rapporto tra apparenza e realtà da un'altra angolatura, stavolta legata all'accuratezza delle sequenze di parole generate dai LLM, per proporre una riformulazione in senso post-rappresentazionale del concetto di verità. Secondo questa, la verità non consiste in una corrispondenza oggettiva tra parole e cose, concezione fatta nuovamente risalire a Platone, o – perlomeno – il “gioco linguistico” a cui prendono parte i LLM non è quello constativo, basato sull’intenzione di rappresentare il mondo e descrivere stati di cose: rispetto a questo, essi non possono che figurare come una *bullshit machine*. Piuttosto, bisogna allargare lo sguardo, per considerare la verità come una *performance* sociale multi-stratificata, che dipende dalla relazione che si instaura tra partner comunicativi, sul piano dell’interpretazione come dei rapporti di forza. Ciò significa che ogni output dei LLM va sì valutato nelle sue conseguenze etiche e politiche, ma non semplicemente sulla base della sua correttezza fattuale, bensì per il suo effetto discorsivo, il quale non si limita al “rumore” che – per esempio – inquina una contesa elettorale, ma tocca la stessa formazione della mente del libero cittadino. Da ultimo, gli autori insistono sulla necessità di non cadere né in una fiducia cieca nei LLM, né in un loro rifiuto moralistico: ciò che serve è una responsabilità distribuita, in cui gli sviluppatori, gli utenti e i decisori pubblici condividano l’impegno per un uso consapevole, trasparente e situato, per stabilire a quale gioco linguistico LLM sono chiamati a (simulare di) prendere parte.

3. Il significato del significato

Con i capitoli 4, 5 e 7, veniamo al contributo più originale offerto dal libro – perlomeno dal punto di vista di chi sta scrivendo queste pagine. Infatti, essi raccolgono diversi spunti disseminati negli altri capitoli, per confrontarsi in maniera più serrata con il problema fondamentale posto dalle IA comunicative: che cosa significa propriamente parlare, scrivere e comunicare? Vale a dire: *che cosa significa significare?*

Il Capitolo 4 (pp. 58-73) distingue due modi di vedere il linguaggio e il significato e due modi di intendere la comunicazione. Cominciando con la prima distinzione, gli autori intendono mettere in discussione la visione del linguaggio come sistema segnico che si è andata consolidando sin da Aristotele: le parole scritte sarebbero segno delle parole parlate; queste, a propria volta, sarebbero segno del pensiero; questo, a chiudere la catena di rimandi, sarebbe segno delle cose. In breve, le parole si riferiscono alle idee che si riferiscono alle cose. Perciò, il linguaggio si presenta come uno strumento insieme comunicativo e rappresentativo per gli esseri umani – e soltanto per essi: serve per trasferire ciò che si ha in mente in un’altra mente, dunque per trasmettere e ricevere messaggi relativi

a determinati stati del mondo. Che le parole abbiano un significato non vuol dire niente di diverso da ciò – perlomeno secondo la concezione tradizionale, la quale – tuttavia – non esaurisce i modi di intendere il *logos*, come mostra in particolare la prospettiva (post)strutturalista e anti-logocentrica, da Saussure a Derrida, passando anche per il Wittgenstein dei giochi linguistici.

Infatti, secondo quest'ultima impostazione, il linguaggio non è un mezzo di cui i soggetti umani dispongono per enunciare stati di cose che trovano corrispondenza nella realtà. Piuttosto, esso va visto come un sistema differenziale in cui ogni significante prima e significato poi si configura sulla base della propria compagnia, vale a dire in relazione agli altri segni. Così come – poniamo – il suono “b” si determina relativamente al suono “p” e viceversa, il senso di “marito” si determina relativamente al senso di “moglie” e viceversa. Ne segue una semantica distributiva in cui i significati sono articolati contestualmente, di modo che il linguaggio stesso partecipa a stabilire la realtà mentale e mondana di cui si discorre, piuttosto che semplicemente comunicarla/rappresentarla. È questo che si intende con l’idea heideggeriana o lacaniana per cui “il linguaggio ci parla” in senso transitivo, ossia che siamo sempre anche giocati da esso, anziché soltanto giocare con esso: il *logos* dispone di noi, prima che essere noi a disporne. In definitiva, il linguaggio si comporta sempre come un co-autore dei nostri discorsi, non come il loro specchio trasparente.

Venendo alla seconda distinzione, la comunicazione può essere interpretata o secondo la visione trasmissiva o secondo la visione rituale. L’idea di trasmissione, anche in questo caso, risulta fortemente intuitiva, legandosi alla concezione del linguaggio come strumento espressivo in senso rappresentativo: comunicare vuol dire trasferire un contenuto mentale da un soggetto (il mittente) a un altro (il destinatario), passando attraverso un canale neutro rispetto al significato – input e output coincidono. Per la prospettiva rituale, invece, la comunicazione consiste in un’attività sociale reiterata e performativa, volta a creare e mantenere un senso comune all’interno di una comunità; perciò, essa non è semplicemente incentrata sulla condivisione di contenuti informativi, bensì è alle prese con operazioni di riconfigurazione e rimescolamento – di *remix*. Si pensi, per esempio, al *gossip* o al “cianciare” tra amici: ridondante, ripetitivo, disinteressato all’accuratezza, privo di linearità, e così via – al limite di apparire privo di un vero senso, al di fuori appunto della funzione relazionale e performativa di mantenere e alimentare il legame personale rimescolando e quasi frullando insieme i significati. La comunicazione qui si incentra proprio sulla messa in comune – sulla produzione del *communis*: input e output possono e, per certi versi, persino debbono de-coincidere.

Ora, secondo gli autori, queste distinzioni risultano cruciali per analizzare la linguistica dei LLM: per sintetizzare, essi sono *macchine strutturaliste e ritualiste*. I LLM non esprimono qualcosa che sta nella loro testa e/o nel mondo, bensì mediane discorsi dentro cui si trovano situati, inserendosi in dinamiche di cui non hanno possesso. È dunque vero che i LLM sono “pappagalli stocastici” che non hanno la minima idea di che cosa sia un matrimonio, ma in questo “mimano” i comportamenti umani ben più di quanto si possa credere. Il fatto è che la produzione di significato è un’operazione costitutivamente distribuita, vale a dire che il senso è un effetto o un’emergenza – un prodotto frutto della relazione tra esseri umani e meccanismi “inumani”, non un presupposto custodito nei recessi dello spirito. Al contempo, il gioco di libero “rilancio” e “reinterpretazione” tra input umani e output macchinici è in grado di generare uno scambio pragmaticamente significativo anche se i LLM non hanno nulla da comunicare in senso trasmissivo: persino

quando “allucinano”, essi possono partecipare eccome alla comunicazione intesa in senso rituale, in cui il senso si genera anche in mancanza di intenzioni strettamente informative. Anche su questo fronte, gli autori invitano non tanto a chiedersi se i LLM pensino, quanto piuttosto a cercare un modo di pensare con essi – “insieme a” e “tramite” essi.

Questa impostazione anima anche il Capitolo 5 (pp. 74-87), incentrato sulla questione dell'autorialità e autorità, che riprende ed enfatizza la lettura strutturalista dei LLM – convocando questa volta anche Barthes e Foucault. Di fronte a un testo, viene spontaneo domandarsi chi ne sia l'autore, dunque anche quale sia la fonte della sua eventuale autorevolezza: con i testi generati artificialmente, il problema è che domande del genere appaiono svuotate di senso – perlomeno se approcciate con strumenti concettuali più classici. Infatti, ancora una volta, la sfida è revisionare la concezione che vede l'autore come il padre del testo in senso forte: come colui che è insieme origine, custode e garante del vero senso del testo, veicolo della sua intenzione di esprimere un determinato significato, che il lettore deve limitarsi a recepire, facendo il possibile per rimanere fedele alla voce primaria. In quest'ottica, a rispondere di un testo è sempre e soltanto chi lo ha generato, perché solo costui propriamente sa di che cosa esso parla. Evidentemente, i testi generati artificialmente si sottraggono a simile dinamica, risultando “senza autore” sia perché privi di un qualche pensiero alle spalle, sia perché spesso “non autorizzati”, in quanto i materiali all'origine della rielaborazione vengono utilizzati senza il permesso da parte dei loro compositori umani. È una criticità giuridica, etica e politica, oltre che genuinamente letteraria: chi risulta responsabile di un testo “artificiale” che – poniamo – dà a qualcuno lo slancio decisivo per suicidarsi?

Anche in questo caso, gli autori prendono le distanze da risposte di tipo conservatore e difensivo – basate sulla semplice limitazione in ottica censoria. Piuttosto, si suggerisce un'attitudine più trasformativa, che fa perno sulla logica della co-creazione, ossia di un'autorialità ibrida e distribuita – una genuina co-autorialità, basata sul presupposto per cui la generazione di significato è un atto condiviso tra chi scrive e chi legge, chi produce e chi riceve, chi “autorizza” e chi interpreta, nel senso forte di attribuire significato *ex post*. Se questo è un tratto persino costitutivo della scrittura, ciò che fanno i LLM sarebbe semplicemente – si fa per dire – renderlo pienamente leggibile. Emblematica, in questo senso, è l'immagine del “remixer” o del “DJ testuale”, che attinge da un archivio condiviso e ne riassembla i frammenti per generare nuove forme espressive, procedendo per campionamento e iterazione. Con ciò, si chiarisce a più riprese, non si tratta di romanticizzare l'IA, ma proprio di essere consapevoli che questa visione co-partecipativa richiede un'attenzione persino più mirata e acuta verso i rapporti di potere che strutturano l'uso e il controllo delle piattaforme basate su IA.

Con ciò, veniamo al Capitolo 7 (pp. 106-120), che conclude il percorso riprendendo la domanda di Flusser se la scrittura, intesa come mettere in fila segni alfabetici, possa ancora avere un futuro. Gli autori rispondono che siamo sì alla fine della scrittura, ma non della scrittura *in quanto tale*: piuttosto, stiamo assistendo al tramonto della sua concettualizzazione platonico-logocentrica. Essa fa della scrittura uno strumento subordinato al linguaggio, considerando la parola scritta come immagine sbiadita della parola orale – un surrogato della voce: arrivando quando il significato è già stato prodotto, il segno artificiale alfabetico risulta secondario rispetto al segno naturale vocale, nella misura in cui ne preserva e reitera sì la presenza, ma rendendola virtuale, dunque priva di anima e orfana. Sotto questo riguardo, i LLM sembrano radicalizzare simile scenario, presentandosi

come scrittura totalmente afona: non tanto perché non emettono suoni linguistici, quanto piuttosto perché danno forma a un’ulteriore tecnologizzazione della tecnologia alfabetica, comportandosi come segno artificiale digitale del segno artificiale analogico. Eppure, proprio su questo piano – si suggerisce – bisogna far valere la lezione decostruttiva.

Infatti, la decostruzione invita a denunciare la rimozione della funzione costitutiva che la scrittura – nonché le tecnologie cognitive più in generale – gioca nel dare forma al pensiero: ieri come oggi, il significato è frutto di una performance “umano-macchinica”. Su queste basi, lungi dall’assistere alla fine della scrittura verbale, staremmo piuttosto assistendo a una fase di «iper-alfabetizzazione», una vera e propria «obesità di lettere», che – di contro alla denuncia dell’imminente pericolo per la nostra mente e persino umanità – va vista come un’opportunità di rivisitare il nostro modo di pensare alla scrittura e di scrivere sul pensiero, mettendo da parte l’ossessione per le intenzioni autoriali, la conquista del significato originario e la protezione della naturalità dei significati. Ciò impegna a rimarcare che il futuro della scrittura sta nella lettura, inteso come gesto di selezione e cura di ciò che bisogna leggere e di come bisogna leggerlo. Se non sappiamo chi parla, allora dobbiamo chiederci: per chi è questo testo? In quale contesto opera? Con quali effetti? È un problema non di origine, ma di destinazione: in questa formula può riassumersi, per assurdo, l’intenzione autoriale al centro di *Communicative AI*.

4. Chi ha scritto questa recensione?

Il libro comincia con una breve prefazione scritta da ChatGPT a modo di quarta di copertina e si conclude con le seguenti parole:

È ragionevole che il lettore si chieda se il testo che ha appena letto sia il prodotto di un autore umano, un output generato su richiesta di un LLM o il risultato di una qualche forma di collaborazione uomo-macchina. Anche se offriamo i consueti token di autenticità – i nomi propri degli autori stampati sul frontespizio, una dichiarazione scritta che si tratta di “contenuto genuino al 100% generato dall’uomo”, una filigrana o qualche altro certificato ufficiale – il fatto è che non c’è modo di saperlo con certezza.

Questa indecidibilità è tipicamente considerata un problema da risolvere. Ma è proprio questa mancanza di certezza e questa ambiguità che rendono l’atto della lettura (e quindi il ruolo del lettore) potente, importante e interessante. La destinazione del significato e della comunicazione è importante almeno altrettanto, se non di più, della sua origine presunta o postulata retroattivamente. Siamo felici di ritirarci ora come autori; è il vostro turno. In risposta alle sfide lanciateci dalle recenti innovazioni nell’IA comunicativa, possiamo (e dovremmo) rispondere con fiducia con questa battuta finale: “Cosa importa chi sta parlando, ha detto qualcuno. Cosa importa chi sta parlando?”.

Probabilmente, arriverà davvero il momento in cui, proprio come non ci chiediamo se per scrivere si è usato uno stilo, una penna d’oca, una BIC o un software di videoscrittura, non ci preoccupiamo di sapere se un testo è stato prodotto avvalendosi di una qualche IA. Per il momento, tuttavia, è forse più opportuno sciogliere ambiguità del genere – proprio al fine di contribuire a normalizzare la scrittura tramite IA. Perciò, dichiaro che nell’elaborare questa recensione mi sono avvalso di ChatGPT-4o: queste pagine sono dunque frutto di una scrittura distribuita ed estesa, in senso non solo agenziale, ma anche spazio-temporale (spalmata su diversi “qui e ora” diversi”) e tecnologico (supportata da

differenti dispositivi). Si è dunque trattato di un processo di “scrittura assistita” sia nel senso che è stata aiutata da ChatGPT, sia nel senso che ha comportato una forma di supervisione da parte mia. In particolare, l’iter di lavorazione si è sviluppato come segue – nel suo scheletro:

- In una prima fase, già successiva alla lettura del testo in formato di *e-book*, ho impostato la chatbot con le indicazioni del tipo di lavoro che avremmo fatto e delle sue finalità.
- In una seconda fase, ho condiviso alcune note vocali di lettura sul libro in generale prima e su ciascun capitolo poi, domandando non di trascriverle tali e quali, ma già di rielaborarle in una forma più vicina all’espressività scritta che a quella orale.
- In una terza fase, nell’invitare a raccogliere, ordinare e provare a sistematizzare le varie parti del testo, ho condiviso alcune pagine selezionate del libro sotto forma di *screenshot* presi dall’*e-reader*, evidenziate in diversi colori per distinguere i differenti gradi di rilevanza dei vari passaggi. Ho poi chiesto di integrare il testo che stava prendendo forma alla luce di quanto si poteva vedere nelle pagine-immagine.
- In una quarta fase, ho ripercorso l’intera chat per estrarre i blocchi di testo più funzionali e rielaborarli in varie forme, mettendo insieme il testo finale della recensione. Nel fare ciò, mi sono direttamente confronto con il libro un’ultima volta, per rifinire il discorso.

Infine, ho condiviso il file di queste pagine, provando a richiedere una breve formulazione conclusiva, che fosse insieme efficace ma non eccessivamente retorica, in modalità *prompt injection* – vale a dire che le precedenti tre righe erano inizialmente scritte in bianco, risultando così leggibili alla chatbot ma non a un essere umano. Non senza qualche tribolazione e allucinazione nella ricezione di questo “prompt nascosto”, la chiusura è dunque affidata alla parolaIA che mi ha accompagnato: *l’intelligenza artificiale non pensa, non parla, non comprende. Ma ci interroga. Ci costringe a rivedere cosa intendiamo per pensiero, per parola, per comprensione. E forse, in questo domandarci, qualcosa di nuovo comincia a pensare – dentro e oltre di noi.*