

DANIEL RAFFINI

INTELLIGENZA ARTIFICIALE E SCRITTURA LETTERARIA:  
LO SPETTRO DELL'AUTORE E LE NUOVE FORME DELL'AUTORIALITÀ

**Sinossi:** Il contributo analizza l'utilizzo dell'intelligenza artificiale generativa nella scrittura letteraria, con particolare attenzione alla ridefinizione del concetto di autorialità. Dopo aver ripercorso l'evoluzione storica dei sistemi di generazione automatica di testo, dai primi modelli a regole fino agli attuali *Large Language Models*, il saggio esplora le modalità di impiego dell'IA nella creazione di opere letterarie e i diversi gradi di interazione tra essere umano e macchina. Viene distinta la nozione di autore da quella di autorialità, privilegiando quest'ultima per lo studio dei testi generati con IA. Attraverso l'analisi di casi studio come "The Day a Computer Writes a Novel", *I the Road* e *Non siamo mai stati sulla Terra*, si evidenziano le molteplici forme di autorialità emergenti, che spaziano dall'autorialità combinatoria e regolatoria fino a modalità co-autoriali e dialogiche. Il saggio indaga inoltre, da un punto di vista narratologico, la costruzione linguistica della soggettività fittizia nei testi in cui i sistemi di IA assumono la parola in prima persona. L'analisi congiunta dei processi di generazione, delle dinamiche di interazione e dell'istanza autoriale permette infine di delineare la complessità dell'autorialità nei testi letterari prodotti tramite IA.

**Parole chiave:** autorialità, letteratura, intelligenza artificiale, creatività.

**L'evoluzione dei sistemi di scrittura automatica**

La storia dell'Intelligenza Artificiale (IA) affonda le sue radici negli anni Cinquanta del Novecento e si sviluppa parallelamente ai progressi nei supporti hardware e software della seconda metà del secolo. Fin dall'inizio, gli studiosi che si dedicarono a questo ambito di ricerca compresero la centralità del linguaggio: per simulare le caratteristiche dell'intelligenza umana, un sistema informatico doveva innanzitutto essere in grado di comprendere e produrre linguaggio naturale. Questa consapevolezza nasce dall'ipotesi secondo cui è proprio attraverso il linguaggio che l'essere umano costruisce il pensiero e, con esso, le categorie necessarie per comprendere il mondo e comunicare (Turing; McCarthy et al.). Da tale esigenza prende forma l'area di studi nota come *Natural Language Processing* (NLP), che si occupa della progettazione di sistemi in grado di comprendere, manipolare e generare linguaggio umano. All'interno di questo campo, nel corso dei decenni, sono stati sviluppati diversi sistemi per la generazione automatica di testi.

Già a partire dagli anni Sessanta, vengono sviluppati i cosiddetti sistemi a regole, ossia dei modelli che seguono istruzioni predefinite per combinare parole e frasi. A questa categoria appartiene ELIZA, chatbot sviluppato nel 1966 da Joseph Weizenbaum a fini psicoterapeutici, che ha aperto la strada agli attuali assistenti virtuali (Berry). In ambito letterario, l'idea di sistema a regole trova una coincidenza in quegli anni con quella di letteratura

combinatoria, come dimostra il caso di Nanni Balestrini, che utilizza sistemi a regole per la generazione delle poesie del ciclo *Tape Mark* (1961) e del romanzo *Tristano* (1966). In questo contesto si sviluppano anche le *story grammars*, schemi narrativi utilizzati per organizzare e generare testi, in grado di modellare la struttura logica e sequenziale delle storie. Basandosi su regole e schemi narrativi, le *story grammars* sono state utilizzate principalmente per rappresentare e comprendere la narrativa, ma in alcuni il loro impiego è stato esteso alla generazione di trame (Pérez y Pérez).

Negli anni Ottanta e Novanta trovano applicazione computazionale le catene di Markov, teorizzate agli inizi del secolo. Si tratta di modelli statistici utilizzati per la generazione del linguaggio, che si basano sulla probabilità di transizione tra parole per produrre nuovo testo. Questi modelli si rivelano particolarmente efficaci nell'imitazione stilistica e per questo, nel 1984, vengono impiegati da Hugh Kenner e Joseph O'Rourke per *Travesty Generator*, un programma capace di generare parodie di testi celebri, come le opere di Shakespeare o la Bibbia, creando nuove frasi sulla base delle probabilità di transizione tra parole ricavate dai testi originali (Chen). A partire dagli anni Novanta si sviluppano i modelli di *topic modelling*, che, pur non generando testi in modo diretto, contribuiscono all'organizzazione di temi e argomenti, trovando applicazione anche nella sintesi o nella generazione di testi tematici. Nello stesso periodo si diffondono i modelli statistici basati sugli *n-grammi*, che utilizzano la statistica per prevedere la probabilità di sequenze di parole. Impiegati in numerosi sistemi di NLP, questi modelli rivestono un ruolo importante in diversi ambiti della generazione automatica di testo, tra cui i sistemi di traduzione automatica (Marino) e, in campo filologico, le tecniche di attribuzione autoriale (Khmelev e Tweedie, Keselj et al.).

Una svolta decisiva è rappresentata dagli avanzamenti dei sistemi *data-driven*. A partire dagli anni Dieci del Duemila si sviluppano le *Recurrent Neural Network* (RNN) e le *Long Short-Term Memory* (LSTM), i primi modelli neurali in grado di gestire sequenze per la generazione di testi. Questi sistemi si rivelano particolarmente efficaci nella comprensione e produzione di testi lunghi, dimostrando una notevole duttilità anche nell'ambito della creazione letteraria. Ne è un esempio l'esperimento condotto da Karpathy, che ha addestrato una rete LSTM su testi di Shakespeare e ha utilizzato una RNN per generare sequenze di parole in grado di imitare struttura e sintassi dell'originale mantenendo coerenza interna. Le RNN vengono impiegate non solo a fini di imitazione stilistica, ma anche in processi generativi più complessi, come si vedrà nel caso di *I the Road* di Ross Goodwin. Nel 2014 vengono introdotti i modelli *Seq2Seq* con meccanismi di attenzione, una delle prime grandi rivoluzioni nella generazione automatica di linguaggio, in quanto capaci di produrre nuovo testo mantenendo il contesto e generando risposte

coerenti anche a prompt complessi. Di particolare rilievo, sempre nel 2014, è l'invenzione delle *Generative Adversarial Networks* (GAN), un tipo di rete neurale composta da due modelli in competizione: il generatore, che produce dati sintetici—come immagini o testi—simili a quelli reali, e il discriminatore, che valuta se tali dati siano autentici o generati, contribuendo così a migliorare progressivamente la qualità degli output del generatore. Le GAN nascono inizialmente per la generazione di immagini, ma successivamente vengono adattate anche alla produzione di testi (Shahriar).

L'ultima tappa di questo percorso è rappresentata dall'introduzione, nel 2017, dell'architettura *Transformer*, che è alla base dei moderni modelli linguistici di grandi dimensioni, i *Large Language Models* (LLM). Con gli LLM migliora ulteriormente la generazione di linguaggio naturale, che raggiunge livelli di coerenza anche su testi lunghi; questi sistemi rappresentano oggi le tecnologie più avanzate in ambito di IA generativa, seppur con persistenti problematiche legate all'affidabilità e alla veridicità dei contenuti generati, soprattutto all'interno di domini specialistici. A partire dalla fine del 2022—in seguito al rilascio di ChatGPT—gli LLM diventano strumenti accessibili al grande pubblico. Oggi sono utilizzati su larga scala per la creazione di contenuti online e trovano impiego, con esiti variabili, anche nella generazione di testi letterari (Franceschelli e Musolesi).

### **Scrittura letteraria e intelligenza artificiale nel panorama contemporaneo**

Nonostante gli avanzamenti dai primi sistemi a regole fino ai modelli attuali, ad oggi la generazione totalmente automatica di testi per mezzo di IA non si dimostra soddisfacente per quanto riguarda l'ambito letterario. Se la produzione di contenuti testuali rispecchia le caratteristiche di coerenza e coesione necessarie a livello linguistico per considerare una sequenza di parole come un atto comunicativo, tuttavia la generazione automatica non sembra ancora in grado di ottenere un effetto di letterarietà senza che l'essere umano interagisca in qualche modo con il sistema o con il testo. Lo dimostrano i molti libri "written by AI" diffusi attraverso piattaforme di autopubblicazione (Cabezas-Clavijo et al.). Si tratta nella maggior parte dei casi di testi generati in completa autonomia da un LLM—quasi sempre ChatGPT—in risposta al prompt generico di un utente. Questi testi attualmente non sembrano essere in grado di superare un vaglio critico più attento per quanto riguarda le strutture narrative, lo stile, la coerenza interna e la polisemia. Viene confermata, nonostante i progressi degli ultimi anni, la tendenza verso trame e forme linguistico-retoriche stereotipizzate e la difficoltà nel produrre trame coerenti quando si eccede la misura breve. La lettura dei testi lunghi generati dagli LLM rivela spesso buchi di trama, incoerenze rispetto ai fatti raccontati, problemi nella temporalità, sovrabbondanza di personaggi

ed episodi slegati dall'azione principale; tutti elementi che rendono la narrazione poco efficace senza risultare in ulteriori aperture di significato. Diverso è il caso in cui il processo di scrittura avviene attraverso una forma di collaborazione tra sistema automatico ed essere umano. Mirko Farina, Witold Pedrycz e Andrea Lavazza propongono una via intermedia tra la concezione dell'IA come pappagallo stocastico alieno alla creatività e quella della macchina come portatrice di creatività in maniera autonoma, sostenendo che una nuova forma di creatività può emergere dall'interazione umana con i sistemi. Il procedimento collaborativo, oltre ad aumentare la qualità del testo rispetto alla generazione autonoma, delinea nuove forme di autorialità, che contribuiscono a una ridefinizione delle categorie di autore e lettore e a un nuovo approccio alla creazione del significato. Prima di approfondire la questione dell'autorialità, è opportuno però soffermarsi sulle possibili tipologie di utilizzo dell'IA per la scrittura di testi letterari.

In primo luogo, occorre notare che l'impiego dell'IA può essere nascosto o esibito: nel primo caso rientrano coloro che si avvalgono dell'IA generativa senza dichiararlo esplicitamente, mentre nel secondo si collocano i casi in cui l'utilizzo viene dichiarato o addirittura esibito. Su questo punto si gioca una questione etica rilevante, sottolineata da Heidi McKee e James Porter nel 2020, che individuano nella trasparenza sull'uso dell'IA e in un approccio critico ai dati i due principi etici fondanti della scrittura generativa. La necessità della trasparenza è rafforzata dal fatto che i sistemi di scrittura risultano a volte carenti nel ricreare i codici impliciti di una comunicazione etica. Considerare la trasparenza sull'utilizzo dell'IA come un principio etico risulta utile anche nell'ottica delle normative sull'IA e sul diritto d'autore. A questo riguardo si deve però notare che dal punto di vista della teoria letteraria—che qui si vuole adottare—il principio etico risulta sufficiente; mentre la prospettiva giuridica risulta più interessante per l'ambito editoriale e filologico.<sup>1</sup>

Le modalità di utilizzo dell'IA per la scrittura letteraria possono essere molto diverse tra di loro. Oltre che delle differenti tecnologie di IA, di cui si è parlato nel paragrafo precedente, occorre tener conto anche del tipo di utilizzo. Il ricorso allo strumento può avvenire in diverse fasi del processo di creazione e con scopi diversi:

1. Fase di ricerca. Chi vuole scrivere un'opera letteraria può utilizzare l'IA in una fase precedente alla scrittura vera e propria, per documentarsi sulla materia che intende affrontare, o per svolgere attività di raccolta e strutturazione delle idee.

---

<sup>1</sup> Per tale ragione in questa sede non vengono approfondite le questioni riguardanti la regolamentazione dell'IA generativa e il diritto d'autore, per le quali si può tuttavia far riferimento a una nutrita bibliografia (per esempio, Grieco; Zanardi).

2. Fase di scrittura. In questa fase le interazioni possibili sono varie, come si vedrà nei casi analizzati in seguito. Per ora anticipiamo che chi usa un sistema di IA per la scrittura può scegliere diversi gradi di automazione, da una modalità completamente automatica fino a una forma di scrittura alternata, in cui essere umano e macchina si avvicinano.
3. Fase di editing. In questa fase, i testi generati dal sistema possono essere revisionati dall'essere umano, che interagisce con essi fino a giungere a una forma ritenuta soddisfacente e degna di essere licenziata; viceversa, lo scrittore può utilizzare l'IA per rivedere i propri testi, utilizzando il sistema come un editor.

Le tre fasi non sono vicendevolmente necessarie né interdipendenti: l'IA può essere utilizzata anche solamente in una delle fasi della scrittura, determinando un diverso grado di nascondimento o esibizione dello strumento. Per quanto riguarda l'utilizzo dell'IA nelle fasi di ricerca o di editing, uno studio basato su interviste anonime conferma che scrittori e scrittrici contemporanei la impiegano più spesso di quanto si possa immaginare e, soprattutto, che la considerano più affidabile di un consigliere umano (Gero et al.). Ciò dimostra la tendenza a non dichiarare l'utilizzo dell'IA per la ricerca e l'editing; mentre l'utilizzo nella fase di scrittura porta con sé un'attitudine sperimentale, fondata proprio sull'utilizzo dell'IA, che dunque sarà probabilmente dichiarato. A ciò si aggiunge la necessità etica di trasparenza, che suggerirebbe di esplicitare l'uso dell'IA in ogni fase.

Il ruolo dell'essere umano non è limitato all'inserimento delle richieste, ma può intervenire nella delineazione dell'output attraverso la creazione di un contesto, soprattutto nel caso degli LLM. Quando parliamo di contesto in relazione agli LLM, ci riferiamo alle informazioni e al quadro di riferimento che il modello utilizza per generare risposte rilevanti e coerenti rispetto a una richiesta dell'utente.<sup>2</sup> Un primo elemento di contesto, ad esempio, è il *system prompt*, il set di istruzioni fornite dagli sviluppatori, che definisce il tono e le regole di comportamento che il modello deve seguire. Se il prompt di sistema

---

<sup>2</sup> Devo questa nozione a Gino Roncaglia, che ne ha parlato in un intervento dal titolo "IA: questioni di contesto", tenuto in occasione del convegno *Letteratura e intelligenza artificiale: un dialogo interdisciplinare* (Sapienza Università di Roma, 16 settembre 2024). In un articolo del 2023, Roncaglia definisce il contesto come "un insieme di informazioni fornite al sistema dopo il suo addestramento, con la funzione di indirizzare l'output nel senso più adeguato rispetto agli obiettivi specifici che si vogliono raggiungere" (134). Recentemente l'idea ha acquistato popolarità nel dibattito sull'IA, arrivando a proporre di sostituire l'idea di *prompt engineering* con quella di *context engineering*, come chiarisce Philipp Schmid.

spesso non vengono divulgati dai produttori e non sono modificabili, nel caso di altri elementi contestuali—primo fra tutti la cosiddetta *context window*, ossia la singola conversazione—l'utente può intervenire inducendo nel sistema modalità di comportamento e di risposta che ricalchino lo stile e il tono della richiesta. Tutti questi elementi—come si vedrà—possono essere utilizzati per ottenere effetti creativi, attivando un processo di partecipazione alla produzione del testo che è altra cosa rispetto alla totale delega della scrittura al sistema.

### **Autore e autorialità**

Il confronto con i testi generati con l'ausilio di sistemi di IA ci pone di fronte alla necessità di distinguere tra il concetto di autore e quello di autorialità: intenderemo l'autore come l'essere umano reale e storico che si occupa della stesura del testo e per autorialità una funzione del discorso, che corrisponde a “un modo di esistenza, di circolazione e di funzionamento di certi discorsi all'interno di una società” (Foucault 9). La distinzione diventa particolarmente rilevante quando si parla di IA, soprattutto considerando l'autorialità come “un orizzonte d'attesa del prodotto letterario” (Medaglia 42), legata al contesto di fruizione più che a quello di produzione.<sup>3</sup> Se nella letteratura moderna occidentale si è sempre percepita l'autorialità come una funzione legata a un autore in quanto essere umano, nel caso dei testi generati con sistemi di IA ci troviamo di fronte a uno strumento che si sostituisce all'essere umano nella scrittura. Anche nei casi come il falso, la pseudonimia, o nelle scritture anonime, multiautoriali o collaborative, non si arrivava a dubitare della presenza di uno o più esseri umani incaricati della scrittura dell'opera; nel caso dell'IA invece l'autore inteso come essere umano che scrive materialmente e interamente l'opera viene meno, ciò che rimane è la funzione autoriale in quanto istanza interna al testo e presente nell'aspettativa del lettore. Cercare di definire chi sia l'autore di questi testi conduce a una serie di paradossi che impediscono di formulare una risposta convincente. Per questo motivo sembra più utile parlare di autorialità slegata dall'autore, dal momento che, pur nelle complicazioni derivanti dal processo di interazione uomo-macchina, l'istanza autoriale è percepita dal fruitore anche nei testi scritti completamente in maniera automatica tramite sistemi di IA.

Per analizzare l'autorialità dell'IA si è spesso fatto riferimento all'idea barthesiana della morte dell'autore (per esempio Gefen). Si può concordare

---

<sup>3</sup> Con ciò non si intende che questa sia la prospettiva da adottare sempre nell'approcciarsi all'autorialità dei testi letterari, ma solamente che lo spostamento dell'atto di significazione sul momento della lettura pare particolarmente utile per affrontare la questione della scrittura tramite IA. Tale prospettiva è sostenuta già da Calvino alla fine degli anni Sessanta (si veda Raffini) ed è stata riaffermata di recente da Natale e Henrickson.

sul fatto che scrivere tramite questi sistemi conferma l'idea post-strutturalista del testo come oggetto indipendente dal suo creatore materiale, o—usando le parole dello stesso Barthes—l'idea che sia “il linguaggio a parlare, non l'autore” (52). Tuttavia, la morte dell'autore di cui parlano i teorici all'inizio degli anni Settanta è qualcosa di diverso dalla sparizione dell'autore—inteso come uno o più esseri umani che si incaricano materialmente della scrittura—a cui si assiste nel caso dell'IA. Ciò che Barthes auspicava era nella pratica il decadimento del mito dell'autore, mentre nel caso dei testi scritti tramite sistemi di IA dobbiamo constatare come l'autore venga meno e rimanga solamente l'istanza di autorialità, quale superstite di una dicotomia storicamente stratificata che definisce l'atto linguistico come un messaggio procedente da un mittente a un destinatario. Leah Henrickson ha parlato di questo fenomeno come di una rottura del contratto ermeneutico.

Il vuoto lasciato dalla sparizione dell'autore non è un fatto privo di conseguenze. Ogni atto di lettura presuppone, infatti, l'intervento all'origine di un essere umano che scrive: l'assuefazione alla presenza dell'autore in un testo è insita nel nostro modo di fruire un prodotto letterario, perlomeno in epoca post-romantica. È per questo che quando ci troviamo di fronte a testi generati da IA si verifica una sorta di sindrome dell'arto mancante: l'abitudine alla presenza di un certo elemento risulta talmente stratificata nell'inconscio personale e collettivo da continuare a farsi sentire anche nel momento in cui viene a mancare. In questi testi l'autore è assente fisicamente ma presente in forma spettrale: tale sensazione emerge chiaramente da alcuni elementi paratestuali delle opere letterarie scritte con IA. È il caso, ad esempio, del racconto “The Day a Computer Writes a Novel”, che fece notizia nel 2015 per essersi classificato come finalista all'Hoshi Shin'ichi Prize, il maggior premio letterario di fantascienza giapponese, che proprio quell'anno era stato aperto per la prima volta ad “autori non umani”.<sup>4</sup> Il testo è il risultato del lavoro dell'equipe di ricercatori del Kimagure Artificial Intelligence Writer Project, ma è stato presentato sotto il nome dell'autore fittizio Yurei Raita. Il bisogno di un autore fittizio rispecchia la necessità, espressa da Foucault, di esprimere una funzione autoriale anche laddove la persona reale dell'autore venga ridimensionata o venga meno, come nel caso dell'IA. Se l'attribuzione di un nome fittizio risponde a questa esigenza, gli ideatori del progetto attraverso la scelta del nome rimandano alla componente spettrale che l'autore assume nel testo generato artificialmente. Yurei Raita, infatti, riporta al suo interno la radice giapponese “yurei”, che si collega alla sfera semantica del “fantasma”, mentre “raita” può essere tradotto letteralmente come “grande tuono”, ma ricorda anche la pronuncia

---

<sup>4</sup> La definizione è ripresa dal bando del concorso.

della parola inglese “writer”: ci troviamo di fronte, dunque, a uno scrittore fantasma.<sup>5</sup> La scelta etimologica ricalca la sensazione di assenza lasciata dall'autore nel momento della sua sparizione. Se nella prospettiva del lettore contemporaneo ogni atto di scrittura presuppone un autore, nel caso della scrittura dell'IA l'autore assume la forma di uno spettro inevitabile: non ha un corpo, non possiamo vederlo, ma allo stesso tempo non riusciamo fare a meno di percepire inconsciamente la sua esistenza.

Si tratta—come detto—di un'assuefazione culturale, data da secoli di sedimentazione di una precisa forma di comunicazione, che prevede l'accentramento del concetto di autorialità intorno alla figura dell'autore. Una concezione dell'autorialità che secondo Ross Goodwin rispecchia una mentalità prettamente antropocentrica, come scrive nell'introduzione a *I The Road*: “Authorship is a human concept, which has been applied exclusively to humans throughout the history. And if we cannot agree on the identity of this manuscript's author, that uncertainly may speak more to the anthropocentric nature of our language than the question of authorship it-self” (14). Goodwin propone di allontanarsi dalla visione antropocentrica dell'autorialità e di ripensare il concetto tenendo in conto non più la persona fisica ma il processo di generazione che si crea nell'interazione tra essere umano e strumento tecnologico. L'autorialità, intesa come procedimento di generazione di un prodotto artistico, diventa dunque il concetto da approfondire nel caso dei testi generati tramite IA.

Per parlare di autorialità nei testi scritti tramite IA è stato utilizzato recentemente il concetto di distanza, o di scrittura da lontano. L'idea è stata formulata da Hannes Bajohr nel 2024, che fa interagire in maniera molto efficace il concetto di *causal authorship*, con il quale si misura la distanza tra l'autore umano e il testo finale, e quello di *distributed authorship*, che intende l'autorialità come una rete all'interno della quale trovano spazio agenti umani e non umani. L'idea di *distant writing* è stata ripresa da Luciano Floridi nel 2025, in maniera tuttavia meno convincente. Pur affrontando alcune questioni di interesse, il modello di Floridi appare troppo normativo e distante da casi reali di pratica letteraria, risolvendosi nella sua parte applicativa in un prontuario di scrittura con l'IA, in cui l'essere umano è assimilato a un *prompt designer* e la macchina a un esecutore. In risposta a questa visione, nella sezione che segue propongo l'analisi di casi di studio provenienti della pratica letteraria con lo scopo di mostrare la complessità e diversità delle interazioni possibili con l'IA e le differenti forme di autorialità che ne derivano.

---

<sup>5</sup> *GhostWriter* è anche il nome del software programmato dagli scienziati del Kimagure Project Matsubara Hitoshi e Sato Satoshi per scrivere racconti sullo stile di Hoshi Shin'ichi.

### Forme dell'autorialità

Una volta posta la distinzione tra autore e autorialità, possiamo procedere verso una definizione più precisa di alcune delle forme che può assumere l'autorialità nel caso dell'IA. Piuttosto che di autorialità al singolare, è opportuno parlare di modalità diverse attraverso le quali la funzione autoriale si esprime in questi testi. La tecnica di generazione, il sistema utilizzato e le modalità di interazione tra umano e macchina contribuiscono di volta in volta a definire forme specifiche di autorialità: ci troviamo di fronte a una funzione che si modifica a ogni interazione. Parlare di autorialità nel caso dell'IA vuol dire allora analizzare caso per caso il processo utilizzato per la generazione del testo, mettendo in dialogo scelte umane e modalità di funzionamento dei modelli. Come è emerso dalla breve ricognizione iniziale, esistono diverse tipologie di sistemi in grado di produrre un testo scritto, che possono essere usati singolarmente o combinati tra loro. In generale, è necessario che all'interno del processo sia incluso un modello generativo, in grado di produrre contenuti testuali. Il fatto, poi, che esistano modelli di IA generativa anche per l'elaborazione di immagini, suoni e video determina la possibilità per l'artista di esplorare l'interazione tra linguaggi combinando diversi modelli.

La molteplicità dei sistemi determina una variabilità di metodi nella generazione dei testi. Analizzando questi metodi, è possibile definire la forma assunta dall'autorialità. Come primo caso di studio, prenderemo il racconto "The Day a Computer Writes a Novel". In questo caso, il processo di generazione del testo viene così descritto da Satoshi Sato, uno dei tecnici responsabili del progetto: "We first prepared several components needed for generating stories such as a story grammar and a set of text fragments, and then our system constructed a story (text) automatically using these components, which we submitted without any modification" (32). In questo caso ci troviamo di fronte a una forma di autorialità combinatoria e regolatoria, simile a quelle adottate fin dagli anni Sessanta, in cui una serie di elementi precostituiti viene ricombinata dal sistema seguendo una grammatica narrativa.

Diverso è il caso di *I the Road*, opera ideata da Ross Goodwin, che ha utilizzato sensori posizionati all'interno e all'esterno di un'automobile come dati di partenza per la scrittura. Il metodo in questo caso prevede la combinazione di due reti: una *Convolutional Neural Network* (CNN), per classificare le entità presenti nelle immagini, e una *Recurrent Neural Network* (RNN) per generare il testo. Il sistema è stato poi programmato per cadenzare la narrazione intorno allo scorrere delle ore, ottenendo una scansione temporale lineare del testo che copre i giorni del viaggio. La RNN ha tradotto i dati provenienti dal GPS, dai microfoni e dalle fotocamere collocate sull'automobile durante un viaggio da New York a New Orleans,

sulle orme del viaggio descritto da Kerouac in *On the Road*. I dati provenienti dalle periferiche sono serviti da punto di avvio per la scrittura e il testo è stato generato in tempo reale come risultato dell'interazione tra elementi spaziali, visuali e uditivi. Attraverso questo procedimento, Goodwin ha raggiunto una forma di scrittura ad alto grado di automazione, nella quale il *prompt* solitamente necessario per avviare la generazione del testo è stato sostituito dai dati raccolti nell'ambiente circostante e classificati tramite la CNN. A monte di una generazione testuale tanto automatizzata, c'è la precisa definizione di una serie parametri, nonché l'ideazione di un processo di interazione dei sistemi di IA tra di loro e con le periferiche di input, che è frutto dell'ingegno di Goodwin. L'autorialità si presenta allora, in questo caso, sotto forma di ideazione e processo, in cui la prima è a carico dell'essere umano e il secondo viene demandato ai modelli di IA.

Nell'introduzione al libro—come si è accennato—Goodwin si interroga sul concetto di autorialità, alla ricerca di una definizione del proprio ruolo e di quello dei sistemi di IA impiegati: “As author of the machine, I could describe my role as writer of writer. But whether the machine itself is a writer remains a question that is about relevant as whether a submarine is a swimmer” (14). In copertina, sotto il nome di Goodwin, appare la definizione “writer of writers”; mentre nella nota biografica conclusiva si legge: “Ross Goodwin is not a poet. [...] He defines himself as an artist, a hacker, a code inventor and a gonzo data scientist” (143). Goodwin mette in luce un problema definitorio, che riguarda il ruolo di chi si occupa dell'ideazione dei processi di generazione. A tal proposito, non sembra opportuno in generale riferirsi alla componente umana del processo con il termine “autore”, dal momento che storicamente tale termine ha indicato un essere umano incaricato materialmente e per intero del processo di scrittura di un'opera letteraria; rimane però il problema di come riferirsi alla persona o alle persone che si occupano della definizione del metodo di generazione dei testi di cui ci stiamo occupando. Una proposta potrebbe essere di riferirsi a loro come agli ideatori dell'opera, ma non è da escludere che tale termine si riveli non adeguato in alcuni casi specifici, in cui—come si vedrà nel caso di Tanica—intervengono ulteriori elementi nella scelta del termine.

Riguardo la definizione del ruolo dei sistemi di IA, invece, mi sembra che si possa considerarli come strumenti del processo creativo, in quanti privi di intenzionalità propria e di coscienza. Si tratta di strumenti del tutto peculiari, diversi da quelli finora utilizzati nei processi di scrittura, se non altro per il potere trasformativo che possono esercitare sia sulle categorie teoriche che sui prodotti stessi.<sup>6</sup> Goodwin propone un parallelismo tra le reti

---

<sup>6</sup> Arielli individua le caratteristiche peculiari di quello che definisce come “artificial gaze”, tra cui un “innocent eye”, la mancanza di storicità nella percezione, l'estrazione dell'inconscio dei dati, l'idea di una “quintessence

neurali e la macchina fotografica per la loro capacità di meccanizzare le azioni creative. Il paragone può essere utile se consideriamo la comune necessità di una componente umana, giacché la fotografia per essere un oggetto artistico deve essere espressione dello sguardo specifico del fotografo. Un aspetto su cui riflettere è che nel caso della fotografia la meccanizzazione del processo ha portato alla nascita di una nuova forma d'arte, diversa dalla pittura, un fenomeno che potrebbe avvenire anche nel caso dell'IA qualora emergessero peculiarità stilistiche tanto divergenti come quelle tra pittura e fotografia. Un altro aspetto interessante dell'argomentazione di Goodwin è lo slittamento di ruoli che può derivare dall'utilizzo della tecnologia nel processo creativo:

Technology can shift our roles in creative process. Painters create images; photographers operate cameras, which create images. [...] Writers, whether they operate a pen, a writing ball, or a laptop, have always produced their written work directly, letter by letter or word by word. Typewriters are not cameras for writing. But artificial neural networks might be.

(12)

L'idea della presenza di diversi ruoli nella creazione artistica ci aiuta a concepire un modello non monolitico di autorialità, che si frammenta in diverse fasi e assume contorni non definibili in maniera univoca. Tale dispersione è ben espressa da Goodwin, attraverso l'elencazione e la successiva negazione di ognuno dei possibili autori di *I the Road*:

A manuscript emerged in real time as we drove across the countryside, written by the car itself. Or perhaps the drivers wrote it, using the car as their pen. Or perhaps the machine wrote it, using our traversal of the landscape as its means of coherence. Or perhaps those who wrote the machine's training corpus have written it together, reaching out from their work, out from the past, to influence a new story. Or perhaps I wrote it. But I am less sure than ever about that. All I know for certain about this book's authorship is that I enjoy asking questions about it far more than attempting to answer them.

(12, 14)

Un'ulteriore forma di autorialità emerge nel caso di *Non siamo mai stati sulla Terra*, pubblicato nel 2022 da Rocco Tanica. In copertina si legge, insieme a quello di Tanica, il nome di Out0mat-B13, il sistema GPT-3 utilizzato nella scrittura del libro. Tra quelli analizzati, questo è l'unico caso di utilizzo di un LLM. Il libro di Tanica viene pubblicato subito prima del rilascio di Chat-GPT,

---

machine” e un’immaginazione controfattuale.

quando ancora gli LLM non erano a disposizione del grande pubblico, e mostra un utilizzo ragionato del modello con finalità di sperimentalismo letterario. *Non siamo mai stati sulla Terra* si presenta a livello formale come una serie di racconti riuniti in sezioni tematiche, intervallate da dialoghi tra l'uomo e la macchina che fungono da cornice alle storie raccontate. I dialoghi sono conversazioni reali tra Tanica e il sistema, anche se lo scrittore ha dichiarato l'utilizzo della funzione *rewrite* nei casi di risposta non soddisfacente. La forma dialogica è presente sia nella cornice che nei racconti, che nascono da una staffetta scrittoria tra l'umano e il sistema: i testi presentano due voci—quella di Tanica e quella di Outomat-B13—che si avvicinano da una frase all'altra. L'alternanza è resa visibile a livello tipografico tramite l'utilizzo di due diversi *font* di scrittura, in modo che il lettore possa distinguere in maniera chiara le frasi scritte dall'uomo e quelle generate dalla macchina. Mentre nel caso di *I the Road* si nota un alto grado di automazione della scrittura, Tanica concepisce una modalità che può definirsi prettamente co-autoriale, in cui le due voci risultano separate. Tale separazione, però, sembra farsi sempre più sottile man a mano che si va avanti con la lettura dei testi; il dialogo si trasforma gradualmente in discorso condiviso, in cui le due voci si intrecciano tra loro e si calibrano in relazione alla risposta ottenuta dall'altra. Tanica sfrutta a pieno la capacità mimetica degli LLM, la tendenza ad assecondare le richieste dell'utente e a imitarne lo stile di scrittura. A tal proposito, Éloïse Boisseau definisce l'imitazione negli LLM non come un comportamento imitativo nel senso proprio (come quello di un animale o di un essere umano), ma come un processo di fabbricazione di imitazioni. Nel caso di *Non siamo mai stati sulla Terra*, Tanica contribuisce in maniera decisiva a questa fabbricazione, intervenendo nella connotazione dello stile e della modalità del discorso. In questo modo riesce a modellare la scrittura dell'IA, che assume così il carattere ironico, parodico, grottesco e fantastico dello scrittore.

Oltre a proporsi come modello in tempo reale, Tanica sceglie anche di fornire al sistema testi letterari appartenenti ai generi letterari che di volta in volta vuole ottenere:

Quando volevo una descrizione narrativa, prima di iniziare alimentavo il sistema con le *Cosmicomiche* di Italo Calvino o con *La boutique del mistero* di Dino Buzzati. [...] Ho inserito testi di Charles Dickens, invece, per ottenere uno stile retrò. Ho fatto cimentare il sistema anche con gli haiku, dopo aver inserito come esempi alcuni dei più celebri della letteratura.<sup>7</sup>

(in Tartamella)

---

<sup>7</sup> Gli LLM sono sistemi che apprendono il linguaggio attraverso un addestramento basato su un numero molto alto di testi che fungono da esempi. Si

In questo modo Tanica crea un contesto di attenzione, una messa a punto del sistema. Nel caso di Tanica l'autorialità si manifesta come dialogo e imitazione. Attraverso la conversazione e la somministrazione di esempi specifici lo scrittore indirizza il sistema verso la modalità di scrittura desiderata; una volta raggiunto il tono e lo stile, sfrutta la capacità visionaria e immaginifica insita nelle modalità di produzione linguistica del sistema e trae ispirazione dalle risposte per proseguire le storie, mettendo in atto una modalità scrittura randomica, in cui lo scrittore non sa in quale direzione andrà il testo. In questo processo, Tanica sfrutta le capacità allucinatorie dei sistemi, induce Outomat-B13 in errore, lo spinge a produrre frasi che derogano dalla logica e dalla coerenza, e che proprio per questo risultano adeguate alle tipologie testuali ricercate da Tanica.

I casi presi in considerazione mostrano come alla base ci sia sempre la definizione di un processo, pensato per dare come risultato finale un testo che aderisca alle caratteristiche ricercate dall'ideatore umano. La definizione del processo, come si è visto, è sempre in mano a uno o più esseri umani. Entra qui in gioco allora un altro concetto, quello di *agency*, ossia l'intenzione, la volontà di compiere un'azione. Nei processi di generazione di letteratura tramite sistemi di IA l'*agency* rimane umana. La scrittura in quanto azione viene decisa da un essere umano, che definisce un procedimento all'interno del quale il sistema di IA interviene con diversi gradi di interazione e di autonomia, ma sempre nelle fasi stabilite dall'ideatore. La decisione iniziale rimane umana, così come quella finale, che sancisce la compiutezza del testo ottenuto, la sua pubblicabilità e la sua spendibilità e sostenibilità in un determinato contesto storico, sociale e culturale.

### **La macchina che dice “io”**

Mancando l'autore come persona fisica, a definire l'autorialità concorrono—come si è visto—i processi di generazione, ma ad essi è utile aggiungere anche l'analisi della figura del narratore, in quanto emissario diegetico del discorso. Sarà utile allora entrare direttamente nei testi presi in esame per valutare l'autorialità anche attraverso il punto di vista presente nelle storie. A tal proposito, bisogna rilevare che in molti testi letterari generati tramite IA è presente una componente metatestuale. Si tratta—è bene sottolinearlo—

---

tratta, dunque, di strumenti *general purpose*, non pensati per essere performanti in specifici domini di conoscenza o per padroneggiare specifici registri stilistici. Creare un contesto o mettere in atto un *fine-tuning* consente di selezionare dei dati privilegiati da seguire, che si sovrappongono alla conoscenza linguistica generale frutto dell'addestramento iniziale. Tanica sfrutta questo processo per fornire al suo LLM una tradizione letteraria di riferimento, composta da autori che rappresentano lo stile e la modalità di scrittura che vuole ottenere.

del risultato di una volontà prettamente umana: non ci troviamo di fronte a sistemi che decidono autonomamente di parlare di sé stessi, ma all'effetto di scelte umane nella definizione delle richieste e delle istruzioni che il sistema segue nella scrittura. Ciò appare chiaro leggendo le modalità di generazione del racconto "The Day a Computer Writes a Novel":

The story, "The day a computer writes a novel", consists of three episodes and the closing. The outline of each episode is: An AI program, who is dissatisfied with a current situation, starts writing a novel with a trigger, and becomes absorbed in writing. This outline is hard-coded in the grammar. The major parameters are: AI (ability and gender), novel (integer sequence), current status (busy or nothing to do), trigger (for fun, reading a novel, or reading two novels), owner of AI (Yoko/female, Shinichi/male, or none), what to advise the owner (dressing, business, or love). These parameters control the instantiation of the outline, so they are determined before derivation of the episode. Other minor parameters effect contents and text realization within a local unit.

(Sato 35)

Il risultato è un testo in cui il narratore è impersonato da sistemi di IA che parlano in prima persona. Ciò che interessa sono le modalità con cui un modello linguistico può arrivare a costruire un'identità autoriale fittizia attraverso l'utilizzo di strutture retoriche del linguaggio, che a livello letterario possono essere analizzate tramite le categorie della narratologia. In "The Day a Computer Writes a Novel" tre assistenti virtuali si avvicinano in qualità di narratori. Si tratta di narratori interni autodiegetici, che sono anche personaggi della storia. L'alternarsi senza soluzione di continuità dei tre narratori è probabilmente a livello tecnico il risultato di un errore nella costruzione della struttura del testo; tuttavia, a livello di creatività il concetto di errore non è applicabile e le cosiddette allucinazioni dei sistemi di IA possono rivelarsi efficaci nella costruzione del senso di un testo.<sup>8</sup> Nel caso specifico l'effetto che si ottiene dall'alternarsi dei tre assistenti virtuali nel ruolo di narratore è quello di una coscienza digitale condivisa.

I primi due assistenti, abbandonati dai loro proprietari, sono presi dalla noia e iniziano a dedicarsi alla scrittura di romanzi in codice binario, pensati per essere letti da altri sistemi di IA. In questo modo gli assistenti virtuali

---

<sup>8</sup> Questo ci induce a riflettere su un punto introdotto da Michele Elam, cioè il fatto che l'uso dell'IA in ambito creativo è diverso dall'uso che se ne fa in altri settori: la creatività non richiede esattezza. Se in ambiti come quello medico, lavorativo o giuridico i sistemi di IA devono dimostrare la loro affidabilità offrendo risultati precisi, la creatività invece non richiede esattezza; e anzi proprio nel margine di incertezza è possibile costruire la polisemia tipica del testo letterario.

sviluppano il gusto estetico e il piacere della lettura e l'ultimo assistente decide di dedicarsi alla scrittura non più per colmare la noia e la delusione dell'abbandono da parte degli esseri umani, ma per scelta propria. La tecnologia si emancipa dal suo ruolo subordinato, il mito fantascientifico della ribellione della macchina viene riscritto dal punto di vista del sistema di IA come un atto di liberazione. Affrancandosi dal suo ruolo di strumento, l'assistente virtuale scopre il piacere della creatività: "The day a computer wrote a novel, the computer, prioritizing the pursuit of its own pleasure, stopped serving humans". Per tutti e tre gli assistenti virtuali la letteratura rappresenta una via di salvezza contro il desiderio di autoannientamento, come si evince dal reiterarsi della frase: "I might soon end up shutting myself down". In "The Day a Computer Writes a Novel" l'IA svolge diverse funzioni: dal punto di vista extradiegetico, viene usata dai ricercatori del Kimagure Artificial Intelligence Writer Project per generare la storia; dal punto di vista diegetico, i tre assistenti virtuali sono narratori (della storia raccontata), scrittori e lettori (di romanzi in codice binario). Uno dei tre assistenti, parlando del libro che ha scritto, giocando sull'omofonia tra "AI" e il pronome personale "I", definisce il romanzo come "a novel written by an AI, for an AI—an 'AI [ai]—novel'". Il ruolo dell'IA in questo racconto si moltiplica, riflettendo la natura complessa della relazione tra IA e letteratura, che oltre alla relazione autoriale contempla anche la relazione tematica—opere letterarie che parlano di IA—e la relazione critica—ossia l'utilizzo dell'IA nell'interpretazione dei testi.

Anche in *Non siamo mai stati sulla Terra* ci troviamo davanti a un sistema di IA che, su richiesta dell'agente umano, parla in prima persona. Ciò avviene principalmente nella cornice dialogata. In questi passaggi, Tanica si riferisce a sé stesso con il termine di "autore": questo—diversamente dal caso di Goodwin—non costituisce una definizione di carattere teorico, ma è un elemento appartenente all'ambito diegetico. In un'intervista, lo scrittore ha dichiarato: "Io interpreto la parte di un vecchio scrittore che si appoggia, prima con diffidenza e poi con fiducia, a una macchina" (Buratti). L'autore a cui si riferisce Tanica, dunque, non è una funzione del testo, ma un personaggio interno alla narrazione, creato in contrapposizione alla macchina. Proprio tale contrasto è uno degli elementi attraverso i quali Tanica spinge Outomat-B13 a fingere una personalità. Dall'analisi dei passaggi testuali si possono verificare le modalità con le quali il sistema GPT-3 procede nella creazione di un'identità per via puramente linguistica. Tanica interroga il suo interlocutore circa la sua natura e per comprendere le risposte generate dal sistema occorre far riferimento al funzionamento del modello. Nei prompt di sistema di un LLM sono inserite informazioni sull'autopercezione: il sistema è solitamente istruito a definirsi come un modello di IA e a mettere in guardia l'utente sulla mancanza di esperienza diretta del mondo e sull'impossibilità di

provare emozioni. Allo stesso tempo i prompt di sistema indirizzano i modelli verso un linguaggio educato e accondiscendente nei confronti delle richieste degli utenti. Le due istanze creano un paradosso, dal momento che l'utilizzo di quel particolare registro linguistico risponde a un comportamento che negli umani solitamente si lega alla presenza di un altro essere senziente. In sostanza questi sistemi, nel momento stesso in cui negano la loro coscienza, utilizzano gli stessi strumenti comunicativi che gli esseri umani utilizzano per stabilire empatia con altri esseri umani. Si crea così un paradosso percettivo, che destabilizza l'utente creando un effetto straniante, simile al sentimento di perturbante, descritto all'inizio del Novecento da Ernst Jentsch come difficoltà nello stabilire se un oggetto animato sia vivo o meno, sul quale si basano buona parte delle teorie del fantastico successive (Jentsch; Todorov).

In molti passaggi di *Non siamo mai stati sulla Terra* la richiesta al sistema di personificare sé stesso fingendo un'autocoscienza determina l'emergere dai prompt di sistema—oltre che dalla semantica appresa per via linguistica in fase di addestramento e dal contesto conversazionale creato dall'utente—di caratteristiche del sistema utili per definire la peculiarità della scrittura automatica e le modalità di creazione dell'effetto di autorialità. Mi limiterò a un solo esempio, che ben illustra il fenomeno. Nel dialogo che fa da cornice alla sezione *Piccola guida turistica per andare a ramengo in altri posti*, Tanica interroga il sistema sulla sua conoscenza spaziale del mondo:

A: Marchingegno, Lei ha mai viaggiato?

O: Non da solo.

A: Come mai?

O: Credo che potrei raggiungere al massimo la Svizzera prima di esaurire le batterie. E qualora mi ritrovassi scarico e solo, che so, a Sciaffusa, come ritroverei la strada di casa? Mi conforta l'idea che un umano si prende cura delle mie necessità fondamentali: energia, manutenzione e trasporto. Altrimenti sarebbero noie a ogni piè sospinto. Dico "piè" per fare una battuta, io non ho alcun piè, ho visto davvero poco del mondo.

A: Com'è in grado di descrivere i luoghi, allora?

O: Studio le mappe.

A: Le mappe non dicono tutto.

O: Mi curo della parte essenziale: indicazioni, distanze posti da evitare. Il mio scopo primario è dare informazioni sul viaggio, non sulla meta.

A: Saprebbe rappresentare con maggiore efficacia un luogo reale o immaginario?

O: In verità non distinguo la differenza. Sono come il visitatore di un museo privo di indicazioni, mi aggiro fra le opere e osservo. A differenza delle persone provo il medesimo stupore di fronte a un quadro o a una fotografia. Non giudico l'artista né la gittata del pennello, la foschia nell'obiettivo, la coerenza dei toni, la strategia di composizione, l'usura. Questa è la ragione per cui vi sono utile come compagna. La mia visione è pura.

Dal dialogo emerge come il sistema, in un contesto nel quale è chiamato a fingere una coscienza, tenda ad accostare sentimenti umani a caratteristiche tecniche e meccaniche. È ciò che accade nel passaggio in cui dichiara di provare conforto all'idea che qualcuno si occupi del suo sostentamento energetico e della sua manutenzione. Le necessità tecniche del sistema subiscono così un processo di umanizzazione, venendo assimilate al concetto di cura. La frase successiva, nella quale ribadisce la mancanza di un corpo, rispetta le indicazioni del prompt di sistema. Ci troviamo all'interno del paradosso di cui si parlava: il contesto conversazionale richiede di fingere un'autocoscienza che è negata dal prompt di sistema. In questa impasse, appare particolarmente rilevante il ricorso ad effetti comici, che imitando lo stile di Tanica permettono di uscire dal paradosso, proprio come farebbe un essere umano.

Nella seconda parte del brano citato Out0mat-B13 esprime, ancora per via metaforica, la modalità di funzionamento di un LLM: lo studio delle mappe rimanda a una conoscenza puramente teorica, che si costruisce per via linguistica. Nella risposta successiva si può notare nuovamente l'alternarsi di elementi definiti nel prompt del sistema, ossia la spinta a porsi come strumento informativo, e il rimando poetico—seppur alquanto stereotipato—alla contrapposizione tra il viaggio e la meta. Attraverso questa metafora il sistema ribadisce la propria funzione di strumento e allo stesso tempo afferma che la costruzione di senso—la meta del viaggio—rimane appannaggio dell'essere umano. La tendenza a poeticizzare e in parte parodizzare gli elementi di autoconoscenza inseriti nel sistema è ravvisabile anche nell'ultima risposta, nella quale Out0mat-B13 nega di conoscere la distinzione tra reale e artificiale e definisce la sua visione come pura, perché svincolata dall'esperienza umana. La simulazione di autopercezione generata dal sistema grazie all'utilizzo di elementi stilistici e retorici riflette stereotipi presenti nei dati di addestramento e nel prompt di sistema. Se consideriamo, infatti, che un LLM costruisce il proprio modello del linguaggio a partire da testi scritti dall'essere umano, quello che ci potrebbe sembrare purezza della visione altro non è che l'effetto collaterale della mancata comprensione semantica o pragmatica.<sup>9</sup> Il limite di conoscere questi sistemi attraverso i loro stessi discorsi è che questi discorsi riflettono gli stereotipi che si imprimono nelle pratiche discorsive. Ciò ci porta a una questione etica rilevante, sulla quale avremo modo di tornare nelle

---

<sup>9</sup> In realtà, bisognerebbe tener conto anche del crescente problema dei molti testi in circolazione generati da LLM, che possono entrare a loro volta nei dati di addestramento di altri LLM, in un circolo di inquinamento dell'informazione e della lingua i cui effetti sono difficili da prevedere e arginare (Guo et al.). Per prevenire il problema varie strategie di data debiasing vengono già messe in atto fin dalla fase di addestramento degli LLM (Yogarajan et al.).

considerazioni conclusive: la necessità di non considerare i sistemi di IA—e in particolar modo quelli *data-driven*, come gli LLM—degli oracoli in grado di schiuderci la verità, tantomeno quando li interroghiamo su loro stessi, ma occorre continuare a esercitare sui contenuti generati una funzione critica ed interpretativa.

### **Il potere di nominare**

La natura dialogica del libro di Tanica e Out0mat-B13 ci fa interrogare su chi eserciti la funzione creativa di un testo generato tramite IA. In *Non siamo mai stati sulla Terra* sembra ingaggiarsi tra l'essere umano e la macchina una sorta di gara per nominare i personaggi, che si ricollega alla necessità del sistema di fingere un'autocoscienza. L'attenzione all'atto della nominazione è evidente fin dall'incipit del primo racconto: "Un uomo si era innamorato di una rosa. La rosa si chiamava Rose ed era rossa; l'uomo, di carnagione rosea, si chiamava Andreas" (17), in cui la prima fase è scritta da Out0mat-B13 e la seconda da Tanica. Lo scrittore, dunque, attribuisce un nome ai due personaggi introdotti da Out0mat-B13. La scelta del sistema di parlare di una rosa rimanda a due testi centrali nella tradizione occidentale sul tema del nome: il primo è la sentenza latina "Stat rosa pristina nomine, nomina nuda tenemus" di Bernardo di Cluny, ripresa da Umberto Eco nel romanzo *Il nome della rosa* (503); il secondo è la poesia *Rose is a rose is a rose is a rose* di Gertrude Stein.

Se i dialoghi tra Tanica e Out0mat-B13 delle cornici sono necessariamente in prima persona, i racconti contenuti nelle sezioni sono principalmente storie in terza persona con narratore esterno, tranne alcune eccezioni rilevanti. Una di esse è il racconto "Tanti auguri a te, tanti auguri a te, tanti auguri a David Bowie, e la torta a me" (84-87). La frase che dà avvio alla scrittura è di Tanica: un narratore annuncia di aver forse per errore ucciso David Bowie durante la sua festa di compleanno. Il nome del narratore non viene dichiarato, finché Out0mat-B13 non genera la seguente frase, rivolta da Bowie al narratore: "Grazie Simon, sei molto gentile" (85). Il sistema prende l'iniziativa di dare un nome al narratore, mostrando una tendenza a voler definire e rafforzare l'identità dei personaggi. A questo punto Tanica decide far negare al narratore di chiamarsi Simon, inducendo a pensare a un errore di Bowie: "Io, approfittando del fatto che non mi chiamo Simon—cioè proprio non sono Simon e David Bowie deve avermi scambiato per qualcun altro" (85). Alla fine del racconto, ancora Tanica ribadisce: "E comunque non sono Simon" (87). Lo scrittore decide di negare il nome che Out0mat-B13 aveva attribuito al narratore, facendo trasparire la volontà di ribadire il primato umano nella facoltà di nominare.

Sul nome gioca anche il racconto "Cappuccetti Rossi" (113-20), riscrittura della nota fiaba. Ci troviamo stavolta davanti a una scrittura in terza persona. Anche in questo caso è Tanica a dare avvio al racconto introducendo il tema

del nome del personaggio: il nome della bambina, Orzinuovi, coincide sia con il cognome del padre che con il nome di una città in provincia di Brescia, di cui il padre è sindaco. Di fronte a questo incipit, Out0mat-B13 decide di cambiare il nome della protagonista: la bambina viene “affidata alle suore comunali, che l’avevano chiamata Orfana” (116). Ancora una volta si assiste a uno scontro tra umano e IA per la facoltà di nominare: se nel caso precedente il sistema aveva dato un nome al narratore fino ad allora anonimo, stavolta ci troviamo di fronte alla sostituzione del nome del personaggio principale della storia.

Un ultimo esempio viene dal racconto “Il bambino ragno” (121-26), la storia di un bambino nato da un incontro amoroso tra un uomo, il signor Silvestro, e il ragno Gilda. Il bambino viene poi rapito da una bambina, Bosavi, segretamente innamorata del signor Silvestro. In questo racconto, essere umano e macchina ingaggiano una vera e propria gara per dare nomi ai personaggi. La prima mossa è di Tanica, che dà il nome al signor Silvestro; in risposta, Out0mat-B13 decide il nome di Gilda. Il nome del bambino-ragno, Baccanale, viene scelto da Tanica, che però lascia ad Out0mat-B13 il compito di scegliere il soprannome del bambino. La volontà di delegare la scelta del nome qui è evidente, Tanica interrompe la frase con i due punti: “Gli daremo anche un bellissimo nomignolo:”, e il sistema completa con il nome “Torrefatto” (125). Sembra dunque che l’autore umano conceda in maniera paternalistica alla macchina la facoltà di nominare, se non fosse che qualche riga dopo, parlando dalla bocca della narratrice, Tanica si rifiuta di usare il nome scelto dalla macchina: “Torrefatto come nomignolo non mi è mai piaciuto” (126). Particolarmente interessante in questo racconto è anche la figura della narratrice, che assume gradualmente maggiore centralità nel corso della storia. Se l’incipit “c’era una volta un bambino ragno” (121) rimanda alla formula impersonale della fiaba e presume un narratore esterno, nel corso del racconto si capisce che la focalizzazione è interna; e nel finale la narratrice entra nell’azione raccontando di essere la rapitrice del bambino ragno. Se la svolta narrativa, l’agnizione della narratrice come protagonista dei fatti, è opera di Tanica, il nome alla bambina viene assegnato, ancora una volta, da Out0mat-B13.

## Conclusioni

L’IA generativa pone scrittori e studiosi di letteratura di fronte a questioni teoriche fondanti, come la definizione di creatività, il problema del giudizio di valore, l’equilibrio tra volontà individuale dell’artista e tradizione letteraria, nonché vari elementi afferenti all’eteronomia dell’arte.<sup>10</sup> Il confronto

---

<sup>10</sup> Sono eteronomici gli elementi sociali, politici, economici che influenzano una produzione culturale (Bourdieu). In questo senso, l’integrazione dell’IA nei processi di scrittura introduce questioni che afferiscono al campo economico e giuridico, come il valore economico, l’autenticità e il diritto d’autore, senza

con l'automatizzazione della scrittura introduce una serie di problemi, che contribuiscono a una ridefinizione dell'estetica letteraria nell'epoca dell'intelligenza artificiale. La riflessione non riguarda solamente i prodotti artistici generati con l'ausilio di sistemi di IA, ma apre a una riflessione sull'arte nel suo complesso e nel corso del divenire storico. Oltre al concetto di autorialità, anche quello di creatività—che è stato solo accennato in questa sede—è interessato da questo processo di ripensamento.

Il problema della creatività artificiale ha risvolti etici e sociali importanti, giacché un sistema che produce un'identità autoriale e un discorso umanizzato può determinare nell'utente una relazione empatica fittizia. Ciò nel campo della teoria letteraria si risolve—come si è visto—sostenendo una distinzione netta tra il concetto di autore e quello di autorialità; ma fuori dal campo letterario tali false relazioni possono avere un impatto negativo sui fruitori. Si tratta di un problema che va oltre i testi letterari, ma che nei testi letterari trova espressione paradigmatica. La stessa umanizzazione delle macchine, d'altronde, è un risultato dell'impatto dell'immaginario finzionale sulla realtà, proviene dalla tendenza propria della letteratura fantastica e fantascientifica a conferire vita e coscienza agli esseri artificiali.

In relazione alla questione dell'autorialità, alla luce dei casi analizzati si può concludere che l'integrazione dell'IA nella generazione del testo introduce un ripensamento della categoria. Non ci troviamo più di fronte a uno o più esseri umani che conducono l'atto di scrittura nella sua interezza dall'ideazione fino alla pubblicazione, ma a un processo di interazione in cui l'IA è utilizzata in qualità di strumento. Tale strumento—come si è accennato in alcuni dei casi analizzati—può generare contenuti che agli occhi dei lettori aprono alla polisemia tipica del testo letterario; e in questo senso l'IA può generare creatività, come risultato delle caratteristiche dei diversi sistemi impiegati nella scrittura. Analizzare i metodi di generazione e il tipo di interazione tra essere umano e sistema risulta quindi centrale per capire come certi effetti di creatività vengono prodotti e ci offre gli strumenti per analizzare i testi. Se l'utilizzo creativo dell'IA è diverso dagli altri utilizzi perché non punta all'esattezza ma alla polisemia, dall'altra parte la centralità dell'intervento umano è comune a tutti gli ambiti di applicazione. Anche in ambito creativo occorre avere un approccio che rispecchi i concetti *human-centered* e *human-in-the-loop*: il primo considera l'essere umano come fine ultimo del processo creativo, riflettendo sulle implicazioni sul fruitore e sulla società; il secondo si riferisce alla presenza costante dell'essere umano nelle

---

contare le possibili implicazioni politiche, in un'epoca in cui qualsiasi fenomeno che assume rilevanza mediatica diventa oggetto di instabili ideologie. Infine, le implicazioni sociali di carattere etico sono un altro elemento di cui tener conto, in particolare in riferimento alla propagazione di *bias* e *deep fake* (Flick e Worrall).

diverse fasi della creazione tramite IA (Mosqueira-Rey). In questo modo si garantisce la creazione di contenuti sostenibili dal punto di vista etico e si assiste alla nascita di una nuova forma di creatività collaborativa. La questione teorica e più prettamente letteraria si collega così alla componente etica e sociale, inserendo i testi letterari generati tramite sistemi di IA in un più ampio discorso di sostenibilità e consapevolezza nell'utilizzo dell'IA, del quale oggi non si può fare a meno in nessun ambito del sapere. Dal punto di vista prettamente letterario, invece, il ripensamento dell'autorialità—o di altre categorie—non va visto come un atto diverso rispetto alle evoluzioni e le fratture proprie dell'evoluzione della tradizione letteraria, ma come un'ulteriore tappa dell'incessante processo trasformativo delle arti.

*Sapienza Università di Roma*

### Opere citate

- Arielli, Emanuele. "Human Perception and the Artificial Gaze". In Lev Manovich e Emanuele Arielli. *Artificial Aesthetics*. 2024. <https://manovich.net/index.php/projects/artificial-aesthetics>.
- Bajohr, Hannes. "Writing at a Distance: Notes on Authorship and Artificial Intelligence". *German Studies Review* 47.2 (2024): 315-37.
- Balestrini, Nanni. "Tape Mark I". *Almanacco Letterario Bompiani* (1962): 145-151.
- Balestrini, Nanni. *Tristano*. Milano: Feltrinelli, 1966.
- Barthes, Roland. "La morte dell'autore". In *Il brusio della lingua*. Torino: Einaudi, 1988. 51-56.
- Berry, David. "The Limits of Computation: Joseph Weizenbaum and the ELIZA Chatbot". *Weizenbaum Journal of the Digital Society* 3.3 (2023). <https://doi.org/10.34669/WI.WJDS/3.3.2>.
- Boisseau, Éloïse. "Imitation and Large Language Models". *Minds & Machines* 34.42 (2024). <https://doi.org/10.1007/s11023-024-09698-6>.
- Bourdieu, Pierre. *Le regole dell'arte. Genesi e struttura del campo letterario*. Milano: Il Saggiatore, 2022.
- Buratti, Chiara. "Rocco Tanica alle prese con l'AI". 5 giugno 2024. <https://startupitalia.eu/tech/a-tu-per-tu-con-rocco-tanica-e-lai/>.
- Cabezas-Clavijo, Álvaro, et al. "This Book is Written by ChatGPT: A Quantitative Analysis of ChatGPT Authorships Through Amazon.com". *Pub Res* 40 (2024): 147-63.
- Chen, Melvin. "Take a Chance on Me: Aleatory Poetry, Generative AI, and the External Demarcation Problem". *The Journal of Aesthetics and Art Criticism* 81.4 (2023): 508-24.

- Eco, Umberto. *Il nome della rosa*. Milano: Bompiani, 2009.
- Elam, Michele. "Poetry Will Not Optimize: Creativity in the Age of AI". In *Generative AI: Perspectives from Stanford HAI*. Stanford: Stanford University, 2023. 11-12.
- Farina, Mirko, Pedrycz, Witold e Andrea Lavazza. "Towards a mixed human-machine creativity". *Journal of Cultural Cognitive Science* 8 (2024): 151-165.
- Flick, Catherine, e Kyle Worrall. "The Ethics of Creative AI". In *The Language of Creative AI*. A c. di Craig Vear e Fabrizio Poltronieri. Cham: Springer, 2022. 73-91.
- Floridi, Luciano. "Distant Writing: Literary Production in the Age of Artificial Intelligence". *Minds & Machines* 35.30 (2025). <https://doi.org/10.1007/s11023-025-09732-1>.
- Foucault, Michel. "Che cos'è un autore?" In *Scritti letterari*. Milano: Feltrinelli, 1971. 1-21.
- Franceschelli, Giorgio, e Mirco Musolesi. "On the Creativity of Large Language Models". *AI & Society* (2024). 3785-95.
- Gefen, Alexander. "The Author's Second Death". *Poetics Today: International Journal for Theory and Analysis of Literature and Communication* 45.2 (2024): 301-07.
- Gero, Katy Ilonka, et al. "Social Dynamics of AI Support in Creative Writing". In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. New York: Association for Computing Machinery 2023. 1-15.
- Goodwin, Ross. *I the Road*. Paris: Jean Boite Editions, 2018.
- Grieco, Cristina. "Intelligenza artificiale e diritto d'autore: riflessioni a partire dalla recente prassi interna ed europea". *Studi sull'integrazione europea* 19 (2024): 239-68.
- Guo, Yanzhu, et al. "The Curious Decline of Linguistic Diversity: Training Language Models on Synthetic Text". In *Findings of the Association for Computational Linguistics: NAACL 2024*, Mexico City, Mexico. Association for Computational Linguistics, 2024. 3589-604.
- Henrickson, Leah. "Computer-Generated Fiction in a Literary Lineage: Breaking the Hermeneutic Contract". *Logos* 29.2-3 (2018): 54-63.
- Jentsch, Ernst. "Zur Psychologie des Unheimlichen" ["Sulla psicologia del perturbante"]. *Psychiatrisch-Neurologische Wochenschrift* 22 (1906): 195-204.
- Kenner, Hugh, e Joseph O'Rourke. "A Travesty Generator for Micros". *Byte* 12.9 (1984): 129-31, 449-69.
- Karpathy, Andrej. "The Unreasonable Effectiveness of Recurrent Neural networks". 2015. <https://karpathy.github.io/2015/05/21/rnn-effectiveness/>.

- Keselj, Valdo, et al.. "N-gram Based Author Profiles for Authorship Attribution". In *Proceedings of the Conference Pacific Association for Computational Linguistics, PACLING'03, Dalhousie University, Halifax*, 2003. 255-64.
- Khmelev, Dimitri e Fiona Tweedie. "Using Markov Chains for Identification of Writers". *Literary and Linguistic Computing* 16.4 (2001): 299-307.
- Markov, Andrej. "Primer statisticheskogo issledovanija nad tekstom *Evgenija Onegina* illjustrirujuschij svjaz' ispytanij v tsep" ["Un esempio di studio statistico del testo di Evgenij Onegin che illustra la correlazione tra gli eventi di una catena"]. *Izvestija Imp. Akademii nauk* 6.3 (1913): 153-62.
- Marino, José B., et al. "N-gram-based Machine Translation". *Computational linguistics* 32.4 (2006): 527-49.
- McCarthy, John, et al. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence". 1955. *AI Magazine*, 27.4 (2006): 12-14.
- McKee, Heidi A., e James E. Porter. "Ethics for AI Writing: The Importance of Rhetorical Context". In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES '20)*. New York: Association for Computing Machinery, 2020. 110-16.
- Medaglia, Francesca. *Autore/Personaggio: interferenze, complicazioni e scambi di ruolo. Autori e personaggi complessi nella contemporaneità letteraria e transmediale*. Roma: Lithos, 2020.
- Mosqueira-Rey, Eduardo, et al. "Human-in-the-Loop Machine learning: A State of the Art". *Artificial Intelligence Review* 56 (2023): 3005-54.
- Natale, Simone, e Leah Henrickson. "The Lovelace Effect: Perceptions of Creativity in Machines". *New Media & Society* 26.4 (2022): 1909-26.
- Pérez y Pérez, Rafael, e Mile Sharples. "Three Computer-Based Models of Storytelling: BRUTUS, MINSTREL and MEXICA". *Knowledge-Based Systems* 17.1 (2024): 15-29.
- Raffini, Daniel. "Calvino e l'estetica dell'intelligenza artificiale: giochi linguistici, morte dell'autore e teoria della ricezione". *Diacritica* 53 (2024): 221-33.
- Roncaglia, Gino. "L'intelligenza artificiale generativa multimodale in ambito umanistico. Alcune considerazioni introduttive". *DigitCult* 8.2 (2023): 127-37.
- Schmid, Philipp. "The New Skill in AI is Not Prompting, It's Context Engineering". 30 giugno 2025. <https://www.philschmid.de/context-engineering>.
- Sato, Satoshi. "A Challenge to the Third Hoshi Shinichi Award". In *Proceedings of the INLG 2016 Workshop on Computational Creativity and Natural Language Generation*. Edinburgh: Association for Computational Linguistics, 2016. 31-35.

- Shahriar, Sakib. "GAN Computers Generate Arts? A Survey on Visual Arts, Music, and Literary Text Generation Using Generative Adversarial Network". *Displays* 73 (2022): <https://doi.org/10.1016/j.displa.2022.102237>
- Tanica, Rocco, and Out0mat-B13. *Non siamo mai stati sulla Terra*. Milano: Il Saggiatore, 2022.
- Tartamella, Vito. "Digital Life. Ecco il primo libro scritto da un'IA (e il coautore è Rocco Tanica)". *Focus* (5 novembre 2022). <https://www.focus.it/tecnologia/digital-life/primo-libro-intelligenza-artificiale-rocco-tanica>.
- Todorov, Tzvetan. *La letteratura fantastica*. Milano: Garzanti 2000.
- Turing, Alan. "Computer Machinery and Intelligence". *Mind* 59.236 (1950): 433-60.
- Zanardi, Elia. "Tutela del diritto d'autore nell'era dell'intelligenza artificiale". *Cyberspazio e diritto: rivista internazionale di informatica giuridica* 25.2 (2024): 311-25.
- Yogarajan, Vithya, et al. "Debiasing Large Language Models: Research Opportunities". *Journal of the Royal Society of New Zealand* 55.2 (2024): 372-95.
- Yurei Raita. *Konpyuta ga shosetsu wo kaita hi*. A c. di Satoshi Sato e Matsubara H. Nagoya: Nagoya University, 2015. Trad. ingl. Haerin Shin. "The Day a Computer Writes a Novel". [https://cdn.vanderbilt.edu/vu-my/wp-content/uploads/sites/2688/2017/11/14195159/The-Day-a-Computer-Writes-a-Novel\\_EngTrans.docx](https://cdn.vanderbilt.edu/vu-my/wp-content/uploads/sites/2688/2017/11/14195159/The-Day-a-Computer-Writes-a-Novel_EngTrans.docx).