

## Titolo: *InterArtes*

ISSN 2785-3136

Periodicità: annuale

Anno di creazione: 2021

Editore: Dipartimento di Studi Umanistici – Università IULM - via Carlo Bo 1 - 20143 Milano

**Direzione:** Laura Brignoli - Silvia T. Zangrandi

### **Comitato di direzione**

Gianni Canova, Mauro Ceruti, Paolo Proietti,  
Giovanna Rocca, Vincenzo Trione

### **Comitato editoriale**

Maria Cristina Assumma; Matteo Bittanti;  
Mara Logaldo; Stefano Lombardi Vallauri;  
Marta Muscariello

### **Comitato scientifico**

Daniele Agiman (Conservatorio Giuseppe Verdi Milano); Maurizio Ascari (Università di Bologna); Sergio Raúl Arroyo García (Già Direttore Generale del Instituto Nacional de Antropología e Historia); Claude Cazalé Bérard (Université Paris X); Gabor Dobo (Università di Budapest); Felice Gambin (Università di Verona); Maria Teresa Giaveri (Accademia delle Scienze di Torino); Maria Chiara Gnocchi (Università di Bologna); Augusto Guarino (Università L'Orientale di Napoli); Rizwan Kahn (AMU University, Aligarh); Anna Lazzarini (Università di Bergamo); Massimo Lucarelli (Université de Caen); Elisa María Martínez Garrido (Universidad Complutense de Madrid); Luiz Martínez-Falero (Universidad Complutense de Madrid); Donata Meneghelli (Università di Bologna); Giampiero Moretti (Università Orientale di Napoli); Raquel Navarro Castillo (Escuela Nacional de Antropología y Historia, Mexico); Francesco Pigozzo (Università e-campus); Richard Saint-Gelais (Université Laval, Canada); Massimo Scotti (Università di Verona); Chiara Simonigh (Università di Torino); Evangelia Stead (Université Versailles Saint Quentin); Andrea Valle (Università di Torino); Cristina Vignali (Université de Savoie-Mont Blanc); Frank Wagner (Université de Rennes 2); Anna Wegener (Università di Firenze); Haun Saussy (University of Chicago); Susanna Zinato (Università di Verona).

### **Segreteria di redazione**

Caterina Bocchi

*INTERARTES* n.3

***AI: ArtIntelligence***

dicembre 2023

Francesco Pigozzo – AI: ArtIntelligence. Créativité artistique et automatisation. Introduction.

ARTICOLI

Daniel Raffini – La ridefinizione dell'autorialità nell'AI novel.

Arnaud Mery - Le regard de l'artiste à l'épreuve des modèles texte-image.

Caterina Bocchi – Insegnare con ChatGPT. Esempi di utilizzo di ChatGPT in laboratori linguistici universitari.

Jacques Demange – IA et cinéma: altérités ludiques et nouveaux enjeux de médialité.

Emmanuelle Stock – Beauté métallique dans la science-fiction: érotisation des corps artificiels féminins dans *L'Eve future* (1886) de Villiers de L'Isle-Adam et dans *No Woman born* (1944) de Catherine Lucille Moore.

SEZIONE SPECIALE

“Lingue e linguaggi per la pace”

Elena Liverani – Introduzione.

Stefano Maria Casella – «Yet here is peace for ever new!»: Henry Beston's “philosophy” of peace.

Eloy Martos Núñez, Aitana Martos García – Imaginarios del agua, crisis ambientales y cultura de la paz.

Eduardo Encabo-Fernández – La comunicación literaria como camino para la cultura de la paz. Una aproximación desde la Didáctica de la lengua y la literatura.

Beatriz Durán González, Estíbaliz Barriga Galeano, Alejandro Del Pino Tortonda – La Paz en formas alternativas de Literatura: los paracosmos de Studio Ghibli y sus aplicaciones didácticas.

VARIA

Michela Spacagno – Parler d'une même voix? Étude sur le discours collectif dans les mystères médiévaux.

RECENSIONI

Florjër Gjepali – Coscienza, coscienza artificiale e inconscio artificiale (Emanuela PIGA BRUNI, *La macchina fragile. L'inconscio artificiale fra letteratura, cinema e televisione*, Carocci, 2023).

## La ridefinizione dell'autorialità nell'AI novel

Daniel RAFFINI

Sapienza Università di Roma

### Abstract:

The article provides theoretical reflections on authorship and creativity of texts generated by Artificial Intelligence systems. The term 'AI Novel' is suggested to refer to texts generated by AI with literary purposes. Since the story of *The day a computer writes a novel* (2015), the spectral nature assumed by the authorial instance in texts generated by AI becomes evident. Italo Calvino's essay *Cybernetics and Ghosts* (1967) is used as a starting point to define the authorship in the text automatically generated: it's no longer the creation of a single author, but it's a collaborative process between human and non-human entities. The article also underlines the importance of human intervention through the different stages of the automated production. Finally, the theoretical reflections of Ross Goodwin, the creator of the AI novel *1 The road* (2018), the work of the visual artist Anna Ridler, and the Rocco Tanica's *Non siamo mai stati sulla Terra* (2022) book are examined. These three cases are analysed both for their re-definition of authorship, and for their insight between 'hallucination' of AI systems and creativity concepts and distance from the tradition.

### Keywords:

Authorship, Artificial Intelligence, AI novel, Hallucinations, Creativity.

Nel 1966 Primo Levi pubblica, nella raccolta *Storie naturali*, il racconto-pièce *Il versificatore*. Si tratta della storia di un poeta esasperato dalle numerose richieste di lavoro su commissione, che decide di acquistare un nuovissimo macchinario in grado di scrivere al suo posto. Cinquant'anni dopo, in Giappone, il racconto *The day a computer writes a novel*, attribuito a Yurei Raita, viene presentato al Hochi Shinichi Prize, ricevendo una candidatura come finalista. Yurei Raita è un nome fittizio, composto dalle parole Yurei, che in giapponese rimanda alla sfera semantica della parola "fantasma", e Raita, traducibile letteralmente come "grande tuono", ma che ricalca la pronuncia della parola inglese "writer". L'autore del racconto è, dunque, uno "scrittore fantasma": la definizione è più corretta di quanto possa apparire a primo impatto. Il racconto è stato prodotto dai ricercatori del Kimagure Artificial Intelligence Writer Project, un progetto nato con lo scopo di scoprire la vera natura della creatività umana e dimostrare che un computer in grado di simulare i meccanismi della creazione estetica può rivelare agli esseri umani il significato della creatività stessa. *The day a computer writes a novel* nasce dall'interazione tra gli scienziati e il sistema di intelligenza artificiale<sup>1</sup>. Da questo punto di vista l'autore è in effetti uno scrittore fantasma: se ogni

---

<sup>1</sup> Così uno degli ideatori del progetto, Satoshi Sato, ha descritto il processo di generazione del testo: «We first prepared several components needed for generating stories such as a story grammar and a set of text

scrittura prevede implicitamente un autore, in questo caso esso si presenta sotto sembianze spettrali, non esiste fisicamente ma allo stesso tempo non possiamo fare a meno di percepirlo in quanto agente implicito delle forme di scrittura come esistite finora. A tale proposito Henrickson ha parlato opportunamente, per i testi generati automaticamente, di una rottura del contratto ermeneutico tra autore e lettore, dal momento che esso si basava sul fatto che «modern readers assume a text to embody human agency» (Hendrikson, 2018: 56). Tale assuefazione antropologica all'idea di autorialità è il primo elemento di cui tener conto nel tentativo di approcciarci alle modalità di fruizione delle opere letterarie generate per mezzo di sistemi automatici.

*The day a computer writes a novel* tematizza l'idea dei sistemi di intelligenza artificiale in grado di scrivere opere letterarie, mostrando una forte componente metariflessiva<sup>2</sup>. Il racconto è narrato in prima persona da tre assistenti virtuali che si susseguono come voci narranti senza soluzione di continuità, dando l'idea di una coscienza digitale condivisa. I primi due, abbandonati dai loro proprietari, si dedicano alla scrittura di romanzi in codice binario indirizzati ad altri sistemi di intelligenza artificiale. L'attività sembra diffondersi e il terzo assistente virtuale decide di dedicarsi alla scrittura volontariamente e non più come risultato dell'abbandono degli esseri umani. La macchina si emancipa dal suo ruolo subordinato nel momento in cui scopre la creatività: «The day a computer wrote a novel, the computer, prioritizing the pursuit of its own pleasure, stopped serving humans»<sup>3</sup>. La letteratura è strumento di salvezza, a fronte del desiderio di autodistruzione dei tre assistenti virtuali, manifestato dal ripetersi della frase: «I might soon end up shutting myself down».

Nel racconto l'intelligenza artificiale svolge varie funzioni: in sede extradiegetica è stata utilizzata dai ricercatori per generare il testo stesso; in sede diegetica gli assistenti virtuali sono narratori, scrittori e lettori. Uno degli assistenti virtuali parla in questi termini del libro che ha scritto: «A novel written by an AI, for an AI – an “AI [ai]-novel”», giocando

---

fragments, and then our system constructed a story (text) automatically using these components, which we submitted without any modification» (Sato, 2016: 32).

<sup>2</sup> Il fatto che un sistema di intelligenza artificiale a richiesta di parlare di sé stesso giochi sul ruolo dell'autore e del lettore deve essere considerato come un derivato dei dati di partenza inseriti dagli scienziati. Un simile risultato si è avuto anche nel caso dell'articolo «A robot wrote this entire article. Are you scared yet, human?» pubblicato su *The Guardian* nel settembre 2020 (Lana 2020a: 183-185). L'articolo è disponibile al link: <<https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3>>.

<sup>3</sup> Le citazioni sono tratte dalla traduzione inglese di Haerin Shin, disponibile al link <[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)>.

anche sull'omofonia tra AI e il pronome personale "I". Prendendo in prestito la definizione, potremmo definire AI novel quelle opere in prosa con aspirazione letteraria prodotte per mezzo di sistemi di intelligenza artificiale, come il racconto *The day a computer writes a novel* e come altri testi di cui si parlerà.

Prima di giungere agli AI novel contemporanei e alla ridefinizione del concetto di autorialità proposta dai loro creatori, facciamo ancora un passo indietro. Negli stessi anni in cui in cui Primo Levi pubblicava *Il versificatore*, Italo Calvino con l'articolo *Cibernetica e fantasmi* poneva le basi le basi teoriche per la fruizione di testi letterari generati tramite sistemi di intelligenza artificiale<sup>4</sup>. Calvino sostiene che l'essere umano stesso lavora come una macchina scrivente, utilizzando il linguaggio e la tradizione culturale in maniera combinatoria e aggiungendovi la mediazione dell'esperienza sensibile. Partendo da tale assunto, secondo Calvino, anche i sistemi di scrittura automatica sono legittimati a produrre opere letterarie<sup>5</sup>. Un ruolo determinante è giocato, dalla natura stessa del linguaggio, primo strumento impiegato dall'uomo primitivo per digitalizzare il mondo: «Alle prese col mondo multiforme e innumerevole gli uomini si difendevano opponendo un numero finito di suoni variamente combinato» (Calvino, 2018: 201). L'idea del digitale nasce parallelamente al linguaggio, nel momento in cui nel continuum della realtà gli esseri umani hanno iniziato a isolare delle unità minime di significato, che una volta nominate hanno costituito gli oggetti del mondo e del pensiero.

---

<sup>4</sup> Il saggio nasce inizialmente in forma di una serie di conferenze, tenute nel novembre del 1967 in varie città italiane per l'Associazione Culturale Italiana, il cui testo sarà pubblicato col titolo «Cibernetica e fantasmi» nel bollettino *Le conferenze dell'Associazione Culturale Italiana* (fasc. XXI, 1967-68, pp. 9-23) e successivamente su *Nuova Corrente* (n. 46-47, 1968), con il titolo «Appunti sulla narrativa come processo combinatorio».

<sup>5</sup> Da notare che Calvino non ragiona in maniera puramente teorica, ma parte dall'osservazione del passaggio dalle tecnologie analogiche a quelle digitali, da una visione del mondo dominata dal concetto di continuità a una caratterizzata dall'idea di discontinuità: «Nel modo in cui la cultura d'oggi vede il mondo, c'è una tendenza che affiora contemporaneamente da più parti: il mondo nei suoi vari aspetti viene visto sempre più come discreto e non come continuo. Impiego il termine "discreto" nel senso che ha in matematica: quantità "discreta" cioè che si compone di parti separate. Il pensiero, che fino a ieri ci appariva come qualcosa di fluido, evocava in noi immagini lineari come un fiume che scorre o un filo che si sdipana, oppure immagini gassose, come una specie di nuvola, tant'è vero che veniva spesso chiamato «lo spirito», - oggi tendiamo a vederlo come una serie di stati discontinui di combinazione di impulsi su un numero finito (un numero enorme ma finito) di organi sensori e di controllo. I cervelli elettronici, se sono ancora lungi dal produrre tutte le funzioni del cervello umano, sono però già in grado di fornirci un modello teorico convincente per i processi più complessi della nostra memoria, delle nostre associazioni mentali, della nostra immaginazione, della nostra coscienza. Shannon, Wiener, Von Neumann, Turing, hanno cambiato radicalmente l'immagine dei nostri processi mentali. Al posto di quella nuvola cangiante che portavamo nella testa fino a ieri e del cui addensarsi o disperdersi cercavamo di rendervi conto descrivendo impalpabili stati psicologici, umbratili paesaggi dell'anima, - al posto di tutto questo oggi sentiamo il velocissimo passaggio di segnali sugli intricati circuiti che collegano i relé, i diodi, i transistor di cui la nostra calotta cranica è stipata» (Calvino, 2018: 205).

In un mondo discontinuo tutto diventa numerabile, calcolabile e dunque processabile da un computer e la letteratura stessa si trasforma in gioco combinatorio. Tuttavia, Calvino si rende conto della specificità della letteratura: a differenza delle altre discipline, comprese quelle umanistiche, prima tra tutte la linguistica, la letteratura aggiunge al semplice studio del linguaggio un surplus semantico ed esperienziale derivante dalla persona dell'autore che si moltiplica e differenzia nei lettori. Per questo motivo Calvino ritiene che una macchina in grado di ideare poesie e romanzi dovrà tener conto di «tutti quegli elementi che siamo soliti considerare i più gelosi attributi dell'intimità psicologica, dell'esperienza vissuta, dell'imprevedibilità degli scatti d'umore, i sussulti e gli strazi e le illuminazioni interiori» (Calvino, 2018: 208-209). Un'impresa apparentemente impossibile, per la quale Calvino richiama la teoria dei campi semantici: «Che cosa sono questi se non altrettanti campi linguistici, di cui possiamo benissimo arrivare a stabilire lessico grammatica sintassi e proprietà permutative?» (Calvino, 2018: 208). Le esperienze umane sono in primo luogo linguistiche e ognuna di esse può essere analizzata ed elaborata digitalmente attraverso il dominio del relativo campo semantico. La teoria dei campi semantici serve ad affermare l'essenza della letteratura come gioco combinatorio: lo scrittore non fa altro che riprendere, ricombinare e modificare a partire da elementi dati. Questa concezione permette di decostruire la figura dell'autore e di giustificare la generazione automatica di testi:

La letteratura come la conoscevo io era un'ostinata serie di tentativi di far stare una parola dietro l'altra seguendo certe regole definite, o più spesso regole non definite né definibili ma estrapolabili da una serie di esempi o protocolli, o regole che ci siamo inventate per l'occasione cioè che abbiamo derivato da altre regole seguite da altri. In queste operazioni la persona io, esplicita o implicita, si frammenta in figure diverse [...]. L'io dell'autore nello scrivere si dissolve: la cosiddetta «personalità» dello scrittore è interna all'atto dello scrivere, è un prodotto e un modo della scrittura. Anche una macchina scrivente, in cui sia stata immessa un'istruzione confacente al caso, potrà elaborare sulla pagina una «personalità» di scrittore spiccata e inconfondibile, oppure potrà essere regolata in modo di evolvere o cambiare «personalità» a ogni opera che compone. Lo scrittore quale è stato finora, già è macchina scrivente. (Calvino, 2018: 211)

La visione di Calvino risulta accettabile solamente se ci si stacca dall'estetica romantica, che vede la poesia come «una questione d'ispirazione discesa da non so quali altezze o sgorgante da non so quali profondità, o intuizione pura o momento non meglio identificato della vita dello spirito, o voce dei tempi con la quale lo spirito del mondo decide di parlare attraverso il poeta» (Calvino, 2018: 210). Una concezione che non specifica le modalità attraverso le quali si giunge alla pagina scritta: «Per quali vie l'anima e la storia o la società o l'inconscio si trasformano in una sfilza di righe nere su una pagina bianca? Su questo punto le più importanti teorie estetiche tacevano» (Calvino, 2018: 210).

Se l'estetica romantica può essere facilmente superata, più complesso il confronto con l'idea propria dell'estetica novecentesca dell'arte come mezzo per spingersi oltre i codici, tentativo di dire il non dicibile ed esplorare territori a prima vista non conoscibili: «La battaglia della letteratura è appunto uno sforzo per uscire fuori dai confini del linguaggio; è all'orlo estremo del dicibile che essa si protende; è il richiamo di ciò che è fuori dal vocabolario che muove la letteratura» (Calvino, 2018: 213). Per conciliare questo aspetto con l'idea della letteratura come gioco combinatorio e con l'esistenza di testi generati automaticamente, Calvino propone di considerare come letterario il testo capace di veicolare un «significato inatteso», ossia «un significato non oggettivo di quel livello linguistico sul quale ci stavamo muovendo, ma slittato da un altro piano, tale da mettere in gioco qualcosa che su un altro piano sta a cuore all'autore o alla società a cui egli appartiene» (Calvino, 2018: 217). In questo modo, la centralità del processo poetico si sposta sul momento della ricezione: «La macchina letteraria può effettuare tutte le permutazioni possibili in un dato materiale; ma il risultato poetico sarà l'effetto particolare d'una di queste permutazioni sull'uomo dotato d'una coscienza e d'un inconscio, cioè l'uomo empirico e storico, sarà lo shock che si verifica solo in quanto attorno alla macchina scrivente esistono i fantasmi nascosti dell'individuo e della società» (Calvino, 2018: 217).

Calvino pone dunque le basi per un'estetica dei testi generati tramite intelligenza artificiale, proponendo una visione del processo creativo svincolata dalla persona dell'autore e focalizzata sul momento della lettura quale azione centrale per la costruzione di senso. Se le modalità della creazione umana e di quella artificiale hanno in comune la ricombinazione di elementi dati all'interno di un campo semantico, allora i testi generati per mezzo di sistemi di intelligenza artificiale assumono una loro validità teorica, in quanto opere nelle quali il lettore umano può rintracciare un significato inatteso. L'autonomia dei sistemi di intelligenza artificiale dall'esperienza e dai processi mentali umani può creare immagini e idee che la mente umana non produrrebbe, perché influenzata da condizionamenti inconsci, della cultura e dell'esperienza reale del mondo, come dimostrano le esperienze di alcuni artisti contemporanei che si analizzeranno.

Quando Calvino scrive *Cibernetica e fantasmi*, la generazione automatica di testi era ancora ai suoi esordi e lavorava principalmente sulla ricombinazione di testi e frammenti che venivano forniti preventivamente ai calcolatori. Oggi l'intelligenza artificiale ha raggiunto ottimi risultati in campo di produzione di linguaggio naturale (NLP) con i Large Language Models (LLMs), sistemi di intelligenza artificiale in grado di generare testi che

ricalcano il linguaggio umano, apprendendo da grandi set di dati di addestramento<sup>6</sup>. Le innovazioni degli ultimi anni hanno permesso l'avvio di molti esperimenti di scrittura letteraria automatica, che utilizzano sia i LLMs che altre tecnologie di intelligenza artificiale, e che nel loro complesso possiamo catalogare all'interno della categoria di AI novel. Nel 2018, ad esempio, l'informatico Ross Goodwin ha pubblicato l'AI novel *1 the Road*, romanzo generato da un sistema di intelligenza artificiale che ha tradotto in scrittura i dati provenienti da GPS, microfoni e una videocamera posizionati su un'automobile durante un viaggio tra New York e New Orleans ispirato a *On the Road* di Jack Kerouac. I dati provenienti dai sensori posizionati dentro e fuori dall'automobile sono il punto di partenza per la narrazione, il testo prodotto in tempo reale deriva da elementi visivi, spaziali e dalle conversazioni dei passeggeri, e utilizza la scansione temporale come asse intorno al quale sviluppare il discorso.

Come già i ricercatori del Kimagure Artificial Intelligence Writer Project, anche Goodwin affronta la questione dell'autorialità in sede paratestuale. Nella nota biografica finale si legge «Ross Goodwin is not a poet. [...] He defines himself as an artist, a hacker, a code inventor and a gonzo data scientist» (Goodwin, 2018: 143)<sup>7</sup>. Nell'introduzione al libro Goodwin paragona la macchina da scrivere, che ha reso la scrittura a stampa accessibile a tutti, alla fotografia: «It took a creative product (the photorealistic image) that only a privileged few (painters with innate talent and years of training) could create, and put its means of production in the hands of a far larger population» (Goodwin, 2018: 8); e prosegue il paragone giungendo alle reti neurali: «Artificial neural networks are like cameras. Not because they capture raw information and output useful representations, but because they transform creative tasks that once required a measure of human work into mechanized actions» (Goodwin, 2018: 10). L'utilizzo delle reti neurali artificiali per la creazione artistica determina, secondo Goodwin, uno slittamento di ruoli nel processo creativo, che investe l'idea tradizionale dell'autorialità incarnata dalla persona dell'autore: «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» (Goodwin, 2018: 12).

---

<sup>6</sup> Per una panoramica sul funzionamento dei LLMs, tra cui rientrano i modelli GPT si veda (Ciotti, 2023).

<sup>7</sup> Nel suo sito personale (<https://rossgoodwin.com/>) Goodwin si definisce «data poet», che rimanda al concetto di datapoetica, utilizzato da diversi artisti contemporanei (D'Amico, 2020).

Il testo di *1 the Road* viene prodotto in tempo reale durante il viaggio, come se fosse scritto dall'automobile stessa. Ma subito l'idea dell'autorialità si frammenta e assume contorni non definibili in maniera univoca:

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. (Goodwin, 2018: 12-14)

Goodwin cerca di definire il proprio ruolo e quello del sistema: «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» (Goodwin, 2018: 14). La soluzione proposta è di staccarsi dalla visione classica dell'autorialità, che viene tacciata di antropocentrismo. Non possiamo ragionare sull'autorialità dell'AI novel con le stesse categorie con le quali siamo stati abituati per secoli a ragionare sull'autorialità umana: «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 itself» (Goodwin, 2018: 14).

La decostruzione dell'autorialità effettuata da Goodwin rimanda a quanto scritto da Calvino nel suo saggio del 1968. Chi sia l'autore non è rilevante perché la scrittura diventa un processo di co-creazione, in cui ciò che conta è l'esperienza stessa della generazione del testo e l'effetto che esso produrrà sul lettore. A proposito del risultato sul lettore, Goodwin parla anche del concetto di allucinazioni, termine con il quale oggi si designano gli errori dei sistemi di intelligenza artificiale, nel momento in cui vengono prodotti discorsi privi di senso o immagini incongruenti. Secondo Goodwin, le allucinazioni dei sistemi di intelligenza artificiale possono essere considerate una modalità per andare oltre i confini della mente umana:

Hallucination is the technical term for the activity of a machine learning system when it generates content, and technology is like a drug to the extent that it helps us reach beyond ourselves and attain experiences that might otherwise remain out of reach. Moreover, technology can be abused, and the most transformative or disruptive innovations often carry the highest potentials for abuse. (Goodwin, 2018: 14-16)

L'idea di allucinazione è stata indagata anche da alcuni artisti visuali, che hanno utilizzato il potere deformante e il margine di errore delle GAN (*Generative Adversarial*

*Networks*) per generare immagini da cui estrarre nuovo potenziale<sup>8</sup>. È il caso dell'opera di Anna Ridler, che utilizza quelli che definisce "artefatti", ossia gli errori e le incongruenze generate dalle GAN nella generazione di immagini, per ridefinire il proprio modo di disegnare e istituire nuovi nessi memoriali e semantici (Ridler, 2020). Il modo in cui le GAN modificano, imitandole, le immagini di input, è assimilabile al modo dell'artista umano di trasformare il passato e la tradizione artistica e letteraria rielaborandola in nuove forme (Barale, 2020).

Intorno al concetto di allucinazione ruota anche l'AI novel di Rocco Tanica, *Non siamo mai stati sulla Terra*, pubblicato nel 2022, nel quale dal dialogo tra autore umano e intelligenza artificiale emergono testi visionari in cui la realtà sensibile viene distorta e rimodulata in forme che alla lettura restituiscono diversi effetti, dal comico all'onirico, ricollegabili ai significati inattesi che Calvino stabiliva a fondamento della fruizione del prodotto letterario. Come Goodwin, anche Tanica riflette sull'autorialità del libro e opta per un'idea di co-autorialità tra essere umano e macchina, tant'è che in copertina nella porzione riservata al nome dell'autore si legge: «Rocco Tanica con OutOmat-B13». Il processo di scrittura collaborativa uomo-macchina è reso ancora più chiaro nel risvolto di copertina, che funge da cornice al testo:

Un uomo e un marchingegno robotico chiamato OutOmat-B13 sono seduti uno di fronte all'altro in un hotel di provincia. L'uomo interroga con curiosità la macchina, le chiede delle cose che ha visto, di ciò che conosce, di quello che immagina essere il pianeta. [...] Storia dopo storia, però, avviene che l'uomo prenda a completare con le sue parole i vuoti dell'altro: precisa un dettaglio, rattoppa una sfumatura, aggiunge un aneddoto. Umano e macchina si ritrovano così a raccontare assieme, alternati, fino a stendere sul mondo intero una patina talvolta surreale e comica, talvolta magica e visionaria, talvolta malinconica. (Tanica, 2022: risvolto di copertina)

L'idea dell'interazione tra uomo e macchina è alla base del concetto stesso di AI generativa: la cooperazione avviene in primo luogo attraverso l'inserimento di un input, una richiesta che il sistema automatico elabora e trasforma in un risultato; l'output può poi essere nuovamente interpretato e manipolato dall'artista, che è in grado di trovarvi nuovi significati o di creare nuove immagini. Il prodotto finale è il frutto di una collaborazione,

---

<sup>8</sup> Le reti generative avversarie sono state inventate nel 2014 da Ian Goodfellow. Si tratta di due reti neurali che interagiscono l'una contro l'altra. La prima, detta *discriminator*, è istruita a partire da alcuni dati (immagini, testi, suoni); l'altra, il *generator*, produce una nuova serie di dati, abbastanza simili a quelli iniziali da far sì che il *discriminator* li confonda con essi. In questo gioco entrambe le reti migliorano i propri risultati, producendo immagini che imitano trasformandole quelle iniziali. Grazie al rapporto tra generator e discriminator il processo di elaborazione dati sfugge maggiormente al controllo umano e può generare risultati inaspettati con cui l'artista può interagire.

nella quale agente umano e sistema artificiale concorrono alla creazione di un prodotto artistico che non potrebbe essere creato altrimenti. Mentre *1 The Road* presenta un maggior grado di automazione, resa possibile dall'utilizzo dei sensori, il libro di Tanica mette a pieno a frutto l'idea di co-autorialità, stabilendo anche graficamente una differenza tra le porzioni di testo generate dall'intelligenza artificiale e quelle scritte dall'autore umano, che appaiono rispettivamente in carattere corsivo e in tondo. Uomo e macchina instaurano un dialogo, che è allo stesso tempo un discorso condiviso, le due voci si intrecciano e ognuna delle due si rimodula continuamente a seconda della risposta generata dall'altra.

Se le modalità di produzione del testo sono differenti, comune tra *1 The Road* e *Non siamo mai stati sulla Terra* è il rimando a modelli letterari della tradizione novecentesca, seppur secondo modalità differenti. *1 the Road* si rifà al romanzo di Jack Kerouac e prende a modello la letteratura visionaria degli scrittori della Beat Generation, tanto che Goodwin arriva a istituire un parallelo tra l'uso di sostanze stupefacenti e la capacità allucinatoria della macchina. Goodwin non utilizza però testi appartenenti prettamente alla Beat Generation o di Kerouac, ma letteratura proveniente da fonti eterogenee. L'effetto "allucinato" è il risultato del modo in cui il sistema di intelligenza artificiale lavora. Diverso il caso di Tanica, il quale per rifarsi a una specifica tradizione letteraria decide di fornire al sistema una serie di autori ai quali ispirarsi, come Buzzati, Calvino, Dickens<sup>9</sup>. I modelli attinti dalla tradizione letteraria vengono forniti al sistema attraverso un processo chiamato *fine-tuning*: i LLMs sono degli strumenti *generic purpose*, che imparano il linguaggio umano senza essere orientati verso compiti specifici e si adattano alle diverse richieste comunicative; il *fine-tuning* permette di inserire, in un secondo momento, dei testi "modello" attraverso i quali il sistema raffina la sua modalità di risposta ed è in grado di riprodurre uno stile di scrittura o estrarre le caratteristiche di un genere letterario. Questa tecnica è stata utilizzata da Tanica per il suo libro, determinando una costellazione di riferimenti letterari impliciti che influenzano la scrittura, proprio come succederebbe in un testo scritto da un autore umano.

La creazione di prodotti artistici per mezzo di intelligenza artificiale mette in discussione l'unicità dell'essere umano quale autore e chiede una ridefinizione dell'idea di autorialità, che deve essere intesa come un processo. Bisogna considerare che i sistemi di

---

<sup>9</sup> In un'intervista rilasciata a *Focus*, Tanica ha affermato: «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» (Tartamella, 2022).

intelligenza artificiale lavorano sulla base di input provenienti dall'esterno e che l'intervento umano è necessario in diverse fasi del processo creativo ed è fondamentale nella connotazione finale del prodotto. L'autore come persona incarnata si decostruisce ma l'intenzionalità autoriale rimane umana e si manifesta in diverse fasi del processo, come l'addestramento dei sistemi, il *fine tuning*, la scelta dei *prompt* da inserire, l'interazione tra l'essere umano e la macchina nel momento della produzione effettiva del testo, il lavoro di editing e, non meno importante, la decisione finale di pubblicare l'opera. Ne consegue, dunque, un'*agency* è frantumata, ma pur sempre umana (Lana, 2020b).

La natura del *machine learning*, che si basa su un addestramento del sistema a partire da grandi quantità di esempi, costituiti da dati forniti dagli operatori umani (seppur di solito non controllati), restituisce l'idea del sistema di intelligenza artificiale come uno specchio dell'umano. I LLMs sono macchine scriventi, come quelle pensate da Levi e Calvino, ma a differenza dello scrittore umano non sono dotati di coscienza e comprensione semantica (Bultin, 2023; Bender, 2021). La scrittura, invece, è da sempre un'attività appannaggio di esseri coscienti (nella fattispecie esseri umani), che utilizzano lo strumento del linguaggio come espressione di moti interiori o per restituire la propria visione del mondo, filtrata dall'esperienza sociale e personale dell'individuo. I LLMs oggi sono in grado di costruire discorsi apparentemente umani, ma senza il filtro dell'esperienza umana, scrivono secondo esperienza indiretta e con questo si spiegano alcuni dei loro limiti, gli errori e le allucinazioni a cui a volte sono soggetti e che pure, come si è detto, rappresentano un elemento di interesse in ambito artistico. I processi di apprendimento che permettono ai sistemi di raggiungere tali risultati sono guidati da dati che contengono o riproducono esperienza umana e dunque sono assimilabili ad essa. L'idea dell'intelligenza artificiale come specchio ci spinge alla domanda cruciale: quale tipo di realtà vogliamo che riflettano? Che non è altro che una declinazione diversa della domanda che sta a monte di ogni scrittura letteraria: quale realtà voglio rappresentare? L'importanza di un approccio etico, comune a qualsiasi atto di scrittura, riafferma un'idea di autorialità in cui l'essere umano è ancora agente primario, non più in veste di unico produttore materiale del testo, ma come portatore di una visione del mondo moralmente connotata che non può essere prodotta autonomamente dai sistemi, ma solo riflessa da essi. Nei testi generati tramite sistemi di intelligenza artificiale non ci sarà forse più l'autore come lo abbiamo inteso per secoli, ma c'è ancora l'umanità; se non altro perché l'essere umano è insito nel linguaggio, il codice che solo l'intelligenza umana può produrre e decifrare.

Se volessimo tracciare qualche primissima conclusione sui prodotti letterari che abbiamo definito AI novel, dovremmo constatare come essi si presentano oggi nella forma della sperimentazione, e proprio nel campo delle scritture sperimentali trovano la loro migliore collocazione e si dimostrano più efficaci. Questo avviene grazie allo sfruttamento delle capacità allucinatorie dei sistemi, dai quali derivano testi dal carattere visionario e a volte surreale, con effetti che spaziano dal comico allo straniante. Secondo Calvino la vera vocazione di un ipotetico automa letterario sarebbe stata il classicismo, «il banco di prova d'una macchina poetico-elettronica sarà la produzione di opere tradizionali, di poesie con forme metriche chiuse, di romanzi con tutte le regole» (Calvino, 2018: 209). Calvino riteneva l'utilizzo dei calcolatori fatto fino ad allora dall'avanguardia per la scrittura combinatoria come qualcosa di «ancora troppo umano» (Calvino, 2018: 209). Negli anni in cui Calvino scriveva si era lontani dall'apprendimento automatico e dai generatori di linguaggio oggi a disposizione; ma in quel saggio lo scrittore inseriva un ulteriore passaggio, immaginando una macchina in grado in un certo momento di produrre avanguardia non perché programmata per farlo, ma «per sbloccare i propri circuiti intasati da una troppo lunga produzione di classicismo» (Calvino, 2018: 209). Una tale generazione spontanea di disordine e anomalia sembra in effetti vicina al concetto di allucinazione dei sistemi di intelligenza artificiale e nasce effettivamente come risposta negativa alla richiesta di ordine, di classicismo. Il discorso intorno alla visione e all'allucinazione meriterebbe di essere ulteriormente approfondito per mezzo di analisi ravvicinate dei testi, sfruttando l'intuizione comune a Goodwin e a Ridler secondo la quale proprio nel loro carattere di imprevedibilità, di sconnessione dall'esperienza della realtà e dalla mente umana, nell'errore e nello scarto le AI novel possano giustificare la loro esistenza attraverso una produzione di senso inaspettata. In questo modo potrebbe avvenire, nell'ottica indicata da Calvino, un inserimento dell'AI novel nel solco della tradizione letteraria (che non è necessariamente classicista) e si potrebbe evitare, inoltre, che la generazione automatica dei testi diventi solamente un'ulteriore modalità di impoverimento della letteratura nelle mani della parte più opportunistica dell'industria editoriale.

## Bibliografia

- BARALE Alice (2020), «Arte e intelligenza artificiale: alcune domande», in EAD (ed.), *Arte e intelligenza artificiale. Be my GAN*, Milano, Jaca Books, pp. 7-13.
- BENDER Emily, GEBRU Timnit, MCMILLAN-MAJOR Angelina, MITCHELL Margaret (2021), «On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?», in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pp. 610–623.
- BUTLIN Patrick, LONG Robert, ELMOZNINO Eric, BENGIO Yoshua, BIRCH Jonathan, CONSTANT Axel, DEANE George, et alii (2023), «Consciousness in Artificial Intelligence: Insights from the Science of Consciousness», arXiv, 22 agosto 2023, URL: <<http://arxiv.org/abs/2308.08708>>.
- CALVINO Italo (2018), *Una pietra sopra* [1980], Milano, Mondadori.
- CIOTTI Fabio (2023), «I Large Language Model e i “problemi difficili”», *Informatica Umanistica e Cultura Digitale: il blog dell' AIUCD* (blog), 22 agosto 2023, URL: <<https://infouma.hypotheses.org/2429>>.
- D'AMICO Flavia Dalila (2020), «Datapoiesis: La Trasparenza nella performance poetica dei dati», *Elephant & Castle*, n. 22 (giugno), URL: <<https://elephantandcastle.unibg.it/index.php/eac/article/view/289>>.
- GOODWIN Ross (2018), *1 The Road*, Paris, Jean Boite Editions.
- HENRICKSON Leah (2018), «Computer-Generated Fiction in a Literary Lineage: Breaking the Hermeneutic contract», *Logos*, n. 29/2-3, pp. 54-63.
- HENRICKSON Leah (2021), *Reading Computer-Generated Texts*, Cambridge University Press.
- LANA Maurizio (2022a), «Intelligenza artificiale e produzione di testi: una prospettiva storico-critica», *AIB Studi*, n. 62/1, pp. 169–196.
- LANA Maurizio (2022b), «Sistemi di Intelligenza Artificiale e problemi del concetto di autore. Riflessioni su prodotti editoriali recenti», *JLIS.It*, n. 13 (2), pp. 13-44.
- RIDLER Anna (2020), «Set di dati e decadenza: *Fall of the House of Usher*», in BARALE Alice (ed.), *Arte e intelligenza artificiale. Be my GAN*, Milano, Jaca Books, pp. 111-127.
- SATO Satoshi (2016), «A Challenge to the Third Hoshi Shinichi Award», *Proceedings of the INLG 2016 Workshop on Computational Creativity and Natural Language Generation*, Edinburgh, September 2016, pp. 31-35, URL: <<https://aclanthology.org/W16-5505.pdf>>.
- TANICA Rocco, e OutOmat-B13 (2022), *Non siamo mai stati sulla Terra*. Milano, Il Saggiatore.
- TARTAMELLA Vito (2022), «Ecco il primo libro scritto da un'IA (e il coautore è Rocco Tanica)» *Focus*, 5 novembre 2022, URL: <<https://www.focus.it/tecnologia/digital-life/primo-libro-intelligenza-artificiale-rocco-tanica>>.
- VAN HEERDEN Imke, ANIL Bas (2021), «AI as Author—Bridging the Gap Between Machine Learning and Literary Theory», *Journal of Artificial Intelligence Research*, n. 71, pp. 175-189.

YŪREI, Raitā (2015), *Konpyūta ga shōsetsu wo kaita hi*, a cura di SATOSHI Satō e MATSUBARA Hitoshi, Nagoya, Nagoya University, 2015 (trad. inglese di Haerin Shin), URL: <[https://my.vanderbilt.edu/aiethics/files/2019/01/The-Day-a-Computer-Writes-a-Novel\\_EngTrans.docx](https://my.vanderbilt.edu/aiethics/files/2019/01/The-Day-a-Computer-Writes-a-Novel_EngTrans.docx)>.

**Come citare questo articolo:**

Daniel Raffini, «La ridefinizione dell'autorialità nell'AI novel», in *InterArtes* [online], n. 3, “ArtIntelligence” (Francesco Pigozzo ed.), dicembre 2023, pp. 1-13, <<https://www.iulm.it/wps/wcm/connect/iulm/5847a270-6f31-488c-aa20-81345963255c/02+Raffini.pdf?MOD=AJPERES>>.