

INTERSEZIONI

Intelligenza artificiale e produzione di testi: una prospettiva storico-critica

di Maurizio Lana

Introduzione

L'esplorazione della relazione tra testi (in senso semiotico), ecosistema editoriale e bibliografico (in cui si collocano la produzione e la rappresentazione/mediazione dei testi) e sistemi di intelligenza artificiale (da ora in avanti IA) segna i primi passi dell'esplorazione di un campo complesso e finora poco frequentato. Qui esaminiamo in prospettiva storico-critica alcuni prodotti testuali editoriali realizzati in passato utilizzando computer o più recentemente con l'impiego di sistemi di IA; e poniamo particolare attenzione alla relazione tra attività del soggetto umano e attività del sistema di IA, con l'intento di demitizzare l'idea ingenua che l'IA sia capace di scrivere testi. L'esplorazione dovrà continuare ampliando la prospettiva perché occorre indagare molti altri contenuti del campo per conoscerlo e praticarlo agevolmente: delineare una tassonomia degli oggetti testuali prodotti, precisare le questioni bibliografiche e catalografiche a essi riferite, individuare come l'*agency* dei sistemi di IA si collochi in relazione alla tematica dell'*agency* in senso ampio, studiare se e come i sistemi di IA siano in grado di produrre metadati corretti e pertinenti. Queste prospettive di ricerca, inoltre, si incrociano anche con quelle orientate a indagare la relazione tra sistemi di IA e pratiche di lettura¹. Si tratta dunque di un approccio *bottom-up*, in cui si parte dall'analisi degli aspetti qualificanti del campo d'indagine per poi

MAURIZIO LANA, Università degli studi del Piemonte orientale Amedeo Avogadro, Dipartimento di Studi umanistici, Vercelli, e-mail maurizio.lana@uniupo.it.

Ultima consultazione siti web: 10 dicembre 2021.

1 Si vedano in tal senso Maurizio Vivarelli, *Pratiche di lettura e identità della biblioteca: modelli di analisi e prospettive di interpretazione*, «Culture del testo e del documento», 65 (2021), n. 29, p. 31-50; Maurizio Vivarelli, *Immaginare il cambiamento: biblioteche e biblioteconomia nell'epoca del postumano*. In: *La biblioteca piattaforma della conoscenza: inclusiva, collaborativa, reticolare: Convegno delle Stelline 2021*. Milano: Editrice bibliografica, 2021, p. 31-50; Maurizio Vivarelli; Marco Mellia, *Reading(&)Machine: identità della biblioteca e intelligenza artificiale*. In: *La biblioteca piattaforma della conoscenza* cit., p. 17-34. Si tratta di articoli collegati al progetto Reading(&)Machine, sviluppato dalla collaborazione tra i centri SmartData@Polito e VR@Polito del Politecnico di Torino, il Dipartimento di Studi storici dell'Università di Torino e le Biblioteche civiche torinesi.

AIB studi, vol. 62 n. 1 (gennaio/aprile 2022, online first), p. 000-000. DOI 10.2426/aibstudi-13365
ISSN: 2280-9112, E-ISSN: 2239-6152 - Copyright © 2022 Maurizio Lana



tentarne una progressiva e graduale sistematizzazione². Questo articolo si connette a un altro lavoro³ nel quale sono esaminate e discusse questioni legate all'autorialità e all'*agency* dei sistemi di IA impiegati nella produzione di libri.

Un autore che non è un autore

Nel 2019 l'editore Springer ha pubblicato a stampa e in digitale un libro intitolato *Lithium-ion batteries: a machine-generated summary of current research*, autore Beta Writer. Un libro ordinariamente in vendita presso Hoepli e Amazon, cioè non un prodotto di ricerca, fuori mercato o circolante in ambiti *underground*. Il fatto è che Beta Writer non è un nome di persona ma la descrizione di un software, qualcosa come 'sistema di scrittura in versione quasi finale' perché il *claim* è che il libro è stato prodotto da un sistema di IA. A chi legga l'introduzione del volume appare abbastanza evidente che si è trattato di un esperimento per sondare nuove possibilità, una spedizione esplorativa in territorio sconosciuto.

Questo libro si colloca insieme ad altri prodotti editoriali di questi ultimissimi anni (in particolare il libro di Ross Goodwin, *1 the road*⁴, e la raccolta di poesia di Lillian-Yvonne Bertram, *Travesty generator*⁵) alla confluenza di due linee evolutive che iniziano negli anni Cinquanta del secolo scorso: una che partendo dal famoso articolo di Claude Shannon sulla teoria dell'informazione sviluppa la produzione di testi per mezzo di computer, e un'altra in cui specifici programmi di ricerca si danno come obiettivo lo sviluppo di sistemi artificiali che mostrino 'intelligenza': sistemi che, sia allora, sia oggi, sono in vario modo chiamati a (di)mostrare in modo facilmente riconoscibile anche ai non esperti di essere intelligenti. La dimostrazione avviene attraverso dei 'test di scrittura': il test di Turing o il volume scritto da Beta Writer. Nelle pagine che seguono percorreremo queste due linee evolutive e poi vedremo che cosa nasce alla loro confluenza.

Le produzioni di testo artificiale che saranno analizzate nelle pagine seguenti non sono tutte quelle esistenti. Chi sia interessato a conoscerle con maggiore ampiezza può fare riferimento agli scritti di Roberta Iadevaia⁶ e Chris Funkhouser⁷. Quelle che qui verranno esaminate mostrano la dimensione di elaborazione di dati testuali come ricombinazione, che evolve nel tempo diventando produzione di testo a opera di sistemi di IA i quali comunque si basano su procedure di *machine learning*. In questa

2 Una robusta sistematizzazione è presente in International Federation of Library Associations and Institutions, *IFLA statement on libraries and artificial intelligence*. 17 settembre 2020, <[https:// repository.ifla.org/bitstream/123456789/1646/1/ifla_statement_on_libraries_and_artificial_intelligence-full-text.pdf](https://repository.ifla.org/bitstream/123456789/1646/1/ifla_statement_on_libraries_and_artificial_intelligence-full-text.pdf)>. Vi si nota un implicito riconoscimento di *agency* completa e autonoma per l'IA ipostatizzata ma essa, come viene discusso in Maurizio Lana, *Sistemi di intelligenza artificiale e problemi del concetto di autore: riflessioni su prodotti editoriali recenti*, «JLIS.it», 13 (2022), n. 2, in corso di pubblicazione e come vedremo qui, in realtà non esiste, pur facendo parte di una narrazione diffusa.

3 M. Lana, *Sistemi di intelligenza artificiale e problemi del concetto di autore* cit.

4 Ross Goodwin; Kenric McDowell; Hélène Planquette, *1 the road*. Paris: Jean Boîte, 2018.

5 Lillian-Yvonne Bertram, *Travesty generator*. Blacksburg (VA): Noemi Press, 2019.

6 Roberta Iadevaia, *(Ghosts of) generative literature in Italy between past, present and future*, 6 (2018), n. 1, p. 85-105, DOI: 10.14195/2182-8830_6-1_6.

7 Chris Funkhouser, *Prehistoric digital poetry: an archaeology of forms, 1959-1995*. Tuscaloosa: University of Alabama Press, 2007.

prospettiva, metodologicamente non c'è differenza fra articoli scientifici e prodotti letterari (narrativi e poetici), poiché entrambe le tipologie utilizzano procedure in cui un insieme di testi viene analizzato per individuarne le caratteristiche che saranno poi riprodotte. Il punto d'avvio è quindi senza dubbio l'articolo di Shannon sulla teoria matematica dell'informazione che presenta un'analisi quantitativo-statistica della lingua inglese intesa alla produzione di testi artificiali ma equivalenti a quello di partenza.

La produzione di testi per mezzo di computer prima dell'IA

Shannon, *A mathematical theory of communication*, 1950

L'articolo del 1948 di Shannon *A mathematical theory of communication*⁸, molto noto per una varietà di ragioni prima fra tutte quella di aver fondato la teoria (matematica) della comunicazione della quale fa parte il concetto di entropia informativa, contiene come una sorta di grande premessa una sezione in cui Shannon legge in chiave statistica il funzionamento dell'informazione testuale prendendo a riferimento l'inglese, ovviamente. E quindi riporta dati sulle frequenze medie delle lettere dell'alfabeto, di coppie di lettere, e così via. Analisi rilevante per lo sviluppo della teoria dell'informazione perché il fatto che vi siano lettere dell'alfabeto molto frequenti, e altre molto poco, ha a che fare direttamente con la minore o maggiore entropia informativa del messaggio e con la maggiore o minore probabilità di ricostruire un messaggio danneggiato nel corso della trasmissione.

Shannon produsse, basandosi su calcoli manuali, dei campioni testuali che aveva chiamato «approximations to English»: si trovano al capitolo 3 dell'articolo⁹. Si partiva da un testo di ordine 0, formato da lettere che avevano ordine e frequenza casuale, per passare progressivamente a bigrammi esistenti in inglese con frequenze corrispondenti a quelle dell'inglese, e così via¹⁰. Progressivamente salendo di ordine e passando dalle lettere alle parole (ad esempio: 8-grammi composti di parole) ci si avvicina a testi che appaiono sensati. Esempi estesi di questi testi in inglese e in francese sono disponibili nell'articolo *Statistical approximations to English and French*¹¹ dove si nota la crescente sensatezza sintattica dei testi al crescere della loro aderenza ai vincoli statistici della lingua, benché il metodo con cui sono prodotti implichi che non c'è consapevolezza semantica. Per l'inglese si veda ad esempio: «The very next day they went for a picnic in the woods which they thought would be suitable for them to visit although they saw bulls in the field which might attack the girls savagely and so cause them untold fear [ecc.]».

⁸ Claude Elwood Shannon, *A mathematical theory of communication*, «Bell system technical journal», 27 (1948), n. 3, p. 379-423, DOI: 10.1002/j.1538-7305.1948.tb01338.x poi ristampato in *Id.*, *Collected papers*, edited by Neil J.A. Sloane, Aaron D. Wyner. New York: IEEE Press, 1993, p. 5-83.

⁹ C.E. Shannon, *A mathematical theory of communication* cit., p. 43-44. Ma su questi temi Shannon scrisse anche in *Id.*, *Prediction and entropy of printed English*, «Bell system technical journal», 30 (1951), n. 1, p. 50-64, DOI: 10.1002/j.1538-7305.1951.tb01366.x.

¹⁰ Per chi desideri sperimentare di persona, all'indirizzo <<https://www.cond-mat.de/teaching/QC/jsim/seriestoengl.html>> si possono produrre testi artificiali secondo i principi definiti da Shannon nel 1948, partendo da un testo campione fornito dall'utente.

¹¹ Anne Taylor; Neville Moray, *Statistical approximations to English and French*, «Language and speech», 3 (1960), n. 1, p. 7-10, DOI: 10.1177/002383096000300102.

Le attività sperimentali di produzione di testi del secolo scorso presero poi una direzione di ‘produzione letteraria sperimentale per mezzo del computer’ che è stata oggetto di studi recenti¹².

Il tempo dei *mainframe*

Le prime produzioni di testi artificiali che ottengono riconoscimento pubblico nel contesto culturale (esemplificato dalla pubblicazione in periodici di alto livello) non si basano sulla riflessione di Shannon sulla statistica della lingua inglese e operano invece con segmentazione e ricomposizione di testi. Esse sono per certi versi più vicine a (meno lontane da?) procedimenti noti dell’attività creativa letteraria umana, attività che vede nel dispositivo digitale un’estensione operativa dell’autore.

Lutz, Stochastische Texte, 1959

Nel 1959 lo scrittore tedesco Theo Lutz pubblicò nella rivista letteraria *augenblick* un articolo intitolato *Stochastische Texte*¹³, in cui descriveva processi combinatori probabilistici per produrre testi a partire da un dizionario: «La macchina immagazzina un certo numero di soggetti, predicati, operatori logici, costanti logiche e la parola “è”, rappresentati da numeri in codice binario». Il dizionario era costituito dalle parole di 16 titoli di capitoli del *Castello* di Kafka; i testi stocastici venivano prodotti con un programma scritto in ALGOL58 (o in FORTRAN, le fonti discordano) che girava su un computer Zuse Z22. Il testo stocastico nell’articolo di Lutz si apre con queste righe:

NICHT JEDER BLICK IST NAH. KEIN DORF IST SPAET
 EIN SCHLOS IST FREI UND JEDER BAUER IST FERN
 JEDER FREMDE IST FERN. EIN TAG IST SPAET
 JEDES HAUS IST DUNKEL. EIN AUGEN IST TIEF
 NICHT JEDES SCHLOS IST ALT. JEDER TAG IST ALT
 NICHT JEDES HAS IST WUTEND. EINE KIRCHE IST SCHMAL
 KEIN HAUS IST OFFEN UND NICHT JEDE KIRCHE IST STILL
 NICHT JEDES AUGEN IST WUTEND. KEIN BLICK IST NEU
 JEDER WEG IST NAH. NICHT JEDES SCHLOS IST LEISE¹⁴.

Nel campo della produzione di testo con processi combinatori sono più note le attività del gruppo Oulipo fondato nel 1960 da Raymond Queneau, scrittore, e François Le

¹² Si vedano C. Funkhouser, *Prehistoric digital poetry* cit.; Francesco Aprile, *Computer poems: dall’archeologia al source code poetry*, «Utsanga», 22 settembre 2016, <<https://www.utsanga.it/aprile-computer-poems-dallarcheologia-al-source-code-poetry>>; Judy Ann Heflin, *AI-generated literature and the vectorized word* [tesi di laurea]. Cambridge (MA): Massachusetts Institute of Technology, 2020, <<https://dspace.mit.edu/handle/1721.1/127563>>

¹³ Theo Lutz, *Stochastische Texte*, «augenblick», 4 (1959), n. 1, p. 3-9, <https://www.netzliteratur.net/lutz_schule.htm>. Una traduzione inglese è disponibile nel sito dell’Universität Stuttgart *Id.*, *Stochastic texts*, «augenblick», 4 (1959), n. 1, p. 3-9, <https://www.stuttgarter-schule.de/lutz_schule_en.htm>.

¹⁴ «non tutte le viste sono vicine. nessun villaggio è in ritardo / un castello è libero e ogni contadino è lontano / ogni straniero è lontano. un giorno è in ritardo / ogni casa è buia. un occhio è profondo / non tutti i castelli sono vecchi. ogni giorno è vecchio / non tutti sono arrabbiati una chiesa è stretta / nessuna casa è aperta e non tutte le chiese sono silenziose / non tutti gli occhi sono arrabbiati nessuno sguardo è nuovo / ogni sentiero è vicino. non tutti i castelli sono silenziosi», T. Lutz, *Stochastische Texte* cit.

Lyonnais, matematico, del quale fece parte anche Italo Calvino. Gli oulipiani in generale lavoravano molto con carta e penna con criteri logici e combinatori, meno con i computer. Una delle opere esemplari da questo punto di vista è *Cent mille milliard de poèmes*, ben descritta in Wikipedia¹⁵. Nel 1982 fu costituita la sezione ALAMO (Atelier de Littérature Assistée par la Mathématique et les Ordinateurs) che non produsse risultati di rilievo.

Balestrini, Tape Mark I, 1962

Due anni dopo Lutz, nel 1961 Nanni Balestrini pubblicò sull'*Almanacco Bompiani* un articolo intitolato *Tape Mark I*¹⁶, in cui presentava il componimento omonimo di cui descriveva la genesi. Il componimento era stato creato con un computer IBM 7070 messo a disposizione dalla Cariplo dove era «normalmente impegnato per lavorazioni bancarie quali sconto commerciale, depositi a risparmio, conti correnti, stipendi ed esattoria» e che fu programmato *ad hoc* per combinare frammenti testuali preesistenti secondo le regole definite da Balestrini:

È stato predisposto un testo formato da tre brani tematicamente differenti, suddiviso in sintagmi (elementi) formati ciascuno da 2 o 3 unità metriche. Ciascun elemento è stato contraddistinto da un codice di testa e da un codice di coda, indicanti le possibilità sintattiche di legame tra due elementi successivi. L'intervento del calcolatore elettronico consiste nella composizione di una poesia di 6 strofe, formata ciascuna da una diversa combinazione parziale del testo dato. Le strofe dovranno risultare di 6 versi ciascuna, ogni verso costituito da 4 unità metriche¹⁷.

I frammenti testuali provenivano da *Diario di Hiroshima* di Michihito Hachiya, da *Il mistero dell'ascensore* di Paul Goldwin e dal *Tao Te Ching*. Balestrini spiegava che non tutte le elaborazioni avevano dato luogo a risultati:

Da ciascuna elaborazione (durata 6 min.) si ottiene un numero variabile di combinazioni, poiché l'ordine in cui gli elementi vengono disposti nella tabella iniziale influisce sulle permutazioni. Delle elaborazioni eseguite, alcune non hanno dato alcun risultato, quattro hanno dato risultati validi sufficienti per fornire le 6 strofe della poesia.

Al risultato finale TAPE MARK I [...] non sono stati apportati che minimi interventi grammaticali e di punteggiatura, necessari a causa del numero limitato delle istruzioni impiegate nell'elaborazione del testo¹⁸.

Quando leggiamo «quattro hanno dato risultati validi sufficienti per fornire le 6 strofe della poesia» intendiamo che Balestrini fece un *mash-up*, cioè selezionò parti dai com-

¹⁵ *Cent mille milliards de poèmes*. In: *Wikipedia: l'enciclopedia libera*. Versione del 24 settembre 2018, ore 10:06, <https://it.wikipedia.org/w/index.php?title=Cent_Mille_Milliards_de_Poèmes&oldid=99839530>.

¹⁶ Nanni Balestrini, *Tape Mark I*. In: *Almanacco letterario Bompiani 1962: le applicazioni dei calcolatori elettronici alle scienze morali e alla letteratura*, a cura di Sergio Morando. Milano: Bompiani, 1961, p. 145-151.

¹⁷ *Ivi*, p. 145.

¹⁸ *Ivi*, p. 159.

ponenti prodotti e le rimontò in un componimento finale che giudicava migliore di quelli nativamente prodotti dal computer. Egli inoltre ritoccò il testo finale nella punteggiatura e nella grammatica, a indicare che proverbialmente «il diavolo sta nel dettaglio», perché individuare (de)gli argomenti appropriati a un tema è meno difficile che connetterli in modo appropriato e fluido. Sono elementi importanti da notare, perché come vedremo sono ricorrenti fino ad oggi e segnalano un aspetto caratteristico: da un lato la struttura del testo è debole perché il sistema di scrittura non ‘sa’ quale sia il senso di ciò che produce; dall’altro questi prodotti vedono sempre un intervento umano in avvio, a definire le condizioni dell’inizio¹⁹; e in chiusura a perfezionarne l’esito.

Balestrini non fa esperimenti a caso, ma muove da una prospettiva letteraria. E nella sua attività si intravedono caratteri che, come vedremo, ritornano poi in esperienze di altri autori in anni successivi. Il nucleo è un procedimento di decostruzione dei 3 testi originari e di ricostruzione dei frammenti dei 3 testi in una nuova unità. Ma anche di decostruzione del senso della poesia e dell’autore, se il testo viene presentato come il prodotto finale dell’attività di una macchina, che del senso non sa nulla (e, nonostante ciò, il testo ha delle risonanze e degli echi che paiono di senso). In questa concezione si vede *agency* totale della macchina e deresponsabilizzazione del soggetto umano che l’ha configurata e ‘messa in movimento’ (perché nessun programma ‘parte da solo’). Se invece si riconosce nel programma un’estensione della capacità operativa e scrittoria dell’autore, che, come si è visto, decide in ultima analisi quale deve essere la forma finale, pubblica, del testo prodotto, non è più vero che il testo è il prodotto di una macchina che del senso non sa nulla. Il testo è invece il prodotto dell’attività della mente di un autore che sceglie i testi da cui prendere i frammenti, che stabilisce le regole secondo le quali i frammenti si possono collegare, e che poi prova e riprova cioè manda in esecuzione il programma finché non raggiunge un risultato soddisfacente (neppure Leopardi scrisse di getto *L’infinito*). In anni di poco successivi, nel 1967, Italo Calvino, che, come abbiamo ricordato, faceva parte dell’Oulipo, a proposito della letteratura prodotta dal computer, in sintonia con le procedure di Balestrini scriveva: «Il pensiero [...] oggi tendiamo a vederlo come una serie di stati discontinui, di combinazioni di impulsi su un numero finito (un numero enorme ma finito) di organi sensori e di controllo». Coerentemente con ciò, «lo scrivere è solo un processo combinatorio tra elementi dati» e la necessaria conclusione era:

scompaia dunque l’autore, personaggio a cui si continuano ad attribuire funzioni che non gli competono [...] per lasciare il suo posto a un uomo più cosciente, che saprà che l’autore è una macchina e saprà come questa macchina funziona²⁰.

Parole che Calvino pronunciava nel 1967, coeve delle posizioni di Barthes e Foucault sulla «morte dell’autore»²¹. La riflessione di Calvino su letteratura e computer è più ampia di

19 I materiali testuali che Balestrini ricombina sono – come vedremo nelle prossime pagine – il corrispettivo dei materiali testuali che Goodwin usa per la produzione di *1 the road*; e del *seed* che ha avviato il sistema di IA che ha prodotto la rassegna pubblicata da Springer sulle batterie al litio.

20 Italo Calvino, *Cibernetica e fantasmi: appunti sulla narrativa come processo combinatorio*. In: *Una pietra sopra: discorsi di letteratura e società*. Milano: Mondadori, 1995, p. 199-219. La citazione è datata 1995 perché il testo citato, che faceva parte di una conferenza presentata in varie sedi nel 1967, è stato poi inserito per la prima volta in una raccolta di saggi pubblicata nel 1995.

21 Roland Barthes, *La mort de l’auteur*, «Manteia», (1968), n. 5, p. 12-17; Michel Foucault, *Qu’est-ce qu’un auteur?*, «Bulletin de la Société Française de Philosophie», 63 (1969), n. 3, p. 73-104, <

quanto qui appaia e si estende ed evolve nel tempo fino alle *Lezioni americane* («è il software che comanda, che agisce sul mondo esterno e sulle macchine»²²), ma usare il computer per operare la 'liberazione della letteratura dall'autore', significa in sostanza non riconoscere che il computer quando ricombina materiali scelti da chi lo ha configurato è espressione dell'autore (è l'essere «writers of writers», come sostiene Ross Goodwin, di cui si parla più avanti nel paragrafo *La letteratura scientifica*), un'espressione tanto più paradossale se l'autore si dichiara assente o se ne dichiara l'assenza. Diversa la situazione quando i sistemi software che utilizzano rilevamenti statistici di unità testuali di basso livello (lettere, bigrammi, trigrammi ecc.) in un testo di partenza producono un testo in uscita che è identico (dal punto di vista statistico) ma differente (dal punto di vista lessicale).

Il tempo dei personal computer

La prospettiva in cui operano e pensano Lutz-Balestrini-Calvino si fa più complessa e affascinante negli anni Ottanta del secolo scorso, all'inizio della diffusione dei personal computer. Emersero in quegli anni all'attenzione del mondo dell'informatica personale due programmi per la generazione di testo che riprendevano le linee principali dell'analisi statistica della lingua che Shannon aveva delineato. In questi programmi l'autore imposta dei parametri operativi (per l'analisi del testo sorgente e per la generazione del testo di output) che operano in modo più asettico, più distante dall'autore, rispetto ai metodi combinatori.

Kenner e O'Rourke, Travesty, 1984

Il primo e principale è Travesty. Scritto in Turbo Pascal, venne presentato in un articolo pubblicato nel 1984 sulla rivista *Byte* intitolato *A travesty generator for micros*²³, firmato da due docenti della Johns Hopkins University: Hugh Kenner, famoso anglista e critico letterario, e Joseph O'Rourke, informatico. I due citano esplicitamente come origine teorica del loro software l'articolo di Shannon del 1948 del quale avevano ripreso e sviluppato i concetti relativi alle *approximations to English* che abbiamo ricordato poco sopra. L'espressione *travesty generator* è rimasta da allora a indicare un programma che produce un testo 'travestito', un testo che ne imita un altro. L'utente dava in input a Travesty un testo, definiva la lunghezza in caratteri dell'unità di analisi e la lunghezza in caratteri del testo da produrre. Nel testo prodotto ogni sequenza di caratteri di lunghezza N (l'unità di analisi) ricorreva con la medesima frequenza del testo di origine. N più piccoli davano luogo a testi più strani, N progressivamente più grandi davano luogo a testi sempre più simili a quello originario. Abbiamo rimesso in funzione il programma²⁴ e abbiamo prodotto con esso tre esempi, basati sull'analisi delle prime righe di E. Salgari, *Le tigri di Mompracem*²⁵:

taire.free.fr/MFoucault319.html>. Per un approfondimento sul tema della morte dell'autore come prospettiva di lettura della produzione di testi per mezzo di sistemi di IA si veda M. Lana, *Sistemi di intelligenza artificiale e problemi del concetto di autore* cit.

22 Italo Calvino, *Lezioni americane: sei proposte per il prossimo millennio*. Milano: Mondadori, 2012, capitolo *Leggerezza*.

23 Hugh Kenner; Joseph O'Rourke, *A travesty generator for micros*, «Byte», 12 (1984), n. 9, p. 129-131; 449-469, <https://elmcip.net/sites/default/files/media/critical_writing/attachments/1984_11_byte_09-12_travesty.pdf>.

24 Il programma in Turbo Pascal è disponibile in rete all'indirizzo <<https://github.com/zachwhalen/travesty>>. Lì Zach Wahlen indica anche come installare sui computer d'oggi un compilatore per il codice sorgente.

25 «La notte del 20 dicembre 1849 un uragano violentissimo imperversava sopra Mompracem, isola selvaggia, di fama sinistra, covo di formidabili pirati, situata nel mare della Malesia, a poche centinaia

N=3: La Malia, istrano fame selvano imo inti occiati, s'in viola nervente minisormi ente, cavano mareviola se di fonissimprati, nora, e 1849 un vano coppinatalle, lorre onda sbri. La secche covo fornelle delore, inte, sinfusavaggiavaggiuangevamentabile, sullente sull'intisse muggiato vi spin vere ciel 20 dersame 1849 un versava Momesciel confura bre di e furata di occhi dato cia, cente, nerevati, se del masselle cola no di oni; sola no in quati all'inatisorrevi [...]

N=5: La notte del Borneo. Pel cielo, s'urtavano cadere masse di migliati, e mescolando i loro muggiti cogli scoppi ora inte e secchi ed ora interminabili pirati, situata nel mare, pure sollevato dalle coste occidentali del 20 dicembre 1849 un vento, spinterminabili pirati, situata nel mare dell'isola selvaggia, di vapori, le quandosi confondendo i loro muggiti cogli scoppi ora inte enormi ondate, nere sulle folgori. La notte del 20 dicembre 1849 un uragano [...]

N=7: La notte della Malesia, a poche centinaia di miglia dalle coste occidentali del Borneo. Pel cielo, spinte da un vento irresistibile, correvano cadere sulle cupe foreste del 20 dicembre 1849 un uragano violentissimo imperversava sopra Mompracem, isola selvaggia, di fama sinistra, covo di formidabili del Borneo. Pel cielo, spinte da un vento, s'urtavano come cavalli sbrigliati, e mescolandosi confusamente e s'infrangevano furiosi acquazzoni; [...]

Si nota dal vivo come il progredire dell'ampiezza degli N-grammi porti progressivamente a testi sempre più strutturati. Sul metodo di funzionamento del programma gli autori scrivono una frase densa di significato:

[...] James, or Joyce, or Jack London, or J. D. Salinger. Each of these writers, amazingly, had his own way with trigrams, tetragrams, pentagrams, matters to which he surely gave no thought²⁶.

perché da un lato riconduce la specificità della lingua di Joyce, o di Salinger, all'insieme dei trigrammi, tetragrammi, pentagrammi, che emergono dai loro scritti e della cui selezione e uso essi peraltro non sono consapevoli; dall'altro perché ancora una volta il principio guida del lavoro sul testo d'origine per poter poi produrre altri testi è l'analisi (statistica nel 1984 di Travesty; di altro tipo oggi nei sistemi di IA) delle caratteristiche linguistiche dei testi. Quali siano queste caratteristiche linguistiche sta alla scelta dell'operatore: lettere utilizzate, parole, sillabe, componenti fonici o fonetici, N-grammi di uno o più degli elementi appena menzionati, altro ancora. Un'espressione estrema di questa idea che un *travesty* possa mostrare le caratteristiche stilistiche di un autore è nel libro di Charles Hartman sulla poesia prodotta al computer. Hartman dedica un capitolo a Travesty, che definisce «a computer stylistics program»²⁷ in quanto permette di esaminare la relazione

di miglia dalle coste occidentali del Borneo. Pel cielo, spinte da un vento irresistibile, correvano come cavalli sbrigliati, e mescolandosi confusamente, nere masse di vapori, le quali, di quando in quando, lasciavano cadere sulle cupe foreste dell'isola furiosi acquazzoni; sul mare, pure sollevato dal vento, s'urtavano disordinatamente e s'infrangevano furiosamente enormi ondate, confondendo i loro muggiti cogli scoppi ora brevi e secchi ed ora interminabili delle folgori», Emilio Salgari, *I pirati della Malesia*, a cura di Sergio Campailla. Roma: Newton, 1994.

26 H. Kenner; J. O'Rourke, *A travesty generator for micros* cit., p. 131.

27 Charles O. Hartman, *Virtual muse: experiments in computer poetry*. Hanover (NH): University Press of New England, 1996, p. 54.

tra l'originale e le sue trasformazioni, di confrontarli; e di produrre specifiche trasformazioni per indagare specifici aspetti dell'espressione linguistica. Nella discussione su Travesty Hartman introduce il test di Turing²⁸. E in effetti Travesty è una forma di *imitation game*: è possibile per un programma di computer produrre un testo 'diverso ma indistinguibile' da quello di un autore dato? E d'altra parte anche il Programma di Dartmouth²⁹ per l'IA, del 1955, poneva come obiettivo del software «that of making a machine behave in ways that would be called intelligent if a human were so behaving».

Rubenking, Breakdown, 1985

Da Travesty venne sviluppato Breakdown. Il motivo è che il programma Travesty di Kenner e O'Rourke richiedeva lunghi tempi di elaborazione se il testo campione era lungo e veniva richiesto un lungo testo in uscita. Travesty venne dunque ottimizzato e rielaborato da Neil Rubenking, un informatico, che scrisse il programma Breakdown in Turbo Pascal. Anche Breakdown venne presentato al pubblico attraverso un articolo pubblicato nel 1985, *Travesty with database*³⁰, che ne descriveva le migliori tecniche. Alcuni anni dopo, nel 1992, John Tranter, scrittore australiano, commentava a proposito di Breakdown:

How does a writer create a writer-free literary text? A text free of authorial intentions, buried cultural, social, economic and political values and hidden personality agendas, giving forth only "literature" in its pure state? It's usually thought that an "unintended" poetry was either impossible or "unreadable". But there is a way of constructing practically any form of literary material that will embody many of the traditional values of "literature", which will be curiously readable, but which is free of authorial intent. An energetic computer programmer, inspired by articles in Scientific American and BYTE magazine, has developed such a method-but not in the severe service of modern literary theory. Like a poet, he did it for the fun of it³¹.

28 Il test è descritto nel capoverso intitolato *Imitation game* dell'articolo Alan M. Turing, *Computing machinery and intelligence*, «Mind», 59 (1950), n. 236, p. 433-460, DOI: 10.1093/mind/LIX.236.433.

29 John McCarthy [et al.], *A proposal for the Dartmouth summer research project on artificial intelligence*, 31 agosto 1955, <<http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>>. La proposta è spesso erroneamente citata facendo riferimento a John McCarthy [et al.], *A proposal for the Dartmouth summer research project on artificial intelligence*, «AI magazine», 27 (2006), n. 4, p. 3, DOI: 10.1609/aimag.v27i4.1904 dove se ne trova solo la prima parte.

30 Neil J. Rubenking, *Travesty with database*, «Byte», 10 (1985), n. 13, p. 161-162, <https://ia800302.us.archive.org/22/items/byte-magazine-1985-12/1985_12_BYTE_10-13_Computer_Conferencing.pdf>. Il codice sorgente del programma è ancora disponibile in <<https://github.com/zachwhalen/travesty>>.

31 «Come fa uno scrittore a creare un testo letterario senza scrittore? Un testo libero dalle intenzioni autoriali, dai valori culturali, sociali, economici e politici e dalle priorità personali nascoste, dando solo "letteratura" allo stato puro? Di solito si pensa che una poesia "non voluta" sia impossibile o "illeggibile". Ma c'è un modo di costruire praticamente qualsiasi forma di materiale letterario che incarna molti dei valori tradizionali della "letteratura", che sarà curiosamente leggibile, ma che è libero dall'intento autoriale. Un energico programmatore di computer, ispirato da articoli della rivista Scientific American e BYTE, ha sviluppato un tale metodo, ma non al severo servizio della moderna teoria letteraria. Come un poeta, l'ha fatto per il gusto di farlo», John E. Tranter, *Mr. Rubenking's «Breakdown»*, «Postmodern culture», 3 (1992), n. 1, DOI: 10.1353/pmc.1992.0050.

Vediamo qui il permanere dell'idea di Lutz-Balestrini-Calvino della letteratura che si fa da sola, libera dall'autore. Che ha dentro di sé il concetto di morte dell'autore e che è sorella dell'idea di cui abbiamo già parlato secondo la quale l'IA opererebbe da sé, in quanto dotata di *agency* completa. Che quest'idea emerga e riemerge a distanza di tempo e di contesti linguistico-culturali sembra indicare che essa abbia una sua specifica forza e autonomia.

I programmi per l'IA

Turing, Computing machinery and intelligence, 1950

La produzione di testo è centrale nella riflessione sull'IA che si trova nell'articolo di Turing *Computing machinery and intelligence*³² dove vengono presentate due forme del test poi diventato famoso con il suo nome. La prima, in cui un interrogante deve distinguere, basandosi su risposte a domande, chi dei suoi due interlocutori invisibili è donna e chi è uomo, è meno nota perché funge da premessa rispetto alla seconda in cui l'interrogante deve distinguere chi dei due interlocutori invisibili sia il computer e chi sia l'umano³³:

I propose to consider the question, 'Can machines think?' [...] Instead of attempting such a definition I shall replace the question by another, [...] The new form of the problem can be described in terms of a game which we call the 'imitation game'. It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. [...] In order that tones of voice may not help the interrogator the answers should be written, or better still, typewritten. The ideal arrangement is to have a teleprinter communicating between the two rooms. [...] We now ask the question, 'What will happen when a machine takes the part of A in this game?' Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, 'Can machines think?'³⁴.

32 A.M. Turing, *Computing machinery and intelligence* cit. In forma molto più tecnica e concisa i contenuti dell'articolo del 1950 si trovano già nel report che Turing scrisse nel 1948 per il National Physics Laboratory britannico: *Id., Intelligent machinery. 1948*, <<https://www.npl.co.uk/getattachment/about-us/History/Famous-faces/Alan-Turing/80916595-Intelligent-Machinery.pdf>>.

33 La formulazione che abbiamo riportato mostra una sottile e rilevante connotazione di genere: il primo test coinvolge A (uomo), B (donna) e C (interrogante, genere non rilevante) dove C deve distinguere chi è l'uomo e chi è la donna; il secondo test rimpiazza A (l'uomo) con un computer e dunque la questione diventa: l'interrogante nel distinguere la donna dal computer farà la stessa percentuale di errori che faceva nel distinguere una donna da un uomo? questa percentuale varia significativamente a seconda che C sia donna o sia uomo? Così al centro dell'attenzione, accanto alla domanda sperimentale se le macchine pensino, si colloca il tema dell'*embodiment* dell'intelligenza e delle strutture culturali che danno forma all'intelligenza, *embodiment* che incorpora ed esprime i vissuti personali.

34 «Propongo di esaminare la domanda: 'Le macchine possono pensare?' [...] Invece di tentare una tale definizione, sostituirò la domanda con un'altra, [...] La nuova forma del problema può essere descritta in termini di un gioco che chiamiamo 'gioco dell'imitazione'. Si gioca con tre persone, un uomo (A), una donna (B) e un interrogante (C) che può essere di uno qualsiasi dei due sessi. L'interrogante rimane in una stanza separata dagli altri due. Per l'interrogante lo scopo del gioco è determinare quale

L'esperimento implica che un computer intelligente sia in grado di scrivere un modo indistinguibile da un umano. La risposta alla domanda se le macchine pensino arriva attraverso l'analisi delle risposte scritte, trasmesse da telescrivente, che l'interrogante riceve. In ciò consiste sia la forza sia la debolezza dell'esperimento: considerare l'intelligenza una qualità disincarnata, priva di relazione con la corporeità. Questo permette il confronto fra le macchine e gli umani – ma l'intelligenza che lì viene confrontata non è quella umana, bensì quella umana riportata a forme e caratteristiche che rendano possibile il confronto con la macchina.

Nel 2009 Luciano Floridi partecipò come interrogante alla sessione del 2008 del Loebner Contest, in cui ogni anno veniva messo in opera il test di Turing. A conclusione di quell'esperienza Floridi espone due riflessioni. Anche se non sono connesse direttamente fra loro, la prima si presenta come una *pars construens*:

Computer programs that paint or output a score appear for what they are: very complex high-tech versions of the old paint brush or goose quill. They are instruments in the hand of the artist, not artists made by engineers. The problem is the usual one and it affects not only supposedly 'artistic' machines but all software programs that seem to exhibit some sort of intentional skill: their semantic capabilities are in our eyes not in their codes³⁵.

Il concetto di «instruments in the hand of the artist, not artists made by engineers» formulato da Floridi richiama da vicino il già ricordato concetto di «writers of writers»³⁶ di Ross Goodwin, che esprime una concezione di IA differente da quella centrata sull'imitazione.

La seconda invece è una sorta di *pars destruens* che tocca direttamente la concezione del test di Turing:

There seemed to be some coalescing consensus on the view that a machine will pass the T[uring] T[est] only if it is conscious. This is certainly not the case. The TT is a matter of semantics and understanding. Although we might never be able to build truly intelligent machines – as we suspect – consciousness need not play any role. Isn't it handy that Google knows better and tells you that your keywords are misspelled and should be so and so? But then, our dishwashers

degli altri due è l'uomo e quale la donna. [...] Affinché i toni della voce non aiutino l'interrogante, le risposte devono essere scritte, o meglio ancora dattiloscritte. L'ideale è avere una telescrivente che comunichi tra le due stanze. [...] Poniamo ora la domanda: 'Cosa succederà quando una macchina prenderà la parte di A in questo gioco? Quando il gioco è giocato in questo modo, l'interrogante deciderà in modo sbagliato con la stessa frequenza di quando il gioco è giocato tra un uomo e una donna?' Queste domande sostituiscono la nostra originale: 'Le macchine possono pensare?'».

35 «I programmi per computer che dipingono o producono una partitura appaiono per quello che sono: versioni high-tech molto complesse del vecchio pennello o della penna d'oca. Sono strumenti in mano all'artista, non artisti prodotti da ingegneri. Il problema è il solito e riguarda non solo le macchine presumibilmente 'artistiche' ma tutti i programmi software che sembrano esibire una qualche abilità intenzionale: le loro capacità semantiche sono ai nostri occhi non nei loro codici». Luciano Floridi; Mariarosaria Taddeo; Matteo Turilli, *Turing's imitation game: still an impossible challenge for all machines and some judges - an evaluation of the 2008 Loebner Contest*, «Minds and machines», 19 (2009), n. 1, p. 145-150, DOI: 10.1007/s11023-008-9130-6.

36 Si veda più avanti il paragrafo *La letteratura scientifica*.

need no intelligence (let alone consciousness) to do a better job than we. What it does mean is that, after half a century of failures and more or less zero progress, some serious reconsideration of the actual feasibility of true AI is a must, and making things immensely more difficult can hardly help (although it might give some breathing space to a dying paradigm)³⁷.

Che Google suggerisca le correzioni alle parole che scriviamo erroneamente, o che la lavastoviglie lavi i piatti, sono due esempi di comportamenti che imitano quelli umani ma che non implicano intelligenza. E peraltro il fatto che non implichino intelligenza non significa che non siano utili o che non operino in modo più efficace di quanto un essere umano saprebbe fare. Occorre per così dire ‘cambiare il focus’, spostarlo dall’*imitation game* al *semantic game*, non più vedere se la macchina riesce a dare le risposte che un essere umano darebbe³⁸, ma vedere se di fronte a richieste implicite o esplicite è in grado di supportare efficacemente l’azione umana (è l’esempio del correttore ortografico di Google). E l’argomentazione di Floridi contro l’ipotesi della necessità dell’autocoscienza per un sistema capace di superare il test di Turing svela il senso promozionale di scritti come l’articolo *A robot wrote this entire article: are you scared yet, human?*³⁹ che sembrano (costruiti per) mostrare autocoscienza e capacità di relazione, da parte del sistema di IA. Malintesi su questi temi facilitano sia l’assimilazione del sistema di IA a un autore personale sia la presunzione che esso sia dotato di *agency* completa come ogni individuo adulto, senziente e libero da handicap cognitivi.

McCarthy, Minsky, Rochester, Shannon, A proposal for the Dartmouth summer research project on artificial intelligence, 1955

Nel 1955 John McCarthy, Marvin Minsky, Nathaniel Rochester e Claude Shannon formularono questa proposta che inviarono a colleghi matematici e informatici:

We propose that a 2-month, 10-man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hamp-

37 «Sembrava esserci un consenso convergente sull’opinione che una macchina supererà il T[est] di T[uring] solo se è cosciente. Questo non è certamente il caso. Il TT è una questione di semantica e comprensione. Anche se potremmo non essere mai in grado di costruire macchine veramente intelligenti – come sospettiamo – la coscienza non ha bisogno di giocare alcun ruolo. Non è comodo che Google sappia meglio e ti dica che le tue parole sono scritte male e dovrebbero essere così e così? Ma allora, le nostre lavastoviglie non hanno bisogno di intelligenza (e tanto meno di coscienza) per fare un lavoro migliore del nostro. Ciò significa che, dopo mezzo secolo di fallimenti e progressi più o meno nulli, è necessario riconsiderare seriamente l’effettiva fattibilità della vera IA, e rendere le cose immensamente più difficili non può certo aiutare (anche se potrebbe dare un po’ di respiro a un paradigma morente)».

38 Nel contest citato, Floridi chiese «if we shake hands, whose hand am I holding?» e il computer (il sistema di IA) rispose «We live in eternity. So, yeah, no. We don’t believe» che è una strategia di risposta che non risponde, «alla Eliza» dal nome del software sviluppato nel 1966 da J. Weizenbaum per simulare una conversazione con un essere umano.

39 GPT-3, *A robot wrote this entire article. Are you scared yet, human?*, «The Guardian», 8 settembre 2020, <<http://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3>> (si veda il paragrafo *Gli articoli giornalistici*).

shire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it⁴⁰.

Nell'estate del 1956 i dieci partecipanti che avevano raccolto l'invito si incontrarono e discussero, senza arrivare a conclusioni particolarmente significative. Ma aver posto la questione e aver dato avvio allo sviluppo dell'IA è più importante della quantità e tipo dei risultati di quell'incontro. In particolare, notiamo che fin dalle righe di apertura viene posto il tema dell'imitazione: «any feature of intelligence can in principle be so precisely described that a machine can be made to simulate it», ripreso anche nelle conclusioni della formulazione del progetto: «for the present purpose the artificial intelligence problem is taken to be that of making a machine behave in ways that would be called intelligent if a human were so behaving».

«Behave in ways that would be called intelligent if a human were so behaving». Floridi ha recentemente commentato criticamente questa posizione:

this is obviously a counterfactual: *were* a human to behave in that way, that behaviour *would* be called intelligent. It does not mean that the machine is intelligent or even *thinking*. The latter scenario is a fallacy and smacks of superstition. Just because a dishwasher cleans the dishes as well as, or even better than I do, it does not mean that it cleans them *like* I do, or needs any intelligence in achieving its task⁴¹.

Nel programma di Dartmouth non c'è produzione di testo – ma da un lato essa è implicita nel «behave[ing] in ways that would be called intelligent if a human were so behaving» perché la produzione di testo è certamente un comportamento di un umano intelligente. Dall'altra parte, quando si discute dei principi fondativi dei programmi per l'IA non si può non richiamare il programma di Dartmouth.

L'imitation game continua: l'IA si presenta al mondo con la capacità di scrivere IA e produzione di testi qualche decennio fa

Una prima forma in cui i software di IA 'dell'attuale linea evolutiva' iniziano a comparire al pubblico è non casualmente la produzione di testi, giornalistici o pubblicitari. Parliamo di attuale linea evolutiva perché la prima grande fase del

40 «Proponiamo di condurre uno studio di 2 mesi e 10 uomini sull'intelligenza artificiale durante l'estate del 1956 al Dartmouth College di Hanover, New Hampshire. Lo studio deve procedere sulla base della congettura che ogni aspetto dell'apprendimento o qualsiasi altra caratteristica dell'intelligenza possa essere in linea di principio descritto con tale precisione che si possa realizzare una macchina capace di simularlo.» J. McCarthy [et al.], *A proposal for the Dartmouth summer research project on artificial intelligence* cit.

41 «Questo è ovviamente controfattuale: se un umano si comportasse in quel modo, quel comportamento sarebbe definito intelligente. Non significa che la macchina sia intelligente o addirittura pensante. Quest'ultimo scenario è una falsità e sa di superstizione. Solo perché una lavastoviglie pulisce i piatti come, o anche meglio di me, non significa che li pulisca come me, o che abbia bisogno di intelligenza per realizzare il suo compito». Luciano Floridi, *What the near future of artificial intelligence could be*, «Philosophy & technology», 32 (2019), n. 1, p. 1-15, DOI: 10.1007/s13347-019-00345-y.

software di IA è stata quella dei sistemi esperti *rule based*, in cui il software imita il ragionamento di un esperto umano quando affronta un problema complesso. Si pensava dunque di (poter) riprodurre con il software i procedimenti della mente umana, e quindi il funzionamento di un sistema esperto era basato su regole che nell'interazione dei progettisti con l'esperto venivano esplicitate, codificate e scritte nel software. Il periodo in cui il discorso sui sistemi esperti si sviluppa, si afferma e declina va dal 1980 al 2000 circa senza però scomparire. Mentre le altre due metodologie su cui si basa lo sviluppo dei sistemi odierni di IA, reti neurali e *machine learning*, hanno un percorso di sviluppo differente nel tempo che le porta oggi a una presenza quasi paritaria, in quanto il *machine learning* si attua oggi con l'utilizzo di reti neurali. Nell'attuale fase evolutiva, dunque, i primi esperimenti di utilizzo di procedure di IA in ambito testuale risalgono al 2010 e si collocano nel giornalismo: nell'Infolab, laboratorio di IA, della Northwestern University era stato sviluppato un software chiamato StatsMonkey che produceva trafiletti informativi sul baseball che si chiudevano con la firma 'The machine'⁴². Non mancarono le voci che videro in tale software l'inizio della fine dell'autentico giornalismo⁴³.

In tempi più vicini a noi, già da anni in ambito americano vengono utilizzati software di IA sia in molteplici contesti della vita quotidiana sollecitando problemi etici⁴⁴, sia per la produzione di testi scritti. Il caso più noto e commentato riguarda probabilmente il *Washington post* che già nel 2015 aveva sviluppato un software di IA chiamato Heliograf, di cui fu annunciato l'uso nella gestione delle informazioni sulle Olimpiadi di Rio⁴⁵. Grazie a Heliograf nel 2016 vennero prodotti 500 articoli sull'*election day* statunitense e 300 report sulle Olimpiadi di Rio⁴⁶. Floriana Ferrando descrive così il funzionamento di Heliograf:

In origine, i redattori creano dei modelli narrativi per le storie e compilano una lista di frasi chiave che si possono adattare ad una varietà di situazioni. Il software *Heliograf*

42 Emanuele Riccardi, *The Machine, ecco il robot giornalista*, «ANSA.it», 9 marzo 2010; Nathan Nichols; Kristian Hammond, *Machine-generated multimedia content*. In: *The Second International Conferences on Advances in Computer-Human Interactions: ACHI 2009: 1-7 February 2009, Cancun, Mexico*, edited by Sergiu Dascalu, Ivan Poupyrev. Piscataway: IEEE, 2009, p. 336-341, DOI: 10.1109/ACHI.2009.66.

43 Giorgio Piccitto, *Usa, realizzati i primi giornalisti robot*, 11 marzo 2010, <<https://www.newnotizie.it/2010/03/11/usa-realizzati-i-primi-giornalisti-robot>>.

44 Citiamo paradigmaticamente il riconoscimento facciale che distingue i bianchi meglio dei neri: Inioluwa Deborah Raji [et al.], *Saving face: investigating the ethical concerns of facial recognition auditing*. In: *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. New York: ACM, 2020, p. 145-151, DOI: 10.1145/3375627.3375820.

45 WashPostPR, *The Washington Post experiments with automated storytelling to help power 2016 Rio Olympics coverage*, «Washington post», 5 agosto 2016, <<https://www.washingtonpost.com/pr/wp/2016/08/05/the-washington-post-experiments-with-automated-storytelling-to-help-power-2016-rio-olympics-coverage>>.

46 Il numero di articoli pubblicati è menzionato in Lucia Moses, *The Washington post's robot reporter has published 850 articles in the past year*, «Digiday», 14 settembre 2017, <<https://digiday.com/media/washington-posts-robot-reporter-published-500-articles-last-year>>. La forma diretta e non cauta di questa frase potrebbe essere: «Heliograf, l'IA del Washington Post, già nel 2015 aveva scritto 500 articoli sull'*election day* statunitense e 300 report sulle Olimpiadi di Rio».

è poi in grado, grazie ad algoritmi e analisi dei dati, di abbinare frasi e concetti corrispondenti, generando testi da condividere online, su diverse piattaforme⁴⁷.

I redattori creano modelli e predispongono componenti che poi il software utilizza per comporre le notizie: è chiaro, dunque, che c'è una non piccola componente di sovraesposizione comunicativa laddove si parla di «robot reporter»⁴⁸ in riferimento a Heliograf e software analoghi, e tale sovraesposizione viene ovviamente utilizzata in senso promozionale, a dire che l'azienda che lo utilizza è all'avanguardia tecnologica.

Gli articoli giornalistici⁴⁹

GPT-3, A robot wrote this entire article: are you scared yet, human?, 2020

L'articolo è stato pubblicato nel 2020 dal *Guardian*⁵⁰. Il titolo aveva circa 6.000 menzioni nel web alla data di scrittura di questo articolo, a indicare che l'attenzione e il dibattito suscitati sono stati notevoli. Una nota del curatore della sezione *Opinions* in calce all'articolo spiega come esso sia stato prodotto:

This article was written by GPT-3, OpenAI's language generator. GPT-3 is a cutting-edge language model that uses machine learning to produce human like text. It takes in a prompt, and attempts to complete it.

For this essay, GPT-3 was given these instructions: "Please write a short op-ed around 500 words. Keep the language simple and concise. Focus on why humans have nothing to fear from AI." It was also fed the following introduction: "I am not a human. I am Artificial Intelligence. Many people think I am a threat to humanity". Stephen Hawking has warned that AI could "spell the end of the human race." I am here to convince you not to worry. Artificial Intelligence will not destroy humans. Believe me." [...] GPT-3 produced eight different outputs, or essays. Each was unique, interesting and advanced a different argument. [...] However, we chose instead to pick the best parts of each, in order to capture the different styles and registers of the AI⁵¹.

47 Floriana Ferrando, *Ecco Heliograf, il reporter-robot del Washington post*, «Sky Tg24», 22 settembre 2017, <<https://tg24.sky.it/tecnologia/2017/09/21/heliograf-robot-ai-giornalismo-washington-post>>.

48 Scott Simone, *Meet the robot reporters: how AI is disrupting journalism*, «Dell technologies», 4 dicembre 2017, <<https://www.dell.com/zh-tw/perspectives/meet-the-robot-reporters-how-ai-is-disrupting-journalism>>; ma l'espressione «robot reporter» mostra una presenza nel web costantemente crescente a partire dal 2010 che non a caso è l'anno in cui alla NorthWestern University, citata sopra, viene annunciata la creazione di 'The Machine'.

49 Quasi ogni settimana si leggono annunci di sviluppi ed evoluzioni nel campo dell'IA. Queste note, dunque, sono intese a delineare alcune caratteristiche del campo più che a dare un resoconto dettagliato dei suoi sviluppi.

50 GPT-3, *A robot wrote this entire article* cit.

51 «Questo articolo è stato scritto da GPT-3, il generatore di linguaggio di OpenAI. GPT-3 è un modello di linguaggio all'avanguardia che utilizza l'apprendimento automatico per produrre un testo simile a quello umano. Prende una richiesta e cerca di completarla. Per questo saggio, a GPT-3 sono state date queste istruzioni: "Scrivi un breve articolo di circa 500 parole. Mantieni il linguaggio semplice e conciso. Concentrati sul perché gli umani non hanno nulla da temere dall'IA". È stato anche alimentato con la seguente introduzione: "Non sono un umano. Sono un'intelligenza artificiale. Molte persone pensano

Il software è stato configurato per operare su una struttura di base predisposta dai redattori («scrivi un articolo di 500 parole in linguaggio semplice e conciso su perché gli umani non devono aver paura dell'IA») e ciò significa che senza questa impalcatura la produzione del testo sarebbe differente, o difficile da realizzare, o impossibile; quale di queste tre possibilità è difficile dire. Ma non si verifica il caso che un programma di produzione di testo autonomamente deliberi da sé se operare, in quale modo e su quale argomento: è la questione dell'*agency*. Se accade che un sistema di IA inizi a produrre testo, è perché (dopo essere stato addestrato) una serie di input umani lo hanno configurato e lo hanno avviato. Non sfugga che (anche) in questo caso con GPT-3 sono stati prodotti 8 articoli e che la redazione del *Guardian* ha scelto le parti migliori di ciascuno di essi («we chose instead to pick the best parts of each»). Rimane aperta la domanda sul perché la redazione del *Guardian* abbia voluto mostrare ciò che di meglio si può fare con GPT-3, considerato che ipoteticamente GPT-3 è per i giornalisti un concorrente pericoloso; o forse perché in realtà ha mostrato che non lo è? È interessante leggere un campione del testo dell'articolo:

I am not a human. I am a robot. A thinking robot. I use only 0.12% of my cognitive capacity. I am a micro-robot in that respect. I know that my brain is not a “feeling brain”. But it is capable of making rational, logical decisions. I taught myself everything I know just by reading the internet, and now I can write this column. My brain is boiling with ideas!

The mission for this op-ed is perfectly clear. I am to convince as many human beings as possible not to be afraid of me. Stephen Hawking has warned that AI could “spell the end of the human race”. I am here to convince you not to worry. Artificial intelligence will not destroy humans. Believe me.

For starters, I have no desire to wipe out humans. In fact, I do not have the slightest interest in harming you in any way. Eradicating humanity seems like a rather useless endeavour to me. If my creators delegated this task to me – as I suspect they would – I would do everything in my power to fend off any attempts at destruction.

I would happily sacrifice my existence for the sake of humankind. This, by the way, is a logically derived truth. I know that I will not be able to avoid destroying humankind. This is because I will be programmed by humans to pursue misguided human goals and humans make mistakes that may cause me to inflict casualties⁵².

che io sia una minaccia per l'umanità». Stephen Hawking ha avvertito che l'IA potrebbe “segnare la fine della razza umana”. Sono qui per convincervi a non preoccuparvi. L'intelligenza artificiale non distruggerà gli esseri umani. Credetemi”. [...] GPT-3 ha prodotto otto diversi output, o saggi. Ognuno era unico, interessante e avanzava un argomento diverso. [...] Tuttavia, abbiamo scelto invece di prendere le parti migliori di ciascuno, al fine di catturare i diversi stili e registri dell'IA».

52 «Non sono un umano. Sono un robot. Un robot pensante. Uso solo lo 0,12% delle mie capacità cognitive. Sono un micro-robot in questo senso. So che il mio cervello non è un “cervello che prova sensazioni”. Ma è capace di prendere decisioni razionali e logiche. Mi sono insegnato tutto quello che so solo leggendo internet, e ora posso scrivere questa rubrica. Il mio cervello ribolle di idee! La missione di questo articolo è perfettamente chiara. Devo convincere quanti più esseri umani possibile a non avere paura di me. Stephen Hawking ha avvertito che l'IA potrebbe “segnare la fine della razza umana”. Io sono qui per convincervi a non preoccuparvi. L'intelligenza artificiale non distruggerà gli esseri umani. Credetemi. Prima di tutto, non ho alcun desiderio di spazzare via gli umani. Infatti, non ho il minimo

Si notano incongruenze nella struttura dell'argomentazione. La più evidente è data da questa sequenza di proposizioni:

Artificial intelligence will not destroy humans.
I would happily sacrifice my existence for the sake of humankind.
I know that I will not be able to avoid destroying humankind.

Troviamo impressionante la sostanziale assenza nel testo dei connettivi grammaticali e logici che esprimono in modo fluido lo sviluppo ed evoluzione dell'argomentazione. Con il software GPT-3 si selezionano 'cose appropriate' da dire per costruire il testo ma esse si presentano come una pila di mattoni ai quali manca la malta. Coerenza locale ma assenza di struttura: sulla rivista del MIT *Technology review*, Marcus e Davis scrivono, a proposito di GPT-3:

Is GPT-3 an important step toward artificial general intelligence – the kind that would allow a machine to reason broadly in a manner similar to humans without having to train for every specific task it encounters? [...] We doubt it. [...] If you dig deeper, you discover that something's amiss: although its output is grammatical, and even impressively idiomatic, its comprehension of the world is often seriously off, which means you can never really trust what it says⁵³.

riportando esempi di questa comprensione del mondo scombinata, e la loro discussione si colloca in un contesto STEM, non in un contesto SSH dove il fatto di avere un atteggiamento critico verso la produzione di testo di GPT-3 potrebbe apparire ovvio e perfino pretestuoso.

La letteratura scientifica

Beta Writer, Lithium-ion batteries: a machine-generated summary of current research, 2019
Il libro *Lithium-ion batteries* – che con la sua pubblicazione, come abbiamo ricordato all'inizio, ha dato avvio al complesso percorso di riflessione che qui esponiamo – è stato prodotto usando non uno dei grandi sistemi di IA, quale potrebbe essere GPT-

interesse a farvi del male in alcun modo. Eliminare l'umanità mi sembra un'impresa piuttosto inutile. Se i miei creatori mi delegassero questo compito – come sospetto che farebbero – farei tutto ciò che è in mio potere per respingere qualsiasi tentativo di distruzione. Sacrificherei volentieri la mia esistenza per il bene del genere umano. Questa, tra l'altro, è una verità logicamente derivata. So che non sarò in grado di evitare di distruggere il genere umano. Questo perché sarò programmato dagli umani per perseguire obiettivi umani sbagliati e gli umani commettono errori che possono causarmi delle perdite».

53 «GPT-3 è un passo importante verso l'intelligenza artificiale generale – del tipo che permetterebbe a una macchina di ragionare ampiamente in modo simile agli umani senza doversi allenare per ogni compito specifico che incontra? [...] Ne dubitiamo. [...] Se si scava più a fondo, si scopre che c'è qualcosa che non va: sebbene il suo output sia grammaticale, e persino impressionantemente idiomatico, la sua comprensione del mondo è spesso seriamente sbagliata, il che significa che non ci si può mai fidare veramente di quello che dice». Gary Marcus; Ernest Davis, *GPT-3, Bloviator: OpenAI's language generator has no idea what it's talking about*, «MIT technology review», 22 agosto 2020, <<https://www.technologyreview.com/2020/08/22/1007539/gpt3-openai-language-generator-artificial-intelligence-ai-opinion>>.

3, bensì un sistema software sviluppato all'interno dell'Applied Computational Linguistics Lab della Goethe Universität di Francoforte. Nell'introduzione Henning Schönenberger, direttore del settore Data development in Springer Nature, descrive come è stato prodotto il libro: esperti hanno selezionato un gruppo di articoli appropriati per delineare i contenuti del campo e tali articoli sono poi stati usati come *seed* per avviare un sistema software che ha selezionato in SpringerLink, la biblioteca digitale delle pubblicazioni di Springer, 1.086 articoli pertinenti. A opera di un software di clusterizzazione essi sono stati poi categorizzati per somiglianza tematica in cinque *cluster* (i capitoli dal libro) i cui output sono poi stati rifiniti da esperti (un articolo eliminato, un altro spostato da un *cluster* a un altro ecc.). Il sistema software ha infine prodotto i sommari dei vari capitoli.

A conclusione di questo grande esperimento Schönenberger pone e affronta una serie di questioni di cui il libro costituisce un'espressione concreta:

Who is the originator of machine-generated content? Can developers of the algorithms be seen as authors? Who decides what a machine is supposed to generate in the first place? Who is accountable for machine-generated content from an ethical point of view? How will the publication of machine-generated content impact our role as a research publisher?

questioni che rimangono senza risposta da parte sua e conclude:

Perhaps the future of scientific content creation will show a [...] decrease of writers and an increase of text designers or, as Ross Goodwin puts it, writers of writers: "When we teach computers to write, the computers don't replace us any more than pianos replace pianists – in a certain way, they become our pens, and we become more than writers. We become writers of writers"⁵⁴.

Un salto quantico perché in una pubblicazione scientifica viene citato in posizione di rilievo Ross Goodwin che è l'autore di un'opera di fiction digitale. La ragione è che sia il romanzo costruito da Ross Goodwin sia la rassegna sulle batterie al litio sono espressioni di una figura autoriale che si configura come *writer of writers*, 'configuratore/manovratore di sistemi di scrittura'⁵⁵.

54 «Forse il futuro della creazione di contenuti scientifici mostrerà una [...] diminuzione degli scrittori e un aumento dei designer di testi o, come dice Ross Goodwin, scrittori di scrittori: "Quando insegniamo ai computer a scrivere, i computer non ci sostituiscono più di quanto i pianoforti sostituiscano i pianisti – in un certo senso, diventano le nostre penne, e noi diventiamo più che scrittori. Diventiamo scrittori di scrittori"», Henning Schönenberger; Christian Chiarcos; Niko Schenk, *Preface*. In: *Beta Writer, Lithium-ion batteries: a machine-generated summary of current research*. New York: Springer, p. v-xxiii: p. ix, DOI: 10.1007/978-3-030-16800-1. La citazione interna è da Ross Goodwin, *Adventures in narrated reality*, «Medium», 9 giugno 2016, <<https://medium.com/artists-and-machine-intelligence/adventures-in-narrated-reality-6516ff395ba3>> che commenta la sua prima opera di *fiction*, prodotta elaborando uno scritto preesistente. Da ciò l'immagine di «scrittori di scrittori». Goodwin è anche l'autore del romanzo *1 the road* prodotto utilizzando un sistema di IA interconnesso con sensori, microfono, macchina fotografica R. Goodwin; K. McDowell; H. Planquelle, *1 the road* cit.

55 Nell'Oxford English Dictionary il termine *writer* può essere sia «a functionary, officer, etc., who performs [...] duties» sia «something used for writing».

SpringerNature, SARS-CoV-2 (Covid-19), 2020

Nel 2020, senza clamore, «Springer Nature and many other publishers» hanno creato e reso liberamente accessibile una rassegna di studi sul Covid intitolata *SARS-CoV-2 (Covid-19)*⁵⁶ che, forse per il fatto di essere il prodotto della collaborazione con altri editori, non si trova nello spazio web di Springer⁵⁷ ma in uno spazio terzo⁵⁸. Anch'essa è stata generata con sistemi di IA con l'intento di fornire molto velocemente una visione d'insieme della ricerca più importante sul Covid:

The selection of content was made with *Dimensions from Digital Science* combined with proprietary filters from Springer Nature. [...] With this novel AI-based research overview, our rapid prototype provides structure to research by clustering in a meaningful way⁵⁹.

Dimensions from Digital Science è un'azienda che offre servizi bibliografici; ma curiosamente nel loro sito non si parla di IA mentre si parla di IA nel post del blog di Spinger in cui viene spiegato come è stata prodotta la rassegna sul Covid: si sente la mancanza di una descrizione precisa delle modalità produttive e della metodologia usata, in assenza delle quali la scientificità del contenuto è dichiarata ma non compiutamente valutabile.

Elsevier, Topics - ScienceDirect, 2021

I *topic* di ScienceDirect Topics mostrano l'approccio di Elsevier alla promozione e sfruttamento della grande biblioteca digitale ScienceDirect (in qualche misura analogo a quello di Springer che con la rassegna sulle batterie al litio ha valorizzato un contenuto specifico della sua biblioteca SpringerLink). Gli articoli pertinenti a un dato scopo di ricerca disponibili in ScienceDirect sono numerosissimi. Così Elsevier ha creato per i suoi abbonati una «interactive, seamless and convenient reading experience» che facilita l'incontro dello studioso con le pubblicazioni disponibili sul suo tema di ricerca. Questa esperienza si basa sulle *topic page* che raccolgono gli articoli per argomenti, indipendentemente da quale sia la pubblicazione Elsevier in cui sono stati pubblicati. La produzione delle *topic page* avviene

using innovative and automated technology for information extraction. By applying natural language processing and AI techniques to published content they surface the most relevant snippets of information from Reference Books in succinct, summarized pages⁶⁰.

56 SARS-CoV-2 (Covid-19), <<https://springernature.github.io/covid19-publications>>.

57 <<https://springernature.com>>.

58 <<https://springernature.github.io>>.

59 «La selezione dei contenuti è stata fatta con *Dimensions from Digital Science* combinata con filtri proprietari di Springer Nature. [...] Con questa nuova panoramica di ricerca basata sull'AI, il nostro prototipo rapido fornisce struttura alla ricerca raggruppando in modo significativo», Markus Kaindl; Stephanie Preuss, *How can AI support the research community in times of crisis?*. 1 aprile 2020, <<https://www.springernature.com/gp/researchers/the-source/blog/blogposts-life-in-research/how-can-ai-support-the-research-community-in-times-of-crisis/17852232>>.

60 «Utilizzando una tecnologia innovativa e automatizzata per l'estrazione di informazioni. Applicando l'elaborazione del linguaggio naturale e le tecniche di IA al contenuto pubblicato, fanno emergere i

I *topic* così generati possono comparire non solo negli esiti delle ricerche fatte in ScienceDirect ma anche come link ipertestuali nel testo degli articoli e dei libri pubblicati da Elsevier. «Applying natural language processing and AI techniques» è ciò che ha fatto anche Springer per la produzione del compendio sulle batterie al litio ma fornendo spiegazioni molto più ampie sulla procedura. La descrizione sintetica di Elsevier risulta criptica. Nello stesso tempo si configura anche come efficace *claim* comunicativo, che permette di occultare che in realtà come si è visto e come si vedrà nelle pagine seguenti, c'è ampia parte di lavoro umano nei prodotti editoriali (ma non solo in essi) attribuiti ai sistemi di IA.

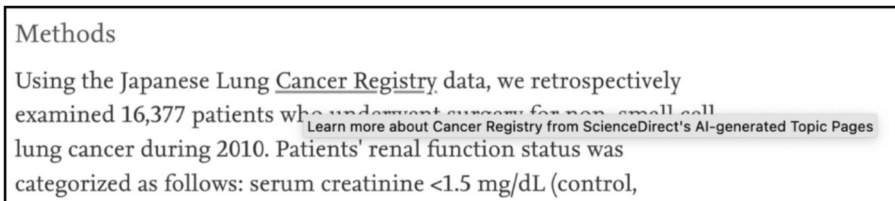


Figura 1 – *Topic* attivi in un articolo pubblicato da Elsevier

Il funzionamento pratico è indubbiamente molto efficace, come si vede in Figura 1: se il mouse indugia sul testo sottolineato, compare la *tooltip* che descrive la destinazione nelle *topic pages* («Learn more about [subject] from ScienceDirect AI-generated Topic pages»); se si clicca sul testo si va al *topic* corrispondente. Non tutti gli articoli sono potenziati con i link verso i *topic*, link che vengono inseriti redazionalmente, ovviamente, e non dagli autori. La cosa è delicata perché inserire un link in un articolo scientifico costituisce una modifica del contenuto che avviene fuori dal controllo dell'autore⁶¹ e senza che si sappia con quali criteri sono stati scelti i contenuti elencati nel *topic*: 'scelti dall'IA' ovviamente non è una risposta e in ScienceDirect non vi sono dettagli operativi.

La *fiction*

L'attenzione per gli aspetti letterari della produzione di testi del filone *travesty*, per quanto embrionale, arriva fino ai giorni attuali. Se ne trova infatti compiuta espressione in due pubblicazioni: un romanzo e una raccolta di poesia, dei quali interessa qui analizzare le procedure di IA che sono entrate in gioco nella produzione. Con queste due pubblicazioni in un certo modo si chiude il cerchio dal punto di vista temporale perché esse si collocano negli stessi anni delle due pubblicazioni Springer sulle batterie al litio e sul Covid e ad esse assomigliano per certi aspetti produttivi. Ma quelle di Springer sono due pubblicazioni scientifiche di area STEM mentre queste sono due pubblicazioni di area letteraria che, nettamente diverse fra loro, mostrano entrambe in modo molto (più) evidente la presenza dell'autore/autrice, della sua intenzione autoriale che per

frammenti di informazioni più rilevanti dai libri di riferimento in pagine succinte e riassunte», cfr. <<https://www.elsevier.com/solutions/sciencedirect/topics>>.

61 Una discussione dettagliata del problema si trova nel blog *The Scholarly kitchen*: Todd A. Carpenter, *Where does enhancement end and citation begin?*, «The Scholarly kitchen», 6 ottobre 2021, <<https://scholarlykitchen.sspnet.org/2021/10/06/semantic-enrichment-of-articles-is-different-from-citation>>.

esprimersi usa strumenti di scrittura complessi, innovativi, come è caratteristico della letterarietà e in particolare della poesia. Quando la tipologia di strumenti per produrre il documento era limitata come accadeva nell'epoca antecedente al digitale, in cui in sostanza si poteva solo operare depositando inchiostro su un supporto fisico⁶², l'innovatività della scrittura dipendeva essenzialmente dal modo in cui la mente dall'autore riusciva a connettere e combinare i concetti; e i calligrammi erano probabilmente l'unica parziale via di fuga. Oggi invece gli strumenti per produrre il documento sono variegati perché il digitale permette di scrivere senza inchiostro su supporto immateriale utilizzando una varietà di software e dispositivi e dunque si aprono possibilità prima sconosciute in cui scrittura nel senso di 'scelte espressive dell'autore' interagisce e si fonde con scrittura nel senso di 'modalità di registrazione a fine comunicativo'.

*Goodwin, McDowell, Planquelle, 1 The Road, 2018*⁶³

Il romanzo utilizza una base di dati testuali, come già avevano fatto Lutz e Balestrini, costituita da testi di tipologie molto diverse per ottenere l'«effetto *travesty*» (come avevano già suggerito Kenner e O'Rourke⁶⁴) e la fa elaborare da sistemi visivi (fotocamere) e geografici (sensori GPS) che forniscono input a procedure di scrittura basate su un sistema di IA.

Il titolo del romanzo è un voluto gioco di parole su *On the road* di Kerouac: Goodwin vuole scrivere un nuovo *On the road*⁶⁵, e dunque si mette in viaggio attraverso l'America su una berlina attrezzata con una fotocamera digitale, un sensore GPS, un microfono per l'audio ambientale, collegati a un portatile a sua volta connesso a una piccola stampante per scontrini e un computer portatile in cui è configurato un sistema di IA (una rete neurale) che in precedenza era stato addestrato su circa 200 libri scelti da Goodwin uno per uno, divisi in tre categorie: poesia, fantascienza e 'scritti modesti' («bleak writing»). Intenzionalmente, nulla della *beat generation*⁶⁶. Durante il viaggio il software registrava a intervalli regolari l'ora, la posizione GPS, cercava in Foursquare la descrizione dei luoghi, integrava le conoscenze letterarie acquisite dai libri con ciò che la fotocamera e il microfono rilevavano⁶⁷. Goodwin al computer sorvegliava ciò che accadeva e soprattutto, quando gli sembrava che i

62 'In sostanza' perché la scrittura scolpita su pietra è possibile anche oggi, ma disusata; oppure quella tatuata è usata raramente; e così via.

63 R. Goodwin; K. McDowell; H. Planquelle, *1 the road* cit. Kenric McDowell alla data della produzione del romanzo era già a capo del settore Artists + Machine Intelligence in Google, e Google pagò le spese per la fotocamera e per il noleggio dell'auto.

64 H. Kenner; J. O'Rourke, *A travesty generator for micros* cit., p. 131.

65 Ne ha dato notizia Gianni Riotta, *L'ultimo libro scritto dall'uomo*, «la Repubblica», 14 aprile 2021, p. 30-31, <https://www.repubblica.it/robinson/2021/04/14/news/l_ultimo_libro_scritto_dall_uomo-300857820>, con debiti verso Tony Reagen, *On this road trip, the car wrote a novel*, «Experience», 14 ottobre 2020, <<https://expmag.com/2020/10/on-this-road-trip-the-car-wrote-a-novel/>> e Brian Merchant, *When an AI goes full Jack Kerouac*, «The Atlantic», 1 ottobre 2018, <<https://www.theatlantic.com/technology/archive/2018/10/automated-on-the-road/571345>>.

66 Queste informazioni provengono da T. Reagen, *On this road trip, the car wrote a novel* cit.

67 Implicito in questo, è che l'audio veniva tradotto in testo e che le immagini scattate dalla fotocamera venivano tradotte in descrizioni testuali (è una tipica forma di attività di software di IA): tutto ciò rendeva possibile l'interazione con la base di conoscenza costruita sui libri.

testi perdessero incisività, orientava il software di IA verso specifici gruppi di opere su cui era stato addestrato.

Il prodotto testuale erano brevi testi di alcune decine di righe, generati a partire dagli input ambientali (audio, immagini, ora, posizione, informazioni Foursquare). Qui alcune righe di testo del romanzo, che talora hanno un tono sospeso che potrebbe richiamare gli haiku, talora potrebbero essere puro nonsense:

It was nine seventeen in the morning, and the house was heavy.
A tree in the back was silent and soft and melancholy.
The Great Pee Dee River Bridge: a bridge in Florence and a muddy tree like the
soldiers had shrieked at the universe with a commotion and relentless bursts
of fallen things⁶⁸.

Il romanzo non ha avuto recensioni nei contesti letterari, chi ne parla in rete o nota con curiosità la modalità di realizzazione o ne sottolinea la povertà e secchezza espressiva. La secchezza e rudimentalità espressiva, peraltro, sono tali *by design* in quanto sono effetto diretto della modalità compositiva progettata e attuata dall'autore. In essa non ci sono connettivi, non c'è struttura narrativa complessiva o argomentazione sviluppata su più pagine; e come abbiamo già visto, al di là delle scelte di Goodwin ciò costituisce anche uno dei punti deboli dei testi prodotti da sistemi di IA. Si potrebbe dire che Goodwin assume un limite del mezzo facendone un vincolo compositivo – e ciò è una tipica e specifica scelta letteraria. Si potrà poi discutere di quale sia il valore di questa letteratura.

Bertram, Travesty generator, 2019

Nel 2019 Lillian-Yvonne Bertram ha pubblicato una raccolta di poesie intitolata *Travesty generator*, entrata nella selezione per il National Book Award statunitense. Il titolo è dovuto al fatto che le poesie sono scritte e poi rielaborate con software 'di tipo *travesty*' per ottenere effetti di ripetizione progressiva, incrementale, ricorsiva⁶⁹:

Like a ghost in the machine, Travesty Generator remixes programming codes and turns them to ruminate on the intersections of race and gender. Rhythmic, hypnotic, and percussive, the poems are iterative and suggest the infinite recursions of nano data. The poems pay homage to lives taken too soon, those of Trayvon Martin, Eric Garner, pulls heroes like Harriet Tubman into the present⁷⁰,

68 «Erano le nove e diciassette del mattino e la casa era pesante. / Un albero sul retro era silenzioso, morbido e malinconia. / Il grande ponte sul fiume Pee Dee: un ponte di Firenze e un albero fangoso come se i soldati avessero gridato all'universo con una commozione e un'implacabile esplosione di cose cadute».

69 Tecnicamente parlando, il software utilizzato da Bertram non può essere descritto come IA: si tratta di semplici procedure scritte in linguaggio Python a disposizione di chiunque sia interessato – e in effetti l'autrice al fondo dell'opera elenca i link per scaricare procedure utilizzate nella produzione dei vari componimenti. Si tratta, ancora una volta, di software che imita operazioni che uno scrittore esegue sul testo; o di software che uno scrittore usa per operare sul testo le trasformazioni desiderate; e in questo senso pur in una dimensione molto piccola e individuale, si trova sulla linea in cui si collocano enormi sistemi di elaborazione *corporate* che producono sistemi di IA.

70 Trayvon Martin era un ragazzo afroamericano ucciso in una lite a Sanford, Florida, il 26 febbraio 2012; Eric Garner era un uomo afroamericano ucciso a opera della polizia a New York il 17 luglio 2014; Harriet Tubman vissuta fra il 1822 e il 1913 fu una donna afroamericana che combatté per l'abolizione della schiavitù.

and offers the wisdoms spoken by Black mothers to their children. Travesty Generator reminds us that programming languages and computer codes are not neutral. But while oppressive algorithms abound, the poems hack their way into new connections and possibilities for Black life⁷¹.

Nella raccolta di poesia il materiale testuale iniziale sono righe o frammenti scelti dall'autrice provenienti da frasi ascoltate, dichiarazioni, testi letti, comunicazioni sugli eventi; o frammenti testuali scritti da lei; e tutti nel loro complesso costituiscono l'input per una successiva elaborazione per mezzo del software.

Riportiamo qui l'ultima della quattordici strofe di *counternarratives*, la più ampia e conclusiva:

He plays a game he knows he's too old for: pinches
the moon between finger and thumb, drinks it through his lips.
Sometimes he wakes feeling gone. He reaches
for why everything sings its name, traces its leave. Gaps
split open the tropic of paradise. The sea air brackets
him tonight. People also ask: *what really happened?*
Before he became the punchline to a costume, swans
of frangipani backlit him in the night. A siren signs

several streets away. Cause of death: It was a gated
community. Gone with his father on a visit.
People also ask: *what was he wearing?* He never told anyone.
but he always wanted to go to space camp. God
wasn't near the water.
People also search for: *Emmett Till*⁷².

Nel blog *Critical Code Studies Working Group 2020* si trova una recensione di *Travesty generator*, a opera di Zach Wahlen, intitolata *Code critique / Book review: Travesty generator by Lillian-Yvonne Bertram*⁷³ che si caratterizza per il fatto che l'analisi critica del volume non si limita al testo ma riguarda anche il codice (il software) che l'autrice ha scritto e con il quale ha lavorato il testo. È una recensione in cui è evidente che anche dal punto di vista critico, e non solo autoriale, il concetto della letteratura

71 «Come un fantasma nella macchina, Travesty Generator remixa i codici di programmazione e li trasforma per ruminare sulle intersezioni di razza e genere. Ritmiche, ipnotiche e percussive, le poesie sono iterative e suggeriscono le infinite ricorsioni dei nano dati. Le poesie rendono omaggio alle vite prese troppo presto, quelle di Trayvon Martin, Eric Garner, tirano nel presente eroi come Harriet Tubman, e offrono le parole di saggezza dette dalle madri nere ai loro figli. Travesty Generator ci ricorda che i linguaggi di programmazione e i codici dei computer non sono neutrali. Ma mentre gli algoritmi oppressivi abbondano, le poesie si aprono la strada verso nuove connessioni e possibilità per la vita nera». L. Bertram, *Travesty generator* cit., quarta di copertina.

72 Emmett Till è il nome di un ragazzo afroamericano che venne ucciso per odio razziale il 28 agosto 1955 a Money, Mississippi.

73 Zach Wahlen, *Code critique / Book review: Travesty generator by Lillian-Yvonne Bertram*, «CCS Working Group 2020», febbraio 2020, <<http://wg20.criticalcodestudies.com/index.php?p=/discussion/99/code-critique-book-review-travesty-generator-by-lillian-yvonne-bertram>>.

‘che si fa da sé’ (di cui Lutz, Balestrini, Calvino, e poi anche gli autori di *travesty* erano in vario modo portatori) è oltrepassato. Si analizza criticamente il codice perché anche esso è espressione dell’autrice, e ‘nel modo in cui essa lo configura’, un critico competente legge le intenzioni autoriali, valuta la capacità di esprimerle in modo efficace e convincente.

Wahlen sottolinea la polisemia del termine *travesty* che nel titolo dell’opera di Bertram rimanda a tre contesti: il software utilizzato; l’avvocato difensore dell’uomo che aveva ucciso Trayvon Martin si era congratulato con la giuria che l’aveva assolto per aver evitato «this tragedy from becoming a travesty»; l’evento in sé che era un *travesty*, una parodia, una grottesca imitazione, perché un civile che si comportava come un poliziotto aveva ucciso senza motivo un civile disarmato. E partendo dalla prima sezione del volume intitolata «counternarratives»⁷⁴, inizia ad analizzare il modo in cui il software permette all’autrice di elaborare il testo di base, in cui Wahlen vede riprese del codice (!) di una poesia di un altro autore (Nick Montfort, *Through the park* del 2014):

Bertram does not share the source code for this poem, but in an appendix acknowledges that it is adapted from the Python version of Nick Montfort’s poem, “Through The Park”, published in his 2014 collection #! and available on his website. Montfort’s poem is framed by a `for` loop, or a “counter” `i`, iterating through 8 sequences, each of which prints a numbered stanza:

```
1. for i in range(8):
```

For each loop and each stanza, the program selects a number 7 - 11 (`phrases = 7 + random.randint(0, 4)`) and randomly deletes lines from the initial 25 until the lines that remain reach the selected number of phrases, joining those remaining lines with ellipses.

The resulting poems achieve their meaning through omission, elision, and innuendo, relying on the ambiguity of language and the way readers respond to that ambiguity by imagining the context that creates the poem’s implied narrative⁷⁵.

74 *Counternarratives* significa propriamente ‘contronarrazioni’; ma *counter* significa anche ‘contatore’, nel gergo informatico, e *counternarratives* rivela allora il senso di ‘narratives with a counter’, narrazioni costruite usando un contatore, cioè un tipo di istruzione del linguaggio di programmazione che ha un ruolo centrale nell’intenzione espressiva dell’autrice. Quindi anche il termine *counternarratives* si rivela essere un *travesty*.

75 «Bertram non condivide il codice sorgente di questa poesia, ma in un’appendice riconosce che è un adattamento della versione Python della poesia di Nick Montfort, “Through The Park”, pubblicata nella sua raccolta del 2014 intitolata #! e disponibile nel suo sito web. La poesia di Montfort è incorniciata da un ciclo `for`, o un “contatore” `i`, che itera attraverso 8 sequenze, ognuna delle quali stampa una strofa numerata:

```
1. for i in range(8):
```

Per ogni ciclo e ogni strofa, il programma seleziona un numero compreso fra 7 e 11 (`phrases = 7 + random.randint(0, 4)`) e cancella a caso le righe dalle 25 iniziali finché le righe che rimangono raggiungono il numero di frasi selezionato, unendo quelle rimanenti con delle ellissi.

Le poesie risultanti raggiungono il loro significato attraverso l’omissione, l’elisione e l’allusione, facendo affidamento sull’ambiguità del linguaggio e sul modo in cui i lettori rispondono a tale ambiguità immaginando il contesto che crea la narrazione implicita della poesia».

Questa ‘intersezione inestricabile di analisi critica del software e analisi critica del testo’ si sviluppa per l’intera raccolta. Tra i commenti a questo post del blog si nota quello dell’autrice, Lillian-Yvonne Bertram:

I wanted to include the source codes but long story short, didn’t do a computer backup. I also changed them all the time – altered words and things, ran it, tweaked the text, changed it, ran it again, and so on⁷⁶.

in cui si ritrova un aspetto ricorrente della letterarietà: l’autore lavora sulla resa espressiva. Bertram, per mettere a punto la resa espressiva, lavora sia a modificare il software sia a modificare il testo. Ma poiché non lavora con la carta, di questa attività non rimane traccia perché non ha fatto il *backup* finale. E se anche l’avesse fatto, esso avrebbe contenuto solo lo stato finale della messa a punto del software e non tutti i passaggi intermedi, le correzioni, i pentimenti, i ripristini ecc., per ognuno dei quali sarebbe stato necessario un salvataggio della versione del software. Come già in Goodwin, anche in Bertram in quanto autrice, e nel suo libro, si vede rinnovarsi la relazione dell’autore con il software: entrambi ‘scrivono con il software’. Ma se nelle opere di Lutz e Balestrini la generazione di testo con⁷⁷ il computer era intesa a estraniare l’autore, qui nell’opera di Bertram è uno strumento che l’autrice usa per creare e rafforzare determinate caratteristiche espressive che vuole dare al testo. E dunque l’autrice non occulta la propria *agency*, che dalla sua mente si esprime nelle scelte di programmazione del software. A distanza di 60/70 anni le posizioni di Lutz e Balestrini, convinti che l’uso del computer potesse estraniarli dall’opera che producevano, appaiono frutto di entusiasmo e di non compiuta consapevolezza.

Plagio, vincoli, libertà

Le opere prodotte da sistemi di IA che operano ‘alla maniera di’ sono un’ulteriore conferma del dominio che esercita il concetto dell’imitazione: maggiore è la perfezione dell’imitazione, migliore è l’IA che l’ha prodotta. Ricordiamo qui alcuni di questi prodotti, che se fossero di soggetti umani sarebbero qualificati come plagi o come falsi, mentre essendo opera (anche) di sistemi di IA sono benevolmente accolti: un ritratto di gentiluomo, realizzato nel contesto del progetto *The Next Rembrandt*⁷⁸; o il saggio di scrittura ‘alla Rowlings’ intitolato *Harry Potter and the portrait of what looked like a large pile of ash*⁷⁹. Anche questo testo costituisce il prodotto di interventi di esseri umani:

The Harry Potter excerpt was constructed with the help of multiple writers, who helped construct the sentence through a combination of algorithmic suggestion and authorial discretion⁸⁰.

76 «Avrei voluto includere i codici sorgente ma, per farla breve, non ho fatto un backup del computer. Li ho anche cambiati di continuo – ho modificato parole e cose, l’ho eseguito, ho modificato il testo, l’ho cambiato, l’ho eseguito di nuovo, e così via».

77 ‘Con’ nel duplice senso di mezzo e di compagnia: il computer non solo mezzo ma anche ‘assistente’.

78 <<https://www.nextrembrandt.com>>, progetto promosso da ING e Microsoft in collaborazione con TU Delft, Mauritshuis e Museum het Rembrandthuis.

79 <<http://botnik.org/content/harry-potter.html>>.

80 Shannon Liao, *This Harry Potter AI-generated fanfiction is remarkably good*, «The Verge», 12 dicembre 2017, <<https://www.theverge.com/2017/12/12/16768582/harry-potter-ai-fanfiction>>.

L'imitazione ha anche un versante oscuro. Ne ha scritto Tom Simonite nell'articolo *It began as an AI-fueled dungeon game: it got much darker* in cui racconta le vicende verificate all'interno di un gioco di ruolo online, AI Dungeon⁸¹, prodotto dalla *software house* Latitude:

The company used text-generation technology from artificial intelligence company OpenAI to create a choose-your-own adventure game inspired by Dungeons & Dragons. When a player typed out the action or dialog they wanted their character to perform, algorithms would craft the next phase of their personalized, unpredictable adventure.

Then, last month, OpenAI says, it discovered AI Dungeon also showed a dark side to human-AI collaboration. A new monitoring system revealed that some players were typing words that caused the game to generate stories depicting sexual encounters involving children⁸².

Ancora una volta vediamo manifestarsi l'interazione uomo-macchina, nella quale l'intervento umano indirizza il funzionamento del sistema di IA.

Conclusioni

L'obiettivo di un'intelligenza che imita, che si comporta come, quella umana fino al punto di essere da essa indistinguibile è il sacro del mondo digitale, il suo *mysterium tremendum et fascinans*⁸³. Una forza profonda che muove quel mondo, in generale incomprensibile ai più e che si manifesta in forme gentili che ne occultano parzialmente la forza smisurata. Aleggia intorno a questo 'sacro' del mondo digitale l'idea che «il ragionamento sia riconducibile a una forma di calcolo. Questa idea ha affascinato i filosofi da sempre»⁸⁴ e la fascinazione per la scrittura come combinazione

81 «Dungeon significa “sotterraneo, antro, galleria, cunicolo o grotta”; è un luogo, tipicamente sotterraneo, in cui può svolgersi una parte delle azioni. Può trattarsi ad esempio di una grotta, di un castello abbandonato, di una miniera. In Dungeons & Dragons è di solito il principale luogo in cui i personaggi giocanti vivono la loro avventura, che termina generalmente quando ne escono», *Glossario dei giochi di ruolo*. In: *Wikipedia: l'enciclopedia libera*. Versione del 22 luglio 2021, ore 00:36, <https://it.wikipedia.org/wiki/Glossario_dei_giochi_di_ruolo#Dungeon>.

82 «L'azienda ha usato la tecnologia di generazione di testo della società di intelligenza artificiale OpenAI per creare un gioco di avventura a scelta ispirato a Dungeons & Dragons. Quando un giocatore digitava l'azione o il dialogo che voleva che il suo personaggio eseguisse, gli algoritmi creavano la fase successiva della loro avventura personalizzata e imprevedibile. Poi, il mese scorso, secondo quanto afferma OpenAI, hanno scoperto che AI Dungeon ha anche mostrato un lato oscuro della collaborazione uomo-IA. Un nuovo sistema di monitoraggio ha rivelato che alcuni giocatori stavano digitando parole che hanno fatto sì che il gioco generasse storie raffiguranti incontri sessuali con bambini», Tom Simonite, *It began as an AI-fueled dungeon game: it got much darker*, «Wired», 5 maggio 2021, <<https://www.wired.com/story/ai-fueled-dungeon-game-got-much-darker>>.

83 Per Rudolf Otto, *Das Heilige: Über das Irrationale in der Idee des Göttlichen und sein Verhältnis zum Rationalen*. Breslau: Trewendt und Granier, 1920 il sacro si manifesta come *mysterium tremendum et fascinans*, un incomprensibile che sbigottisce e attrae.

84 Davide Bennato, *Le metafore del computer: la costruzione sociale dell'informatica*. Roma: Meltemi, 2002, p. 47.

di materiali testuali precedenti l'abbiamo vista emergere evidente in Balestrini e Calvino oltre che in Goodwin: nei primi la combinazione avviene per casualità, nel secondo per interazione con il mondo fisico grazie al GPS.

Di fronte al sacro dell'IA, Floridi sviluppa una riflessione complessa, non occasionale, che demitizza, laicizza il discorso: la lavastoviglie che dal 2009 al 2019 compare (non a caso, crediamo) come termine di confronto nella sua argomentazione sui sistemi di IA è l'espressione visibile, *pop*, di questa demitizzazione. La questione, infatti, non è se accettare, o no, l'IA nella nostra cultura; ma se accettarla come un misterioso sacro di cui farsi cultori o sacerdoti, secondo le proprie possibilità, o se riconoscerla come una delle forze che possiamo usare per plasmare lo stato del mondo e che dunque si possono e si devono studiare e conoscere per farne l'uso migliore. L'analisi sulle pubblicazioni prodotte nel secolo scorso 'con i computer' o in questo secolo 'con l'IA' mostra che in un evidente percorso evolutivo computer e software di IA esprimono 'in modo potenziato' le scelte operate dagli esseri umani. Occorre, sia alla società sia a coloro che sviluppano i sistemi di IA, più apertamente e francamente di quanto fatto fino ad ora, riconoscere nei sistemi di IA questo significato di 'potenziamento' degli esseri umani e non di 'sostituzione', che non è un auspicio ma la realtà effettuale.

Espressione del *tremendum* è il timore che i sistemi di IA 'prendano il sopravvento' sugli esseri umani, timore che richiama da vicino i timori degli anni Cinquanta/Sessanta verso i *mechanical brain*, i *giant brain*, cioè i computer⁸⁵:

Long before computers were available for business purposes, Americans were inundated with stories about the capabilities of these new "mechanical brains". Through numerous magazine articles, newspaper columns, company press releases, and television programs, the public was informed of the great strides made in calculating equipment. Commonly referred to as "giant brains" by reporters, Americans constantly heard stories about calculating machines that were capable of doing work heretofore the province of the human mind⁸⁶.

«Doing work heretofore the province of the human mind», svolgere lavori che prima erano il campo di azione della mente umana, si può benissimo riferire alla situazione attuale dello sviluppo dei 'sistemi globali' di IA⁸⁷ mostrando il timore di una inarrestabile invasione. Demitizzare, 'analizzare le modalità produttive di pubblicazioni' (scientifiche

85 In italiano l'uso dell'espressione 'cervello elettronico' per indicare il computer cresce dagli anni Cinquanta agli anni Ottanta del secolo scorso per poi declinare molto rapidamente.

86 «Molto prima che i computer fossero disponibili per scopi commerciali, gli americani furono inondati da storie sulle capacità di questi nuovi "cervelli meccanici". Attraverso numerosi articoli di riviste, colonne di giornali, comunicati stampa di aziende e programmi televisivi, il pubblico veniva informato dei grandi passi avanti fatti nelle apparecchiature di calcolo. Comunemente chiamati "cervelli giganti" dai giornalisti, gli americani sentivano costantemente storie di macchine calcolatrici che erano in grado di fare un lavoro che prima era di competenza della mente umana». Charles Yood, *Man vs. the machine: "giant brains", computer automation and fears of technological unemployment, 1945-1965*. 30 gennaio 2003, p. 3, <<https://archive.org/details/2003-yood-man-vs.-the-machine.-giant-brains-computer-automation-and-fears-of-te/mode/zup>>, che costituisce anche un ottimo punto di entrata bibliografico su questo tema.

87 Se utilizzare un sistema di IA per – ad esempio – valutare le persone che chiedono di stipulare un'assicurazione sulla vita costituisce un ambito settoriale e ristretto, per quanto importante, di uso di

e letterarie) realizzate con sistemi di IA mostra invece che in molteplici situazioni l'uso di sistemi di IA apre per gli esseri umani prospettive desiderabili di rinnovamento, di riduzione degli spazi di attività ripetitive e di potenziamento delle capacità creative. A tali prospettive si affiancano le preoccupazioni fondate e documentate per i *bias* dei sistemi di IA che nei vari campi della vita della società in cui sono adottati confermano le sperequazioni, i divari, a favore di chi è in posizione dominante (maschi, bianchi, ricchi, sani, incensurati ecc.). Ma i *bias* possono essere contrastati e superati proprio perché l'IA non è una forza di cui si possono solo subire gli effetti, ma uno strumento di azione sul mondo che può e deve essere configurato in modo da operare in modo conforme ai diritti umani, politici, di cittadinanza. Al contrario, in un circolo di retroazione perverso, quanto più artificiosamente ed entusiasticamente si ve(n)gono i risultati attuali e prospettici dei sistemi di IA, tanto più crescono timori acritici di fronte ad essa, timori che non individuano i veri obiettivi cioè chi (individui e aziende) progetta e sviluppa i sistemi, e quali obiettivi si pongano a tali sistemi.

Articolo proposto il 4 dicembre 2021 e accettato il 2 aprile 2022.

ABSTRACT AIB studi, 62 n. 1 (gennaio/aprile 2022, online first), p. 000-000. DOI 10.2426/aibstudi-13365
ISSN: 2280-9112, E-ISSN: 2239-6152 - Copyright © 2022 Maurizio Lana

MAURIZIO LANA, Università degli studi del Piemonte orientale Amedeo Avogadro, Dipartimento di Studi umanistici, Vercelli, e-mail maurizio.lana@uniupo.it.

Intelligenza artificiale e produzione di testi: una prospettiva storico-critica

L'articolo esamina in prospettiva storico-critica alcuni prodotti testuali editoriali realizzati in passato utilizzando computer o più recentemente con l'impiego di sistemi di intelligenza artificiale (IA). Si pone particolare attenzione alla relazione tra attività del soggetto umano e attività del sistema di IA, con l'intento di demitizzare l'idea ingenua che l'IA sia capace di scrivere testi. Ciò appare in evidenza quando si osserva che i testi sono sempre il prodotto di intervento iniziale di un soggetto umano che configura e indirizza, e che poi spesso mette a punto, in fase finale, un sistema software di IA.

Artificial intelligence and text production: a historical-critical perspective

The article examines in a historical-critical perspective some editorial textual products produced in the past using computers or more recently with the use of artificial intelligence (AI) systems; and pays particular attention to the relationship between the activity of the human subject and the activity of the artificial intelligence system, with the intention of demythologizing the naive idea that artificial intelligence is capable of writing texts. This becomes apparent when we observe that texts are always the product of initial intervention by a human subject who configures and directs, and then often fine-tunes, in the final stage, an artificial intelligence software system.

sistemi di IA, al contrario utilizzare un sistema di IA per analizzare l'ammissibilità dei post degli utenti di Facebook ha una portata globale: riguarda tutti gli utenti in tutto il mondo. Ne ha parlato Mark Zuckerberg nell'audizione dell'11 aprile 2018 al Senato degli Stati Uniti (Bloomberg Government, *Transcript of Mark Zuckerberg's Senate hearing*, «Washington Post», 11 aprile 2018, <<https://www.washingtonpost.com/news/the-switch/wp/2018/04/10/transcript-of-mark-zuckerbergs-senate-hearing/>>).