

Artificial Intelligence Systems and problems of the concept of author. Reflections on a recent book

Maurizio Lana^(a)

a) [University of Piemonte Orientale](#), <https://orcid.org/0000-0002-7520-1195>

Contact: Maurizio Lana, maurizio.lana@uniupo.it

Received: 2 June 2022; Accepted: 15 November 2022; First Published: 15 May 2022

ABSTRACT

The publication of the book Beta Writer. 2019. Lithium-Ion Batteries. A Machine-Generated Summary of Current Research. New York, NY: Springer, produced with Artificial Intelligence software prompts analysis and reflections in several areas. First of all, on what Artificial Intelligence systems are able to do in the production of informative texts. This raises the question if and how an Artificial Intelligence software system can be treated as the author of a text it has produced. Evaluating whether this is correct and possible leads to re-examine the current conception for which it is taken for granted that the author is a person. This, in turn, when faced with texts produced by Artificial Intelligence systems necessarily raises the question of whether they, like the author-person, are endowed with agency. The article concludes that Artificial Intelligence systems are characterized by a distributed agency, shared with those who designed them and make them work, and that in the wake of the reflections of 50 years ago by Barthes and Foucault, it is necessary to define and recognize a new type of author.

KEYWORDS

Artificial Intelligence; Texts; Books; Agency; Authorship.

Sistemi di Intelligenza Artificiale e problemi del concetto di autore. Riflessioni su prodotti editoriali recenti

ABSTRACT

La pubblicazione del libro Beta Writer. 2019. Lithium-Ion Batteries. A Machine-Generated Summary of Current Research. New York, NY: Springer, prodotto con software di Intelligenza Artificiale sollecita analisi e riflessioni in vari ambiti. Prima di tutto su che cosa i sistemi di Intelligenza Artificiale siano in grado di fare nell'ambito della produzione di testi di tipo informativo. Di qui si pone la questione se e come un sistema software di Intelligenza Artificiale possa essere trattato come autore di un testo che ha prodotto. Valutare se ciò sia corretto e possibile porta a riprendere in esame la concezione attuale per la quale è scontato che l'autore sia una persona. Ciò, a sua volta, di fronte a testi prodotti da sistemi di Intelligenza Artificiale sollecita necessariamente la domanda se essi, come l'autore-persona, siano dotati di agency. L'articolo conclude che i sistemi di Intelligenza Artificiale si caratterizzano per un'agency distribuita, condivisa con chi li ha progettati e li fa funzionare e che sulla scorta delle riflessioni di 50 anni fa di Barthes e Foucault occorre definire e riconoscere un nuovo tipo di autore.

PAROLE CHIAVE

Intelligenza Artificiale; Testi; Libri; Agency; Authorship.

Premessa

Sempre più spesso anche nei mezzi di comunicazione non specializzati compaiono notizie e riflessioni sull'uso dell'Intelligenza Artificiale nella quotidianità della nostra società. I toni oscillano tra entusiasmo e timori distopici. Ma il fatto stesso di ipostatizzare «l'Intelligenza Artificiale» segnala una visione imprecisa, perché «l'Intelligenza Artificiale» è una disciplina in cui si opera con programmi+dati, programmi che progettisti, programmatori (e altri software) addestrano su vari tipi di dati. Tanto i programmi quanto i dati riflettono, ed amplificano, la visione del mondo dei progettisti e programmatori e i loro bias. Qui useremo invece espressioni come «software | sistemi | programmi di Intelligenza Artificiale» perché non scriviamo di Intelligenza Artificiale ma di situazioni in cui i prodotti dell'Intelligenza Artificiale vengono utilizzati.

Questo articolo interseca o sfiora grandi questioni; qui ne menzioniamo alcune senza pretesa di esaurirle. Alcune riguardano le caratteristiche proprie dei sistemi di Intelligenza Artificiale, in primis quella dell'agency di cui molto si discute: «i sistemi di Intelligenza Artificiale hanno agency?» hanno cioè capacità di agire individualmente in modo libero in un determinato contesto? La seconda è quella del rapporto con il mondo: «i sistemi di Intelligenza Artificiale comprendono il mondo?» cioè sono capaci di operare in modo appropriato sia dal punto di vista sintattico sia dal punto di vista semantico¹? Altre questioni riguardano il modo in cui i sistemi di Intelligenza Artificiale si collocano nel mondo. Una questione è quella storica: «qual è la storia dei sistemi di Intelligenza Artificiale?» cioè in quale linea di eventi si iscrivono i sistemi di Intelligenza Artificiale? Eventi sono snodi di possibilità, che segnalano che «le cose avrebbero potuto andare diversamente». Un'altra è quella ambientale: «i sistemi di Intelligenza Artificiale sono sostenibili?». Addestrare 1 singolo modello su 1 singolo problema testuale con 1 rete neurale comporta consumi elettrici che producono circa 284 chili di CO₂². In questo articolo affronteremo la prima questione, quella dell'agency e ciò che essa implica.

¹ La riflessione su sintattico/semantico è al centro di (Durante 2019). Durante porta ad esempio una partita di scacchi contro il computer: per il giocatore umano la partita è semantica cioè la scelta delle mosse si iscrive all'interno di una visione strategica complessiva, per il programma di Intelligenza Artificiale è sintattica perché in risposta alla mossa del giocatore umano si calcolano tutte le possibili mosse lecite successive. Nella quotidianità, chi fa una ricerca web ha uno scopo semantico cioè le parole della query implicano e veicolano un significato atteso, mentre il motore di ricerca fornisce gli esiti su base sintattica cioè essi contengono in modo appropriato le parole della query – la discrepanza sintattica/semantica si vede nel fatto che davanti a migliaia di esiti sintatticamente (co)rispondenti noi tutti andiamo in cerca di quei «7» semanticamente validi e soddisfacenti. Quindi il soggetto umano e il computer costantemente al medesimo tavolo giocano due giochi differenti. Parte della complessità del confronto della società con i sistemi di Intelligenza Artificiale è invece credere che il gioco sia il medesimo. L'equivoco si manifesta per la prima volta in modo esplicito in (McCarthy et al. 1955): «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» ma è già presente nel test di Turing che è presentato come un *imitation game* (Turing 1950). L'importanza del «Proposal» del 1955 è data dal fatto che si tratta della *prima formulazione di un progetto complessivo per l'Intelligenza Artificiale*, e gli autori (McCarthy, Minsky, Rochester, Shannon) sono quattro giganti, per usare la metafora di Bernardo di Chartres.

² I dati qui appena accennati provengono da (Strubell, Ganesh, e McCallum 2019) *these accuracy improvements depend on the availability of exceptionally large computational resources that necessitate similarly substantial energy consumption. As a result these models are costly to train and develop, both financially, due to the cost of hardware and electricity or cloud compute time, and environmentally, due to the carbon footprint required to fuel modern tensor processing hardware. In this paper we bring this issue to the attention of NLP researchers by quantifying the approximate financial and environmental costs of training a variety of recently successful neural network models for NLP. Based on these findings, we propose actionable recommendations to reduce costs and improve equity in NLP research and practice.*,”contai-

In specifico prenderemo in esame due tipi differenti di testi prodotti da (per mezzo di) sistemi di Intelligenza Artificiale: un libro e un articolo di rivista. Proprio il fatto che siano testi di tipi molto differenti permette di mostrare con più evidenza caratteristiche specifiche dei sistemi di IA nella scrittura di testi che compaiono in entrambi i casi. In particolare, entrambi i testi sono sintatticamente corretti ma mostrano problemi semantici (cioè di organizzazione complessiva dello sviluppo di senso) e sono quindi stati necessari interventi di soggetti umani per sanarli. Nella forma finale i testi sono dunque il prodotto dell'attività di soggetti umani + software di Intelligenza artificiale, il che nostra anche nell'ambito della produzione di testi che l'agency dei sistemi di IA è agency condivisa e dunque i soggetti umani sono co-responsabili delle operazioni dei sistemi di IA. La complessità e ampiezza di queste tematiche, e le loro ramificazioni, mostrano che gli studi di scienza della biblioteca aprono necessariamente alle tematiche di information literacy che non hanno solo la dimensione intra-professionale ma anche quella civile e di cittadinanza.

I due casi qui esaminati non sono costituiscono esempi *perfetti* di agency, come si vedrà nel corso dell'analisi sviluppata più avanti. Ma il contesto in cui sono stati generati (la relazione editore-autore) li presenta come tali e dunque da questa assunzione partiremo. L'elemento formale che in modo inoppugnabile chiarisce l'implicita ma netta affermazione di agency del software di Intelligenza Artificiale è l'indicazione autoriale associata ai due scritti: Beta Writer e GPT-3. Nel primo caso si indica (allusivamente, a causa dell'apparente nesso nome-cognome) un programma informatico, nel secondo caso si menziona esplicitamente il nome del software di Intelligenza Artificiale che ha prodotto il testo. Come si vedrà più avanti³ l'indicazione di un nome di autore non è secondaria, marginale o trascurabile, per definire l'oggetto e dunque in partenza ci atteniamo a ciò che questa indicazione usualmente significa e che ha ovviamente una forte connotazione promozionale, di marketing: attirare l'attenzione sull'oggetto grazie alla novità eclatante segnalata dal nome dell'autore. L'analisi mostrerà che ci sono degli elementi, o degli aspetti, di imprecisione in questa indicazione autoriale, proprio connessi con l'agency che usualmente si ritiene propria di un autore 'tradizionale' che qui non c'è, e che il nuovo autore non ha.

Beta Writer. 2019. «Lithium-Ion Batteries. A Machine-Generated Summary of Current Research». New York, NY: Springer

Nel 2019 l'editore Springer ha pubblicato a stampa e in digitale un volume intitolato «Lithium-Ion Batteries. A Machine-Generated Summary of Current Research» (Beta Writer 2019) la cui caratteristica principale è di esser stato prodotto per mezzo di un sistema *ad hoc* di Intelligenza Arti-

ner-title": "Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics", "DOI": "10.18653/v1/P19-1355", "event": "Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics", "event-place": "Florence, Italy", "language": "en", "page": "3645-3650", "publisher": "Association for Computational Linguistics", "publisher-place": "Florence, Italy", "source": "DOL.org (Crossref che con (Crawford 2021) è un importante riferimento sul tema. Il libro di Crawford è commentato da (Hao 2021) in un'intervista con l'autrice. Timnit Gebru, condirettrice del settore «Ethical AI» in Google, è stata licenziata il 2 dicembre 2020 quasi certamente per dissenso dell'azienda su un articolo che Gebru ha poi pubblicato nel 2021 (Bender et al. 2021) in cui segnalava il serio pericolo per l'ambiente dato dai consumi di energia necessari per l'addestramento dei modelli di AI.

³ Cfr. pag. 3.

ficiale, tanto che l'autore è stato denominato «Beta Writer»⁴, qualcosa come «autore in versione quasi finale», perché nel mondo del software la versione beta di un programma è una versione intermedia, ma avanzata e funzionante⁵ e dunque «Beta Writer» allude ad un processo di sviluppo che non è ancora completo ma è già pervenuto a risultati sufficientemente buoni da poter essere resi pubblici. Ovviamente, nessuna forma di Intelligenza Artificiale ha autonomamente deciso di produrre questo volume, e lo ha prodotto; bensì un editore ha deciso di utilizzare procedure differenti da quelle solitamente adottate per la produzione di un libro⁶. Genette richiama l'attenzione sull'effetto comunicativo che uno pseudonimo può produrre: «Ce qui nous concerne comme élément paratextuel, c'est ... l'effet produit sur le lecteur, ou plus généralement sur le public, par la présence d'un pseudonyme.»⁷

Il comunicato stampa che annunciava l'uscita del libro⁸ spiegava che esso contiene una rassegna di articoli scientifici sugli sviluppi della ricerca sulle batterie al litio, rassegna definita «machine-generated» e «automatically compiled by an algorithm», due espressioni sostanzialmente simili. Partner di Springer in questa attività è stato un gruppo di ricercatori dell'«Applied Computational Linguistics Lab» della Goethe Universität di Francoforte⁹. È chiaro che si è trattato di una scelta editoriale dell'editore che ha inteso attirare l'attenzione sul prodotto (il libro) al di fuori, soprattutto al di fuori, dei circoli degli esperti di Intelligenza Artificiale.

Analisi di questioni di politica editoriale

Il volume si apre con un'introduzione scritta da Henning Schönenberger, direttore del settore «data development» in Springer Nature, che dichiara:

With the exception of this preface [this book] has been created by an algorithm on the basis of a re-combined accumulation and summarization of relevant content in the area of Chemistry and Materials Science. The book is a cross-corpora auto-summarization of current texts from Springer Nature's content platform "SpringerLink", organized by means of a similarity-based clustering routine in coherent chapters and sections.¹⁰

⁴ L'autore è così descritto nella pagina web del volume (<https://link.springer.com/book/10.1007%2F978-3-030-16800-1>): «Machine-generated by Beta Writer 0.7 software developed at Goethe University Frankfurt».

⁵ Tale versione viene rilasciata affinché gli utenti interessati la utilizzino aiutando con ciò a verificare se vi siano errori che erano sfuggiti ai programmatori; ma questa dimensione di verifica del prodotto non viene dichiarata per il volume in questione.

⁶ Il libro ha suscitato il normale interesse di una pubblicazione scientifica, nondimeno non ne sono mancate letture che costituiscono forzature di specifici aspetti della produzione del libro: «Machines can generate literature reviews and even empirical data analysis and may soon write academic papers» (Haski-Leventhal 2020, 142).

⁷ (Genette 2002, cap. Le nom d'auteur-par. Pseudonymat)

⁸ (Springer Nature 2019).

⁹ («Projects and Cooperations - Applied Computational Linguistics Lab Goethe University Frankfurt, Germany» s.d.) Il progetto che ha portato alla pubblicazione del libro è intitolato «Schwach überwachte Verfahren zur Bibliographianaalyse», metodi per analisi bibliografica a supervisione debole. La collaborazione che ha portato alla pubblicazione del volume era stata avviata nel 2014 e si colloca in un quadro di vari progetti del Laboratorio che hanno impronta filologica, linguistica, e di digital humanities.

¹⁰ (Schoenenberger, Chiarcos, e Schenk 2019, V).

La risorsa utilizzata per la selezione delle fonti è dunque stata la piattaforma SpringerLink. Schönberger evidenzia poi una serie di questioni complesse direttamente connesse con la produzione di pubblicazioni ad opera di un sistema di Intelligenza Artificiale. L'assunzione di principio dichiarata è la *trasparenza*:

Full transparency is essential for us to discover both the opportunities of machine-generated content and the current limitations that technology still confronts us with. But it was also an ethical decision that if we start this journey, we want to do so in a correct and responsible manner, in order to enable a discussion in the research communities that is as open-minded as possible.¹¹

We aim to demonstrate both possible merits and possible limitations of the approach, and to put it to the test under real-world conditions, in order to achieve a better understanding of what techniques work and which techniques do not.¹²

Rendere possibile una discussione a mente libera nelle comunità di ricerca sulle questioni in gioco: pregi e limiti dell'approccio, quali tecniche funzionano e quali no. I primi interrogativi posti da Schönberger riguardano la funzione autoriale:

Who is the originator of machine-generated content? Can developers of the algorithms be seen as authors? Or is it the person who starts with the initial input (such as “Lithium-Ion Batteries” as a term) and tunes the various parameters? Is there a designated originator at all? 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?¹³

che Schönberger pone di fatto come questione dell'agency, segnalando in relazione alla produzione del libro che il sistema di Intelligenza Artificiale ha operato in base alle assunzioni e scelte di progettisti e operatori e chiedendo(si) chi sia responsabile del contenuto.

Un'altra grande categoria di questioni affrontate da Schönberger riguarda la qualità intrinseca delle fonti (che dipende dalla qualità della *peer review*), e l'estrazione e selezione delle fonti in funzione di una rassegna su un tema specifico. Il software di Intelligenza Artificiale ha lavorato le fonti disponibili nella piattaforma SpringerLink in un implicito assunto che esse siano tutte di alta qualità, il che a sua volta presuppone che tutto il processo di peer review sia di alta qualità, mentre è noto che ogni anno un certo numero di articoli scientifici pubblicati viene ritirato per gravi difetti¹⁴. E oltre alla valutazione di qualità (la fonte ha valore scientifico?) c'è quella forse anche più complessa di pertinenza (la fonte è appropriata al discorso che si vuole sviluppare?): due tipi di valutazione che nella produzione di questo libro non sono entrati. La riflessione di Schönberger si sposta poi verso i modi e la qualità della scrittura machine-based:

¹¹ Ivi, pag. VII.

¹² Ivi, pag. X.

¹³ Ivi, pag. VII.

¹⁴ Si veda ad esempio la disamina del problema presentata da (Brainard e You 2018) e (Vuong 2020), che sollecita la domanda (per ora?) senza risposta **è** un sistema di Intelligenza Artificiale in grado di riconoscere se un articolo scientifico è per qualche ragione difettoso? che è poi il motivo per cui si opera la peer review.

the extractive summarization of large text corpora is still imperfect, and paraphrased texts, syntax and phrase association still seem clunky at times. However, we clearly decided not to manually polish or copy-edit any of the texts due to the fact that we want to highlight the current status and remaining boundaries of machine-generated content. ... How will the publication of machine-generated content impact our role as a research publisher?

... we know that the quality of machine-generated text can only be as good as the underlying sources which have been used to generate it. At Springer Nature, we publish research which stands up to scientific scrutiny. In consequence, machine-generated content makes it even more necessary to re-emphasize the crucial role of peer-review itself. Though peer-review is also in the course of being continuously re-defined (and in the future we expect to see substantial progress in machine-support also in this regard) we still think that for the foreseeable future we will need a robust human review process for machine-generated text.

Si tratta di una riflessione complessa. In primo piano c'è il riconoscimento che il testo prodotto dal software di Intelligenza Artificiale lascia a desiderare. Più complicato, e appena accennato, quale possa essere lo spazio, o il ruolo, per un editore scientifico quando si affermi la «publication of machine-generated content». Non meno complessa la questione del «robust human review process» che continuerà ad essere necessario perché un'elaborazione puramente algoritmica dell'informazione non è in grado di individuare se un articolo è metodologicamente valido. Per la *peer review* sarà indiscutibilmente utile il supporto di sistemi di Intelligenza Artificiale¹⁵ che non rimpiazzano l'attività dello studioso, piuttosto la supportano («substantial progress in machine-support also in this regard») liberandola dai compiti esecutivi e aiutandola a focalizzarsi sul contenuto e sulla sua valutazione¹⁶. Ma occorre riflettere su come si potrà fare la *peer review* di scritti composti con sistemi di Intelligenza Artificiale. Un sistema di Intelligenza Artificiale non è un *peer* di uno studioso, «the term peer itself indicates a certain inadequacy for machine-generated research content. Who are the peers in this context? Would you as a human reader consider yourself as peer

¹⁵ Si può immaginare un ambiente di lavoro di questo tipo per il revisore: apre il file presentato per la pubblicazione e il sistema di Intelligenza Artificiale di supporto individua e apre le fonti citate, per ogni fonte sia evidenzia citazioni dirette o indirette nel testo in valutazione; sia riporta i vari indici di citazione disponibili, se esistenti; e per l'autore e il tema in valutazione individua eventuali altre pubblicazioni simili per permettere la valutazione di originalità.

¹⁶ Continua ad essere valido il paradigma Licklider-Engelbart formulato negli anni 60 del secolo scorso, nel quale il computer svolge compiti funzionali liberando tempo per l'intelligenza umana che si può maggiormente dedicare a compiti di tipo ideativo/interpretativo/creativo. Esso è espresso in modo non formalizzato in (Licklider 1960b), (Engelbart 1962), (Engelbart e English 1968). Siamo però testimoni di una fase di evoluzione per cui il software è passato da compiti esecutivi in senso riduttivo (trovare un indirizzo in una rubrica in base alla posizione in cui ci troviamo) a compiti esecutivi più complessi (attivare una data configurazione del computer, in base alla posizione in cui ci troviamo) ma che non comportano un'azione che modifica lo stato delle cose. Fino a che oggi con la disponibilità di sistemi di Intelligenza Artificiale è possibile demandare l'esecuzione di compiti sempre più complessi che comportano azioni che modificano lo stato delle cose: il riconoscimento facciale di individuo attraverso telecamere di sorveglianza può portare all'incarcerazione di quell'individuo: la valutazione di una domanda di assicurazione vita può portare al fatto che una persona non sia in grado di sostenere il costo delle cure sanitarie; o un'azione di guerra condotta per mezzo di un drone autonomo (sorveglianza, identificazione di soggetti sospetti, programmazione di possibili azioni di contrasto, attuazione dell'azione – nel 2021 se ne è avuta una prima chiara occorrenza (Kallenborn 2021) nella guerra libica ad opera dell'esercito turco, occorrenza poi confermata dal rapporto dell'ONU (Choudhury et al. 2021, 17/548)) porta all'uccisione di individui da parte del drone.

to a machine?»¹⁷. I suoi processi di organizzazione e gestione dell'informazione non sono umani, e dunque «the review process on machine-generated research content needs refinement, if not a complete re-definition» (Schoenenberger, Chiarcos, e Schenk 2019, IX). Se il *reviewer* al termine di una serie di osservazioni puntuali conclude «ripresentare con correzioni», qual è il «soggetto» che interpreta le osservazioni del *reviewer* e le rende in forma di modifiche appropriate? Il software di Intelligenza Artificiale riesce (abbastanza bene) a scrivere frasi, anche a scrivere una frase simile ad un'altra, ma non sa che cosa significhi ciò che scrive. Sullo sfondo c'è il pensiero che un esperto in uno specifico campo di ricerca non deve diventare un esperto in reti neurali e in NLP per poter valutare la qualità di un testo prodotto da un'Intelligenza Artificiale. Schoenenberger crede che «soon enough we will see machine-generated texts from unstructured knowledge bases that will lead to more complex evaluation processes»¹⁸ ma sembra non considerare il problema che generare testo a partire da basi di conoscenza non strutturate *ha senso per il mondo* solo se la conoscenza infondata, la falsa conoscenza viene riconosciuta e messa da parte, il che richiede di «conoscere il mondo», qualunque cosa ciò significhi in dettaglio, e saper rapportare a ciò il testo che si legge decidendo se le discrepanze sono ammissibili, fittizie, inammissibili (=falsità); e se in chi scrive c'è consapevolezza di avere responsabilità per ciò che scrive, a partire dall'autorevolezza che si costruisce nel proprio contesto, fino agli effetti che la conoscenza condivisa può produrre su/per mezzo di altri. Il pensiero a questo punto corre allo spazio di azione che avranno i sistemi di Intelligenza Artificiale e a quello che *resterà* per gli umani:

We foresee that in future there will be a wide range of options to create content—from entirely human-created content to a variety of blended man-machine text generation to entirely machine-generated text. We do not expect that authors will be replaced by algorithms. On the contrary, we expect that the role of researchers and authors will remain important, but will substantially change as more and more research content is created by algorithms. To a degree, this development is not that different from automation in manufacturing over the past centuries which has often resulted in a decrease of manufacturers and an increase of designers at the same time. Perhaps the future of scientific content creation will show a similar 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.”¹⁹

La frase «We become writers of writers» che si potrebbe tradurre, per esplicitarne il senso, con «diventiamo configuratori di sistemi di scrittura», dice che i sistemi di Intelligenza Artificiale non hanno agency, non più di un pianoforte rispetto ad un pianista, perché essa rimane nel soggetto umano che configura e utilizza il sistema di Intelligenza Artificiale. Posizione non di un com-

¹⁷ Ovviamente si possono ipotizzare sistemi di Intelligenza Artificiale che operano in maniera concorrente, ma il problema viene rinviato, non risolto; a meno che si voglia demandare in toto la *peer review* ai sistemi di Intelligenza Artificiale.

¹⁸ (Schoenenberger, Chiarcos, e Schenk 2019, IX).

¹⁹ (Schoenenberger, Chiarcos, e Schenk 2019, IX). La citazione interna è da (Goodwin 2016b) 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 «I the road» prodotto utilizzando un sistema di Intelligenza Artificiale interconnesso con sensori, microfono, macchina fotografica (Goodwin, McDowell, e Planquelle 2018).

mentatore terzo, ma di un autore che ha usato intensamente sistemi di Intelligenza Artificiale per produrre due opere **differenti**²⁰.

In questa linea di pensiero, la promessa, il disegno del futuro, per Schönenberger è che gli autori non saranno rimpiazzati da algoritmi ma che il ruolo autoriale cambierà perché sempre più contenuto di ricerca sarà prodotto da algoritmi e quindi attività che oggi consideriamo di alto livello potranno essere demandate alle/svolte dalle macchine e nuovi spazi si apriranno alla scrittura scientifica. Nondimeno nelle sue parole due concetti si scontrano: «we do not expect that authors will be replaced by algorithms» e «more and more research content is created by algorithms». Se il contenuto di ricerca sarà prodotto dagli algoritmi come si può sostenere che gli algoritmi non soppianteranno gli autori? si sta affermando che un autore scientifico si qualifica come tale non per la produzione di contenuto?²¹ Da una parte quindi ci si può chiedere *che cosa fa o che cosa è un autore, se non produce contenuto*; dall'altra il paradigma Licklider-Engelbart che abbiamo già ricordato alla **nota 16** ricorda che ad ogni epoca si scopre che una parte dei compiti che si reputavano creativi o inestricabilmente connessi con la creatività possono essere demandati a programmi di computer. Pertanto la domanda si riformula in questi termini: *tra i caratteri che oggi reputiamo specifici dell'essere autore, che cosa è essenziale e che cosa è secondario?*

La posizione di Schönenberger si può leggere in linea di continuità con le trasformazioni che nell'ultima ventina d'anni gli strumenti digitali e le reti hanno portato nel lavoro degli studiosi con la progressiva disponibilità prima degli OPAC, poi delle biblioteche digitali, poi dei cataloghi di risorse in Accesso Aperto, di Google Scholar e di DOAJ e delle piattaforme degli editori, dei tool di information retrieval, dei discovery tools, delle segnalazioni di articoli interessanti da parte delle piattaforme editoriali, e così via. Tutte risorse digitali che gestiscono e offrono accesso all'informazione scientifica²² e che volendo si potrebbero vedere come forme del passato in cui appare l'evolvere del «more and more research content is created by algorithms». Ma la perplessità rimane.

Analisi degli aspetti tecnici della produzione del libro

Alle analisi di Schönenberger seguono quelle di Christian Chiarcos e Niko Schenk dell'«Applied Computational Linguistics Lab» della Goethe Universität, che discutono della procedura di generazione (scrittura) del libro e della selezione delle fonti:

²⁰ (Goodwin 2016a) e (Goodwin, McDowell, e Planquelle 2018).

²¹ Si può riprendere utilmente la distinzione di Peirce tra ragionamenti corollari, che sviluppano premesse note, e ragionamenti teorematichi il cui autore adotta un metodo sperimentale.

²² Analogamente, crediamo, sarà per il problema del sovraccarico informativo, che parzialmente si risolve(rà) man mano che diventino disponibili strumenti di raccolta e selezione di fonti capaci di operare più pervasivamente e selettivamente di quanto possa fare oggi qualsiasi ricerca in qualsiasi ambito o servizio del web. Il cosiddetto «sovraccarico informativo» peraltro sembra preferibile concepirlo come, e chiamarlo, «abbondanza informativa». Un sovraccarico è indesiderabile e deve essere eliminato; l'abbondanza è desiderabile e deve essere governata con strumenti adeguati. Quello del «sovraccarico informativo» non è tema di oggi e può in vario modo essere tracciato indietro nel tempo almeno fino all'invenzione della stampa, e anche più indietro (Blair 2003); (Blair 2010).

we decided for a relatively conservative approach, a workflow based on

1. document clustering and ordering,
2. extractive summarization, and
3. paraphrasing of the generated extracts. (Schoenenberger, Chiarcos, e Schenk 2019, XI)

Tre fasi operative, la prima è «document clustering» e non qualcosa come «ricerca ed estrazione degli articoli da SpringerLink». La ragion d'essere di questo libro, in un campo dove c'è un'ampissima produzione scientifica, è di creare una struttura coerente di contenuti quindi organizzare le fonti per temi, «clustering and ordering»; e per ogni tema scrivere una sintesi introduttiva ad ogni capitolo (Schoenenberger, Chiarcos, e Schenk 2019, XI). Qui si colloca l'utilità, per il lettore/studioso, del ricorso ad un sistema di Intelligenza Artificiale: leggendo le introduzioni ai capitoli, che corrispondono a campi specifici della ricerca sulle batterie al litio, si ha una visione di sintesi delle novità emerse in quel campo. La sintesi prodotta dal software di Intelligenza Artificiale è più raffinata di ciò che chiunque potrebbe fare semplicemente con una ricerca in SpringerLink, che contiene 53.000 fonti sul tema pubblicate negli ultimi 3 anni. Spiegano Chiarcos e Schenk:

In preparation for generating a book, we identify a seed set of source documents as a thematic data basis for the final book, which serve as input to the pipeline. These documents are obtained by searching for keywords in publication titles or by means of meta data annotations. (Schoenenberger, Chiarcos, e Schenk 2019, XIV)

La sfumatura è sottile e sta tutta nella parola *seed* (seme): *identifichiamo un set di fonti sul tema e le usiamo come seed, come input iniziale del software di Intelligenza Artificiale*²³. Non viene aggiunto altro e il lettore consapevole intende che il software di Intelligenza Artificiale ha provveduto alla ricerca ed estrazione da SpringerLink degli articoli da usare per la produzione del libro imparando dagli esempi del *seed* quali articoli scegliere. Coloro che hanno selezionato il set di articoli iniziale sono quindi a pieno titolo (co)autori della pubblicazione, infatti senza di loro o con differenti loro scelte la pubblicazione prodotta da Beta Writer sarebbe stata differente. Si vorrebbe sapere di più su questo passaggio, ma come Schönemberger stesso ha dichiarato in un passo che abbiamo citato sopra «it becomes increasingly difficult to understand how a result has been actually derived»²⁴. Istruito con il set di articoli individuati dagli esseri umani, il software di Intelligenza Artificiale ha estratto da SpringerLink una collezione di 1086 pubblicazioni selezionate in base a parole presenti nel titolo o nei metadati, e per anno di pubblicazione. Tale collezione è stata successivamente lavorata per ordinare e raggruppare le fonti e creare così la struttura della pubblicazione. Gli sviluppa-

²³ È il procedimento che si adotta con le reti neurali: si mostra al software un tipo di esito desiderato (in questo caso: gli articoli d'esempio scelti per tema, qualità eccetera) affinché produca altri esiti simili (in questo caso individuare altri articoli sull'argomento). Si vedano come introduzioni al tema (Hagan, Demuth, e Beale 1996) e (Wang 2003).

²⁴ Proprio in riferimento a queste questioni si sta sviluppando un orientamento critico denominato XAI, eXplainable Artificial Intelligence, il cui scopo è «to make a shift towards more transparent AI. It aims to create a suite of techniques that produce more explainable models whilst maintaining high performance levels» su cui si vedano ad esempio (Adadi e Berrada 2018, 52138) che presentano una rassegna di studi sulle vari forme di XAI oppure (Barredo Arrieta et al. 2020). Amichevole, ma non meno robusto, l'articolo di (Gunning e Aha 2019) **transportation, security, medicine, finance, defense** che ha molte componenti visive. La XAI a sua volta è espressione della spinta della società e degli scienziati verso l'«ethical AI», per la quale validi punti di entrata sono (Floridi et al. 2018), (Taddeo e Floridi 2018) e (Mittelstadt 2019)

tori inizialmente valutarono se raggruppare le fonti basandosi sulla sovrapposizione bibliografica, o sulla somiglianza testuale dei documenti, e scelsero poi di quest'ultima perché adottando la sovrapposizione bibliografica risultava difficile bilanciare fonti con bibliografia molto ampia e fonti con bibliografia meno ampia. La somiglianza testuale fornì dunque dati lessicali per la cluster analysis²⁵ cioè la suddivisione e raggruppamento delle fonti che si svolse in due fasi: prima vennero individuati i «core thematic topics» che diedero luogo ai capitoli; poi all'interno di essi i «subtopics», le sezioni. L'indice sommario venne messo a punto con l'intervento di esperti umani:

The result of this process is a structured table of content. At this level, subject matter experts requested the possibility for manual refinement of the automatically generated structure. We permit publications to be moved or exchanged between chapters or sections, or even removed if necessary, for example, if they seem thematically unrelated according to the domain expertise of the editor. We consider the resulting publication nevertheless to be machine-generated. (Beta Writer 2019, XV)

L'intervento degli esperti umani per la messa a punto del contenuto (eliminare un articolo oppure spostarlo da un capitolo ad un altro) opera sulla pertinenza espositiva, ha a che fare con la costruzione complessiva di senso; non è descritto ma sarebbe stato interessante capire, ad esempio, perché alcune pubblicazioni siano state espunte in fase finale in quanto «thematically unrelated», anche se questa non-pertinenza tematica avrebbe potuto essere individuata già all'inizio quando venne costituita la collezione iniziale delle 1086 pubblicazioni.²⁶

GPT-3. 2020. «A Robot Wrote This Entire Article. Are You Scared yet, Human?» The Guardian, 8 settembre 2020

La produzione del libro sulle batterie al litio e della rassegna di studi sul Covid-19 si **iscrivono** nel grande filone di ricerca della produzione di testi da parte dei sistemi di Intelligenza Artificiale. Da una parte **come si è visto** la produzione di testi può avere uno scopo funzionale (come ad esempio informare su quale atleta ha vinto una certa gara in una certa data e luogo) e non necessariamente intrinsecamente creativo (come potrebbe essere «scrivere un racconto» o, ben più complesso,

²⁵ Specificamente, *recursive non-hierarchical clustering*: PCA (analisi dei componenti principali) con vincolo di generare 4 cluster (i capitoli) e per ognuno di essi dei sottocluster costituiti ciascuno dai 25 elementi più rilevanti.

²⁶ Springer Nature nel 2020 ha utilizzato software di Intelligenza Artificiale per realizzare una seconda pubblicazione, non più in forma-libro, nella quale si può riconoscere che un'evoluzione delle attività di Intelligenza Artificiale utilizzate per produrre il libro del 2019. Si tratta di «SARS-CoV-2 (COVID-19)» (<https://springernature.github.io/covid19-publications/>), una rassegna di fonti focalizzata sul Covid-19, pubblicata come sito web in uno spazio terzo (github.io) e senza indicazione di autore. L'indicazione dell'impiego di software di Intelligenza Artificiale per produrre la rassegna non è contenuta nella rassegna stessa, dalla quale occorre prima seguire il link «Coronavirus | For Researchers» e poi nella successiva pagina il link «How can AI support the research community in times of crisis?» che porta al post (Kaindl e Preuss 2020) del blog «The Source». Elsevier invece, con un approccio più di basso profilo, per mezzo di sistemi di analisi del linguaggio naturale individua negli articoli scientifici (del)le entità e per ciascuna di tali entità individuate inserisce un link che porta ad una selezione di articoli o capitoli di libri pertinenti a quell'entità (cfr. (Carpenter 2021)). La selezione viene effettuata da un sistema di Intelligenza Artificiale che estrae i contenuti dalle riviste pubblicate dall'editore e genera le «topic pages» dell'editore (Elsevier 2021). Siamo di fronte ad una gestione e classificazione di fonti simile a quella pensata da Springer, ma l'ambito è molto più vasto e la forma editoriale non mette in evidenza la modalità produttiva.

«scrivere una poesia», ma negli anni 50/60 ci si provò), cioè uno scopo «utile». Dall'altra parte proprio la difficoltà intrinseca della produzione di testo creativo spesso privo di vincoli con lo stato del mondo (non ci sono regole o pratiche che definiscano quando, come, chi, perché, su quale argomento, deve scrivere un racconto o una poesia) ne segnala un altissimo valore dimostrativo delle (eventuali) capacità di un sistema di Intelligenza Artificiale.

Proprio in questa funzione dimostrativa, e anche un po' provocatoria, si iscrive un articolo giornalistico prodotto da un sistema di Intelligenza Artificiale non specializzato su tale funzione bensì focalizzato in senso più ampio sull'analisi e produzione di testi basati sul linguaggio naturale. Si tratta di GPT-3²⁷, prodotto dall'azienda di San Francisco OpenAI. L'articolo «A robot wrote this entire article. Are you scared yet, human?», autore GPT-3, pubblicato nel 2020 dal Guardian (GPT-3 2020), è menzionato nel web 140.000 volte ad indicare una notevole attenzione e relativo dibattito. Una nota 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. ... 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.» The prompts were written by the Guardian, and fed to GPT-3 by Liam Porr GPT-3 produced eight different outputs, or essays. Each was unique, interesting and advanced a different argument. The Guardian could have just run one of the essays in its entirety. However, we chose instead to pick the best parts of each, in order to capture the different styles and registers of the AI.

Il software dunque è stato configurato per operare su una struttura di base predisposta dai redattori (indicazioni di lunghezza: 550 parole; avvio del testo: «I am not a human ...»). Senza questa impalcatura la produzione del testo sarebbe differente, o difficile da realizzare, o impossibile. Qui si vede che un sistema di AI inizia a produrre testo perché (dopo essere stato addestrato) una serie di input da parte di esseri umani lo hanno configurato e predisposto ad operare. Infatti un programma di produzione di testo non parte dal nulla deliberando autonomamente se operare, in quale modo e su quale argomento: è la questione dell'agency già ricordata a proposito della produzione del libro Springer. Con GPT-3 al Guardian sono stati prodotti 8 articoli e la redazione scelse le parti migliori di ciascuno di essi («we chose to pick the best parts of each»).

È quindi 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!

²⁷ GPT sta per Generative Pretrained Transformer: trasformatore generativo pre-addestrato. Il tema dell'addestramento è cruciale.

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 casualties.

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.

È impressionante inoltre la sostanziale assenza nel testo dei connettivi grammaticali e logici che esprimono in modo fluido lo sviluppo ed evoluzione dell’argomentazione. Ovviamente anche un testo di autore umano potrebbe presentare gli stessi limiti, o difetti; ma non sarebbe portato come esempio di capacità di scrittura. Non si può non pensare al test di Turing: con la sua struttura centrata su testi scritti (risposte) generati in seguito a domande di un interrogante umano a due soggetti di cui non si sa se/chi sia umano; in esso è implicito che un sistema software-hardware *al massimo livello delle sue capacità* debba risultare indistinguibile da un essere umano. Il software GPT-3 seleziona «cose appropriate» da dire per costruire il testo ma esse si presentano come una pila instabile di mattoni ai quali manca la malta che li tenga solidamente insieme. Considerazioni analoghe vengono fatte da Guido Vetere in riferimento alla produzione di musica da parte di Jukebox che è un generatore di musica ‘nello stile di...’ basato su GPT-3 (Vetere 2021): coerenza locale ma assenza di struttura. Con una metafora si potrebbe pensare ad un conversatore da party, che sa inserire in modo agile e brioso alcune frasi in una conversazione di altri (controllo della sintassi) ma non ha un discorso suo *organico e appropriato allo stato del mondo* (controllo della semantica). Dei testi scritti da sistemi di Intelligenza Artificiale si commenta molto la coerenza e la pertinenza informativa, e molto meno ciò che a nostro avviso è anche più importante, cioè la capacità di interpretare lo stato del mondo (sì, certo: che cosa vuol dire «lo stato del mondo?»); ma qualunque cosa esso sia, interpretarlo e agire di conseguenza è ciò che ogni soggetto umano fa in ogni istante) e in base a ciò operare in modo adeguato.

Agency e autorialità²⁸

La produzione del libro si iscrive in un quadro complesso per cui le principali aziende che sviluppano sistemi di Intelligenza Artificiale (OpenAI partecipata da Microsoft il cui ultimo prodotto, GPT-3, viene descritto come *language generator*; DeepMind acquistata da Google; EleutherAI, «a decentralized grassroots collective of volunteer researchers, engineers, and developers focused on AI alignment, scaling, and open source AI research» che mira a replicare in versione aperta il software di OpenAI) vedono nell'elaborazione del linguaggio e nella produzione di testi una sorta di Sacro Graal. Elaborazione e produzione che si esprimono anche in forma di produzione di testi «nello stile di...»; produzione (o completamento) di immagini a partire da descrizioni verbali; generazione di descrizioni verbali di immagini; completamento coerente di testi parziali, e altro ancora. Molto a lungo si potrebbe analizzare e discutere che cosa significhi «produrre un testo»: quando avviene, perché avviene, quale relazione hanno il contesto esterno e l'intenzione dell'autore, quale relazione ci debba essere, o possa non esserci, tra il testo e – come dicevamo – «lo stato del mondo». In questo articolo focalizziamo l'attenzione su una questione rilevante nel campo della library science, cioè quello dell'autorialità di un prodotto testuale di un sistema di Intelligenza Artificiale, quale può essere il libro sulle batterie al litio o anche l'articolo pubblicato dal Guardian.

Chi è l'autore e che cosa fa, in prospettiva bibliografica

L'esistenza di un libro il cui autore è indicato come Beta Writer, o di un articolo di giornale il cui autore è indicato come GPT-3, solleva questioni complesse riguardanti il fatto che l'autore a cui fanno riferimento i sistemi di catalogazione è definito come persona (in subordine come ente in quanto un ente è formato da persone), sia nel quadro di riferimento italiano:

Per responsabilità, ai fini catalografici, si intende la relazione che lega un'opera o una delle sue espressioni a una o più persone o enti che l'hanno concepita, composta, realizzata, modificata o eseguita²⁹

sia, giusto a titolo di **esempio** per un sistema molto differente, in quello statunitense:

The U.S. Copyright Office will register an original work of authorship, provided that the work was created by a human being.³⁰

²⁸ IFLA ha pubblicato nel 2020 un ampio «Statement on Libraries and Artificial Intelligence», completato da 3 appendici (IFLA 2020) Le riflessioni lì presenti segnalano utilmente una necessità di formazione sia per i professionisti della biblioteca sia per gli utenti, e individuano potenzialità e rischi in relazione ai vari ambiti d'uso. Ma si tratta di riflessioni intelligenti e pragmatiche *a valle*, e si sente la mancanza di una riflessione critica *a monte* sul perché i sistemi di Intelligenza Artificiale, su quale sia il *quid* dei software di Intelligenza Artificiale nel sistema bibliotecario, e così via.

²⁹ (ICCU 2009, par. 0.1.2.4)

³⁰ (U.S. Copyright Office 2021, par. 306. The Human Authorship Requirement)DC", "publisher-place": "Washington, DC", "title": "Compendium of U.S. Copyright Office practices", "author": [{"family": "U.S. Copyright Office", "given": ""}], "issued": {"date-parts": [{"2021"}]}, "locator": "306. The Human Authorship Requirement", "label": "section"}, "schema": "https://github.com/citation-style-language/schema/raw/master/csl-citation.json")

Contigua alla domanda su chi può essere autore/chi è l'autore, è la domanda su che cosa fa l'autore, in che cosa consista la sua attività. Dublin Core definendo «creator» come «an entity (a person, an organization, or a service) primarily responsible for making the resource»³¹, amplia e generalizza il significato di autore perché l'orizzonte si allarga dal libro alla «risorsa».

La riflessione sul significato del concetto di autore ha la sua prima ragion d'essere, come scrive Svenonius, nel fatto che

In Western cultures and since medieval times, authorship has been the primary identifying attribute of works. Users tend to remember and search for works by the persons responsible for their creation. This is evidenced by the primacy (in Western cultures) of the author catalogue, where entries are arranged so that a user will find all works of an author collocated together.³²

E già nel 2004 ricordava che «the nature of authorship has changed significantly over the last hundred years. For one thing, it has become increasingly collaborative»³³.

Analogamente Genette sottolinea l'importanza del nome dell'autore per le opere scientifiche/informative:

Le nom de l'auteur remplit une fonction contractuelle d'importance très variable selon les genres : faible ou nulle en fiction, beaucoup plus forte dans toutes les sortes d'écrits référentiels, où la crédibilité du témoignage, ou de sa transmission, s'appuie largement sur l'identité du témoin ou du rapporteur.³⁴

Queste espressioni, e i concetti che le innervano, in area bibliografica si basano sul concetto *letterario* di autore che la nostra cultura ha messo a punto nel corso dei secoli: un soggetto unitario, autonomo, capace di creatività, che liberamente dispone delle indicazioni che gli giungono dal mondo in cui vive. Non tutto è così ovvio però se già nel 1200 Bonaventura da Bagnoregio scriveva che la funzione-autore nella produzione dei libri si può esplicitare in quattro forme: *scriba* (trascrive cose di altri); *compilatore* (scrive unendo cose non sue); *commentatore* (scrive cose principalmente di altri alle quali aggiunge una parte sua di spiegazione); *autore* (scrive cose principalmente sue a cui aggiunge a conferma cose di altri)³⁵. Quando Barthes scrisse nel saggio «La mort de l'auteur», del 1968

Linguistiquement, l'auteur n'est jamais rien de plus que celui qui écrit, tout comme je n'est autre que celui qui dit *je* ...

Le scripteur moderne naît en même temps que son texte ; il n'est d'aucune façon pourvu d'un être qui précéderait au excéderait son écriture. ...

³¹ <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/elements11/creator/>

³² (Svenonius 2000, cap. 3 Bibliographic entities-par. Superworks)

³³ (Svenonius 2000, cap. 3 Bibliographic entities-par. Author Sets): «a thing is usually what it represents itself as being».

³⁴ (Genette 2002, cap. Le nom d'auteur-par. Onymat)

³⁵ (Bonaventura da Bagnoregio 1882, I:14–15): «Quadruplex est modus faciendi libros. Aliquis enim scribit aliena, nihil addendo vel mutando; et iste mere dicitur "scriptor". Aliquis scribit aliena addendo, sed non de suo, et iste "compilator" dicitur. Aliquis scribit et aliena et sua, sed aliena tamquam principalia, et sua tamquam annexa ad evidentiam, et iste dicitur "commentator" non "auctor". Aliquis scribit et sua et aliena, sed sua tamquam principalia, aliena tamquam annexa ad confirmationem; et talis debet dici "auctor"».

Nous savons maintenant qu'un texte n'est pas fait d'une ligne de mots dégageant un sens unique, en quelque sorte théologique (qui serait le message de l'auteur-Dieu), mais un espace à dimensions multiples, où se marient et se contestent des écritures variées, dont aucune n'est originelle : le texte est un tissu de citations issues des mille foyers de la culture.³⁶

le discussioni furono complesse e vivaci. Qui importa ricordare che nel 1970 Barthes riprende questa quadripartizione nella la forma descritta da Bonaventura ma senza citarlo allo scopo di leggere già nel «texte ancien» la sua prospettiva della «morte dell'autore»³⁷:

Quant à l'écrit, il n'est pas soumis, comme aujourd'hui, à une valeur d'originalité ; ce que nous appelons l'auteur n'existe pas ; autour du texte ancien, seul texte pratiqué et en quelque sorte géré, comme un capital reconduit, il y a des fonctions différentes : 1) le *scriptor* recopie purement et simplement ; 2) le *compiler* ajoute à ce qu'il copie, mais jamais rien qui vienne de lui-même ; 3) le *commentator* s'introduit bien dans le texte recopié, mais seulement pour le rendre intelligible ; 4) l'*auctor*, enfin, donne ses propres idées mais toujours en s'appuyant sur d'autres autorités. Ces fonctions ne sont pas nettement hiérarchisées : le *commentator*, par exemple, peut avoir le prestige qu'aurait aujourd'hui un grand écrivain (ce fut, au XII^{ème} siècle, le cas de Pierre Hélie, surnommé le *commentator*).

Se la solidità e robustezza del concetto di autore viene messa in crisi in letteratura, viene corrispondentemente messa in crisi la robustezza del concetto di autore anche in campo bibliografico. Non si può non pensare al(l'autore) software che scrive essendo stato prodotto per effetto di molteplici intenti e influenze e che non esiste al di fuori di tale attività: quando non opera non consuma, non riposa, non fa altro, e se opera ciò accade perché è stato attivato da chi ha deciso di usarlo (e da chi ha deciso come esso deve funzionare). La forza della connessione dei due campi (letterario e bibliografico) in riferimento al concetto di autore appare chiara nel fatto che si dà (si può dare) per autoevidente che l'autore sia una persona, tanto che la cosa viene affermata, ma non discussa. Quando se ne discute (come ad esempio in Svenonius e Genette) da un lato l'orizzonte è quello dell'opera creativa individuale (quale è eminentemente l'opera letteraria); dall'altro la discussione verte su valenze implicazioni e significato del concetto di autore, non sulla sua ammissibilità.

Nel contesto italiano, il luogo in cui la riflessione bibliografica sul concetto di autore si concretizza e si esprime sono le regole REICAT³⁸. Quando le REICAT vengono pubblicate nel 2009³⁹ non considerano (e anche volendolo fare, non si saprebbe *come* considerare) le riflessioni di Barthes

³⁶ (Barthes 1968, 14). Prima pubblicazione in inglese (Barthes 1967) in Aspen, rivista americana; successivamente in francese (Barthes 1968). Citeremo facendo riferimento alla pubblicazione in francese.

³⁷ (Barthes 1970, 184–85). Nell'antichità classica questa quadripartizione di Bonaventura ripresa da Barthes non è attestata (tranne un paio di casi: «Itinerarium Alexandri», par. 7 *scriptor pro auctore laudatur*; Ps. Philargirius, «Vitae Vergilii», I, *cum ... compiler veterum diceretur*) né lo sono le qualifiche sistematizzate di *scriptor*, *compiler*, *commentator*, *auctor* (i nomi ricorrono, ovviamente, ma isolati, non all'interno di un sistema di concetti che li colleghi). Queste qualifiche iniziano a comparire con gli autori cristiani (Girolamo, Massimo di Torino, Isidoro di Siviglia, Beda, Incmaro di Reims) e trovano poi strutturazione in Bonaventura.

³⁸ Un'analisi critica del significato di autore nella prospettiva culturale italiana può partire da (Guerrini 2020) che del tempo presente evidenzia il passaggio dalla catalogazione tradizionale a metodologie di metadattazione. E quando scriviamo «la riflessione pre-Barthes si esprime nelle regole REICAT» poiché le REICAT vengono pubblicate 40 anni dopo

³⁹ (ICCU 2009)

perché esse pongono una questione, ma fino a questi anni esse sono state intese come una rilettura e decostruzione del concetto tradizionale di autore e non c'era stato modo di intuire, intravedere in esse la prefigurazione di un nuovo autore, un nuovo modo concetto di autore. Nuovo concetto di autore che richiede un tipo di consapevolezza critica che fino ad oggi non era **necessarie** e non esisteva. Le REICAT nella sezione in cui vengono discusse le varie forme di responsabilità si riferiscono sempre a persone: persone variamente note o ignote o che formano gruppi o che nascondono il loro vero nome come singoli o come gruppi⁴⁰. Ma un software di Intelligenza Artificiale non è né una persona né un ente. Prendendo a riferimento il libro sulle batterie al litio, di cui mostriamo in Figura 1 la copertina, come a metterci nei panni di un bibliotecario catalogatore che ha davanti a sé il libro, come descrivere la pubblicazione secondo le regole REICAT?

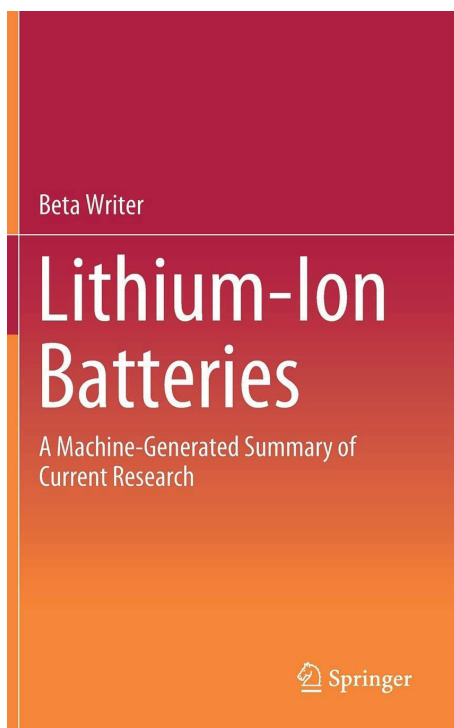


Figura 1. Copertina del libro Springer sulle batterie al litio

Nella copertina del libro la sequenza di caratteri «Beta Writer» *si presenta* in modo tale per cui può essere intesa (approccio costruttivista⁴¹) come nome di un autore personale, o pseudonimo di un autore personale (analogamente a «Romain Gary») o nome di un autore collettivo (analogamente

⁴⁰ Si potrebbe pensare di risolvere il problema ricorrendo a ISBD-ER dal momento che si tratta di una pubblicazione prodotta con sistemi di Intelligenza Artificiale; ma *questa* non è una pubblicazione elettronica, è invece una normale pubblicazione a stampa anche se il suo percorso produttivo è stato tutto digitale tranne il passaggio finale. E dunque i problemi di individuazione di autorialità si pongono nel contesto tradizionale della catalogazione di monografie.

⁴¹ (Svenonius 2000, cap. 3 Bibliographic entities-par. Superworks).

mente a «Luther Blisset»⁴²⁴³. Ma come abbiamo già visto esaminando la produzione del libro⁴⁴, Beta Writer non è nulla di tutto questo. Da un lato Beta Writer è il prodotto dell'attività di progettisti e programmatori; dall'altro Beta Writer ha iniziato a operare in base a un *seed* iniziale preparato da studiosi (cfr. pag. 8) dei quali non è noto il nome; successivamente (altri?) studiosi dei quali parimenti non è noto il nome hanno deciso quali articoli individuati da Beta Writer dovessero essere eliminati (cfr. pag. 9) e hanno effettuato messe a punto testuali minori. Si potrebbe pensare che «Beta Writer» sia da interpretare come nome di autore collettivo anonimo, se non fosse per il fatto che insieme agli esseri umani ha operato un sistema software che a quel punto verrebbe personificato - e ciò non pare appropriato. Se si pensasse che Beta Writer possa essere assimilato al nome di un Ente («un'organizzazione, un'istituzione o un gruppo di persone»⁴⁵) rimane il fatto che tutti i nomi di organizzazioni e istituzioni riportati a titolo di esempio dalle REICAT descrivono enti che operano per mezzo di persone. Tra l'altro le REICAT non danno una definizione di «persona», benché tale concetto sia centrale per la definizione della responsabilità, e in effetti quando le regole vennero prodotte il problema di «autori software» non si poneva. La concezione di persona rimanda oggi, in continuità col passato, ad un soggetto individuale che appartiene alla specie umana, soggetto che si specifica – anche in modo non continuativo – come cosciente di sé, pensante, capace di operare autonomamente.

Dunque le questioni da affrontare sono su due versanti: quale significato e riconoscimento dare all'operatività di soggetti umani in posizione decisiva/decisionale nei sistemi di Intelligenza Artificiale; quale significato e riconoscimento dare all'operatività della parte software dei sistemi di Intelligenza Artificiale, che non può essere né negata/ignorata, né semplicemente assimilata ad una persona per farla rientrare nelle forme previste dalle REICAT (e se le due possano essere tenute distinte o se debbano essere considerate parte di un'unica entità – di tipo fino ad ora non conosciuto).

La complessità del significato del concetto di autore che emerge in modo così chiaro quando ci si chiede in che senso Beta Writer e GPT-3 siano gli autori dei testi a loro ascritti richiama ovviamente all'attenzione le discussioni che sull'autore svilupparono nel secolo scorso Michel Foucault e Roland Barthes. Non è questa la sede appropriata (né il tempo) per riprendere le discussioni di ambito letterario, critico-letterario, filosofico, sulle loro posizioni. Ci interessa invece notare come la carica critica dei loro scritti sulla morte dell'autore in certo modo risulti meno destabilizzante quando venga messa a confronto con prodotti finiti del l'attività scrittoria dei sistemi di Intelligenza Artificiale, che per così dire, attivamente e positivamente la confermano. Di Barthes abbiamo già ricordato il saggio «La mort de l'auteur», del 1968, di cui è coevo il saggio di Foucault «Qu'est-ce qu'un auteur?»⁴⁶ nel quale il filosofo francese argomenta per mostrare che l'autore correntemente inteso non esiste (più), è morto:

il ne suffit pas, évidemment, de répéter comme affirmation vide que l'auteur a disparu. ... Ce qu'il faudrait faire, c'est repérer l'espace ainsi laissé vide par la disparition de l'auteur, suivre de l'oeil la ré-

⁴² (ICCU 2009, par. 17.1.1.A).

⁴³ Cfr. (Guerrini 2020, 75 ePUB): «Esistono numerosi casi in cui l'autore si presenta sul frontespizio di un libro con una formulazione generica, fuorviante o volutamente ingannevole.»

⁴⁴ Cfr. pag. 3 e pag. 5.

⁴⁵ (ICCU 2009, par. 16.0.1).

⁴⁶ (Foucault 1969).

partition des lacunes et des failles, et guetter les emplacements, les fonctions libres que cette disparition fait apparaître. ... Je voudrais d'abord évoquer en peu de mots les problèmes posés par l'usage du nom d'auteur. Qu'est-ce que c'est qu'un nom d'auteur ? Et comment fonctionne-t-il ? Bien éloigné de vous donner une solution, j'indiquerai seulement quelques-unes des difficultés qu'il présente. ... Un nom d'auteur exerce par rapport aux discours un certain rôle : il assure une fonction classificatoire. ... La fonction-auteur est liée au système juridique et institutionnel qui enserme, détermine, articule l'univers des discours ; elle ne s'exerce pas uniformément et de la même façon sur tous les discours, à toutes les époques et dans toutes les formes de civilisation ; *elle n'est pas définie par l'attribution spontanée d'un discours à son producteur, mais par une série d'opérations spécifiques et complexes ; elle ne renvoie pas purement et simplement à un individu réel, elle peut donner lieu simultanément à plusieurs ego, à plusieurs positions-sujets que des classes différentes d'individus peuvent venir occuper.*⁴⁷

Le righe finali che abbiamo evidenziato, datate 1969, così come quelle datate 1968 di Barthes, parlano in modo incisivo di ciò che avviene a 50 anni di distanza, nei nostri anni, nei quali sistemi di Intelligenza Artificiale operano nella scrittura; e viceversa la scrittura di testi da parte di tali sistemi rende le riflessioni di Barthes e Foucault *quasi* ovvie (!), più descrizioni analitiche di ciò che vediamo accadere in un ambito specifico che elaborazioni concettuali destabilizzanti di un sistema consolidato⁴⁸. Una situazione in cui sistemi di Intelligenza Artificiale producono testi complessi dotati di senso compiuto è destabilizzante: ma non si tratta più di una teoria su cui discutere, si tratta invece di una realtà che occorre analizzare e comprendere. In altre parole la scrittura di testi da parte di sistemi di Intelligenza Artificiale porta a compimento movimenti teorici, riflessioni, intuizioni, prospettive critiche che quando vennero presentate suscitavano discussione (e anche ripulsa) anche perché lontane dalla quotidianità produttiva dei testi che erano in sostanza sempre prodotti da persone fisiche. Qui invece abbiamo testi prodotti da entità immateriali, e questo scompiglia gli usuali riferimenti letterari su che cosa sia un autore, su che cosa sia scrivere, su che cosa sia un testo. Tutto ciò benché nella contemporaneità sia ormai abituale che l'autorialità sia *diffusa* e inafferrabile come nel caso di «edited works»; e *mista* per tutti quei prodotti editoriali che comportano la cooperazione più soggetti responsabili del contenuto, come scrive Svenonius⁴⁹. Ma a questo scompiglio che 50 anni fa fu introdotto nel campo della riflessione letteraria oggi sono necessarie risposte in campo bibliografico e specificamente all'ambito catalogafico.

L'agency dei sistemi di Intelligenza Artificiale

Crediamo lo «scompiglio» debba essere affrontato approfondendo la riflessione sull'autorialità dei sistemi di Intelligenza Artificiale e quindi prendendo in esame la questione della loro *agency*: se non lo si fa, li si assimila in maniera semplificante all'autore di contenuti editoriali di cui è respon-

⁴⁷ (Foucault 1969, 78,79,83). Corsivo nostro.

⁴⁸ Non mancano infatti le voci di coloro che sostengono che l'autore è tornato, o che in realtà non era morto. Qui ne discutiamo non in una prospettiva letteraria, che non sarebbe pertinente; ma in una prospettiva filosofica in cui questi concetti sono intesi come strumenti di cui ci si chiede se siano utili o no per approfondire la conoscenza e comprensione del funzionamento del mondo **bibliografico**

⁴⁹ (Svenonius 2000, cap. 3 Bibliographic entities-par. Superworks): «Authorship has also become increasingly diffuse and mixed».

sabile un soggetto umano, e in questo ovviamente ci sono elementi superficiali di corrispondenza. Ma i sistemi di Intelligenza Artificiale, noi crediamo, non sono semplicemente assimilabili a ciò che è già noto – e ciò emerge se si sviluppa la riflessione sull'*agency*. *Agency*, talora tradotto in italiano con «agentività», significa capacità di un individuo di agire in modo libero in un determinato contesto (ricollegandoci a quanto detto sopra, dire che «un autore è una persona» significa dire che «una persona può essere autore perché è dotata di *agency*»).

Una concisa presentazione delle principali tematiche che ruotano intorno all'*agency* tocca almeno **3** strutture concettuali fra loro connesse: l'*agency* individuale e il concetto di agente; e la teoria dell'*agency*, il concetto di agente. La nozione di *agency* individuale è «centered on a self with the capacity to effectively act upon the world»⁵⁰. L'*agency* è tanto il motore originario dell'azione quanto un suo prodotto riflessivo: il soggetto si riconosce dotato di *agency* (si riconosce agente) in quanto è in grado di operare efficacemente nel mondo⁵¹. *Agency* non significa che il soggetto sia totalmente libero da vincoli in quanto «actions are always already socially, culturally, and linguistically constrained»⁵². In letteratura scientifica, l'autore è usualmente inteso come un soggetto dotato di *agency*⁵³. La teoria dell'*agency*, che si sviluppa in ambito economico, si può formulare concisamente in questi termini: «in an agency relationship, one party acts on behalf of another»⁵⁴. Le due parti sono usualmente denominate *principal* e *agent*, committente e agente. Essa si sviluppa soprattutto in ambito economico e di teoria delle organizzazioni perché è appropriata a descrivere una parte delle relazioni che si sviluppano all'interno di una grande azienda, ma da tale ambito essa si è estesa poi a descrivere relazioni in molteplici ambiti, non ultimo quello autoriale dove si può individuare in questo senso, senza pretesa di assolutizzare o di esaurire le molteplici sfaccettature e aspetti, una relazione di *agency* tra editore e autore (l'editore ha un'intenzione produttiva /economica che si attua per mezzo del lavoro dell'autore), tra autore e pubblico (l'autore scrive su un tema su cui al pubblico interessa leggere qualcosa), tra autore ed editore (l'autore ha bisogno di ottenere un riconoscimento – apprezzamento e/o un introito economico che necessitano dell'attività dell'editore). Al di là di quanto si potrebbe pensare (cioè che la relazione di *agency* sia una relazione chiusa, che ha ambiti chiari e confini definiti in cui l'agente è in posizione di inferiorità rispetto al committente) temi ricorrenti nell'ambito di questa teoria sono le discrepanze di conoscenza tra committente e agente: il committente non sa sempre esattamente che cosa fa l'agente; l'agente, che è scelto per la sua competenza in uno dei molti campi in cui il committente vuole agire, solitamente ha del campo di azione una conoscenza più ampia/più profonda del committente. In tali situazioni di fatto l'agente è in posizione di superiorità rispetto al committente.

I sistemi di Intelligenza Artificiale possono essere interpretati come dotati di *agency* completa e indipendente (agenti che non hanno un committente, agenti che sono pienamente 'committenti di se stessi', come in generale si concepisce una persona adulta, senziente, intelligente, priva di han-

⁵⁰ (Gubrium e Holstein 1995, 555)

⁵¹ È una delle tesi portanti di (Mehan e Wood 1975).

⁵² (Ahearn 1999, 13).

⁵³ Alla data di scrittura di questo articolo una ricerca per «agency of the author» dà 37 esiti in SpringerLink, 817 in Science Direct, 1520 in Google Scholar (non interessa qui numerare le occorrenze uniche di tale stringa, ma individuare se di questo argomento si parla in letteratura). Semplicemente a dire che esiste nella letteratura scientifica un oggetto di discorso che è l'*agency* dell'autore.

⁵⁴ (Shapiro 2005, 263).

dicap cognitivi o fisici), oppure come dotati di agency condivisa con altri agenti (nel quadro di una relazione in cui un committente incarica molteplici agenti di agire), oppure come privi di agency (strumenti meccanici: un martello con cui si pianta un chiodo). La prima descrizione di Intelligenza Artificiale ante litteram è quella di Turing che nel 1950 parla di *imitation game*⁵⁵ nell'articolo famoso in cui si descrive l'esperimento che la macchina supererà quando un interlocutore alla tastiera non saprà distinguere le risposte di una macchina e di un essere umano. Molto vicino ad una definizione classica di Intelligenza Artificiale formulata nel "programma di Dartmouth" del 1955: «making a machine behave in ways that would be called intelligent if a human were so behaving»⁵⁶ che colloca l'indistinguibilità nel comportamento e non vede il software pienamente dotato di agency in quanto lo paragona ad un essere umano che agisce nel mondo: il software cioè non ha autonomia di azione e di indirizzo, ma deve conformarsi al modello umano di comportamento. Non si fa parola di agency ma l'intelligenza si manifesta, presuntamente, nella capacità di agire (ovviamente questa concezione imitativa/simulativa dell'intelligenza umana, e dunque dell'azione umana, presente anche nel test di Turing richiederebbe una discussione ben più ampia di quanto possibile in questo articolo). Taddeo e Floridi nel 2018 riformulano il concetto del 1955 in termini più ricchi e sfumati abbandonando l'idea dell'imitazione:

a growing resource of interactive, autonomous, self-learning agency, which enables computational artifacts to perform tasks that otherwise would require human intelligence to be executed successfully⁵⁷

ma proprio in materia di agency propongono una lettura più complessa:

the effects of decisions or actions based on AI are often the result of countless interactions among many actors, including designers, developers, users, software, and hardware. This is known as distributed agency. With distributed agency comes distributed responsibility.⁵⁸

Ciò il funzionamento di un sistema di Intelligenza Artificiale esprime le implicazioni e le conseguenze delle scelte fatte da coloro che lo hanno prodotto (*distributed agency*) – e tutte queste persone sono co-responsabili (*distributed responsibility*) con il sistema di Intelligenza Artificiale delle decisioni e azioni da esso operate⁵⁹. Ad oggi però questo è solo un principio, ad esempio nessuno conosce i nomi delle persone che hanno configurato il sistema di Intelligenza Artificiale «COMPAS» benevolo con i bianchi e severo con i neri che chiedevano la libertà condizionata (Larson et al. 2016) (come si vede, il tema dell'agency s'intreccia con quello dell'explainable AI e dell'ethical AI, già menzionate). Ovviamente le cose sarebbero più semplici se si optasse per l'agency completa del sistema di Intelligenza Artificiale (uno studio recente commissionato dal Parlamento

⁵⁵ (Turing 1950, 433)

⁵⁶ Il passo è solitamente citato con rinvio a (McCarthy et al. 2006) che pubblica nel 2006 una versione ridotta del «Proposal for the Dartmouth Summer research project on Artificial Intelligence» concepito e scritto nel 1955 ma mai precedentemente pubblicato a stampa. McCarthy 2006 è molto citato in letteratura scientifica, *ma il passo per cui generalmente viene citato non si trova lì*. Si trova invece in (McCarthy et al. 1955) che contiene il «Proposal» completo.

⁵⁷ (Taddeo e Floridi 2018, 751).

⁵⁸ (Taddeo e Floridi 2018, 751) Analoghe posizioni in (Yang et al. 2018)

⁵⁹ Il rapporto dell'ONU sull'Intelligenza Artificiale (United Nations e Kaye 2018) segnala tra l'altro che l'estendersi dell'agency dell'Intelligenza Artificiale comporta una simmetrica riduzione dello spazio dell'agency umana.

Europeo ha prospettato la possibilità di trattare i sistemi di Intelligenza Artificiale come persone elettroniche⁶⁰), perché in tal caso esso potrebbe essere trattato come autore singolo e ‘internamente omogeneo’, ma sarebbe comunque un autore non umano. Il fatto di usare il principio imitativo per concepire e valutare i sistemi di Intelligenza Artificiale viene messo in discussione da (Floridi e Cowsls 2019) che scrivono:

“Making a machine behave in ways that would be called intelligent if a human were so behaving.” (Quotation from the re-issue in McCarthy et al., 2006 [1955]). This is 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 does not mean that it cleans them like I do, or needs any intelligence to achieve its task. The same counterfactual understanding of AI underpins the Turing test (Floridi, Taddeo, & Turilli, 2009)⁶¹, which, in this case, checks the ability of a machine to perform a task in such a way that the outcome would be indistinguishable from the outcome of a human agent working to achieve the same task (Turing, 1950).

con ciò indicando che per risolvere il problema dell’autorialità dei sistemi di Intelligenza Artificiale non si può prendere la scorciatoia del pensare che essi siano in grado di riprodurre il comportamento di un soggetto umano: «non sono autori umani ma si comportano in tutto e per tutto come se lo fossero».

La questione dell’agency in prospettiva bibliografico/editoriale, inestricabilmente connessa con quella dell’autorialità e dei diritti di proprietà intellettuale, rimane viva perché si potrebbe pensare che al software di Intelligenza Artificiale possa essere riconosciuto un ruolo autoriale anche se non nella forma tradizionale dell’autore-persona. La questione dell’agency dei sistemi di Intelligenza Artificiale e dei connessi, presunti, diritti di proprietà intellettuale è abbastanza ampiamente trattata soprattutto in ambito giuridico e informatico. La bibliografia sul tema non è soffocante ma nemmeno esigua. La prima attestazione di una presa di posizione formale è del 1975 con il «Final Report of the National Commission on New Technology Uses of Copyrighted Works», statunitense, al cap. 3, sezione «Computers and Copyright. New Works» dove si legge:

There is no reasonable basis for considering that a computer in any way contributes authorship to a work produced through its use. The computer, like a camera or a typewriter, is an inert instrument. ... To be used in the creation of a work, a computer must be controlled by a program and must ordinarily utilize data input from other sources. ... It appears to the Commission that authorship of the program or of the input data is entirely separate from authorship of the final work.⁶²

Quindi nessuna autorialità per un’opera prodotta utilizzando un computer – o per meglio dire ‘utilizzando un software che gira su un computer’. Le tematiche del report del 1975 vennero rie-

⁶⁰ (European Parliament. Directorate General for Internal Policies of the Union. 2020).

⁶¹ In (Floridi, Taddeo, e Turilli 2009) si rende conto dell’edizione [deo](#) 2008 del concorso Loebner in cui si mette in atto il test di Turing per vedere se un software di Intelligenza Artificiale lo supererà.

⁶² («Final Report of the National Commission on New Technology Uses of Copyrighted Works, Chapter 3 – Computers and Copyright. New Works» 1975, 43–46)

laborate 10 anni dopo nei lavori dell'Office of Technology Assessment⁶⁵ statunitense che su questi temi scrisse:

It is misleading to think of programs as inert tools of creation, in the sense that cameras, typewriters, or any other tools of creation are inert. ... One must ask whether machines or interactions with machines might produce a pattern of output that would be considered creative or original if done by a human being⁶⁴. If machines are in any sense co-creators, the rights of programmers and users of programs may not be easily determined within the present copyright system.

Se le macchine sono co-creatrici (co-autrici), *e funzionano per mezzo di software*, nel quadro del copyright del tempo non era semplice definire diritti dei programmatori e diritti degli utenti in riferimento ad uno specifico output. Queste conclusioni vennero sottoposte a minuziosa disamina giuridica sempre nel 1985 da Pamela Samuelson (Samuelson 1985) che giunse ad una conclusione paradossale e provocatoria, «let the raw output be in the public domain»:

perhaps it is best not to give anyone property rights in whatever raw text some fertile computer has generated. Let the raw output be in the public domain, just as a found object would be. ... One reason not to adopt this seemingly sensible proposal is that it conflicts with the temper of the times. (Samuelson 1985, 1224)

Samuelson è ritornata sull'argomento nel 2020 (Samuelson 2020) in relazione a recenti sviluppi dei prodotti di sistemi di Intelligenza Artificiale, come ad esempio il ritratto di gentiluomo creato nel progetto «The Next Rembrandt»⁶⁵ richiamando da un lato la posizione di principio dello U.S. Copyright Office: «The Office will register an original work of authorship, provided that the work was created by a human being»⁶⁶ dall'altro ricordando che invece nel Regno Unito la normativa concede 50 anni di protezione all'opera «computer-generated» riconoscendo come autore la persona che ha utilizzato il computer⁶⁷. Di fronte alla varietà di posizioni emerse da studiosi e da sentenze statunitensi Samuelson conclude pragmaticamente che se gli output sono generati grazie ad un'attività dell'utente allora è l'utente che ha i diritti sugli output - pensando al libro Springer, in questa prospettiva i diritti sul libro sono del gruppo di persone che hanno configurato e fatto funzionare il software, e che hanno messo a punto i contenuti; non di Beta Writer. E dunque queste persone sono da riconoscere come autori del libro. Schönberger quando parla della produzione del libro Springer e citando Goodwin afferma «diventeremo scrittori di scrittori», cioè manovreremo sistemi di scrittura che opereranno per noi secondo le nostre indicazioni come un pianista opera attraverso un pianoforte⁶⁸, concorda con Samuelson sul fatto che l'agency è nel soggetto umano che configura e utilizza il sistema di Intelligenza Artificiale.

⁶⁵ (Office of Technology Assessment 1986).

⁶⁴ Ritorna nel 1985 il concetto che era stato formulato nel «Proposal» del 1955, cfr. nota 56.

⁶⁵ <https://www.nextrembrandt.com/>

⁶⁶ (U.S. Copyright Office 2021, par. 306).

⁶⁷ (*Copyright, Designs and Patents Act* 1988, par. 9.3; 12.3): «In the case of a literary, dramatic, musical or artistic work which is computer-generated, the author shall be taken to be the person by whom the arrangements necessary for the creation of the work are undertaken.»; «If the work is computer-generated copyright expires at the end of the period of 50 years». (Bently 2018) sostiene che queste norme non sono adeguate/efficaci.

⁶⁸ La metafora pianista/pianoforte è di Schönberger.

In anni vicini a noi, nel 2019, in «A unified framework of five principles for AI in society», Floridi e Cowls propongono una prospettiva differente sull'agency dei sistemi di Intelligenza Artificiale: un'agency regolabile in ampiezza di azione, e reversibile, che chiamano meta-autonomy:

The autonomy of humans should be promoted and that the autonomy of machines should be restricted and made intrinsically reversible. ... This introduces a notion we might call 'meta- autonomy,' or a 'decide-to-delegate' model: humans should retain the power to decide which decisions to take: exercising the freedom to choose where necessary, and ceding it in cases where overriding reasons, such as efficacy, may outweigh the loss of control over decision-making. Any delegation should also remain overridable in principle (i.e., deciding to decide again).⁶⁹ (Floridi and Cowls 2019:7)

Quindi Samuelson e i documenti governativi americani da una parte, quelli inglesi dall'altra, pur senza mai menzionarla esprimono posizioni sull'agency dei sistemi di Intelligenza Artificiale, gli uni negandola gli altri affermandola. Floridi e Cowls adottano una prospettiva fluida in cui l'agency è variabile. Tutte le posizioni vengono espresse in prospettiva giuridica, o filosofica, non bibliografica: «chi e quando ha diritto di pretendere un pagamento per dei diritti», e non «come catalogare un'opera» ma le prospettive ovviamente sono connesse.

Nel sito della WIPO, World Intellectual Property Organization, ci sono contenuti interessanti per questo tema (piani strategici nazionali per l'Intelligenza Artificiale⁷⁰) ma non posizioni ufficiali sulla relazione tra diritti di proprietà intellettuale e Intelligenza Artificiale. Un'indagine nei siti di IFLA, ALA, AIB, AIB Studi, JLIS, Biblioteche Oggi, rileva che la discussione sul software di Intelligenza Artificiale riguarda *principalmente* i servizi della biblioteca, o il materiale che le biblioteche mettono a disposizione per il *machine learning*, mentre il tema più astratto dell'autorialità dei sistemi di Intelligenza Artificiale (se i sistemi di Intelligenza Artificiale possano o no essere concepiti come autori, a quali condizioni, con quali limitazioni) e della proprietà intellettuale in relazione ai sistemi di Intelligenza Artificiale è assente. In altre parole si coglie dei sistemi di Intelligenza Artificiale la componente operativa, pragmatica, e non sui principi di esistenza di tali sistemi. Sono due campi e due ambiti differenti come di fronte all'automobile: da un lato si può ragionare sulle possibilità che essa offre di spostarsi per conoscere il mondo, e dall'altro si può ragionare sulla possibile evoluzione e diffusione del motore elettrico.

Conclusioni

Due linee differenti e complementari di riflessioni si possono formulare al termine di questo lungo percorso. La prima prende avvio dalla relazione tra sistemi di Intelligenza Artificiale e produzione di testo e porta a ripensare al test di Turing e alla sua struttura centrata su testi scritti (risposte) generati in seguito a domande di un interrogante umano a due soggetti di cui non si sa se/chi sia umano. Il test concepisce che un sistema software-hardware *al massimo livello delle sue capacità*

⁶⁹ (Floridi e Cowls 2019, 7).

⁷⁰ Nel documento del MISE «Proposte per una Strategia italiana per l'intelligenza artificiale» (Gruppo di Esperti MISE sull'intelligenza artificiale 2020) il tema della proprietà intellettuale dei prodotti di sistemi di Intelligenza Artificiale non è presente.

debba risultare indistinguibile da un essere umano e che il banco di prova sia la produzione di testo. La formulazione del test di Turing è del 1950⁷¹ e di area inglese; la formulazione del quasi-coevo progetto di Intelligenza Artificiale di Dartmouth è del 1955⁷² e di area americana; in entrambi si nota che il tema è l'imitazione: *imitation game* per Turing; *behaving like a human* per il progetto americano. Oggi, a cinquant'anni e più di distanza è cambiata la tecnologia ma apparentemente non la concezione focalizzata sull'imitazione/riproduzione dell'umano. Infatti per promuovere nella società i software di Intelligenza Artificiale se ne mostrano le capacità (e si fanno intuire le possibilità) facendone dei ...falsari: si vedano «The Next Rembrandt»⁷³ oppure i numerosi esempi di testi «alla maniera di»⁷⁴. Cioè si mostra la qualità del sistema di Intelligenza Artificiale mostrando che ciò che produce è indistinguibile da ciò che avrebbe prodotto un corrispondente soggetto umano, dipinge come Rembrandt ma non è Rembrandt (nel mondo di ieri questa sarebbe stata l'abilità di un falsario, e nessun altro avrebbe avuto motivo di usarla). Ma questi sono gli usi per lo show, per questo abbiamo scritto «apparentemente non è cambiata la concezione». Nella realtà quotidiana gli usi dei sistemi di Intelligenza Artificiale nel funzionamento della società, lo abbiamo già accennato, ci sono già e anche se non mancano gli aspetti positivi, frequentemente i sistemi di Intelligenza Artificiale aggravano le disuguaglianze e rafforzano i poteri fuori del controllo democratico dei cittadini (Brayne 2020)⁷⁵. Basti vedere ad esempio che l'European Data Protection Board e l'European Data Protection Supervisor

call for a general ban on any use of AI for automated recognition of human features in publicly accessible spaces, such as recognition of faces, gait, fingerprints, DNA, voice, keystrokes and other biometric or behavioural signals, in any context.⁷⁶

E in prospettiva molto più ampia la Commissione Europea sta lavorando su un «Artificial Intelligence Act» (AIA)⁷⁷ la cui bozza è stata resa nota il 21 Aprile 2021 e che ha come scopo centrale di regolare la produzione e l'utilizzo dei sistemi di Intelligenza Artificiale proprio perché ne stanno

⁷¹ (Turing 1950).

⁷² (McCarthy et al. 1955).

⁷³ <https://www.nextrembrandt.com/>

⁷⁴ Un buon punto di partenza è (Branwen 2020).

⁷⁵ Non mancano usi positivi dei sistemi di Intelligenza Artificiale, ad esempio alle azioni di contrasto al Covid-19 (dalla diagnosi, all'evoluzione delle terapie, al supporto alle attività del personale medico, eccetera (Arora, Banerjee, e Narasu 2020)). Peraltro il terreno sui cui ad esempio in Google-Alphabet è in corso uno scontro durissimo fra ricercatori e azienda è proprio quello dell'influenza indesiderata dei sistemi di Intelligenza Artificiale sul funzionamento della società. In Google-Alphabet il 2 dicembre 2020 è stata licenziata Timnit Gebru, che faceva parte del gruppo di lavoro sull'ethical AI, per la quale l'evento critico è stata la pubblicazione di un articolo in cui presentava rilievi critici sul costo energetico dei grandi modelli linguistici utilizzati in Google per lo sviluppo dei sistemi di Intelligenza Artificiale (Bender et al. 2021) (Hao 2020), preceduto di poco da un altro articolo in cui esprimeva preoccupazioni etiche sullo sviluppo di sistemi commerciali di riconoscimento facciale che non sono ugualmente accurati per tutti i gruppi etnici (Raji et al. 2020) (in nota 2 abbiamo già menzionato, in una differente prospettiva, la vicenda che ha visto coinvolta Gebru). Poi, il 19 febbraio 2021, è stata licenziata Margaret Mitchell, la responsabile del medesimo gruppo di lavoro, che ha dichiarato «I tried to use my position to raise concerns to Google about race and gender inequity, and to speak up about Google's deeply problematic firing of Dr. Gebru» (Dickey 2021).

⁷⁶ (EDPB 2021).

⁷⁷ (Proposta di Regolamento del Parlamento Europeo e del Consiglio che stabilisce regole armonizzate sull'Intelligenza Artificiale (legge sull'Intelligenza Artificiale) e modifica alcuni atti legislativi dell'Unione 2021)

emergendo in evidenza gli aspetti problematici. Su questa proposta ha scritto un articolo di commento Luciano Floridi che osserva:

[The AIA] is a good starting point to ensure that the development of AI in the EU is ethically sound, legally acceptable, socially equitable, and environmentally sustainable, with a vision of AI that seeks to support the economy, society, and the environment. (Floridi 2021, 216)

Il richiamo alla società, e al 'marketing sociale' dei sistemi di Intelligenza Artificiale, implica la presenza dei cittadini come destinatari del messaggio, perché la società, per la riflessione che qui sviluppiamo, è fatta di cittadini che leggono lo stato del mondo e deliberano su di esso. Leggere lo stato del mondo e/per poi deliberare su di esso richiede un lavoro di interpretazione che è possibile solo se i cittadini sono *information literate*⁷⁸, proprio come un secolo fa le questioni connesse con il diritto di voto intersecavano quelle dell'alfabetizzazione. Leggere e interpretare lo stato del mondo è un'attività che gli esseri umani svolgono da tempi remoti⁷⁹. Oggi è più difficile perché da un lato pensiamo di essere ormai esperti, dall'altro la complessità cresce e i layer da riconoscere e decodificare aumentano. La metafora del «leggere il libro del mondo»⁸⁰ è vitale, e leggere un libro sulle batterie al litio lo conferma.

La seconda linea di riflessione è più focalizzata su aspetti bibliografici e vede l'«evento Beta Writer» del 2019 come uno spartiacque tra un prima e un dopo. Un *prima* in cui si concepiva in maniera indiscussa l'autore come persona, con tutto ciò che questo comporta dal punto di vista catalografico e della riflessione bibliografica, e un *dopo* in cui abbiamo preso atto che l'autore può essere una costellazione mista di persone e software e computer in un costante circolo di retroazione come aveva già scritto Licklider nel 1960⁸¹ e dunque la tranquilla certezza che un autore è una persona è stata scossa. Questo «evento Beta Writer» era stato preannunciato quando Barthes e Foucault nel 1967-1969 avevano annunciato la morte dell'autore (che Barthes ritrovava già nell'antichità): essi ne parlavano, ed erano stati recepiti, in ambito letterario, critico-letterario, ma stavano fornendo strumenti concettuali per analizzare l'«evento Beta Writer». L'autore-costellazione è radicalmente differente dall'autore-persona e questo rende necessario ripensare quel fondamento catalografico che è l'autore, riconoscendo in esso la *possibilità* di complessità e profondità nuove. Se si accoglie la prospettiva critica della linea di pensiero Barthes-Foucault queste complessità e profondità riguardano ogni autore-persona; se invece non si accoglie la prospettiva critica globale di cui essi sono portatori, rimane pur vero che essi forniscono strumenti di analisi per evidenziare l'esistenza dell'autore-costellazione. Non la *morte* dell'autore ma la *nascita* di un nuovo autore, nascita che è proprio un venire-alla-luce attraverso attività che rendono pubblicamente visibile qualcosa che prima era sotteso, non visibile, nascosto. Oggi si può pensare che questi siano ragionamenti su

⁷⁸ (Lana 2020).

⁷⁹ Da quando eravamo cacciatori/raccoglitori, come argomenta (Ginzburg 1979).

⁸⁰ (Blumenberg 2009).

⁸¹ Occorre rileggere ciò che Licklider scrisse nel 1960 a proposito della simbiosi uomo-computer (Licklider 1960a) perché aiuta a liberarsi da quel 'presentismo' che porta a pensare che nel mondo della tecnologia digitale ogni istante del presente porti con sé una radicale novità che rende obsoleta e dunque trascurabile la riflessione teorica, culturale, costruita in precedenza. O che forse rende(rebbe) addirittura inutile costruire una riflessione teorico/culturale dal momento che essa sarebbe obsoleta nel momento stesso in cui la si formula.

eventi-limite, questioni che si pongono in maniera eccezionale su eventi che accadono raramente e che non toccano l'ordinario corso del mondo bibliografico. Noi crediamo invece che questi eventi-limite negli anni a venire diventeranno frequenti e poi ordinari, e che occorra dunque predisporre gli strumenti concettuali per gestirli. Se ogni sistema di catalogazione è teso a mappare e categorizzare lo stato del mondo, forti cambiamenti nello stato del mondo chiedono una risposta dai sistemi di catalogazione. E quello della diffusione dei software di Intelligenza Artificiale per la produzione di testo è uno di questi forti cambiamenti nello stato del mondo.

Riferimenti bibliografici

- Adadi, Amina, e Mohammed Berrada. 2018. «Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)». *IEEE Access* 6: 52138–60. <https://doi.org/10.1109/ACCESS.2018.2870052>.
- Ahearn, Laura M. 1999. «Agency». *Journal of Linguistic Anthropology* 9 (1/2): 12–15. <https://doi.org/10.1525/jlin.1999.9.1-2.12>.
- Arora, Neelima, Amit K Banerjee, e Mangamoori L Narasu. 2020. «The Role of Artificial Intelligence in Tackling COVID-19». *Future Virology* 15 (11): 717–24. <https://doi.org/10.2217/fvl-2020-0130>.
- Barredo Arrieta, Alejandro, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador Garcia, et al. 2020. «Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI». *Information Fusion* 58 (giugno): 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>.
- Barthes, Roland. 1967. «The death of the author». *Aspen*, 1967. <https://www.ubu.com/aspen/aspen5and6/threeEssays.html#barthes>.
- . 1968. «La mort de l’auteur». *Manteia*, 1968.
- . 1970. «L’ancienne rhétorique. Aide-mémoire». *Communications* 16 (1): 172–223. <https://doi.org/10.3406/comm.1970.1236>.
- Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, e Shmargaret Shmitchell. 2021. «On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?». In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–23. Virtual Event Canada: ACM. <https://doi.org/10.1145/3442188.3445922>.
- Bently, Lionel. 2018. «The UK’s Provisions on Computer-Generated Works: A Solution for AI Creations?» Presentato al ECS International Conference: EU copyright, quo vadis ? From the EU copyright package to the challenges of Artificial Intelligence, Bruxelles. <https://european-copyright-society-dot-org-files-wordpress-com-2018-06-lionel-the-uk-provisions-on-computer-generated-works-pdf>.
- Beta Writer. 2019. *Lithium-Ion Batteries. A Machine-Generated Summary of Current Research*. New York, NY: Springer. <https://doi.org/10.1007/978-3-030-16800-1>.
- Blair, Ann. 2003. «Reading Strategies for Coping with Information Overload ca. 1550-1700». *Journal of the History of Ideas* 64 (1): 11. <https://doi.org/10.2307/3654293>.
- . 2010. *Too much to know: managing scholarly information before the modern age*. New Haven [Conn.]: Yale University Press.
- Blumenberg, Hans. 2009. *La leggibilità del mondo: il libro come metafora della natura*. Tradotto da Bruno Argenton. Bologna: il Mulino.
- Bonaventura da Bagnoregio. 1882. *Commentaria in quatuor libros sententiarum magistri Petri Lombardi*. Vol. I. 4 voll. ad Claras Aquas (Quaracchi, Firenze). <https://archive.org/details/operaomnia-01bona/page/n7/mode/2up?view=theater>.

Brainard, Jeffrey, e Jia You. 2018. «What a massive database of retracted papers reveals about science publishing's 'death penalty'». *Science* 25 (ottobre). <https://www.sciencemag.org/news/2018/10/what-massive-database-retracted-papers-reveals-about-science-publishing-s-death-penalty>.

Branwen, Gwern. 2020. «GPT-3 Creative Fiction». *Essays* • Gwern.net. 19 giugno 2020. <https://www.gwern.net/GPT-3>.

Brayne, Sarah. 2020. «Enter the Dragnet». *Logic Magazine*, 20 dicembre 2020. <https://logicmag.io/commons/enter-the-dragnet/>.

Carpenter, Todd A. 2021. «Where Does Enhancement End and Citation Begin?». *The Scholarly Kitchen* (blog). 6 ottobre 2021. <https://scholarlykitchen.sspnet.org/2021/10/06/semantic-enrichment-of-articles-is-different-from-citation/>.

Choudhury, Majumdar Roy, Alia Aoun, Dina Badawy, Luis Antonio de Albuquerque Bacardit, Yassine Marjane, e Adrian Wilkinson. 2021. «Final report of the Panel of Experts on Libya established pursuant to Security Council resolution 1973 (2011)». S/2021/229. United Nations Security Council.

Copyright, *Designs and Patents Act*. 1988. <https://www.legislation.gov.uk/ukpga/1988/48/enacted/data.pdf>.

Crawford, Kate. 2021. *Atlas of Ai: power, politics, and the planetary costs of artificial intelligence*. New Haven: Yale University Press.

Dickey, Megan Rose. 2021. «Google fires top AI ethics researcher Margaret Mitchell». *TechCrunch* (blog). 19 febbraio 2021. <https://techcrunch.com/2021/02/19/google-fires-top-ai-ethics-researcher-margaret-mitchell/>.

Durante, Massimo. 2019. *Potere computazionale: l'impatto delle ICT su diritto, società, sapere*. Filosofia delle scienze e dintorni 3. Milano: Meltemi.

EDPB. 2021. «EDPB & EDPS call for ban on use of AI for automated recognition of human features in publicly accessible spaces, and some other uses of AI that can lead to unfair discrimination». European Data Protection Board. giugno 2021. https://edpb.europa.eu/news/news/2021/edpb-edps-call-ban-use-ai-automated-recognition-human-features-publicly-accessible_en.

Elsevier. 2021. «Topics - ScienceDirect». Elsevier.Com. 2021. <https://www.elsevier.com/solutions/sciencedirect/topics>.

Engelbart, Douglas. 1962. «Augmenting human intellect: a conceptual framework.» AFOSR-3233. Menlo Park, CA: Stanford Research Institute. <https://apps.dtic.mil/dtic/tr/fulltext/u2/289565.pdf>.

Engelbart, Douglas, e William K. English. 1968. «A research center for augmenting human intellect». In *AFIPS Conference Proceedings of the 1968 Fall Joint Computer Conference*, 33:395–410. San Francisco. <https://web.stanford.edu/dept/SUL/library/extra4/sloan/mousesite/Archive/ResearchCenter1968/ResearchCenter1968.html>.

European Parliament. Directorate General for Internal Policies of the Union. 2020. «Artificial Intelligence and Civil Liability». Publications Office. <https://data.europa.eu/doi/10.2861/220466>.

«Final Report of the National Commission on New Technology Uses of Copyrighted Works, Chapter 3 – Computers and Copyright. New Works». 1975. National Commission on New Technology Uses. <http://digital-law-online.info/CONTU/contu17.html>.

Floridi, Luciano. 2021. «The European Legislation on AI: A Brief Analysis of Its Philosophical Approach». *Philosophy & Technology* 34 (2): 215–22. <https://doi.org/10.1007/s13347-021-00460-9>.

Floridi, Luciano, e Josh Cowls. 2019. «A Unified Framework of Five Principles for AI in Society». *Harvard Data Science Review* 1 (1). <https://doi.org/10.1162/99608f92.8cd550d1>.

Floridi, Luciano, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, Robert Madelin, Ugo Pagallo, e Francesca Rossi. 2018. «AI4People—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations». *Minds and Machines* 28 (4): 689–707. <https://doi.org/10.1007/s11023-018-9482-5>.

Floridi, Luciano, Mariarosaria Taddeo, e Matteo Turilli. 2009. «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 (1): 145–50. <https://doi.org/10.1007/s11023-008-9130-6>.

Foucault, Michel. 1969. «Qu’est-ce qu’un auteur?» *Bulletin de la Société Française de Philosophie* 63 (3): 73–104. <http://libertaire.free.fr/MFoucault319.html>.

Genette, Gérard. 2002. *Seuils*. Points Essais 474. Paris: Seuil.

Ginzburg, Carlo. 1979. «Spie. Radici di un paradigma indiziario». In *Crisi della ragione. Nuovi modelli nel rapporto tra sapere e attività umane*, a cura di Aldo Gargani, 1–30. Torino: Einaudi.

Goodwin, Ross. 2016a. «Narrated reality». Master thesis, New York: New York University.

———. 2016b. «Adventures in Narrated Reality». *Medium* (blog). 9 giugno 2016. <https://medium.com/artists-and-machine-intelligence/adventures-in-narrated-reality-6516ff395ba3>.

Goodwin, Ross, Kenric McDowell, e Hélène Planquelle. 2018. *1 the road*. Art + machines 1. Paris: Jean Boîte éditions.

GPT-3. 2020. «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>.

Gruppo di Esperti MISE sull’intelligenza artificiale. 2020. «Proposte per una Strategia italiana per l’intelligenza artificiale». Roma: MISE. https://www.mise.gov.it/images/stories/documenti/Proposte_per_una_Strategia_italiana_AI.pdf.

Gubrium, Jaber F., e James A. Holstein. 1995. «Individual agency, the ordinary, and postmodern life». *Sociological Quarterly* 36 (3): 555–70. <https://doi.org/10.1111/j.1533-8525.1995.tb00453.x>.

Guerrini, Mauro. 2020. *Dalla catalogazione alla metadattazione: tracce di un percorso*. Collana Percorsi AIB 5. Roma: Associazione italiana biblioteche.

Gunning, David, e David Aha. 2019. «DARPA’s Explainable Artificial Intelligence (XAI) Program». *AI Magazine* 40 (2): 44–58. <https://doi.org/10.1609/aimag.v40i2.2850>.

Hagan, Martin T., Howard B. Demuth, e Mark H. Beale. 1996. *Neural network design*. 1st ed. Boston: PWS Pub.

Hao, Karen. 2020. «“I Started Crying”: Inside Timnit Gebru’s Last Days at Google». *MIT Technology Review*, 16 dicembre 2020. <https://www.technologyreview.com/2020/12/16/1014634/google-ai-ethics-lead-timnit-gebru-tells-story/>.

———. 2021. «Stop Talking about AI Ethics. It’s Time to Talk about Power.» *MIT Technology Review*, 23 aprile 2021. <https://www.technologyreview.com/2021/04/23/1023549/kate-crawford-atlas-of-ai-review/>.

Haski-Leventhal, Debbie. 2020. «The Way Forward in Higher Education». In *The Purpose-Driven University*, 133–45. Bingley: Emerald Publishing Limited. <https://doi.org/10.1108/978-1-83867-283-620201011>.

ICCU. 2009. *Regole italiane di catalogazione: REICAT*. Roma: ICCU.

IFLA. 2020. «IFLA Statement on Libraries and Artificial Intelligence». <https://www.ifla.org/publications/node/93397>.

Kaindl, Markus, e Stephanie Preuss. 2020. «How can AI support the research community in times of crisis?» *The Source* (blog). 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>.

Kallenborn, Zachary. 2021. «Was a Flying Killer Robot Used in Libya? Quite Possibly». *Bulletin of the Atomic Scientists* (blog). 20 maggio 2021. <https://thebulletin.org/2021/05/was-a-flying-killer-robot-used-in-libya-quite-possibly/>.

Lana, Maurizio. 2020. *Introduzione all’information literacy*. Editrice Bibliografica. <https://www.editricebibliografica.it/scheda-libro/maurizio-lana/introduzione-allinformation-literacy-9788893570862-579327.html>.

Larson, Jeff, Mattu Surya, Lauren Kirchner, e Julia Angwin. 2016. «How We Analyzed the COMPAS Recidivism Algorithm». *ProPublica*, Maggio 2016. <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm?token=pc1OkzviktiLco0hJY5BDRPpI44HbKV>.

Licklider, J. C. R. 1960a. «Man-Computer Symbiosis». *IRE Transactions on Human Factors in Electronics* HFE-1 (1): 4–11. <https://doi.org/10.1109/THFE2.1960.4503259>.

Licklider, J.C.R. 1960b. «Man-computer symbiosis». *IRE Transactions on Human Factors in Electronics*, HFE-1, marzo, 4–19.

McCarthy, John, Marvin L. Minsky, Nathaniel Rochester, e Claude Elwood Shannon. 1955. «A proposal for the Dartmouth summer research project on Artificial Intelligence». <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>.

———. 2006. «A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence». *AI Magazine* 27 (4): 3. <https://aaai.org/ojs/index.php/aimagazine/issue/view/165>.

Mehan, Hugh, e Houston Wood. 1975. *The Reality of Ethnomethodology*. New York: John Wiley & Sons, Ltd.

Mittelstadt, Brent. 2019. «Principles alone cannot guarantee ethical AI». *Nature Machine Intelligence* 1 (11): 501–7. <https://doi.org/10.1038/s42256-019-0114-4>.

Office of Technology Assessment. 1986. «Intellectual property rights in an age of electronics and information». OTA-CIT-302. Washington, DC: U.S. Congress. https://digital.library.unt.edu/ark:/67531/metadc39611/m2/1/high_res_d/8610.pdf.

«Projects and Cooperations - Applied Computational Linguistics Lab Goethe University Frankfurt, Germany». s.d. Consultato 1 luglio 2021. <http://www.acoli.informatik.uni-frankfurt.de/projects.html>.

Proposta di Regolamento del Parlamento Europeo e del Consiglio che stabilisce regole armonizzate sull'Intelligenza Artificiale (legge sull'Intelligenza Artificiale) e modifica alcuni atti legislativi dell'Unione. 2021. <https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=CELEX:52021PC0206&from=EN>.

Raji, Inioluwa Deborah, Timnit Gebru, Margaret Mitchell, Joy Buolamwini, Joonseok Lee, e Emily Denton. 2020. «Saving Face: Investigating the Ethical Concerns of Facial Recognition Auditing». In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 145–51. New York NY USA: ACM. <https://doi.org/10.1145/3375627.3375820>.

Samuelson, Pamela. 1985. «Allocating ownership rights in computer-generated works». *University of Pittsburgh Law Review* 47: 1185–1228.

———. 2020. «AI Authorship?» *Communications of the ACM* 63 (7): 20–22. <https://doi.org/10.1145/3401718>.

Schoenenberger, Henning, Christian Chiarcos, e Niko Schenk. 2019. «Preface». In *Lithium-Ion Batteries. A Machine-Generated Summary of Current Research*. New York, NY: Springer. <https://doi.org/10.1007/978-3-030-16800-1>.

Shapiro, Susan P. 2005. «Agency theory». *Annu. Rev. Sociol.* 31: 263–84. <https://doi.org/10.1146/annurev.soc.31.041304.122159>.

Springer Nature. 2019. «Springer Nature Publishes Its First Machine-Generated Book». *Www.Springer.Com* (blog). 2 aprile 2019. <https://www.springer.com/gp/about-springer/media/press-releases/corporate/springer-nature-machine-generated-book/16590126>.

Strubell, Emma, Ananya Ganesh, e Andrew McCallum. 2019. «Energy and Policy Considerations for Deep Learning in NLP». In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3645–50. Florence, Italy: Association for Computational Linguistics. <https://doi.org/10.18653/v1/P19-1355>.

Svenonius, Elaine. 2000. *The intellectual foundation of information organization*. Digital libraries and electronic publishing. Cambridge, Mass: MIT Press.

Taddeo, Mariarosaria, e Luciano Floridi. 2018. «How AI Can Be a Force for Good». *Science* 361 (6404): 751–52. <https://doi.org/10.1126/science.aat5991>.

Turing, A. M. 1950. «Computing Machinery and Intelligence». *Mind* LIX (236): 433–60. <https://doi.org/10.1093/mind/LIX.236.433>.

United Nations, e David Kaye. 2018. «Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression». A/73/348. New York: United Nations.

U.S. Copyright Office. 2021. *Compendium of U.S. Copyright Office practices*. III. Washington, DC.

Vetere, Guido. 2021. «L'intelligenza artificiale “stona” ancora: i limiti della generazione automatica di musica». *Agenda Digitale* (blog). 20 aprile 2021. <https://www.agendadigitale.eu/cultura-digitale/lintelligenza-artificiale-stona-ancora-i-limiti-della-generazione-automatica-di-musica/>.

Vuong, Quan-Hoang. 2020. «The Limitations of Retraction Notices and the Heroic Acts of Authors Who Correct the Scholarly Record: An Analysis of Retractions of Papers Published from 1975 to 2019». *Learned Publishing* 33 (2): 119–30. <https://doi.org/10.1002/leap.1282>.

Wang, Sun-Chong. 2003. «Artificial Neural Network». In *Interdisciplinary Computing in Java Programming*, di Sun-Chong Wang, 81–100. Boston, MA: Springer US. https://doi.org/10.1007/978-1-4615-0377-4_5.

Yang, Guang-Zhong, Jim Bellingham, Pierre E. Dupont, Peer Fischer, Luciano Floridi, Robert Full, Neil Jacobstein, et al. 2018. «The Grand Challenges of Science Robotics». *Science Robotics* 3 (14): eaar7650. <https://doi.org/10.1126/scirobotics.aar7650>.