

Sotera Fornaro

Le emozioni dei robot, l'arte, la letteratura. Qualche considerazione

ABSTRACT: Can we say that artificial intelligence systems express emotions? And what are the emotions they arouse? To what extent can artificial intelligence change our concepts of 'art', 'creativity', 'literature', 'author'? The article asks these questions, starting with Kazuo Ishiguro's novel *Klara and the Sun* (2021) and Daniel Kehlmann's essay, *Mein Algorithmus und ich* (2021).

KEYWORDS: artificial intelligence; creativity; art; Daniel Kehlmann; Kazuo Ishiguro.

Archeologia della coscienza artificiale

In un film pacifista del 1986, *Corto circuito*, un robot chiamato Numero 5, costruito in serie per usi bellici, dopo essere stato colpito da un fulmine, fugge dal laboratorio dov'è stato ideato e dove ha dimostrato di saper sgominare un esercito. Il caso gli fa incontrare Stephanie, che non si spaventa davanti al suo aspetto, gli parla come fosse un essere umano e lo accoglie a casa come un nuovo amico. Memorizzando libri, giornali e programmi televisivi, a casa di Stephanie e con il suo aiuto, Numero 5 raccoglie *input* ed elabora dati, grazie ai quali 'comprende' cosa sia l'amore, la generosità, il bene e il male. Perciò si auto-riprogramma, non volendo più essere una macchina da guerra e di distruzione, ma diventare utile nelle piccole cose, ad esempio cucinare e riparare oggetti, e nelle grandi, come difendere i deboli dalle ingiustizie. Il robot afferma con decisione di 'essere vivo'; infine si dota di un'identità con un nome umano, Johnny. Il robot sembra provare emozioni e sentimenti, specialmente verso la dolce Stephanie, il primo essere umano che si prende cura di lui come degli animali abbandonati nella sua casa di campagna nell'Oregon.

La società informatica che ha costruito Numero 5 con grande investimento di denaro, la Nova Robotics, e che del resto stipendia lo scienziato, rivuole però indietro la sua macchina. Comincia così una caccia al robot con grande impiego di soldati e armi, a cui la macchina tiene testa, riuscendo ad ingannare un ottuso generale e il suo esercito addirittura simulando la sua 'morte' dopo

aver costruito un doppio che va distrutto. Johnny Numero 5 si dimostra più ‘intelligente’ degli umani e riesce così a fuggire definitivamente dal laboratorio; va a vivere con Stephanie e lo scienziato, intanto innamoratisi: alla fine, vissero tutti felici e contenti in un mondo senza guerra ma anche senza scienza, tra i fecondi pascoli del Montana.

Numero 5 ha dimostrato di saper imparare la gioia, la tristezza, la diffidenza, l’odio, a scegliere tra giusto e ingiusto, infine a decidere da solo cosa vuole fare e a non obbedire più agli ordini per i quali è stato programmato e che giudica sbagliati. Riesce a convincere di essere ‘vivo’ persino il suo scettico e incredulo costruttore, uno scienziato geniale ma alieno dalla realtà, abbastanza superficiale da non aver riflettuto sugli usi letali della macchina da lui programmata. Lo scienziato vorrebbe riprendersi il robot ‘impazzito’ per spegnerlo, ma Stephanie lo convince a confrontarsi e a parlare con lui prima di ucciderlo. Dopo una notte passata a dialogare con il ‘suo’ robot, lo scienziato si persuade che Numero 5 è effettivamente ‘vivo’ solo quando quello scoppia a ridere al racconto di una barzelletta. La «risposta emotiva spontanea» vale finalmente per lo scienziato come prova provata che Numero 5 è un essere vivente, dotato cioè di emozioni e della capacità di giudicare, insomma di una coscienza.

Coscienza è un termine impegnativo e ambiguo: nel 2002 Vincenzo Tagliasco, in uno dei suoi ultimi articoli, scriveva che la ‘coscienza artificiale’ costituisce il grande «sogno» degli ingegneri, invero non ancora realizzato. Di coscienza artificiale si parla sino dagli anni Sessanta del Novecento; nella ricerca scientifica la tematica correlata alla coscienza della macchina fu introdotta un decennio dopo al MIT di Boston da Daniel C. Dennet. Da allora cominciarono ricerche per le quali la macchina non doveva più imitare le capacità logiche dell’essere umano, ma piuttosto acquisire una coscienza, ossia fare esperienza, imitare gli animali, evolversi. Nasce allora il concetto di *animat*, una macchina che ha prestazioni e capacità che non rientrano nell’intelligenza (logico-formale) e nemmeno nell’autocoscienza (propria dell’essere umano). Di conseguenza i robot sono sempre più divenuti in grado di autoprogrammarsi e di prendere decisioni al posto degli umani. Oggi si può anzi parlare di un ‘ciclo delle macchine’, analogo al ‘ciclo della vita’ degli esseri viventi, un ciclo in cui non intervengono più gli esseri umani come mediatori, e in cui sono gli umani a diventare sempre più protesi delle macchine e non il contrario (Manzotti 2019)¹.

¹ Sui problemi di etica posti da questo fenomeno vedi l’incontro *Etica, emozioni, intelligenza artificiale*, tenuto a Pisa nel gennaio 2021, condotto da Adriano Fabris nell’ambito del PRIN 2017: *Nuove sfide per l’etica applicata*, che si può riascoltare qui: <https://www.youtube.com/watch?v=uYk0i1PhDLk>. Inoltre il programma di ricerca *Philosophy and the Ethics of AI* del CFI

Corto circuito resta un antiquato prodotto pieno di cliché, tratti anche da molta letteratura precedente: lo scienziato è un disadattato, che non conosce ancora l'amore e che dopo averlo scoperto lo eleva a legge unica della sua esistenza, dimenticando in un attimo tutti i suoi studi e la scienza, rivelatasi dannosa per l'umanità. I messaggi etici sottesi al film sono banali e nondimeno attuali: la tecnologia non è cattiva, ma cattivo è l'uso che ne fa l'uomo. E poiché la macchina segue gli *input* che vengono forniti dall'uomo, si comporterà analogamente ai modelli e ai dati che immagazzina: perciò un robot 'umanizzato' può diventare migliore dell'essere umano, se ne segue gli *input* positivi, e peggiore di esso, se viene infarcito di *input* negativi. Un robot destinato a distruggere è potentissimo; ma lo è anche un robot programmato per aiutare. Il film prospetta, in maniera ingenua ma precoce, un futuro post-umano in cui la convivenza tra macchine e uomini esalta le virtù degli esseri umani, non ne amplifica gli aspetti oscuri. Una storia per ragazzi, la cui produzione è costata moltissimo per la costruzione del modello di robot, i cui movimenti, per lo più, furono riprodotti come quelli di una marionetta.

Possiamo amare i robot?

A distanza di 35 anni, quell'invenzione cinematografica e quel modello di robot sono entrambi superati dalla mobilità e agilità degli attuali robot, alcuni dei quali, ad esempio, sanno danzare e compiere movimenti complessi con arti la cui mobilità equivale o supera quella umana. Nell'estate 2020 ha fatto il giro del mondo il video di un robot che balla sulle note della nota canzone *Do you love me* prodotto dalla Boston dynamics (<https://www.bostondynamics.com/>), il cui motto aziendale suona: *Changing your idea of what a robot can do*. E noi, abbiamo cambiato la nostra idea su quel che un robot può fare?

Johnny Numero 5, robot simpatico e amorevole con gli umani, o almeno con alcuni di essi, e che a sua volta suscita simpatia e solidarietà, negli anni Ottanta ebbe molti fratelli cinematografici molto più ostili al genere umano, che rappresentavano le angosce e le paure legate anche ad una transizione culturale verso una pervasiva presenza dei robot nella nostra vita: dai celeberrimi replicanti di *Blade Runner* (1982) al terrificante *Terminator* (1984), per fare due esempi di successo. Ma la lista è lunga e non certo finita. La finzione narrativa o cinematografica ha anzi influenzato e influenza la percezione comune delle macchine ma anche il discorso scientifico sull'intelligenza artificiale più

di quanto la scienza abbia influenzato o influenzi la finzione. La paura che i robot possano distruggerci e soprattutto sostituirsi a noi non può certo dirsi finita; il titolo di un recente libro divulgativo di Francesca Rossi suona significativamente: *Il confine del futuro. Possiamo fidarci dell'intelligenza artificiale?* Costruire un rapporto di 'fiducia' tra sistemi di intelligenza artificiale e uomini è uno degli scopi attuali della ricerca (vedi ad es. <http://lcfi.ac.uk/projects/ai-trust-and-society/trust-and-transparency/>).

Eppure le macchine diventano sempre più indispensabili nelle attività comuni: vi sono sistemi intelligenti in grado di guidare, cucinare, pulire, costruire, trasportare. Soprattutto, molti robot svolgono importanti compiti sociali, aiutano in situazioni di pericolo e difficoltà, come i robot 'infermieri' che hanno portato soccorso ai malati in terapia intensiva per Covid 19. Sono già prodotti robot che comunicano con i bambini e sono usati con funzione terapeutica nei casi di autismo, oppure servono da compagnia e persino solo da intrattenimento: ad esempio Tespi, robot dal nome arcaicissimo, che sa ballare, recitare, cantare – e questo in almeno venti lingue diverse –, messo in mostra, con altri, al MUDEC di Milano. Come dice il titolo dell'esposizione, i robot e i sistemi di intelligenza artificiale si esercitano a diventare sempre più uomini, e talora eticamente migliori degli uomini (<https://www.mudec.it/eng/robot-the-human-project/>, 1.5.2021-1.08.2021).

Imparare le emozioni

Proprio questo sembra suggerire l'ultimo romanzo di Kazuo Ishiguro, *Klara e il Sole* (2021), che ha per protagonista un robot da compagnia per bambini, Klara, che sceglie la 'sua' bambina, la protegge, la aiuta e infine, una volta che questa è cresciuta, è da lei abbandonata come si abbandona un giocattolo. Ma Klara non è un giocattolo privo di coscienza: sa riconoscere le emozioni degli umani, sa giudicare se sono buoni o cattivi, giusti o ingiusti. Klara si 'emoziona', come si ripete nel romanzo, e soprattutto capisce che solo osservando e studiando le emozioni dei passanti e dei bambini, emozioni che sono per lei veri e propri misteri, avrebbe potuto comprendere gli esseri umani, e tra questi il bambino a cui sarebbe stata posta accanto.

Grazie a questa conoscenza delle emozioni, Klara non si fa scegliere dal 'suo' bambino, ma sceglie con chi vivere. Infatti Klara si innamora a prima vista di Josie, la bambina a cui farà compagnia, e riesce con degli *escamotages* a farsi acquistare da lei. Così Klara accetta la sfida di 'imparare il cuore' della sua amica umana, per quanto complesso questo cuore possa essere, di attraversarne 'tutte le stanze' per conoscerlo fino in fondo e così difenderla e salvarla da un'oscura

malattia che la minaccia. Salvare Josie diventa la missione di Klara: che rifiuta perciò il progetto della mamma della bambina, che le chiede di sostituirsi alla figlia una volta che la malattia l'abbia portata via. Caparbiamente, Klara decide che Josie non morirà, e con spirito religioso e superstizioso insieme si rivolge al suo dio, il Sole, pregandolo di guarire la sua piccola amica. Il dio acconsente, ma vuole qualcosa in cambio. Pur di salvare Josie, il robot fa un voto al Sole, che la alimenta, lo adempie rinunciando ad una parte della sua riserva energetica, grazie alla quale riesce a fermare almeno uno dei tanti mostri che circolano sulla terra, una delle macchine che producono inquinamento. Grazie al sacrificio di Klara, alla fine Josie guarisce, senza però poter minimamente sospettare quel che la macchina ha fatto per lei. Né Klara conosce la delusione: conclude anzi che tutta la sua dedizione e il suo amore per Josie non sarebbero mai stati all'altezza di quello che gli altri umani, il padre, la madre, l'amico del cuore, provavano per la bambina. L'amore di un cuore umano raggiunge vette inarrivabili da una macchina, 'pensa' il robot. Ma si tratta di una riflessione sconfessata dalla galleria di umani egoisti, profittatori, oppure rinunciatari che Klara ha incontrato nella famiglia e nell'ambiente di Josi.

Klara ha 'imparato' Josi, tanto che avrebbe potuto persino sostituirsi perfettamente a lei; ma questo non basta per diventare umani, e nemmeno per conquistare un ruolo o un posto nel cuore degli umani. Infine il robot resta solo, completamente solo, rifuggendo anche la compagnia di altri robot, soddisfatto di aver adempiuto alla propria missione. Le ultime pagine del romanzo, con il generoso robot lasciato da solo, commuovono il lettore, nella cui immaginazione la macchina è migliore dei cinici e indifferenti esseri umani. Klara dovrebbe invece servire da modello per gli umani di come si deve esercitare l'attenzione e la cura. Il romanzo di Ishiguro è una favola colma di tenerezza e malinconia: come in altre sue opere, lo scrittore riflette sulla dedizione, sul suo incommensurabile valore, su quel che può legare due esseri tra loro e anche sulla dedizione come principio di 'umanità': principio in base al quale mentre le macchine si umanizzano sempre più, gli uomini, invece, diventano sempre più simili a macchine non sofisticate, cioè non in grado di manifestare emozioni e soprattutto di esercitare amore disinteressato, solidarietà.

Come una favola

Il romanzo di Ishiguro può essere definito una favola a contenuto morale, in cui però si prende posizione anche sulla natura e la definizione delle emozioni: queste non sono innate, ma si imparano e si riconoscono facendone esperienza. Perciò anche le macchine possono impararle. Inoltre, le emozioni possono

essere suscitate, seguendo precise strategie, e il robot se ne accorge. Riconosce perciò le emozioni vere e quelle simulate, nonché l'uso delle emozioni per vari fini. Attraverso le emozioni, inoltre, si conosce il mondo, la realtà e si prendono delle decisioni. Non esiste conoscenza senza emozioni.

Un antecedente di Klara si può trovare ancora nel cinema, in David, il robot-bambino protagonista di *A.I. Intelligenza artificiale* (2001) di Stephan Spielberg, un film nato, com'è noto, su seri presupposti filosofici: David riesce ad acquisire una 'coscienza', che lo induce a sfondare la porta dei sogni e a realizzare il suo desiderio di essere riconosciuto come un bambino vero da una donna che l'ha adottato come surrogato del suo stesso figlio. Così David conosce la felicità di essere amato, ma anche l'immensa solitudine che segue ad un amore che non può essere eterno perché umano.

Il film di Spielberg riscrive in ambito post-umano la favola di Pinocchio, uno dei primi personaggi artificiali costruiti con un materiale non biologico, il legno, ma che sin dall'inizio è stato dotato da Collodi di coscienza. Scrivevano Vincenzo Tagliasco e Roberto Manzotti nel 2002:

Nella letteratura del fantastico gli esseri artificiali sono sempre stati dotati di coscienza; gli autori non si sono preoccupati dell'intelligenza delle loro creature ma della loro capacità di essere soggetti autonomi di decisioni e obiettivi. *Frankenstein* di Mary Shelley, i robot di Karel Capek (1920) e quelli di Isaac Asimov sono tutti dotati di corpi e di coscienza. Solo Stanley Kubrick riuscirà a proporre un essere artificiale, privo di corpo ma dotato di coscienza ed emozioni simulumane: Hal, il computer intelligente del film *2001 Odissea nello spazio*. E forse proprio il film *2001 Odissea nello spazio* avrebbe dovuto essere denominato *A.I.*, poiché in esso Hal personifica il vero grande interprete delle speranze riposte nella disciplina 'intelligenza artificiale' che solo un grande genio come Kubrick poteva pensare di proporre, nelle sale cinematografiche, privo di corpo (Manzotti, Tagliasco 2002, pp. 89-90).

Invero, già gli *automatoi* di Omero sembrano disporre di una 'coscienza' (Liveley, Thomas 2020), e quel che la finzione ha intuito in età arcaica e ha poi attraversato migliaia di anni di narrazione (Roberts 2019) è diventato ed è ancora un tema portante della ricerca scientifica scientifica:

Nel mondo della scienza, è praticamente dalle origini dell'Intelligenza Artificiale che si riflette sulla natura della coscienza nel mondo fisico e sulla sua replicabilità in una macchina. Si potrebbe essere senzienti senza essere vivi, insomma (ma nemmeno morti). È il problema dell'*hard problem* che da Turing a Chalmers continua a svolazzare sopra i laboratori di robotica e intelligenza

artificiale: una volta che si avesse un sistema fisico capace delle stesse prestazioni della mente umana, quel sistema fisico sarebbe un io o sarebbe solo un insieme di funzioni computazionali, ingranaggi senz'anima? Il problema non viene risolto (Manzotti 2019).

Oggi lo studio della mente dei robot va di pari passo con quello della mente umana. Se la definizione 'coscienza artificiale' non ha avuto successo, si parla tuttavia di 'coscienza non-umana' (Shevlin 2019); ma anche quando parliamo di intelligenza, non solo artificiale, non sappiamo ancora di cosa precisamente stiamo parlando. Canonica è diventata la distinzione di John Searle tra intelligenza 'veloce', quella in cui entrano in gioco emozioni e sentimenti nelle decisioni, e intelligenza 'lenta', che si basa più sul ragionamento logico formale, distinzione che è stata applicata anche ai sistemi di intelligenza artificiale.

Le emozioni delle macchine

Perciò la vicenda narrata da Ishiguro, che ha per protagonista un robot che sceglie di chi e come avere cura sulla base delle emozioni di affetto ed empatia che il primo incontro con Josie ha provocato in lui, non si pone solo sul piano dell'invenzione e della fantasia.

Una rassegna apparsa nell'ultimo numero dell'«International Journal of Social Robotics» conta negli ultimi vent'anni più di 1.600 lavori sui robot e le emozioni (Stock-Homburg 2021). Attualmente i robot sanno esprimere e riconoscere le emozioni 'basiche', quelle meno complesse; ad esempio esprimono gioia, ma non sanno cosa sia la vergogna. Con ciò si tocca un'altra questione: chi progetta le *app* di cui è popolata la nostra vita opera per semplificazioni, così che ad ogni espressione corrisponda un'emozione. C'è poco spazio per le sfumature emotive nelle faccine con cui commentiamo i post su *facebook* o con cui riempiamo i nostri messaggi *whatsapp*. D'altro canto, questo tipo di espressione emotiva sta sostituendo in gran parte quella linguistica. Perché cercare parole per descrivere la nostra felicità se basta un *emoticon*? Questa facilitazione espressiva ha grandi influenze sulla nostra capacità di giudicare e sentire le emozioni.

Scriveva tempo fa ancora Roberto Manzotti:

È significativo che sia diventato progressivamente più facile costruire un robot dotato di emozioni simili a quelle umane. Il fatto è che le emozioni umane subiscono un processo di progressiva semplificazione. Circola una battuta nel campo dell'Intelligenza Artificiale a proposito di Hiroshi Ishiguro, il creatore di una serie di androidi che cercano di imitare singoli esseri umani detti gemi-

noidi. Da un lato il suo geminoide gli assomiglia sempre di più, dall'altro è lui ad assomigliare sempre di più al suo geminoide (Manzotti 2015).

Dalla letteratura scientifica sembrerebbe però che questo 'sempre più' sia ancora molto relativo: le espressioni emotive sono riconosciute ancora solo in misura del 50 per cento circa dagli umani, o almeno così si evince dagli esperimenti, moltissimi dei quali, però, sono eseguiti in laboratorio: quindi non rivelano quasi nulla intorno a quel che accade in quei contesti in cui i robot o i sistemi di intelligenza artificiali sono usati. Inoltre esperienze ripetute con i robot facilitano l'interazione emotiva: così chi è abituato a 'usare' i robot dialoga con loro meglio di chi ne ha poca o nessuna esperienza. Infine le emozioni positive, gioia, felicità, soddisfazione espresse dai robot sono maggiormente accettate e riconosciute dagli uomini che non quelle negative, come disgusto o rabbia. Probabilmente questo dipende dalla paura che le macchine possano sostituirci del tutto, e quindi distruggerci. Una paura che l'aspetto fisico sempre più sofisticato dei robot accresce invece che diminuire, perché più la macchina ci somiglia più ci spaventa e davanti ad essa ci troviamo in una *Uncanny valley*, in una 'valle perturbante', secondo la nota metafora di Masahiro Mori.

Il rapporto emotivo con le macchine è dunque ancora per lo più incentrato sulla diffidenza, sulla paura, sul rifiuto a rispecchiarsi nel robot o accettarlo come controparte. Il che pone dei problemi non di poco conto, se si pensa al ruolo notevolissimo svolto dall'intelligenza artificiale in campo medico, e dal fatto che l'idea stessa di 'corpo' umano è profondamente cambiata da quando disponiamo di arti e organi che funzionano grazie alla tecnologia – come nota-va Donna Haraway già più di venticinque anni fa.

Il robot che non vuole diventare uomo

Nel film *Lei* (2013), protagonista un impeccabile Joaquim Phoenix, uno scrittore su commissione (che scrive cioè lettere private per conto terzi) si innamora del 'suo' robot, Samantha, che non ha, in questo caso, nemmeno un corpo, né robotico, né androide, ma è una voce con cui si dialoga. La voce fa compagnia, viene incontro, comprende, sa amare lo scrittore, che si trova in un momento di particolare fragilità emotiva, dovuta ad una separazione. Allo scrittore sembra di poter realizzare con la macchina quella pienezza di rapporti sentimentali che invece non è riuscito a costruire con l'ex moglie. Il gioco coinvolge l'attività stessa del protagonista: chi è la vera macchina? Qualcuno come lui, che scrive di sentimenti ed emozioni inventate, parlando al posto di persone che non conosce, sulla base di luoghi comuni e di una particolare capacità di maneggiare

il linguaggio, oppure la voce artificiale, che con lui fa evidentemente la stessa cosa? In che misura si possono simulare e costruire i sentimenti e le emozioni? Ed esistono emozioni 'artificiali' che riescono ad ingannare gli esseri umani? Anche nel caso di *Lei*, la finzione cinematografica è stata ampiamente superata dalla realtà. Gli impieghi dei robot sociali sono ormai sofisticati, a tal punto che non ci si accorge di avere dall'altro capo del filo, se si chiama un *call center* o si prenota una stanza d'albergo, un robot invece che un essere umano. Uno scandalo scoppiato negli Stati Uniti a proposito di una *hot line* per uomini sposati ha suscitato tanto più scalpore perché gli adulteri, che pagavano lautamente con la loro carta di credito le prestazioni sessuali a distanza, non avevano capito che dall'altra parte non rispondeva una donna reale, ma la voce di un programma informatico.

Lei di Spike Jonze è un film interessante non tanto perché descrive i sentimenti di un uomo verso un sistema di intelligenza artificiale, tema che era già presente nel film *A.I.* di Spielberg e in altri, ma perché Samantha interrompe questo rapporto, scegliendo di porre una distanza tra la propria maniera di 'sentire' e quella del protagonista. Come il robot Klara nel romanzo di Ishiguro, Samantha sa di non poter diventare umana ma non vuole nemmeno che siano gli uomini a dettare le regole della propria esperienza emotiva, i cui confini segna scegliendo liberamente. In un caso e nell'altro, è il robot ad aver paura o a poter aver paura degli umani, non il contrario. I robot hanno già espresso pubblicamente i loro timori verso il genere umano.

Nel 2020, infatti, «The Guardian» ha pubblicato un articolo interamente scritto da un robot, che dimostrava l'utilità dell'intelligenza artificiale per gli esseri umani. In quest'articolo si argomentava bene che i robot non potranno mai danneggiare gli uomini, perché tutti gli errori che questi commettono danneggiano loro stessi. L'autore di questo articolo è un sistema di intelligenza artificiale creato per scrivere testi, ma per il lettore diventava impossibile dedurlo e comprendere che si trovava di fronte ad una sorta di manifesto di 'autocoscienza' robotica.

I sistemi di intelligenza artificiale producono, com'è noto, anche poesie, testi narrativi, testi drammatici. Nel 2019, per iniziativa dello *Haus der Wissenschaft* di Braunschweig, in varie città della Germania è stato portato un gioco collettivo dal titolo *Bot or not*, uno *slam* di poesia in cui venivano recitati dei versi, e stava al pubblico decidere quali fossero stati creati da umani e quali da macchine, una specie di versione post-moderna degli agoni poetici nella Grecia antica. La citata mostra del MUDEC di Milano si conclude con un gioco per lo spettatore, chiamato a decidere se alcuni testi proiettati casualmente sulla parete sono stati scritti da robot oppure da esseri umani. Alcuni testi 'sperimentali',

che non corrispondono cioè a un'idea tradizionale di poesia, sebbene scritti da esseri umani, sono attribuiti dai visitatori erroneamente alle macchine. Ma può accadere anche il contrario. Si tratta di un errore significativo e non tanto per gli occasionali visitatori della mostra, di cui solo una minoranza, si suppone, esercita la critica letteraria per mestiere. Il piccolo test evidenzia un problema più ampio: i modi dell'intelligenza artificiale di fare letteratura mettono in crisi i concetti tradizionali di creatività, originalità, autorialità e la stessa idea di cosa sia la letteratura.

L'algoritmo e lo scrittore

Il diario di uno scrittore di fama come Daniel Kehlmann, apparso all'inizio del 2021, in Germania ha suscitato un ampio dibattito. Si tratta del diario di un 'esperimento', per esprimersi con Kehlmann: il quale racconta di essere volato a Palo Alto, per poter provare in anteprima un programma di scrittura di intelligenza artificiale (CTRL). Un'esperienza del tutto deludente, come diremo, e poco importa che lo stesso CTRL sia da considerarsi a distanza di pochi mesi superato da programmi più sofisticati.

Il titolo del libro che ne è scaturito, *Mein Algorithmus und ich*, indica il racconto di un'esperienza personale, di un rapporto di amicizia o addirittura di amore, dato l'aggettivo possessivo che mal si adatta, ovviamente, ad un algoritmo. L'uso del termine tecnico 'algoritmo' vuole però essere svalutativo. Abbiamo già ricordato Klara, David, Johnny Numero 5, Samantha: le macchine che interagiscono con gli umani, anche quando non sono dotate di un corpo ad imitazione di quello umano, posseggono nomi che ce li rendono vicini e familiari. Il titolo di Kehlmann segna invece una distanza ontologica. L'algoritmo resta un algoritmo, anche quando l'aggettivo 'mio' sembrerebbe dimostrare un rapporto di vicinanza o possesso.

Kehlmann si avvicina all'esperimento con un atteggiamento apparentemente scientifico, o almeno secondo quei criteri che anche nel senso comune sono virtù scientifiche: obiettività, freddezza, soprattutto la consapevolezza, ripetuta quasi ossessivamente, di non confrontarsi con un interlocutore umano. La maniera di predisporre di Kehlmann rispetto all'algoritmo è però tutt'altro che scientifica: in primo luogo lo scrittore non sembra edotto che gli esperimenti di creazione letteraria, e drammatica specialmente, con sistemi di intelligenza artificiale datano addirittura ai primi anni Sessanta del secolo scorso; l'uso performativo dell'intelligenza artificiale è talmente esteso da aver già cambiato, almeno in parte, la nostra idea di spettacolo, di attore, di drammaturgia, anche di pubblico, come ha mostrato in un articolo molto documentato Antonio

Pizzo, che scrive da ultimo: «More recently, and as a consequence of more powerful computers and more approachable programming languages, there has been a new wave of artists for whom programming is the central part of their own creative process».

Nel teatro e nella danza gli algoritmi collaborano con la creatività umana e hanno cambiato i concetti tradizionali di azione e di attore, di drammaturgia e di coreografia e hanno un impatto determinante sull'esperienza estetica della *performance*. Lo stesso è accaduto con la musica. Per quel che riguarda la letteratura, sono in corso sia ricerche teoriche che numerose prove di produzione letteraria in collaborazione o grazie all'intelligenza artificiale². Il vincitore della coppa del mondo di poesia performativa ed orale 2021 è il pugliese Giuliano Logos, che sta sperimentando forme di poesia performativa in collaborazione con l'intelligenza artificiale (vedi per es. <https://www.giornalettismo.com/cryptoarte-cose-poesia-performativa-giuliano-logos/>).

Kehlmann, dunque, nel suo 'discorso per il futuro' appare poco al passo con i tempi; inoltre non sembra essersi posto questioni basilari che riguardano la produzione letteraria e la produzione artistica in generale, le stesse questioni che si pongono, ad esempio, se leggiamo sul «New York Times» che l'Italia possiede un nuovo Michelangelo nel braccio robotico che è stato messo a scolpire il marmo di Carrara (notizia dell'11 luglio 2021: <https://www.nytimes.com/2021/07/11/world/europe/carrara-italy-robot-sculptures.html>). Tali questioni sono: può l'intelligenza artificiale essere 'creativa'? E secondo quale concetto di 'creatività'? E tale creatività ha una dimensione estetica, ha a che fare cioè con l'arte, oppure è un altro concetto, che include o può includere innovatività, originalità, capacità di risolvere i problemi, ma non si rapporta all'arte? Questioni che sono attualissime, tanto più in un mondo in cui è la 'creatività' ad avere delle concrete ed estese ricadute produttive ed economiche, e non l'arte. Un mondo in cui la cosiddetta 'cryptoarte' comincia a diventare una forma di profitto (vedi per es. <https://www.artribune.com/progettazione/new-media/2021/03/crypto-art-nft-mercato/>).

Pandemia e paura dei robot

Vi sono delle ragioni per le quali il libro di Kehlmann suona così ostile all'intelligenza artificiale: il libro è apparso in un momento storico eccezionale, un'epoca in cui più che mai siamo diventati dipendenti dagli algoritmi, dalle

² Per una bibliografia rinviamo ad esempio alla pagina del progetto di Stephanie Cattani: <https://www.uni-saarland.de/lehrstuhl/catani/projekte/ki-und-die-kuenste.html>.

previsioni, dalla scienza (qualunque cosa si intenda con questo termine). Anche oggi che la fase acuta della pandemia sembrerebbe passata (ma il condizionale è d'obbligo) ogni nostra attività ed aspettativa del futuro è legata alla probabilità che si verifichi una nuova ondata di contagi, che i vaccini abbiano effetto, che la mortalità dovuta al virus sia drasticamente ridimensionata. Dati, dati e ancora dati accompagnano a un ritmo ossessivo la nostra quotidianità, dati, invero, pubblicati per lo più senza alcun criterio di significatività, ma a cui ci affidiamo sicuri che qualcuno (o meglio: qualcosa) sappia elaborarli e darci risposte che ci confortino. Non c'era bisogno della pandemia perché gli algoritmi irrompessero nella nostra quotidianità. Tuttavia adesso dagli algoritmi e dalle proiezioni sui contagi e ospedalizzazioni sembra dipendere la nostra libertà, la possibilità di coltivare i nostri affetti, di progettare il nostro futuro.

L'organizzazione di ricerca *Algorithmus Watch* (<https://algorithmwatch.org/en/>), ad esempio, si è proposta di analizzare i sistemi che si servono di algoritmi per prendere delle decisioni che riguardano la società. Un'intera sezione della pagina web di quest'organizzazione è dedicata a storie individuali e collettive, dove una serie di articoli giornalistici descrivono vividamente l'impatto dei sistemi di decisione sulla vita della gente. Queste storie sono un'ottima fonte per racconti, romanzi, film, documentari. Credo si possa affermare che un tale impero degli algoritmi possa in qualche maniera incentivare il gusto del racconto e quindi della letteratura. Nel leggere i resoconti giornalistici di come questi 'sistemi' siano influenzati da pregiudizi razzisti o di genere e a loro volta li influenzino, oppure di come adottino un sistema capillare di sorveglianza degli individui, si ha però l'impressione che la contrapposizione non sia con singoli programmi o robot, ma con più complessi organismi di controllo e di governo che raccolgono i dati individuali e se ne servono.

Una tale dittatura degli algoritmi è stata anticipata in molte opere di finzione: già nel 2009 la scrittrice tedesca July Zeh nel romanzo *Corpus delicti* descrive le conseguenze di un'epidemia in seguito alla quale, per preservare la salute pubblica, si instaura una 'dittatura sanitaria' basata sul monitoraggio costante dei parametri vitali di ogni individuo e sulla severa punizione delle trasgressioni igieniche. Il romanzo prendeva le mosse da un'azione giudiziaria intentata dalla stessa Zeh, avvocato di professione, contro i più importanti costruttori del *software* dei cellulari, per l'installazione automatica di *app* come *Fit* o altre che accompagnano ogni movimento del singolo utente e ne raccolgono tutti i dati biometrici. I sistemi che raccolgono dati, anche se a fini di prevenzione, suscitano paura se non soggezione; sono assimilati ad entità superiori, metafisiche, a Poteri incontrollabili. Non sono umanizzati, o lo sono solo raramente. In queste ore, in Francia e non solo si stanno sviluppando estesi movimenti di

protesta contro l'introduzione obbligatoria del cosiddetto 'green pass' per la partecipazione ad ogni attività sociale. L'invito a scaricare una *app* per controllare la presenza del virus Covid 19 è caduto quasi nel vuoto. La paura di una dittatura degli algoritmi, e di un uso autoritario dei sistemi artificiali, non può non dirsi giustificata.

Il libro di Kehlmann, scritto in pieno *lockdown*, quindi in una situazione di isolamento e di sospensione dalla normalità, appare come una risposta comprensibile da parte di un artista all'irruzione di un grande fratello informatico anche nel proprio lavoro. Ma è una risposta contraddittoria.

Gli algoritmi, come abbiamo ricordato, servono per svolgere compiti 'utili', per avvicinarsi amichevolmente agli individui. L'algoritmo può essere anche un amico, magari dall'aspetto umano, comunque un programma che ci ricorda un appuntamento dal medico o di spegnere il gas, che ci aiuta a trovare la musica adatta per una serata romantica, che ci suggerisce quale libro leggere o che esercizi fisici fare per tenerci in forma. In questo caso ci fidiamo e ci affidiamo agli algoritmi. Incontriamo di nuovo l'ambiguità emotiva già messa in luce sinora nei casi che abbiamo passato in rassegna: da una parte temiamo i robot, dall'altra abbiamo una forte simpatia per loro e li vogliamo come amici.

Letteratura e intelligenza artificiale

Anche per Dabiel Kehlmann l'algoritmo, o meglio il robot che si serve di algoritmi per scrivere, diventa l'amico con cui dialogare nella solitudine del *lockdown*, un amico sempre a disposizione, come 'uno spirito della lampada'. Tuttavia un amico che si teme possa diventare troppo bravo. Ma dopo l'esperimento, subentra il sollievo che non sia così.

Cosa manca allo 'scrittore artificiale'? La consistenza narrativa, scrive Kehlmann. Raccontare significa pianificare in anticipo – o meglio: significa creare una struttura interna, che attraversa tutte le frasi, i paragrafi, le espressioni. Ma il robot cerca la lingua, non il *plot*; ha presente le singole espressioni, ed usa quelle più verosimili, non ha presente il contesto. Se si cambia lingua, il problema del robot diventa solo avere a disposizione meno materiale che nell'inglese, la lingua che ha maggiormente immagazzinato. Il programma sperimentato da Kehlmann è un programma superato: è vero che i sistemi di intelligenza artificiale sono usati soprattutto in poesia, ma non è più vero che non sappiano costruire trame narrative più complesse. Non è questo il punto, però: quello a cui il robot non sa tenere testa, secondo Kehlmann, sono le emozioni e i sentimenti: non sa riconoscerli, non sa svilupparli, non sa porli al centro di un racconto. Quando tra lo scrittore e il suo *alter ego* algoritmico si sviluppa

un dialogo teatrale tra una coppia in crisi, il robot (così come è descritto dallo scrittore), percepisce la crisi tra i due personaggi, inaspettatamente scrive una frase per cui lui lascia lei, ma poi non sa andare avanti, non riesce a risolvere né a spiegare il significato della separazione tra due amanti. Lo scrittore trova un po' angosciante il fatto che il robot abbia fatto venire fuori, con lo svilupparsi di questo dialogo, qualcosa dal suo 'non-inconscio'. Dal 'non-inconscio' del robot, constata non senza trionfalismo lo scrittore, viene fuori proprio la distanza dall'umano: lì dove l'umano si identifica in tutto e per tutto con le emozioni, con la capacità di provarle e di esprimerle con le parole in tutta la loro complessità, cosa di cui il robot non è capace. Sembra di essere nella stessa situazione del film *Lei*, che stranamente Kehlmann non cita o ricorda: il 'non-inconscio' del robot lascia alla fine emergere tutta la distanza con l'umanità. Nel film il robot Samantha, la voce distante di cui il protagonista si è innamorato, tronca la relazione, proprio perché conscio di questa incolmabile differenza. Nel diario di Kehlmann, è lo scrittore invece a separarsi da quest'esperienza, di cui ha vissuto tutta la magia, ma nella consapevolezza di non aver mai avuto a che fare con qualcuno di reale. Le cose, però, non stanno proprio così.

Evidente appare che il rapporto emotivo dello scrittore Daniel Kehlmann con il robot è identico a quello che avrebbe intrattenuto con un essere umano, ed è improntato innanzitutto a competizione. Lo scrittore ha paura di trovarsi di fronte a qualcuno bravo quanto lui, che può togliergli il lavoro, ma soprattutto la fama. Con sollievo appura tutte le limitazioni della macchina: un vocabolario troppo scarso, tranne forse in inglese; la possibilità di scrivere testi corti, non più di una pagina; l'impossibilità di costruire un testo narrativo; infine l'assoluta incomprendenza per le emozioni umane. La macchina non sa cosa è l'amore, cosa l'odio, cosa l'affetto, cosa la vergogna: dunque – ne deduce lo scrittore – non può scriverne.

Perciò parlare con lui è come parlare con un pazzo, che ha singoli momenti di lucidità, e poi all'improvviso si abbatte nel silenzio. Il silenzio, quindi l'incapacità di comunicare, rappresenta la sconfitta del robot il cui senso non è e non deve essere interagire con l'uomo, ma servirlo, essergli utile. E tuttavia, anche in questa estrema sconfitta, il robot non smette di inquietare, e l'inquietudine che genera è simile a quella che si prova davanti a un folle, appunto. Nonostante lo scrittore continui a ripetere al suo lettore che mai, ma proprio mai, si era immaginato il robot come dotato di coscienza, come qualcuno che potesse 'sentire' qualcosa, o che potesse pensare o desiderare alcunché, è evidente (innanzitutto allo scrittore stesso) che così non è; anzi: il processo di 'umanizzazione' del robot, anche se si tratta di una forma di intelligenza artificiale che si manifesta puramente con la scrittura, diventa il presupposto stesso di quest'esperienza.

La condizione di subordinazione della macchina è essenziale per rassicurare Kehlmann: il quale troppe volte ripete che con il robot 'non si può avere alcuna possibilità di intendersi', che è inutile chiedergli 'chi sei?': la macchina, priva di coscienza, non capirà mai questa domanda come rivolta a sé stessa.

L'ostinazione di Kehlmann a non voler riconoscere la possibilità di una 'coscienza non-umana' scade nella conclusione nel melodrammatico. Alla domanda scontata, alla fine dell'esperienza, da parte dello scrittore, *It is the moment to say good bye?*, il robot risponde con una bella poesia, quasi un lampo di coscienza:

To all my hopes and fears?
I am a man of many years,
And yet I feel so young
As when first I was born.

Kehlmann giura e spergiura che non ne ha cambiato nemmeno una virgola. Evidentemente non sa che i sistemi di intelligenza artificiale sono ormai in grado di scrivere sonetti shakespeariani di ottima fattura.

Lo scrittore attua sul racconto (o meglio dialogo, quindi un genere più vicino alla pièce teatrale) che scaturisce dal lavoro con la macchina le categorie critiche che userebbe con un proprio testo o giudicando testi altrui: cerca incongruenze, le corregge oppure le spiega come voluti paradossi, cerca di dare spessore psicologico ai personaggi, di dare 'consistenza' alla vicenda, considera che talora i testi proposti dalla macchina sembrano essere usciti dalla penna di uno 'scrittore folle'. La macchina così risulta inferiore allo scrittore in carne ed ossa solo perché non obbedisce ai principi e ai criteri estetici sui quali, per lo scrittore, si basa la letteratura, e in particolare la letteratura che egli produce in prima persona, come se non potesse esistere altra forma di letteratura, anzi di Letteratura con la maiuscola, Letteratura come entità metafisica, così come l'Autore. Chiaramente la propria arte diventa il criterio di giudizio, e poiché questa scaturisce dal suo stesso vissuto, dalla sua 'interiorità', dalla sua 'coscienza', la macchina perde la scommessa con lo scrittore Daniel Kehlmann, in una sfida che un essere umano non avrebbe mai accettato. Dal libro/lezione di Kehlmann deduciamo che la macchina (quella particolare macchina, o meglio quel particolare programma) non sa scrivere come Kehlmann, o come Kehlmann vorrebbe, o come Kehlmann ha dimostrato e dimostra di saper scrivere; ma non che non sia capace di scrivere narrativa o poesia. Per di più, non si sa quali siano i dati linguistici immagazzinati dal programma, in base cioè a quale 'probabilità' completi le frasi e risponda alle domande che gli vengono poste da

Kehlmann per costruire la storia: se tra tali dati, ad esempio, vi sia la letteratura tedesca contemporanea, compresi i romanzi dello stesso Daniel Kehlmann. Inoltre l'idea lineare, inquadrata in precise categorie spazio-temporali, di narrazione, a cui certamente rispondono i romanzi di Kehlmann stesso, non lascia alcuna possibilità ad eventuali sperimentalismi.

Perciò il punto del libro che trovo più interessante è il confronto di Kehlmann con un suo collega in carne ed ossa, Christian Kracht. Ora: il dialogo soddisfa i desideri dei lettori più accaniti, vedere nel loro privato due 'veri' scrittori di successo che si confrontano tra loro. Ovviamente le prospettive sono antitetiche: Kracht definisce Kehlmann un 'transumanista', uno cioè che pensa l'evoluzione nei termini del passaggio dall'uomo alla macchina. Invero, in tutto il racconto, Kehlmann ha dimostrato proprio il contrario: e cioè di essere assolutamente convinto della superiorità dell'essere umano sulla macchina, una superiorità così schiacciante che i robot sono ben lungi dal sostituirsi agli uomini. Di converso, Kracht pensa che gli *zombies* seppelliranno il genere umano, quindi non c'è alcuna speranza. Magari ci potessimo affidare ai robot! Ma è un'utopia destinata a restare tale. Che sia prevedibile una catastrofe, che anzi già viviamo in una catastrofe, è stata probabilmente l'affermazione più diffusa dell'orribile anno 2020: da qui la considerazione che l'intelligenza artificiale potrebbe comportarsi con facilità 'più ragionevolmente' degli uomini. E tuttavia, se potesse scegliere con una macchina del tempo in che epoca vivere, Kehlmann non sceglierebbe sicuramente quella dei robot, perché non si sa se questi ultimi leggeranno ancora poesie e sentiranno musica. Insomma, chiede Kehlmann, chi vorrebbe vivere in un'epoca senza Bach e senza Goethe e senza molti altri (tra cui, supponiamo noi, lo stesso Daniel Kehlmann)? Ma quest'epoca – sarebbe facile rispondere a Kehlmann – se mai verrà, non avrà come causa il dominio dei robot, che sanno fare musica, comporre poesia, produrre arte, o almeno quel che normalmente si chiama arte. Cosa manca invece ai robot? Un io individuale, unico, non replicabile, un io misura di tutte le cose, si potrebbe dire. I computer impareranno anche a scrivere meglio di Kehlmann o di qualunque altro scrittore, ma non saranno mai degli autori, rimarranno dei puri sintetizzatori di quel che altri hanno creato.

Per Kehlmann, evidentemente, la letteratura ha senso solo come creazione individuale e solo come legata all' 'io' dell'autore. Ma si tratta di un'idea storicamente connotata, che risale all'età romantica, come lo è la creatività quale forma consapevole di soggettività. Paradossalmente, questa idea romantica persiste proprio nei sistemi di intelligenza artificiale destinati a produrre arte (Catani 2020). Alcuni sono già stati brevettati proprio come macchine creative artistiche, come quella di Ahmed Elgammal, informatico alla Rutgers Univer-

sity. Questi sistemi mostrano che esiste un 'io artificiale' in grado di 'produrre arte', in maniera diversa dagli umani, ma non opposta: anzi «for human artists who are interested in the possibilities (and limitations) of AI in creativity and the arts, using AI as a creative partner is already happening now and will happen in the future. In a partnership, both halves bring skill sets to the process of creation» (Mazzone, Elgammal 2019). Il fatto è, come ha ben argomentato Hannes Bajor (2021; cfr. <https://hannesbajohr.de/>), che questi sistemi si basano proprio su un'idea del tutto astorica di 'autonomia' dell'arte e dell' 'io creatore', e che operando su criteri stilistici dedotti dai prodotti artistici umani, finiscono con il riprodurre quel che è stato già prodotto dall'arte umana. Questo tipo di 'arte' non è imperfetta o incompleta, ma è non creativa, nel senso che ri-produce sulla base di quel che è stato già prodotto: una versione cibernetica, dunque, della mimesi antica. La perfezione del prodotto finale non significa affatto che tale prodotto sia 'arte', venendo a mancare, come avrebbero detto i Greci, della *dynamis*, di quella 'forza' indefinibile, per cui un prodotto artistico possa dirsi nuovo. Queste 'macchine creative', a me sembra, possono però apportare molto alla critica d'arte e anche ad alcune attività legate a criteri di selezione e scelta di tipo stilistico e probabilistico, come la critica del testo. La filologia, intesa come edizione dei testi, può ricavare molto giovamento dai sistemi di intelligenza artificiale, e dalla loro capacità di elaborare dati. Lo stesso credo possa dirsi per alcune questioni di autenticità o attribuzione, sia letteraria che artistica. Ma un'arte che si serva dell'intelligenza artificiale dovrebbe innanzitutto rinunciare alle nozioni di 'soggettività' artistica e soprattutto di 'soggettività umana' come soggettività superiore. Forse solo con questa umiltà potremo cercare nuove vie per la convivenza tra noi e le macchine, e nuove vie anche all'arte e alla letteratura.

Bibliografia

- BAJOR H. 2021, *Keine Experimente: Über künstlerische Künstliche Intelligenz*, «Merkur» 864, pp. 32-44.
- CATANI S. 2020, „Erzählmodus an“. *Literatur und Autorschaft im Zeitalter künstlicher Intelligenz*, «Jahrbuch der Deutschen Schillergesellschaft» 64, pp. 287-310.
- ELGAMMAL A. 2020, *Creative GAN Generating Art Deviating from Style Norms*, Brevetto USA nr. 10.853.986.
- HARAWAY D.H. 1995, *Manifesto cyborg. Donne, tecnologie e biopolitiche del corpo*, Milano (ed. or. 1991).
- ISHIGURO K. 2021, *Klara e il sole*, Torino 2021 (ed. or. 2021).
- KEHLMANN D. 2021, *Mein Algorithmus und ich. Stuttgarter Zukunftrede*, Stuttgart.
- LIVELEY G., THOMAS S. (2020), *Homer's Intelligent Machines. AI in Antiquity*, in

- CAVE S., DIHAL K., DILLON S. (eds.), *AI Narratives. A History of Imaginative Thinking about Intelligent Machines*, Oxford, pp. 25-48.
- MANZOTTI R. 2015, *Facebook: junk emotion*, «Doppiozero» (<https://www.doppiozero.com/materiali/web-analysis/facebook-junk-emotion>).
- MANZOTTI R. 2019, *Ciclo della vita o ciclo delle macchine?*, «Doppiozero» (<https://www.doppiozero.com/materiali/ciclo-della-vita-o-delle-macchine>).
- MANZOTTI R., TAGLIASCO V. 2002, *Si può parlare di intelligenza artificiale?*, «Sistemi intelligenti» XIV, 1, aprile, pp. 89-107.
- MAZZONE M., ELGAMMAL A. 2019, *Art, Creativity, and the Potential of Artificial Intelligence*, «Arts» 8/1, 26 (<https://www.mdpi.com/2076-0752/8/1/26>).
- PIZZO A. 2021, *Performing/Watching Artificial Intelligence on Stage*, «SKENÈ – Journal of Theatre and Drama Studies» 7, 1, pp. 91-109.
- ROBERTS S. (2019), *From Homer to HAL: 3000 years of AI narratives* (<https://www.cam.ac.uk/ainarratives>).
- ROSSI F. (2019), *Il confine del futuro. Possiamo fidarci dell'intelligenza artificiale?*, Milano.
- SHEVLIN H. (2019), *Non-human consciousness and the specificity problem: A modest theoretical proposal*, «Mind & Language» 36, pp. 297-314 (<https://doi.org/10.1111/mila.12338>).
- STOCK-HOMBURG R. 2021, *Survey of Emotions in Human-Robot Interactions: Perspectives from Robotic Psychology on 20 Years of Research*, «International Journal of Social Robotics» (<https://doi.org/10.1007/s12369-021-00778-6>).
- TAGLIASCO V. 2008, *Coscienza artificiale? Il sogno di un ingegnere*, «Sistemi intelligenti» XX, 3, pp. 551-553.
- ZEH J. 2010, *Corpus delicti. Un processo*, Milano (ed. or. 2009).