Università degli Studi Suor Orsola Benincasa

ANNALI

2022

*



ANNALI

2022

*

SUOR ORSOLA
BENINCASA
UNIVERSITÀ EDITRICE

Direttori
Piero Craveri
Vincenzo Omaggio

Direttore responsabile
Arturo Lando

Comitato scientifico
Lucio d'Alessandro
Francesco M. De Sanctis
Enricomaria Corbi
Tommaso Edoardo Frosini
Emma Giammattei
Aldo Sandulli
Paola Villani

Redazione
Luciana Trama (responsabile)
Emanuele Garzia
Flavia Soprani

Gli articoli negli «Annali» dell'Università Suor Orsola Benincasa sono preliminarmente sottoposti a un giudizio di referaggio nel rispetto dell'anonimato sia dell'autore che dei revisori (a doppio cieco). La valutazione è affidata a uno o più specialisti del tema trattato, designati dal Direttore degli «Annali» anche fra i componenti del Comitato Scientifico, che rivestano, o abbiano rivestito, la posizione di professore universitario nelle Università italiane o straniere ovvero si distinguano come esperti di chiara fama.

I revisori giudicano, entro 30 giorni dal ricevimento dell'elaborato, sull'opportunità di accogliere il lavoro scientifico tenendo conto dei parametri valutativi di cui alla vigente legislazione e formulando un giudizio motivato – senza modifiche o previo apporto di modifiche – di accoglimento o di rigetto. Il Direttore, infine, su propria responsabilità, può decidere di non assoggettare a revisione i lavori di particolare pregio o di autori di particolare prestigio.

—

Progetto grafico
Flavia Soprani

Realizzazione e stampa
Carmine Marra

© Università degli Studi Suor Orsola Benincasa, Napoli 2022
Tutti i diritti sono riservati

ISSN *online* 2281-3241
ISSN *press* 2037-5867

SOMMARIO

7 MARIA AURINO
Il museo come medium *per sperimentare e progettare itinerari educativi e formativi. Oltre la fruizione. Tra creatività e riflessività*

49 DANIELA CARDONE
Rauschenberg cancella Johns. L'arte fuori dal quadro

67 CLELIA CASTELLANO
La negoziazione delle forme soggettive femminili al di fuori delle norme apparenti nella micro società delle donne dell'Alto Atlante

87 DANILA BARDI ANTONELLA GRITTI GENNARO CATONE ROBERTA VACCA
Uno spazio di ascolto per gli studenti universitari: percezione del servizio di Counselling *e atteggiamento verso la relazione di aiuto*

111 ANNA GIANNETTI PASQUALE ROSSI FRANCESCA CASTANÒ
La Croce di Lucca: fondazione, committenza e storie di una chiesa napoletana

147 FEDERICA IZZO ALESSANDRA STORLAZZI PAOLA VILLANI
L'evoluzione dei percorsi narrativi: dagli studi classici al contesto aziendale

167 MARCO NIELI
Le fonti latine e brittoniche dell'Arthur pre-galfridiano: un aggiornamento sullo stato della ricerca e alcune prospettive di interpretazione contestuale

205 ANTONIO PERRI
Decostruire il circuit de la parole*? Note in difesa di un «minimalismo» saussuriano*

231 CIRO RANISI
Érec et Énide *e il romanzo «matrimoniale»*

253 ILENIA TAMBURRO
Il Discobolo Lancellotti: una nuova lettura dell'opera

293 NATASCIA VILLANI
L'intelligenza artificiale e le sue sfide etiche

NATASCIA VILLANI

L'intelligenza artificiale e le sue sfide etiche

Ci troviamo di fronte ad una nuova tecnologia che sta annullando i confini tra il fisico, il digitale e il biologico: l'intelligenza artificiale. L'articolo nel tratteggiarne gli elementi ripercorre il dibattito attuale sull'impatto e sui possibili rischi dei sistemi di IA. Negli ultimi anni ci si è interrogati sulle sfide etiche e sull'utilizzo dell'IA secondo principi fondamentali di una buona innovazione.

Parole chiave: intelligenza artificiale, etica del digitale, algoretica, agere, intelligere

We are faced with a new technology that is blurring the boundaries between the physical, the digital and the biological artificial intelligence. In outlining its elements, the article traces the current debate on the impact and possible risks of AI systems. In recent years, there have been questions about the ethical challenges and the use of AI according to basic principles of good innovation.

Keywords: words, artificial intelligence, digital ethics, algoretic, agere, intelligere

«Di tutte le nuove avvincenti sfide che ci troviamo ad affrontare al giorno d'oggi, comprendere e definire la nuova rivoluzione tecnologica è la più affascinante e suggestiva, poiché comporta una vera e propria trasformazione per l'umanità. Sia-

mo alle soglie di una rivoluzione che sta cambiando drasticamente il modo in cui viviamo, lavoriamo e ci relazioniamo con gli altri. Se ne valutiamo la portata e la complessità, quella che io considero la quarta rivoluzione industriale è qualcosa con cui l'uomo non ha mai dovuto confrontarsi prima d'ora». È questo l'*incipit* del testo di Klaus Schwab, *La quarta rivoluzione industriale*[1], il quale rivolge un invito a tutti gli esperti di varie discipline ad interrogarsi su cosa questo cambiamento significhi e su come affrontarlo. Occorre avere una visione condivisa del cambiamento in atto: ci troviamo di fronte ad una nuova tecnologia che sta annullando i confini tra il fisico, il digitale e il biologico: l'intelligenza artificiale.

Secondo Nicholas Carr ogni tecnologia è un'espressione della volontà umana al fine di estendere il potere e il controllo sul mondo circostante: sulla natura, sul tempo, sulla distanza e gli uni sugli altri[2]. Vi sono tecnologie che ampliano la forza fisica, la destrezza o la capacità di recupero; altre che estendono la portata o l'acutezza dei nostri sensi, altre ancora che ci aiutano a dare una nuova forma alla natura per assecondare le nostre necessità e i nostri desideri.

Esiste poi una quarta categoria che si potrebbe definire di «tecnologie intellettuali»[3]: «Questa locuzione si riferisce a tutti gli strumenti che usiamo per ampliare o rafforzare le nostre facoltà mentali [...] che usiamo per esprimerci, per dare una forma alla nostra identità personale e pubblica e per coltivare le relazioni con gli altri»[4].

[1] K. Schwab, *La quarta rivoluzione industriale*, Franco Angeli, Milano 2016.

[2] N. Carr, *Internet ci rende stupidi? Come la rete sta cambiando il nostro cervello*, Raffaello Cortina Editore, Milano 2011.

[3] Carr prende a prestito un termine usato con un significato leggermente diverso dall'antropologo sociale Jack Goody e dal sociologo Daniel Bell (N. Carr, *Internet ci rende stupidi?*, cit., p. 68).

[4] *Ibidem*. Sulla interazione con le tecnologie vedi F. Parisi, *La tecnologia che siamo*, Codice Edizioni, ed. Kindle, Torino 2019.

Questa definizione, pur se meno generica di quella di Wanda Orlikowski, («insieme di regole e risorse messe in atto nella pratica»[5]), può comunque allinearsi al ragionamento condotto da Federico Cabitza sulla «intelligenza artificiale» e su come dovremmo porci eticamente in relazione con essa: anziché considerare i sistemi di IA quali «agenti vivi, esistenze vere e proprie che si integrano con naturalezza e grazia nel nostro ambiente domestico, urbano e lavorativo»[6], come potrebbe suggerire l'espressione «intelligenza artificiale», occorre risalire faticosamente verso l'idea che tale espressione riguardi «manifestazioni e modalità del fare umano»[7].

Per secoli storici e filosofi hanno ricostruito e discusso il ruolo della tecnologia nel dar forma alla civiltà. Da un lato i deterministi, secondo cui il progresso tecnologico è visto come una forza autonoma al di fuori del controllo umano[8]; dall'altro gli strumentalisti, che ne minimizzano il potere ritenendoli manufatti neutrali asserviti ai desideri inconsapevoli dei loro utilizzatori[9]. «Se l'esperienza della società moderna ci insegna qualcosa», osserva il politologo Langdon Winner, «è che le tecnologie non sono soltanto ausili per l'attività umana ma anche forze potenti che agiscono per dare forma e significato nuovi

[5] W.J. Orlikowski, *Using technology and constituting structures. A practice lens for studying technology in organisations*, in «Organisation Science», 11 (2000), pp. 404-428.

[6] E. Sadin, *Critica della ragione artificiale. Una difesa dell'umanità*, Luiss University Press, Roma 2019.

[7] F. Cabitza, Deus in machina*? L'uso umano delle nuove macchine, tra dipendenza e responsabilità*, in L. Floridi, F. Cabitza, *Intelligenza artificiale. L'uso delle nuove macchine*, Bompiani, Milano 2021, ed. Kindle, p. 11.

[8] La visione deterministica nella sua concezione più estrema considera gli esseri umani poco più che «organi sessuali del mondo della macchina» (M. McLuhan, *Gli strumenti del comunicare*, Garzanti, Milano 1986, p. 61), fino a far diventare gli uomini stessi superflui.

[9] «La tecnologia è la tecnologia», ha affermato il critico dei media James Carey, «uno strumento per la comunicazione e il trasporto attraverso lo spazio, nulla di più» (J.W. Carey, *Communication as Culture. Essays on Media and Society*, Routledge, New York 2008, p. 107).

a quell'attività»[10]. A volte i nostri strumenti fanno ciò che noi diciamo loro di fare. Altre volte siamo noi a adattarci alle loro richieste[11]. Questo emerge in modo evidente nelle tecnologie intellettuali e in modo particolare nell'intelligenza artificiale.

Secondo Kate Crawford, prima di interrogarsi sull'impatto dell'IA nella società attuale, occorre sfatare due miti. Il primo mito è credere che i sistemi non umani (qualsiasi tecnologia intellettuale) «siano qualcosa di analogo alla mente umana». Il secondo mito «è che l'intelligenza sia qualcosa di indipendente, una sorta di elemento naturale, distinto dalle forze sociali, culturali, storiche e politiche»[12].

Questi miti sono particolarmente potenti nel campo dell'intelligenza artificiale, dove la convinzione che l'intelligenza umana possa essere formalizzata e riprodotta dalle macchine è stata dominante sin dalla metà del XX secolo[13]. Già allora Joseph Weizenbaum, pioniere dell'IA e creatore di Eliza, il primo *chatbot* o *software* progettato per simulare una conversazione con un essere umano, non era d'accordo con questa visione ritenendo estremamente riduttivo un concetto di intelligenza in cui gli esseri umani fossero concepiti come semplici

[10] L. Winner, *Technologies as forms of life*, in *Readings in the Philosophy of Technology*, a cura di D.M. Kaplan, Rowman & Littlefield, Lanham 2004, p. 105.

[11] Cfr.: N. Carr, *Internet ci rende stupidi?*, cit., p. 72.

[12] K. Crawford, *Né intelligente né artificiale. Il lato oscuro dell'IA*, Il Mulino, Bologna 2021, ed. Kindle, p. 12. Vedi anche: F. Cabitza, *Deus in machina?*, cit. «È utile evitare di considerare macchine e strumenti come qualcosa di ontologicamente stabile e fisso, a dispetto della loro materialità; o come qualcosa dotato di "agenzia", cioè di una capacità autonoma di azione; questo perché farlo potrebbe indurci nell'errore di attribuire alle macchine anche un'identità e, perfino, una volontà autonoma» (p. 13).

[13] *Ibidem*. Alla domanda se le macchine potessero pensare, Alan Turing, in un articolo del 1950 (*Computing Machinery and Intelligence*, in «Mind», 236 (1950), pp. 433-460) pose la questione in altri termini: una macchina può comportarsi in modo da sembrare intelligente (*imitation game*), cioè da sembrare ad un osservatore esterno un essere umano? Secondo Cabitza ponendo questa domanda Turing riprende la *Seconda meditazione* di Cartesio per proporre il primato del comportamento sull'essere (vedi F. Cabitza, *Deus in machina?*, p. 21).

sistemi di elaborazione delle informazioni e dal quale discendesse la «grande fantasia perversa» per cui gli scienziati dell'IA avrebbero potuto creare una macchina in grado di apprendere «come un bambino»[14].

Questa è stata una delle dispute centrali della storia dell'IA a partire dagli anni Sessanta[15]. Hubert Dreyfus, preoccupato della deriva che aveva preso la discussione, nel suo lavoro *What Computers Can't Do*[16], sottolineò una basilare differenza: mentre i computer richiedono che i processi e i dati siano espliciti e formalizzati, l'intelligenza e l'esperienza umana si basano in larga misura su numerosi processi inconsci e subconsci.

«Di conseguenza, gli aspetti meno formali dell'intelligenza devono essere estratti, eliminati o approssimati per i computer, rendendoli pertanto incapaci di elaborare le informazioni sulle situazioni come fanno gli esseri umani»[17].

Da quegli anni sono cambiate molte cose nell'IA, e tra queste il passaggio dai sistemi simbolici alle tecniche di apprendimento automatico. «Sotto molti aspetti le antiche controversie su ciò che l'IA può fare sono finite nel dimenticatoio e lo scetticismo si è dissolto»[18]. Negli ultimi anni l'intelligenza artificiale si è rapidamente espansa sia come oggetto di studi accademici che nell'industria. Oggi un piccolo numero di potenti società tecnologiche implementa su scala planetaria sistemi di intelligenza artificiale che ancora una volta vengono salutati come pari o addi-

[14] Cfr. J. Weizenbaum, *On the impact of the computer on society. How does one insult a machine?*, in «Science», 176 (1972), p. 612.

[15] Pietra miliare furono le conferenze tenute nel 1961 presso il MIT (*Massachusetts Institute of Technology*) con il titolo *Management and the Computer of the Future* per discutere i rapidi progressi compiuti dall'informatica digitale. Al termine delle conferenze, John McCarthy, inventore del termine *Intelligenza artificiale*, si spinse a sostenere l'illusorietà delle differenze tra lavoro umano e lavoro delle macchine.

[16] H.L. Dreyfus, *Che cosa non possono fare i computer. I limiti dell'intelligenza artificiale*, Armando, Roma 1988.

[17] K. Crawford, *Né intelligente né artificiale*, cit., p. 13.

[18] *Ibidem*.

rittura superiori all'intelligenza umana[19]. Ma che cosa si intende per intelligenza artificiale? Già a partire dagli anni Sessanta ci si riferiva a quel campo di indagine di intersezione delle nuove discipline impegnate a ridurre la complessità del mondo fisico a meccanismi informazionali retroazionati o a modelli simbolici e computabili: cibernetica e informatica. Da allora l'IA è stata definita in molti modi e non esiste una definizione unitaria su cui tutti concordano[20]. Dopo sessantacinque anni l'*High-Level Expert Group on Artificial Intelligence* (AI Hleg) nominato dalla Commissione Europea per redigere un documento di orientamenti etici per un'IA affidabile, definisce l'Intelligenza artificiale come «sistemi che mostrano un comportamento intelligente analizzando il proprio ambiente e compiendo azioni, con un certo grado di autonomia, per raggiungere specifici obiettivi»[21]. Tale definizione sarà poi ampliata al termine del documento nel seguente modo: «I sistemi di intelligenza artificiale (IA) sono sistemi *software* (ed eventualmente *hardware*) progettati dall'uomo che, dato un obiettivo complesso, agiscono nella dimensione fisica o digitale percependo il proprio ambiente attraverso l'acquisizione di dati, interpretando i dati strutturati o non strutturati raccolti, ragionando sulle conoscenze, o elaborando le informazioni derivate da questi dati e decidendo le migliori azioni da intraprendere per raggiungere l'obiettivo dato. I sistemi di IA possono usare regole simboliche o apprendere un modello numerico, e possono anche

[19] Si pensi che nel 2021 i *venture capitalist* americani hanno spostato in questo settore 52,9 miliardi di dollari, cifra tre volte più grande di quello cinese (17 miliardi di dollari) e dieci volte più grande del corrispettivo britannico (4,6 miliardi di dollari) (fonte: *www.agendadigitale.eu*).

[20] Un rapporto ormai datato (S. Legg, M. Hunter, *A collection of definition of intelligence*, in *Advances in Artificial General Intelligence. Concepts, Architecture, and Algorithms*, a cura di B. Goertzel e P. Wang, IOS press, Amsterdam 2007, pp. 17-24) elencava già 18 definizioni di IA. Il numero è in crescita (vedi S. Russel, P. Norvig, *Intelligenza artificiale. Un approccio moderno*, Pearson, Milano-Torino 2021).

[21] AI Hleg, *Una definizione di IA: principali capacità e discipline*, Commissione Europea, Bruxelles 2019, p. 1.

adattare il loro comportamento analizzando come l'ambiente è influenzato dalle loro azioni precedenti»[22].

Il documento rileva che, nonostante i numerosi studi utilizzare il termine intelligenza, proprio per la sua vaghezza, è meno adeguato del termine «razionalità»[23], intesa quale capacità di scegliere la migliore azione da intraprendere per conseguire un determinato obiettivo alla luce di alcuni criteri da ottimizzare e delle risorse a disposizione. Se consideriamo l'IA come una forma di automazione, qualcosa di riconducibile a semplici operazioni di lettura e scrittura di simboli, «sgombriamo il campo da diversi equivoci e spettri, come quelli dell'identità, della coscienza e della autonomia»[24].

Si potrebbe dire con le parole di uno dei più grandi esperti di robotica, Rodney Brooks: «ma è solo computazione». Infatti la razionalità di un sistema di IA si esprime attraverso il seguente processo: «percependo tramite sensori l'ambiente in cui è immerso, e dunque raccogliendo e interpretando i dati, ragionando su ciò che viene percepito o elaborando le informazioni desunte dai dati, decidendo quale sia l'azione migliore e agendo di conseguenza attraverso i suoi attuatori, infine modificando eventualmente l'ambiente. I sistemi di IA possono usare regole simboliche o apprendere un modello numerico e possono anche adattare il loro comportamento analizzando come l'ambiente è influenzato dalle loro azioni precedenti»[25].

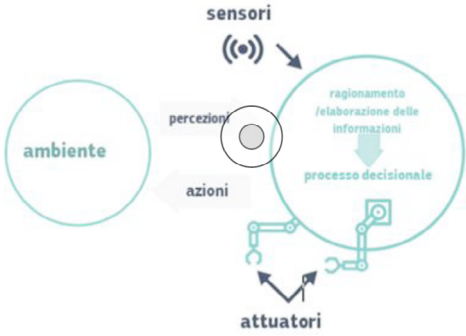
Lo schema riportato nel documento esemplifica un sistema di IA. Ogni sistema di IA interagisce con l'ambiente in quan-

[22] *Ivi*, p. 6.

[23] Massimo Chiriatti preferisce parlare invece di «inconscienza», perché gli algoritmi, eseguendo regole che imparano autonomamente dai dati, producono risultati senza alcuna comprensione e coscienza di ciò che stanno facendo (M. Chiriatti, *Incoscienza artificiale. Come fanno le macchine a prevedere per noi*, Luiss University Press, Roma 2021, ed. Kindle, p. 22).

[24] F. Cabitza, *Deus in machina?*, cit., p. 19.

[25] AI Heg, *Una definizione di IA*, cit., p. 1.



to necessita di dati[26] raccolti attraverso sensori, dati che il sistema acquisisce come *input*. Tali dati devono essere trasformati in informazioni che il modulo ragionamento/elaborazione delle informazioni possa comprendere per poi elaborare un processo decisionale. È questo l'elemento cardine di un sistema di IA che si tradurrà poi in azione mediante attuatori. «Un sistema razionale che apprende è un sistema razionale che, dopo aver eseguito un'azione, valuta la nuova condizione dell'ambiente (attraverso la percezione) per misurare il successo dell'azione e quindi adatta le proprie regole di ragionamento e i propri metodi decisionali»[27]. Percezione, ragionamento/processo decisionale e attuazione sono, schematicamente, le tre fasi del processo su cui si basano la maggior parte delle tecniche usate per costruire sistemi di IA.

Sintetizzando al massimo tutte queste tecniche possono essere organizzate in due gruppi principali che afferiscono alla capacità di ragionamento e di apprendimento. Il gruppo di tecniche che afferisce al ragionamento/processo decisionale comprende la rappresentazione delle conoscenze e il ragionamento basato sulle conoscenze, la pianificazione, la programmazione,

[26] Tali dati possono essere strutturati o non strutturati a seconda se sono organizzati seguendo uno schema noto o no.

[27] AI HLEG, *Una definizione di IA*, cit., p. 3.

la ricerca e l'ottimizzazione. Tali tecniche permettono di svolgere il ragionamento sui dati provenienti dai sensori. Fanno parte di questo processo la formulazione di inferenze tramite regole simboliche, la pianificazione e la programmazione di attività, la ricerca tra una vasta gamma di soluzioni e l'ottimizzazione tra tutte le possibili soluzioni a un problema. Il gruppo di tecniche che fa capo all'apprendimento include l'apprendimento automatico (*Machine Learning* / ML), le reti neurali (*Artificial neural networks* / ANN), l'apprendimento profondo (*Deep Learning* / DL), e molte altre tecniche di apprendimento.

Grazie ad esse un sistema di IA può imparare a risolvere problemi che non possono essere definiti in modo preciso, o il cui metodo di soluzione non può essere descritto con regole di ragionamento simboliche. Esempi di tali problemi sono quelli relativi alle capacità di percezione, come la comprensione della parola e del linguaggio, la visione artificiale o la previsione del comportamento[28].

Chiunque non sia stupito di fronte ad una tale rivoluzione digitale non ne ha afferrato la portata «Stiamo parlando di un nuovo capitolo della storia umana»[29]. Siamo sulla soglia di una nuova epoca e il prezzo di occupare un posto così speciale lo si paga con incertezza da un lato, ma anche con la responsabilità di gettare solide basi per il futuro. A tal fine si può prendere spunto da quanto detto dal cardinale Carlo Maria Martini in

[28] Un'altra disciplina pertinente è la robotica che può essere definita come una «IA in azione nel mondo fisico» (anche detta *IA incorporata*). Il robot è una macchina fisica studiata per affrontare le dinamiche, le incertezze e le complessità del mondo fisico. Nell'architettura di controllo del sistema robotico sono generalmente integrate capacità di percezione, di ragionamento, di azione, di apprendimento nonché di interazione con altri sistemi.

[29] L. Floridi, *Etica dell'intelligenza artificiale. Sviluppi, opportunità, sfide*, Raffello Cortina Editore, Milano 2022, p. 11. Sui pericoli che tale trasformazione porta con sé in campo politico a livello globale si veda: S. De Luca, Hic sunt leones. *La democrazia nell'era dei social media, dei big data e dell'intelligenza artificiale*, in «Rivista di Politica», 2 (2020), pp. 147-159.

uno dei dialoghi da lui promossi in *Le cattedre dei non credenti*: «Sono così minacciose tutte le tecnologie del virtuale? L'intero cammino verso l'intelligenza artificiale finirà per svalutare il valore della persona, riducendola a pura meccanica? O, invece, saranno i valori dell'uomo a indurre la scienza ad aprire nuovi fronti grazie alle conquiste tecnologiche? [...] Tutto questo costituisce uno scenario molto incoraggiante, purché l'intelligenza umana rimanga padrona dei processi»[30].

Le sfide etiche poste dai sistemi di IA impegnano la discussione attuale di ricercatori di varie discipline.

Tra queste Luciano Floridi[31], nel suo ultimo libro, argomenta in modo convincente che l'IA separa la risoluzione efficace dei problemi e l'esecuzione corretta (l'*agere*) dal comportamento intelligente (*intelligere*). Grazie a tale scissione l'IA può incessantemente «colonizzare lo spazio sterminato di problemi e compiti [...] è proprio quando smettiamo di cercare di produrre intelligenza umana che possiamo sostituirla con successo in un numero crescente di compiti»[32].

Tale divorzio si presenta eticamente problematico in quanto gli agenti artificiali «sono sufficientemente informati, "smart", autonomi e in grado di compiere azioni moralmente rilevanti indipendentemente dagli esseri umani che li hanno creati»[33]. Tale scissione ha modificato una delle equazioni su cui si è sempre

[30] C.M Martini, *Le cattedre dei non credenti*, Bompiani, Milano 2015, ed. Kindle, p. 832. Per Paolo Benanti una governance delle tecnologie legate all'IA è un appello obbligante alle coscienze, «uno spazio dove le considerazioni antropologiche ed etiche, in un mutuo scambio e dialogo, devono divenire forze efficaci per plasmare e guidare l'innovazione tecnologica, rendendola autentica fonte di sviluppo umano» (P. Benanti, *Le macchine sapienti. Intelligenze artificiali e decisioni umane*, Marietti, Torino 2018, p. 130).

[31] L. Floridi, *Etica dell'intelligenza artificiale*, cit.

[32] *Ivi*, p. 52.

[33] L. Floridi, J.W. Sanders, *On the morality of artificial agents*, in «Minds and Machines», 14 (2004), pp. 349-379. Nel campo della sanità digitale o nei mercati finanziari tutti i dati rilevanti sono leggibili da macchine cosicché decisioni e azioni possono essere compiute automaticamente da applicazioni e attuatori.

fondata la valutazione morale: l'*agere* è sempre stato associato all'intenzionalità, in quanto l'agire significa passare dalla sfera interiore della conoscenza e della decisione alla sfera esterna degli enti. L'agire in tal senso è un «fenomeno umano originario».

Romano Guardini spiega questo concetto confrontando le azioni dell'uomo con le attività svolte dagli animali. L'uomo agisce sulla base della conoscenza e della decisione.

Egli con l'azione si spinge nella realtà, la coglie, ne dispone, le dà forma, ma il suo nucleo resta interiore, perché parte dall'intenzione. Mediante l'azione l'intenzione diventa storia: lo spazio interiore supera o precede la storia, sebbene agendo l'intenzione si sottoponga al tribunale della storia. Non c'è moralità che possa restare confinata nell'intenzione, deve entrare nella realtà e scontrarsi con le resistenze che la realtà pone. Diversamente nell'animale le azioni sono totalmente connesse alla vita animale corrispondente e si prolungano nell'ambiente circostante. Tale ambiente è armonizzato alla vita dell'animale: «l'animale è propriamente il concreto essere vivente più il suo ambiente, è l'animale nel suo ambiente [...] l'agire dell'animale appartiene a questo contesto e proprio la perfezione di questo agire è un segno del fatto che ha il carattere di funzione» [34].

Si potrebbe dire che una cosa molto simile accade nei sistemi di IA: le loro azioni sono connesse all'ambiente. Negli anni quaranta e cinquanta il computer era una stanza e l'interazione era fisica; negli anni settanta siamo usciti per sederci davanti a esso e la relazione divenne semantica; oggi siamo entrati nuovamente in esso sotto forma di «infosfera» [35], che ci circonda

[34] R. Guardini, *Etica. Lezioni all'Università di Monaco (1950-1962)*, Morcelliana, Brescia 2021, p. 163. Ad esempio la tela del ragno è esattamente come deve essere in base al peso e al tipo di movimento del ragno, al tipo degli insetti da catturare, in base alle effettive possibilità di fissazione, ecc.

[35] In tale infosfera l'assunto è che un agente possa essere artificiale. Ecco perché ci viene chiesto regolarmente di dimostrare che non siamo dei robot, cliccando su Captcha (*Completely Automated Public Turing test to tell Computers and Humans Apart*).

ed è sempre più adattata all'IA. Siamo, per usare l'espressione di Floridi, sempre più *onlife*. È l'ambiente che è progettato per essere compatibile con i robot, è una infosfera in cui ci sentiamo più ospiti analogici che digitali. Pertanto «non c'è da stupirsi che i nostri agenti artificiali si comportino sempre meglio. È il loro ambiente»[36].

Il rischio etico di questo avvolgimento è che gli esseri umani possono diventare parte del meccanismo senza averne consapevolezza e che, disegnando un mondo digitale a misura dell'*agere* artificiale e non dell'*intelligere* umano, le tecnologie potrebbero modellare i nostri ambienti fisici e concettuali spingendoci ad adattarci a loro che è il modo più semplice per far funzionare le cose. Ritornando all'invito di Martini che Floridi accoglie «il design umano intelligente dovrebbe svolgere un ruolo maggiore nel plasmare il futuro delle nostre interazioni con gli artefatti "smart" attuali e futuri e gli ambienti che condividiamo con loro. Dopotutto è un segno di intelligenza far lavorare la stupidità per noi»[37]. Il dibattito etico prese forma già a partire dagli anni Sessanta[38], anche se solo negli ultimi anni è aumentata l'esigenza di riflessioni sull'impatto dell'IA nella società. Già lo scrittore di fantascienza Isaac Asimov in un breve racconto del 1942[39], aveva elaborato le sue ormai famigerate «Tre

[36] L. Floridi, *Etica dell'intelligenza artificiale*, cit., p. 58.

[37] *Ivi*, p. 63.

[38] Si pensi al dibattito tra Wiener e Samuel. Wiener sosteneva che «machines can and do transcend some o limitations of their designers, and that in doing so they may be both effective and dangerous» (N. Wiener, *Some moral and technical consequences of automation*, in «Science», 131 (1960), pp. 1355-1358); Samuel invece «The machine is not a threat to mankind, as some people think. The machine does not possess a will, and its so-called "conclusions" are only the logical consequences of its input, as revealed by the mechanistic functioning of an inanimate assemblage of mechanical and electrical parts» (A. Samuel, *Some moral and technical consequences of automation – A refutation*, in «Science», 132 (1960), pp. 741-742).

[39] Nel racconto *Runaround* (tradotto in italiano con *Girotondo* o *Circolo vizioso*, inserito nella antologia *Io, Robot*, Bompiani, Milano 1979) l'autore formula le sue tre leggi: 1. un robot non può recar danno a un essere umano, né

Leggi della Robotica», e ai giorni d'oggi Frank Pasquale propone nuove leggi formulate per coloro che progettano, controllano e usano i sistemi di IA[40].

Negli ultimi anni molte sono state le iniziative promosse da varie organizzazioni per stabilire principi etici nell'interesse dell'IA socialmente vantaggiosa[41]. I principi emersi mostrano una loro coerenza «che è impressionante e rassicurante»[42]. Anche *La pontificia accademia per la vita* sotto la spinta del presidente, mons. Vincenzo Paglia, ha dimostrato uno specifico interesse per le nuove tecnologie, dedicando il biennio 2019-2020 alla robo-etica e alle questioni etico-antropologiche connesse alle cosiddette *intelligenze artificiali*.

L'esito di questo cammino ha trovato un primo momento di sintesi nella *Rome Call for AI Ethics*[43] quale documento di

può permettere che, a causa del proprio mancato intervento, un essere umano riceva danno; 2. un robot deve obbedire agli ordini impartiti dagli esseri umani, purché tali ordini non contravvengano alla Prima Legge; 3. un robot deve proteggere la propria esistenza, purché questa autodifesa non contrasti con la Prima o con la Seconda Legge.

[40] 1. I sistemi robotici e le AI devono essere complementari ai professionisti e non sostituirli; 2. I sistemi robotici e le AI non devono contraffare l'umanità; 3. I sistemi robotici e le AI non devono intensificare una corsa agli armamenti a somma zero; 4. I sistemi robotici e le AI devono sempre indicare l'identità dei loro creatori, controllori e proprietari (F. Pasquale, *Le nuove leggi della robotica. Difendere la competenza umana nell'era dell'intelligenza artificiale*, Luiss University Press, Roma 2021).

[41] Tra le maggiori iniziative: i *Principi di Asilomar per l'IA* (Future of Life Institute, 2017); la *Dichiarazione di Montréal per l'IA responsabile* (Università di Montréal, 2017); i principi generali offerti nella seconda versione di *Ethically Aligned Design. A Vision for Prioritizing Human Wellbeing with Autonomous and Intelligent Systems* (IEEE, 2017); i *Principi di partenariato sull'IA* (*Partnership on IA*, 2018).

[42] L. Floridi, *Etica dell'intelligenza artificiale*, cit., p. 95. Floridi nel fare un'analisi di tutti i principi emersi li raggruppa avvicinandoli ai 4 principi fondamentali usati in bioetica: beneficenza (ogni tecnologia è creata a beneficio dell'umanità); non maleficenza (mettere in guardia contro le conseguenze negative dell'uso dell'IA); autonomia (per evitare che la crescita di autonomia artificiale possa minare l'autonomia umana); giustizia (promuovere la prosperità, preservare la solidarietà ed evitare l'iniquità).

[43] Firmata a Roma il 28 febbraio 2020 dal presidente di Microsoft, Brad Smith, dal vice presidente di IBM, John Kelly III, dal direttore Generale della

impegni condiviso, in cui, in forma breve e sintetica, si offrono alcune linee per un'etica dell'Intelligenza Artificiale e per promuovere, secondo il termine coniato da Paolo Benanti, una «algor-etica»[44], ovvero lo sviluppo e l'utilizzo dell'IA secondo i principi, fondamentali di una buona innovazione. «Se le macchine riescono a surrogare l'uomo in tante decisioni, dobbiamo chiederci con quali criteri può avvenire questa surrogazione. In altre parole, se la macchina commette un errore chi è responsabile?

L'etica diventa il *guard rail* che ci permette di vivere in modo più sicuro con queste macchine sapienti. L'etica, però, è una questione di valori difficili da comunicare alle macchine, che funzionano sulla base di numeri. E allora bisogna mettere insieme algoritmi ed etica. Da qui nasce un nuovo termine, l'algoretica, la nuova disciplina che vorrebbe rendere le macchine capaci di computare principi tipicamente umani. Un percorso che coinvolge più discipline: non bastano più filosofia, tecnologia, informatica, serve la contaminazione»[45].

Finora le linee tracciate sono: «trasparenza», affinché i sistemi di IA possano essere comprensibili; «inclusione», affinché tutti possano beneficiare e a tutti gli individui possano essere offerte le migliori condizioni possibili per esprimersi e svilupparsi; «responsabilità», di coloro che progettano e implementano soluzioni di IA; «imparzialità», affinché siano salvaguardate l'equità e la dignità umana; «affidabilità», perché i sistemi di IA devono essere in grado di funzionare in modo affidabile; «sicurezza e *privacy*», affinché sia rispettata la *privacy* degli utenti.

Fao, Qu Dongyu, e dal ministro per l'Innovazione tecnologica, Paola Pisano (governo Conte) e con la partecipazione dell'allora presidente del Parlamento europeo, David Sassoli.

[44] P. Benanti, *Oracoli. Tra algoretica e algocrazia*, Luca Sossella editore, Roma 2018.

[45] E. Coen, *Al futuro serve l'algoretica. Colloquio con Paolo Benanti*, «L'Espresso», 23/2/2020, p. 77.

Il punto di partenza per Benanti è che non dobbiamo cadere nella logica del datismo, o come lui la definisce del *dataismo*, la nuova religione che irretita da una logica del fare, considera l'intero universo come un flusso di dati. È la «nuova narrazione universale che legittima un nuovo principio di legittimità: gli algoritmi e i Big Data»[46]. Pertanto il ruolo delle riflessioni etiche non dovrà incentrarsi nel trovare soluzioni tecniche, ma nel rendere presente la domanda critica sul senso dell'umano che è quello che deve guidare le decisioni della nostra vita. «Con un computer possiamo trasformare quasi tutti i problemi umani in statistiche, grafici, equazioni.

La cosa davvero inquietante, però, è che così facendo creiamo l'illusione che questi problemi siano risolvibili con i computer»[47]. Lehonard Gerd per evidenziare la specificità dell'essere umano e le sue qualità parla di androritmi che dobbiamo assolutamente preservare anche se, nel confronto con sistemi, computer e robot non biologici, ci comunicano un senso di goffaggine, complessità, lentezza, pericolo o inefficienza[48]. C'è il rischio che questa «fallibilità», cifra della nostra umanità, sia considerata come nociva ai valori fondamentali[49].

Mentre la lentezza degli umani non è un *bug*, un errore dell'evoluzione o uno svantaggio rispetto alle macchine; è piuttosto ciò che ci differenzia, il tempo della coscienza, la possibilità di farci domande, di chiederci cosa «sentiamo»[50]. Massimo Chiriatti in modo provocatorio ha iniziato il suo testo con un

[46] P. Benanti, *Le macchine sapienti. Intelligenze artificiali e decisioni umane*, Marietti, Torino 2018, p. 98.

[47] N. Yehia, Homo cyborg. *Il corpo postumano tra realtà e fantascienza*, Eleutheria, Milano 2005, p. 15.

[48] L. Gerd, *Tecnologia vs Umanità. Lo scontro prossimo venturo*, Egea, Milano 2019, ed. Kindle, p. 693. In contrapposizione all'acronimo Stem, l'autore ne conia coniato un altro: Core che sta per creatività/compassione, originalità, reciprocità/responsabilità ed empatia.

[49] P. Benanti, *Le macchine sapienti*, cit., p. 108.

[50] M. Chiriatti, *Incoscienza artificiale*, cit., p. 19.

brano generato da una IA per porci davanti alla questione in cui ci imbattiamo quotidianamente: l'Uomo vede l'Intelligenza Artificiale come una macchina in grado di prendere le sue decisioni, ma si sbaglia, perché è solo un calcolatore di simboli, anche se sempre più sofisticato; l'Intelligenza Artificiale vede l'Uomo come un insieme di numeri, ma si sbaglia, perché la coscienza è incomputabile[51].

Siamo di fronte a sfide senza precedenti e inedite: problemi nuovi richiedono soluzioni nuove. È questo l'invito di papa Francesco.

> «Un mondo migliore è possibile grazie al progresso tecnologico se questo è accompagnato da un'etica fondata su una visione del bene comune, un'etica di libertà, responsabilità e fraternità, capace di favorire il pieno sviluppo delle persone in relazione con gli altri e con il creato»[52].

[51] *Ivi*, p. 17 e ss.

[52] Discorso del santo padre Francesco ai partecipanti al seminario *Il Bene comune nell'era digitale*, promosso dal Pontificio Consiglio della Cultura e dal Dicastero per il Servizio dello Sviluppo Umano Integrale, 27 settembre 2019.