

JOHN C. LENNOX

2004

**DIO, L'INTELLIGENZA ARTIFICIALE
È IL FUTURO DELL'UMANITÀ**

Titolo originale:

Originally published in English under the title
"2084"

Published by arrangement with

The Zondervan Corporation L.L.C.

a subsidiary of HarperCollins Christian Publishing, Inc.

3900 Sparks Dr. SE, Grand Rapids, Michigan 49546 – U.S.A.

Copyright c 2020 by John C. Lennox

All rights reserved

Edizione italiana:

"2084"

Dio, l'Intelligenza Artificiale e il futuro dell'umanità

© ADI-Media

Via della Formica, 23 - 00155 Roma

Tel. 06 2251825 - 06 2284970

Cell. 388 7334503

Email: adi@adi-media.it

Internet: www.adimedia.it

Servizio Pubblicazioni delle

Chiese Cristiane Evangeliche

"Assemblee di Dio in Italia"

Maggio 2021 - Tutti i Diritti Riservati

Traduzione: A cura dell'Editore - M.C.

Tutte le citazioni bibliche, salvo che

non sia indicato diversamente, sono tratte

dalla Bibbia Versione Nuova Riveduta - Ed. 2006

Società Biblica di Ginevra - Svizzera

Stampa: Rotomail Italia S.p.A. - Vignate (MI)

ISBN 978 88 3306 202 0

INTRODUZIONE

Questo libro è un tentativo di rispondere alle domande sulla direzione che l'umanità sta prendendo riguardo agli sviluppi in campo tecnologico, bioingegneristico e, in particolare, dell'intelligenza artificiale. Ci si chiede: "Saremo in grado di costruire la vita artificiale e la superintelligenza?"; "gli umani saranno in grado di modificare sé stessi per diventare completamente diversi?"; "quali implicazioni emergono dallo sviluppo della intelligenza artificiale (IA) alla luce della nostra visione del mondo e di Dio?".

Spero che il titolo, di ispirazione orwelliana, non sembri pretenzioso poiché il mio libro non è un romanzo distopico né, tantomeno, io sono George Orwell. In realtà, il titolo mi fu suggerito da un collega di Oxford, il professor Peter Atkins, mentre eravamo impegnati, da prospettive diverse, in un dibattito dal titolo: "La scienza può spiegare tutto?".* Sono debitore, a questo mio collega, dell'idea, alla luce di parecchi incontri pubblici anche piuttosto serrati, volti ad affrontare varie tematiche riguardanti la scienza e Dio.

* Si veda John C. Lennox, *La scienza può spiegare tutto?* ADI-Media, Roma, gennaio 2020. N.d.E.

Mi considero in debito anche con altre persone, fra cui la Dr.ssa Rosalind Picard del Media Laboratory del MIT per i suoi puntuali commenti. Fra gli altri includo i professori David Cranston, Danny Crookes, Jeremy Gibbons, il Dr. David Glass e il mio, sempre prezioso assistente di ricerca, il Dr. Simon Wenham.

Le mie competenze personali riguardano essenzialmente la matematica e la filosofia della scienza, non l'IA e il lettore, soprattutto quello esperto in materia, potrebbe rimanere perplesso di fronte al fatto che io decida di "invadere" il campo altrui. Chiarisco subito che le mie intenzioni, poiché mi sembra che ci siano vari livelli di approccio e coinvolgimento rispetto al tema dell'IA: ci sono i pensatori "pioneristici" e gli esperti che effettivamente elaborano i software usati nei sistemi di IA; poi ci sono quelli che comprendono le potenzialità dell'IA e che lavorano a nuove applicazioni; e infine ci sono altri che, anche in assenza di una formazione scientifica rigorosa, si occupano di divulgare considerazioni sul valore e l'impatto sociale, economico ed etico dell'IA.

È chiaro che non è necessario sapere come si costruisce un veicolo o un'arma autonomi per potersi formare un'opinione sulle conseguenze etiche del loro utilizzo. Non serve saper programmare un sistema di tracciamento degli acquisti artificiale per farsi un'opinione sulla questione della privacy.

Infatti, è grande l'attenzione da parte dei lettori più riflessivi nell'ambito della divulgazione scientifica a livello pubblico. È proprio a questo livello che vorrei posizionare il presente libro e sono debitore verso tutti quelli che, da addetti ai lavori, hanno già trattato questi argomenti nei loro scritti.

Capitolo 1

I CONFINI DEL DIBATTITO

Noi umani siamo insaziabilmente curiosi. Sin dagli albori della storia, ci poniamo domande riguardanti le nostre origini e il destino finale: “Da dove provengo? Dove sto andando?”. L'importanza di queste domande è ovvia e, mentre la risposta alla prima modella la concezione di chi siamo, la seconda ci fornisce gli obiettivi per cui vivere. Prese assieme, le risposte a queste domande fissano la nostra visione del mondo, la narrazione che dà significato alla nostra esistenza.

Il problema è che queste non sono domande semplici, e ce ne rendiamo conto dal fatto che vengono fornite molte risposte, spesso contraddittorie. Ma questo non ha rappresentato per gli uomini un ostacolo poiché, nel corso dei secoli, sono state formulate risposte suggerite dalla scienza, dalla filosofia, dalla religione, dalla politica, etc.

Due dei più famosi scenari futuristici sono descritti nei romanzi *Il mondo nuovo*, di Aldous Huxley (pubblicato nel 1931), e *1984*, di George Orwell (pubblicato nel 1949).

Entrambi i libri sono stati considerati, in tempi diversi, molto influenti: per esempio, nel 2005, la rivista *Time*, ha inserito *1984* nell'elenco dei cento migliori romanzi in lingua inglese pubblicati fra il 1923 e il 2005. Entrambi i racconti sono distopici, infatti, secondo la definizione dell'*Oxford English Dictionary*, essi “descrivono un luogo o una condizione immaginaria che è la più brutta possibile”, in quanto “si prefigurano situazioni, sviluppi, assetti politico-sociali e tecnologici altamente negativi”.^{*} In ogni caso, gli scenari assai inquietanti che sono descritti nei due romanzi risultano tutt'altro che simili e le loro differenze, sulle quali torneremo in seguito, sono state brevemente spiegate dal sociologo Neil Postman nel suo libro *Divertirsi da morire. Il discorso pubblico nell'era dello spettacolo*:

Orwell immagina che saremo sopraffatti da un dittatore. Nella visione di Huxley non sarà il Grande Fratello a toglierci l'autonomia, e a privarci della cultura e della storia. La gente sarà felice di essere oppressa e adorerà la tecnologia che libera dalla fatica di pensare.

Orwell temeva che i libri sarebbero stati banditi; Huxley, non che i libri fossero vietati, ma che non ci fosse più nessuno desideroso di leggerli. Orwell temeva quelli che ci avrebbero privato delle informazioni; Huxley, quelli che

* Secondo la definizione di “*distopia*” in <http://www.treccani.it/vocabolario/distopia2/>. N.d.E.

*ce ne avrebbero fornite troppe, fino a ridurci alla passività e all'egoismo. Orwell temeva che la verità sarebbe stata manipolata; Huxley che la verità sarebbe diventata irrilevante. Orwell temeva che la nostra sarebbe stata una civiltà di schiavi; Huxley, che sarebbe stata una cultura di cialtroni... In breve, Orwell temeva che saremmo stati distrutti da ciò che odiamo, Huxley, da ciò che amiamo.*¹

Orwell ha introdotto le idee di una sorveglianza a tappeto in uno stato totalitario, di “controllo mentale” e di una “neolingua”; oggi queste si legano sempre più allo sviluppo dell’intelligenza artificiale, soprattutto con il tentativo di elaborare una tecnologia informatica in grado di replicare le potenzialità della mente umana attraverso lo sviluppo di una mente imitativa. Non sorprende, infatti, che vengano investiti miliardi di dollari nella ricerca di sistemi di IA, poiché c’è un grande interesse su dove tutto questo porterà. Per esempio, da un lato, il miglioramento della qualità della vita in virtù dell’assistenza digitale, dell’innovazione medica e del potenziamento umano, mentre dall’altro ci sono i timori per la perdita di posti di lavoro e la comparsa di uno scenario orwelliano legato a una capillare sorveglianza delle società.

Molti dei traguardi raggiunti finora nel campo dell’IA sono riconducibili alla costruzione di sistemi in grado di fare una sola cosa che, di norma, richiederebbe comunque l’uso dell’intelligenza umana. Ma il lato che si presta maggiormente

1. Neil Postman, *Amusing Ourselves to Death: Public Discourse in the Age of Show Business*, Penguin, New York, 2006, pp. XIX-XX (trad. it. *Divertirsi da morire. Il discorso pubblico nell'era dello spettacolo*, Marsilio, 2002).

alla speculazione (al momento) è quello della ricerca, molto più ambiziosa, tesa ad approntare sistemi in grado di fare tutto ciò che è prerogativa dell'intelligenza umana, vale a dire una *Intelligenza Artificiale Generale* (IAG) che, sostengono alcuni, sorpasserà l'intelligenza umana in un tempo relativamente breve, intorno al 2084 o forse prima. Alcuni ritengono che l'IAG, se mai ci arriveremo, agirà come un dio, mentre altri intravedono in essa l'immagine di un despota totalitario.

Per introdurre questi argomenti che suscitano un sempre più grande interesse, con tutte le speranze e i timori che originano, faccio riferimento a tre bestseller contemporanei. I primi due sono stati scritti dallo storico israeliano Yuval Noah Harari; in *Sapiens. Da animali a dei: breve storia dell'umanità*, viene trattata, come si può comprendere dal titolo, la domanda sull'origine dell'umanità; in *Homo Deus. Breve storia del futuro*, invece, il tema è chiaramente il futuro dell'umanità. Il terzo libro, *Origin* di Dan Brown, è un romanzo come quelli di Huxley e Orwell e si concentra sull'uso dell'IA per rispondere alle nostre due domande. Ci si aspetta che sarà letto da milioni di persone (soprattutto giovani) che saranno influenzate dalle idee proposte sotto forma di thriller. Poiché il libro riflette chiaramente la prospettiva del suo autore riguardo questi temi, rappresenta un interessante punto di partenza per la nostra ricerca.

Inoltre, sono consapevole che per alcuni la fantascienza abbia rappresentato lo stimolo giusto per intraprendere la propria carriera, volta allo studio della scienza vera e propria. In ogni caso, non va trascurato un fatto importante che ci richiama alla cautela: Dan Brown afferma di essersi avvalso della scienza per giungere alle proprie conclusioni e quindi,

sebbene il libro sia un'opera di fantasia, dobbiamo soppesare attentamente le argomentazioni e le conclusioni cui giunge l'autore per essere in grado di fornire un'analisi veritiera.

Questo è molto importante se consideriamo che, per ammissione stessa del Brown, la motivazione di base che ha ispirato il suo romanzo era quella di affrontare una domanda fatidica: "Dio sopravvivrà alla scienza?". È la stessa domanda che, in varie forme, mi ha personalmente indotto a scrivere diversi libri. Il mio lavoro di ricerca mi ha portato alla conclusione che senza dubbio Dio sopravvivrà alla scienza, ma, d'altro canto, mi ha fatto sorgere un dubbio, vale a dire se l'ateismo riuscirà a sopravvivere alla scienza.

Uno dei personaggi principali del libro *Origin* di Dan Brown, Edmond Kirsch, è un informatico miliardario, esperto in intelligenza artificiale, che afferma di aver trovato la risposta alla domanda sulle origini della vita e sul destino dell'uomo. Kirsch vuole usare questi risultati per raggiungere l'obiettivo di un'intera vita, cioè "utilizzare la verità scientifica per distruggere il mito delle religioni",² in particolare le tre fedi abramitiche: ebraica, cristiana e islamica, concentrandosi, quasi inevitabilmente, sul cristianesimo. Le sue risoluzioni, una volta presentate al mondo, si riveleranno fondate sulla sua competenza nel campo dell'intelligenza artificiale, mentre la sua visione del futuro implica la mutazione tecnologica degli esseri umani.

È utile notare che l'idea che l'umanità sarà cambiata dalla tecnologia non è auspicata unicamente da storici o da scrittori di romanzi, ma da alcuni dei più rinomati scienziati. Per

2. Dan Brown, *Origin*, Doubleday, New York, 2017, p. 53 (trad. it. *Origin*, Mondadori, Milano, 2018).

esempio, l'astronomo britannico Lord Rees afferma: “Possiamo essere certi che le intelligenze dominatrici dei prossimi secoli non avranno alcuna comprensione emotiva dell'essere umano, nonostante possano avere una comprensione algoritmica del nostro comportamento”³

Sulla stessa linea, Rees ha anche affermato: “La capacità di pensare astrattamente, da parte dei cervelli biologici, è ciò che ha permesso l'emergere della cultura e della scienza. Ma questa attività, che finirà per coprire decine di millenni, sarà un precursore delle più potenti intelligenze dell'era post-umana inorganica. In un futuro lontano, le menti che riusciranno a comprendere pienamente l'universo non saranno quelle umane, ma quelle delle macchine”⁴

Questo argomento sarà sempre centrale poiché interessa non soltanto chi è direttamente coinvolto nella ricerca nel campo dell'IA, ma anche i matematici e gli scienziati di altre discipline i cui studi risentono sempre più dell'influenza di queste tematiche. A conferma di tutto ciò, poiché le idee e gli effetti della ricerca sull'IA finiranno per condizionare inevitabilmente la vita di ciascuno di noi, molte persone, pur non essendo scienziati, si stanno cimentando nello sviluppare riflessioni e vari approfondimenti sul tema. Le implicazioni sono tali che richiedono il coinvolgimento di filosofi,

3. Martin Rees, *On the Future: Prospects for Humanity*, Princeton University Press, Princeton, NJ, 2018, p. 7.

4. Martin Rees, “Astronomer Royal Martin Rees: How Soon Will Robots Take Over the World?”, in *The Telegraph*, 23 Maggio 2015, www.telegraph.co.uk/culture/hay-festival/11605785/Astronomer-Royal-Martin-Rees-predicts-the-world-will-be-run-by-computers-soon.html.

eticisti, teologi, esegeti, romanzieri e artisti, nell'ambito di un dibattito veramente ampio. Dopotutto, non bisogna essere fisici nucleari o climatologi per discutere dell'impatto dell'utilizzo dell'energia nucleare o dei cambiamenti climatici.

CHE COS'È L'INTELLIGENZA ARTIFICIALE?

Prendiamo in considerazione i robot. La parola *robot* deriva da *robota*, una parola ceca (e russa) che significa *lavoro*. Un robot è una macchina progettata e programmata da un'intelligenza umana per assolvere, in linea di massima, un singolo compito che, normalmente, dovrebbe svolgere un essere umano intelligente. In questo senso, il comportamento del robot simula l'intelligenza umana, circostanza che ha fatto sorgere un importante dibattito sull'opportunità di considerare o no il robot "intelligente", anche se quell'intelligenza non è ciò che noi intendiamo quando pensiamo alle capacità cognitive dell'essere umano (ma questa è già, di per sé, un'altra questione).

Il termine "IA" fu coniato, nel 1956, durante un campus estivo tenutosi nel dipartimento di matematica della Dartmouth University organizzato da John McCarthy, il quale affermò che: "L'IA è il lavoro scientifico e ingegneristico volto a creare macchine intelligenti"⁵ Il termine ora viene usato per descrivere sia le macchine intelligenti sia le conoscenze scientifiche e tecnologiche che mirano alla costruzione di quest'ultime.

5. John McCarthy, *What Is Artificial Intelligence?* www-formal.stanford.edu/jmc/whatisai.pdf.

La ricerca in quest'area si sviluppa lungo due direttrici principali: in generale, la prima coincide con il tentativo di comprendere i processi del pensiero e del ragionamento umani mediante modelli generati al computer. La seconda si concretizza nello studio del comportamento umano e nel tentativo di costruire delle macchine in grado di imitarlo. La differenza è sostanziale: una cosa è creare una macchina capace di “simulare” una mano umana che alza un oggetto, altra cosa è creare una macchina che sia in grado di replicare i pensieri di un essere umano mentre alza un oggetto. È molto più facile creare la prima rispetto alla seconda e, se l'utilità è l'obiettivo cui si tende, la simulazione è ciò cui dobbiamo mirare. Facciamo un esempio: l'industria aeronautica è tesa alla realizzazione di macchinari che volano, ma non implica la costruzione di un cervello elettronico simile a quello di un uccello affinché l'aereo voli esattamente come farebbero gli uccelli, vale a dire battendo le ali.⁶

L'idea di costruire macchine che possano simulare aspetti del comportamento umano (e animale) ha alle spalle una lunga storia. Duemila anni fa, il matematico greco Erone di Alessandria costruì una vasca adornata da alcuni uccellini meccanici che cantavano e da un gufo che, quando girava la testa, faceva cessare il canto degli uccellini. Nel corso dei secoli, in tanti sono rimasti affascinati dall'idea di creare rudimentali robot (detti automata) in grado di replicare alcuni aspetti della vita. Importanti collezioni di esemplari molto sofisticati di questi automata possono essere ammirate al Museo della Scienza di Londra, al Museo di Storia dell'Arte

6. Si veda, Stuart Russell e Peter Norvig (a cura di), *Artificial Intelligence: A Modern Approach*, Harlow: Pearson Education, 2016, pp. 1-5.

di Vienna e al Museo Speelklok di Utrecht. Sebbene nel diciannovesimo secolo l'interesse per la costruzione di queste macchine andò a scemare, rimase vivo nelle opere di fantasia come il romanzo *Frankenstein* di Mary Wollstonecraft Shelley del 1818, che rappresenta una pietra miliare della fantascienza sin dagli albori di questo genere narrativo.

Una delle più importanti attività umane quotidiane è il calcolo numerico e sono stati fatti grandi sforzi per automatizzare questo processo: nel diciassettesimo secolo, il matematico francese Blaise Pascal realizzò un calcolatore meccanico⁷ per aiutare suo padre, un ufficiale delle tasse, con i noiosi conteggi. Nel diciannovesimo secolo, Charles Babbage pose le basi della computazione programmabile inventando prima la macchina differenziale, che svolgeva le addizioni e, in seguito, la macchina analitica, che fu il primo calcolatore programmabile. Per questo motivo Babbage è riconosciuto come il padre dei moderni computer.

Durante la seconda guerra mondiale, il geniale scienziato inglese Alan Turing utilizzò una sofisticata tecnologia informatica per realizzare uno strumento, noto come la “Bomba”, che permise a lui e alla sua squadra di Bletchley Park di decifrare il codice tedesco “Enigma”, utilizzato per le comunicazioni militari segrete. Le invenzioni di Turing e il suo lavoro teorico hanno condotto alle concezioni del “learning machine”, vale a dire “all'apprendimento automatico”. Secondo Turing, una macchina in grado di conversare con gli umani, senza che questi ultimi si rendano conto che si tratti di una macchina,

7. La prima macchina calcolatrice fu costruita intorno al 1623 dal tedesco Wilhelm Schickard, professore di ebraico e di astronomia.

potrebbe superare il cosiddetto “imitation game” (meglio noto come “Test di Turing”) ed essere considerata intelligente. Ad ogni modo, come vedremo in seguito, questo approccio ha incontrato alcune obiezioni da parte di diversi filosofi.

Nello stesso periodo (1951) Marvin Minsky (cofondatore del laboratorio di ricerca sulla IA del MIT) e Dean Edmonds, realizzarono il primo computer a rete neurale. Altri storici traguardi che attirarono l'attenzione pubblica furono il computer Deep Blue della IBM che nel 1997 batté il campione del mondo di scacchi Garry Kasparov, e il programma di Google AlphaGo che, nel 2016, sconfisse un giocatore professionista di “Go”. L'importanza dell'IA è stata riconosciuta con l'assegnazione, nel 2018, del Premio Turing (considerato il “Nobel dell'informatica”) a un trio di ricercatori che hanno gettato le basi dell'attuale progresso dell'IA, in particolare nel campo del “deep learning” (“apprendimento profondo”, branca dell'apprendimento automatico).

I primi robot e i primi sistemi di intelligenza artificiale erano sprovvisti di ciò che chiamiamo “apprendimento automatico”, alla cui base c'è l'idea dell'algoritmo che può essere di vario tipo (simbolico, matematico, etc....).⁸ La parola “algoritmo” deriva dal nome del famoso matematico, astronomo e geografo persiano Muhammad ibn Mūsā al- Khwārizmī (ca. 780-850).⁹

* Gioco da tavolo strategico di origine orientale. N.d.E.

8. È utile sottolineare che i primi sistemi di IA non ricorrevano agli algoritmi.

9. Si veda Jeffrey Aronson, “When I Use a Word... Algorithms”, in *BMJ Opinion*, 11 Agosto 2017, <https://blogs.bmj.com/bmj/2017/08/11/jeffrey-aronson-when-i-use-a-word-algorithms>.

Oggi un algoritmo viene definito come “procedimento di calcolo esplicito e descrivibile con un numero finito di regole che conduce al risultato dopo un numero finito di operazioni, cioè di applicazioni delle regole [...] insieme di istruzioni che deve essere applicato per eseguire un’elaborazione o risolvere un problema”.¹⁰ Il concetto stesso di algoritmo viene fatto risalire, però, ai tempi dell’antica Babilonia (1800-1600 a.C.) e l’eminente informatico Donald Knuth della Stanford University, pubblicando questi algoritmi babilonesi, affermò che “i calcoli descritti nelle tavolette babilonesi non sono propriamente delle soluzioni a problemi individuali specifici, ma si tratta di procedure generali per risolvere una intera categoria di problemi”.¹¹ Questa è la caratteristica principale di un algoritmo: quando se ne comprende l’operatività, lo si può usare per risolvere intere classi di problemi e non solamente uno.

Uno dei più famosi esempi (che abbiamo studiato anche a scuola) è quello dell’algoritmo euclideo, cioè la procedura usata per trovare il massimo comun divisore (MCD) fra due numeri interi positivi. Questo fu descritto per la prima volta da Euclide nel suo scritto “*Elementi*” intorno al 300 a.C. È un algoritmo efficiente che, in vari modi, è usato ancora oggi dai computer. Esso consiste in una serie di divisioni e calcolo dei resti fino a raggiungere il risultato. Questa operazione si può comprendere al meglio con il seguente esempio, sebbene sia importante ricordare che l’algoritmo si applica a una coppia di numeri interi.

10. Secondo la definizione di “algoritmo” presente in <http://www.treccani.it/vocabolario/algoritmo/>.

11. Donald E. Knuth, “Ancient Babylonian Algorithms”, in *Communications of the ACM*, vol. 15, n. 7 (luglio 1972), pp. 672–73.

Supponiamo di dover calcolare il MCD fra 56 e 12. Questi sarebbero i passaggi da seguire:

1. *Passaggio 1* - Dividere il numero più grande per quello più piccolo:
 $56 \div 12 = 4$, con il resto di 8.
2. *Passaggio 2* - Dividere il dividendo, 12, per il resto ottenuto nel passaggio precedente:
 $12 \div 8 = 1$, con il resto di 4.
3. *Passaggio 3* - Continuare il passaggio 2 fino a quando non ci saranno più resti (in questo caso c'è soltanto un passaggio in più):
 $8 \div 4 = 2$ (nessun resto).

In questo caso il MCD è 4.

È facile tradurre questo processo in un codice di programmazione software e implementarlo in un computer. Una rapida ricerca su internet vi farà scoprire che sono migliaia gli algoritmi di ogni tipo che oggi vengono usati in qualsiasi ramo della scienza, dell'ingegneria, della medicina. La robotica è un esempio fra i principali, poiché i robot sono in genere progettati per eseguire un singolo compito a ripetizione.

In un tipico sistema di IA contemporaneo, gli algoritmi rilevanti sono incorporati nei software che classificano, filtrano e selezionano frammenti dei vari dati che in esso vengono introdotti. Uno degli approcci utilizzati prevede, in qualche maniera, la simulazione delle funzioni neurali nella corteccia umana (le reti neurali). In termini generali, un sistema di que-

sto tipo può utilizzare i dati per “apprendere” (apprendimento automatico, machine learning)¹² per riconoscere, identificare e interpretare modelli digitali quali immagini, suoni, voci, testi o dati. Un altro approccio usa applicazioni informatiche che si rifanno alla logica della *probabilità bayesiana** per analizzare, da una prospettiva statistica, le informazioni disponibili al fine di stimare la probabilità di una particolare ipotesi. In poche parole, un sistema di apprendimento automatico incamera le informazioni pregresse e, sulla base di queste, prende decisioni o elabora predizioni quando vengono introdotte nuove informazioni.

È importante notare che gli algoritmi sono programmati esplicitamente per ricorrere all'approssimazione attraverso l'ottimizzazione numerica e, in molti casi, sono provvisti di esempi di input e output e di criteri di arresto come guida per l'ottimizzazione. Gli algoritmi non “macinano chilometri” senza l'intervento umano che guida l'intero processo. L'intervento umano è cosciente, quello della macchina no.

12. L'*Enciclopedia dell'Intelligenza Artificiale* definisce il “machine learning” (apprendimento automatico) come “un'area dell'Intelligenza Artificiale relativa ai metodi adattivi computazionali quali le Reti Neurali Artificiali e gli Algoritmi Genetici”. Juan Ramon Rabunal Dopico, Julian Dorado de la Calle, Alejandro Pazos Sierra, a cura di, *Encyclopedia of Artificial Intelligence*, Information Service Reference, Hershey, PA, 2009, p. 666.

* La *probabilità bayesiana* è un'interpretazione del concetto di probabilità, in cui, anziché la frequenza o la propensione di qualche fenomeno, la probabilità è interpretata come aspettazione razionale. Il termine *bayesiano* deriva dal matematico e teologo del XVIII secolo Thomas Bayes, che fornì la prima trattazione formale di un problema non banale riguardante l'analisi dei dati statistici usando quella che è ora nota come inferenza bayesiana. N.d.E.

Nei primi prodotti di intelligenza artificiale venivano implementati degli algoritmi per la risoluzione di singoli problemi, ma con la più recente IA questo non avviene più: si utilizzano degli algoritmi di apprendimento generale, i quali, poi, “apprendono” la soluzione applicandola al problema. Spesso gli sviluppatori umani non conoscono l’effettivo algoritmo adoperato per risolvere il problema e non sanno in che modo il sistema sia arrivato a quella conclusione. I vecchi programmi che giocavano a scacchi erano del primo tipo descritto (anche Deep Blue), mentre il software che gioca a “Go” è del secondo tipo.

Facciamo degli esempi di sistemi di IA, alcuni dei quali già noti al grande pubblico:

- Amazon usa algoritmi che tracciano tutti i prodotti che milioni di persone comprano online. Analizzando questo enorme database, vengono confrontate le liste dei nostri acquisti con prodotti simili che ancora non possediamo. Di fatto, gli algoritmi usano metodi statistici per selezionare quei prodotti acquistati da persone “come te” e finiscono per essere evidenziati durante la nostra navigazione.
- Un algoritmo informatico può vagliare un database composto da candidature di lavoro e suggerire gli individui più adatti per una determinata posizione lavorativa. Sono soggetti all’analisi da parte di intelligenze artificiali soprattutto quei posti di lavori che richiamano il maggior numero di candidati: l’IA conduce i primi colloqui nei quali non si limita a registrare i dati del candidato, le sue risposte alle domande, ma anche le sue

reazioni emotive che vengono filmate e analizzate per determinare l'esito di quello specifico colloquio.

- L'IA viene impiegata con successo nella progettazione di edifici più efficienti dal punto di vista energetico, di elettrodomestici connessi con "l'internet delle cose", e di sistemi di trasporto integrati.
- I sistemi di Intelligenza Artificiale sono già adoperati in campo medico attraverso l'analisi di database riguardanti migliaia di lastre ai polmoni (per avere uno storico dei differenti stati clinici in cui i polmoni possono trovarsi), e fornendo anche una analisi medica professionale sullo stato di salute dei polmoni stessi. Il sistema confronta le lastre dei vostri polmoni con quelle presenti nel database per capire se siete affetti da una determinata malattia. Più precisamente, il sistema tende a creare delle statistiche a partire dal modello visivo dell'immagine ai raggi x e lo confronta con altri modelli estratti dal database. Questo tipo di sistemi, in certi casi, può effettuare delle diagnosi più accurate rispetto a quelle dei medici in carne e ossa.
- Gli astronomi hanno utilizzato l'intelligenza artificiale per "insegnare" alle macchine a identificare i lampi radio veloci provenienti da galassie lontane attraverso l'analisi di grandi database di segnali raccolti dai radiotelescopi. Hanno già scoperto (dati del settembre 2018) settantadue nuovi campioni che saranno studiati per le ricerche sulle intelligenze extraterrestri. Inoltre, l'IA viene usata anche per il riconoscimento automatico delle galassie.
- I veicoli a guida autonoma, essendo macchine non coscienti, hanno generato, fin da subito, problemi etici le-

gati ai principi con i quali vengono programmati, soprattutto riguardo a ciò che questi veicoli dovrebbero cercare di evitare. È interessante notare che i sistemi di auto pilotaggio degli aerei non sono completamente autonomi, questo perché le compagnie aeree non si assumono un simile rischio in quanto, in caso di disastro aereo, ci sarebbero ben pochi dubbi su chi ricada la responsabilità, malgrado il sistema abbia agito in modo da far morire alcune persone per salvarne altre.

- I sistemi di riconoscimento facciale sono molto sviluppati. Un'applicazione piuttosto divertente è l'uso della tecnologia di riconoscimento facciale mediante IA in un pub per individuare chi sarà il prossimo a essere servito al bancone per poter, quindi, evitare il fenomeno dei furbetti che saltano la fila. Oggi le telecamere a circuito chiuso sono molto diffuse e servono alle forze di polizia per controllare le attività criminali. Ma questo tipo di sorveglianza può essere usato anche come forma di controllo sociale; vedremo in seguito le maggiori implicazioni etiche che scaturiscono dall'utilizzo di queste applicazioni.
- I sistemi di armi autonome, cioè in grado di scegliere e attaccare l'obiettivo autonomamente, hanno sollevato notevoli problemi etici e sono oggetto di acceso dibattito a livello internazionale.

Questa breve lista è sufficiente per capire come questi progressi tecnologici facciano sorgere dilemmi etici che vanno dalla condotta in campo finanziario al controllo sociale e della privacy. Il pericolo che s'intravede è quello del-

l'affermazione di una mentalità in forza della quale, se qualcosa può essere fatto allora deve essere fatto, trascurando la necessaria riflessione sui risvolti etici che una tale scelta comporta. Ad ogni modo, va riconosciuto che le questioni etiche occupano un'importanza sempre maggiore per chi è coinvolto nello sviluppo dell'IA. Una, in modo particolare, è rappresentata da una questione di ordine preliminare: come si può inserire la dimensione etica in un algoritmo che, di per sé, è privo di cuore, anima e mente?

Molti dei sistemi di IA citati in precedenza, sono progettati per compiere uno specifico compito come guidare, diagnosticare una malattia, o fare delle previsioni sulla base dei dati raccolti; in casi come questi si parla di intelligenza artificiale debole. Ma, poiché tutte le varie applicazioni in campo di IA sono, almeno fino ad ora, catalogabili come deboli, si usa il termine di "tecnologie cognitive" per indicare i progressi ottenuti nel campo delle machine intelligenti.

Il filosofo svedese Nick Bostrom e il ricercatore e scrittore americano Eliezer Shlomo Yudkowsky hanno commentato:

Gli algoritmi delle intelligenze artificiali implementati fino ad oggi, che eguagliano o superano le capacità umane, sono caratterizzati dal fatto di essere volutamente programmati con una competenza ristretta a un singolo ambito. Deep Blue, per esempio, divenne campione mondiale di scacchi, ma non sapeva giocare a dama né guidare un veicolo né fare una scoperta scientifica. I moderni algoritmi somigliano alla vita biologica ma non all'essere umano: un'ape mostra competenza nel costruire un alveare; un castoro è competente nel costruire dighe; ma un'ape non co-

*struisce dighe e un castoro non sa creare un alveare. Un uomo, invece, dall'osservazione, può imparare a fare entrambe le cose, e questa rimane un'abilità unica tra gli esseri viventi.*¹³

Un'altra fonte di confusione nel dibattito può essere rappresentata da parole di uso quotidiano quali *apprendimento*, *pianificazione*, *ragionamento*, *intelligenza* usate, dagli informatici, per descrivere tecnicamente le qualità di una macchina. Questo ci induce a considerare i sistemi di intelligenza artificiale più sofisticati di quanto non siano realmente, poiché questi termini sono usati in un senso molto più ristretto rispetto a quello di uso comune. La conseguenza è questa: i media tendono a “drammatizzare” i risultati ottenuti in questo campo apparendo estremamente ottimisti o decisamente catastrofici. Il professore di Ingegneria Informatica Danny Crookes della Queens University di Belfast scrive:

Le attuali tecnologie che stanno cominciando a preoccupare le persone per la loro capacità di controllare e manipolare le masse non sono realmente molto intelligenti. Anzi, non devono esserlo poiché la loro forza sta nel riuscire a gestire grandi quantità di dati, creando profili individuali e individuando dei modelli sia nel comportamento di un individuo sia in quello di una popolazione. I regimi nazisti e

13. Nick Bostrom, Eliezer Yudkowsky, “The Ethics of Artificial Intelligence”, in *Cambridge Handbook of Artificial Intelligence*, a cura di Keith Frankish, William M. Ramsey Cambridge University Press, Cambridge, 2014, p. 318.

*comunisti facevano tutto questo manualmente su una scala minore, mentre oggi la tecnologia permette di farlo su scala globale. Per quanto sia notevole e preoccupante, non si tratta di vera e propria intelligenza. Il cosiddetto “apprendimento profondo” è ciò che in questo momento va di moda nella ricerca sull’IA, ma non c’è nulla di particolarmente inedito: ora la potenza di calcolo è in grado di gestire reti neurali multi-strato (profonde) che esistono, sulla carta, da decenni.*¹⁴

Il professor Joseph McRae Mellichamp, dell’Università dell’Alabama, ha espresso il suo punto di vista intervenendo a una conferenza all’Università di Yale alla quale era presente, fra vari pionieri dell’intelligenza artificiale, anche il premio Nobel Sir John Eccles (famoso per la scoperta delle sinapsi): “Ritengo che si potrebbero evitare parecchi dibattiti se i ricercatori dell’IA ammettessero che ci sono differenze fondamentali fra l’intelligenza della macchina e quella umana, differenze che nessuna ricerca potrà colmare”. In altre parole, per parafrasare un intervento¹⁵ di Mellichamp, quando si parla di “intelligenza artificiale”, la dimensione “artificiale” è assolutamente reale.

14. Da una conversazione privata con l’autore (2019), riportata con autorizzazione.

15. Da una conversazione privata (2018) su uno scritto presentato alla Artificial Intelligence and Human Mind conference (Yale University, 1986), riportata con autorizzazione. Il resoconto di quella conferenza è consultabile qui: www.aaai.org/ojs/index.php/aimagazine/article/view/601.

Il professor Crookes pone enfasi sulla necessità di essere realisti in proposito:

Siamo ancora molto, molto lontani dal creare un'intelligenza simil-umana. L'impatto delle elaborazioni a base di dati ha portato le persone a pensare che ci si stia avvicinando al livello dell'intelligenza umana ma, secondo me, non ci siamo nemmeno vicini. Anzi, si potrebbe quasi dire che il vero e proprio progresso in campo di IA sia rallentato negli ultimi anni e che ci sia minore ricerca realmente indirizzata verso la IA, poiché la maggior parte degli investimenti vengono fatti nel campo della pubblicità e i ricercatori seguono il denaro.

Le sfide riguardanti la comprensione dei processi cognitivi umani sono enormi e, per quel che vale, riesco a individuare due problemi fondamentali che devono ancora essere superati: 1) anche se fossimo in grado di descrivere le regole con cui avviene il ragionamento umano, come potremmo riuscire a ricavarne una formulazione astratta da applicare come regola generale che descriva il ragionamento? 2) In che modo un computer può creare e sostenere un modello mentale interno del mondo reale? Come fa una persona cieca a visualizzare il mondo e a ragionare su di esso? Gli uomini hanno l'abilità versatile di visualizzare oggetti esterni, di ragionare su di essi e sulla loro esistenza. Questa versatilità, comune a tutti gli esseri umani, è fenomenale: è il requisito chiave della vera intelligenza, ma è assente nei sistemi di IA.

Ciò che voglio dire è che dobbiamo stare attenti anche solo a supporre che l'umanità abbia la capacità intellettuale

*di creare un'intelligenza che rivaleggi con quella umana, per non parlare poi di superarla, non importa quanto tempo abbiamo a disposizione.*¹⁶

Dobbiamo tenere bene a mente questi ultimi concetti, ora che analizzeremo il modo in cui Dan Brown impiega nel suo romanzo l'intelligenza artificiale (debole) per risolvere le due grandi questioni filosofiche.

16. Da una conversazione privata (2019).

INDICE

<i>Prefazione dell'Editore italiano</i>	5
<i>Introduzione</i>	7
1. I confini del dibattito	9
2. La prima grande domanda: da dove veniamo?	31
3. La seconda grande domanda: dove stiamo andando?	47
4. Intelligenza artificiale debole: il futuro è davvero luminoso?	59
5. Intelligenza artificiale debole: forse il futuro non è così luminoso	71
6. Il potenziamento umano	95
7. Intelligenza artificiale generale: il futuro è oscuro?	109
8. Le verità del libro di Genesi: che cos'è un essere umano?	131
9. Le origini del senso morale umano	157
10. Il vero <i>Homo Deus</i>	183
11. Lo shock futuro: il ritorno dell'Uomo che è Dio	199
12. <i>L'Homo Deus</i> nel libro di Apocalisse	223
13. Il tempo della fine	243