

Quant | Q1

Andrea Daniele Signorelli

Rivoluzione Artificiale

**L'uomo nell'epoca
delle macchine intelligenti**

Ledizioni

© Ledizioni LediPublishing
Via Alamanni 11 Milano
<http://www.ledizioni.it>
e-mail: info@ledizioni.it

Prima edizione Ledizioni: Aprile 2019
Il volume è stato pubblicato originariamente da Informant nel mese di
giugno 2017

Andrea Daniele Signorelli, *Rivoluzione artificiale. L'uomo nell'epoca delle
macchine intelligenti*

ISBN cartaceo 9788867059386
ISBN eBook 9788867059614

Responsabile Collana: Massimo Colasurdo
Illustrazioni di copertina: Luca Lorenzoni

Le riproduzioni a uso differente da quello personale potranno avvenire,
per un numero di pagine non superiore al 15% del presente volume, solo a
seguito di specifica autorizzazione rilasciata da Ledizioni, Via Alamanni 11 –
20141 Milano, e-mail: info@ledizioni.it

INDICE

1. L'intelligenza artificiale è tra noi **p. 7**

- 1.1 Dai filtri antispam ai robot avvocati: la AI è già una realtà
- 1.2 Le origini dell'intelligenza artificiale
- 1.3 Come imparano le AI: machine learning e deep learning
- 1.4 La catena di montaggio dell'intelligenza artificiale
- 1.5 L'evoluzione delle AI: il futuro prossimo
- 1.6 L'evoluzione delle AI: arriveremo alla superintelligenza artificiale?
- 1.7 I limiti qualitativi e tecnologici delle AI

2. Etica artificiale **p. 47**

- 2.1 La coscienza delle AI
- 2.2 Diritti umani per i robot?
- 2.3 Il diritto dei robot alla vita e alla procreazione
- 2.4 La libertà dalla schiavitù
- 2.5 Il diritto di parola e di voto
- 2.6 Le responsabilità di una AI

3. I pericoli per l'uomo **p. 75**

- 3.1 Il rischio esistenziale
- 3.2 Il pregiudizio dell'algoritmo
- 3.3 Se lo sbirro è un algoritmo
- 3.4 La grande corsa alle armi autonome
- 3.5 Il collasso della realtà

4. Il futuro del lavoro **p. 93**

- 4.1 I robot e la disoccupazione di massa
- 4.2 La via d'uscita del reddito universale
- 4.3 Le nuove utopie: l'Accelerazionismo
- 4.4 Le nuove utopie: il Comunismo Automatizzato di Lusso

1 L'INTELLIGENZA ARTIFICIALE È TRA NOI

1.1 DAI FILTRI ANTISPAM AI ROBOT-AVVOCATI: LA AI È GIÀ UNA REALTÀ

La domanda che Alan Turing si è posto nel 1950¹ (“Le macchine possono pensare?”) ottiene ancora oggi la stessa risposta di allora: no, non possono. Le si può però imbottire di una quantità tale di dati da costringerle a metterli in relazione tra loro identificando eventuali collegamenti. Un meccanismo che entra in azione, per fare un esempio, ogni volta che caricate una foto di gruppo su Facebook per pubblicarla e, immediatamente, il social network vi suggerisce nomi e cognomi delle persone che in quella foto compaiono. L'intelligenza artificiale di Facebook ha infatti analizzato tutti i dati a sua disposizione – in questo caso, le foto che avete postato nel corso degli anni e i profili dei vostri amici – utilizzandoli per fare delle previsioni sulla loro identità.

La tecnologia di Facebook per il riconoscimento facciale (*face recognition*) dimostra soprattutto una cosa: l'intelligenza artificiale è già qui, la vediamo all'opera ogni giorno, la utilizziamo in continuazione, spesso senza nemmeno saperlo. Certo, le AI (*artificial intelligence*) non hanno le sembianze di HAL 9000 o di Skynet, non sembrano essere sul punto di prendere coscienza di se stesse e di ribellarsi all'uomo dando vita a un nuovo ordine mondiale robotico; eppure sono una realtà che sta avendo un impatto enorme in ogni ambito, grazie all'abilità che accomuna ognuno di questi software: analizzare una miriade di dati per imparare a distinguerli e associarli; un meccanismo che permette a Gmail di individuare ed eliminare la spam, ad Amazon di suggerirvi i libri che potrebbero piacervi o a Google Assistant di organizzarvi il calendario.

1 Alan Turing, *Computing machinery and intelligence*, 1950. Il saggio è reperibile a questo link: <http://www.loebner.net/Prizef/TuringArticle.html>

L'esempio di Google Assistant porta inevitabilmente a parlare di una delle applicazioni più promettenti dell'intelligenza artificiale, sulla quale tutti i principali attori della Silicon Valley stanno investendo massicciamente: gli assistenti virtuali che rispondono alle nostre richieste ("mostrami dei film che potrebbero piacermi") e che sono in grado di anticipare i nostri desideri e necessità avendo imparato a conoscerci alla perfezione (ricordandovi, per esempio, che il mercoledì siete soliti fare una telefonata alla mamma). Google Assistant, Siri di Apple, Cortana di Windows e Alexa di Amazon sono solo alcuni degli esempi che dimostrano la diffusione degli assistenti personali, sia vocali come Alexa, sia *text-based* come Replika; con il quale interloquiamo attraverso un'interfaccia grafica simile a quella degli sms o dei servizi di *instant messaging*.

Descritti così, gli assistenti virtuali (soprattutto quelli che utilizzano il testo scritto) potrebbero non sembrare niente di rivoluzionario: i *chatbot*, i programmi che simulano il dialogo tra esseri umani, sono tutto tranne che una novità; anzi, richiamano direttamente la celebre ELIZA², software creato da Joseph Weizenbaum nel 1966 per fare la parodia di uno psicoterapeuta³. Se i bot che dialogano con le persone sono un'invenzione che ha più di 50 anni, per quale ragione oggi si pensa che possano essere la *next big thing* del mercato tecnologico?

La ragione è semplice: le potenzialità di questi *bot*, oggi, sono infinitamente superiori a quanto avrebbero potuto esserlo anche solo pochi anni fa; tutto merito di computer dalla potenza di calcolo sempre maggiore in grado di analizzare la quantità enorme di dati che viene prodotta quotidianamente attraverso il web, i social network e in generale la *internet of things*.

Siamo ancora agli inizi: gli assistenti testuali hanno bisogno di operatori umani che intervengono nei casi più complessi e ancora

2 Si può parlare con ELIZA a questo link: <http://psych.fullerton.edu/mbirnbaum/psych101/Eliza.htm>

3 Il chatbot è diventato famoso anche perché alle domande risponde sempre con ulteriori domande: per esempio, chiedere a ELIZA "Come stai?" provocherà sicuramente una sua risposta di questo tipo: "Perché sei interessato a sapere quale sia il mio umore?". Un espediente parodistico per imitare uno psicoterapeuta.

non ne esiste uno in grado di organizzarci interamente un viaggio o una vacanza; gli assistenti vocali come Alexa di Amazon o Google Home hanno invece ancora parecchi difetti e funzionalità limitate. Ma è solamente questione di tempo: grazie alle potenzialità dell'intelligenza artificiale (privacy permettendo), i *bot* del futuro succhieranno informazioni dai profili social, dalle app e dal nostro comportamento sul web, per imparare a conoscerci e rispondere in modo semplice, colloquiale e coerente alle nostre richieste.

Nonostante gli ambiti in cui l'intelligenza artificiale può trovare applicazione siano pressoché illimitati, sono soprattutto gli assistenti virtuali a suscitare attenzione (e un po' di inquietudine); forse perché sono ciò che più da vicino ricorda il modo in cui ci siamo immaginati i robot/software del futuro (basti pensare a *Her* di Spike Jonze), oltre a essere il settore sul quale sono piovuti investimenti enormi da parte di tutti i protagonisti della Silicon Valley (da Google a Facebook, da Apple ad Amazon, da Microsoft a IBM). O forse è che questa applicazione rappresenta l'esempio più semplice ed evidente di come la AI stia invadendo il mercato di massa e sia ormai coinvolta in ogni operazione quotidiana.

Ma la stessa tecnologia si sta facendo largo anche in ambito professionale, dove sono già all'opera software che assistono gli studi legali – analizzando database sterminati e trovando correlazioni con casi simili a quello su cui sta lavorando un avvocato – e software che assistono i medici per individuare con precisione estrema i primi segni di comparsa, per esempio, del tumore della pelle (com'è il caso di Watson, il software di IBM).

Tutto ciò pone una grossa sfida per la nostra società: se i robot e i software sono in grado di svolgere una vastissima gamma di lavori con una precisione anche superiore a quella umana (e con costi molto più ridotti), se il loro impiego non riguarda più solamente lavori scarsamente qualificati, ma anche professioni intellettuali o altamente specializzate, che fine faranno i lavoratori umani? Lo spettro della disoccupazione di massa si sta facendo largo nella nostra epoca, ponendo interrogativi che necessitano una risposta urgente.

Il punto è che fermare o anche solo rallentare questa "rivoluzione robotica" (che riguarda però anche software immateriali) è

speranza vana: il percorso sembra irreversibile. Il che non significa che l'uomo debba rassegnarsi a diventare superfluo, ma che deve probabilmente iniziare a immaginare una società diversa, in cui il lavoro non sia più l'aspetto dominante delle nostre vite. Una società post-lavoro, insomma, che potrebbe sorgere grazie a misure come il reddito universale (per non parlare delle tesi più estreme, come il Comunismo Automatizzato di Lusso).

L'avvento di intelligenze artificiali sempre più evolute non avrà un enorme impatto solo sul mondo del lavoro, ma ci costringerà ad affrontare importanti questioni etiche, richiamando in servizio attivo la filosofia. Non è un caso che, oggi, uno dei pensatori più ascoltati quando si parla di temi tecnologici sia un filosofo, un italiano a Oxford: Luciano Floridi. D'altra parte, i temi etici che riguardano il mondo delle AI non sono solo di grande rilevanza, ma anche molto pressanti.

Nel 2022, per esempio, potrebbero iniziare a circolare per le città le prime auto autonome commerciali, ovvero veicoli che si guidano da soli. Questo significa che sarà un software a decidere come comportarsi in caso di imprevisti: se sterzare all'improvviso per evitare un ostacolo, ma rischiando di investire dei pedoni, o se tirare dritto con il rischio di ferire il conducente. Siamo sicuri di volerci sedere in un veicolo che, in un certo senso e in alcune condizioni, sarebbe di fatto programmato per ucciderci?

Un timore che impallidisce davanti agli inquietanti progressi compiuti da software che stanno imparando a utilizzare la memoria, che scoprono come cifrare i loro messaggi e che sono sempre più vicini a formulare ragionamenti di tipo umano. Ma fino a dove possono arrivare le AI? Possono evolvere da algoritmi in grado di svolgere un solo compito per prendere la forma di *vere* intelligenze artificiali, capaci di competere con l'uomo sotto ogni punto di vista e magari, un giorno, lasciarselo alle spalle? La grande attenzione dedicata al tema da serissimi scienziati dimostra come preoccupazioni di questo genere non appartengano più solo al regno della fantascienza.

E allora, come reagiremo quando davanti a noi compariranno intelligenze artificiali che si comportano *come* esseri davvero intelli-

genti, *come* esseri che pensano, *come* – in definitiva – esseri dotati di coscienza? Dovremo indagare a fondo per capire se la loro sia vera coscienza o solo una simulazione di essa, oppure il fatto stesso che si comportino da esseri coscienti significa che debbano essere trattati come tali? E infine: se esseri artificiali intelligenti e coscienti inizieranno a circolare per le nostre città, non dovremo preoccuparci di garantire loro alcuni diritti basilari? Così facendo, però, rischieremo di far venir meno la ragione stessa per cui li stiamo creando: farli lavorare al servizio dell'uomo.

Gli interrogativi presentati dalla prossima era delle macchine coinvolgono gli aspetti più complessi della nostra società e mostrano come la politica sia in ritardo su un tema che rischia di diventare prioritario nel giro di pochi anni. Un'analisi dell'impatto che le intelligenze artificiali avranno sulla nostra società, però, non può che partire dalle basi; dal capire come questi algoritmi imparino a distinguere e associare dati diventando intelligenti al punto da sconfiggere il campione di un gioco complesso e astratto come Go e suscitando timori (più o meno giustificati) sulla possibilità che, un domani, robot super-intelligenti si possano liberare dalla schiavitù e soppiantare i padroni.

1.2 LE ORIGINI DELL'INTELLIGENZA ARTIFICIALE

Si parla sempre di *intelligenza* artificiale, ma il termine rischia di essere fuorviante e di far pensare a macchine in grado di ragionare, di conoscere una materia, di avere consapevolezza di ciò che stanno facendo. Il fatto che una macchina impari a riconoscere la presenza di un gatto nelle foto, però, non significa che sappia che cos'è un gatto; allo stesso modo, il computer che ha battuto Lee Sedol a Go⁴ non aveva la benché minima idea di che cosa stesse facendo (e lo stesso vale per lo storico, ma diverso, esempio riguardante le partite a scacchi tra Deep Blue di IBM e Gary Kasparov).

4 <http://qz.com/639952/googles-ai-won-the-game-go-by-defying-millennia-of-basic-human-instinct/>

Questi software, insomma, non sono in grado di pensare; sono semplicemente capaci di processare una quantità tale di dati da riuscire a metterli in relazione tra loro, identificando eventuali collegamenti o calcolando statisticamente quale mossa di un determinato gioco abbia la maggior probabilità di avere successo. I metodi utilizzati per ottenere questi risultati sono principalmente il *machine learning* (apprendimento automatico) e la sua più recente evoluzione, il *deep learning* (apprendimento approfondito), una branca del *machine learning* (sviluppata soprattutto grazie al lavoro svolto da scienziati informatici come Andrew Ng e Geoffrey Hinton) che sfrutta numerosi strati di elaborazione per scomporre gerarchicamente i dati su livelli di astrazione crescente.

L'intelligenza artificiale è una realtà del presente che segnerà indelebilmente il nostro futuro. Eppure, le aspettative riposte oggi in questa tecnologia sono, incredibilmente, le stesse che ci si attendeva sessant'anni fa. Nel 1958, un articolo del New York Times⁵ presentava infatti una nuova meraviglia tecnologica, titolando: «Il cervello elettronico che insegna a se stesso: nel giro di un anno sarà in grado di percepire, riconoscere e identificare ciò che lo circonda, senza bisogno di controllo o addestramento da parte dell'uomo». Le promesse potenzialità di quel cervello elettronico, però, erano ancora più vaste: avrebbe dovuto imparare a pensare come gli umani, diventare cosciente di sé e, un giorno, sarebbe potuto partire per visitare «altri pianeti come una sorta di esploratore spaziale meccanico».

Insomma, le aspettative erano decisamente elevate; perché allora ci sono voluti quasi sessant'anni per trasformare parzialmente in realtà quelle promesse? Per capirci qualcosa, dobbiamo prima fare un altro passo indietro. Negli anni Quaranta, i biologi stavano sviluppando le prime teorie per spiegare come l'intelligenza e l'apprendimento fossero il risultato dei segnali trasmessi tra i neuroni nel cervello umano. La tesi fondamentale – che poi è quella valida ancora oggi – era che i collegamenti tra alcuni neuroni si rafforzassero

⁵ <https://www.nytimes.com/1958/07/13/archives/electronic-brain-teaches-itself.html>

sero attraverso la frequenza delle comunicazioni. È il meccanismo che fa sì che la prima volta che vi cimentate in una nuova azione siete incerti, ma col passare del tempo acquistate sempre più confidenza. I tentativi riusciti rafforzano il collegamento tra i neuroni coinvolti in una specifica azione.

Sulla base di queste nuove conoscenze, nel 1943 due pionieri come Warren McCulloch e Walter Pitts pubblicarono un paper⁶ in cui mostravano come un semplice sistema di neuroni artificiali potesse essere in grado di eseguire delle funzioni logiche basilari. Almeno in teoria, questo sistema poteva imparare nello stesso modo in cui impariamo noi: usando l'esperienza ed eseguendo quei tentativi ed errori che, come detto, rafforzano o indeboliscono le connessioni tra neuroni. «Il sistema neurale artificiale proposto dai due ricercatori avrebbe funzionato come il cervello, modificando le relazioni numeriche tra i neuroni artificiali sulla base dei tentativi e degli errori», si legge in un più recente articolo sempre del New York Times⁷. «Quindi, non avrebbe dovuto essere pre-programmato con regole fisse; piuttosto, avrebbe riscritto se stesso per riflettere gli schemi che emergevano dai dati assorbiti dalla macchina».

Quello proposto da McCulloch e Pitts era un sistema artificiale evuzionista; un modello che negli anni Cinquanta venne portato avanti dai (pochi) ricercatori convinti che il modo migliore per arrivare a una intelligenza artificiale fosse quello di ricalcare i meccanismi dell'apprendimento umano e permettere al “cervello elettronico” di imparare autonomamente, scovando i pattern all'interno dei dati che gli vengono forniti. Una modalità flessibile, che consente di modificare le competenze in base ai dati.

Ma i ricercatori che utilizzavano questo modello “dal basso”, dai dati, erano una piccola minoranza. Dall'altra parte delle barricate scientifiche si trovavano gli scienziati convinti che una vera intelligenza artificiale potesse nascere solo se istruita dall'alto, fornita cioè di tutte le regole necessarie a portare a termine il suo compito: una logica “creazionista”, meglio nota come simbolica. Questo sistema

6 <http://www.cse.chalmers.se/~coquand/AUTOMATA/mcp.pdf>

7 <https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html>

funzionava in modo esattamente opposto a quello evoluzionista: prevedeva infatti che le macchine venissero indottrinate con tutte le regole necessarie per portare a termine un compito. Per tradurre dall'italiano all'inglese, per esempio, sarebbe stato necessario fornire al computer tutte le regole grammaticali e i vocaboli delle due lingue, per poi chiedere loro di convertire una frase da una lingua all'altra.

Il modello simbolico è un sistema che ha portato a qualche successo (come sa bene Gary Kasparov, sconfitto a scacchi da Deep Blue di IBM; che adottava questa tecnica), ma ha grossi limiti: prima di tutto, richiede agli esseri umani un lavoro enorme; inoltre, funziona solo in quei campi che hanno regole molto chiare: la matematica o, appunto, gli scacchi. Nella traduzione – con tutte le sue sfumature, eccezioni e importanza del contesto – non raggiunge livelli nemmeno lontanamente accettabili. Ma proprio il fatto che questa tecnica eccellesse in settori nei quali è richiesta parecchia intelligenza, come la matematica o gli scacchi, veniva considerato un segnale molto promettente da parte dei sostenitori della AI simbolica.

Il modello simbolico, quindi, era quello dominante dell'epoca. Nonostante questo, fu il modello evoluzionista ad affascinare Frank Rosenblatt, psicologo dell'università di Cornell che nel 1956, partendo dalle teorie di McCulloch e Pitts, ideò la prima macchina in grado di simulare a livello software e hardware il funzionamento dei neuroni. Il sistema a cui diede vita venne ribattezzato Mark I Perceptron, più comunemente noto solo come Perceptron⁸. Ed eccolo, finalmente, il “cervello elettronico che insegna a se stesso” di cui parlava il New York Times: una macchina gigantesca, fittamente aggrovigliata da cavi e composta da motori e manopole collegati a 400 rilevatori di luce. Il tutto, per simulare il comportamento di otto neuroni.

«Ognuno di questi neuroni riceveva una parte dei segnali emessi dai rilevatori di luce, li combinava e a seconda di quale fosse il risultato, emetteva un 1 o uno 0», si legge sulla MIT Tech Review⁹.

8 <https://www.pearsonhighered.com/assets/samplechapter/0/1/3/1/0131471392.pdf>

9 <https://www.technologyreview.com/s/540001/teaching-machines-to-under->

«Messe assieme, queste cifre fornivano la ‘descrizione’ di ciò che il Perceptron aveva visto. Inizialmente, i risultati furono pessimi, ma Rosenblatt utilizzò un metodo chiamato ‘apprendimento supervisionato’ per addestrare il Perceptron a generare risultati in grado di distinguere correttamente forme diverse». E così, per esempio, contrassegnare correttamente un triangolo o un quadrato.

Oggi, i network neurali usano algoritmi complicatissimi, simulano il comportamento di milioni di neuroni e hanno miliardi di connessioni tra di loro; ma il funzionamento alla base è lo stesso del Perceptron: dati, apprendimento supervisionato, tentativi ed errori. Quando una prova va a buon fine, il legame tra i neuroni che hanno portato al risultato corretto si rafforza; quando invece fallisce, si indebolisce.

Il progetto di Rosenblatt attirò l'attenzione della marina USA, che decise di finanziarlo. E fu forse per assicurarsi questi finanziamenti che Rosenblatt fece promesse esagerate, affermando che il suo Perceptron sarebbe stato a breve in grado di riconoscere le persone e compiere, col tempo, tutte le straordinarie imprese citate dall'articolo del New York Times. Quando venne svelato il primo prototipo, nel 1958, il Perceptron era un macchinario da 5 tonnellate che occupava una stanza intera. Ma sapeva solamente distinguere la destra dalla sinistra.

Probabilmente lo stesso Rosenblatt era consapevole di quanto assurde fossero le aspettative riposte nel Perceptron. Nella prefazione al suo libro *Principles of Neurodynamics*¹⁰ (1962), racconta: «Il perceptron, originariamente, doveva essere un nome generico per una varietà teorica di reti neurali, ma ha la sfortuna di indicare invece un pezzo di hardware. È solo con grande difficoltà che riesco a convincere i suoi divulgatori, armati di buone intenzioni, a non usare la P maiuscola. E quando mi chiedono, Come sta andando il Perceptron oggi?, sono tentato di rispondere, Molto bene, grazie. E invece come vanno il Neurone e l'Elettrone?».

stand-us/

10 <http://mpclab.com/2016/09/frank-rozenblatt-principles-of-neurodynamics-perceptrons-and-the-theory-of-brain-mechanisms/>