

Dall'Intelligenza alla Comunicazione artificiale? Sociologia e possibilità teoriche

Elena Esposito, *Comunicazione artificiale. Come gli algoritmi producono intelligenza sociale*, Egea, Milano, 2022, pp. 256 (*Artificial Communication. How Algorithms Produce Social Intelligence*, The MIT Press, Cambridge, Mass., 2021)

Parole chiave

Intelligenza artificiale, algoritmi, comunicazione.

Matteo Finco è assegnista di ricerca presso il Dipartimento di Scienze Sociali ed Economiche, Università di Roma La Sapienza (matteo.finco@uniroma1.it)

Nella pervasiva tematizzazione dell'intelligenza artificiale tende a riprodursi, tanto nei media e nell'opinione pubblica quanto a livello politico-istituzionale, la classica contrapposizione tra rischi e pericoli di una tecnologia emergente (di cui si possono soltanto tentare di prevedere gli effetti) da un lato, ed entusiasmo per le possibilità che essa offre già nel presente, dall'altro. Questa tensione viene modulata in maniera variegata a seconda dell'osservatore, dando vita a una pluralità di posizioni, che oscillano fra l'apertura incondizionata e la prudenza, fra l'impegno e l'indifferenza. Quello che non viene mai messo in

dubbio è però il concetto stesso di *intelligenza* artificiale, usato fra l'altro per indicare un insieme di differenti tecnologie (*machine learning*, *deep learning*, *natural language processing*, ecc.). Si parte cioè dall'idea che le macchine (o meglio gli algoritmi) abbiano raggiunto uno stadio di sviluppo tale non solo da eseguire istantaneamente un volume immane di operazioni logico-matematiche per cui sono state programmate, ma che consenta ormai loro di produrre risultati inaspettati, le cui premesse non sono date in partenza, risolvendo così una serie di problemi altamente complessi attraverso risposte e soluzioni creative. Su tale presupposto, le scienze umane e sociali si sentono chiamate a fornire un contributo interpretativo e critico, che possa stare al passo con l'avanzamento tecnologico e il suo sfruttamento commerciale, ma anche ad accompagnarne le applicazioni in vari ambiti concernenti il bene pubblico (salute, tutela dell'ambiente, ecc.) – da cui ad esempio lo sviluppo di un'etica, di una filosofia e di un diritto dell'intelligenza artificiale, con le loro varianti (es. algoretica). In gioco c'è, naturalmente, la tutela di valori, libertà e diritti e concetti (es. autonomia) caratteristici della tradizione moderna, e soprattutto la necessità di affrontare le implicazioni concrete dell'AI per soggetti, organizzazioni e istituzioni, fornendo agli stessi un orientamento per il suo utilizzo.

Il libro di Elena Esposito (che in parte raccoglie e amplia alcuni saggi recenti ed è stato pubblicato prima in inglese e poi in italiano) si segnala, in particolare, per due ragioni. Anzitutto, offre un deciso cambio di prospettiva – gravido di conseguenze – sull'AI: quella degli algoritmi non sarebbe in realtà intelligenza, ma *comunicazione*. Ovvero, la comune analogia fra le loro prestazioni e l'intelligenza umana (trascurando il fatto che quest'ultima può essere scomposta in differenti tipologie: intelligenza logico-matematica, verbale-linguistica, emotiva, sociale, ecc.) non solo non sarebbe necessaria, ma risulterebbe fuorviante. Questo perché gli algoritmi – più in generale, le macchine – *non comprendono*. Non capiscono alla maniera degli esseri umani, cioè non attribuiscono un senso ai dati che raccolgono: invece *prevedono* (comportamenti, preferenze, tendenze, ecc.) e ciò consente loro di *sembrare* intelligenti, quando in realtà stanno funzionando sulla base

di un procedimento (un numero finito di passaggi) per cui sono stati programmati. Essi si nutrono, attraverso i Big Data (cioè una massa sterminata di dati), delle differenze generate dai comportamenti e dalle scelte passate degli utenti. Ma lo fanno producendo informazioni nuove. In altre parole, l'intelligenza che essi usano è quella degli esseri umani, ovvero le loro preferenze, le loro risposte tipiche, ma anche la loro imprevedibilità, per poi rielaborare tutto ciò. Organizzando e sistematizzando i dati, trovano schemi e strutture ricorrenti (*pattern*) e correlazioni potenzialmente istruttive. La loro efficienza si basa proprio sul fatto di rinunciare in partenza al tentativo di comprendere alla maniera umana (attribuendo significato alle situazioni e a contenuti specifici, dando vita a rappresentazioni astratte, generalizzazioni, ecc.), puntando invece sul trattamento dei dati, ovvero sull'elaborazione delle osservazioni degli utenti stessi. Basti pensare all'algoritmo di Google (PageRank), che attribuisce alle pagine Web indicizzate una data importanza sulla base principalmente dell'uso stesso del motore di ricerca da parte degli utenti e della maniera in cui vengono messi in collegamento fra loro i contenuti (*link*) – e non invece ad esempio sulla fedeltà del contenuto, sulla reputazione della fonte o su altri criteri ragionati. In altre parole, gli algoritmi imparano dalle scelte fatte dagli utenti e le riflettono: così viene prodotta una *contingenza virtuale*, risultante dal confronto fra l'utente e le istruzioni dell'algoritmo, sviluppate in maniera contingente e in una forma che risulta informativa, a partire dall'elaborazione dei comportamenti contingenti di altri utenti. Non solo: l'algoritmo continua a imparare, di fronte a comportamenti non tipici e risposte inedite, a loro volta risultato del confronto con la macchina stessa. Le macchine agiscono così non come semplici strumenti (*tool*), ma come *partner della comunicazione*, che rispondono alle richieste degli utenti con informazioni nuove, di cui non è possibile ricostruire la genesi. Si pensi soltanto ai principali assistenti vocali digitali (Siri, Alexa, Cortana): si possono fare loro richieste imprevedibili, non determinate in partenza, a cui essi tentano di rispondere utilizzando dati originati dalle fonti più diverse – con risultati non sempre soddisfacenti, naturalmente.

La seconda ragione della rilevanza del volume di Esposito consiste in un doppio movimento che esso realizza. Da un lato, fornisce un contributo che è insieme genuinamente sociologico e transdisciplinare: è un lavoro di sociologia della comunicazione e del mondo digitale, ma anche una riflessione sull'evoluzione della memoria sociale e dunque delle modalità in cui la società produce conoscenza; si interroga sulle forme di individualizzazione e personalizzazione che la comunicazione artificiale rende possibili, ma anche sulle sue implicazioni per il diritto all'oblio (un'istanza che emerge in conseguenza della possibilità di *ricordare* molto, cioè dalla difficoltà di *dimenticare*) – e dunque per la *privacy*; osserva le modalità di previsione del futuro mettendo a confronto fra loro il metodo statistico e le procedure algoritmiche, rintracciando inoltre i punti in comune fra queste ultime e quelle di divinazione nelle società arcaiche. Al tempo stesso, questo lavoro si discosta da una postura diffusa nelle scienze umane e sociali, che, oltre a comprendere e descrivere, problematizzare e porre ulteriori questioni, vuole anche e soprattutto criticare e fornire risposte e soluzioni concrete, utili e meglio ancora immediatamente spendibili. Il testo si sforza invece di osservare le possibilità offerte e già *messe in atto* attraverso la comunicazione artificiale, non trascurando i problemi evidenti, ma soprattutto – grazie in primo luogo all'attenzione alle questioni tecniche e comunicative – segnalando quelli meno evidenti (se si rimane all'interno di un atteggiamento classicamente umanistico), ma che emergono proprio grazie allo spostamento dell'attenzione sulla comunicazione e sulle implicazioni che ne derivano.

Qualche esempio al riguardo. Rispetto al primo punto (nuove opportunità) è particolarmente interessante il capitolo sulla fotografia e il suo utilizzo nei *social network*: la moda dei *selfie* e la possibilità di pubblicazione istantanea, se rischiano di spostare l'interesse dell'individuo/utente dall'esperienza (anzitutto corporeo-sensitiva) del momento presente alla condivisione *online*, al tempo stesso offrono nuove possibilità di elaborazione dell'esperienza stessa e della propria identità. Riguardo al secondo aspetto (problematiche inedite), si pensi alla questione dei *bias*, ovvero dei pregiudizi dei programmatori (generalmente maschi

bianchi) che gli algoritmi rischiano di riprodurre e amplificare. Si tratta di rischi che non vanno certo sottovalutati e per i quali vanno ricercate soluzioni tecniche, normative e giuridiche. Tuttavia, potrebbe destare altrettanta preoccupazione il fatto che il funzionamento delle macchine può essere distorto proprio perché esse *non* riflettono i valori dei loro creatori: si tratta del *data bias*, cioè del fatto che i dati trovati online, e riferiti al comportamento degli utenti, ne riflettono i pregiudizi e, visto che l'accesso alla rete non è neutrale né egualitario, rischiano di produrre aberrazioni (ad esempio costringendo, come è successo, a ritirare dei servizi che riproducevano le risposte razziste o sessiste apprese dagli utenti). Ancora, in ambito giuridico, se l'AI può essere utilissima nell'analisi della giurisprudenza come premessa per la produzione di decisioni (la cosiddetta giurisprudenza meccanica), al tempo stesso essa fatica a riprodurre (oltre che a comprendere) quell'ambiguità caratteristica della comunicazione umana, che proprio nel diritto risulta essenziale per orientarsi nella inevitabile varietà dei casi e dei contesti – da cui appunto argomentazioni giuridiche spesso vaghe e ambigue, ma che se non fossero tali implicherebbero la limitazione di possibilità future di interpretazione (la quale a sua volta non si esaurisce nella questione della trasparenza del procedimento logico seguito).

Il volume ha dunque il merito di segnalare alcuni snodi e passaggi logici essenziali per la comprensione del tema, anche volendo rifiutarne l'argomentazione di fondo, ovvero il mutamento di prospettiva dall'intelligenza alla comunicazione artificiale. Infatti, da un lato il testo illustra alcuni meccanismi fondamentali che stanno alla base delle tecnologie (es. l'uso delle liste, la visualizzazione e l'interpretazione nell'analisi digitale dei testi) e che rimangono solitamente in ombra; dall'altro, evidenzia una serie di questioni che riguardano il rapporto tra individui e macchine, l'accesso alle informazioni, lo sfruttamento dei dati e l'orientamento al futuro. Servendoci di alcuni argomenti affrontati nel volume, mettiamo di seguito in rilievo alcuni temi di cui le scienze sociali potrebbero farsi carico.

A questo proposito, essenziale è ad esempio la differenza fra *dati* e *informazioni*. Gli algoritmi infatti trattano dati (cioè unità discrete – o differenze – raccolte nelle situazioni e nei contesti più diversi) che di per sé non sono informativi, ma che lo diventano quando inseriti in un contesto specifico (si pensi ai dati sulla salute ricavati dai dispositivi indossabili, che possono essere utilizzati in contesti molto diversi da quelli per cui erano stati originariamente rilevati).

Fondamentale è anche il rapporto tra *in-trasparenza* e *complessità*. Dispositivi digitali, algoritmi e Big Data sono infatti in-trasparenti: non ne comprendiamo la logica di funzionamento – talvolta ciò vale anche per i loro programmatori. Tuttavia essi aumentano notevolmente la complessità, fornendo distinzioni, connessioni, correlazioni, previsioni – appunto in maniera opaca – che però sono reali e offrono possibilità inedite (da cui conseguono anche rischi) che occorre valutare, il che di per sé rappresenta un aggravio decisionale per gli individui e le organizzazioni.

Le categorie di cognizione, coscienza e intelligenza acquistano poi uno spessore differente, a partire non solo dalla consapevolezza che l'elevata capacità di calcolo e le procedure inedite degli algoritmi sono costitutivamente diverse dalla cognizione umana (come mostrano anche le acquisizioni più recenti delle neuroscienze, sottolineando ad esempio l'inestricabilità fra sentimenti ed emozioni – che, almeno al momento, le macchine non hanno – e cognizione stessa): se l'intelligenza umana si distingue per una elevata dose di creatività e arbitrarietà nell'attribuzione del senso e per la possibilità dell'individuo di autoservarsi, le macchine muovono invece da istruzioni fornite loro in origine. Soprattutto, funzionano senza la consapevolezza di farlo: dunque non hanno autocoscienza e non possono auto-osservarsi. Da cui l'inapplicabilità di concetti quali responsabilità e intenzionalità. Il fatto che possano produrre risposte imprevedibili non le rende per questo libere nel senso normalmente attribuito al termine.

Anche la dinamica tra *ricordo* e *dimenticanza* andrebbe ridiscussa. L'enorme capacità delle macchine di immagazzinare dati non ha molto a che vedere con la memoria umana e neanche con quella sociale

nella forma che conosciamo: altamente selettive, dipendenti dal contesto, esse riattivano possibilità in un primo tempo rimosse, rielaborando continuamente i ricordi e determinando cosa rimane sullo sfondo come latenza. La memoria consiste, in questo senso, in una costante oscillazione tra ricordo e dimenticanza. Con gli algoritmi, invece, ricordare diventa più facile e più economico (per questo deleghiamo alle macchine la conservazione di molte informazioni che sappiamo di poter recuperare facilmente), mentre dimenticare diventa quasi l'eccezione e richiede uno sforzo ulteriore: si pensi soltanto ai motori di ricerca e alle già menzionate pretese di oblio, che richiedono lo sviluppo di specifiche strategie di offuscamento.

Di conseguenza, a cambiare sono il concetto di *memoria* stesso, che da statico (conservazione dei dati) diviene dinamico (costruzione delle informazioni) e quello di *apprendimento*, tanto in riferimento alle macchine quanto agli esseri umani. Le prime infatti non *imparano* come i secondi, proprio perché non attribuiscono significati alle cose e non identificano errori: invece i loro calcoli, i *pattern* e le correlazioni identificate, nonché le previsioni generate, possono – oppure no – in seguito risultare significativi per gli utenti. I quali dal canto loro, attraverso algoritmi, Big Data, link, *tag* e motori di ricerca, possono accedere a informazioni non soltanto nuove, ma che neanche stavano cercando – il che, ancora una volta, aumenta la complessità, ma anche le possibilità di apprendimento.

Un'altra questione rilevante attiene alla ricerca scientifica, in particolare all'elaborazione teorica. L'AI può infatti indurre a pensare che l'accesso a una mole imponente di dati e l'elevata capacità di calcolo ed elaborazione consentano di per sé di raggiungere risultati inarrivabili per la mente umana e che dunque la costruzione di teorie e di schemi interpretativi rappresenti oramai uno stadio superato nel processo di conoscenza. In questo modo, le teorie generali diverrebbero obsolete se non inutili. Anche perché ormai non sarebbe tanto importante capire il *perché* dei fenomeni, quanto *misurarli*: è questa in sostanza la discussa proposta di Chris Anderson, ex direttore della rivista Wired, espressa in un articolo di qualche anno fa, intitolato proprio *The End of Theory*.

Trovare correlazioni – appunto il lavoro degli algoritmi – sarebbe sufficiente poiché, con molti dati a disposizione, i numeri parlerebbero da sé. La proposta tuttavia non regge per vari motivi. Fra i più evidenti, c'è, da un lato, il fatto che gli algoritmi stessi si fondano su teorie (statistiche e matematiche, nonché sulla *computer science*) e che nella lettura e nell'analisi dei dati si fa ricorso a teorie, necessarie per selezionarli e interpretarli (i dati vanno infatti sempre ripuliti prima dell'uso – il cosiddetto *data cleaning*); dall'altro lato, il passaggio dal focus sul *perché* dei fenomeni al *come* e al *quanto* – appunto, il passaggio dalla spiegazione causale alla correlazione – riduce la comprensione all'interesse per il mero *funzionamento*, ovvero all'applicabilità immediata. Inoltre, il *come* esprime sostanzialmente degli effetti, i quali, anche quando non risultano riconducibili ad una catena causale, sono comunque espressione di una *funzione*. Ciò significa che rinunciare ad indagare il *perché* dei fenomeni implica l'impossibilità di distinguere tra funzioni latenti e manifeste e soprattutto di inquadrare i fenomeni all'interno di processi più ampi, di rintracciare analogie fra ordini differenti e di identificarne le caratteristiche strutturali. In definitiva, la correlazione può descrivere un fenomeno, ma più difficilmente ne evidenzia il *sensò*. Questo, tuttavia, non dovrebbe sorprendere: gli algoritmi infatti “pretendono di prevedere il futuro” (p. 146) e si disinteressano del passato, a partire dal presupposto che la previsione non dipende dall'averlo compreso, poiché i *pattern* identificati non si basano sul significato, ma sul contesto che di volta in volta si analizza. Concentrandosi non sulla media degli elementi (come fa la statistica), ma su specifici soggetti (sulla base di una specifica forma di *individualizzazione algoritmica*, che combina l'orientamento al contesto alla personalizzazione) di cui viene osservato il comportamento, l'interesse si focalizza su ciò che essi potranno fare in futuro (si pensi a quel padre indignato per la pubblicità di prodotti per neonati ricevuta dalla figlia adolescente, e che ha poi scoperto che la ragazzina era effettivamente incinta, come l'algoritmo all'origine del messaggio pubblicitario aveva ‘intuito’, per così dire, a partire dalle ricerche fatte online dalla stessa). La strategia messa in atto attraverso gli algoritmi è quella di rispondere all'inconoscibilità del futuro attraverso il tentativo di prevederlo, senza necessariamente comprenderne le possibili cause remote (il fatto che sul Web mi venga

suggerito l'acquisto di un determinato libro non si basa sul caso né su possibili interessi in astratto, ma sulle ricerche e gli acquisti effettuati in precedenza, combinando queste informazioni con gli acquisti fatti da profili simili al mio).

Capire questi meccanismi è essenziale perché quello della previsione è un ambito tanto problematico quanto attuale, soprattutto nel momento in cui se ne considerino i possibili effetti di auto-avveramento (comprerò effettivamente quello che mi viene suggerito, nonostante non ci avessi mai pensato prima?). Anche qui, che si tratti di comunicazione *artificiale* e non *umana* non dovrebbe preoccupare: i consigli di lettura di un amico non sono in fondo meno arbitrari. Ciò che conta è insomma che la comunicazione funzioni, offrendo all'utente delle possibilità inedite: gli algoritmi realizzano appunto questo passaggio dalla *verità* alla *performatività* (p. 157). Più preoccupanti sono invece, ad esempio, le pretese di prevedere dove, quando e da chi verrà commesso un crimine, perché ciò potrebbe portare le autorità a ignorare altri fattori importanti. Se la sfida è quella di "combinare la previsione algoritmica individuale con l'apertura al futuro" (p. 163), anche le scienze sociali dovrebbero occuparsene.

Occorre cioè, come questo libro riesce a fare (in maniera tanto più apprezzabile in quanto ciò non risulta né da una postura di stampo normativo, etico o critico), mettere in evidenza i lati oscuri dell'intelligenza/comunicazione artificiale, segnalando ulteriori piste di indagine e riflessione e offrendo sguardi e prospettive differenti da quelle del *mainstream*. Occorre dunque osservare, rintracciando regolarità e analogie fra fenomeni e tecnologie differenti, rotture rispetto al passato, reazioni e comportamenti adattivi – in una parola, fare teoria, o almeno dissodare il terreno perché ciò sia poi possibile. Intrecciando l'analisi del rapporto circolare fra mutamenti semantici e strutturali della società a quella della loro relazione con gli sviluppi tecnologici emergenti – senza porre questi ultimi in una posizione subordinata e strumentale rispetto ai primi –, emerge così la necessità *pratica e concreta* di capire come comunicare in maniera efficiente con macchine incomprensibili, o quantomeno in-trasparenti. A ulteriore conferma che forse la fine della teoria non è così vicina come si potrebbe pensare.