

PENSARE GLI ALGORITMI: QUALE INTELLIGENZA?

Estratto dall'*Introduzione* del volume della sociologa **Elena Esposito**:

Comunicazione artificiale. Come gli algoritmi producono intelligenza sociale, Milano, Egea 2022.

«I recenti **algoritmi** che lavorano con deep learning e big data stanno diventando sempre più **bravi a fare sempre più cose**: generano informazioni in modo rapidissimo e preciso, stanno imparando a guidare le macchine in maniera più sicura e affidabile degli autisti umani, sanno rispondere alle nostre domande, fare conversazione, comporre musica, leggere libri e scrivere testi interessanti, appropriati e – se occorre – anche divertenti.

Nell'osservare questi progressi, però, difficilmente siamo del tutto rilassati, e non solo perché **ci preoccupiamo** dei bias, degli errori, delle minacce alla privacy o delle possibili intenzioni maligne delle aziende o dei governi. Anzi, più gli algoritmi diventano bravi, più aumenta il nostro disagio. Un recente articolo pubblicato su *The New Yorker* descrive l'esperienza del giornalista con Smart Compose,¹ una funzione di Gmail che suggerisce come finire le frasi mentre le si sta digitando. L'algoritmo ha completato i messaggi del giornalista in modo così appropriato, conforme al suo stile e pertinente che lui si è ritrovato a imparare dalla macchina non solo che cosa avrebbe scritto, ma anche che cosa avrebbe dovuto scrivere o avrebbe potuto voler scrivere. E non gli è piaciuto per niente.

Questa esperienza, estremamente comune nell'interazione con le presunte macchine intelligenti, viene chiamata «uncanny valley»²: un'inquietante **sensazione di disagio** che compare nei casi in cui una macchina sembra essere troppo simile a un essere umano - o allo stesso osservatore. Vogliamo che la macchina supporti i nostri pensieri e i nostri comportamenti, ma quando ritroviamo *nella* macchina i nostri pensieri e i nostri comportamenti non ci sentiamo affatto a nostro agio. Oggi ciascuno di noi comunica abitualmente con dei programmi automatizzati (dei bot) senza badarci: quando compriamo online i biglietti d'aereo, quando chiediamo assistenza sul web, giochiamo ai videogame e in moltissime altre occasioni.³ Ciononostante, quando riflettiamo e dibattiamo sugli algoritmi, discutiamo ancora della possibilità che una macchina superi il test di Turing,⁴ e che possa arrivare a una singolarità tecnologica o a una super-intelligenza che superi le capacità umane.⁵ Ci confrontiamo con le macchine e non ci piace che vincano loro. Cerchiamo di costruire delle macchine intelligenti, ma poi non ci chiediamo solo se ci siamo riusciti, ma anche se le macchine non stiano diventando *troppo* intelligenti.

Ma è davvero di questo che ci dobbiamo preoccupare? Anche se ci possiamo sentire a disagio quando le macchine sembrano assomigliarci troppo, riteniamo che il rischio fondamentale degli algoritmi è che essi possano competere con l'intelligenza umana? Questo libro parte dall'ipotesi che **l'analogia tra le prestazioni degli algoritmi e l'intelligenza umana non solo non sia necessaria, ma sia fuorviante**, anche se il ragionamento che ci sta dietro è di primo acchito plausibile. Oggi in effetti molti algoritmi sembrano essere in grado di «pensare» e di comunicare. Nella comunicazione come noi la conosciamo i partner sono sempre stati degli esseri umani, e gli esseri umani sono dotati di intelligenza. Se ora il nostro interlocutore è un algoritmo, viene spontaneo attribuirgli le caratteristiche degli esseri umani. Se la macchina sa comunicare autonomamente come un essere intelligente, si pensa, deve essere anch'essa intelligente, per

¹ John Seabrook, «Can a Machine Learn to Write for the New Yorker?», *The New Yorker*, 14 ottobre 2019, <https://www.newyorker.com>.

² Masahiro Mori, «The Uncanny Valley», *IEEE Robotics and Automation*, 19, n. 2, 2012, pp. 98-100

³ Sulla capacità degli algoritmi di condurre in modo «naturale» delle conversazioni complesse, si veda Chris Welch, «Google Just Gave a Stunning Demo of Assistant Making an Actual Phone Call», *The Verge*, 8 maggio 2018, www.theverge.com.

⁴ Alan Turing ha proposto il test che porta il suo nome in «Computing Machinery and Intelligence» per valutare la capacità di una macchina di esibire un comportamento intelligente. La macchina supera il test se un osservatore non riesce a distinguere i suoi contributi a una conversazione in linguaggio naturale da quelli di un partner umano.

⁵ Si vedano per esempio Ray Kurzweil, *The Singularity Is Near*, New York, Viking Books, 2005; Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies*, Oxford, Oxford University Press, 2014.

quanto magari in modo diverso dagli esseri umani. Sulla base di questa analogia, la ricerca si è concentrata sui parallelismi e sulle differenze tra l'intelligenza umana e le prestazioni delle macchine, osservando gli eventuali limiti e facendo dei confronti.⁶ Ma è davvero opportuno continuare a perseguire questa analogia?

Il fatto che possiamo comunicare con le macchine non implica che esse abbiano una loro intelligenza che deve essere spiegata (magari spiegando anche i misteri dell'intelligenza «naturale»), ma che, innanzitutto, *la comunicazione* sta cambiando. L'oggetto della ricerca presentata in questo libro non è l'intelligenza, che è e rimane un mistero, ma la comunicazione, che possiamo osservare e di cui sappiamo già molto. Sappiamo, per esempio, come la comunicazione sia cambiata nel corso dei secoli e con l'evoluzione della società, passando dalla semplice interazione tra persone presenti in un medesimo luogo a forme sempre più flessibili e inclusive, che consentono anche di comunicare con partner distanti nello spazio e nel tempo, via via più inaccessibili, in contesti anonimi e impersonali. Alla comunicazione vocale e gestuale si sono aggiunte la scrittura a mano e poi a stampa, la trasmissione di immagini fisse, in movimento e anche tridimensionali.

Con l'evoluzione della comunicazione, il ruolo degli esseri umani è cambiato profondamente: oggi non c'è bisogno che i partner condividano lo stesso spazio fisico, non c'è bisogno di sapere chi sono e perché comunicano, e non c'è nemmeno bisogno di sapere che cosa intendono e di tenerne conto. Leggiamo e comprendiamo il libretto di istruzioni della lavastoviglie senza sapere chi lo ha scritto e senza identificarci con il suo punto di vista; interpretiamo un'opera d'arte senza essere vincolati alla prospettiva e all'intenzione dell'autore.⁷ Non è necessario che le informazioni siano memorizzate nella mente di qualcuno (nessuno conosce a memoria il Codice civile), e in tutti i casi di finzione ci identifichiamo con i personaggi dei romanzi e dei film sapendo che non sono mai esistiti e che non sono gli emittenti della comunicazione. L'idea di comunicazione come condivisione di contenuti il più possibile identica tra le menti dei partecipanti è già irrealistica da molti secoli, nella pratica se non nella teoria. Nella maggior parte dei casi, gli emittenti e i riceventi non si conoscono, non conoscono la reciproca prospettiva, il contesto e i vincoli della controparte, e non ne avvertono l'esigenza. Anzi, questa «intrasparenza» consente gradi altrimenti impensabili di libertà e di astrazione.

Che le forme della comunicazione cambino non è una novità e non è nemmeno un enigma. Si tratta piuttosto di individuare e comprendere le differenze e le continuità tra le forme nuove e quelle già familiari. Attualmente però l'autonomia della comunicazione dai processi psichici dei partecipanti si è spinta un passo più avanti. Occorre un concetto di comunicazione che sia in grado di tener conto anche della possibilità che il partner comunicativo non sia un essere umano ma un algoritmo. Il risultato, che può essere osservato già oggi, è una condizione in cui disponiamo di informazioni di cui spesso nessuno può ricostruire né comprendere la genesi, ma che ciononostante non sono arbitrarie. Le informazioni generate autonomamente dagli algoritmi non sono affatto casuali e sono del tutto controllate, ma non dai processi della mente umana.⁸

Come possiamo controllare questo controllo, che per noi può essere anche incomprensibile? Questa è, a mio parere, la vera sfida che ci pongono oggi le tecniche di machine learning e l'uso dei big data» (Esposito, 2022, 1-4).

Elena Esposito nel volume *Comunicazione artificiale. Come gli algoritmi producono intelligenza artificiale* sviluppa «questa prospettiva indagando l'utilizzo degli algoritmi in diverse aree della vita sociale. Che cosa si vede, non si vede o si vede diversamente se ci si riferisce alla comunicazione anziché all'intelligenza?» e mette in discussione «l'adeguatezza della classica

⁶ Da John R. Searle, «Mind, Brains and Programs», *Behavioral and Brain Sciences*, 3, n. 3, 1980, pp. 417-457, fino a Reza Negarestani, *Intelligence and Spirit*, Cambridge, MA, Urbanomic/Sequence Press, 2018.

⁷ Umberto Eco, *Opera aperta*, Milano, Bompiani, 1962.

⁸ Che siano controllate ovviamente non vuol dire affatto che siano corrette, neutrali o che debbano essere accettate senza riserve e senza critica. Come mostrano le dinamiche del feedback, la presenza di controllo non esclude rischi, manipolazioni e risultati negativi. D'altro canto anche il controllo umano, come è noto, non è certo garanzia di successo, e nemmeno di razionalità.

metafora dell'intelligenza artificiale e dei suoi derivati come le reti neurali per analizzare i recenti sviluppi delle tecnologie digitali e del web. La nuova generazione di algoritmi, quelli che in forme differenti hanno dato impulso all'utilizzo dei big data e ai progetti connessi, non cercano di riprodurre artificialmente i processi dell'intelligenza umana. Questa, sostengo, non è una rinuncia né una debolezza, ma la base della loro incomparabile efficienza nell'elaborazione dell'informazione e nell'interazione con gli utenti. Per la prima volta le macchine sono in grado di produrre delle informazioni mai pensate da una mente umana e di fungere da partner comunicativi interessanti e competenti – non perché siano diventate intelligenti, ma perché non cercano più di esserlo. **I processi che guidano gli algoritmi sono completamente differenti dai processi della mente umana**, e in effetti nessuna mente umana o combinazione di menti umane potrebbe riprodurli e spesso nemmeno comprenderli. Però l'intelligenza umana rimane irrinunciabile. **Gli algoritmi che autoapprendono sono in grado di calcolare, combinare ed elaborare differenze con un'efficienza strabiliante, ma non sono in grado di produrle autonomamente.** Le trovano sul web. Tramite i big data, gli algoritmi si «nutrono» delle differenze generate (consapevolmente o inconsapevolmente) dalle persone e dal loro comportamento per produrre delle informazioni nuove, sorprendenti e potenzialmente istruttive. I processi degli algoritmi partono dall'intelligenza e dall'imprevedibilità (dalla contingenza) degli utenti per rielaborarle e operare in modo intelligente come partner comunicativi, senza alcun bisogno di essere essi stessi intelligenti» (Esposito, 2022, 5).

L'autrice nel corso del volume esplora le conseguenze di questa condizione nell'utilizzo pratico degli algoritmi, in particolare fa riferimento ai seguenti aspetti inerenti all'utilizzo di

- *liste* che consentono di organizzare i dati e di gestire le informazioni anche senza comprenderle (Cap. 2);
- *visualizzazioni* nelle *digital humanities* come una tecnica che serve a rendere significativi i risultati delle procedure incomprensibili dell'elaborazione algoritmica dei testi (Cap. 3);
- *profilazione* digitale e modalità di individualizzazione realizzate dagli algoritmi (Cap. 4);
- *memoria* algoritmica e diritto all'oblio (Cap. 5);
- *foto* digitali che sembrano prodotte per sfuggire alla pressione del presente (Cap. 6).

Elena Esposito chiude il suo libro esplorando **il futuro della previsione algoritmica** e si interroga su come gestire l'incertezza.

«Di fronte alla crescente “intrasparenza” di algoritmi sempre più efficienti, diversi osservatori cominciano a pensare che le macchine siano incomprensibili innanzitutto perché non c'è niente da comprendere, e non c'è niente da comprendere perché le macchine non comprendono. **Gli algoritmi sembrano intelligenti e funzionano come se fossero intelligenti non perché comprendono ma perché sono in grado di prevedere.** In riferimento ai software di scrittura automatica, Ilya Sutskever, ricercatore capo di OpenAI, lo afferma esplicitamente: «Se una macchina [...] potesse disporre di dati e di potenza di calcolo sufficienti per prevedere alla perfezione [...], questo equivarrebbe alla capacità di comprendere».⁹

La previsione è il nuovo orizzonte della ricerca sulle forme artificiali di intelligenza, in un contesto che cambia radicalmente i termini della questione: quando si lavora con gli algoritmi non ci si occupa di spiegare ma di prevedere, non di individuare delle relazioni causali ma di trovare delle correlazioni, non di gestire l'incertezza del futuro ma di scoprirne le strutture (i pattern). Il mondo però rimane incerto, il futuro resta aperto e l'uso degli algoritmi deve comunque essere spiegato. È qui, a mio parere, che si pongono oggi la questione del controllo e la difficoltà fondamentale degli algoritmi: **come gestire l'impatto di queste procedure indipendenti dal significato in una società in cui i significati, la contingenza e l'incertezza sono tuttora risorse preziose**» (Esposito, 2022, 6-7).

* Documento introduttivo al seminario “Educarci per educare all'intelligenza artificiale” prodotto per uso interno.

⁹ Citato in Seabrook, «Can a Machine Learn to Write for the *New Yorker*?», cit.