

Gruppo Interdisciplinare su Scienza, Tecnologia e
Società (GI-STTS) dell'Area della Ricerca di Pisa del CNR

Intelligenze al confine tra biologico e artificiale
Seminario aperto al pubblico e on-line

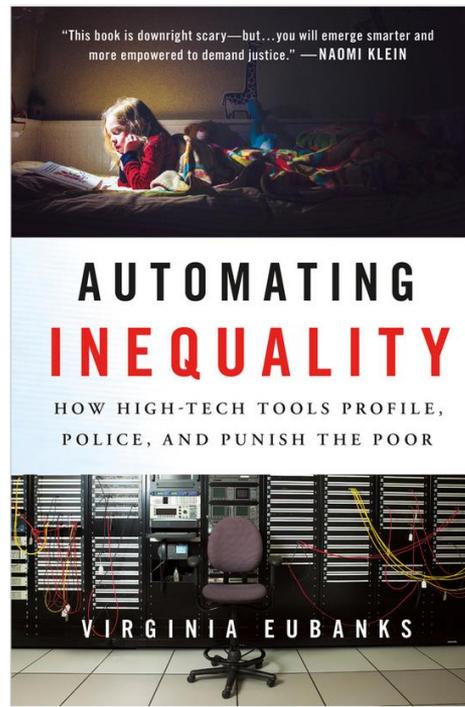
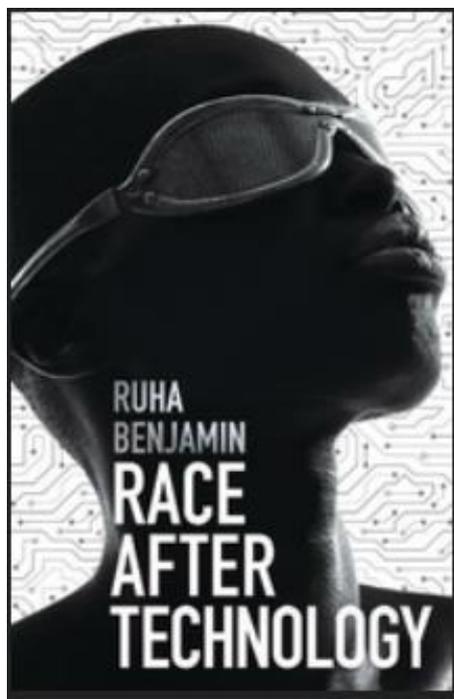
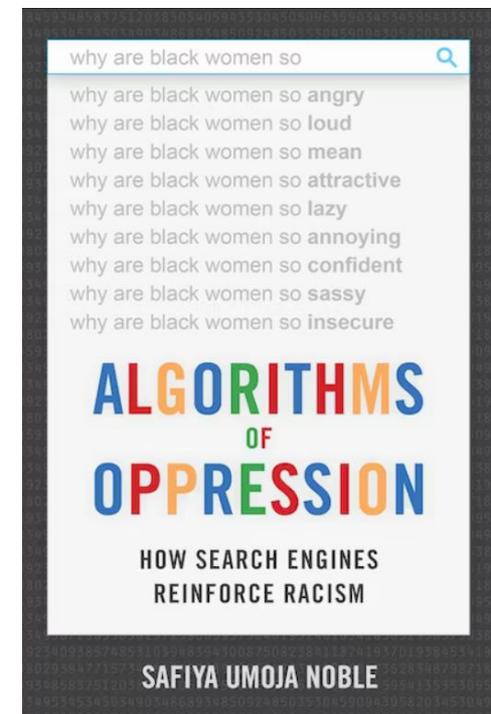
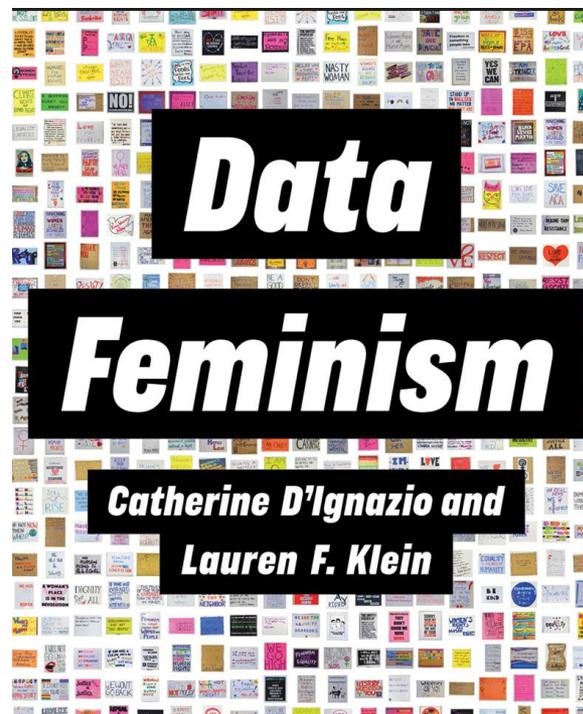
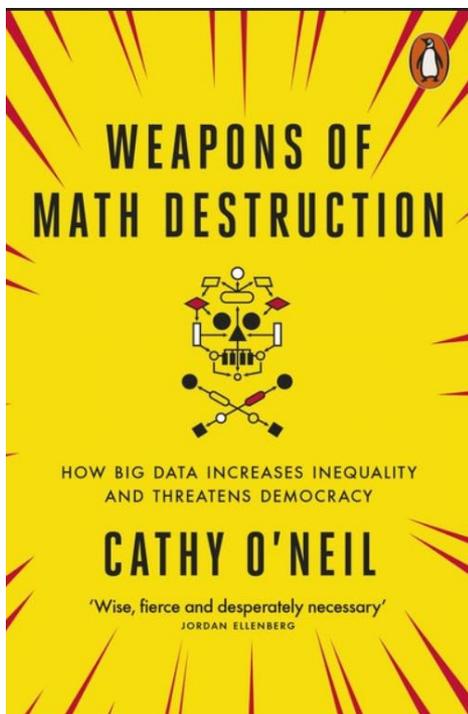
Daniela Tafani

**"Intelligenza artificiale", pensiero
magico e cattura culturale**

20 settembre 2023



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/)



Tre fallacie

Il mito dell'oggettività algoritmica si è dissolto, negli ultimi anni, di fronte alla vasta documentazione dei danni, delle discriminazioni e delle ingiustizie prodotti dalle decisioni fondate su statistiche automatizzate.

Resta ancora intatta, tuttavia, la tendenza a credere che tutti i sistemi fondati sull'«intelligenza artificiale» funzionino:

- fallacia della funzionalità dell'IA;
- fallacia degli esempi tratti dal futuro o dalla fantascienza;
- fallacia del primo passo.

Intelligenza artificiale e narrazioni

«Intelligenza artificiale», oggi:

1. **Una tecnologia**, entro uno spettro che va dai sistemi esperti, fino, teoricamente, all'intelligenza sintetica. Al momento, per arrivare a quest'ultima, non abbiamo nemmeno le domande giuste.
2. **Un atto linguistico**, una narrazione che deriva in gran parte dalla fantascienza su computer e robots con capacità straordinarie e tendenze apocalittiche, un discorso che evoca una precisa percezione sociale dell'IA come tecnologia. Marketing.

E' utilizzata per esercitare controllo, o assicurarsi profitto, senza assumersi responsabilità.

Non mera evasione dalla responsabilità: creazione di asimmetrie di potere.

Es. lavoratori a distanza di un call center e «sorveglianza con un sistema di IA».

Peter R. Lewis, Stephen Marsh, Jeremy Pitt, *AI vs «AI»: Synthetic Minds or Speech Acts*, «IEEE Technology and Society Magazine», 2021, <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9445758>

IA ristretta e IA generale

Narrow AI (o weak AI) i sistemi di IA ristretta sono sistemi che possono eseguire uno o pochi compiti specifici.

General AI (o strong AI): un sistema di IA generale è un sistema che può eseguire la maggior parte delle azioni che gli esseri umani possono compiere.

Tutti i sistemi di IA attualmente in uso sono esempi di IA ristretta: funzionano per i compiti particolari per i quali sono stati programmati, a condizione che ciò che incontrano non sia troppo diverso da quello che hanno sperimentato in precedenza.

G. Marcus, E. Davis, *Rebooting AI. Building Artificial Intelligence We Can Trust*, New York, Pantheon Books, 2019.

La formula di marketing «intelligenza artificiale»

I sistemi di intelligenza artificiale sono oggi in grado di svolgere alcuni specifici compiti, che erano stati, finora, prerogativa dei soli esseri umani. Nell'entusiasmo per i sistemi di apprendimento automatico, che hanno consentito questi genuini progressi, le grandi aziende tecnologiche hanno colto l'opportunità per un'espansione illimitata di prodotti e servizi «intelligenti».

Hanno diffuso e messo in commercio, con la formula di marketing «intelligenza artificiale», sistemi di apprendimento automatico, per lo svolgimento di attività che tali sistemi non sono in grado di svolgere o che semplicemente non sono possibili.

Tra i prodotti di questo genere ci sono

1. le auto a guida autonoma,
2. i sistemi di ottimizzazione predittiva,
3. i generatori di linguaggio naturale.

D. Tafani, *Sistemi fuori controllo o prodotti fuorilegge? La cosiddetta «intelligenza artificiale» e il risveglio del diritto*, «Bollettino telematico di filosofia politica», 2023, pp. 1-13, <https://btfp.sp.unipi.it/it/2023/05/sistemi-fuori-controllo-o-prodotti-fuorilegge>

Intelligenza artificiale e cattura culturale

Cattura del regolatore (*regulatory capture*) e cattura culturale (*cultural capture*): colonizzazione dello spazio dell'intermediazione scientifica, come già accaduto con l'industria del tabacco e quella dei combustibili fossili.

I detentori di grandi monopoli intellettuali finanziano la ricerca sui temi dell'intelligenza artificiale, per assicurarsi che siano coerenti con il loro modello di business

- gli esiti
- l'inquadramento teorico
- il tono.

In evidente conflitto di interessi, le grandi aziende tecnologiche finanziano l'ambito di ricerca interdisciplinare dell'etica dell'intelligenza artificiale, che si occupa dei danni e delle sofferenze che esse stesse provocano, con la commercializzazione dei sistemi di IA.

A. Saltelli, D.J. Dankel, M. Di Fiore, N. Holland, M. Pigeon, *Science, the Endless Frontier of Regulatory Capture*, in «Futures», 2022, n. 135

L'«etica» come merce

Politiche neoliberali di definanziamento pubblico delle università.

Mercato con un'elevata domanda di «prodotti etici», ossia di ricerche sull'etica dalle caratteristiche e dagli esiti prefissati: i ricercatori diventano così «i fornitori di un servizio in questa nuova economia della virtù» e sono indotti alla «complicità con sistemi e attori che cercano di operazionalizzare l'etica per proteggere i propri interessi».

L'«etica dell'intelligenza artificiale» è assimilabile dunque a una merce, a «un'esca per catturare la fiducia» dei cittadini, un mero discorso.

La funzione di tale discorso è quella di tutelare un modello di business fondato sulla sorveglianza e sulla raccolta di dati e metadati individuali.

Non è l'azienda che diventa etica, ma l'«etica» (ossia un mero discorso sull'etica) che diventa un asset aziendale.

Economies of Virtue: The Circulation of 'Ethics' in AI, a cura di T. Phan, J. Goldenfein, D. Kuch, M. Mann, Amsterdam, Institute of Network Cultures, 2022, <https://networkcultures.org/blog/publication/economies-of-virtue-the-circulation-of-ethics-in-ai/>

«Intelligenza artificiale» e narrazioni

A tutela del loro modello di business, i giganti della tecnologia hanno diffuso una famiglia di narrazioni che costituiscono ormai gli assiomi indiscussi di qualsiasi discorso pubblico:

- l'idea che sia davvero disponibile, oggi, in senso proprio, un'«intelligenza artificiale»;
- l'antropomorfizzazione delle macchine e la deumanizzazione delle persone;
- il principio di inevitabilità tecnologica («se non lo facciamo noi, lo farà qualcun altro»);
- la logica del fatto compiuto;
- i miti dell'eccezionalismo tecnologico e del vuoto giuridico;
- il principio di innovazione;
- la prospettiva del soluzionismo tecnologico: questioni sociali trattate quali problemi di disciplinamento e controllo, passibili di soluzioni tecniche; un set di problemi e soluzioni che individua nel *design* tecnico il livello appropriato per la soluzione di tutti i problemi e pone al riparo dalla discussione le decisioni di business;
- tema dei *bad actors*, anziché dei rapporti di sfruttamento o di potere strutturali;
- le ideologie TESCREAL.

Intelligenza artificiale e pensiero magico

Le narrazioni sull'intelligenza artificiale presentano tre caratteristiche proprie del pensiero magico:

1. l'animismo: la tendenza a concepire in termini antropomorfici alcuni oggetti della tecnologia;
2. la mossa, da prestigiatori, di mostrare un risultato, o un effetto, nascondendone allo stesso tempo le cause concrete e i costi;
3. l'assunto che sia possibile prevedere il comportamento futuro di ogni singola persona (come i responsi dell'astrologia, le predizioni algoritmiche a proposito di future azioni o prestazioni umane sono fondate su correlazioni alle quali non corrisponde alcun nesso causale).

D. Tafani, *What's wrong with "AI ethics" narratives*, in «Bollettino telematico di filosofia politica», 2022, pp. 1-22, <https://commentbfp.sp.unipi.it/daniela-tafani-what-s-wrong-with-ai-ethics-narratives>

Antropomorfizzazione e deumanizzazione

La tendenza spontanea ad antropomorfizzare gli oggetti della tecnologia è sfruttata dalle aziende per indurre la convinzione che i loro prodotti siano in grado di comprendere, conoscere, consigliare o giudicare in senso proprio.

Correlativamente, il cervello è assimilato a un computer, gli esseri umani a macchine e i volti delle persone a codici a barre, così da giustificare il fatto che le persone siano effettivamente trattate come cose, nei processi in cui si decidono le loro sorti sulla base di statistiche automatizzate, a partire da etichette e collezioni di dati che – a qualsiasi titolo e pur senza alcun riferimento a contesti e significati – li riguardino.

← Tweet



Sam Altman  @sama · 04 dic 22

i am a stochastic parrot, and so r u

 175

 284

 1.812



<https://twitter.com/sama/status/1599471830255177728>

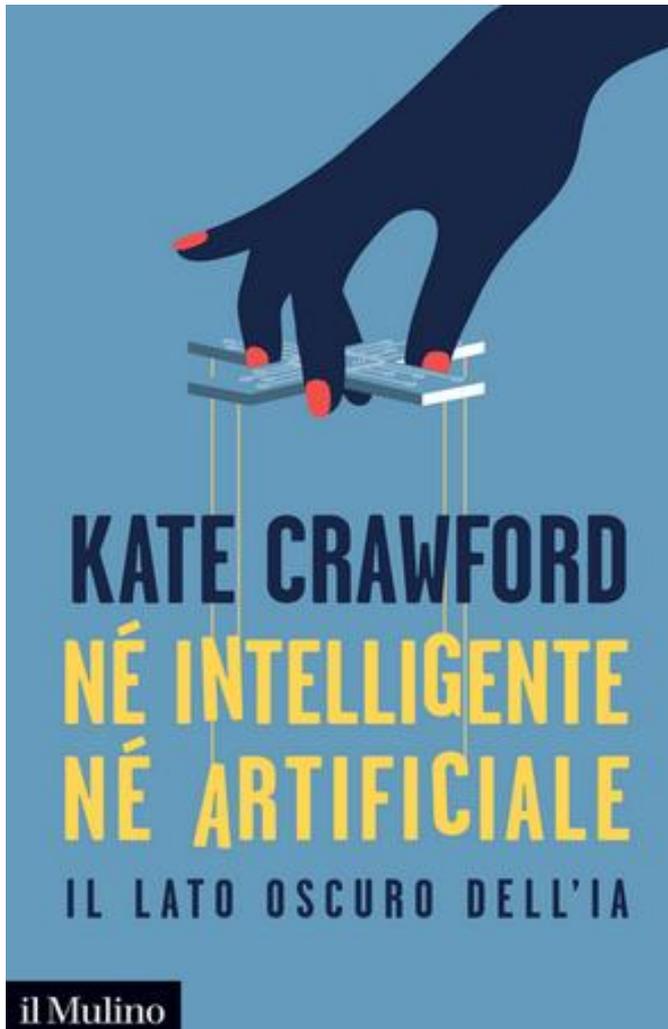
La fuga dalla responsabilità

Fuga dei giganti della tecnologia dalle loro responsabilità per gli effetti dannosi dei sistemi di apprendimento automatico:

- appelli all'eccezionalità delle nuove tecnologie;
- tesi di un vuoto di responsabilità (*responsibility gap*);
- proposta di una responsabilità distribuita anche tra gli utenti e le vittime
- proposta di attribuire un ruolo decisivo agli esseri umani coinvolti nei processi automatizzati (*human on the loop*), affidando a una persona il compito di intervenire, con prontezza fulminea, nei casi di emergenza, o di rettificare il responso imperscrutabile di un sistema automatico.

L'introduzione di un inverosimile controllo umano svolge la funzione di legittimare l'uso di sistemi pericolosi, attribuendo all'essere umano il ruolo di capro espiatorio.

Quanto simili strategie aziendali siano analoghe alle truffe in senso stretto, è manifesto nei casi in cui le vittime siano colpevoli *by design*, come pare sia stato specificamente previsto nei veicoli Tesla attraverso un meccanismo di disattivazione della funzione Autopilot un secondo prima dell'impatto, così da far risultare che il veicolo, al momento dello schianto, fosse affidato al passeggero umano.



1. Dati
2. Potenza di calcolo
3. Algoritmi

Altre «estrazioni»:

4. **Terre rare:** ad es., per raffinare una tonnellata di terre rare, il processo produce 75.000 litri di acqua acida e una tonnellata di residui radioattivi.
 5. **Energia:** ad es., l'esecuzione di un solo modello di elaborazione del linguaggio naturale ha prodotto 300.000 chilogrammi di anidride carbonica (quanto 5 auto a gas nel loro intero ciclo di vita, produzione compresa, o 125 voli andata e ritorno da New York a Pechino).
- **Lavoro:** colonialismo digitale; sfruttamento del lavoro, lavoro che compromette la salute mentale



THE ORIGINAL CURE ALL

RELIEVES INSTANTANEOUSLY

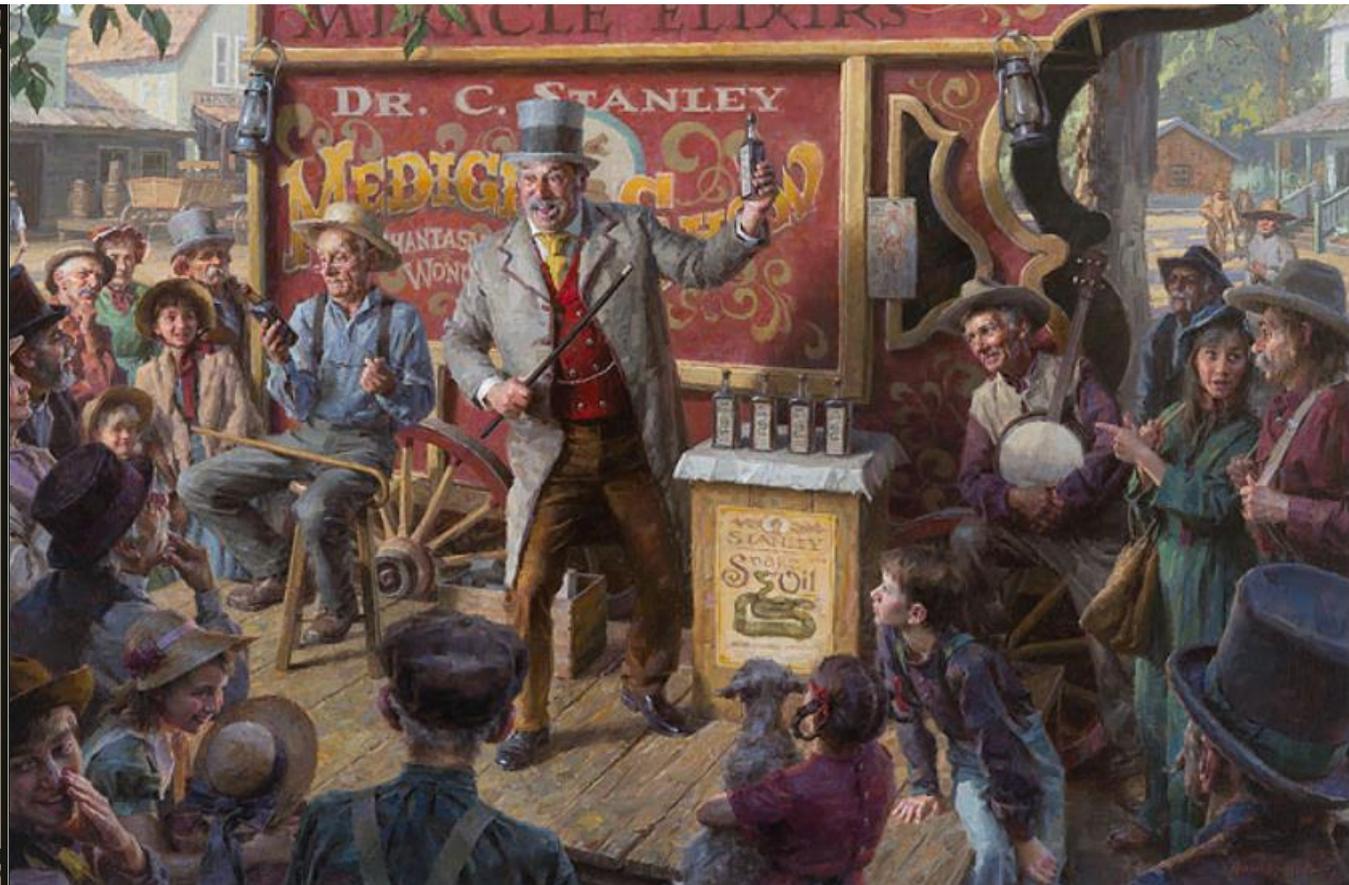
And Cures: Headaches, Neuralgia, Cough, Cold, Sneezing, Hiccups, Gout, Gonorrhoea, Dyptheria, Dampiang, Mumps, Measles, Whooping cough, Tuberculosis, And even Bowden's Malady.



MITCHELL'S SNAKE OIL PURE PROOF CURE ALL LINAMENT

Providing the Finest in do-it-yourself health care since 2966

FOR BLINDNESS TRY OUR RATTLESNAKE OIL!



<https://aisnakeoil.substack.com/>

Non possiamo predire il futuro comportamento di singoli esseri umani. Nemmeno utilizzando l'IA.

Per la previsione dei risultati sociali, l'IA non è significativamente migliore di un calcolo manuale che utilizzi solo alcune caratteristiche.

Le applicazioni di IA che promettono di predire esiti sociali sono fondate su pseudoscienze, sono «olio di serpente».

A una verifica puntuale, tali sistemi risultano così poco affidabili, per la previsione di eventi e azioni individuali, che alcuni ricercatori suggeriscono piuttosto il ricorso a una lotteria tra le persone ammissibili.

Quando il responso dei sistemi di apprendimento automatico sia utilizzato a fini decisionali, la decisione produce ciò che pretende di prevedere: se il genere predice una paga più bassa e il colore della pelle predice la probabilità di essere fermati dalla polizia, con il passaggio dalla previsione alla decisione tale profilazione sociale si autoavvera.

A. Wang, S. Kapoor, S. Barocas, A. Narayanan, Against Predictive Optimization: On the Legitimacy of Decision-Making Algorithms that Optimize Predictive Accuracy, October 4, 2022, pp. 1-29, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4238015

L'etica dell'intelligenza artificiale come via di fuga dalla regolazione giuridica

La famiglia di narrazioni diffuse dai giganti tecnologici - chiamata "etica dell'IA" - rimuove l'idea che etichettare le persone come cose e trattarle come tali equivalga a negare loro il riconoscimento di qualsiasi diritto, danneggiando infallibilmente gli individui più deboli, e debba quindi essere vietato.

Invece di rifiutare semplicemente le decisioni statistiche automatizzate, le narrazioni sull'etica dell'IA riformulano la questione come un problema di equità algoritmica e di allineamento dei valori, come se

- l'unico problema fosse costituito da pregiudizi singoli ed emendabili;
- gli algoritmi potessero essere dotati delle competenze umane necessarie per formulare giudizi morali;
- i valori morali incorporati nei sistemi di ML potessero essere semplicemente scelti dagli ingegneri e tradotti in termini computazionali.

Etica dell'intelligenza artificiale: le condizioni non soddisfatte 1/2

L'etica dell'intelligenza artificiale richiederebbe una serie di condizioni, nessuna delle quali è attualmente soddisfatta:

1. dal punto di vista etico, sarebbe necessario individuare un'etica normativa che non ammetta l'esistenza di genuini dilemmi morali - e che quindi contenga i criteri per la soluzione di tutti i conflitti morali apparenti - e che sia condivisa in modo sufficientemente ampio da rendere pubblicamente ammessa la sua implementazione nei sistemi informatici.
2. dal punto di vista metaetico, sarebbe necessario affrontare la questione della traducibilità in termini computazionali dell'etica normativa adottata, o almeno di un suo sottoinsieme coerente.

Etica dell'intelligenza artificiale: le condizioni non soddisfatte 2/2

Anzitutto, **dovrebbero essere soddisfatti tutti i requisiti non morali dell'etica:** il giudizio morale richiede infatti

- la capacità di agire non solo secondo le leggi, ma anche secondo la rappresentazione delle leggi;
- il ragionamento logico,
- una comprensione autentica del linguaggio,
- la capacità di distinguere una connessione causale da una mera correlazione
- la famiglia di intuizioni e procedure di ragionamento incluse nel senso comune umano.

La capacità di formulare giudizi morali richiederebbe infatti un'intelligenza artificiale generale (AGI) e attualmente nessuno ha un'idea realistica non solo di come realizzarla, ma neppure di quali siano le domande di ricerca.

Oggi, i sistemi di intelligenza artificiale

1. non sono in grado di formulare neppure i giudizi morali più banali e condivisi, cioè di respingere alternative universalmente considerate come moralmente ripugnanti,
2. mostrano, con l'aumento della potenza e delle prestazioni, un aumento, anziché una diminuzione, della tossicità, ossia della tendenza a riprodurre e amplificare le discriminazioni riflesse nei dati di partenza.

Le auto a guida autonoma e il dilemma del trolley

Allo stato attuale della tecnologia, le auto «a guida autonoma» non sono a guida autonoma.

L'idea che, in caso di incidente, la tecnologia delle auto a guida autonoma ponga dilemmi etici straordinari - come le scelte che le auto stesse dovrebbero compiere - piuttosto che le consuete questioni di sicurezza, trasparenza, controllo e cautela, nasce dalla considerazione antropomorfa di tali veicoli come agenti morali.

Narrazione generata, a fini di marketing, dai produttori.

Operazione consueta, nel campo dell'etica dell'IA: la distrazione.

Quanto ai veicoli Tesla già in circolazione, di cui è stata millantata l'autonomia completa (con video rivelatisi poi mere costruzioni cinematografiche), gli sforzi dell'azienda (che vende, a migliaia di dollari, un optional denominato «Full self driving capability») si concentrano oggi nel sostenere, nelle sedi giudiziarie, che «il fallimento nella realizzazione di un ideale obiettivo a lungo termine» non equivalga a una frode.

H. Jin, *Tesla video promoting self-driving was staged, engineer testifies*, «Reuters», January 18, 2023, <https://www.reuters.com/article/tesla-autopilot-video-idTRNIKBN2TW1EB> ; P. Valdes-Dapena, *Tesla: Our 'failure' to make actual self-driving cars 'is not fraud'*, December 12, 2022, <https://edition.cnn.com/2022/12/12/business/tesla-fsd-autopilot-lawsuit/index.html>

Il principio dell'inevitabilità tecnologica

Il progresso tecnologico è considerato inarrestabile e di ogni singola tecnologia si dà per scontato che sia «qui per restare» o che «se non lo facciamo noi, lo farà qualcun altro». Qualsiasi dibattito ha luogo perciò entro la «logica del fatto compiuto», così che la possibilità di non costruire affatto alcuni sistemi o di non utilizzarli per alcune finalità non possa essere neppure contemplata.

«Il mito dell'inevitabilità tecnologica, politica e sociale è un potente tranquillante per la coscienza. Il suo servizio è quello di togliere la responsabilità dalle spalle di tutti coloro che ci credono veramente. Ma, in realtà, ci sono degli attori! [...] La reificazione di sistemi complessi che non hanno autori, di cui sappiamo solo che ci sono stati dati in qualche modo dalla scienza e che parlano con la sua autorità, non permette di porsi domande di verità o di giustizia.»

Joseph Weizenbaum, *Computer Power and Human Reason. From Judgment to Calculation*, San Francisco, W.H. Freeman & Company, 1976, pp. 241, 252, https://archive.org/details/computerpowerhum0000weiz_v0i3

Il principio di innovazione

Il principio di innovazione – fondato sul tacito assunto che qualsiasi innovazione tecnologica sia foriera di competitività e occupazione e debba perciò essere assecondata, anche a scapito del principio di precauzione – è in genere la maschera dietro la quale grandi aziende rivendicano la tutela dei loro concreti interessi economici.

Gli studi più recenti mostrano, del resto, che i monopoli del capitalismo intellettuale ostacolano qualsiasi innovazione, per quanto dirompente e benefica, che non si adatti al loro modello di business, promuovendo principalmente un'innovazione tossica che estrae o distrugge valore, anziché produrlo.

A. Saltelli, D.J. Dankel, M. Di Fiore, N. Holland, M. Pigeon, *Science, the Endless Frontier of Regulatory Capture*, in «Futures», 2022, n. 135.

I miti dell'eccezionalismo tecnologico e del vuoto giuridico

La tesi che le leggi vigenti non si applichino ai prodotti basati su sistemi di «intelligenza artificiale», in virtù della loro novità e straordinarietà, e che servano dunque nuove leggi, scritte ad hoc per ciascuna tecnologia, serve a dar luogo a una corsa che vedrà il legislatore perennemente in affanno, nel rincorrere le più recenti novità tecnologiche, le quali saranno dunque commercializzabili eslege.

Il problema non è l'assenza di una risposta giuridica, ma piuttosto il dubbio che può esistere sul significato di tale risposta (sarebbe allora più corretto parlare di vaghezza giuridica, in particolare in assenza di precedenti).

D. Tafani, *Sistemi fuori controllo o prodotti fuori legge? La cosiddetta «intelligenza artificiale e il risveglio del diritto*, <https://btfp.sp.unipi.it/it/2023/05/sistemi-fuori-controllo-o-prodotti-fuorilegge/>

La «bolla giuridica»

«Bolla giuridica» (Marco Giraudo): le grandi compagnie tecnologiche hanno fondato il loro modello di business sull'appropriazione e la commercializzazione dei dati personali, in violazione di diritti giuridicamente tutelati, scommettendo su un successivo “salvataggio giuridico”, in nome dell'inarrestabilità dell'innovazione tecnologica.

<https://www.cambridge.org/core/journals/journal-of-institutional-economics/article/on-legal-bubbles-some-thoughts-on-legal-shockwaves-at-the-core-of-the-digital-economy/62964F5EE993E8A7CC60EB4833FA71CA#article>

Il risveglio del diritto

In una Dichiarazione congiunta, la Federal Trade Commission e altre tre agenzie federali statunitensi hanno reso noto che, negli interventi di applicazione delle leggi vigenti, esse intendono utilizzare «con vigore» le loro prerogative «per proteggere i diritti degli individui, indipendentemente dal fatto che le violazioni del diritto avvengano attraverso mezzi tradizionali o tecnologie avanzate».

La pretesa che i nuovi sistemi di intelligenza artificiale siano emersi in un «vuoto giuridico» è respinta come infondata, sulla base della elementare constatazione che i codici non prevedono che l'intelligenza artificiale sia esentata dalle leggi.

Quanto al generale non funzionamento di intere categorie di prodotti che le aziende presentano come fondati su sistemi di «intelligenza artificiale», la Commissione esprime la propria consapevolezza di tali pratiche e la decisione di assimilarle a pubblicità ingannevole o a frode.

https://www.ftc.gov/system/files/ftc_gov/pdf/remarks-of-chair-lina-m-khan-re-joint-interagency-statement-on-ai.pdf

Grazie per l'attenzione

daniela.tafani@unipi.it

UNIVERSITÀ DI PISA

