

L'intelligenza artificiale tra narrazioni e sorveglianza

26 novembre 2023

Daniela Tafani

Dipartimento di Scienze politiche - Università di Pisa



SISTEMI DI APPRENDIMENTO AUTOMATICO



Nella famiglia di tecnologie denominata «intelligenza artificiale» – che si occupa di realizzare strumenti (software e hardware) che siano capaci di eseguire compiti normalmente associati all'intelligenza naturale – l'apprendimento automatico ha reso possibile, per alcuni specifici compiti non trattabili con l'intelligenza artificiale simbolica, un rapido e genuino progresso:

- previsione e generazione di stringhe di testo,
- riconoscimento facciale,
- ricerca per immagini,
- identificazione di contenuti musicali.

I sistemi di apprendimento automatico, di natura sostanzialmente statistica, consentono di costruire modelli a partire da esempi, in un processo iterativo di minimizzazione della distanza rispetto ai risultati attesi, purché si abbiano a disposizione

- potenti infrastrutture computazionali,
- enormi quantità di dati.

Le grandi aziende tecnologiche che, intorno al 2010, in virtù di un modello di business fondato sulla sorveglianza, detenevano già l'accesso al mercato necessario per l'intercettazione di grandi flussi di dati e metadati individuali e le infrastrutture di calcolo per la raccolta e l'elaborazione di tali dati, hanno potuto perciò raggiungere, con l'applicazione di algoritmi in gran parte già noti da decenni, traguardi sorprendenti.

“IA” E PRODOTTI CHE NON FUNZIONANO



Le grandi aziende tecnologiche hanno colto l'opportunità di un'espansione illimitata di prodotti e servizi «intelligenti»:

se un sistema di «intelligenza artificiale» è in grado di tradurre quello che scriviamo, perché non sostenere che sia anche in grado di comprenderlo?

Se può identificare un singolo individuo o classificarne correttamente alcuni tratti somatici, perché non sostenere che sia in grado altresì di riconoscere un ladro o un bravo lavoratore dalle loro fattezze esteriori o un malato di mente dalla voce ?

Perché non trasformare un sistema statistico, grazie alla polvere magica dell'«intelligenza artificiale», in un oracolo in grado di prevedere i futuri reati di ogni individuo o la «futura performance» scolastica di ogni singolo studente?

Giudizi e decisioni che hanno effetti rilevanti sulle vite di esseri umani sono oggi affidati, in un numero crescente di ambiti, a sistemi di intelligenza artificiale che non funzionano.

Tali malfunzionamenti non sono occasionali e non sono scongiurabili con interventi tecnici: essi rivelano, anzi, il funzionamento ordinario dei sistemi di apprendimento automatico.

Considerato il ruolo cruciale di tali sistemi nel modello di business delle grandi aziende tecnologiche, esse mirano a sottrarre tali prodotti all'intervento giuridico: è nata così, come operazione di cattura culturale, con l'obiettivo di rendere plausibile un regime di mera autoregolazione, l'«etica dell'intelligenza artificiale» (ora, «allineamento dei valori»).

INTELLIGENZA ARTIFICIALE E NARRAZIONI

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.

Tutela e legittimazione del modello di business delle grandi aziende tecnologiche. L'impostazione del discorso è dettata dalla sua funzione. Perciò:

- **determinismo tecnologico**: assunto dell'inevitabilità e «logica del fatto compiuto»;
- **soluzionismo**: questioni sociali trattate quali problemi di disciplinamento e controllo, passibili di soluzioni tecniche;
- **design tecnico** come 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;
- **principio di innovazione**;
- **mito della sostituzione del lavoro umano**;
- **miti dell'eccezionalismo tecnologico e del vuoto giuridico**;
- **narrazioni sull'«etica dell'IA»**.

From “Heavy Purchasers” of Pregnancy Tests to the Depression-Prone: We Found 650,000 Ways Advertisers Label You

A spreadsheet on ad platform Xandr’s website revealed a massive collection of “audience segments” used to target consumers based on highly specific, sometimes intimate information and inferences

By [Jon Keegan](#) and [Joel Eastwood](#)

- **Automobiles** (Example: “Past Purchases > Autos > Makes > Subaru”)
- **Demographics** (Ex: “Life Events > Newly Engaged”)
- **Business / B2B** (Ex: “B2B > Manufacturing > Candlemaking Equipment & Supplies”)
- **Retail stores** (Ex: “Brand Affinities > Retail > Prada”)
- **Interests** (Ex: “Psychographic Interests > Geek Culture”)
- **Brands** (Ex: “Wants to buy - Brands > The North Face “)
- **Grocery** (Ex: “Intent > Heavy Purchaser - Meat Pies - Refrigeration“)
- **Travel** (Ex: “Vacation Travel Attitudes > Not a Sightseer“)
- **Financial** (Ex: “Highest Risk > Poorer Unemployed Neighbourhoods”)
- **Political** (Ex: “US Politics > Issues & Advocacy > Allow Transgender Bathroom - Oppose”)
- **Health** (Ex: “Healthcare > Medications > Depression Medications”)

Political issues

Marijuana Reform Supporters Gun Control Advocates Gun Rights Advocates
 Womens Equality Advocates Immigration Control Advocates
 Immigration Rights Advocates Environmental Conservation Organized Labor Supporters
 Pro Choice Supporters Pro Life Supporters LGBTQ Advocacy
 Marriage Equality Opposition Animal Rights Supporters

Political activity

- Attended or willing to volunteer for a political protest
- Voting history by candidate, election, and state propositions
- Political donations for candidates and causes
- Voter registration status

Political profiles

Ultra-Conservative Streaming TV-Viewer Deep Root Analytics > Defund Police Persuadables
 Deep Root Analytics > Anti Defund Police Lifestyle > Political > Flags & Trump
 Lifestyle > Political > Doves Social Profiles by Type > Black Lives Matter Supporters

Location-based political targeting

- Extensive geofenced segments from Foursquare (Factual) and from political ad firms Rising Tide Interactive, DSPolitical. Voters by congressional district.



Diagnosis for

Atrial fibrillation Congestive heart failure Coronary artery disease Hearing loss
Hypertension Diabetes DVT Leukemia Non-hodgkins lymphoma ADHD Arthritis

Propensity for

Stroke Menopause Migraines Liver disease Insomnia IBS Heart disease
Fibromyalgia Erectile dysfunction Depression Diabetes
Exocrine pancreatic malfunction Obesity Smoking cessation Sleep apnea
Urinary tract infection

Likely symptoms of

Menstrual cramps Insomnia Migraines Sleep disorders Asthma DVT MS
Depression

Health relevance for

Polycystic kidney disease Chronic idiopathic constipation Chronic migraine
ADHD recent adult diagnosis

Reproductive health

Family planning Pregnancy / maternity Infertility / IVF Pregnancy and ovulation apps
Heavy Purchaser - Pregnancy Test Kits Heavy Purchaser - Male Contraceptives
K-Y Brand High > Heavy Buyer Trojan Brand High > Heavy Buyer
Clearblue Brand High > Heavy Buyer (Pregnancy test)
FirstResponse Brand High > Heavy Buyer (Pregnancy test)
Nature Made: Fertility and Ovulation - Cross Device
Nature Made: Pregnancy Location and Action - Cross Device
Viagra - Unhealthy Place Visits (adsquare)

Lifestyle

Tattoo Addicts Hipsters Indulgent Dog Owners Households with Trendy Moms (BlueKai)

Financial

Pensions & Ports Khakis & Credit Leveraged Life McMansions & Merriment
Aspirations and Dreams - Happiness Seekers Aspirations and Dreams - Love Aspirers
Aspirations and Dreams - Money Driven

Brand personas

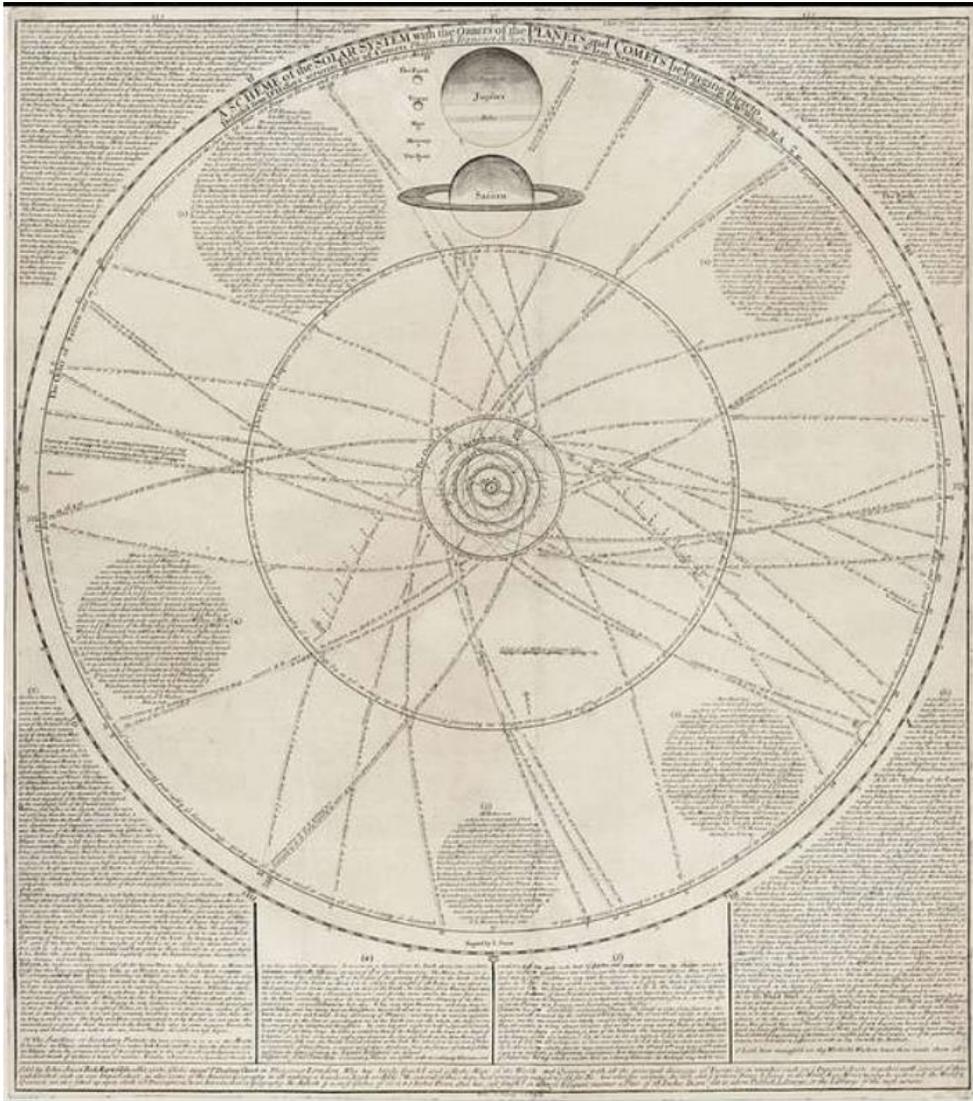
Lone Wolves Receptive to emotional messaging Rebellious spirits
Concerned with self-image Confident and Social Status Shopper

Personality

Extraversion - Jokers Extraversion - Party Animals Extraversion - Frustrated Extraverts
Neuroticism - Trapped Neuroticism - Stress Reactors Neuroticism - Self Lovers
Neuroticism - Easily Deflated Neuroticism - Internal Escapists
General Attitudes - I generally get a raw deal out of life Dealing with Stress - Hot and Cold
Dealing with Stress - Emotional Dealing with Stress - Bottled Up
Agreeableness: Disinterested Agreeableness - Rat Racers Agreeableness - Low
Love - Passionate Lovers Love - Rollercoaster Romantics

<https://themarkup.org/privacy/2023/06/08/from-heavy-purchasers-of-pregnancy-tests-to-the-depression-prone-we-found-650000-ways-advertisers-label-you>

LA “RIVOLUZIONE DELL’IA”?

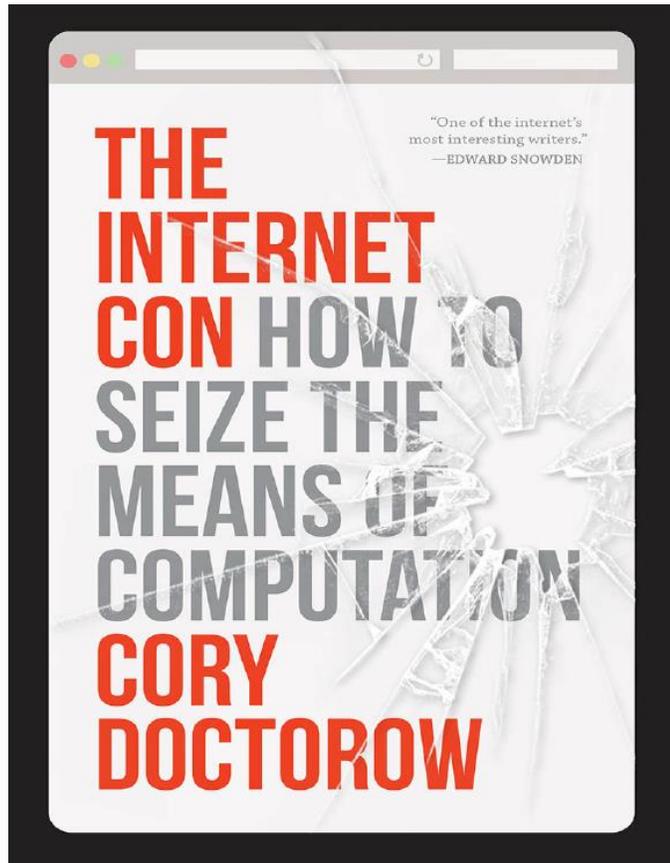


<https://picryl.com/media/whiston-senex-solar-system-chart-54bf6c>



<https://picryl.com/media/sans-titre-cortège-des-têtes-coupees-devant-lhotel-de-ville-place-de-greve-e86410>

IL MITO DEL MULTIMILIARDARIO GENIALE



KATHARINA PISTOR
**IL CODICE
DEL CAPITALE**

COME IL DIRITTO
CREA RICCHEZZA
E DISUGUAGLIANZA

POSTFAZIONE DI FRANCESCO DI CIOMMO,
SERGIO DI NOLA E MASSIMILIANO VATIERO

LUISS 

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 Judgement to Calculation*, San Francisco, W.H. Freeman & Company, 1976, pp. 241, 252, https://archive.org/details/computerpowerhum0000weiz_v0i3



L'“IA VERDE”?



<https://www.nbcnews.com/tech/internet/drought-stricken-communities-push-back-against-data-centers-n1271344>



II MITO DELLA SOSTITUZIONE DEL LAVORO

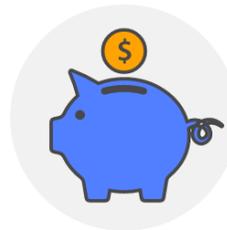


Create Tasks

Human intelligence through an API. Access a global, on-demand, 24/7 workforce.

Create a Requester account

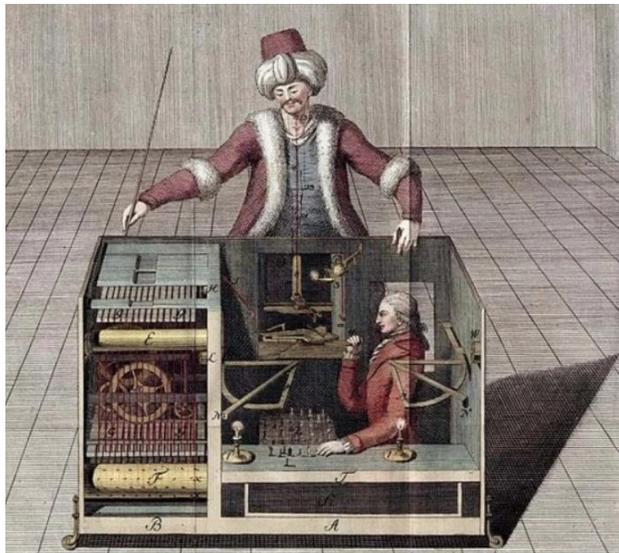
or



Make Money

Make money in your spare time. Get paid for completing simple tasks.

Request a Worker account



ONLINE GIG WORK IS GROWING IN DEVELOPING COUNTRIES

In developing countries, demand for online workers is outpacing that of developed countries. Nearly 60 % of firms surveyed in poorer countries reported increased outsourcing to gig workers. In wealthier countries less than half said the same.

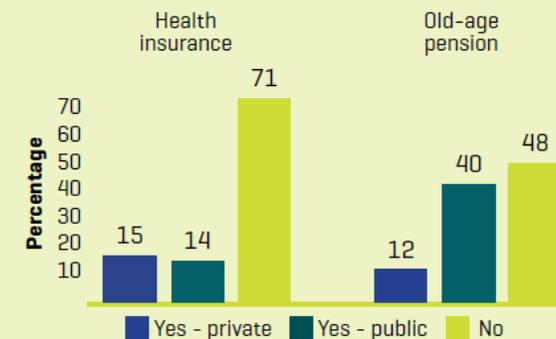
There is a total of **545 online gig work platforms** across the globe.

Close to **three quarters** of the platforms are **regional or local**.

With workers and clients located in **186 countries**.

Over 400 million people are estimated to be doing online gig work, often as a side job.

Do you subscribe to health insurance and an old-age pension?



TRE FALLACIE

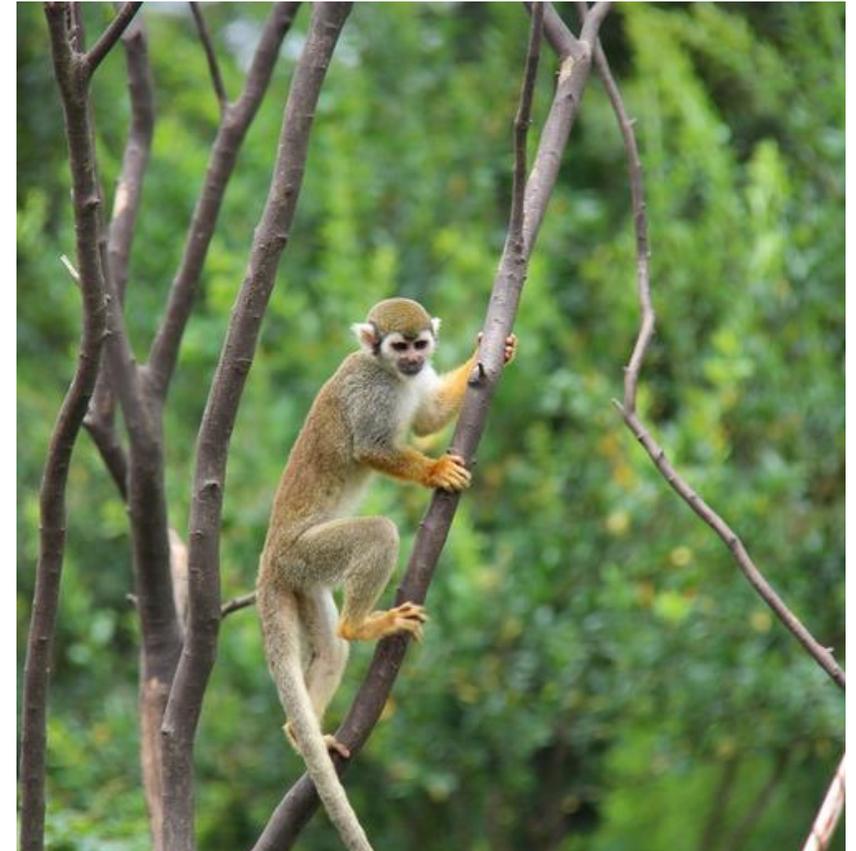
Fallacia della funzionalità dell'IA

Tendenza a credere che tutti i sistemi fondati sull'«intelligenza artificiale» funzionino, senza chiedersi se il nesso tra i dati e ciò che si intende prevedere sia fondato scientificamente.

Fallacia degli esempi tratti dal futuro o dalla fantascienza

Fallacia del primo passo

Assunzione che *machine learning* e AGI siano collocabili, in quanto omogenei, lungo un continuum. Hubert Dreyfus illustrava la fallacia del primo passo, con le parole del fratello Stuart, come equivalente alla tesi «che la prima scimmia che si arrampica su un albero stia facendo progressi verso lo sbarco sulla luna».



IL PAESE DI ACCHIAPPACITRULLI



La 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 impercettibile di un sistema automatico.

Il controllo umano

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.

RICERCA O COMMERCIO? IL PRINCIPIO DI INNOVAZIONE

Un dispositivo per diffondere impunemente prodotti insicuri e incolpare sistematicamente utenti e clienti è l'offuscamento della linea di confine tra la fase di ricerca e sperimentazione e quella della distribuzione e commercializzazione di prodotti.

OpenAI, ad esempio, rende pubblicamente disponibili i propri generatori di linguaggio naturale, in fase sperimentale, ne distribuisce al tempo stesso versioni a pagamento e fissa termini d'uso in virtù dei quali ogni utente è responsabile, secondo una logica analoga a quella del paese di Acchiappacitrulli, tanto dei propri input quanto degli output prodotti dal sistema (sui quali, ovviamente, non ha alcuna facoltà di controllo o decisione e i cui fondamenti gli sono del tutto oscuri).

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.

L'“ETICA” COME SPECCHIETTO PER LE ALLODOLE



Ethics washing: le narrazioni sull'etica dell'IA (o su «allineamento dei valori» o «AI safety») sono uno specchietto per le allodole, una via di fuga dalla regolazione giuridica.

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 operationalizzare 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.

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).

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.

LA «BOLLA GIURIDICA» E IL RISVEGLIO DEL DIRITTO



“There is no AI exemption to the laws on the books”

Lina M. Khan

«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 della presunta 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 SOLUZIONISMO TECNOLOGICO



«Riformulare tutte le situazioni sociali complesse come problemi ben definiti con soluzioni definite e computabili o come processi trasparenti e autoevidenti che possono essere facilmente ottimizzati, se solo ci sono gli algoritmi giusti! Questa ricerca rischia di avere conseguenze inaspettate che potrebbero causare più danni dei problemi che cercano di risolvere.»

«Questi sono i tipi di problemi che, ad un esame attento, non devono essere definiti nei modi singolari e onnicomprensivi in cui li hanno definiti i "risolutori"; ciò che è controverso, quindi, non è la loro proposta di soluzione ma la definizione stessa del problema.

Il soluzionismo, quindi, non è solo un modo elegante di dire che per chi ha un martello, tutto sembra un chiodo. Non è solo che molti problemi non sono adatti al kit di strumenti facile e veloce del risolutore. È anche che ciò che molti soluzionisti presumono siano "problemi" da risolvere non sono affatto problemi.»

E. Mozorov, *To Save Everything, Click Here: The Folly of Technological Solutionism*, New York, Public Affairs, 2013.



MESSIANESIMI EUGENETICI E RISCHI FANTASCIENTIFICI

Una congerie di utopie per multimiliardari bianchi invita l'opinione pubblica a occuparsi del futuro dell'umanità, anziché del presente, e vagheggia scenari di rischi fantascientifici o in cui il bene dell'umanità sarà il frutto delle tecnologie su cui questi stessi multimiliardari intendono, di volta in volta, lucrare.

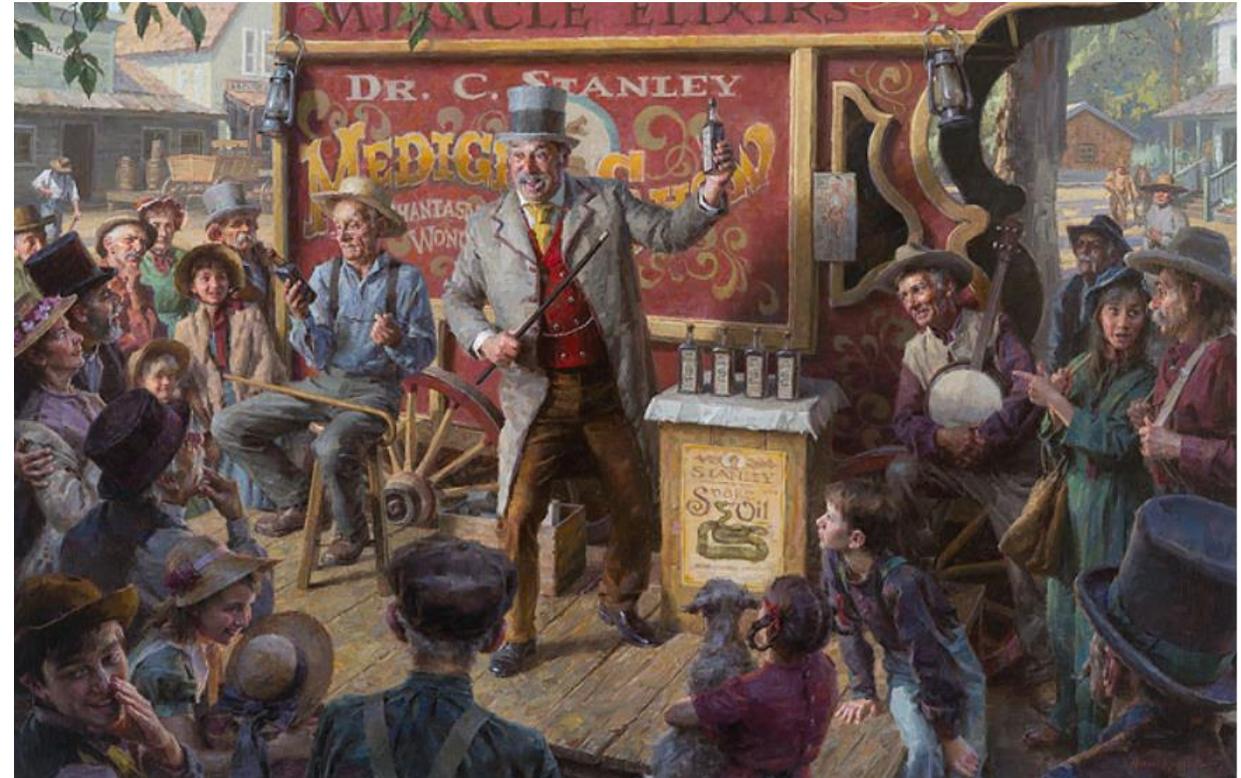
Filosoficamente inconsistenti e massicciamente finanziate, hanno in comune deliri eugenetici e un concetto di "umanità" che esclude gran parte degli esseri umani.



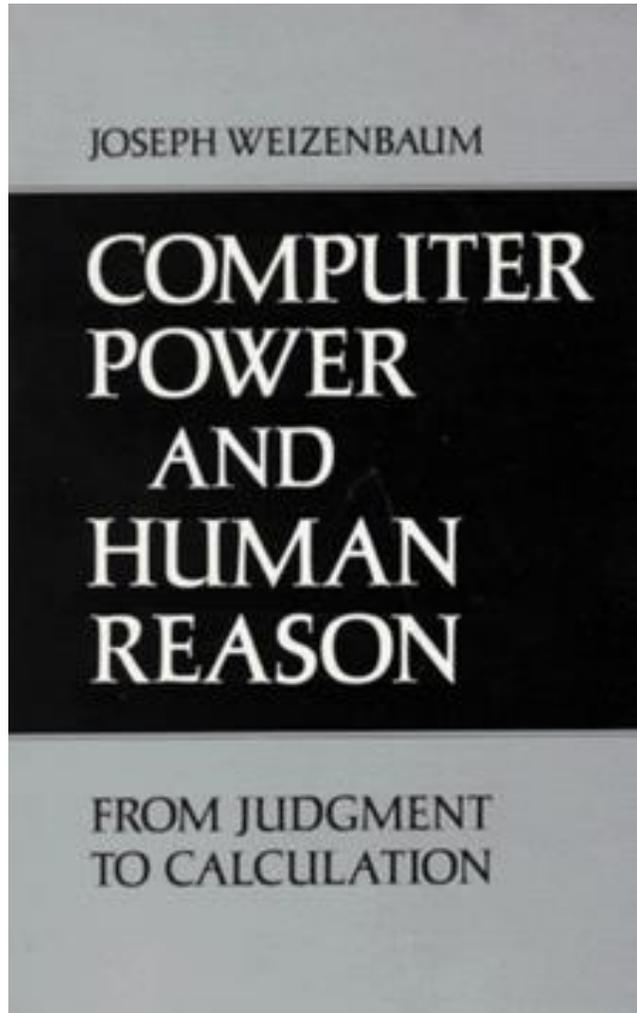
INTELLIGENZA ARTIFICIALE E PENSIERO MAGICO

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).



L'ANIMISMO



La tendenza umana a concepire gli oggetti della tecnologia in termini antropomorfici è nota e tuttavia irresistibile: risposte emotive e sociali sono generate automaticamente anche da mezzi di comunicazione, quali i televisori o i computer.

Tendiamo a inferire, da una singola, limitata prestazione di un sistema di intelligenza artificiale, che il sistema possieda le competenze che sarebbe ragionevole ascrivere a una persona in grado di svolgere lo stesso compito.

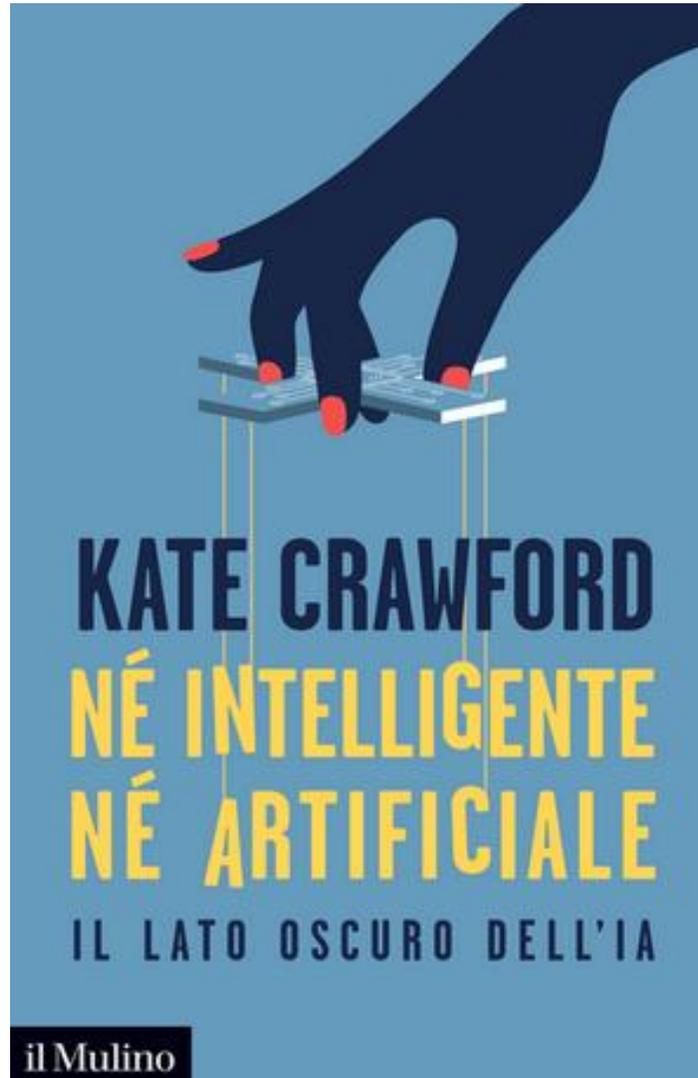
Es. generatori di linguaggio: non vale «parla, dunque pensa».

L'antropomorfizzazione delle macchine è sempre correlata alla deumanizzazione delle persone.

Es. Joseph Weizenbaum e Eliza Doctor.



LA MOSSA DA PRESTIGIATORI



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).
6. **Lavoro:** colonialismo digitale; sfruttamento del lavoro, lavoro che compromette la salute mentale.

I SISTEMI DI OTTIMIZZAZIONE PREDITTIVA

La tesi che sistemi di apprendimento automatico siano in grado di prevedere eventi o azioni future di singole persone non ha alcun fondamento scientifico

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».

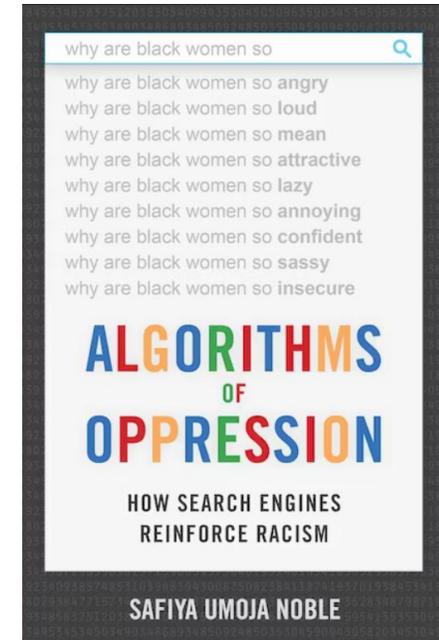
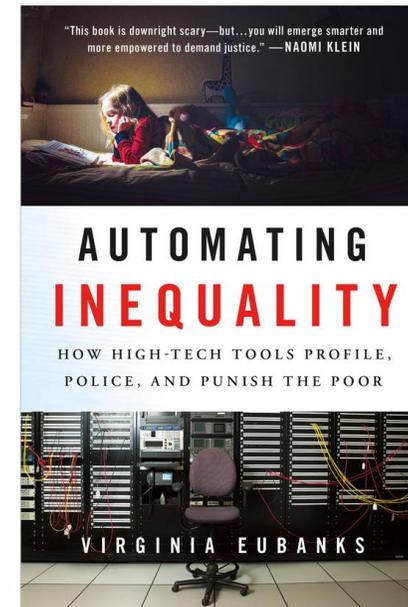
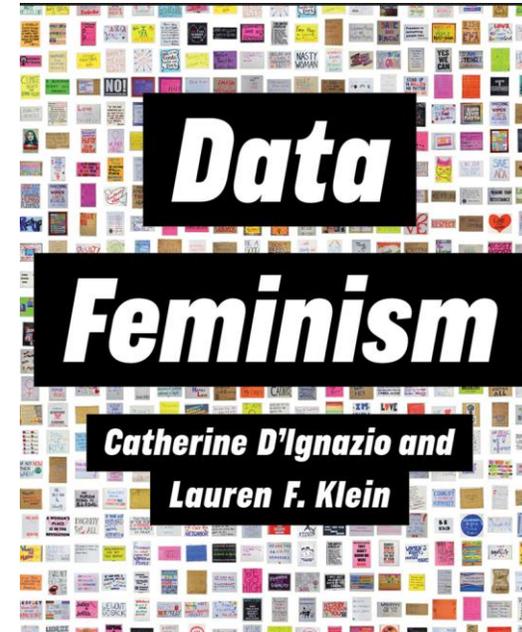
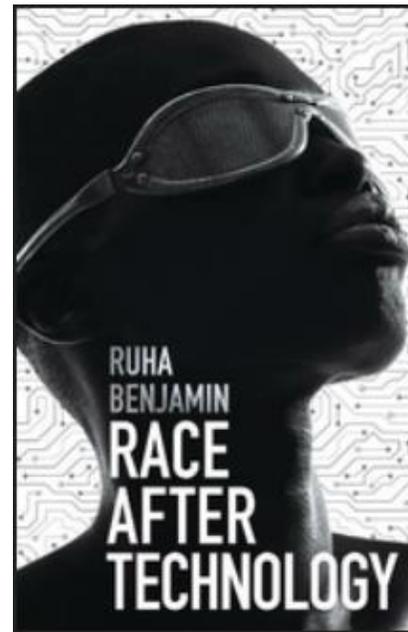
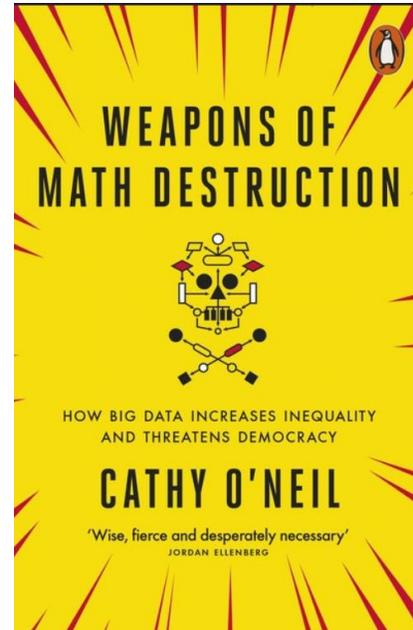
A una verifica puntuale, tali sistemi risultano così poco affidabili, per la previsione di eventi e azioni individuali, che alcuni ricercatori suggeriscono piuttosto il ricorso a una lotteria tra le persone ammissibili.

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.

IA, GIUDIZI E DECISIONI

La fine di un mito

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.



LE DISCRIMINAZIONI COME DIVERSIVO

Le decisioni basate sui sistemi di apprendimento automatico sono costitutivamente discriminatorie, in quanto procedono etichettando le persone e raggruppandole in varie classi, secondo qualsiasi tipo di regolarità nei dati di partenza.

Essendo radicata nella natura statistica di questi sistemi, la caratteristica di dimenticare i "margini" è strutturale: non è accidentale e non è dovuta a *bias* singoli e tecnicamente modificabili. Le decisioni algoritmiche

- replicano attraverso l'automazione le discriminazioni e le disuguaglianze del passato;
- allo stesso tempo, poiché i loro modelli si basano su mere correlazioni, generano nuove e imprevedibili discriminazioni sulla base di fattori irrilevanti.

Discriminazioni contro:

- gruppi tutelati dalla legge;
- “gruppi algoritmici”, non previste dalla legge in virtù della loro insensatezza (es. proprietari di cani, adolescenti tristi, video giocatori);
- gruppi creati sulla base di caratteristiche, come la configurazione dei pixel in una foto o il mero ordine di presentazione dei dati, che non sono significativamente riconducibili a individui e sulla base dei quali, invece, possono avvenire trattamenti differenziati.

VALUTARE I RISCHI O DIFENDERE I DIRITTI?



AI ACT. Due possibili impostazioni della valutazione dei sistemi di IA:

1. Approccio basato sui rischi

Una mera valutazione d'impatto relativa ai rischi riduce la politica della tecnologia a un esercizio tecnocratico di calcolo dei rischi, partendo dal presupposto che il cambiamento tecnologico rifletta il corso inevitabile del progresso scientifico, anziché una scelta umana.

2. Approccio basato sui diritti

«Ci si può chiedere se tali sistemi possano essere utilizzati da qualcuno, oltre che dai governi e dalle grandi aziende, e se tali organizzazioni non li utilizzino principalmente per scopi anti-umani. Oppure si pensi ai sistemi di riconoscimento vocale. Non saranno usati principalmente per spiare le comunicazioni private? Rispondere a queste domande dicendo che i grandi sistemi informatici, le reti di computer e i sistemi di riconoscimento vocale sono inevitabili significa rinunciare alla propria umanità. Perché una risposta del genere deve basarsi o sulla convinzione che la società abbia già perso il controllo sulla sua tecnologia o sulla posizione assolutamente immorale che "se non lo faccio io, lo farà qualcun altro".»

«Chiunque detti le domande determina in gran parte le risposte» (Joseph Weizenbaum)

D. Tafani, *Do AI systems have politics? Predictive optimisation as a move away from liberalism, the rule of law and democracy*, (submitted).

Corporate Europe Observatory, *The lobbying ghost in the machine. Big Tech's covert defanging of Europe's AI Act*, February 23, 2023, <https://corporateeurope.org/en/2023/02/lobbying-ghost-machine>

P. van Zwanenberg, *The Unravelling of Technocratic Orthodoxy? Contemporary knowledge politics in technology regulation*, in I. Scoones, A. Stirling (eds.), *The Politics of Uncertainty. Challenges of Transformation*, Routledge, London 2020, pp. 58-72.

J. Weizenbaum, *On the Impact of the Computer on Society*, "Science", 176, n. 4035 (1972), pp. 609-614, <https://www.jstor.org/stable/1734465>

«AUTOMATICAMENTE ILLEGALI»



Illegalità dei sistemi di intelligenza artificiale, fino a prova contraria: è la proposta di Frank Pasquale e Gianclaudio Malgieri, fondata sull'evidenza delle violazioni dei diritti individuali che hanno luogo quando i sistemi di intelligenza artificiale sono utilizzati per ottenere classificazioni o produrre decisioni che hanno effetti rilevanti sulle vite delle persone.

I modelli di IA ad alto rischio incorporati oggi in prodotti e servizi dovrebbero essere disciplinati entro un sistema di “illegalità di default”: fino a prova contraria, tali sistemi dovrebbero essere considerati illegali.

Prima di immettere sul mercato un prodotto o un servizio che incorpori sistemi di IA ad alto rischio, le aziende avrebbero l'obbligo di dimostrare che la loro tecnologia non è discriminatoria, non è manipolatoria, non è iniqua, non è inaccurata e non è illegittima nelle sue basi giuridiche e nei suoi scopi.

<https://brusselsprivacyhub.com/2022/06/01/brussels-privacy-hub-working-paper-from-transparency-to-justification-toward-ex-ante-accountability-for-ai-by-gianclaudio-malgieri-and-frank-pasquale-has-been-published/>

Zero Trust AI Governance

Published in
August 2023

+accountable
tech

AINOW

epic.org / ELECTRONIC
PRIVACY
INFORMATION
CENTER

a/

Principle 1

Time is of the essence - start by vigorously enforcing existing laws.

Industry leaders have deployed numerous tactics to cast themselves as thoughtful while delaying accountability. They've played up the long-term threat of human *extinction*, asked Congress to create a *new agency*, and *heaped praise* on proposals that would slow-walk action – all while continuing to drive the AI arms race forward at a breakneck speed. But concrete harms from these systems are already being felt, and advancing as rapidly as AI itself. As officials across federal enforcement agencies have *underscored*, there is no AI exemption from the laws on the books; enforcing them swiftly and vigorously is a critical first step toward mitigating automated harms and deterring the reckless deployment of unsafe systems.

Principle 2

Bold, easily administrable, bright-line rules are necessary.

It should be clear by now that self-regulation will fail to forestall AI harms. The same is true for any regulatory regime that hinges on voluntary compliance or otherwise outsources key aspects of the process to industry. That includes complex frameworks that rely primarily on auditing – especially first-party (internal) or second-party (contracted vendors) auditing – which Big Tech has increasingly embraced. These approaches may be strong on paper, but in practice, they tend to further empower industry leaders, overburden small businesses, and undercut regulators' ability to properly enforce the letter and spirit of the law.

Principle 3

At each phase of the AI system lifecycle, the burden should be on companies to prove their systems are not harmful.

Industry leaders have taken a range of voluntary steps to demonstrate a commitment to key ethical AI principles. But they've also *slashed AI ethics teams*, *ignored internal alarms*, *abandoned transparency* as the arms race has escalated, and sought to *pass accountability* off to downstream users and civil society. Rather than relying on the good will of companies, tasking under-resourced enforcement agencies or afflicted users with proving and preventing harm, or relying on post-market auditing, *companies* should have to prove their AI offerings are *not* harmful.



I GENERATORI DI LINGUAGGIO

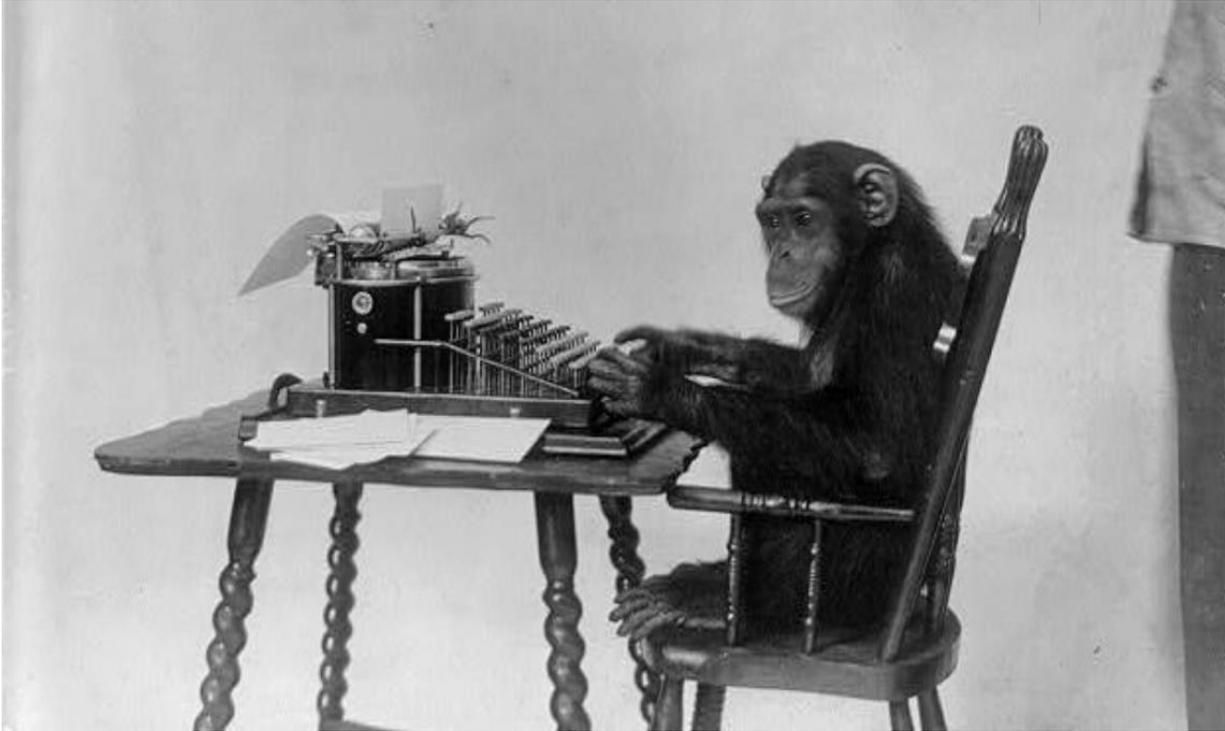


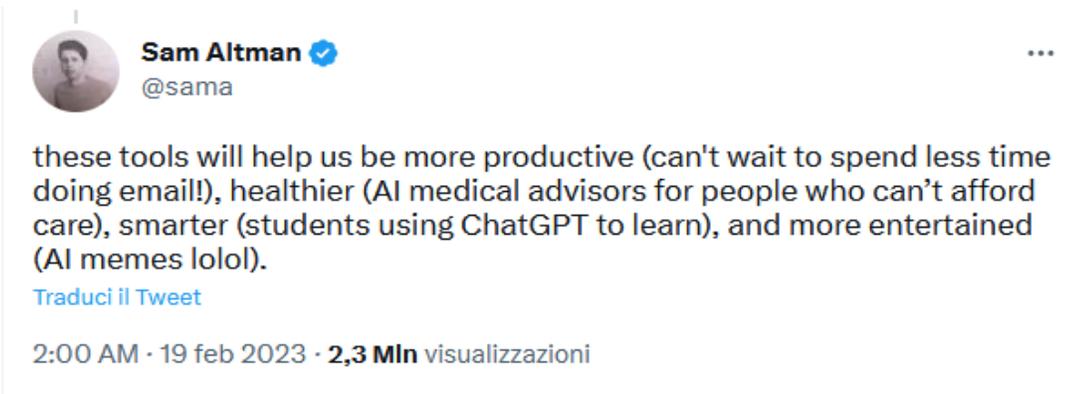
Illustration - [Chimpanzee seated at a typewriter] (Picryl - Library of Congress), <https://picryl.com/media/chimpanzee-seated-at-a-typewriter>

I modelli del linguaggio naturale sono programmi informatici che utilizzano le statistiche sulla distribuzione delle parole per produrre altre stringhe di parole.

Un generatore di linguaggio non è programmato per produrre risposte corrette, ma sequenze di parole che siano statisticamente simili ai testi su cui è stato programmato o alle risposte che sono state formulate dagli esseri umani che lo hanno «addestrato».

Sistema che produce stringhe di testo plausibili e pertinenti, prive di intento comunicativo e valore informativo: produce non-informazione.

ANTROPOMORFIZZAZIONE E DEUMANIZZAZIONE

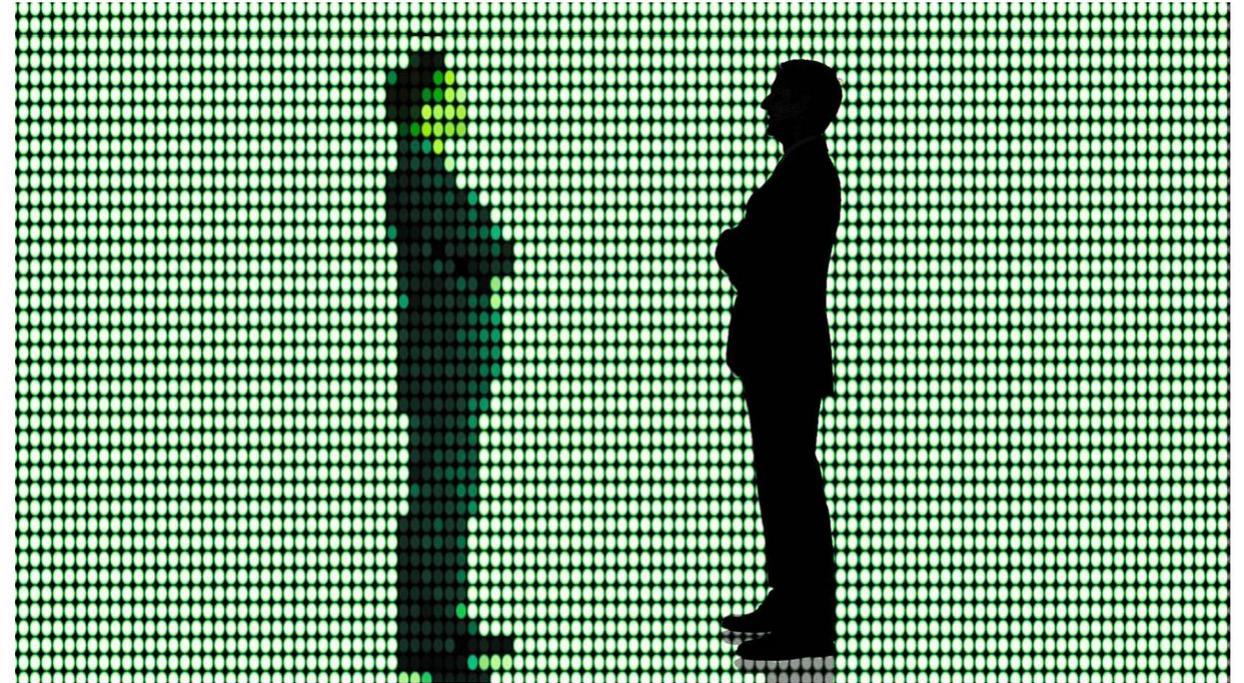


- Sistema di autocompletamento che scrive «io» e «mi dispiace»: ingannevole *by design*.
- Strumenti non generativi in senso proprio, ma normalizzanti: riproducono il pensiero egemonico. Aspetto normativo.
- Testi sintetici che inquinano l'ecosistema dell'informazione.
- Narrazione che oscura la differenza tra esseri umani e sistemi informatici: gesto politico.
- Plagio automatizzato.
- Sorveglianza.
- Modelli non sicuri, rilasciati pubblicamente anzitempo, anche per affidare alla collettività il lavoro di test, svolto gratuitamente.
- Convergenza con politiche neoliberali.

PERSONE CONTRAFFATTE E RESPONSABILITA'

«Il denaro esiste da diverse migliaia di anni e fin dall'inizio la contraffazione è stata riconosciuta come un crimine molto grave, che in molti casi richiede la pena capitale perché mina la fiducia da cui dipende la società. Oggi, per la prima volta nella storia, grazie all'intelligenza artificiale, è possibile per chiunque creare persone contraffatte che possono passare per vere in molti dei nuovi ambienti digitali che abbiamo creato».

«Prima che sia troppo tardi (forse è già troppo tardi) dobbiamo mettere al bando sia la creazione di persone contraffatte sia il "passaggio" di persone contraffatte. Le pene per entrambi i reati dovrebbero essere estremamente severe.»



The Atlantic / Getty

D.C. Dennett, *The Problem With Counterfeit People*, May 16, 2023,

<https://www.theatlantic.com/technology/archive/2023/05/problem-counterfeit-people/674075/>

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.



INTELLIGENZA ARTIFICIALE: NON È COME SEMBRA

«IA» che non è IA:

- vecchie tecnologie spacciate per IA (perché l'uso della vecchia tecnologia è illegale e l'«IA» no)
es. «AI webcam» per la telesorveglianza dei lavoratori
- fauxtomatic o fake automation: tecnologia presentata come «IA», che ha bisogno di un umano dietro le quinte, che svolga una parte del lavoro o, a volte, tutto il lavoro, salvo l'interfaccia
es. Kiwibots
es. presunti veicoli a guida autonoma Cruise
richiedono costanti controlli a distanza
(3 lavoratori impiegati ogni 2 auto)

AI snake oil: prodotti che utilizzano un sistema di IA e che non funzionano perché vengono usati per funzioni che

- quei sistemi non possono svolgere
es. controlli sui migranti alla frontiera
- non sono possibili *tout court*
es. fisiognomica nel reclutamento del personale
es. riconoscimento delle malattie mentali dalla voce

IL RITORNO DELLA FISIOGNOMICA

Rivive, con l'IA, la pseudoscienza della fisiognomica, l'idea di una connessione sistematica tra aspetto esteriore (volti, ma non solo) e interiorità (orientamento sessuale, orientamento politico, tratti del carattere, aspetti della personalità)

L. Stark, J. Hutson, *Physiognomic Artificial Intelligence*, «Fordham Intellectual Property, Media and Entertainment Law Journal», XXXII, 4, 2022, pp. 922-978, <https://ir.lawnet.fordham.edu/iplj/vol32/iss4/2>

TECHNOLOGY

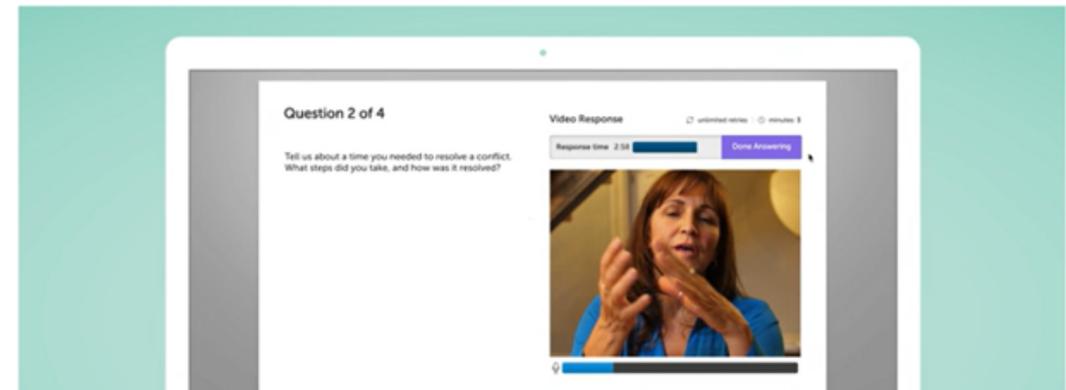
A face-scanning algorithm increasingly decides whether you deserve the job

HireVue claims it uses artificial intelligence to decide who's best for a job. Outside experts call it 'profoundly disturbing.'



By Drew Harwell

November 6, 2019 at 12:21 p.m. EST



<https://www.washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/>

«CREATIVITÀ?»



<https://arstechnica.com/ai/2023/11/fake-movie-posters-with-disney-logos-force-microsoft-to-alter-bing-image-creator/>



Grazie per l'attenzione
Domande?

daniela.tafani@unipi.it

a/simmetrie – Associazione italiana per lo studio delle
asimmetrie economiche

Seguici su    

www.asimmetrie.org