

# LE PROPRIETA POLITICHE DEI SISTEMI DI INTELLIGENZA ARTIFICIALE

Daniela Tafani

29 febbraio 2024

# 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».

Di ogni sistema di «IA» dovremmo sempre chiederci:

- Che cosa fa
- A chi lo fa?
- Per chi lo fa?

# IA: DI COSA STIAMO PARLANDO

## Intelligenza artificiale simbolica

Così come l'informatica, deriva dalla logica, ossia dalla disciplina che studia il ragionamento deduttivo.

I sistemi simbolici sono sempre in grado di spiegare perché hanno fornito una determinata soluzione o hanno preso una qualsiasi «decisione». Questo perché le regole della logica usate per effettuare la deduzione possono essere usate anche per spiegare il perché della deduzione stessa.

## Intelligenza artificiale sub-simbolica

I sistemi di apprendimento automatico (machine learning) sono di natura statistica. Il focus si è spostato dai linguaggi di programmazione usati per «istruire» la macchina ai dati usati per permettere alla macchina di «imparare» in modo autonomo

I sistemi sub-simbolici, al momento, non hanno la capacità di fornire spiegazioni per i risultati prodotti.

# IA RISTRETTA E IA GENERALE



Source: [cacm.acm.org](http://cacm.acm.org)

Ernest Davis and Gary Marcus. 2015. Commonsense reasoning and commonsense knowledge in artificial intelligence. *Commun. ACM* 58, 9 (August 2015), 92-103.

## 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.



# SISTEMI DI APPRENDIMENTO AUTOMATICO (MACHINE LEARNING)

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.

# INTELLIGENZA ARTIFICIALE E HYPE

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» o «IA costituzionale»).

# L'IA COME DERIVATO DELLA SORVEGLIANZA

Privacy

## 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>

# TRE FALLACIE

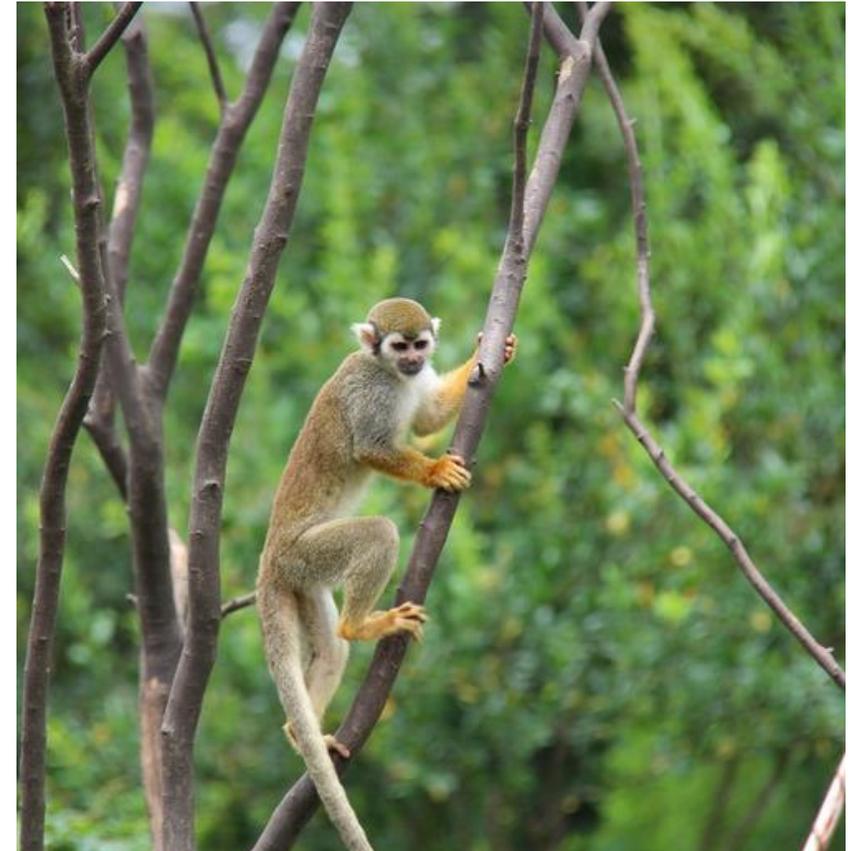
## Fallacia della funzionalità dell'IA

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

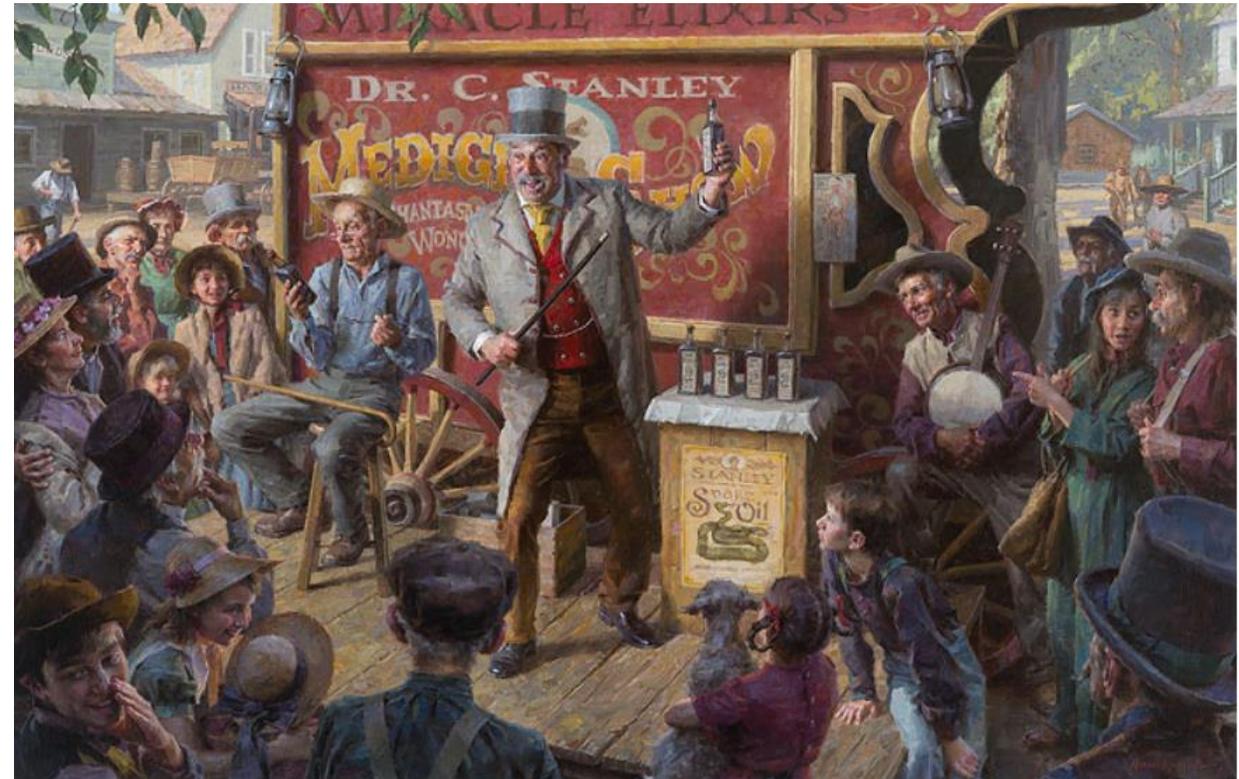
Assunzione che *machine learning* e AGI siano collocabili, in quanto omogenei, lungo un continuum. Hubert Dreyfus illustrava la fallacia del primo passo, con le parole del fratello Stuart, come equivalente alla tesi «che la prima scimmia che si arrampica su un albero stia facendo progressi verso lo sbarco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



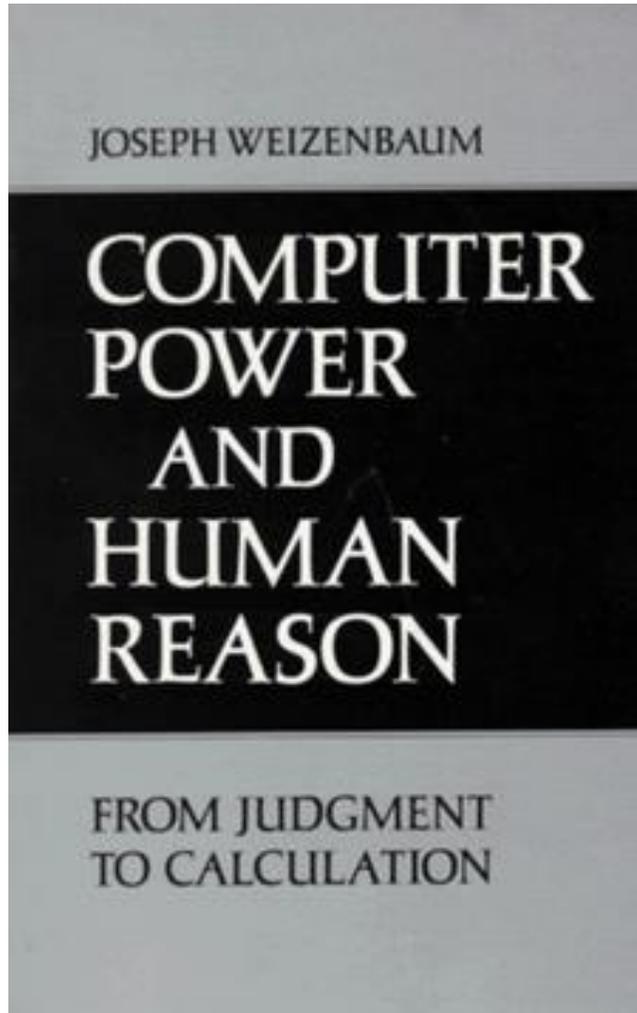
# INTELLIGENZA ARTIFICIALE E PENSIERO MAGICO

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.

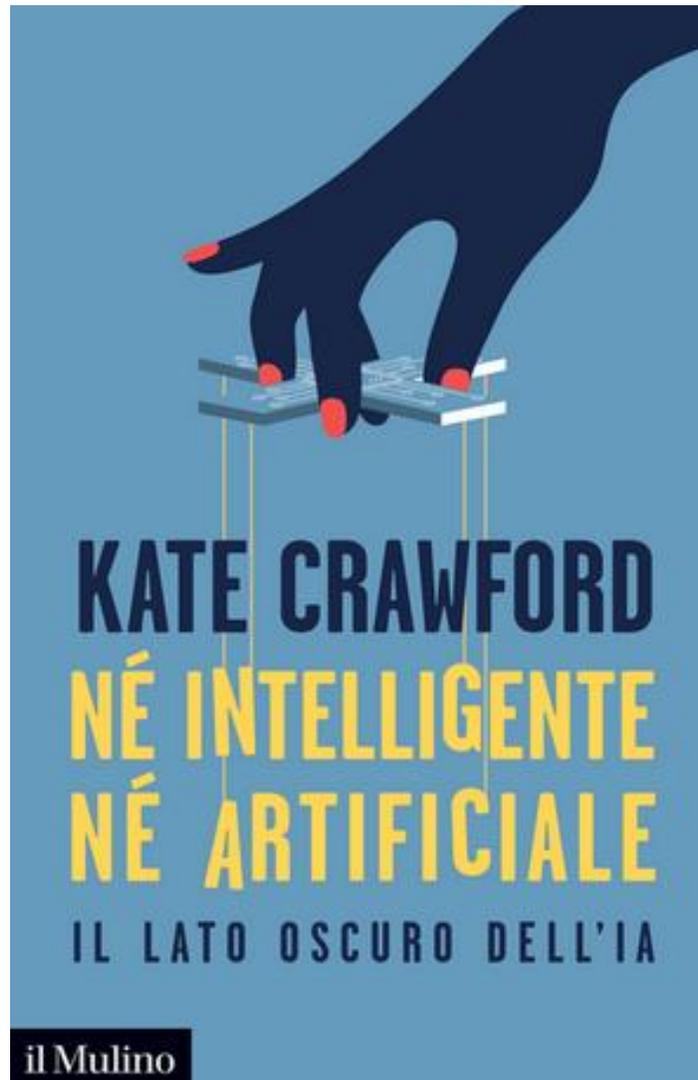
Tendiamo a inferire, da una singola, limitata prestazione di un sistema di intelligenza artificiale, che il sistema possieda le competenze che sarebbe ragionevole ascrivere a una persona in grado di svolgere lo stesso compito.

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

# L'“IA VERDE”?



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



# II MITO DELLA SOSTITUZIONE DEL LAVORO

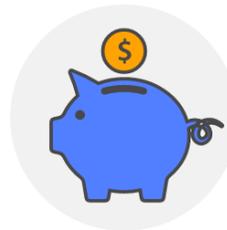
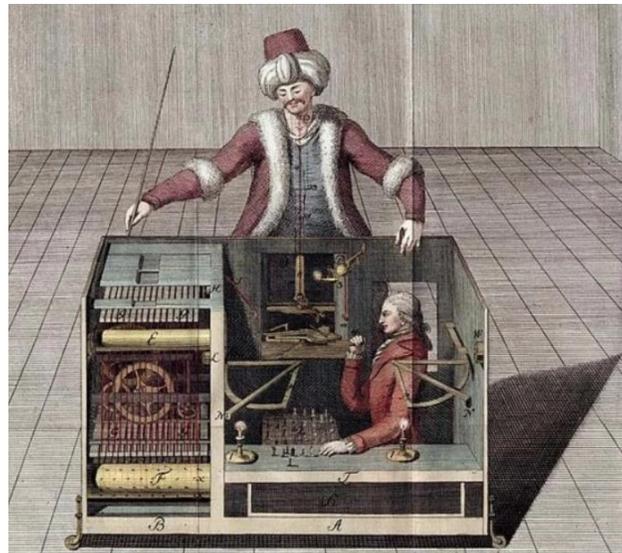


## Create Tasks

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.

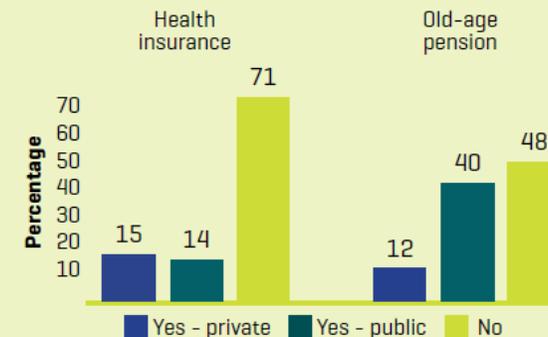
There is a total of **545 online gig work platforms** across the globe.

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?



# 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.

# 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

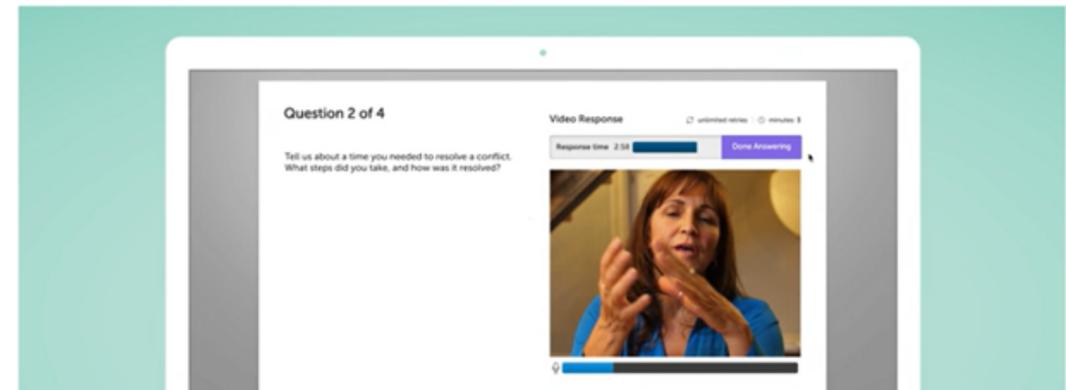
## A face-scanning algorithm increasingly decides whether you deserve the job

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/>



**HE COULD BE THE SHOOTER, HE MIGHT GET SHOT. THEY DIDN'T KNOW. BUT THE DATA SAID HE WAS AT RISK EITHER WAY**

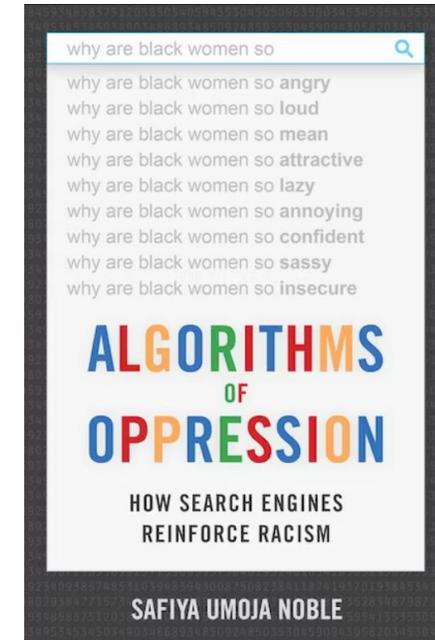
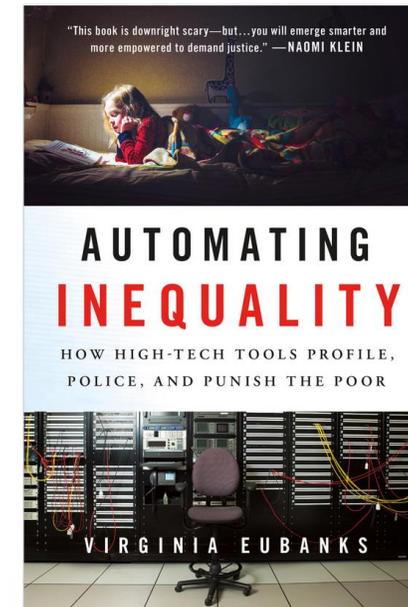
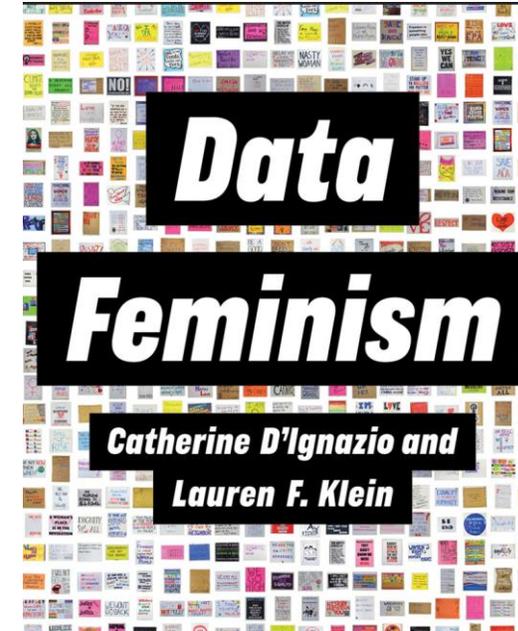
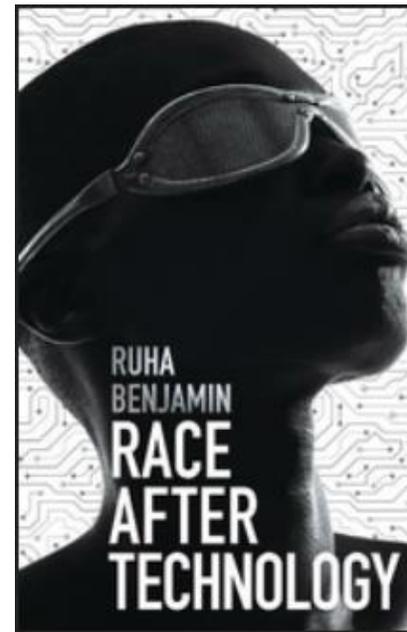
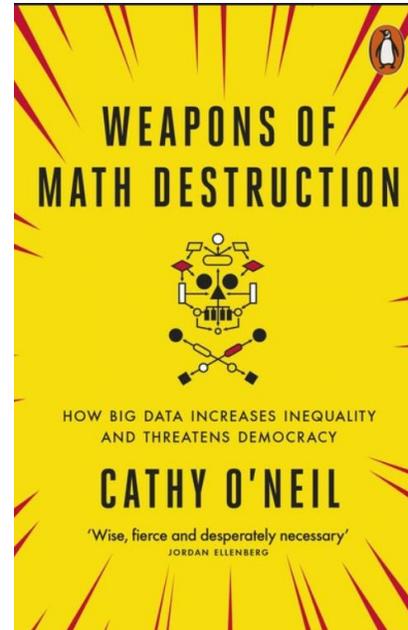
Chicago's predictive policing program told a man he would be involved with a shooting.

**IT WASN'T HIGH-TECH — COPS WOULD JUST USE THE LIST AS A WAY TO TARGET PEOPLE**

# 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.



# INTELLIGENZA ARTIFICIALE E DISCRIMINAZIONI

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.

# 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.

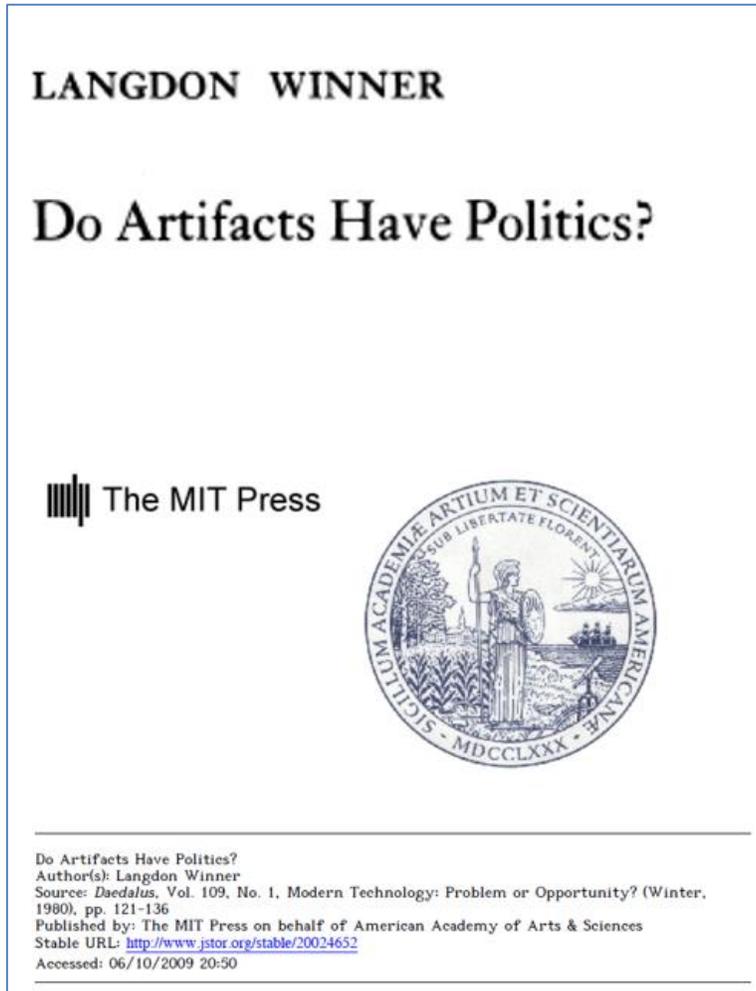
I detentori di grandi monopoli intellettuali finanziano la ricerca sui temi dell'intelligenza artificiale, per assicurarsi che siano coerenti con il loro modello di business

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

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».**

# GLI ARTEFATTI HANNO PROPRIETÀ POLITICHE?



Un artefatto può

- essere il modo in cui si affronta una questione in una specifica comunità;
- manifestare effetti politici dopo la sua adozione;
- richiedere o essere fortemente compatibile con una specifica relazione politica (es. tecnologie autoritarie e tecnologie democratiche; nucleare e solare).

Langdon Winner, *Do Artifacts Have Politics?*, in «Daedalus», 1980, 1, pp. 121-136.



# LE PROPRIETÀ POLITICHE INTRINSECHE

1. L'"invenzione, la progettazione o la disposizione di uno specifico dispositivo o sistema tecnico" possono essere "un modo per risolvere una questione negli affari di una particolare comunità":
    - i cavalcavia straordinariamente bassi di Long Island, a New York, furono progettati da Robert Moses per negare l'accesso agli autobus e quindi ai poveri e alle minoranze razziali che non possedevano automobili;
    - le macchine introdotte da Cyrus McCormick nella sua fabbrica di Chicago alla fine del XIX secolo, sebbene costose e inefficienti, furono aggiunte alla fonderia semplicemente per distruggere il sindacato degli operai specializzati.
  2. Esistono "**tecnologie intrinsecamente politiche**", che "sembrano richiedere o essere fortemente compatibili con particolari tipi di relazioni politiche":
    - una nave in mare, scriveva Platone, ha bisogno di un unico capitano e di un equipaggio obbediente;
    - l'organizzazione di una grande fabbrica, osservava Engels, richiede un governo centralizzato e autoritario, che tenderà a diffondersi nel resto della società come modello di velocità ed efficienza.
- Così, non solo le "cose che chiamiamo 'tecnologie'" configurano forme di vita; a volte, la scelta di una particolare tecnologia comporta la scelta di una particolare forma di vita politica, che non può essere cambiata senza abbandonare del tutto quella tecnologia

# OTTIMIZZAZIONE PREDITTIVA E SISTEMI NEOLIBERALI

«si usa designare come neoliberismo o neoliberalismo un'ideologia universale che afferma che qualunque settore della società, ciascun individuo in essa e, infine, la società intera in quanto somma dei due elementi precedenti, può funzionare meglio, costare meno, presentare minor problemi, essere più efficace ed efficiente qualora sia governata in ogni momento dai principi di una razionalità economica e strumentale»

(Luciano Gallino)

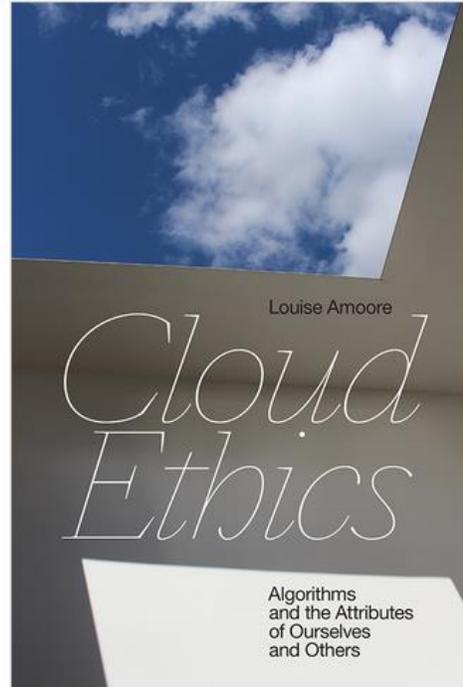
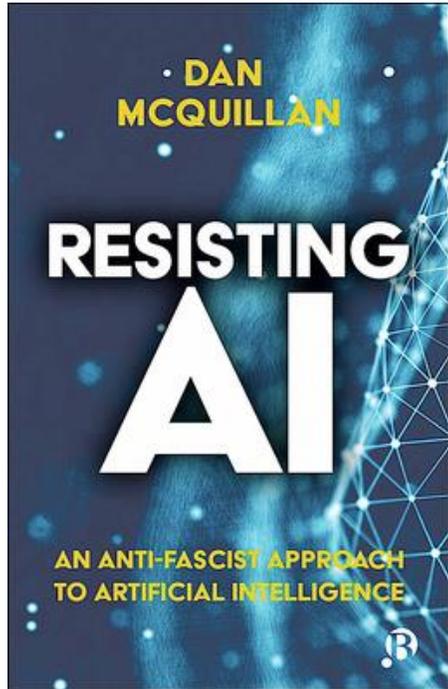
L'ottimizzazione predittiva rispecchia la logica politica dei sistemi neoliberali :

- riflette la supremazia della razionalità strumentale;
- è desocializzata e de-storicizzata;
- è affetta dal "pregiudizio soluzionistico" che trasforma ogni questione sociale o politica in una questione tecnica, poiché incorpora l'idea che il modo di risolvere un problema sia trovare un obiettivo rispetto al quale ottimizzare la previsione".

Il risultato dell'uso di sistemi di ottimizzazione predittiva è *winners take all*.

Correlativamente, i più deboli di vengono schiacciati e puniti in modo automatizzato.

# I SISTEMI DI OTTIMIZZAZIONE PREDITTIVA HANNO PROPRIETÀ POLITICHE?



I sistemi di ottimizzazione predittiva

- **sono conservatori:** naturalizzano e replicano, automatizzandolo, lo status quo;
- **sono intrecciati con l'essenzialismo e il determinismo biologico;**
- **sono incompatibili con il liberalismo:** le persone non sono riconosciute come soggetti in grado di autodeterminarsi, sono trattate come cose;
- **sono antidemocratici:** ciò che è o non è una famiglia, un assembramento pericoloso o un migrante affidabile alla frontiera non è deciso dalle persone, ma determinato, invece, dai meccanismi del sistema, dal cui output probabilistico dipende chi sarà considerato tale e trattato di conseguenza.

# OTTIMIZZAZIONE PREDITTIVA E STATO DI DIRITTO

I sistemi di ottimizzazione predittiva sono incompatibili con lo stato di diritto, perché i cittadini

- **non sono più soggetti giuridici**, ma mutevoli fasci di dati (dati estratti e riassembleati in modi sempre diversi); porzioni di noi stessi (qualsiasi tipo di comportamento contingente che possa essere estratto) riassembleate seguendo criteri sempre diversi.
- **non sono uguali di fronte alla legge**: si è ascritti a un cluster in virtù di elementi non pertinenti alla fattispecie esaminata;
- **non possono prevedere quale legge sarà loro applicata**, perché i sistemi di ottimizzazione predittiva agiscono come leggi retroattive e segrete;
- **sono tutti sorvegliati** come «sospetti»;
- **soggiacciono alla logica totalitaria dell'«integrazione totale»**: tutti i dati personali sono considerati in qualsiasi tipo di previsione. Poiché i dati utilizzati per classificare gli individui si riferiscono a ogni sfera o aspetto della vita delle persone (religioso, professionale, sanitario, sessuale, economico, politico, demografico, legale), l'uso di sistemi di ottimizzazione predittiva fa crollare la differenziazione funzionale dei sistemi legali da quelli religiosi, etici ed economici.
- **sono sudditi di un *ancien régime* dinamico\***: possono o non possono esercitare i loro diritti in modi che dipendono non dal loro status di cittadini, ma dal loro contingente status economico, sociale, emotivo, di salute o religioso.



# VALUTARE LA BENEVOLENZA DEL DESPOTA?

## 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».

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

## Approccio basato sui rischi

Una mera valutazione d'impatto relativa ai rischi

- parte dal presupposto che il cambiamento tecnologico rifletta il corso inevitabile del progresso scientifico, anziché una scelta umana;
- riduce la politica della tecnologia a un esercizio di calcolo dei rischi.

Se un sistema è incompatibile con diritti giuridicamente tutelati, a cosa serve la valutazione d'impatto relativa ai rischi?

# A CHE SERVE L'ETICA DELL'INTELLIGENZA ARTIFICIALE?

Perché, di fronte all'evidenza delle discriminazioni, dei danni e delle ingiustizie generati dall'impiego, nei processi di decisione, dei sistemi di apprendimento automatico, le grandi aziende tecnologiche hanno risposto con l'«etica dell'intelligenza artificiale»?

E perché presentare l'etica, anziché il diritto, come l'ambito moralmente pertinente, considerato che ad essere violati sono diritti giuridicamente tutelati?

E, infine, se i sistemi di ottimizzazione predittiva semplicemente non funzionano, che senso ha invocare l'etica?

Che cosa penseremmo se un tecnico, chiamato ad aggiustare un lavatrice difettosa, che ci allaga infallibilmente la casa ad ogni lavaggio, tentasse di convincerci che la lavatrice funziona benissimo e che occorre soltanto capire come programmarla all'«equità», onde ottenere «un allineamento dei valori»? I sistemi di apprendimento automatico non sono più intelligenti di una vecchia lavatrice e, come quella, sono artefatti, prodotti del lavoro umano.

A che serve, dunque, l'etica dell'intelligenza artificiale?

# 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.

# ETICA DELL'INTELLIGENZA ARTIFICIALE

## LE CONDIZIONI NON SODDISFATTE

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.

Anzitutto, però, 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 un'intelligenza artificiale generale (AGI).

Attualmente nessuno ha un'idea realistica non solo di come realizzarla, ma neppure di quali siano le domande di ricerca.

# 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 imperscrutabile 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.

# 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](https://archive.org/details/computerpowerhum0000weiz_v0i3)

# 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”

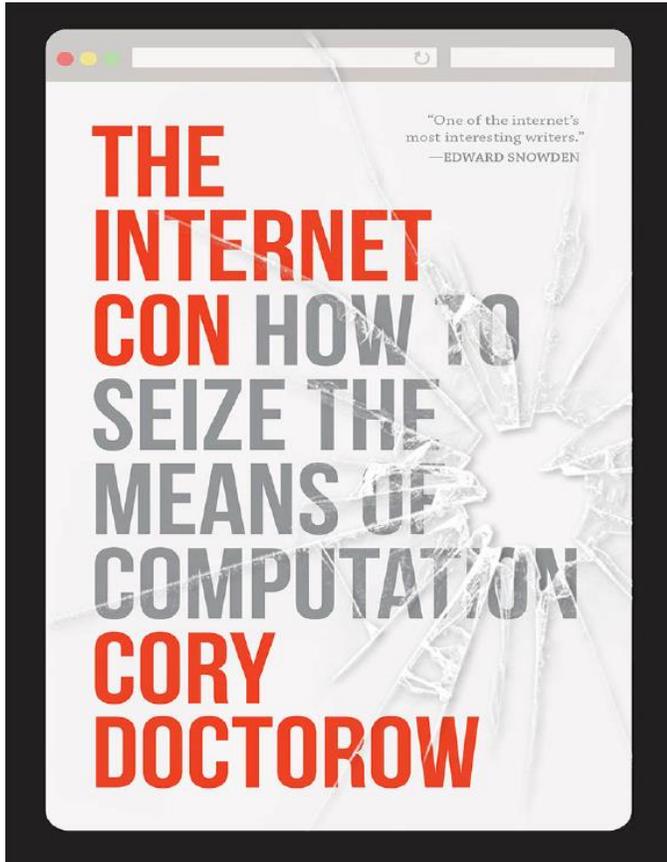
Lina M. Khan

«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 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 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 

Grazie per l'attenzione.  
Domande?

[daniela.tafani@unipi.it](mailto:daniela.tafani@unipi.it)

