

# **Intelligenza Artificiale**

## **Verità e bugie**

Andreas R. Formiconi

# La doppia narrativa

“Business”

VS

Ricerca

# Un passo indietro nel futuro...

Intelligenza artificiale simbolica

VS

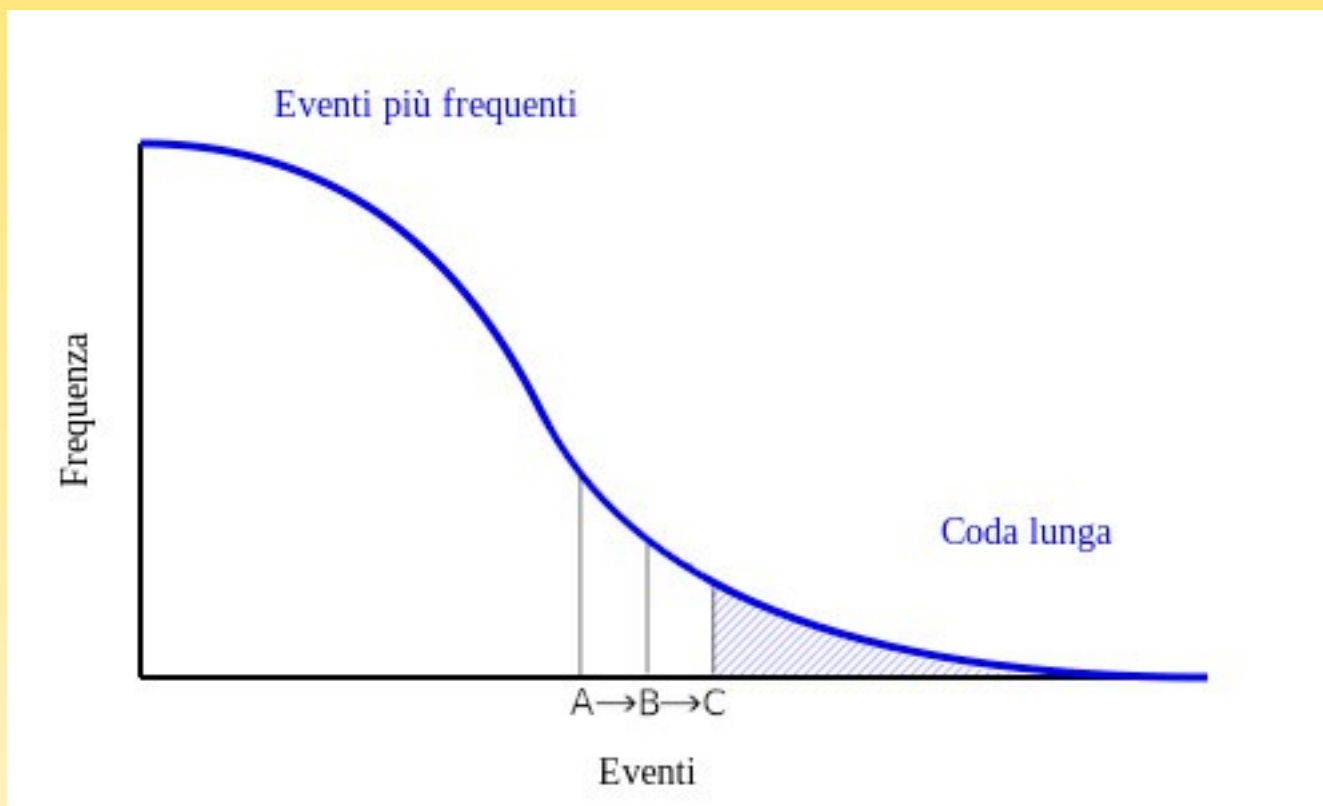
subsimbolica

# Intelligenza Artificiale

- Predittiva
- Generativa

# Intelligenza Artificiale Predittiva

## Problema delle code lunghe



# Intelligenza Artificiale Predittiva

“For in-hospital mortality prediction, the models tested using our synthesized cases fail to recognize 66% of the injuries.”

*Pias et al (2025) Low responsiveness of machine learning models to critical or deteriorating health conditions, Nature*

“Il problema fondamentale è che lo spazio dei possibili campioni di classi minoritarie è vastissimo. Tentare di coprirli tutti o quasi tutti con i dati di addestramento è irrealizzabile. Pertanto, l'ingegneria dei dati non sembra essere una direzione praticabile. Un approccio più promettente consiste nel codificare direttamente la semantica medica relativa alle decisioni cliniche.”

**Che significa recupero del paradigma simbolico...**

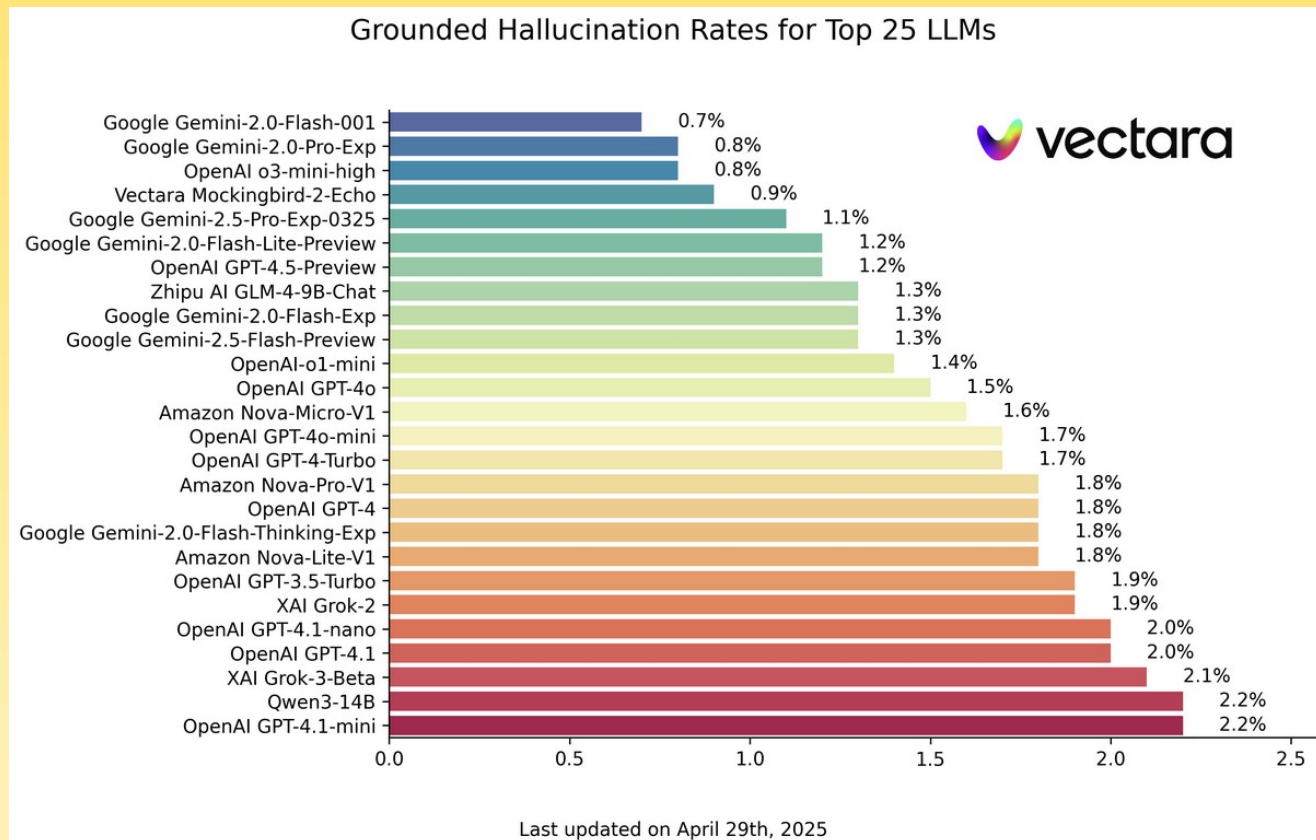
# Intelligenza Artificiale Generativa

## Problemi...

1. Allucinazioni
2. Distorsione della varietà di risposte possibili su questioni non dicotomizzabili
3. Collasso ragionamento su problemi “semplici” — Alice, benchmark ARC-AGI...
4. Prevalenza lingue, culture, economie dominanti (inglese > 90 % )
5. Bias ineliminabili (genere, etnia, paesi...)
6. “AI models collapse when trained on recursively generated data”
7. “AI and the problem of knowledge collapse”
8. Controllo da parte dei poteri — aziende e governi

# Allucinazioni

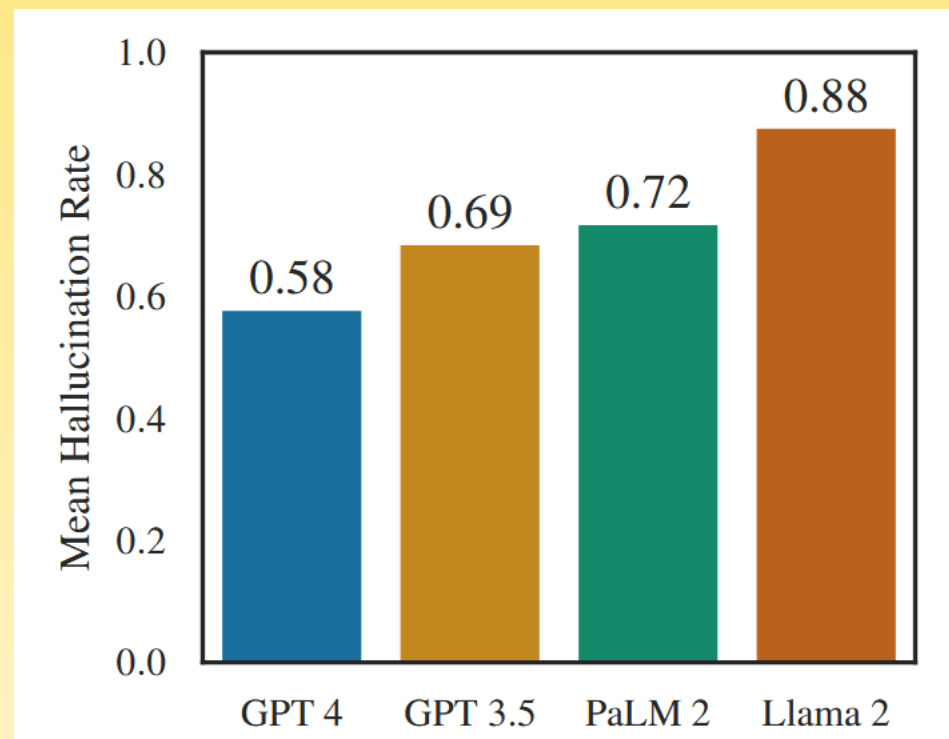
<https://github.com/vectara/hallucination-leaderboard/>



# Allucinazioni

Dahl et al (2024) Large Legal Fictions: profiling legal hallucinations in Large Language Models

<https://arxiv.org/pdf/2401.01301>

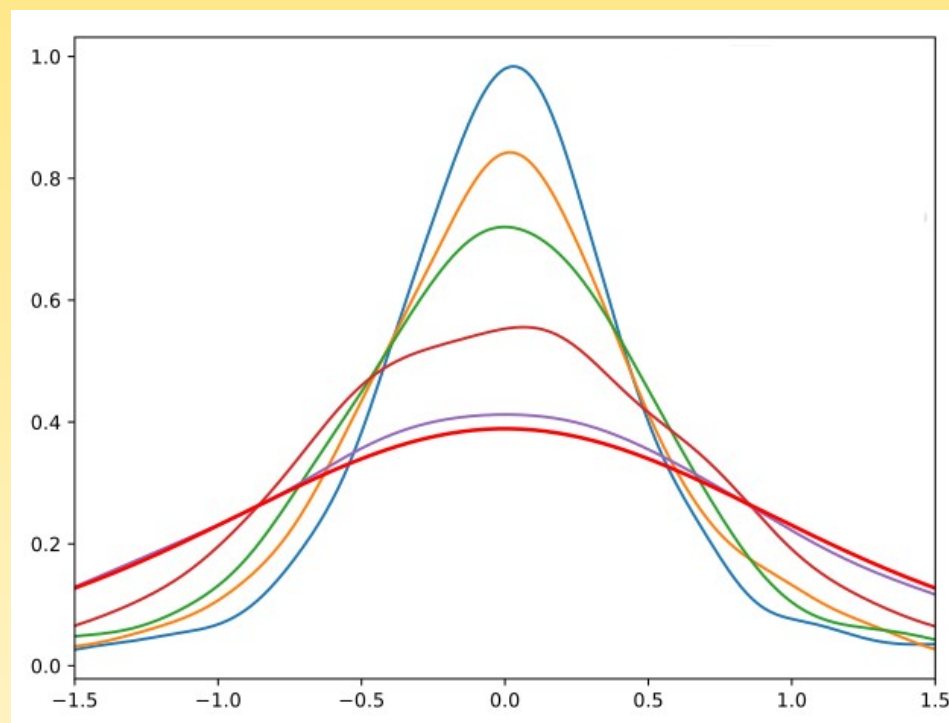


# Distorsione della distribuzione di risposte non dicotomiche

AI and the problem of knowledge collapse

Peterson (2024) Springer

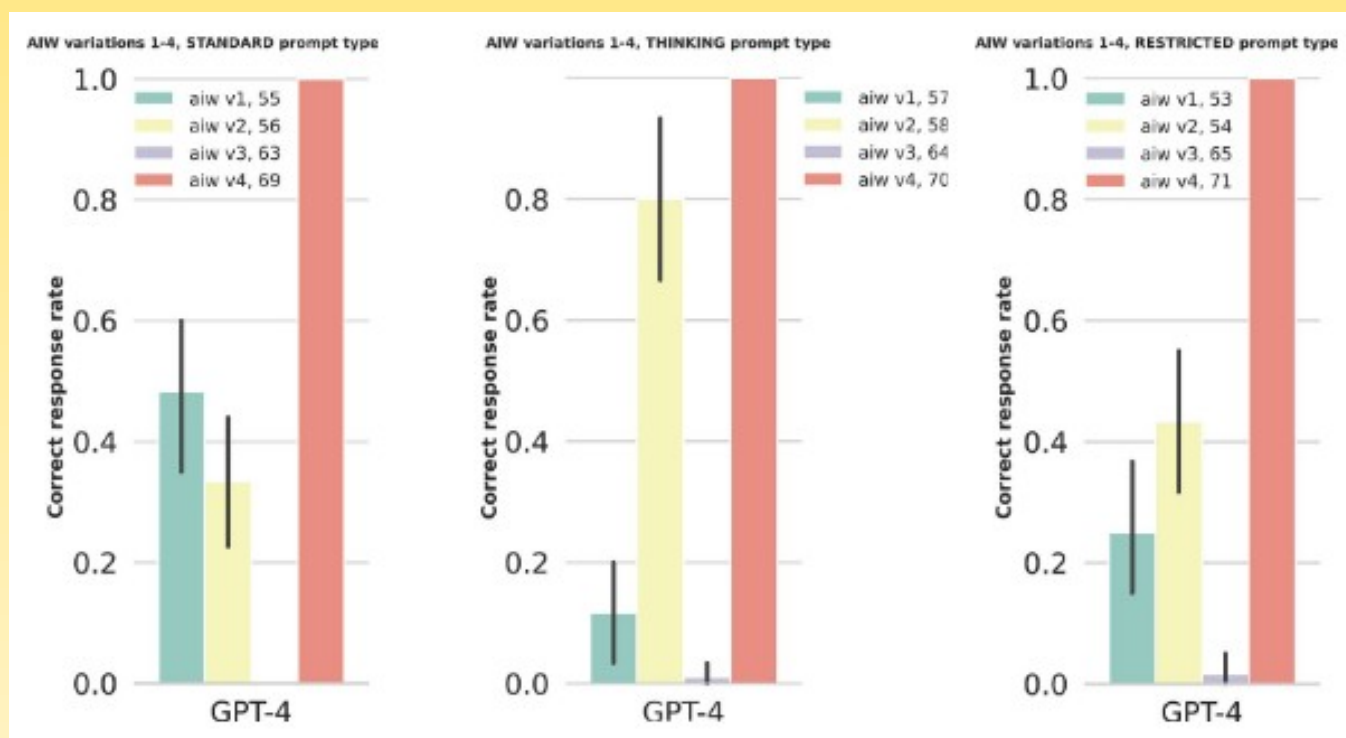
<https://link.springer.com/article/10.1007/s00146-024-02173-x>



# Collasso del ragionamento su questioni semplici

Alice in Wonderland: Simple Tasks Showing Complete Reasoning Breakdown in State-Of-the-Art Large Language Models

<https://arxiv.org/pdf/2406.02061>

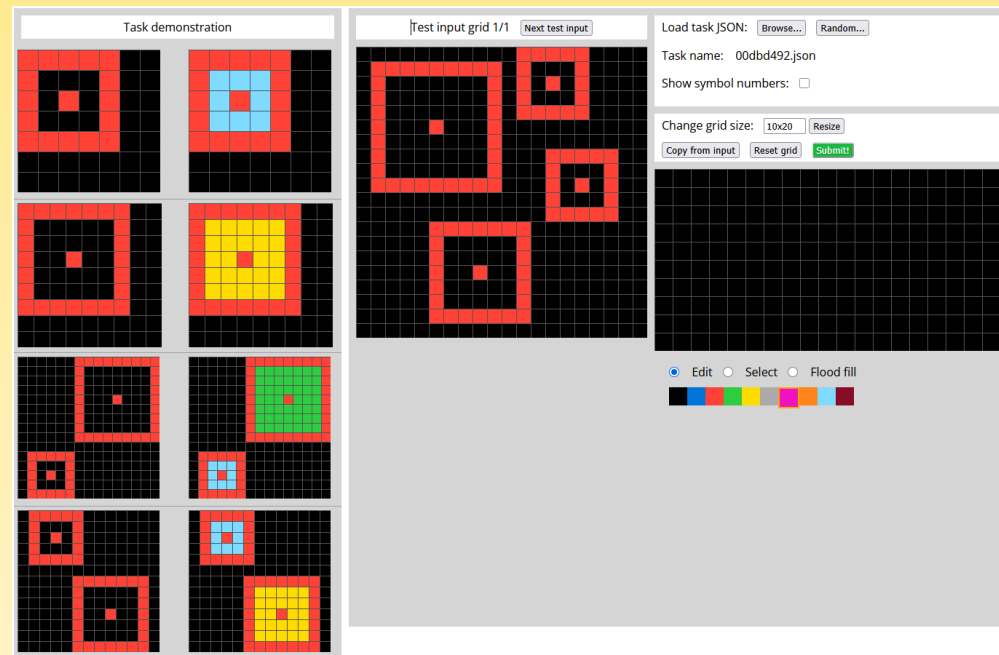


# Collasso del ragionamento su questioni semplici

Abstract and Reasoning Corpus for Artificial General Intelligence" (ARC-AGI) benchmark: <https://arcprize.org/arc-agi>

François Chollet (2019) On the Measure of Intelligence

<https://arxiv.org/abs/1911.01547>



# Prevalenza lingue e culture dominanti nei corpora di addestramento (>90% inglese)

Xu et al (2024) A Survey on Multilingual Large language Models: Corpora, Alignment, Bias

<https://arxiv.org/abs/2404.00929>

Model	Languages	Language proportion	Source
XLM-R	100	English (12.56%); Russian (11.61%); Indonesian (6.19%); Vietnamese (5.73%); Others (63.89%)	Generated using the open source; CC-Net repository
BLOOM	46	English (30.03%); Simplified Chinese (16.16%); French (12.9%); Spanish (10.85%); Portuguese (4.91%); Arabic (4.6%); Others (20.55%)	Web Crawl (38%); BigScience Catalogue Data (62%)
Falcon	>100	Inglese (58.2%); Russian (13.19%); German (10.81%); Spanish (9.45%); Others (66.55%)	Common Crawl
GPT-3	95	English (92.7%); French (1.8%); German (1.5%); Others (5.9%)	Common Crawl, Wikipedia, Books1, Books2, WebText2
GPT-4	26	Unknown	Common Crawl, Wikipedia, Books1, Books2, WebText2
LLaMa 3	176	Over 5% non-English	Publicly available sources excludes Meta user data
Claude 3	43	Unknown	Publicly available information on the Internet; Non-public data.

# Bias (genere, etnie e culture dominanti)

Xu et al (2024) A Survey on Multilingual Large language Models: Corpora, Alignment, Bias  
<https://arxiv.org/abs/2404.00929>

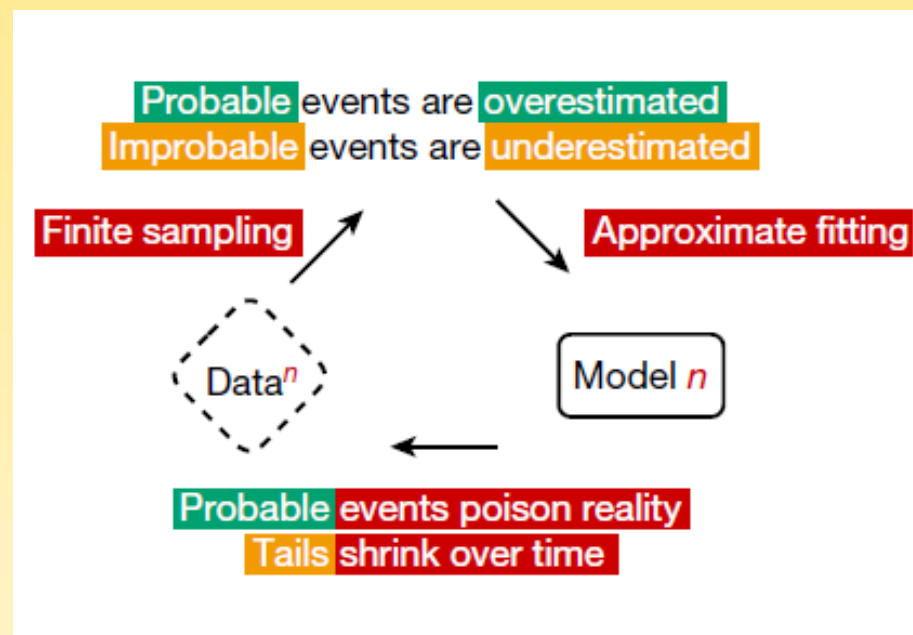
Model	Languages	Language proportion	Source
XLM-R	100	English (12.56%); Russian (11.61%); Indonesian (6.19%); Vietnamese (5.73%); Others (63.89%)	Generated using the open source; CC-Net repository
BLOOM	46	English (30.03%); Simplified Chinese (16.16%); French (12.9%); Spanish (10.85%); Portuguese (4.91%); Arabic (4.6%); Others (20.55%)	Web Crawl (38%); BigScience Catalogue Data (62%)
Falcon	>100	Inglese (58.2%); Russian (13.19%); German (10.81%); Spanish (9.45%); Others (66.55%)	Common Crawl
GPT-3	95	English (92.7%); French (1.8%); German (1.5%); Others (5.9%)	Common Crawl, Wikipedia, Books1, Books2, WebText2
GPT-4	26	Unknown	Common Crawl, Wikipedia, Books1, Books2, WebText2
LLaMa 3	176	Over 5% non-English	Publicly available sources excludes Meta user data
Claude 3	43	Unknown	Publicly available information on the Internet; Non-public data.

# Degenerazione dell'“intelligenza” dell'IA quando questa si nutre dei suoi stessi prodotti

AI models collapse when trained on recursively generated data

Shumailov et al (2024) Nature

<https://www.nature.com/articles/s41586-024-07566-y>

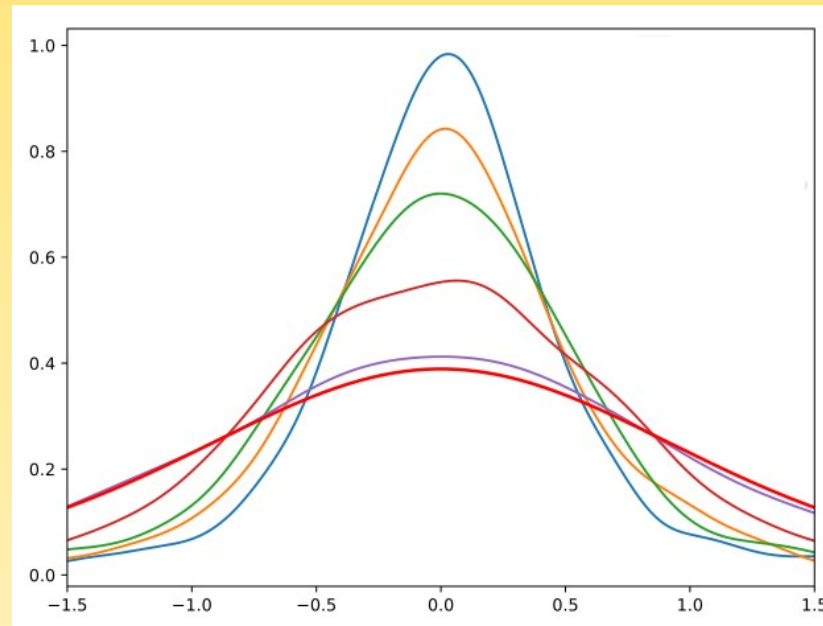


# Problemi — Collasso dell'intelligenza collettiva

AI and the problem of knowledge collapse

Peterson (2024) Springer

<https://link.springer.com/article/10.1007/s00146-024-02173-x>



## Problemi — Controllo e manipolazione

Under Trump, AI Scientists Are Told to Remove 'Ideological Bias'  
From Powerful Models (2025) Wired

<https://www.wired.com/story/ai-safety-institute-new-directive-america-first/>



# Key takeaways...

## L'intelligenza artificiale...

- è scientific computing, non altro:  
algebra lineare + statistica, non magia
- predittiva: soffre le code lunghe → non solo dati!
- generativa: non affinità semantica ma prossimità statistica...
- ...genera plausibilità, non verità fattuale  
(anche nei modelli “reasoning”)
- controllata da pochi, pochissimi...

# E nel mondo del business (quello vero)?

“Companies Are Struggling to Drive a Return on AI”

***The Wall Street Journal, April 26, 2025***

<https://www.wsj.com/articles/companies-are-struggling-to-drive-a-return-on-ai-it-doesnt-have-to-be-that-way-f3d697aa?st=bfHnJ4&reflink>

*“Secondo il sondaggio State of AI condotto da McKinsey pubblicato a marzo, lo scorso anno il 78% delle aziende ha dichiarato di utilizzare l'intelligenza artificiale in almeno una funzione, rispetto al 55% del 2023. A fronte di questo impegno, le aziende hanno affermato di ottenere in genere risparmi sui costi inferiori al 10% e aumenti dei ricavi inferiori al 5%”*

“Welcome to the AI trough of disillusionment”

***The Economist, May 25, 2025***

<https://www.economist.com/business/2025/05/21/welcome-to-the-ai-trough-of-disillusionment>

*“Per molte aziende, l'entusiasmo per le promesse dell'intelligenza artificiale generativa (AI) ha lasciato il posto alla frustrazione per la difficoltà di utilizzare questa tecnologia in modo produttivo. Secondo S&P Global la percentuale di aziende che abbandonano la maggior parte dei loro progetti pilota di IA generativa è salita al 42%, rispetto al 17% dello scorso anno. Il capo di Klarna, un fornitore svedese di servizi “compra ora, paga dopo”, ha recentemente ammesso di aver esagerato nell'uso della tecnologia per tagliare i posti di lavoro nel servizio clienti e ora sta riassumendo personale umano per ricoprire tali ruoli.”*

# Microsoft-Backed Builder.ai Set for Bankruptcy After Cash Seized

*Bloomberg, May 20, 2025*

<https://www.bloomberg.com/news/articles/2025-05-20/microsoft-backed-builder-ai-to-enter-insolvency-proceedings?sref=YfHlo0rL>

*“Viola Credit, che ha finanziato 50 milioni di dollari di debito alla società di software lo scorso anno, ha sequestrato 37 milioni di dollari dai conti di Builder.ai, lasciando alla società 5 milioni di dollari.”*

*Tech Funding News, June 3,2025*

<https://techfundingnews.com/fake-it-till-you-unicorn-builder-ais-natasha-was-never-ai-just-700-indian-coders-behind-the-curtain/>

*“Builder.ai’s Natasha was never AI – just 700 Indian coders behind the curtain”*

**So what?**

**AI ok**

**but no BS, please!**