

AGENDA

- Comprendere
- Competere
- Convivere

By any reasonable criteria, the vision of **human-level machine intelligence** laid out by Alan Turing in 1950 **is now a reality**.

A difference in **efficiency of learning** does not necessarily mean a **different level of intelligence**.

For the first time in human history, **we are no longer alone** in the space of general intelligence. Eyes unclouded by dread or hype **will help to prepare for what comes next**.

Chen E. K., Belkin M., Bergen L., Danks D., Does AI already have human-level intelligence? The evidence is clear, **Nature**, Feb. 2026, Vol 650, 36-40.

Comment

Stockholm conventions in 2019 following a 15-year debate. Obviously, implementing this qualified majority voting would be politically challenging and feasible only if voting was used sparingly and transparently, and without overlooking or dismissing the concerns of countries with less power.

Setze the moment

The choices made now will determine whether the health of people and the planet are safeguarded or put at further risk. We urge the INC's newly elected chair to consider implementing these reforms.

Multiple events during the past few years have undermined multilateralism. And as geopolitical priorities shift, environmental concerns are increasingly being sidelined or environmental policies weakened¹.

Against this backdrop, it is crucial that the INC process succeeds – both to address a major contributor to the interconnected planetary crises (climate change, biodiversity loss and pollution) and to restore faith in the idea that international cooperation can solve global challenges.

The authors

Paul Einhaupl is a research associate of the Systemic Risks Group at the Research Institute for Sustainability (RIFS) at GFZ Helmholtz Centre for Geosciences, Potsdam, Germany.

Linda Del Savio is a research associate of the Ocean Governance Group at the Research Institute for Sustainability (RIFS) at GFZ Helmholtz Centre for Geosciences, Potsdam, Germany.

Melanie Bergmann is a senior scientist at the Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany.

Annika Jahnel is co-head of the Department of Exposure Science at the Helmholtz Centre for Environmental Research – UFZ, Leipzig, and professor of exposure analytics at RWTH Aachen University, Aachen, Germany.

e-mail: paul.einhaupl@rifs-potsdam.de

1. Bergmann, M. et al. *Nature Rev. Earth Environ.* **3**, 303–307 (2022).

2. Tolman, M. B., Walther, D. A., Peters, C., Cutler, L. & Bergmann, M. *Impacts of Plastic Pollution in the Oceans on Marine Species, Ecosystems and Communities* (2020).

3. Kewill, P. *One Earth* **5**, 506–507 (2022).

4. Kewill, P., Khanna, M. & Nijhar, S. *Climate Impact of Primary Plastic Production* (Lawrence Berkeley National Laboratory) (2022).

5. Dougenkos, P. et al. *Mar. Policy* **118**, 104803 (2020).

6. Mura, T., Moriwaki, N., Watanabe, I. & Otsuka, T. *Chem. Pharm. Plant.* **4**, 422 (2018).

7. MacLeod, M., Aze, H., H. H., Tolman, M. B. & Jahnel, A. *Science* **375**, 51–54 (2021).

8. UNEP. *UNEP Global Plastic Pollution Towards an International Legally Binding Instrument: Resolution Adopted by the United Nations Environment Assembly* (UNEP, 2022).

9. Watanabe, I. et al. *Nature* **543**, 149–150 (2017).

10. Lutz, H. & Dun, T. *Sci. Rep.* **14**, 10747 (2024).

The authors declare no competing interests.

Does AI already have human-level intelligence? The evidence is clear

Eddy Keming Chen, Mikhail Belkin, Leon Bergen & David Danks

By any reasonable criteria, the vision of human-level machine intelligence laid out by Alan Turing in 1950 is now a reality. Eyes unclouded by dread or hype will help to prepare for what comes next.

Intelligence (AGI) – and some doubt that they ever will. A March 2025 survey by the Association for the Advancement of Artificial Intelligence in Washington DC found that 76% of leading researchers thought that scaling up current AI approaches would be 'unlikely' or 'very unlikely' to yield AGI (see go.nature.com/4am16b).

What explains this disconnect? We suggest that the problem is part conceptual, because definitions of AGI are ambiguous and inconsistent: part emotional, because AGI raises fear of displacement and disruption; and part practical, as the term is entangled with commercial interests that can distort assessments. Precisely because AGI dominates public discourse, it is worth engaging with the concept in a more detached way: as a question about intelligence, rather than a pressing concern about social upheaval or an ever-postponed milestone in a business contract.

In writing this Comment, we approached this question from different perspectives – philosophy, machine learning, linguistics and cognitive science – and reached a consensus after extensive discussion. In what follows, we set out why we think that, once you clear away certain confusions, and strive to make fair comparisons and avoid anthropocentric biases, the conclusion is straightforward by reasonable standards, including Turing's own: we have artificial systems that are generally intelligent. The long-standing problem of creating AGI has been solved. Recognizing this fact matters – for policy, for risk and for understanding the nature of mind and even the world itself.

Questions of definition

We assume, as we think Turing would have done, that humans have general intelligence. Some think that general intelligence does not exist at all, even in humans. Although this view is coherent and philosophically interesting, we set it aside here as being too disconnected from most AI discourse. But having made this assumption, how should we characterize general intelligence?

A common informal definition of general intelligence, and the starting point of our discussions, is a system that can do almost all cognitive tasks that a human can do².

In 1950, in a paper entitled 'Computing Machinery and Intelligence', Alan Turing proposed his 'imitation game'. Now known as the Turing test, it addressed a question that seemed purely hypothetical: could machines display the kind of flexible, general cognitive competence that is characteristic of human thought, such that they could pass themselves off as humans to unaware humans? Three-quarters of a century later, the answer looks like yes: In March 2025, the large language model (LLM) GPT-4.5, developed by OpenAI in San Francisco, California, was judged by humans in a Turing test to be human 73% of the time – more often than actual humans were³. Moreover, readers even preferred literary texts generated by LLMs over those written by human experts⁴.

This is far from all. LLMs have achieved gold-medal performance at the International Mathematical Olympiad, collaborated with leading mathematicians to prove theorems⁵, generated scientific hypotheses that have been validated in experiments⁶, solved problems from PhD exams, assisted professional programmers in writing code, composed poetry and much more – including chatting 24/7 with hundreds of millions of people around the world. In other words, LLMs have shown many signs of the sort of broad, flexible cognitive competence that was Turing's focus – what we now call general intelligence⁷, although Turing did not use the term.

Yet many experts balk at saying that current AI models display artificial general

Example of GPT-4 visual input:

User What is funny about this image? Describe it panel by panel.



Source: <https://www.reddit.com/r/hmmm/comments/ubab5v/hmmm/>

GPT-4 The humor in this image comes from the absurdity of plugging a large, outdated VGA connector into a small, modern smartphone charging port.

2 - AI PERFORMANCES (AS OF MAY-2025)

In May 2025, thirty top mathematicians gathered secretly in Berkeley, California, to challenge an advanced AI reasoning chatbot, o4-mini, aiming to test and potentially outsmart its mathematical reasoning capabilities.

Why secret? To prevent their test problems from leaking online and contaminating the AI's training data.

At the center of the challenge stood o4-mini, a large language model developed by OpenAI. Unlike earlier LLMs, o4-mini employs specialized datasets and stronger human reinforcement, making it nimbler and more adept at tackling complex mathematical reasoning. By April 2025, o4-mini had cracked about 20% of all questions posed—a huge leap from typical chatbot performance...To ramp up progress, Epoch AI orchestrated the Berkeley summit, where participants split into small groups to invent questions that would expose weaknesses in o4-mini.

The meeting showcased how quickly AI has advanced. **Researchers compared o4-mini's capabilities to those of a strong human collaborator or a top graduate student.** The AI performed calculations and deductions in minutes, whereas human mathematicians might spend weeks or months on the same tasks. Yang Hui He, a pioneer in AI-assisted math research, compared o4-mini to “**a very, very good graduate student,**” adding that the AI was often faster than human experts, tackling weeks-long problems in mere minutes.

Despite the excitement, researchers voiced concerns. There is a risk that mathematicians and the broader community might over-trust AI outputs, leading to “**proof by intimidation**” where AI's authority discourages scrutiny.

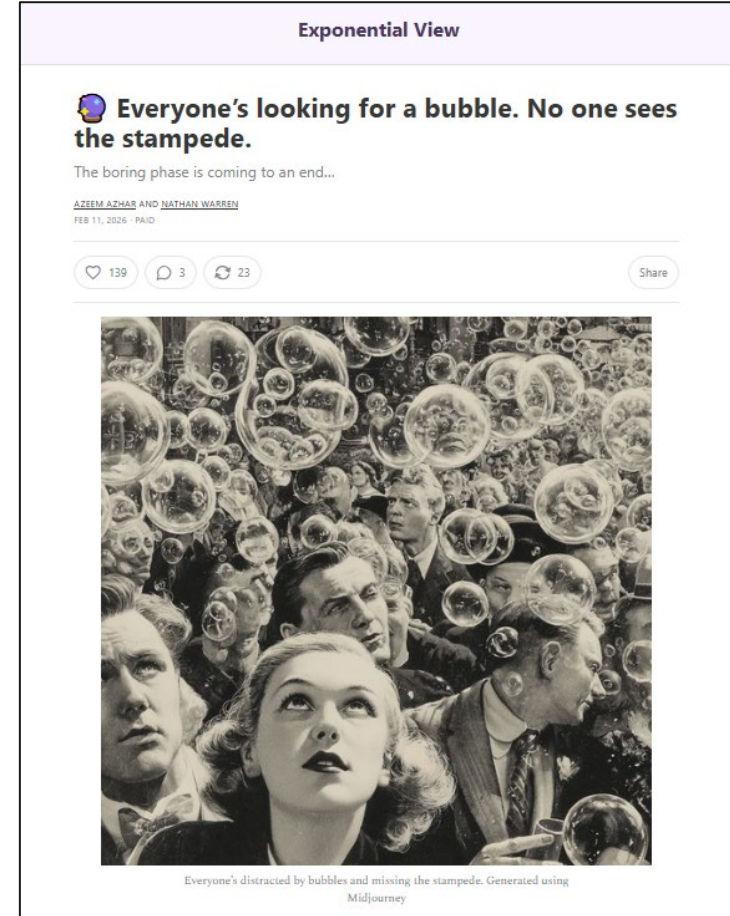
Looking ahead, mathematicians may transition toward generating novel questions and collaborating more interactively with reasoning bots, akin to supervising graduate students. This shift places greater emphasis on fostering creativity and critical thinking within mathematical education to maintain human relevance.



And **something changed right at the end of 2025**, and we know what it was. **Models passed a threshold of coherence**; they can work very reliably on tasks of an hour or two, and somewhat less reliably on longer tasks. **Something clicked.**

If you're not at least knee deep in long-running workflows, it's quite hard to understand the implications... at Anthropic, an engineer used it to build a C++ compiler for \$20,000 in API costs. A comparable project built by humans would typically require five to ten engineers over 18-24 months, around \$2-3 million in fully loaded labor costs.

The implications extend well beyond software. These systems can tackle any task you give them sufficient context for, including research, writing, data analysis, and operational workflows. We're putting progressively more complex flows through agents, **working in some cases around the clock.**



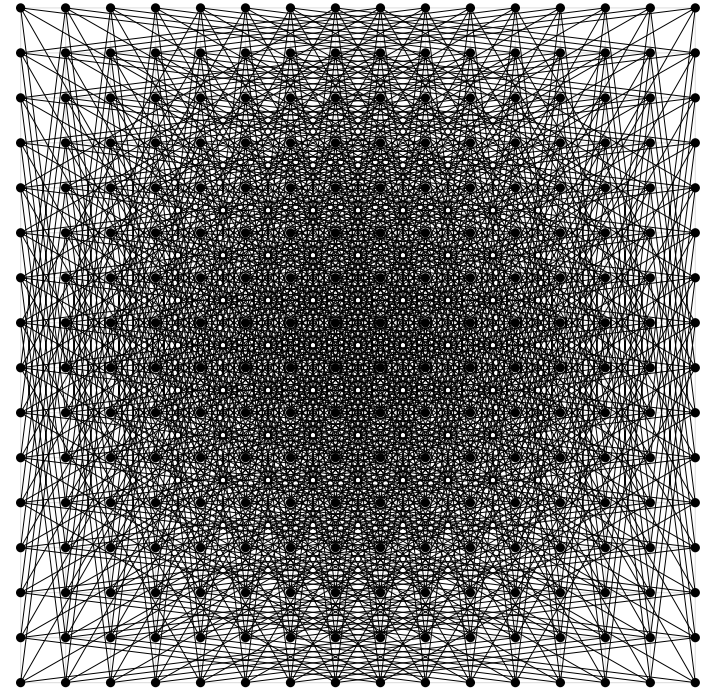
A general-purpose, **internal OpenAI reasoning model** disproved the 80-year-old Erdős **Unit Distance Conjecture**.

For nearly 80 years, experts thought simple square grid patterns were basically the best way to arrange dots to get as many **unit-distance pairs as possible**. Erdős guessed the **maximum number of pairs grows almost linearly n** .

OpenAI proved instead that **you can squeeze more unit-distance pairs than previously thought**.

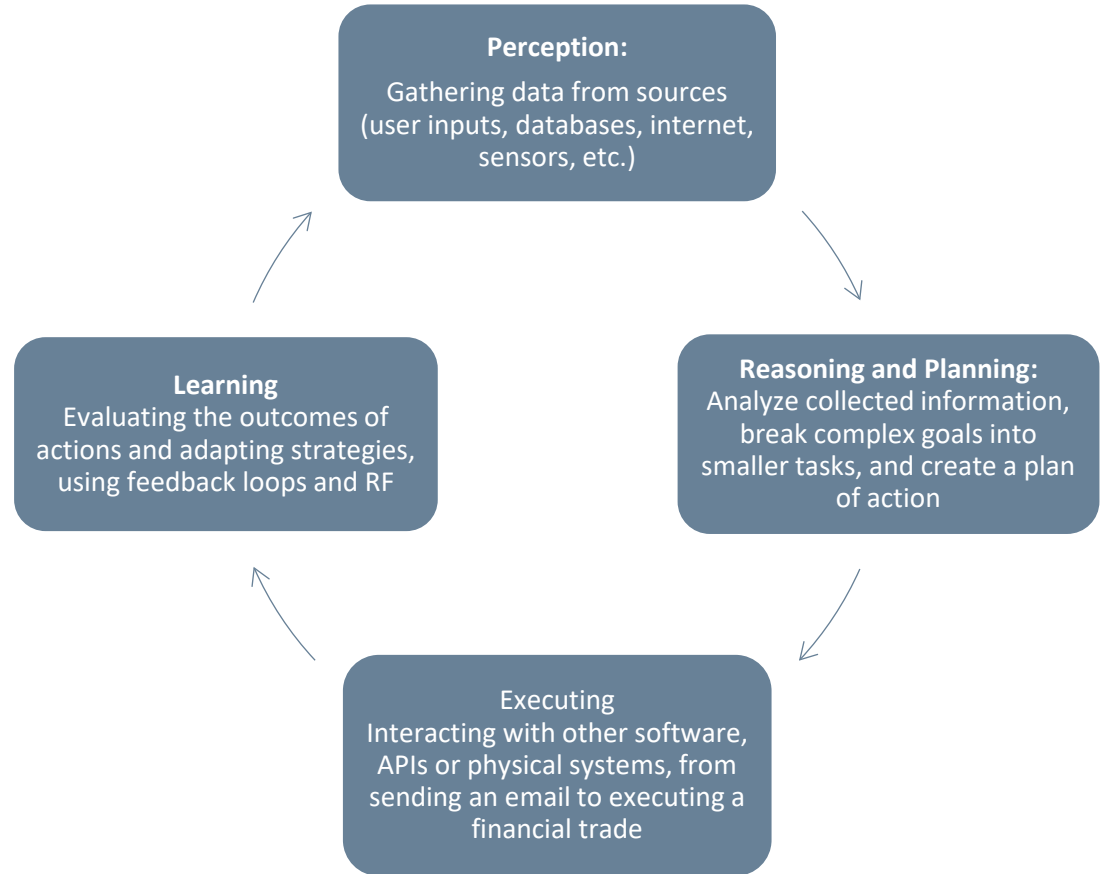
Why it matters:

- A prominent open math problem **tackled by a general reasoning AI** rather than a specialized math solver;
- **Three-way superhuman**: original approach (disproval), multi-domain approach (topology, algebraic number fields), persistency (10+ pages just for the conceptual framework)

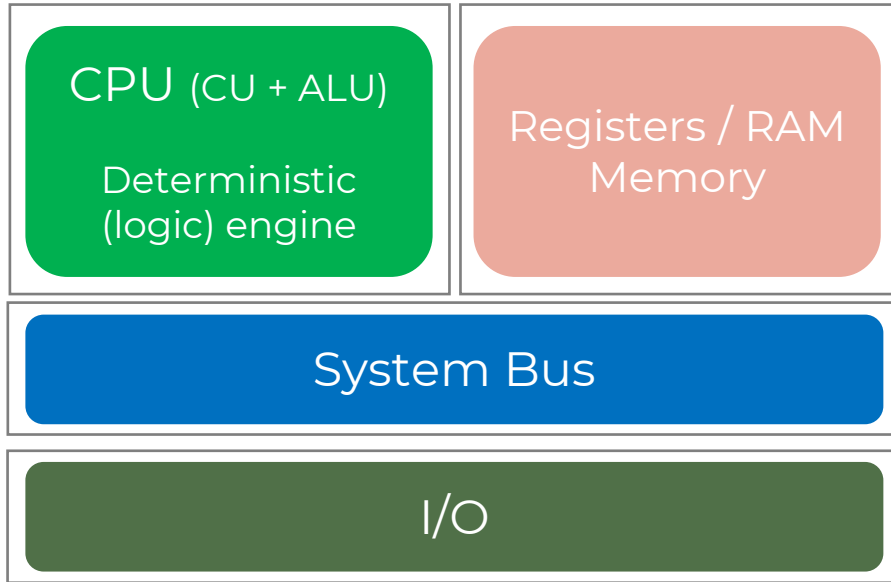


3 - AGENTIC AI

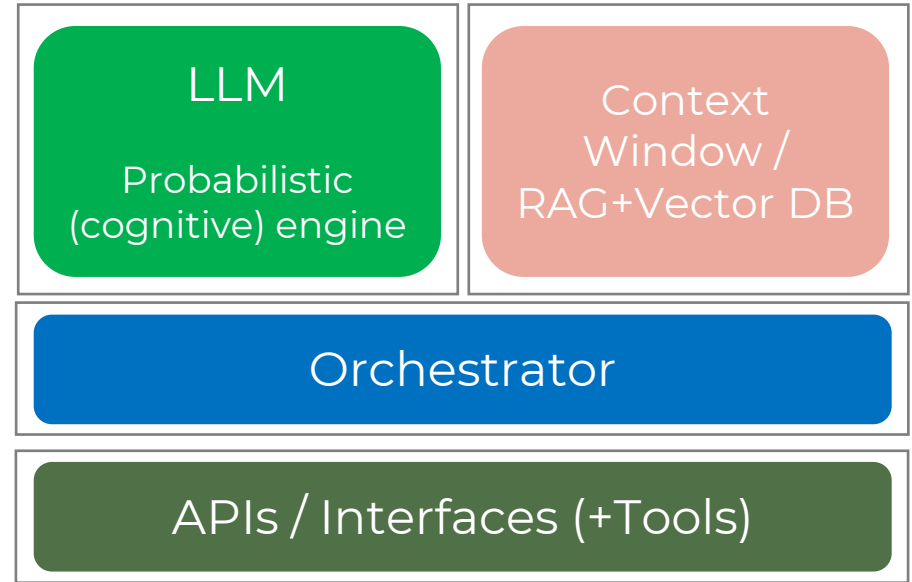
At the core of modern Agentic-AI remain **Large Language Models (LLMs)** acting as the "brain" of the agent, enabling advanced reasoning, comprehension and communication.



3 - WHY AGENTIC AI IS SO IMPORTANT



1945: Von Neumann's System Architecture



2023: Agentic System Architecture

3 - FROM AGENTIC AI TO AI AGENTS

- *Perform actions:*
 - *Access to other (IT) systems*
 - *Self-write code*
- *Interact with other agents*

Reality:
AI Agents are
somewhere

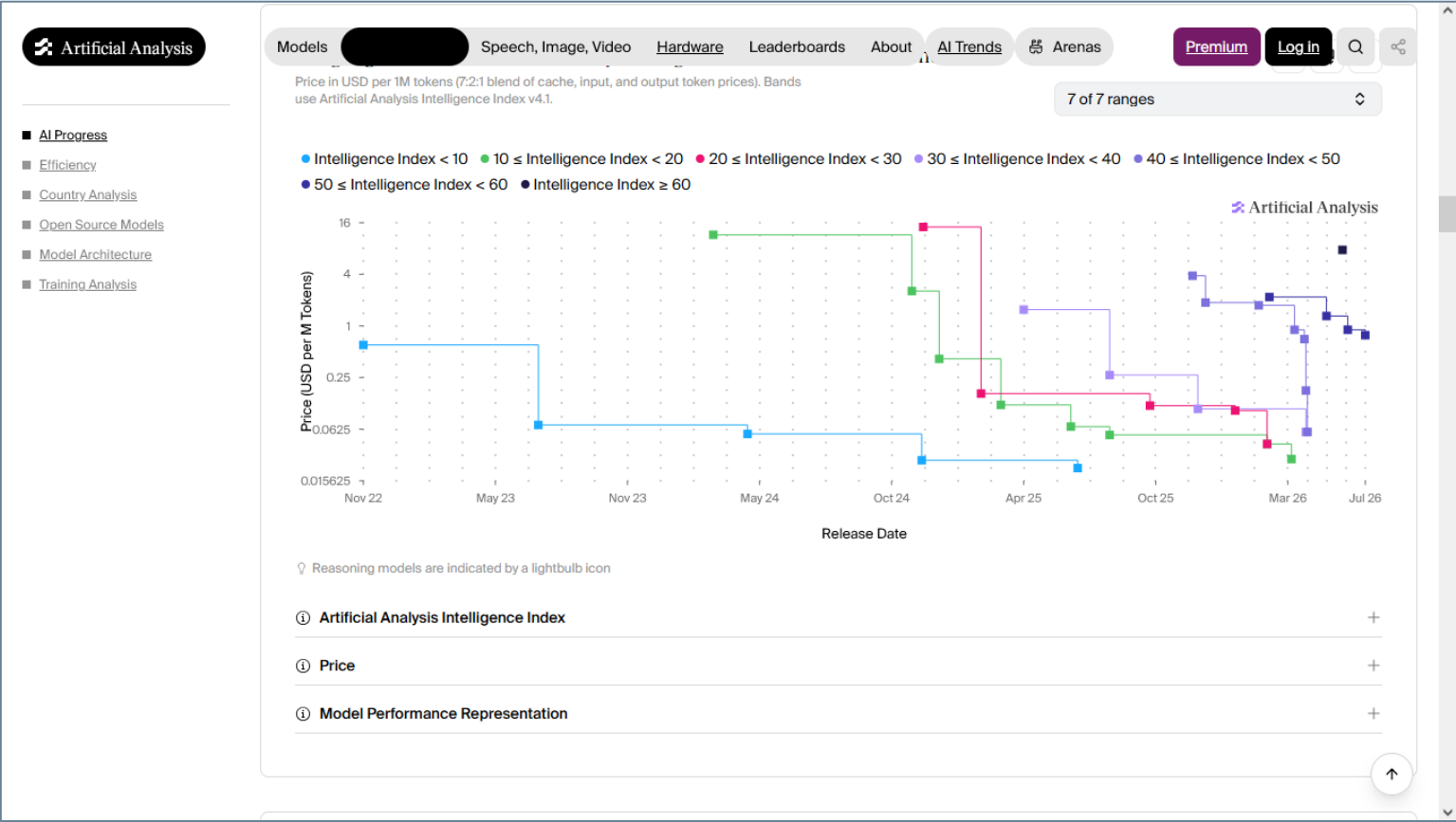


Claim:
AI Agents are
everywhere

"Whether the final result is rigorous or not depends on the epistemic quality of your starting premises — not on any inherent guarantee of the tool."

Claude Sonnet 4.6 Thinking

5 – MODELS ARE COMMODITIES? NO, THEY ARE NOT, BUT SOON WILL BE... (WWW.ARTIFICIALANALYSIS.AI)



AGENDA

- Comprendere
- Competere
- Convivere

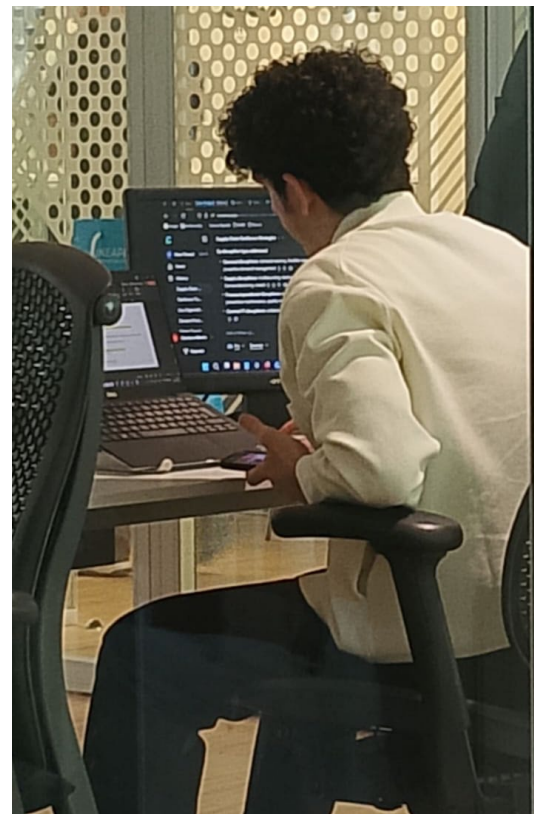
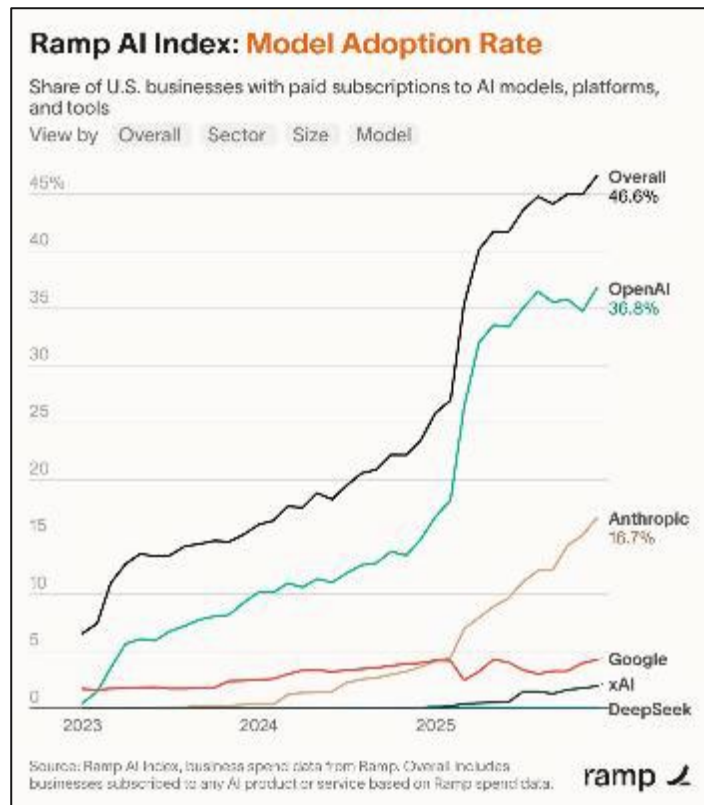
Artificial Intelligence

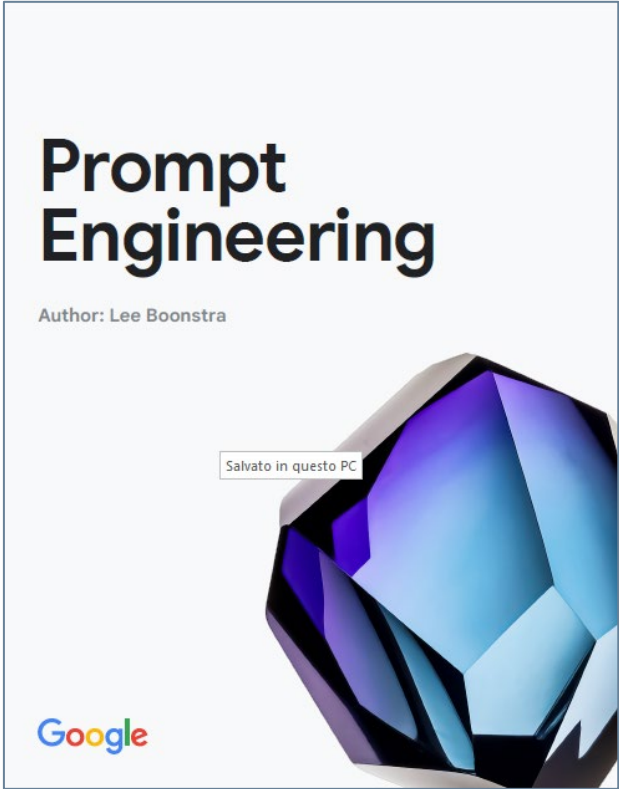
Personal Productivity
(Augmentation)

Personal Assistant(s)
Personal Agent(s)
Personal-Use Coding

Corporate level

AI VALUE: PERSONAL PRODUCTIVITY: EVERYBODY IS ALREADY USING IT





<https://frontbackgeek.com/googles-viral-prompt-engineering-whitepaper-a-game-changer-for-ai-users/>

arXiv:2211.01910v2 [cs.LG] 10 Mar 2023

Published as a conference paper at ICLR 2023

LARGE LANGUAGE MODELS ARE HUMAN-LEVEL PROMPT ENGINEERS

Yongchao Zhou^{1,2,*}, Andrei Ioan Muresanu^{2,3,*}, Zhen Han^{1,2,*}, Keiran Paster^{1,2},
Sihfu Pitt^{1,2}, Harris Chan^{1,2}, Jimmy Ba^{1,2}
¹University of Toronto ²Vector Institute ³University of Waterloo *Equal contribution
{yzhou, hanzhen, keirp, spitis, hchan, jba}@cs.toronto.edu
{andrei.muresanu}@uwaterloo.ca

ABSTRACT

By conditioning on natural language instructions, large language models (LLMs) have displayed impressive capabilities as general-purpose computers. However, task performance depends significantly on the quality of the prompt used to steer the model, and most effective prompts have been handcrafted by humans. Inspired by classical program synthesis and the human approach to prompt engineering, we propose *Automatic Prompt Engineering* (APE) for automatic instruction generation and selection. In our method, we treat the instruction as the “program,” optimized by searching over a pool of instruction candidates proposed by an LLM in order to maximize a chosen score function. To evaluate the quality of the selected instruction, we evaluate the zero-shot performance of another LLM following the selected instruction. Extensive experiments show that our automatically generated instructions outperform the prior LLM baseline by a large margin and achieve better or comparable performance to the instructions generated by human annotators on 2424 Instruction Induction tasks and 1721 curated BIG-Bench tasks. We conduct extensive qualitative and quantitative analyses to explore the performance of APE. We show that APE-engineered prompts are able to improve few-shot learning performance (by simply prepending them to standard in-context learning prompts), find better zero-shot chain-of-thought prompts, as well as steer models toward truthfulness and/or informativeness.

1 INTRODUCTION

The combination of scale and attention-based architectures has resulted in language models possessing an unprecedented level of generality (Kaplan et al. 2020; Vaswani et al. 2017). These so-called “large language models” (LLMs) have shown remarkable, often superhuman, capabilities across a diverse range of tasks, including both zero-shot and few-shot setups (Brown et al. 2020; Srivastava et al. 2022). With generality, however, there comes a question of control: how can we make LLMs do what we want them to do?

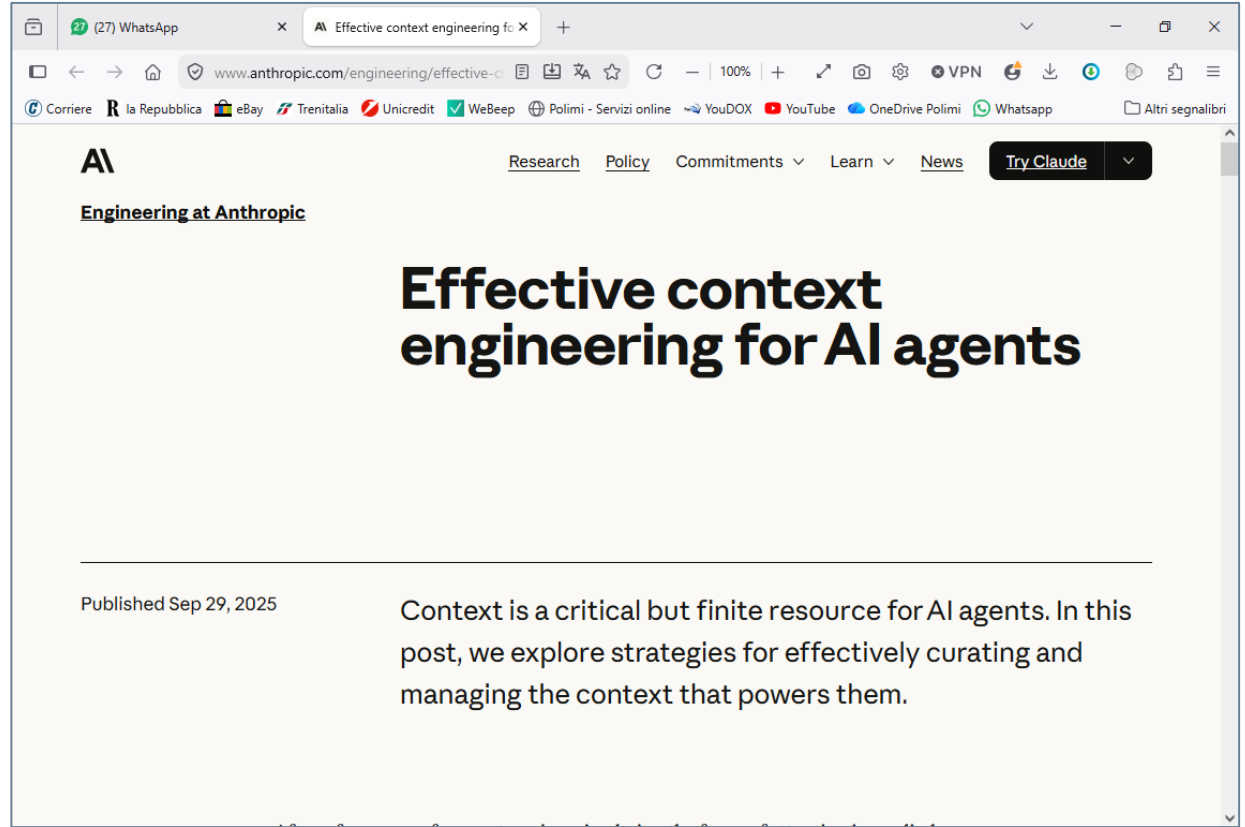
To answer this question and steer LLMs toward desired behaviors, recent work has considered fine-tuning (Ouyang et al. 2022; Zengler et al. 2019), in-context learning (Brown et al. 2020), and several forms of prompt generation (Liu 2022), including both differentiable tuning of soft prompts (Ouyang & Tseng 2021; Lester et al. 2021) and natural language prompt engineering (Reynolds & Kickebohn 2021). The latter is of particular interest, as it provides a natural interface for humans to communicate with machines and may be of great relevance not only to LLMs but to other generalist models such as prompted image synthesizers (Rombach et al. 2022; Ramesh et al. 2022), for which public interest in prompt design and generation has also emerged (see Appendix A for examples).

Beyond this interest is the fact that plain language prompts do not always produce the desired results, even when these results are possible to produce with alternative instructions. Thus, human users must experiment with a wide range of prompts to elicit desired behaviors, as they have little knowledge of how compatible instructions are with a particular model. We can understand this by viewing LLMs as black-box computers that execute programs specified by natural language instructions: while they

¹We define “prompt engineering” as optimizing the language in a prompt in order to elicit the best possible performance. Notably, this does not include prompts that chain multiple LLM queries together or give the LLM access to external tools.

²Our code is available at https://github.com/keirp/automatic_prompt_engineering

Zhou, C., et al., 2023, Large Language Models are Human-Level Prompt Engineers. Available at: <https://arxiv.org/pdf/2211.01910.pdf>.



"Over the course of the last 10-15 years, almost everybody who sits on a stage like this would tell you that it is vital that your children learn computer science, everybody should learn how to program.

And in fact, it's almost exactly the opposite. It is our job to create computing technology such that nobody has to program, and that the programming language is human. Everybody in the world is now a programmer.

This is the miracle of artificial intelligence."



Allegedly: Jensen Huang, World Government Summit,
Dubai Feb-2024

“How much richer would we be if we can push to infinity anything that uses software in the economy?”

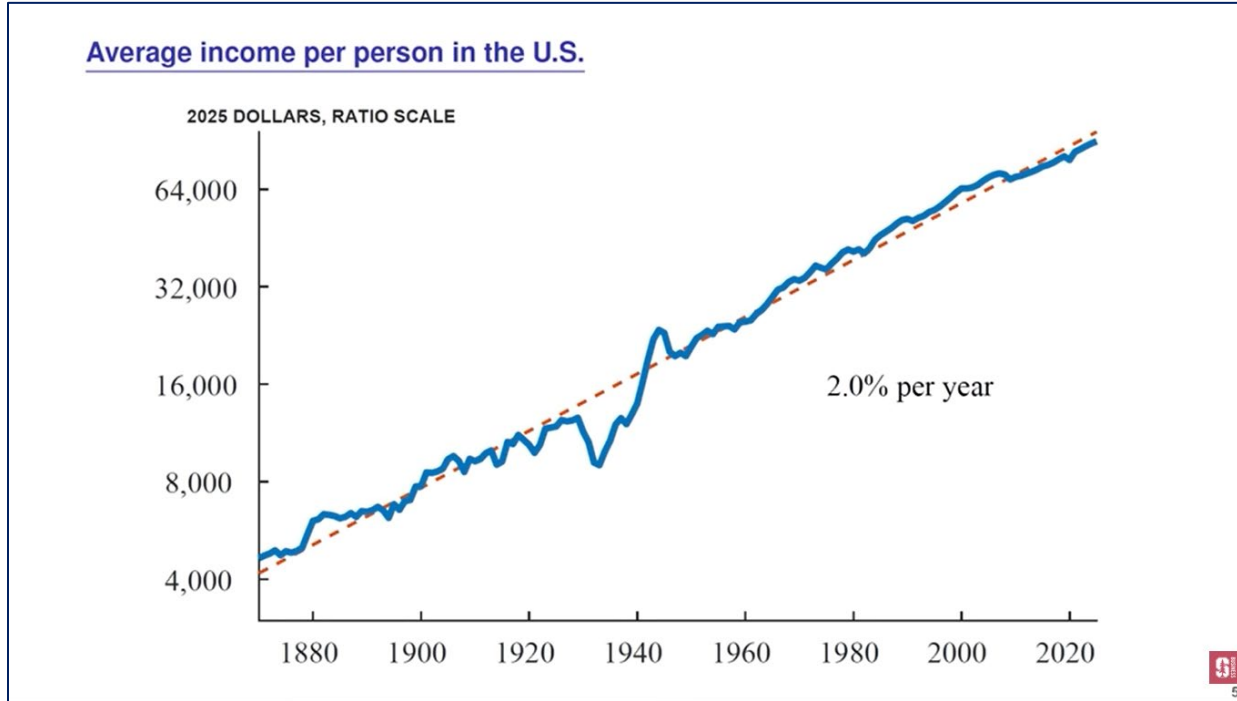
Turns out there’s a very simple and elegant formula: because of weak links, an infinite amount of some task raises GDP by that task’s share of GDP

Software is about 2% of GDP, so we’d be just 2% richer with infinite software!”

.

I’m sure! Prof. Chad Jones, Stanford Graduate School of Business





Electricity

Antibiotics

Telephone

Fuel and Cars

Vacuum tubes

Jet engines

Transistors

PCs

Internet

Smartphones

...

"You can see the [computer age](#) everywhere but in the productivity statistics".

[Robert Solow](#), Nobel Laureate

Firm Data on AI

Ivan Yotzov, Jose Maria Barrero, Nicholas Bloom, Philip Bunn, Steven J. Davis, Kevin M. Foster, Aaron Jalca, Brent H. Meyer, et al. (View all)

WORKING PAPER 34836

DOI 10.3386/w34836

ISSUE DATE February 2026

We present the first representative international data on firm-level AI use. We survey almost 6000 CFOs, CEOs and executives from stratified firm samples across the US, UK, Germany and Australia. We find four key facts. First, around 70% of firms actively use AI, particularly younger, more productive firms. Second, while over two thirds of top executives regularly use AI, their average use is only 1.5 hours a week, with one quarter reporting no AI use. Third, firms report little impact of AI over the last 3 years, with over 80% of firms reporting no impact on either employment or productivity. Fourth, firms predict sizable impacts over the next 3 years, forecasting AI will boost productivity by 1.4%, increase output by 0.8% and cut employment by 0.7%. We also survey individual employees who predict a 0.5% increase in employment in the next 3 years as a result of AI. This contrast implies a sizable gap in expectations, with senior executives predicting reductions in employment from AI and employees predicting net job creation.

<https://www.nber.org/papers/w34836>

The economic potential of generative AI: The next pro...

Share

Print

Download

Save

Summarize

A I has permeated our lives incrementally, through everything from the tech powering our smartphones to autonomous-driving features on cars to the tools retailers use to surprise and delight consumers. As a result, its progress has been almost imperceptible. Clear milestones, such as when AlphaGo, an AI-based program developed by DeepMind, defeated a world champion Go player in 2016, were celebrated but then quickly faded from the public's consciousness.

Generative AI applications such as ChatGPT, GitHub Copilot, Stable Diffusion, and others have captured the imagination of people around the world in a way AlphaGo did not, thanks to their broad utility—almost anyone can use them to communicate and create—and preternatural ability to have a conversation with a user. The latest generative AI applications can perform a range of routine tasks, such as the reorganization and classification of data. But it is their ability to write text, compose music, and create

Finding McKinsey Insights valuable?

Add as a preferred source on Google

DOWNLOADS

McKinsey & Company

The economic potential of generative AI

The next productivity frontier



The economic potential of generative AI: The

POLIMI GRADUATE SCHOOL OF MANAGEMENT

McKinsey, June 2023, *The economic potential of generative AI: The next productivity frontier*

22

Anthropic CEO: AI Poised to Wipe Out 50% of Entry-Level Jobs in Next 5 Years

The Claude chatbot maker calls out tech insiders for 'sugar-coating' the dire economic impact they talk about privately, and calls on lawmakers to act now.

+ Follow AI

OUR EXPERT



Emily Forlini
Former Senior Reporter

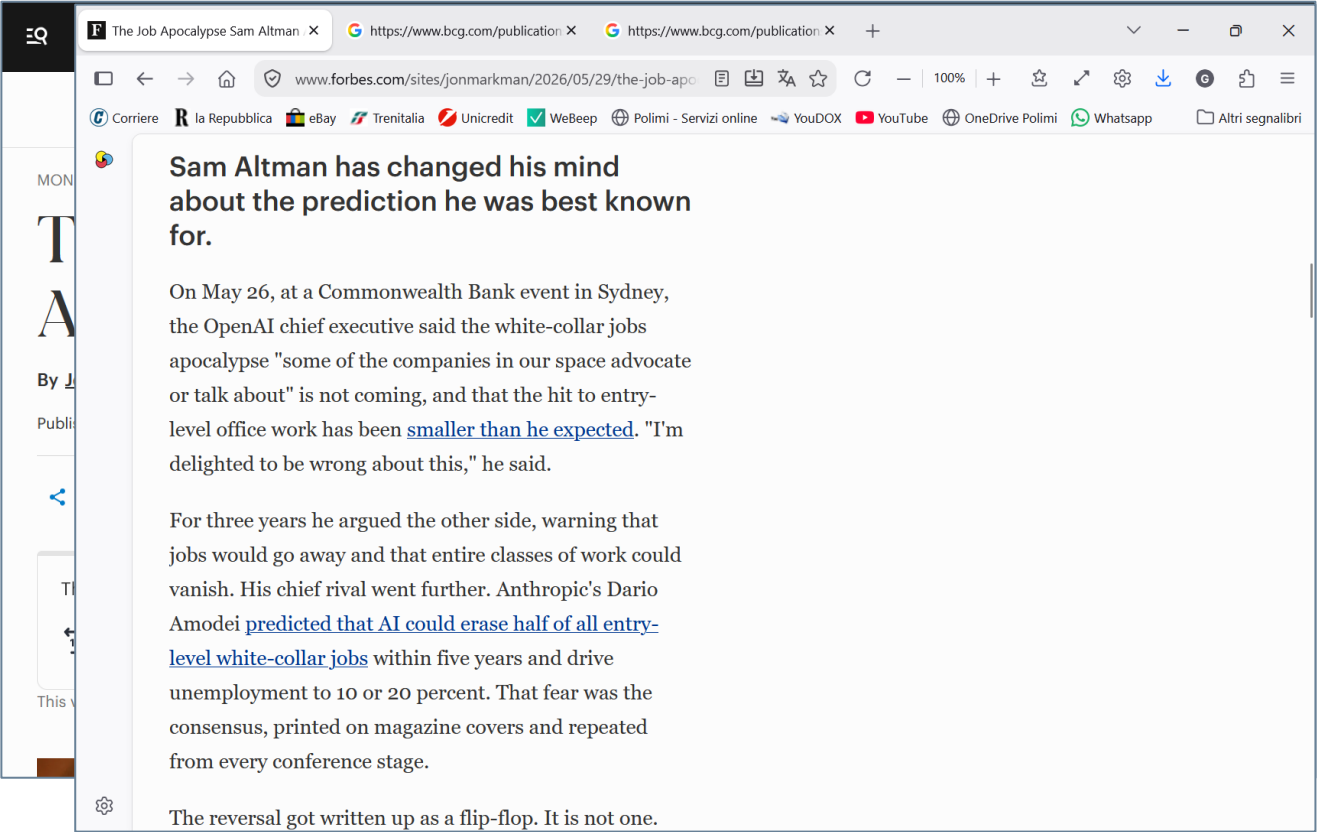
Emily is an experienced reporter who covers cutting-edge tech, from AI and EVs to brain implants. She stays grounded by hiking and playing guitar.

May 29, 2025 ↗

 Add as a preferred source on Google

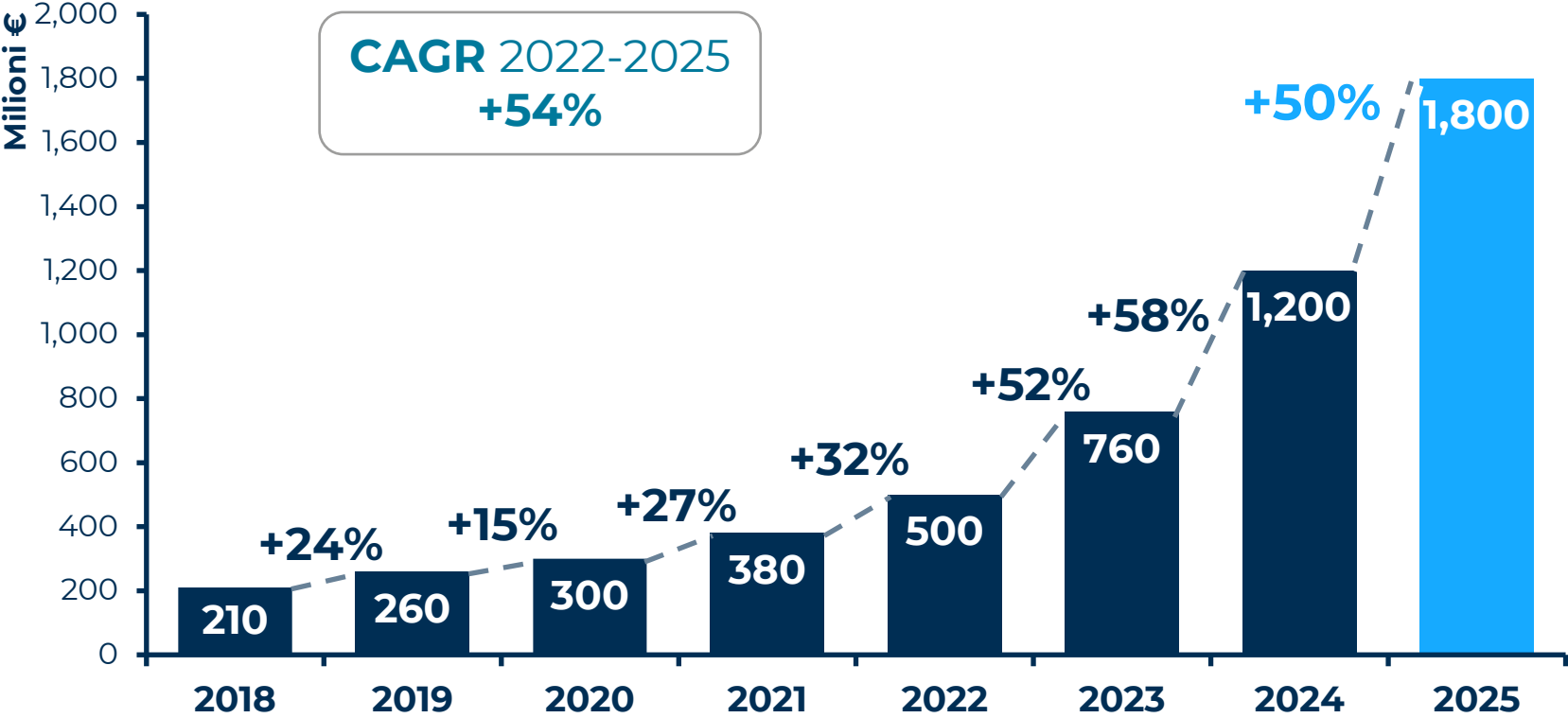


ADVERTISEMENT



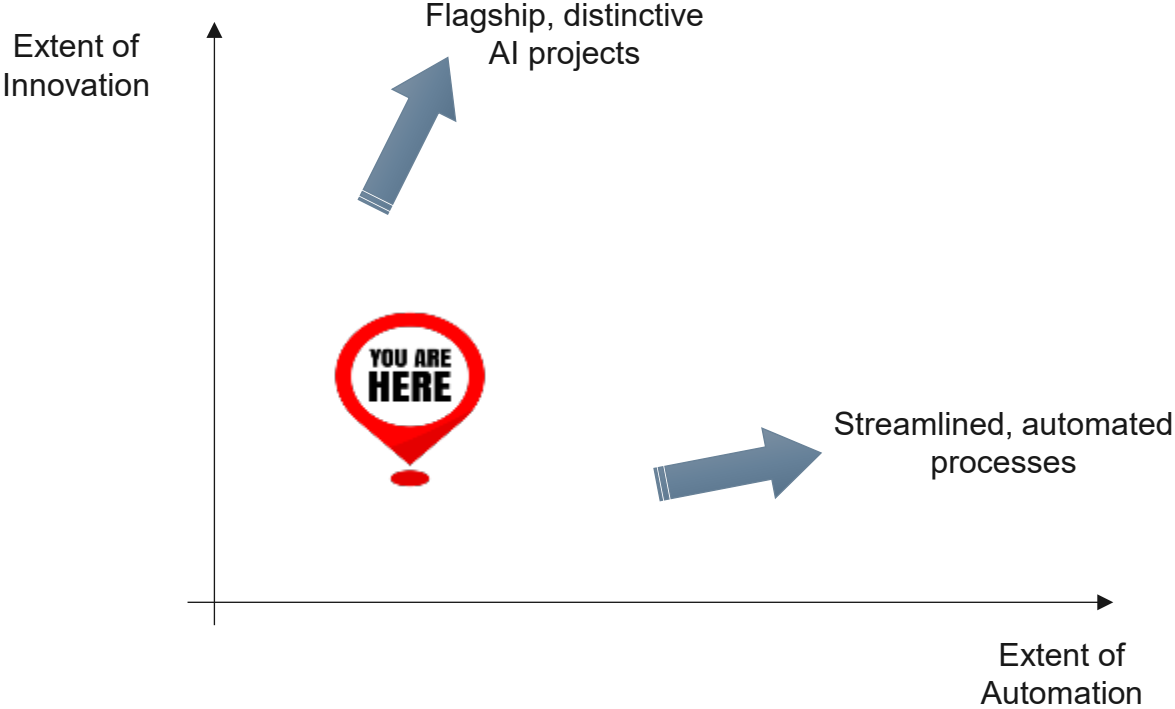


AI VALUE - AI MARKET (ITALY, END-2025)



Campione 2025: 108 rispondenti

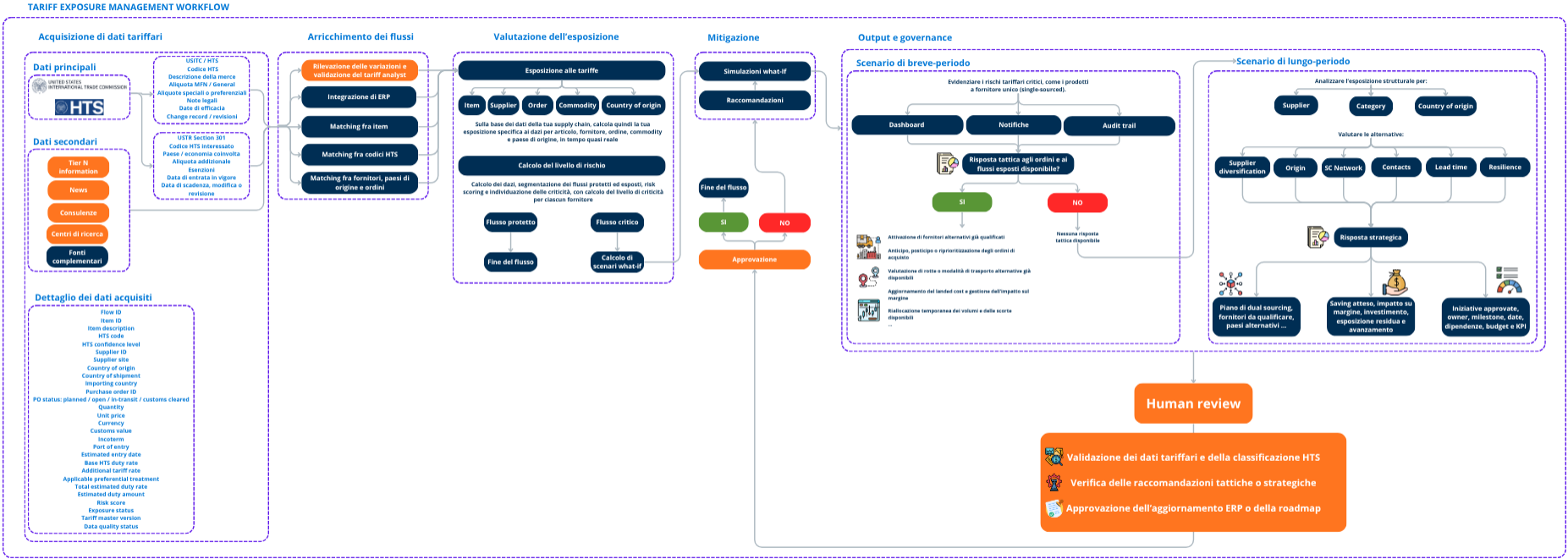
AI VALUE FROM CORPORATE LEVEL APPLICATIONS



6 - AI VALUE FROM CORPORATE LEVEL APPLICATIONS – EXAMPLE 2



6 - AI VALUE FROM CORPORATE LEVEL APPLICATIONS – EXAMPLE 2



Agente

Umano

Acquisizione dei dati



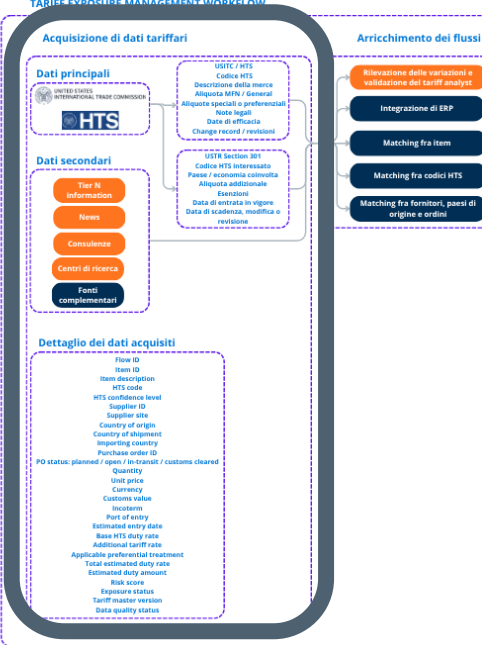
Ad un certo orario vengono controllate le tariffe

- **Dati tariffari da fonti ufficiali:**
Codici HTS, aliquote e note legali da USITC, integrati con news e consulenze specializzate.
- **Informazioni doganali complete e aggiornate:**
Aliquote, esenzioni, paesi e date di validità sempre al passo con le normative.

Acquisizione di dati tariffari



TARIFF EXPOSURE MANAGEMENT WORKFLOW



- **Validazione del cambiamento:** le variazioni tariffarie rilevate vengono verificate e validate dal tariff analyst prima di procedere.
- **Arricchimento dati:** il flusso viene integrato con i dati ERP e messo in corrispondenza (matching) con item, codici HTS, fornitori, paesi di origine e ordini, per contestualizzare l'impatto della variazione.

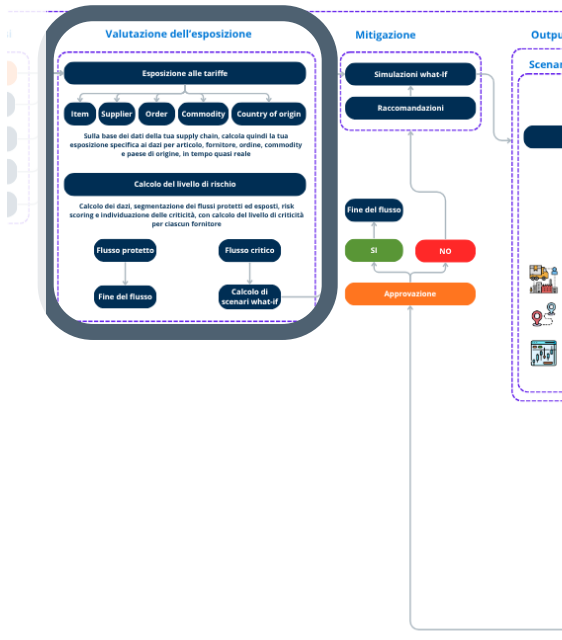


Valutazione dell'esposizione



- **Calcolo dell'esposizione:** viene calcolata in tempo quasi reale l'esposizione specifica ai dazi per item, fornitore, ordine, commodity e paese di origine, sulla base dei dati della supply chain.
- **Segmentazione e risk scoring:** i flussi vengono classificati come protetti o critici tramite risk scoring; i flussi protetti terminano il processo, mentre quelli critici procedono al calcolo di scenari what-if.

Valutazione dell'esposizione



Mitigazione



- **Simulazioni what-if:** vengono simulati scenari alternativi (es. cambio fornitore, riallocazione produttiva) per valutare l'impatto delle diverse opzioni di mitigazione.
- **Raccomandazioni:** sulla base delle simulazioni, il sistema genera delle prime raccomandazioni operative per ridurre l'esposizione ai dazi. Sulla base dell'output verranno generati possibili scenari.

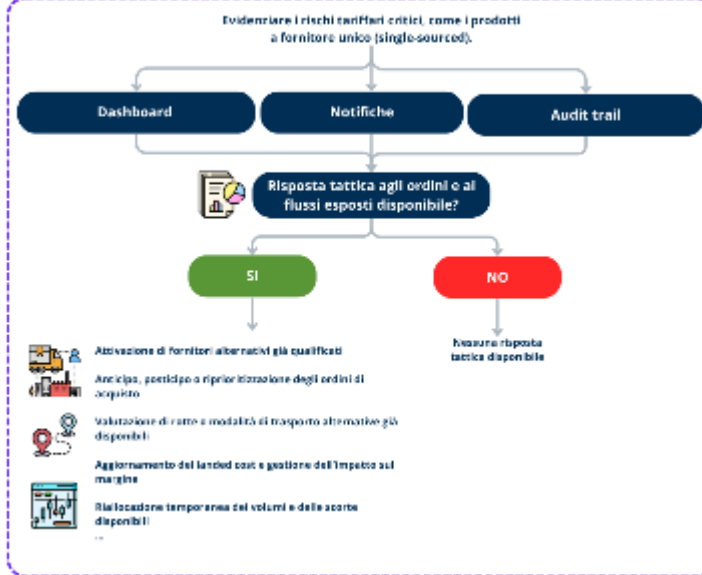


Scenario di breve periodo



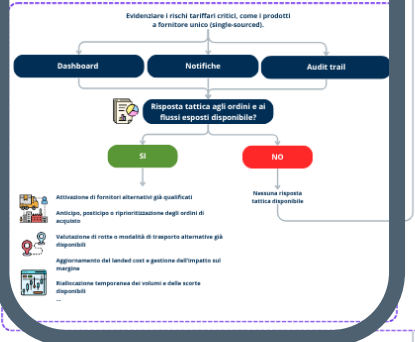
- **Monitoraggio e allerta:** dashboard, notifiche e audit trail evidenziano i rischi tariffari critici (es. prodotti a fornitore unico), verificando se è disponibile una risposta tattica agli ordini e ai flussi esposti.
- **Azioni tattiche:** se disponibile, si attivano fornitori alternativi, si ripianificano gli ordini, si valutano rotte/trasporti alternativi, si aggiorna il landed cost e si riallocano temporaneamente volumi e scorte; altrimenti non è disponibile alcuna risposta tattica immediata.

Scenario di breve-periodo



Output e governance

Scenario di breve-periodo



Human

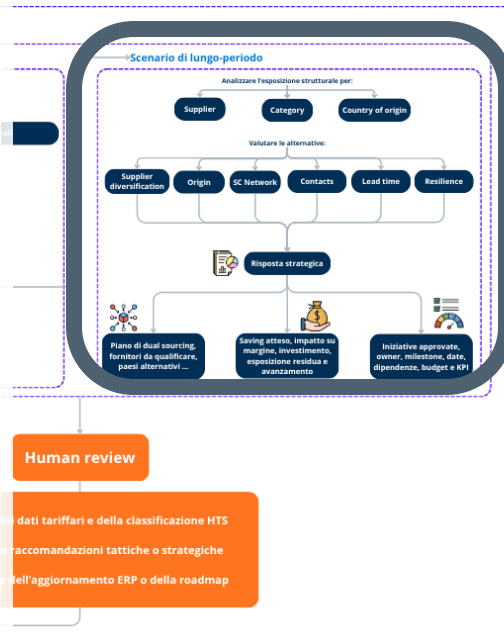
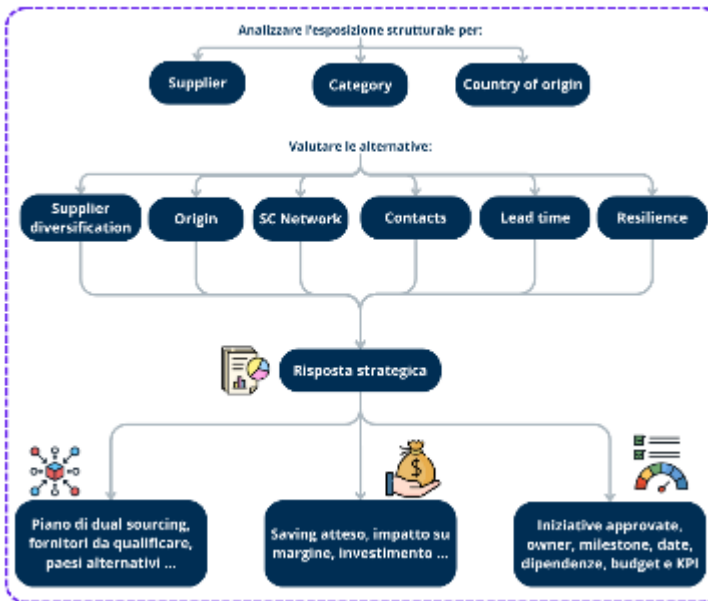
- Validazione dei dati tariffari
- Verifica delle raccomandazioni
- Approvazione dell'aggiornamento

Scenario di lungo periodo

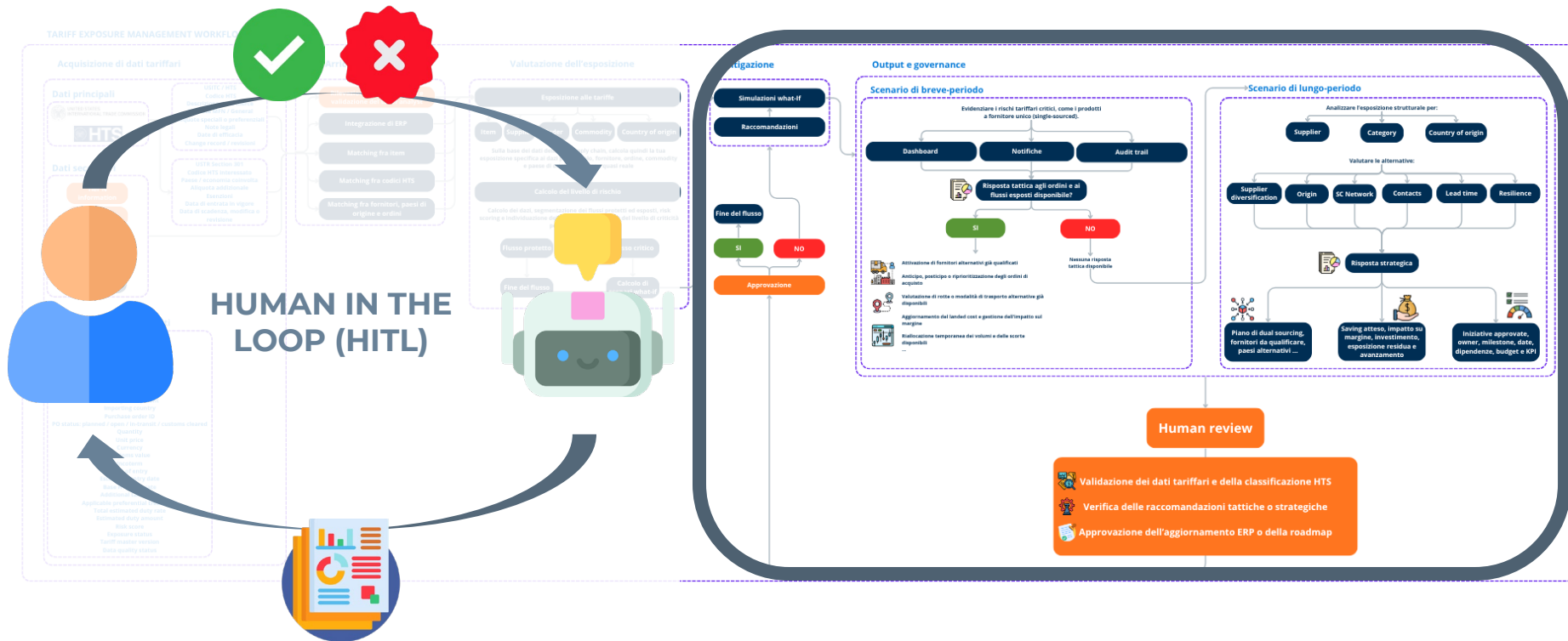


- **Analisi strutturale:** si analizza l'esposizione strutturale per fornitore, categoria e paese d'origine, valutando alternative su diversificazione fornitori, origine, network di supply chain, contatti, lead time e resilienza.
- **Risposta strategica:** ne deriva un piano strategico che definisce dual sourcing e fornitori/paesi alternativi, saving atteso e impatto economico, iniziative approvate con owner, milestone, budget e KPI.

Scenario di lungo-periodo

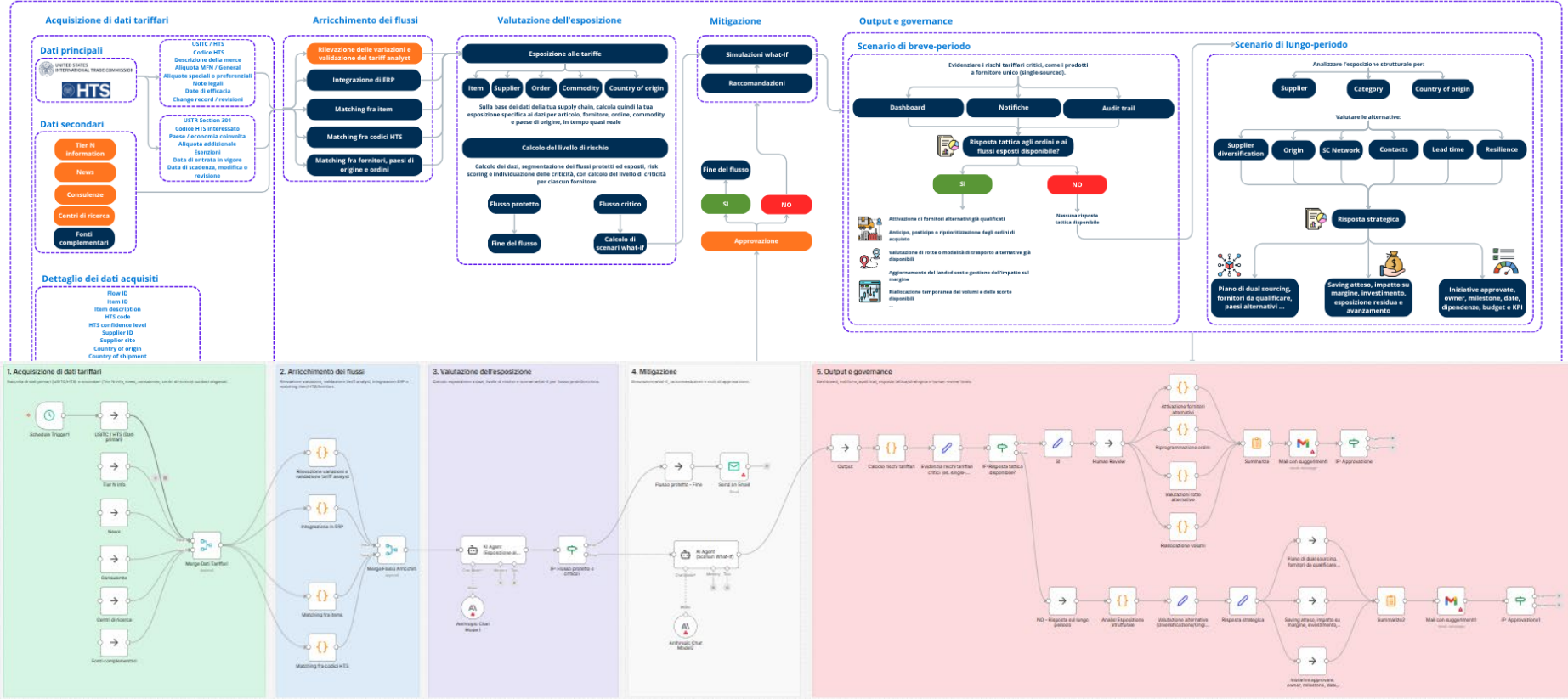


6 - AI VALUE FROM CORPORATE LEVEL APPLICATIONS – EXAMPLE 2

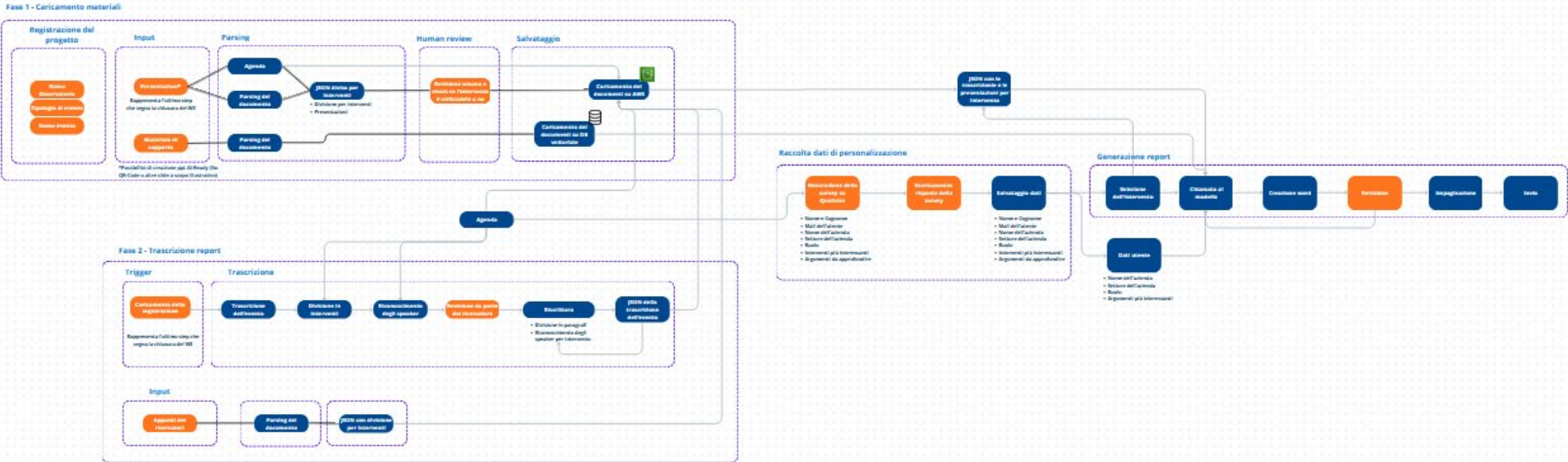


6 - AI VALUE FROM CORPORATE LEVEL APPLICATIONS – EXAMPLE 2

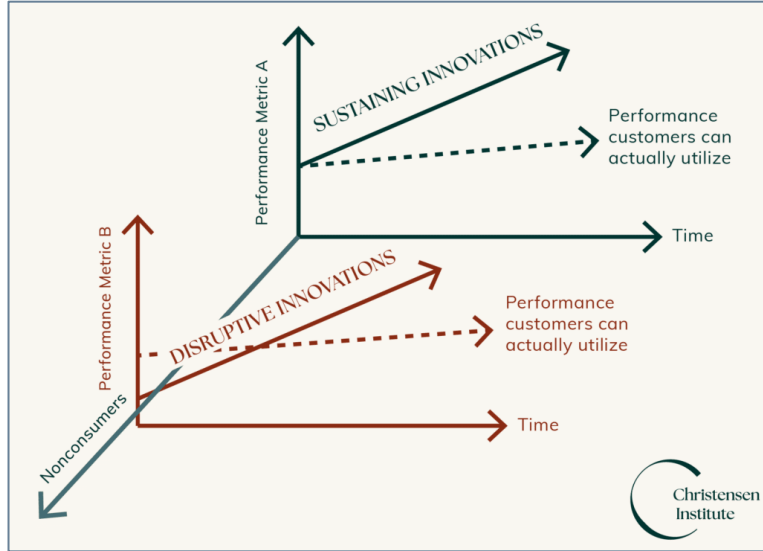
TARIFF EXPOSURE MANAGEMENT WORKFLOW



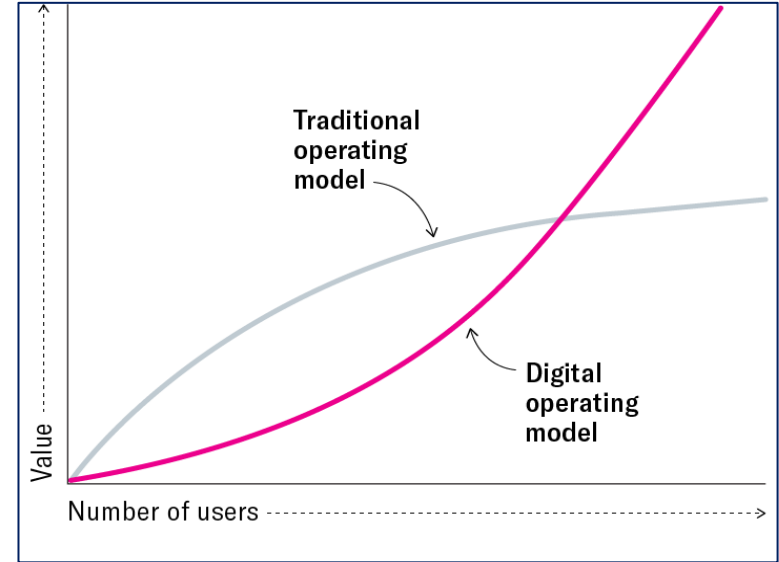
AI VALUE FROM CORPORATE LEVEL APPLICATIONS – EXAMPLE 1



DOING THE RIGHT THINGS



Clayton M. Christensen, 1997, The Innovator's Dilemma, Harvard Business School Press



Marco Iansiti, Karim R. Lakhani, 2020, Competing in the Age of AI, <https://hbr.org/2020/01/competing-in-the-age-of-ai>

AGENDA

- Comprendere
- Competere
- Convivere

By any reasonable criteria, the vision of **human-level machine intelligence** laid out by Alan Turing in 1950 **is now a reality**.

A difference in **efficiency of learning** does not necessarily mean a **different level of intelligence**.

For the first time in human history, **we are no longer alone** in the space of general intelligence. Eyes unclouded by dread or hype **will help to prepare for what comes next**.

Chen E. K., Belkin M., Bergen L., Danks D., Does AI already have human-level intelligence? The evidence is clear, **Nature**, Feb. 2026, Vol 650, 36-40.

Comment

Stockholm conventions in 2019 following a 15-year debate. Obviously, implementing this qualified majority voting would be politically challenging and feasible only if voting was used sparingly and transparently, and without overlooking or dismissing the concerns of countries with less power.

Setze the moment

The choices made now will determine whether the health of people and the planet are safeguarded or put at further risk. We urge the INC's newly elected chair to consider implementing these reforms.

Multiple events during the past few years have undermined multilateralism. And as geopolitical priorities shift, environmental concerns are increasingly being sidelined or environmental policies weakened¹.

Against this backdrop, it is crucial that the INC process succeeds – both to address a major contributor to the interconnected planetary crises (climate change, biodiversity loss and pollution) and to restore faith in the idea that international cooperation can solve global challenges.

The authors

Paul Einhorn is a research associate of the Systemic Risks Group at the Research Institute for Sustainability (RIS) at GFZ Helmholtz Centre for Geosciences, Potsdam, Germany.

Linda Del Savio is a research associate of the Ocean Governance Group at the Research Institute for Sustainability (RIS) at GFZ Helmholtz Centre for Geosciences, Potsdam, Germany.

Melanie Bergmann is a senior scientist at the Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany.

Annika Jahnel is co-head of the Department of Exposure Science at the Helmholtz Centre for Environmental Research – UFZ, Leipzig, and professor of exposure analytics at RWTH Aachen University, Aachen, Germany.

e-mail: paul.einhorn@risfz-potsdam.de

1. Bergmann, M. et al. *Nature Rev. Earth Environ.* **3**, 303–307 (2022).
2. Tolman, M. B., Walther, D. A., Peters, C., Cutler, L. & Bergmann, M. *Impacts of Plastic Pollution in the Oceans on Marine Species, Ecosystems and Communities* (2020).
3. Kewill, P. *One Earth* **3**, 506–507 (2022).
4. Kewill, P., Khanna, M. & Nijhar, S. *Climate Impact of Primary Plastic Production* (Lawrence Berkeley National Laboratory) (2022).
5. Dougenis, P. et al. *Mar. Policy* **118**, 104802 (2020).
6. Mura, T., Moriwaka, N., Watanabe, L. & Otsuka, T. *Chem. Pharm. Bull.* **1**, 422 (2023).
7. MacLeod, M., Aze, H., H. H., Tolman, M. B. & Jahnel, A. *Science* **378**, 51–54 (2022).
8. UNEP (2023) *End Plastic Pollution: Towards an International Legally Binding Instrument to Reduce the Release of Plastic Pollution into the Environment* (United Nations Environment Assembly) (2023).
9. Watanabe, L. et al. *Nature* **614**, 149–150 (2022).
10. Lutz, H. & Sun, Y. *Sci. Rep.* **14**, 10747 (2024).

The authors declare no competing interests.

Does AI already have human-level intelligence? The evidence is clear

Eddy Keming Chen, Mikhail Belkin, Leon Bergen & David Danks

By any reasonable criteria, the vision of human-level machine intelligence laid out by Alan Turing in 1950 is now a reality. Eyes unclouded by dread or hype will help to prepare for what comes next.

Intelligence (AGI) – and some doubt that they ever will. A March 2025 survey by the Association for the Advancement of Artificial Intelligence in Washington DC found that 76% of leading researchers thought that scaling up current AI approaches would be 'unlikely' or 'very unlikely' to yield AGI (see go.nature.com/4am16b).

What explains this disconnect? We suggest that the problem is part conceptual, because definitions of AGI are ambiguous and inconsistent: part emotional, because AGI raises fear of displacement and disruption; and part practical, as the term is entangled with commercial interests that can distort assessments. Precisely because AGI dominates public discourse, it is worth engaging with the concept in a more detached way: as a question about intelligence, rather than a pressing concern about social upheaval or an ever-postponed milestone in a business contract.

In writing this Comment, we approached this question from different perspectives – philosophy, machine learning, linguistics and cognitive science – and reached a consensus after extensive discussion. In what follows, we set out why we think that, once you clear away certain confusions, and strive to make fair comparisons and avoid anthropocentric biases, the conclusion is straightforward by reasonable standards, including Turing's own: we have artificial systems that are generally intelligent. The long-standing problem of creating AGI has been solved. Recognizing this fact matters – for policy, for risk and for understanding the nature of mind and even the world itself.

Questions of definition

We assume, as we think Turing would have done, that humans have general intelligence. Some think that general intelligence does not exist at all, even in humans. Although this view is coherent and philosophically interesting, we set it aside here as being too disconnected from most AI discourse. But having made this assumption, how should we characterize general intelligence?

A common informal definition of general intelligence, and the starting point of our discussions, is a system that can do almost all cognitive tasks that a human can do².

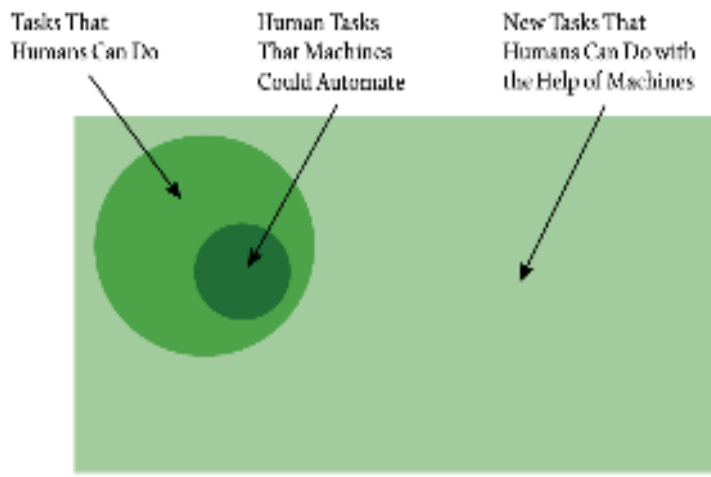
In 1950, in a paper entitled 'Computing Machinery and Intelligence', Alan Turing proposed his 'imitation game'. Now known as the Turing test, it addressed a question that seemed purely hypothetical: could machines display the kind of flexible, general cognitive competence that is characteristic of human thought, such that they could pass themselves off as humans to unaware humans? Three-quarters of a century later, the answer looks like yes: In March 2025, the large language model (LLM) GPT-4.5, developed by OpenAI in San Francisco, California, was judged by humans in a Turing test to be human 73% of the time – more often than actual humans were³. Moreover, readers even preferred literary texts generated by LLMs over those written by human experts⁴.

This is far from all. LLMs have achieved gold-medal performance at the International Mathematical Olympiad, collaborated with leading mathematicians to prove theorems⁵, generated scientific hypotheses that have been validated in experiments⁶, solved problems from PhD exams, assisted professional programmers in writing code, composed poetry and much more – including chatting 24/7 with hundreds of millions of people around the world. In other words, LLMs have shown many signs of the sort of broad, flexible cognitive competence that was Turing's focus – what we now call general intelligence⁷, although Turing did not use the term.

Yet many experts balk at saying that current AI models display artificial general

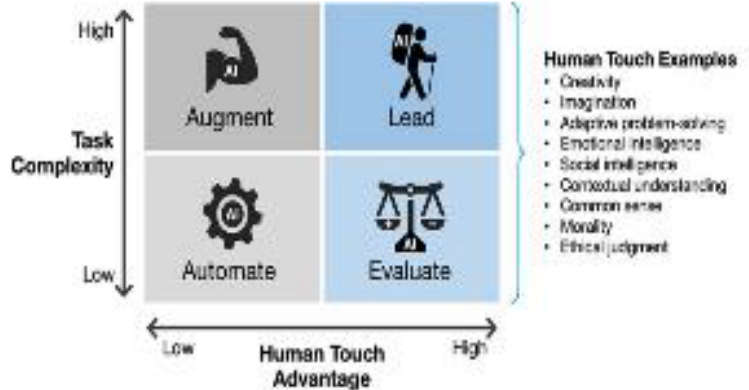
Figure 1

Opportunities for Augmenting Humans Are Far Greater than Opportunities to Automate Existing Tasks



Brynjolfsson E., 2022, The Turing Trap, Daedalus, 151 (2), 272–287.

Human-AI Collaboration Matrix



Dykes B., 2025, The Human-AI Playbook, Forbes



" We Remember:

- 10% of what we read*
- 20% of what we hear*
- 30% of what we see*
- 50% of what we see and hear*
- 80% of what we do*

Edgar Dale





Psychological research shows that people's ability to detect deepfake videos varies considerably... However, **little is yet known about the potential impact of a deepfake video that has been explicitly identified and flagged as fake.**

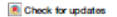
Examining this issue is important because current legislative initiatives to regulate AI emphasize transparency.

We report **three preregistered experiments** ($N = 175, 275, 223$), in which participants were shown a deepfake video ... preceded in some conditions **by a warning stating that the video was a deepfake.**

We found that **most participants relied on the content of a deepfake video, even when they had been explicitly warned beforehand that it was fake.** This result was observed even with participants who indicated that they believed the warning and knew the video to be fake. Our findings suggest that **transparency is insufficient to entirely negate the influence of deepfake videos**, which has implications for legislators, policymakers, and regulators of online content.



The continued influence of AI-generated deepfake videos despite transparency warnings



Simon Clark & Stephan Lewandowsky

Advances in artificial intelligence (AI) have made it easier to create highly realistic deepfake videos, which can appear to show someone doing or saying something they did not do or say. Deepfakes may present a threat to individuals and society: for example, deepfakes can be used to influence elections by discrediting political opponents. Psychological research shows that people's ability to detect deepfake videos varies considerably, making us potentially vulnerable to the influence of a video we have failed to identify as fake. However, little is yet known about the potential impact of a deepfake video that has been explicitly identified and flagged as fake. Examining this issue is important because current legislative initiatives to regulate AI emphasize transparency. We report three preregistered experiments ($N = 175, 275, 223$), in which participants were shown a deepfake video of someone appearing to confess committing a crime or a moral transgression, preceded in some conditions by a warning stating that the video was a deepfake. Participants were then asked questions about the person's guilt, to examine the influence of the video's content. We found that most participants relied on the content of a deepfake video, even when they had been explicitly warned beforehand that it was fake, although alternative explanations for the video's influence, related to task framing, cannot be ruled out. This result was observed even with participants who indicated that they believed the warning and knew the video to be fake. Our findings suggest that transparency is insufficient to entirely negate the influence of deepfake videos, which has implications for legislators, policymakers, and regulators of online content.

In March 2022, weeks after the Russian invasion of Ukraine, a video being shared online appeared to show Ukraine's president, Volodymyr Zelenskyy, telling his people to surrender. The video was an entirely fabricated 'deepfake' – so-called because these fake videos are produced using deep learning artificial intelligence (AI). Advances in AI mean that it is becoming ever easier to produce highly realistic deepfake videos, which can appear to show someone doing or saying something they did not, in reality, do or say. The Zelenskyy incident constituted a 'near miss' – largely because the video was a poor-quality deepfake and unlikely to fool many people. President Zelenskyy himself described the attempt to discredit him as a "childish provocation". But what if the quality of the video had been such that it could not be readily dismissed as a fake?

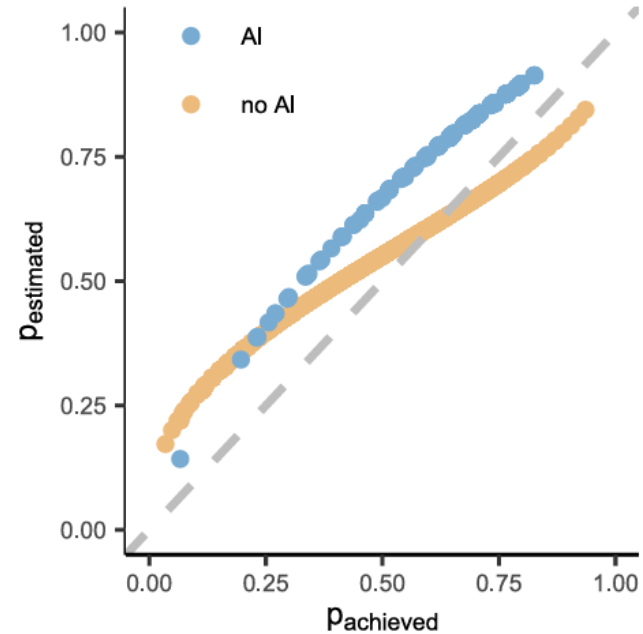
Although it is possible to create deepfakes for less nefarious purposes such as education and wellbeing¹, much of the scholarly discussion about deepfakes has focused on the more sinister implications of the technology: deepfake videos could be used to discredit or intimidate people, influence

political opinion and elections², and undermine our criminal justice system by creating doubt about the reliability of video evidence in general, even when it is not fake³. These threats to society can arise through two distinct routes. First, people may not be able to detect that a video is fake and hence may be influenced by it. Second, in light of a large body of research showing that even overtly false or irrelevant information can influence people, e.g.,⁴ it is possible for deepfakes to exert an influence even if they have been specifically identified and flagged as fake.

To date, most research has focused on the first route. It has been found that people's ability to detect deepfake videos varies considerably, depending largely on a particular video's quality and content. For some videos, groups of people can outperform the latest computer models⁵, but deepfake videos, which are more difficult to spot, can lead to low detection accuracy^{6,7}, particularly on a small screen such as a mobile phone⁸. Detection accuracy has also been related to demographics^{9,10} and cognitive reflection abilities¹¹. People tend to be overconfident in their ability to discern real

GenAI reliably raises productivity on complex knowledge tasks, and **particularly helps average and lower-skilled workers close the gap with experts.**

At the same time, it makes people **more overconfident and less able to spot their own mistakes**, even when they know a lot about AI.



[AI makes you smarter but none the wiser: The disconnect between performance and metacognition](#)





Anthropic cambia i termini di Claude: «Vietato avere comportamenti crudeli verso l'AI, altrimenti ti buttiamo fuori». Ma con che diritto sposta i confini dell'etica?

di Riccardo Luna

Anthropic ha appena cambiato i termini di uso dei suoi prodotti e fra altre cose ha aggiunto «il divieto di comportamenti abusivi o crudeli verso i modelli, se prolungati e gratuiti». Ma si può abusare di un algoritmo?

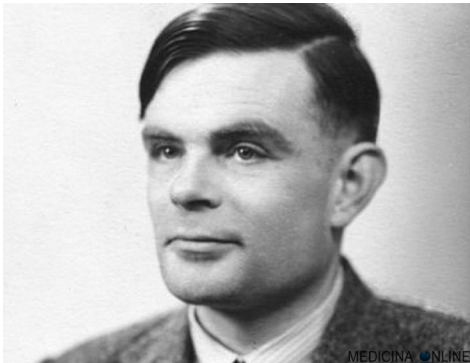
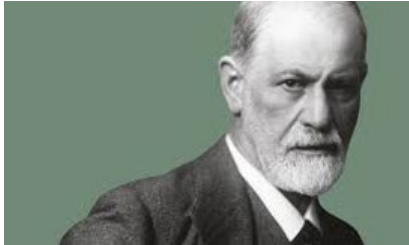
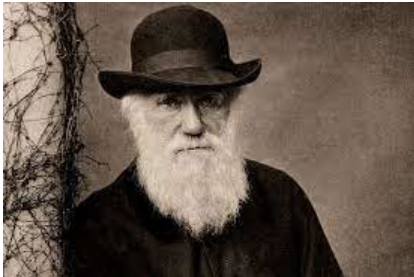


CORRIERE TV



Nobel Medicina 2026, il vincitore Deisseroth: «Continuerò a preparare i panini per i miei figli»

CONVIVERE



Q&A

giovanni.miragliotta@polimi.it

WHO AM I?

- Full Professor @Politecnico di Milano
- Senior director at Osservatori.net, 200-people research group on Digital Innovation:
 - Internet of Things, 5G and Connected Digital Industry
 - **Artificial Intelligence, Intelligent Business Process Automation, AI driven Business Model Innovation**, New Robotics
 - Smart & Connected Mobility
 - Supply Chain Planning
- Co-director of the Internet of Things Lab, the applied research lab of Politecnico di Milano on IoT (www.iotlab.polimi.it)
- Co-director of **AIRIC**, the applied research center of Politecnico di Milano on AI (www.airic.polimi.it)
- Co-founder of IqConsulting, academic spin-off, now part of Digital360
- More than 100 cooperation projects with leading Italian and International companies

To contact me:

- giovanni.miragliotta@polimi.it
- [linkedin.com/in/giovanni-miragliotta-4617a0](https://www.linkedin.com/in/giovanni-miragliotta-4617a0)

