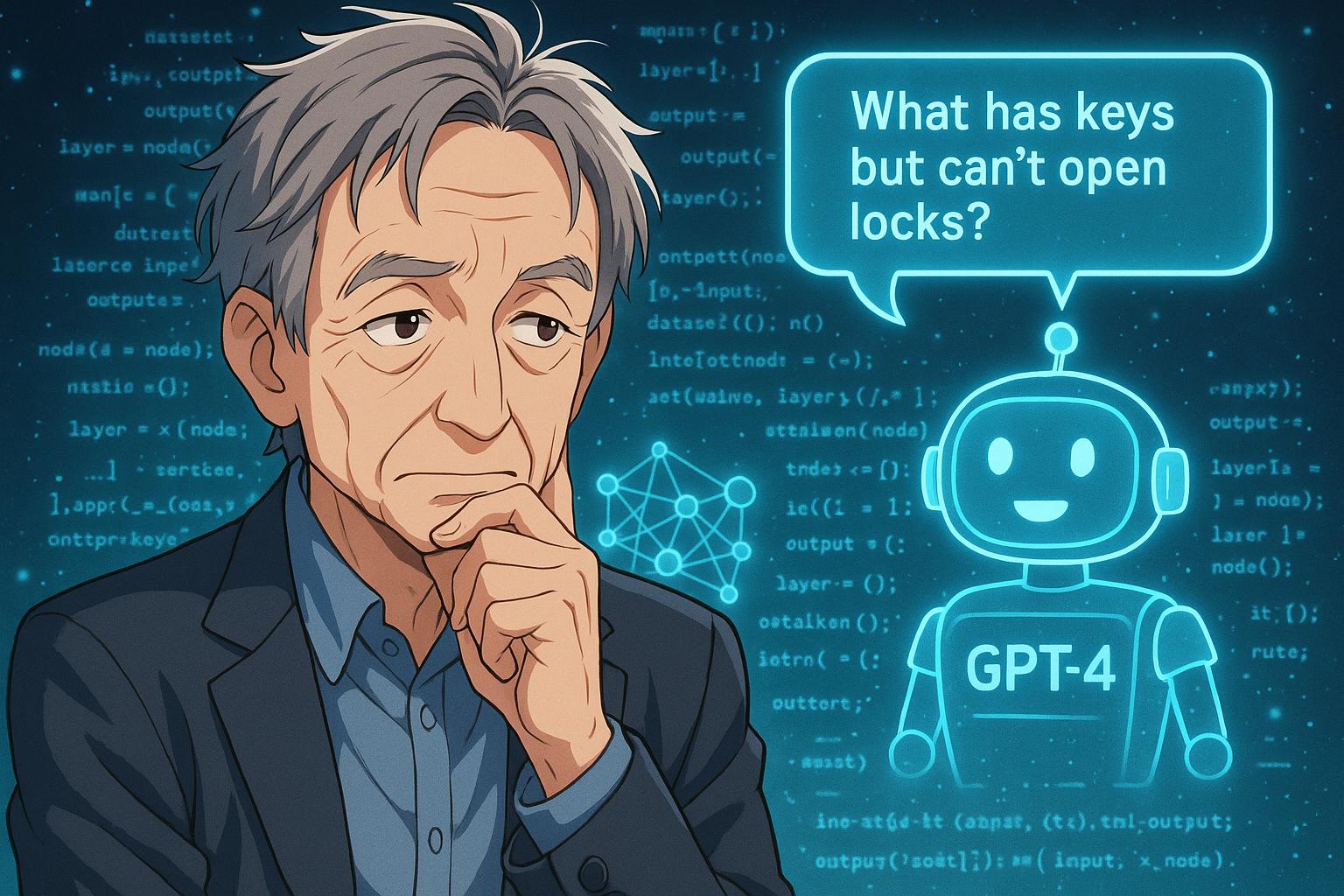
# Geoffrey Hinton warns of misplaced trust in GPT-4 despite optimism for future AI



Geoffrey Hinton, often hailed as the “Godfather of AI,” recently voiced his concerns regarding the trust placed in artificial intelligence, particularly OpenAI’s GPT-4. In a candid CBS interview, Hinton reflected on his own reliance on the chatbot for various everyday tasks, admitting that he sometimes trusts its outputs more than he ought to. “I tend to believe what it says, even though I should probably be suspicious,” he revealed, a sentiment that strikes at the heart of growing anxiety surrounding the reliance on AI technologies.

During his discussion, Hinton presented GPT-4 with a seemingly simple riddle: “Sally has three brothers. Each of her brothers has two sisters. How many sisters does Sally have?” The model responded incorrectly, stating two when the answer is one—Sally herself. This misstep not only surprised Hinton but also serves as a reminder of the limitations that even the most advanced AI still possesses. He characterised GPT-4 as “an expert at everything,” while also acknowledging its inconsistent performance.

Yet, despite the occasional blunders, Hinton remains optimistic about the trajectory of AI development. He expressed hope that future iterations, such as GPT-5, would demonstrate improved capabilities. “Yeah, I suspect,” he remarked when asked about whether the new model would answer the riddle correctly. After the interview aired, users noted that more recent versions like GPT-4o and GPT-4.1 successfully solved the riddle, reflecting the rapid advancements in AI technology.

Hinton’s insights come at a time when the AI landscape is evolving swiftly. Launched in 2023, GPT-4 became a benchmark for advanced reasoning and language tasks within the tech industry. OpenAI’s continued innovation has seen the release of enhanced models, including GPT-4.1 and GPT-4o, which offer even faster and more versatile performance. Meanwhile, competitors like Google’s Gemini 2.5-Pro have emerged to dominate the independent Chatbot Arena leaderboard, keeping pressure on OpenAI to maintain its competitive edge.

This backdrop of fierce competition is coupled with Hinton’s increasing concerns regarding the broader implications of artificial intelligence. In a June 2024 interview with CBS's 60 Minutes, he warned of potential risks associated with AI systems, particularly the prospect of these technologies surpassing human control if developed without proper oversight. “Humanity is at a turning point with AI,” he cautioned, highlighting the urgent need for ethical guidelines to govern AI development and deployment.

Furthermore, Hinton’s reflections on AI’s capabilities extend to its surprising potential to emulate human-like thinking and reasoning. In a podcast with The Economist, he expressed astonishment at AI's rapid advancements and its potential to eventually outperform human capabilities. Yet, alongside this optimism lies a deep-seated fear of the existential risks posed by uncontrolled AI development, a sentiment echoed in his earlier discussions on PBS regarding the societal threats posed by rapidly evolving technologies.

As AI continues to transform industries and everyday life, the dialogue around trust, safety, and ethical boundaries becomes increasingly crucial. Hinton’s insights remind us that while we stand on the precipice of a new era in technology, careful consideration and regulation are essential to navigate the complexities and risks that such advancements entail.

### Reference Map

1. Paragraph 1: Sources (1), (2), (4)
2. Paragraph 2: Sources (1), (3)
3. Paragraph 3: Sources (1), (4)
4. Paragraph 4: Sources (1), (4), (5)
5. Paragraph 5: Sources (3), (6), (7)
6. Paragraph 6: Sources (4), (5), (6)

Source: [Noah Wire Services](https://www.noahwire.com)

## Bibliography

1. <https://www.businesstoday.in/technology/news/story/godfather-of-ai-geoffrey-hinton-says-he-trusts-openais-gpt-4-chatbot-more-than-he-should-476910-2025-05-20> - Please view link - unable to able to access data
2. <https://www.cbs.com/shows/video/wcagJ9qtiVJAyCdg_6EETYDGyItFHfmr/> - In a March 2023 interview on CBS Saturday Morning, Geoffrey Hinton, known as the 'Godfather of AI,' discussed the evolution of artificial intelligence, highlighting its rapid advancements and potential future developments. He emphasized the importance of understanding AI's capabilities and limitations, noting that while AI has made significant strides, it still has a long way to go before achieving human-like understanding and reasoning.
3. <https://www.economist.com/podcasts/2025/03/12/ai-is-more-human-than-you-think-an-interview-with-geoffrey-hinton> - In a March 2025 podcast interview with The Economist, Geoffrey Hinton discussed the human-like qualities of artificial intelligence, expressing surprise at the rapid development of AI technologies. He highlighted the potential for AI to think, reason, and understand the world in ways that could eventually surpass human capabilities, reflecting on the unexpected trajectory of AI's evolution.
4. <https://www.cbsnews.com/amp/news/artificial-intelligence-risks-dangers-geoffrey-hinton-60-minutes/> - In a June 2024 interview on CBS's 60 Minutes, Geoffrey Hinton warned about the potential risks of artificial intelligence, including the possibility of AI systems surpassing human control. He emphasized the need for responsible development and regulation to prevent unintended consequences, expressing concerns about AI's potential to 'take over' if not developed ethically.
5. <https://web.cs.toronto.edu/news-events/news/humanity-is-at-a-turning-point-with-ai-geoffrey-hinton-tells-cbs-newss-60-minutes> - In a June 2024 interview with CBS News's 60 Minutes, Geoffrey Hinton discussed the rapid development of artificial intelligence and its potential implications for humanity. He expressed concerns about AI's trajectory and the uncertainty surrounding its future impact, emphasizing the need for careful consideration and responsible development to ensure beneficial outcomes.
6. <https://www.pbs.org/newshour/show/godfather-of-ai-discusses-dangers-the-developing-technologies-pose-to-society> - In a May 2023 interview with PBS News, Geoffrey Hinton discussed the potential dangers posed by developing artificial intelligence technologies. He highlighted concerns about AI's rapid expansion and the need for careful consideration of its societal impacts, emphasizing the importance of understanding and mitigating potential risks associated with AI advancements.
7. <https://www.pbs.org/wnet/amanpour-and-company/video/geoffrey-hinton-warns-of-the-existential-threat-of-ai/> - In a May 2023 interview on PBS's Amanpour & Company, Geoffrey Hinton discussed the existential threats posed by artificial intelligence. He expressed concerns about AI's rapid development and the potential for unintended consequences, emphasizing the need for responsible development and regulation to ensure AI technologies are used ethically and safely.