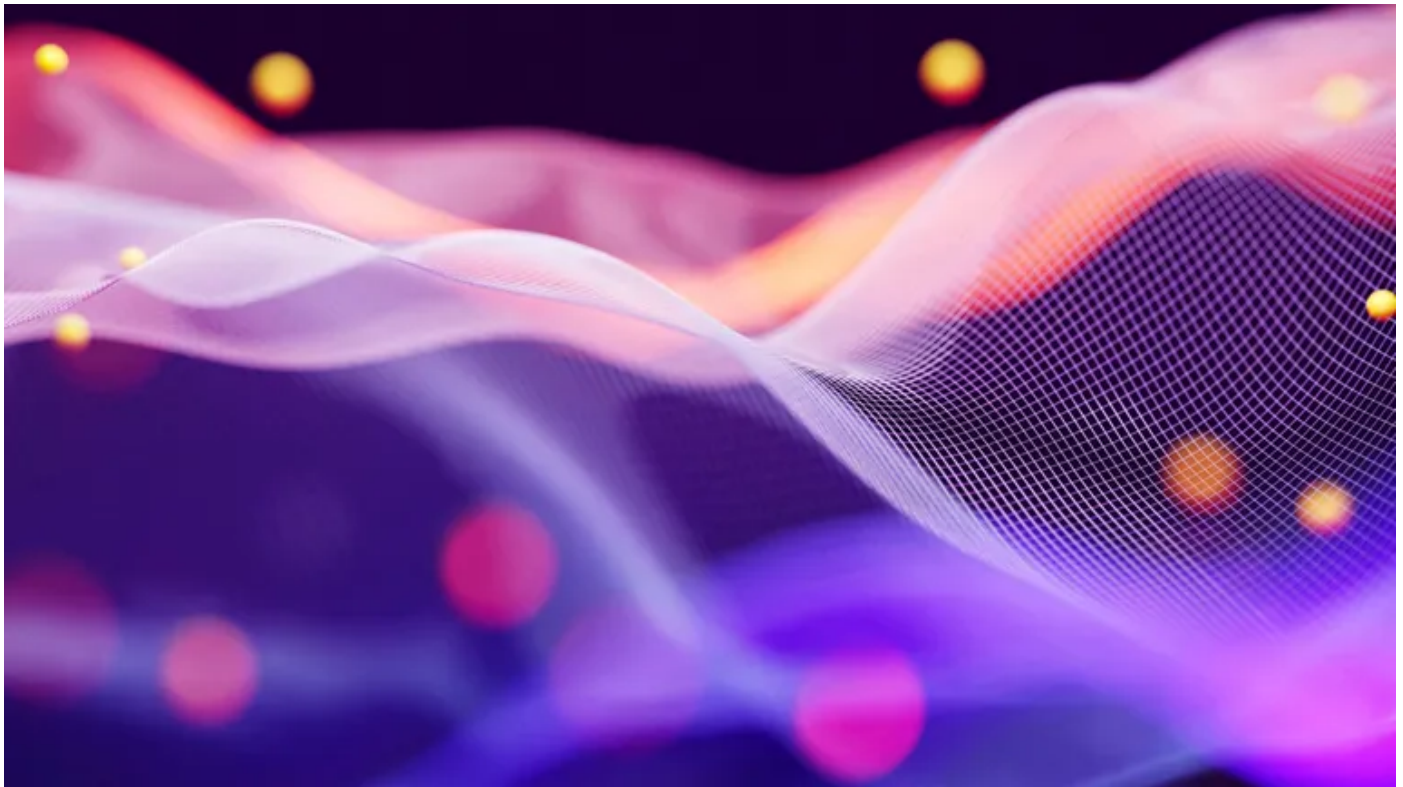


# AI has grown beyond human knowledge, says Google's DeepMind unit

A new agentic approach called 'streams' will let AI models learn from the experience of the environment without human 'pre-judgment'.

April 18, 2025 at 2:38 a.m. PT



worawit chutrakunwanit/Getty Images

The world of [artificial intelligence](#) (AI) has recently been preoccupied with advancing [generative AI](#) beyond simple tests [that AI models easily pass](#). The famed Turing Test [has been "beaten" in some sense](#), and controversy rages over whether the newest models are being built to [game the benchmark](#)

[tests](#) that measure performance.

The problem, say scholars at Google's DeepMind unit, is not the tests themselves but the limited way AI models are developed. The data used to train AI is too restricted and static, and will never propel AI to new and better abilities.

**Also: [OpenAI's most impressive move has nothing to do with AI](#)**

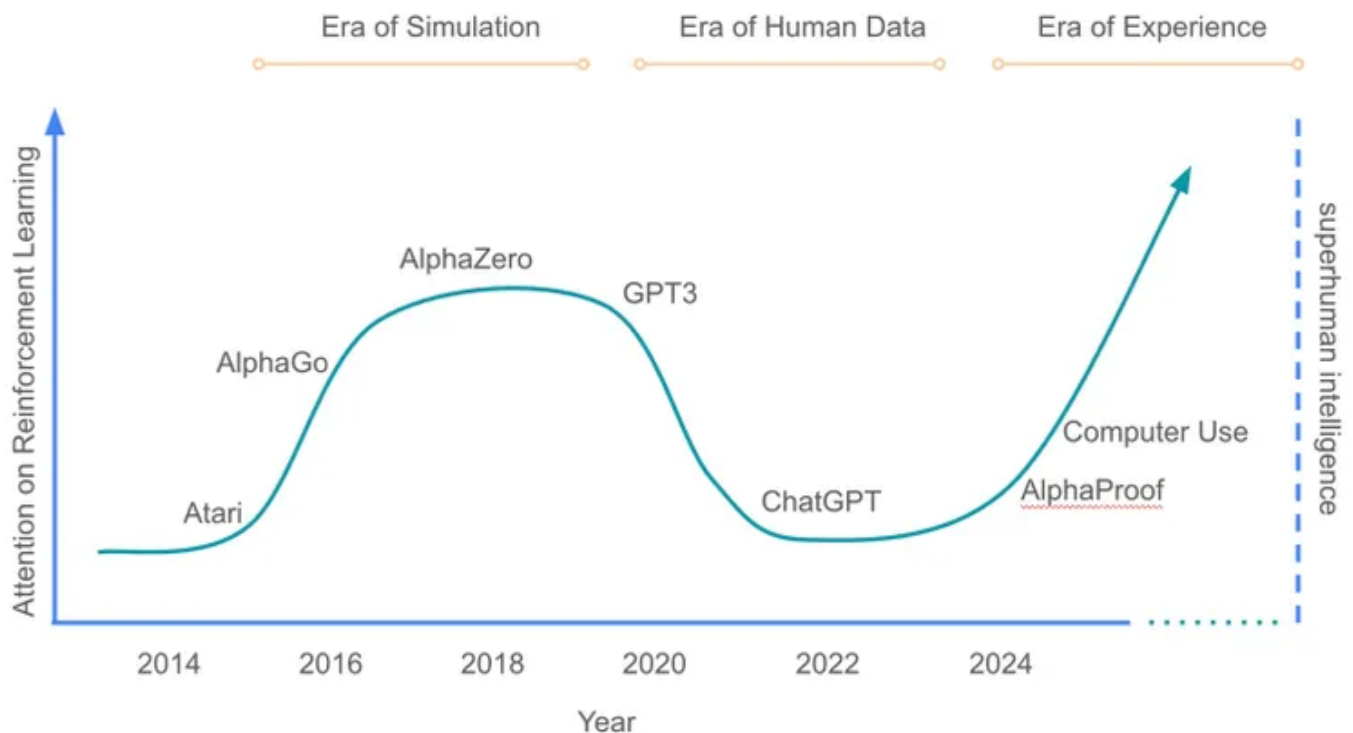
In a paper [posted by DeepMind last week](#), part of a forthcoming book by MIT Press, researchers propose that AI must be allowed to have "experiences" of a sort, interacting with the world to formulate goals based on signals from the environment.

"Incredible new capabilities will arise once the full potential of experiential learning is harnessed," write DeepMind scholars David Silver and Richard Sutton in the paper, *Welcome to the Era of Experience*.

The two scholars are legends in the field. Silver most famously led the research that resulted in [AlphaZero](#), DeepMind's AI model that beat humans in games of Chess and Go. Sutton is one of [two Turing Award-winning developers](#) of an AI approach called reinforcement learning that Silver and his team used to create AlphaZero.

**Also: [With AI models clobbering every benchmark, it's time for human evaluation](#)**

The approach the two scholars advocate builds upon reinforcement learning and the lessons of AlphaZero. It's called "streams" and is meant to remedy the shortcomings of today's large language models (LLMs), which are developed solely to answer individual human questions.



Google DeepMind

Silver and Sutton suggest that shortly after AlphaZero and its predecessor, AlphaGo, burst on the scene, generative AI tools, such as [ChatGPT](#), took the stage and "discarded" reinforcement learning. That move had benefits and drawbacks.

**Also: [OpenAI's Deep Research has more fact-finding stamina than you, but it's still wrong half the time](#)**

Gen AI was an important advance because AlphaZero's use of reinforcement learning was restricted to limited applications. The technology couldn't go beyond "full information" games, such as Chess, where all the rules are known.

Gen AI models, on the other hand, can handle spontaneous input from humans never before encountered, without explicit rules about how things are supposed to turn out.

However, discarding reinforcement learning meant, "something was lost in this transition: an agent's ability to self-discover its own knowledge," they write.

Instead, they observe that LLMs "[rely] on human prejudgment", or what the human wants at the prompt stage. That approach is too limited. They suggest that human judgment "imposes "an impenetrable ceiling on the agent's performance: the agent cannot discover better strategies underappreciated by the human rater.

## Newsletters

### ZDNET Tech Today

ZDNET's Tech Today newsletter is a daily briefing of the newest, most talked about stories, five days a week.

Subscribe

[See all](#)

Not only is human judgment an impediment, but the short, clipped nature of prompt interactions never allows the AI model to advance beyond question and answer.

"In the era of human data, language-based AI has largely focused on short interaction episodes: e.g., a user asks a question and (perhaps after a few thinking steps or tool-use actions) the agent responds," the researchers write.

"The agent aims exclusively for outcomes within the current episode, such as directly answering a user's question."

There's no memory, there's no continuity between snippets of interaction in prompting. "Typically, little or no information carries over from one episode to the next, precluding any adaptation over time," write Silver and Sutton.

**Also: [The AI model race has suddenly gotten a lot closer, say Stanford scholars](#)**

However, in their proposed Age of Experience, "Agents will inhabit streams of experience, rather than short snippets of interaction."

Silver and Sutton draw an analogy between streams and humans learning over a lifetime of accumulated experience, and how they act based on long-range goals, not just the immediate task.

"Powerful agents should have their own stream of experience that progresses, like humans, over a long time-scale," they write.

Silver and Sutton argue that "today's technology" is enough to start building streams. In fact, the initial steps along the way can be seen in developments such as [web-browsing AI agents, including OpenAI's Deep Research](#).

"Recently, a new wave of prototype agents have started to interact with computers in an even more general manner, by using the same interface that humans use to operate a computer," they write.

The browser agent marks "a transition from exclusively human-privileged communication, to much more autonomous interactions where the agent is able to act independently in the world."

**Also: [The Turing Test has a problem - and OpenAI's GPT-4.5 just exposed it](#)**

As AI agents move beyond just web browsing, they need a way to interact

and learn from the world, Silver and Sutton suggest.

They propose that the AI agents in streams will learn via the same reinforcement learning principle as AlphaZero. The machine is given a model of the world in which it interacts, akin to a chessboard, and a set of rules.

As the AI agent explores and takes actions, it receives feedback as "rewards". These rewards train the AI model on what is more or less valuable among possible actions in a given circumstance.

The world is full of various "signals" providing those rewards, if the agent is allowed to look for them, Silver and Sutton suggest.

"Where do rewards come from, if not from human data? Once agents become connected to the world through rich action and observation spaces, there will be no shortage of grounded signals to provide a basis for reward. In fact, the world abounds with quantities such as cost, error rates, hunger, productivity, health metrics, climate metrics, profit, sales, exam results, success, visits, yields, stocks, likes, income, pleasure/pain, economic indicators, accuracy, power, distance, speed, efficiency, or energy consumption. In addition, there are innumerable additional signals arising from the occurrence of specific events, or from features derived from raw sequences of observations and actions."

To start the AI agent from a foundation, AI developers might use a "world model" simulation. The world model lets an AI model make predictions, test those predictions in the real world, and then use the reward signals to make the model more realistic.

"As the agent continues to interact with the world throughout its stream of experience, its dynamics model is continually updated to correct any errors in its predictions," they write.

## **Also: [AI isn't hitting a wall, it's just getting too smart for benchmarks, says Anthropic](#)**

Silver and Sutton still expect humans to have a role in defining goals, for which the signals and rewards serve to steer the agent. For example, a user might specify a broad goal such as 'improve my fitness', and the reward function might return a function of the user's heart rate, sleep duration, and steps taken. Or the user might specify a goal of 'help me learn Spanish', and the reward function could return the user's Spanish exam results.

The human feedback becomes "the top-level goal" that all else serves.

The researchers write that AI agents with those long-range capabilities would be better as AI assistants. They could track a person's sleep and diet over months or years, providing health advice not limited to recent trends. Such agents could also be educational assistants tracking students over a long timeframe.

"A science agent could pursue ambitious goals, such as discovering a new material or reducing carbon dioxide," they offer. "Such an agent could analyse real-world observations over an extended period, developing and running simulations, and suggesting real-world experiments or interventions."

## **Also: ['Humanity's Last Exam' benchmark is stumping top AI models - can you do any better?](#)**

The researchers suggest that the arrival of "thinking" or "reasoning" AI models, such as [Gemini](#), [DeepSeek's R1](#), and [OpenAI's o1](#), may be surpassed by experience agents. The problem with reasoning agents is that they "imitate" human language when they produce verbose output about steps to an answer, and human thought can be limited by its embedded



assumptions.

"For example, if an agent had been trained to reason using human thoughts and expert answers from 5,000 years ago, it may have reasoned about a physical problem in terms of animism," they offer. "1,000 years ago, it may have reasoned in theistic terms; 300 years ago, it may have reasoned in terms of Newtonian mechanics; and 50 years ago, in terms of quantum mechanics."

The researchers write that such agents "will unlock unprecedented capabilities," leading to "a future profoundly different from anything we have seen before."

However, they suggest there are also many, many risks. These risks are not just focused on AI agents making human labor obsolete, although they note that job loss is a risk. Agents that "can autonomously interact with the world over extended periods of time to achieve long-term goals," they write, raise the prospect of humans having fewer opportunities to "intervene and mediate the agent's actions."

On the positive side, they suggest, an agent that can adapt, as opposed to today's fixed AI models, "could recognise when its behaviour is triggering human concern, dissatisfaction, or distress, and adaptively modify its behaviour to avoid these negative consequences."

**Also: [Google claims Gemma 3 reaches 98% of DeepSeek's accuracy - using only one GPU](#)**

Leaving aside the details, Silver and Sutton are confident the streams experience will generate so much more information about the world that it will dwarf all the Wikipedia and Reddit data used to train today's AI. Stream-based agents may even move past human intelligence, alluding to the arrival



of [artificial general intelligence](#), or super-intelligence.

"Experiential data will eclipse the scale and quality of human-generated data," the researchers write. "This paradigm shift, accompanied by algorithmic advancements in RL [reinforcement learning], will unlock in many domains new capabilities that surpass those possessed by any human."

Silver also explored the subject in a DeepMind [podcast this month](#).

## **Artificial Intelligence**