

Science & technology

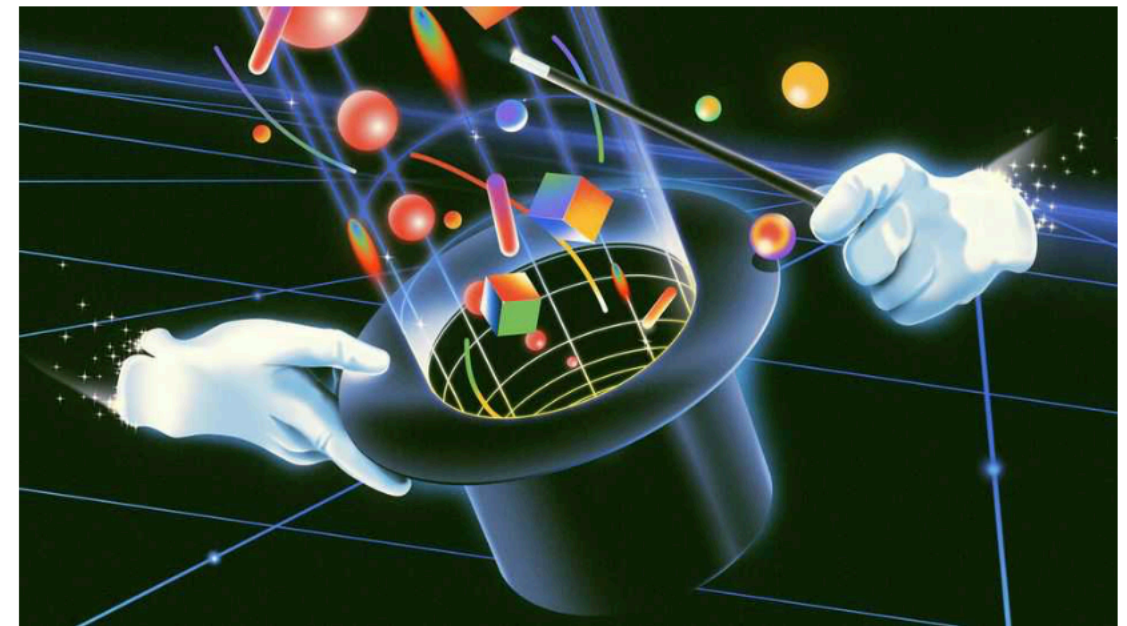
- AI tools are being prepared for the physical world
- Will one-stop blood tests for cancer save lives?
- Marks left by Stone Age humans were surprisingly complex
- Should you be fibremaxxing?

Science & technology | Reality beckons

AI tools are being prepared for the physical world

The race to build world models is on

February 26th 2026



PROJECT GENIE, an experimental artificial-intelligence model released by Google in January, is a jaw-dropping technical achievement. Give the tool a prompt—an

image, say, or a brief snippet of text—and it will generate an interactive world for the user to explore. Type in a straightforward request, and the result is a realistic simulation. Start with a painting by Georges Seurat, by contrast, and you can wander through a Sunday in the park in perfect pointillist style.

Project Genie may feel like a video game, but its makers claim it is something much more profound. They call it a “world model”, an essential tool to help AI systems make sense of the complex, unpredictable physical spaces into which many will eventually be put to work. The company argues that a future where humanoid robots pop to the shops to pick up ingredients before cooking dinner, or self-driving cars navigate country roads, would not be possible without world models.

The concept dates back to a 1943 book by Kenneth Craik, a Scottish psychologist who

suggested that organisms carried a “small-scale model” of the world inside their head, to test hypotheses on before carrying them out in reality. Having some grasp of how the world works is a necessary step before making plans about how to change it. Without one, any living thing would be forced into a purely reactive life—flinching from pain, reaching for food, and little more.

Giving that same ability to AI systems was a promising area of research as far back as the 1990s, before large language models (LLMs) sucked away the world’s attention. Now that attention is back.

There are three main approaches being explored to build world models. One natural starting-point is AI video generators. Generating a coherent video depends on simulating a coherent world—if the laws of reality change between frames, the output would be nonsensical. Such rudimentary

world models can fill in details of the world beyond what they have been fed: give one a picture of a maze and it will be able to draw a route through it; present it with a photo of hands holding a jar and it will accurately model the movements required to open it.

Project Genie is the culmination of this approach. Its usefulness becomes apparent when one imagines pairing it with a different AI—a robotic shopkeeper, say—that is trying to learn how to operate in the physical world. The billions of hours of training data essential for such a task would be much harder to obtain from the real world than from a model that can simulate the environment. And, if the simulations are accurate enough, the system can use the data to train itself.

But even the most realistic video of the world cannot capture every detail that a person would pick up on. The broken freezer at the back of the shop causing the

fresh fish to rot is not caught on camera, for example, nor is the associated smell. Even objects that are not directly visible are beyond it. Generate the contents of one aisle, for example, and the neighbouring ones do not exist for the model until the user enters them. That makes it harder to simulate complex environments, or let multiple users move in the same model.

Another approach to building world models, therefore, seeks to create full 3D environments rather than 2D simulations. Fei Fei Li, a computer scientist at Stanford University, is leading an approach she calls spatial intelligence. In her view, world models must be interactive, multimodal (capable of interpreting prompts) and consistent. Video-based systems can clear the first two hurdles but balk at the third. Project Genie, for instance, runs for a maximum of 60 seconds before its simulations start fraying at the edges.

Dr Li's startup, World Labs, has built a world model called Marble that can create digital versions of 3D worlds which are internally consistent and complete. That means it is possible to, for instance, have several users inside the same world. What's more, spaces are not hallucinated afresh each time the user looks around; instead, they are created in their entirety from the off. World Labs is pitching its product to architects, who could use it to dream up a space and explore it virtually before sending it to a 3D printer.

Yann LeCun, Meta's former chief AI scientist, thinks world models can be built in a different, less literal, way. To him, focusing on real spaces is a distraction. After all, many AIs will have to navigate virtual mazes such as HR systems or legal documents rather than physical spaces such as shops. He believes that giving AIs the tools to consistently model environments of both kinds is an important step towards

making them useful. In his view, an AI could use an LLM to interact with such a world model in order to help it carry out tasks, whether in the real world or on a computer.

That approach, called a Joint-Embedding Predictive Architecture (JEPA), would allow an AI to simulate complex features of the real world. Existing world models focus on what is just about to happen, rather than events that might (or might not) happen in the distant future. Humans think ahead all the time: gauging the weather before deciding whether to leave the house with an umbrella; factoring in the risks of being late for an important meeting when choosing which train to catch; and so on. Crucially, these decisions can be made quickly, without needing to visualise every single second of the day. Current world models have no such shortcut.

Dr LeCun has been exploring the potential of a JEPA system since 2022, and in November 2025 he left Meta to work on this problem full time. His startup, Advanced Machine Intelligence, plans to turn his ideas into reality, starting with a partnership with Nabla, a health-tech startup. He says the goal is a system which uses its own world model to work out “what sequence of actions will optimally accomplish a task that I’m setting”.

But what if these complicated approaches are superfluous? If existing generative AI systems can already do useful things in the real world, then maybe they already contain some kind of world model within them. That’s the view of Ilya Sutskever, an OpenAI cofounder, and many of his former colleagues still at the lab. Training a large language model is, he said in 2023, no more than “learning a world model”. Compressing all the information contained on the internet down into a few hundred gigabytes

of numbers is possible only if a system “learns” the underlying principles behind that information.

There is some evidence he may be right. In 2023 a language model trained on a list of moves in the game Othello was shown to have reflected the board state within its own neural network—even though it had never seen an Othello board nor been taught the rules of the game. It was a detailed enough representation that the researchers could identify specific parts of the neural network that stored the colour of individual pieces. That meant they could make specific tweaks to change its perception of the game, an unprecedented level of control over an LLM’s calculations.

Bigger language models are likely to have more complex world models inside—if only researchers could find them. Anthropic, an AI lab, has been leading research into “interpretability” of its Claude models,

finding clusters of artificial neurons that correspond to anything from feelings of guilt to the Golden Gate bridge. And reaching in and changing them, as in the Othello example, causes corresponding changes to the subsequent behaviour of those models. That suggests the systems aren't simply stringing words together: they have a consistent understanding of physical features in the real world, which they draw on to answer questions. It sounds suspiciously like what you would expect from an internal world model.

Not everyone agrees. LLMs, Dr Li argues, are just “wordsmiths in the dark”. Being able to use language to describe the world, she says, does not mean they have a grounded understanding of it. Like a student who has only read about a foreign country, there's a missing piece of knowledge that can't be patched with books, she says. Whichever approach will prove most

effective, there is little doubt that AI is about to pay the real world a visit. ■

Curious about the world? To enjoy our mind-expanding science coverage, sign up to [Simply Science](#), our weekly subscriber-only newsletter.

This article was downloaded by [zlibrary](#) from <https://www.economist.com/science-and-technology/2026/02/25/ai-tools-are-being-prepared-for-the-physical-world>