What is Genie 4? Google DeepMind's real-time world model with native audio

21 November 2026 · 7 min read · blogs.iictp.org/what-is-genie-4-world-model

Short answer: Genie 4 is Google DeepMind's world model, a model that generates an interactive environment frame by frame in response to your controls and prompts. Announced as a research preview on 27 October 2026, it opened to eligible Google AI Ultra subscribers in Project Genie on 18 November and to developers through the Gemini API and Google AI Studio on 19 November. What is new, according to Google DeepMind, is native audio generated with the world, including character dialogue, plus mid-session text edits steered by a Gemini model in the loop.

TL;DR The short version

- Genie 4 is Google DeepMind's world model, generating an interactive environment frame by frame from your controls and prompts.
- Google DeepMind says it is its first real-time world model with native audio, including dialogue, and it is available to developers through the Gemini API and AI Studio.
- Coherence holds over short sessions but details can drift, so treat it as a prototyping and agent-testing tool, not a physics simulator.

Read online, listen to it, or change the text size: https://blogs.iictp.org/what-is-genie-4-world-model

What is Genie 4?

A world model predicts what an environment looks like next given what has happened and what the user does. Unlike a video model that renders a fixed clip, it runs interactively: you move with keyboard controls, and the model generates the following frames in real time. Google DeepMind describes Genie 4 as its most realistic and interactive world model so far, and positions world models as useful for training environments for robots and agents as well as for interactive experiences.

How it works

- **Inputs.** A text prompt, a reference image, or both, per the Gemini API announcement. Users then navigate with standard keyboard controls and steer with live text prompts.
- **Output.** Continuous 720p worlds at 24 frames per second with synchronised native audio, from ambient sound and effects to character dialogue, according to Google.
- **Gemini in the loop.** Google DeepMind says Genie 4 builds on the Gemini family: Gemini Omni anchors appearance and behaviour in the real world, and Gemini 3.8 Flash acts as a guiding agent that observes the scene, breaks goals into steps, checks incoming frames to verify each step, and adapts the plan when the user intervenes.

- **Prompt adaptation.** When you type a change such as new weather or a new character, Gemini adapts the prompt to the surrounding scene, and between inputs it monitors and steers the scene to keep it coherent.

Google has not published the architecture, training data or model size in the posts we read, so this section describes the system's behaviour, not its internals.

What is new compared with Genie 3

- **Native audio.** Google DeepMind calls Genie 4 its first real-time world model to generate audio natively, including interactive dialogue.
- **Realism.** Google claims higher-fidelity visuals, physics (aerodynamics and collisions are given as examples) and more natural human motion.
- **Mid-session editing.** Typing a new prompt during a session changes the running world.
- **Genie 3 is still offered** in Project Genie for longer runs: up to 3 minutes per generation (up from 60 seconds), world chaining, and seeding from 360-degree panorama images.

Limitations reported by Google DeepMind

- Coherence holds over short sessions, but visual details and object permanence can drift as a simulation runs longer, after several complex prompt edits, or while navigating the scene.
- Complex physical interactions, real-time audio accuracy, legible text, spatial accuracy and long-term consistency are listed as open research problems. The availability post adds camera latency.
- All generated video and audio carries SynthID watermarking, and inputs and outputs go through safety screening, per Google.

Our reading: for agent or robotics training, drift and imprecise physics matter more than for entertainment. Treat Genie 4 worlds as a source of varied scenarios to test against, not as a physics simulator whose outcomes you can trust.

How to try Genie 4

1. Consumers: eligible Google AI Ultra subscribers can use Genie 4 in Project Genie at labs.google/projectgenie; Google AI Pro members can join a waitlist on the same page.
2. Developers: Genie 4 is available through the Gemini API, via the Interactions API, and in Google AI Studio, which has an interactive canvas and starter apps.
3. Check the Gemini API documentation and pricing page for the current model ID, limits and costs before building, as they were not stated in the announcement.

Project Genie also adds a Project Genie agent that builds worlds from a conversation and a community gallery where worlds can be published and remixed.

What it means for engineers and technical teams

- Prototyping: spatial concepts and interactive scenes can be sketched from a prompt or a single image without a game engine.

- Agent testing: a promptable environment lets you generate many scenario variations, but outcomes are generated, not simulated from physical laws, so validate anything that matters in a real simulator or the real world.
- Provenance: SynthID watermarks are applied to outputs; if you publish generated media, keep your own record of what was generated. See responsible AI in practice.

Frequently asked questions

- **What is Genie 4?** A real-time, interactive world model from Google DeepMind that generates explorable environments with native audio from text or image prompts.
- **Can developers use Genie 4?** Yes. Google says it is available through the Gemini API (Interactions API) and Google AI Studio since 19 November 2026.
- **What resolution does Genie 4 generate?** Continuous 720p at 24fps, according to Google's developer announcement.
- **Is Genie 4 a physics engine?** No. It generates plausible next frames; Google lists complex physical interactions and long-term consistency as open problems.

Related reading

- Multimodal AI explained
- AI agents explained
- Responsible AI in practice
- AI Practitioner certification

Sources were read on 21 November 2026. Preview APIs, model IDs, prices and access rules change often, so check the linked documentation before you build on them.

> Related credential: AI Practitioner. Explore the exam facts and syllabus.

Sources and further reading

1. Google DeepMind (Hadsell, R., Rivas, D., Qi, H.). Genie 4: A new generation of world simulation. 2026-10-27. deepmind.google/blog/genie-4-a-new-generation-of-world-simulation
2. Google (Rivas, D., Kumar, P.). Genie 4 is now available. 2026-11-18. blog.google/innovation-and-ai/models-and-research/google-deepmind/genie-4-now-available-project-genie
3. Google. Genie 4 is now available on the Gemini API. 2026-11-19. blog.google/innovation-and-ai/technology/developers-tools/genie-4-in-gemini-api

Published by IICTP, an independent technical credential body, on 21 November 2026. The current version of this post is at https://blogs.iictp.org/what-is-genie-4-world-model. More guides: blogs.iictp.org.