

World Models and Scientist AI Rise as Claude and Microsoft Push Scale

AI News Digest

2026-03-14

World Models and Scientist AI Rise as Claude and Microsoft Push Scale

By AI News Digest • March 14, 2026

Leading AI researchers sharpened the debate over what comes after today’s LLMs, with Yann LeCun pushing world models, Yoshua Bengio arguing for “scientist AI,” and Geoffrey Hinton and Gary Marcus warning that governance is lagging. At the same time, Anthropic expanded Claude’s context window, Microsoft advanced next-generation AI infrastructure, and Sakana AI showed more ambitious research automation.

The clearest signal today: leading researchers are arguing about what should come after today’s LLMs

The biggest theme was not a single model release, but a widening debate among top AI researchers about what kind of systems should come next—and how urgently governance needs to catch up [1, 2, 3, 4].

LeCun lays out a world-model agenda through AMI Labs

Yann LeCun said he has left Meta and is building Paris-based **AMI Labs** around **Advanced Machine Intelligence**, arguing that the next major leap will come from systems that understand the real world through **hierarchical world models**, not from scaling LLMs alone [1]. He pointed to **JEPA** and **Video JEPA** as core building blocks, saying recent self-supervised methods can surpass fully supervised systems and that Video JEPA has shown early signs of learned “intuitive physics” [1].

Why it matters: This is a concrete post-LLM research and company-building agenda from one of the field’s most influential researchers [1].

Bengio pairs “scientist AI” with a governance push

Yoshua Bengio said his nonprofit **Law Zero** is building a “**scientist AI**”: systems designed for understanding rather than hidden goals, with the aim of making them trustworthy enough to veto unsafe actions from other AI systems [2]. He said Canada is supporting the effort with funding, people, and compute, while he separately warned—through his work on the **International AI Safety Report**—that current harms already include deepfakes and fraud, with frontier risks extending to cyberattacks, bioweapons misuse, misalignment, and loss of control [2, 5].

“The ideal is pure intelligence without any goals.” [2]

Why it matters: Bengio is making a two-part case at once: safer AI likely needs different training objectives, and the institutions around AI need to move faster too [2, 5].

Hinton and Marcus, from different angles, say the governance window is still open—but narrowing

Geoffrey Hinton said AI may surpass human intelligence soon, but stressed that humans still have agency because “we’re still making them” and can still change how these systems are built [3]. Gary Marcus argued that current LLMs remain unreliable enough to threaten democracy through misinformation and deepfakes, and called for global governance, AI-generated-content labeling, public literacy, and better detection tools [6, 4].

Why it matters: Even across researchers who disagree on technical direction, there is growing overlap on one point: capability progress is outrunning verification and governance [3, 6, 4].

Frontier products and infrastructure kept stretching the frontier

Anthropic makes 1M context mainstream in Claude 4.6

Anthropic made the **1 million token context window** generally available for **Claude Opus 4.6** and **Claude Sonnet 4.6** [7, 8]. The company also removed the API long-context price increase, dropped the beta-header requirement, made **Opus 4.6 1M** the default for Claude Code users on Max, Team, and Enterprise plans, and now supports up to **600 images** in one request [8].

Why it matters: This is not just a bigger number on a benchmark card; Anthropic is trying to make extreme context cheaper and more normal in everyday developer use [7, 8].

Microsoft brings NVIDIA’s Vera Rubin NVL72 into cloud validation

Microsoft said it is the **first cloud** to bring up an **NVIDIA Vera Rubin NVL72** system for validation, a step toward next-generation AI infrastructure [9]. In separate remarks, Satya Nadella described the AI data-center buildout as a **“token factory”** whose job is to turn capital spending into return on invested capital [10].

“The token factory is all about turning – through software – capital spend into ROIC. That’s the job.” [10]

Why it matters: The competitive frontier is still being fought on supply, utilization, and economics—not only on model quality [9, 10].

Research tools are moving from assistants toward discovery systems

Sakana AI pushes evolutionary search toward automated science

In a detailed discussion of **Shinka Evolve**, Sakana AI described an open-source system that uses LLMs to mutate, rewrite, and evaluate programs with a more sample-efficient evolutionary search process, including model ensembling and bandit-style selection across frontier models [11]. The speaker said it improved on the circle-packing result shown in the AlphaEvolve paper with very few evaluations, would have ranked second on one ALE Bench programming task, and that **AI Scientist V2** has already reached the point of generating workshop-level papers by shifting from linear experiment plans to agentic tree search [11].

Why it matters: The research frontier is inching away from AI as a coding copilot and toward AI as an iterative search-and-experiment engine [11].

Bottom line

Today’s mix of commentary, launches, and research points to two races running in parallel: one toward more scale, longer context, and heavier infrastructure, and another toward AI that is more grounded, causal, and governable [7, 9, 1, 2].

Sources

1. “Perspectives on IA” : conf. de Yann LeCun, WinterWeek – Graduate School – Univ. Gustave Eiffel
2. Fireside Chat with Yoshua Bengio – AI Safety, Governance & the Future of AI | IASEAI ’26
3. A chilling warning by AI pioneer, Geoffrey Hinton.
4. VIRTUAL HEADLINER: In Conversation with Gary Marcus

5. AI safety and democratic governance of powerful AI systems - Yoshua Bengio
6. Discussing the implausibility that AI will approach human intelligence by the end of 2025
7. X post by @claudeai
8. X post by @alexalbert___
9. X post by @satyanadella
10. X post by @sequoia
11. Solving the Wrong Problem Works Better - Robert Lange