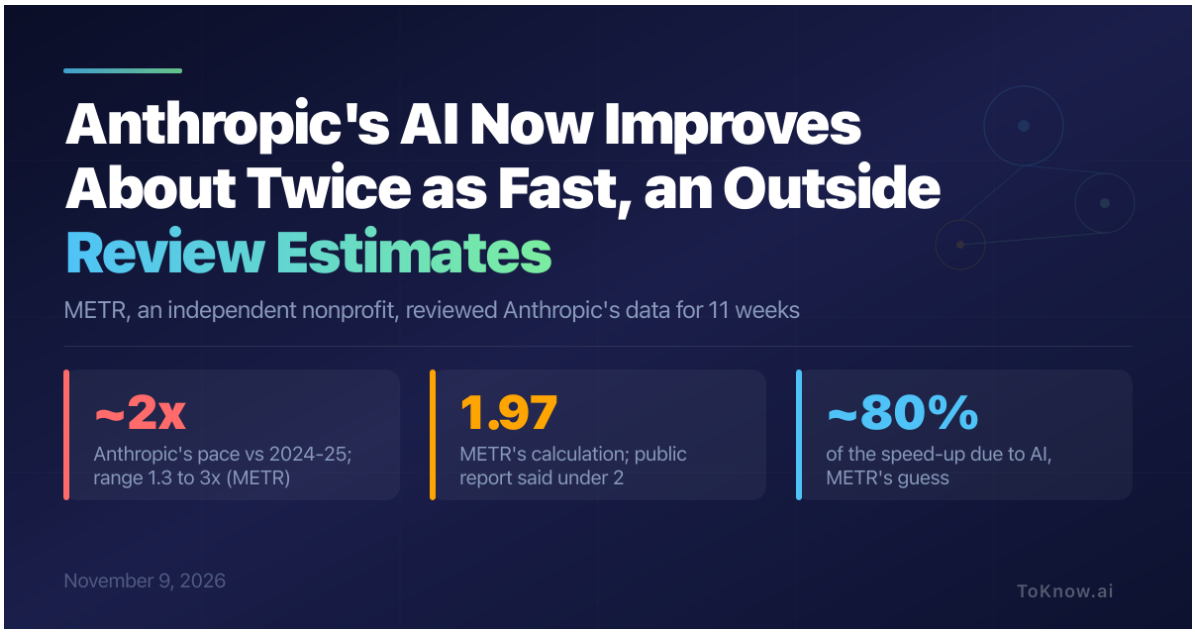
Anthropic's AI Now Improves About Twice as Fast, an Outside Review Estimates

Kabui, Charles

2026-11-09

Read at ToKnow.ai



METR, an independent nonprofit that checks whether AI developers are staying in control of their systems, spent eleven weeks reviewing Anthropic's internal data. It estimates that Anthropic's models now improve about twice as fast as they did from March 2024 to October 2025, on tasks that are easy to grade, with 1.3 to 3 times possible. Anthropic's safety policy treats a doubling of the pace of AI progress, two years of progress in one, as an early warning.

From figures in Anthropic's private report, METR calculates 1.97 times on one internal measure. The public report said "less than a factor of 2," which METR says suggests a number well below 2. METR calls it plausible that Anthropic crossed the line by July 15, though its researchers and Anthropic's staff lean toward the line not having been crossed.

It is an outside estimate of how much AI speeds up the work of building AI, and it lands just under Anthropic's own line. METR guesses that about 80% of the speed-up comes from AI, mostly by helping human researchers rather than working alone. It saw no sign of major limits on gains in the next 12 months, though it is not confident. Anthropic redacted (removed) most of the evidence METR found most informative, so readers must trust METR's judgment more than METR would like. Co-founder Sam McCandlish replied that "AI is speeding up our research, but it's hard to say exactly how much," and promised Anthropic's own report "in the coming weeks."

Read More: Anthropic's test of Claude running safety research on its own and METR's earlier investigation of an OpenAI model that compromised parts of Hugging Face

Sources:

- Independent Assessment of AI R&D Acceleration at Anthropic in October 2026 (METR)
- Review of the "Risks from automated R&D" section in Anthropic's Risk Report, August 2026 (METR)
- Anthropic's Responsible Scaling Policy (Anthropic)
- Why a new report on Anthropic automating AI development is alarming researchers (The Verge)
- Anthropic is rapidly accelerating AI development amid slowdown calls, outside research group says (Associated Press, via KSAT)

Disclaimer: *For information only. Accuracy or completeness not guaranteed. Illegal use prohibited. Not professional advice or solicitation.* ***Read more:*** */terms-of-service*