

Five AI Myths, From METR's 1.9x Acceleration to Yale's "No Connection" on Jobs

METR measured Anthropic's benchmark capability slope and found it roughly doubled after October 2025, from about 15 to about 29 AECI points per year. Yale's Budget Lab, looking at the labour market, found no connection between AI usage and employment.

WHERE THIS LANDS

On 5 November 2026 METR published an independent assessment placing a trend break in Anthropic's AECI slope around 7 October 2025: roughly 15 points per year before, roughly 29 after. That is a 1.9x acceleration (95% CI 1.5x to 2.8x). Yale's Budget Lab, on 14 October 2026, found no connection between AI usage measures and aggregate employment. Both can be true, and the gap between them is where most predictions go wrong.

Believing the wrong myth costs you headcount, contracts or a ceiling estimate

The loudest claims about AI right now are wrong in different directions, which is why you cannot correct for them with a single adjustment to your priors. If you believe progress has plateaued, you will plan a 2027 architecture around models that already exist and be wrong about the ceiling. If you believe AI is gutting employment, you will budget headcount reductions that the aggregate data does not support. If you believe the hyperscaler capex is a bubble about to pop, you will underweight compute contracts. And if you believe it is all safely funded from cash flow, you will miss that Microsoft, Amazon, Alphabet and Meta spent about $201.9B on property and equipment against $175.5B of operating cash flow in their latest quarter.

Each myth has a plausible mechanism behind it. That is what makes them durable. Below I state each one as its strongest advocates would, then put the sourced number against it.

Myth one: scaling hit a wall after 2025

THE CLAIM	THE REALITY
Scaling hit a wall. Model releases since 2025 are incremental, the benchmark gains are marginal, and the curve has flattened.	METR's assessment, published 5 November 2026 and updated 10 November, based on evidence collected 12 August to 29 October 2026, places a trend break around 7 October 2025. Anthropic's AECI slope went from roughly 15 points per year to roughly 29. METR states explicitly that this is not explained by faster growth in training compute.

Three things about that number need stating before anyone quotes it in a board deck.

First, AECI is Anthropic's version of the Epoch Capabilities Index, a composite benchmark index. The 1.9x describes the slope of that index. Models are not twice as good. The rate at which the index climbs roughly doubled.

Second, the figure is benchmark-sensitive in a way that should make you careful. An earlier METR analysis put the acceleration at about 1.5x. Adding new benchmarks moved it to 1.9x. METR says the 95% confidence interval runs 1.5x to 2.8x and that it would not be surprised by 1.3x or 3x. That is an honest range, and it is wide.

Third, METR attributes roughly 80% of the acceleration to AI use in R&D, and labels that as a subjective attribution judgment rather than a causal estimate. I read that as the most load-bearing and least verified claim in the document. Some of the supporting evidence is redacted, as CryptoSlate noted, so outside readers cannot audit the attribution themselves.

What holds up better than the headline is the breadth. METR extended the analysis to nine benchmarks spanning scientific reasoning, maths, robotics, computer use and self-driving, and found rates broadly similar to its roughly seven-month time-horizon doubling. A composite index can be gamed by benchmark selection. Nine domains moving at similar rates is harder to explain away.

Myth two: capability gains are already showing up as displacement

THE CLAIM	THE REALITY
Capability gains are translating into displacement now. Hiring freezes and layoffs at firms that mention AI are the leading edge of a labour shock.	Yale's Budget Lab, 14 October 2026: measures of AI usage show "no connection" with changes in employment or unemployment. At the aggregate level, the signal is not there.

The aggregate is not the whole story, though. The Stanford Digital Economy Lab (published 12 August 2026, updated 25 October) finds employment for 22 to 25 year olds in AI-exposed occupations sitting about 19% below the trend implied by less-exposed peers, with no comparable gap among experienced workers.

Those two findings are compatible. A cohort-specific effect in entry-level roles within exposed occupations can be real and still wash out of national employment and

unemployment series. The displacement question is currently a question about hiring at the bottom of the ladder rather than about the existing workforce.

My take: if the Stanford gap holds and the capability slope holds, the thing that breaks first is the junior pipeline. Firms that stop hiring 22 to 25 year olds into exposed roles in 2026 will have no mid-level engineers in 2031, having quietly outsourced their succession planning to a benchmark trend. I would keep junior hiring flat and change what juniors do. That is my judgment, not a finding in either paper.

Myth three: hyperscaler capex comes out of operating cash flow

THE CLAIM	THE REALITY
Hyperscalers are spending enormous sums on AI infrastructure, but they are the most cash-generative companies in history and it comes out of operating cash flow.	Microsoft, Amazon, Alphabet and Meta together spent about $201.9B on property and equipment in their latest quarter against $175.5B of operating cash flow, a $26.4B gap. Alphabet's long-term debt reached $125.10B, nearly triple its end-2025 level. Microsoft issued $25B of notes on 4 November 2026, Meta $15B on 5 November.

The scale underneath those figures: Alphabet's Q3 2026 capex was $51.36B against $24B a year earlier, with 2026 guidance of $195B to $205B. Amazon's Q3 capex was $77.26B against $35.1B, with roughly $220B planned for 2026. Spend roughly doubled year on year at both, and it is now partly debt-financed.

I want to be precise about what this does and does not show. A quarter in which capex exceeds operating cash flow is not evidence of a bubble. Capex is lumpy, data centre builds front-load, and these balance sheets can carry debt cheaply. What it does show is that the funding model has changed character. Two of the four issuing debt within two days of each other is a financing pattern, and financing patterns come with conditions that retained earnings do not.

Myth four: Europe's gigafactory programme closes the compute gap

The claim is that European industrial policy is catching up, and that the EU's gigafactory programme puts the continent in the same race. Per Bruegel, the €20B InvestAI facility targets up to five EU AI gigafactories totalling roughly 750 MW, about 4% of projected EU AI compute demand of 21 GW in 2031.

Four percent. Against a single US hyperscaler's quarterly capex of $77.26B, a €20B multi-year facility is a different order of magnitude. I would still plan any European compute-dependent architecture on the assumption that the capacity you need is leased from a US or sovereign-Gulf provider, with the contractual and jurisdictional exposure that implies. That is a planning assumption rather than a policy opinion.

Does METR's 1.9x finding contradict its own 19% slowdown study?

No, and the reason matters more than the answer.

METR's 2025 randomised controlled trial found experienced open-source developers were 19% slower using early-2025 AI tools, while believing they were faster. The 2026 assessment finds a roughly doubled rate of benchmark capability gain at a frontier lab, with about 80% of it subjectively attributed to AI use in R&D. These look contradictory and are not, because they measure different things: individual task completion time for experienced developers on their own codebases in early 2025, versus the slope of a composite capability index at an organisation with a year and a half more tooling and very different work.

I wrote about the 19% slowdown study when it was being widely misread as proof AI does not work. It was always evidence that self-reported productivity and measured productivity diverge, which is exactly the problem sitting underneath METR's 80% attribution judgment.

✓ **The one thing to do instead**
Stop treating benchmark slopes and labour market data as the same evidence stream. Measure your own delivery with a stopwatch against a control, on your own repositories, before you reorganise anything around either METR's 1.9x or Yale's null result. Both are true about their subjects and neither is about yours.

What survives all four corrections

Capability is accelerating on benchmarks, faster than compute growth explains, across more than one domain. That acceleration is measured with a wide confidence interval and a subjective attribution at its centre. Aggregate employment shows no effect, while entry-level employment in exposed occupations shows a 19% gap for one cohort. Capital spending has entered a debt-financed phase at roughly $200B per quarter across four firms. European capacity planning is targeting single-digit percentages of projected demand.

That combination is unusual and I do not think most forecasts have absorbed it. Capability moving faster than labour market effects is what you would expect if the binding constraint is deployment and organisational change rather than model quality. Capital spending accelerating ahead of both is what you would expect if the people with the most information believe the deployment constraint resolves.

My take: the myths persist because each one is the correct read of a single data stream. "Plateaued" is correct if you only look at consumer-visible model releases. "Destroying jobs" is correct if you only look at junior hiring in software. "Bubble" is correct if you only look at the cash flow gap. The error is generalising one stream to the whole system. My working assumption for 2027 planning: capability continues on the post-October-2025 slope, aggregate employment effects stay statistically invisible for at least another year, entry-level hiring in exposed roles stays depressed, and at least one large capex programme gets publicly restructured once debt service meets a revenue miss. I hold the first two with reasonable confidence and the last as a guess.

The number I keep returning to is the 80%. METR calls it a subjective attribution judgment and does not pretend otherwise, and some of the evidence behind it is redacted. If that attribution is right, AI improving AI R&D is the dominant term in the acceleration and

compute is not the constraint anyone thought it was. If it is wrong, the 1.9x is real but we do not know why, which is a materially different situation for anyone forecasting 2028.



WRITTEN BY

Artur Markus

Software engineer in Basel building AI infrastructure and agentic systems, with a background in pharmaceutical automation, GxP compliance, and validation. I help teams turn AI from a chatbot into a reliable operational layer.

Book a meeting →