More capable work. A harder test of progress.

Review: 21 October 2026, Sydney | Latest verified publication: 19 October 2026

By 31 December 2031, capable organisations are likely to commission substantial digital projects with far less direct human execution. The strongest new evidence broadens the work AI can attempt; the pace of useful adoption depends on whether outputs can be checked, actions controlled and capacity financed.

The 2031 direction is retained. This week strengthens selected mechanisms; it does not establish general autonomy or justify a new broad-intelligence arrival date.

Five clocks, five different questions

CLOCK	CURRENT JUDGEMENT	WHAT WOULD MOVE IT
Capability	Stronger selected capability. Legacy-code work and a checkable research result broaden the case. [W01-W03]	Independent fresh-project completion at accepted quality.
Autonomy & reliability	Useful delegation, uneven reliability. New task gains do not establish dependable whole-project autonomy. [W08, W12]	Human intervention, repair and consequential failure rates.
AI research acceleration	Bounded gains supported. Whole-cycle acceleration and its attribution remain uncertain. [W04, W11]	Validated research output per total cost, with a counterfactual.
Economy & distribution	Selective restructuring. Hiring pressure, complementary demand and financing constraints coexist. [W05-W07]	Output, labour intensity, hiring and commissioned capacity.
Safety & control	More concrete scrutiny and restrictions. Net safety improvement remains unmeasured. [W08-W10]	Completed external assessments and verified remediation.

What changed in this edition

A result we could check. The seed recipe reproduced all five historical NIST constants. That validates the arithmetic, not the discovery process, runtime or general research competence. [W02-W03]

A more demanding labour story. The newly published IMF study supports relative hiring pressure, but also stabilisation and complementary demand. Its observations end in April 2026, before the latest models. [W05]

Access has its own trajectory. Finance, power connections and operational controls can slow or redirect deployment even while capability improves. Institutional responses can also remove constraints. [W06-W10]

Strongest challenge: these results may favour tasks with cheap, clear checks. If quality assurance and repair consume the savings in messier work, the central delegation forecast is too strong.

THE FORECAST / COMMITMENTS WE CAN REVISIT

From here to 2031

Through 2027	More software, analysis and administrative work becomes delegable. The commercial test is the full cost of a finished, acceptable outcome. Hiring can adjust before total employment.
2028-2029	Capable adopters reorganise delivery around smaller teams and combined roles. Research feedback is the major branch point; finance and power access affect who can scale.
2030-2031	Substantial digital delegation is ordinary among capable adopters. Ownership, market access, public capacity and bargaining power shape whether lower production costs improve lives.

Rolling five-year view: 21 October 2026-21 October 2031. The original 31 December 2031 endpoint and earlier deadlines remain intact.

Four prospective public-evidence tests

These are low-confidence subjective probabilities of qualifying public studies, not underlying capability. The flagship 2031 adoption claim remains qualitative and unscored; these four tests do not validate it. Full criteria are frozen in the pack.

QUALIFYING PUBLIC EVIDENCE	BY	CHANCE
Halved human input for accepted digital projects across three organisations.	2027	**40%**
Two independent studies: 1.5x research results per total cost, including question selection.	2027	**20%**
Two independently audited research programmes: 2x output per cost over six months.	2029	**25%**
Two-country longitudinal evidence: 20% lower labour intensity for experienced professionals.	2029	**50%**

No resolved scores yet. Score launch forecasts and a separate 90-day snapshot; compare with a disclosed 50% baseline and the unchanged launch estimate. Four correlated events cannot establish calibration. T05 and T06 remain unscored.

Numerical continuity and the next decisive test

The September AC / TED-AI / ASI medians (2029.6 / 2033.5 / 2037) remain historical records pending full re-elicitation, not fresh forecasts. The next decisive evidence is repeated completion of useful projects with less total human effort after failed attempts, checking and repair are counted.

Sources and evidence status

W01 OpenAI: coding in low-resource programming languages
W02 Anthropic: ECDSA curve origins
W03 NIST: original prime-curve seed constants
W04 METR: assessing AI-driven research acceleration
W05 IMF: labour markets in Asia and the Pacific
W06 IEA: Powering Artificial Intelligence
W07 IMF: financial stability and AI financing
W08 Anthropic: unintended actions in evaluations and internal use
W09 Anthropic: access for independent evaluators
W10 NIST, NSA and CISA: incident-readiness practices
W11 TasteVal: experimental research judgement
W12 METR: task-completion time horizons
W13 AI Futures Project: August timelines update
W14 IMF: World Economic Outlook, October 2026

Source inspection does not replicate provider experiments or establish causal effects. Full URLs, limits, protocols, revision history and the seed-check script accompany this edition.