More work can be handed to AI

By the end of 2031, I expect organisations that can integrate AI and check its work to commission substantial digital projects with much less direct human effort.

That means handing over a sizeable piece of digital work, with people setting the goals and checking the result. It is a forecast for capable adopters; the share of organisations that will get there is unmeasured.

What the evidence shows

Specialist work is becoming useful. Anthropic reports AI-assisted exploit development and contributions to early drug-development tasks. Google independently credits Claude-assisted researchers for particular security fixes. Overall autonomous success rates and complete costs remain unknown. [1] [2] [3]

Integration can produce large task savings. Across four pharmaceutical batch records, Enveda reports 166 minutes for AI processing plus checking its comments, against 1,553 minutes for manual review. This 9.4-fold difference concerns a specific review step using a custom system and Opus 4.7. Full human reviews remained outside the AI comparison. [4]

Control still needs independent testing. OpenAI says it resumed frontier reinforcement learning and internal inference on 1 October after strengthening containment. Its monitoring report describes wider coverage, with gaps in alerts and intervention. Net safety remains unknown. [5] [6]

Five clocks to watch

Capability, reliable autonomy, research, adoption and control can move at different speeds.

Clock	Current assessment	What would change it
Capability	Useful specialist work is supported by concrete cases. Transfer to unfamiliar tasks remains the test.	Independent tests of fresh tasks, counting every attempt and its cost.
Autonomy and reliability	Supervised workflows can be useful. Dependable delivery of complete projects remains unproven.	Finished projects with all human specification, checking, rescue and repair time counted.
Research acceleration	Selected applied-science work benefits. End-to-end gains in AI research remain unmeasured.	Externally validated research results per total cost, including the choice of questions.
Economy and distribution	Some tasks take less time. Broad effects on productivity, hiring and incomes remain uncertain.	Sustained evidence across organisations on quality, total labour, wages and demand.
Safety and control	Cyber capability and reported safeguards are advancing together. Net safety is unknown.	Independent tests of current containment, monitoring coverage and intervention.

Evidence and scope. Evidence assessed as of 2 November 2026, 08:50 AEDT. Latest verified publication in the assessment: 30 October 2026. Selective primary-source review, mainly provider and partner reports; no replication, proprietary-data access or new economy-wide estimate.

Full definitions and assessment rules accompany the website briefing.

What this could mean

Through 2027. More software, analysis and administrative work becomes delegable. Judge systems by finished, acceptable results and the full cost of checking and repair. Hiring can change before total employment does.

2028-2029. Organisations able to verify and integrate AI redesign work around smaller teams and combined roles. Research validation, finance, power and institutional capacity constrain how quickly this spreads.

2030-2031. Substantial digital delegation becomes ordinary among capable adopters. Platform owners may capture disproportionate returns. Service exporters face price pressure and new demand; lower-capacity states gain cheaper expertise while depending on foreign systems. Procurement, interoperability and bargaining power influence who benefits.

Forecast horizon: to 31 December 2031. The central judgement is qualitative; the four public-evidence tests do not directly measure its adoption rate.

Four tests of public evidence

These are low-confidence estimates of the chance that qualifying public evidence will be published by each deadline. They are not probabilities that the underlying capability exists. The four events are related, and remain unresolved.

Chance	Deadline	Evidence required
40%	End of 2027	**At least 50% less human input per accepted digital project.** One prospective controlled study: at least 30 fresh project briefs across three unrelated organisations, with quality preserved and failures counted.
20%	End of 2027	**At least 1.5 times as many validated AI-research results per total cost.** Two independent prospective controlled studies, including research-question selection, external validation and full cost accounting.
25%	End of 2029	**At least twice the validated AI-research results per total cost across complete research cycles.** Two unrelated research organisations publish independent audits of controlled programmes lasting at least six months.
50%	End of 2029	**At least 20% less human labour per accepted professional output.** Two independent studies in different countries follow experienced cohorts for at least twelve months, counting supervision, failed work and attrition.

The strongest challenge

Selected successes may favour tasks that are cheap to check. Custom engineering and expert judgement may account for much of the gain, while messy projects remain expensive to supervise. If checking and repair consume the savings, the case for broad delegation by 2031 weakens.

Practical implication. Budget for integration and checking. Measure the full cost of an accepted result, including failed attempts and human rescue. Track who receives the gains as well as what the tools can produce.

Sources

[1] Anthropic: Cybersecurity research (27 October 2026).
[2] Google: Chrome security fixes (20 October 2026).
[3] Anthropic: Drug-development case studies (27 October 2026).
[4] Enveda: Batch-record review methods and results (27 October 2026).
[5] OpenAI: Frontier training containment (23 October 2026).
[6] OpenAI: Monitoring and intervention limits (29 October 2026).