AI value depends on the cost of accepted work

Review, rework and deployment costs can absorb gains from faster drafting.

RESEARCH BRIEFING · TRANSFORMATION · 11 NOVEMBER 2026

Companies should assess AI against the cost of a completed, accepted outcome. Faster drafts create value only when review, correction, integration and recurring fees leave the full workflow better off.

Key points

• Measure cost per accepted outcome, including human review.

• Keep historical research separate from estimates for today’s tools.

• Name the operating action that converts saved time into a benefit.

The evidence points in different directions

In a 2023 experiment with 758 consultants, Harvard Business School reported speed gains of more than 25% on tasks suited to GPT-4. METR’s 2025 trial found a 19% slowdown across 246 repository tasks completed by 16 experienced developers. Different tasks, users and tools produced different results. [1][2]

METR warned in February 2026 that selection and measurement problems made its follow-up an unreliable estimate of current productivity. The early result is a historical finding, not a verdict on current coding tools. These studies should not be averaged into a company-wide return forecast. [3]

Worked example: count the review

Assume 4,000 cases a month take 12 minutes each. An AI-assisted process takes six minutes to prepare, three to review and one in average rework: 10 minutes in total. The two-minute reduction releases about 133 hours a month.

At an assumed loaded labour cost of €45 an hour, the released capacity is worth €6,000. Subtract €3,000 in monthly tooling and support: the illustrative balance is €3,000 before setup costs. If only half the capacity is used productively, that recurring benefit is zero.

All inputs are hypothetical. The calculation values capacity; it does not establish cash savings. Cash effects need an operating decision such as reduced overtime, avoided contractor spending or additional accepted output.

Run a bounded comparison

Choose one workflow with a measurable acceptance standard. Record volume, case difficulty, elapsed time, labour time and errors. Compare the existing process with AI-assisted work, including difficult and failed cases.

Report accepted-output cost, correction minutes, material-error rate and the share of released capacity put to use. Include setup, integration and monitoring in the payback calculation. Stop or redesign the pilot if review effort cancels the drafting gain.

Finance should approve the conversion plan

Agree the baseline and the benefit mechanism before procuring more licences. A time saving counted as reduced cost cannot also be counted as additional output from the same hours. The expansion decision needs both a quality result and a credible use for the released capacity.

Sources & research notes

01 / Harvard Business School AI Institute / Jagged Technological Frontier (2023)

https://aiinstitute.hbs.edu/navigating-the-jagged-technological-frontier/
758 consultants; experimental tasks and early GPT-4. Results are not firm-wide returns or estimates for current tools.

02 / METR / Experienced developer productivity study (10 July 2025)

https://metr.org/blog/2025-07-10-early-2025-ai-experienced-os-dev-study/
16 developers, 246 repository tasks; early-2025 tools and experienced maintainers.

03 / METR / Experiment design update (24 February 2026)

https://metr.org/blog/2026-02-24-uplift-update/
Selection and measurement problems make the follow-up an unreliable estimate of current productivity.

Independent TACITUS analysis. Research checked 11 November 2026. Published findings retain their dates and limits. Proposed curricula, tests, schedules and worked calculations are labelled analytical examples, not client results.