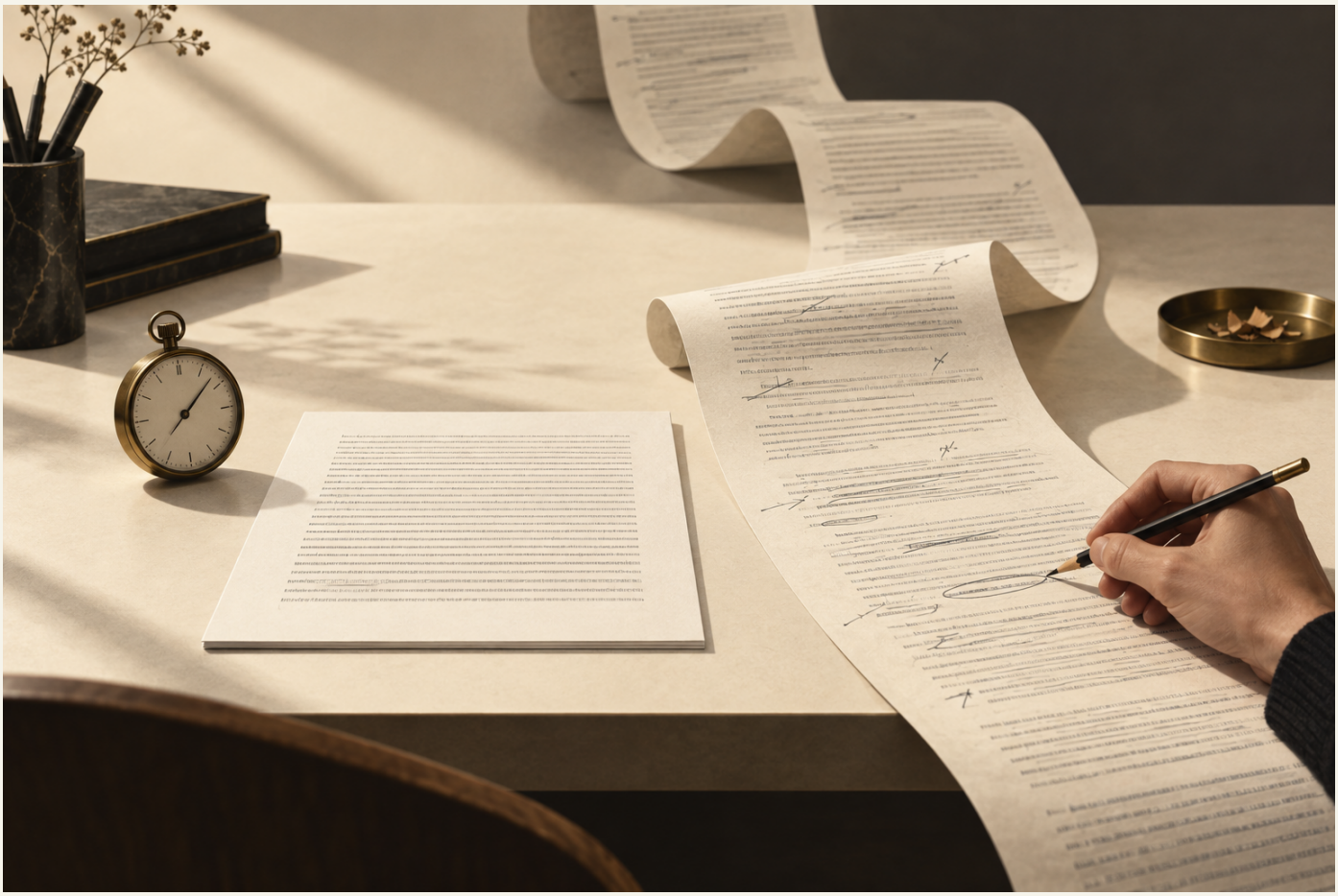
FIELD NOTE / MEASURING AI AT WORK

How to Tell When AI Is Actually Saving You Time

Generation speed is only useful when trust is cheap.

Don Long · Founder, SummitMark · Helping people and organizations become capable with AI



The generation clock stops before the trust work is finished.

A faster first draft is only a productivity gain if the whole task gets easier - including the work required to trust the result.

01 / START TO TRUSTWORTHY FINISH

The real clock starts after AI finishes.

AI can turn a 20-minute task into 30 seconds. Sometimes that is exactly as good as it sounds: the answer is right, the check is quick, and you move on.

Now change the task. The AI still takes 30 seconds, but reviewing takes eight minutes. You find a mistake, reopen the source, correct it, and check the rest because your confidence has fallen.

ILLUSTRATIVE EXAMPLE FROM THE ARTICLE

00:30
GENERATE

then

08:00+
VERIFY, CORRECT, RECHECK

The point is not to treat those example times as a benchmark. It is to count from the start of the task to a result someone can responsibly use.

How much human effort disappeared from start to trustworthy finish?

01 / HOW EXPENSIVE IS TRUST?

Formatting a document is easy to verify: you can look at it. A 200-page contract analysis that says "no meaningful risks found" is different. The output may arrive quickly; establishing whether it is safe to rely on may not.

02 / HOW EXPENSIVE IS BEING WRONG?

A weak headline suggestion is cheap to discard. A bad financial recommendation, customer decision, or employment judgment carries a different cost. As the consequence rises, preserve more human attention around the result.

02 / A PRACTICAL CLASSIFICATION

Green removes work. Yellow adds a cleanup tax.

GREEN

Trust is cheap, mistakes are cheap, and faster production creates little downstream burden. Reformatting information, extracting obvious fields, making a first-pass outline, or drafting routine internal copy can fit. If a 20-minute task takes one minute and needs 30 seconds to check, the time saved is plain. Use the AI.

YELLOW

The gain is real, but review and correction still take attention. In the article's example, a 30-minute task becomes a two-minute AI first pass and ten minutes of careful review: 12 minutes total, not a 28-minute saving. The remaining 18 minutes are the honest comparison.

Automation Cleanup Tax: the human attention required after AI has technically finished.

42%

of respondents' AI-use time went to troubleshooting errors and iterating on prompts.

35%

of respondents' AI-use time went to work they classified as directly productive.

BambooHR survey: 1,608 U.S. full-time salaried desk workers; online, self-reported responses collected June 26 - July 15, 2026. Vendor-sponsored research; these are shares of AI-use time, not independently observed time logs.

BambooHR, "Redesigning Work: AI's Performance Review" (Sept. 1, 2026).

03 / RED: THE BOTTLENECK MOVES

More output can mean more work.

Red tasks are hard to verify, costly when wrong, or likely to create a new queue. Keep the workflow human-led even when AI helps inside it. The scarce resource may become judgment about what deserves to survive.

META INTERNAL METRICS REPORTED BY REUTERS / YEAR OVER YEAR

+220%

code changes to internal platforms and infrastructure used by employees.

+36%

changes that resulted in new or upgraded features reaching Meta users.

Reuters also reported internal posts showing major technical and security incidents up 40%, with staff firefighting time up 70%. These figures are company-internal measures reported by Reuters. They do not isolate AI as the cause of every incident, and code-change volume is not a direct measure of useful software.

The output increased. So did the work of making sure it was safe and useful.

A FRONT-LINE EXAMPLE

WIRED reported claims adjusters describing AI misclassifying insurance claims and hallucinating details in summaries. One adjuster said staff had to reroute cases and correct errors that sometimes reached customers or their attorneys. The same article includes straightforward administrative uses workers found helpful. The task - and the check around it - matters.

Reuters, Aug. 26, 2026. Internal company measures; the report does not establish that AI caused every incident.

WIRED, Aug. 31, 2026. Interviewed-worker accounts, not a controlled test of AI causation.

04 / LEARNING TO VERIFY

Can people learn to spot the 5% that is wrong?

Valon, a mortgage-servicing software company, says it temporarily removed AI access for new non-engineering hires so they could learn the work and build enough judgment to question and verify AI output. The company's announcement frames this as a temporary, blunt policy; managers restore access when a person can verify the work. It is an operating choice, not evidence that the policy has already improved performance.

THE THREE QUESTIONS BEFORE AUTOMATION

01 Trust

How hard is it to know whether the output is right?

02 Failure

What is the cost if a mistake survives review?

03 Next scarce thing

If production grows, who reads, compares, integrates, and decides?

Sometimes automation removes work. Sometimes it simply makes another part of the workflow scarce.

Valon, "We turned off AI access for new hires" (Aug. 18, 2026). Company post describing policy and intent.

05 / FROM IMPRESSION TO EVIDENCE

Measure the task to its trustworthy end.

Give comparable tasks to two workflows: one completed directly by a person, and one where AI makes the first pass and a person verifies it. Compare the whole process, not just generation speed.

01	**Total human time**	Did effort fall from task start to trustworthy finish?
02	**Review time**	How expensive was it to establish trust?
03	**Corrections required**	How much cleanup did the first pass create?
04	**Errors that survive review**	What was missed, and what would relying on it cost?
05	**Exceptions and escalations**	Did automation move work somewhere else?
06	**Confidence in the result**	Would the person responsible actually act on it?

Not "AI won." Not "humans won." A map of where AI creates leverage – and where its speed hides the number that matters.

REPLY WITH TWO EXAMPLES

One effortless task. One with a cleanup tax.