The Journeyman Age

AI, the Printing Press, and the Price of Standing Behind Work

A position / working paper on naming and navigating the present information revolution

***Aaron Long** · August 2026 · **Draft v4** (core edition) — revised under the orders of an adversarial review (VERDICT.md, 14 Aug 2026). Every headline number independently re-verified 14 Aug 2026; entries carried from the earlier research base are tagged as such in* `data.py`*; formal claims computed by* `verification_bottleneck.py`

Abstract

This paper takes seriously a deceptively simple analogy: that artificial intelligence is to our century what the printing press was to the fifteenth. It reconstructs what happened between 1450 and 1710 — not the storybook version but the one historians now tell, in which print's effects ran through institutions, took generations to stabilize, and produced witch-hunt manuals alongside scientific journals. It maps the 2020–2026 empirical record onto that history and tests the analogy against the cases that break it: the internet, whose trust institutions never arrived, and East Asia, which held movable type for centuries without a Reformation. The central claim comes in two parts with different standings, and the paper holds them apart. The **mechanism claim**, for which task-level evidence exists: the printing press collapsed the cost of copying knowledge; AI collapses the cost of the competent first draft — the candidate application of knowledge — and, *at the task level*, shifts the binding constraint from generating work to standing behind it. The **market claim**, which is a staked prediction, not a finding: that this relocation will surface in wages, hiring, and institutions as adoption deepens — with named indicators, first readings (one supporting, one contrary), and stated failure conditions (§4.5, §5). A small formal exercise — an accounting identity, presented as such — locates where the empirical question lives. Robotics extends the collapse toward physical work on a timeline nobody knows. The paper closes by defending the name it proposes for the period: the **Journeyman Age** — accurate in both directions, at the machine and at the human — and states the condition under which the name dies.

What this object is. Position/working paper. AI-assisted throughout: research sweeps, verification, drafting, figures, build scripts — and the adversarial review itself, which was an **AI review commissioned by the author and conducted in-pipeline** (ATTACK_NOTES.md, DEFENSE_NOTES.md, VERDICT.md, all retained in the project record). That process is not a substitute for external human review, which this paper has not yet had. The human author owns the claims. Draft lineage: v1 (markdown); v2 (PDF, AI adversarial review, separate session); v3 (pipeline port, re-verified numbers); v4 (this draft, revised under VERDICT.md). Drafts v1–v3 carried the subtitle "AI, the Printing Press, and the Making of a Second Renaissance" — changed by order of the verdict (§11 convicts the name it borrowed). Sources and verification status for every number: `data.py` (V = primary-source verified, V2 = secondary naming the primary, U = unverified, hedge required). Conflict note: an AI-assisted paper about AI's effects discloses that its production is itself an instance of its subject.

1. Introduction: The Question Behind the Question

Every generation that lives through a technological rupture reaches for a historical mirror. The question “is AI like the printing press?” is really three questions wearing one coat. A mechanical question: does AI change the economics of knowledge the way print did? A social question: will it run the same course — democratization, then authority crisis, then a slow rebuilding of institutions? And an existential question: what happens to human skill, mastery, and meaning when the thing that made expertise scarce stops being scarce?

The mirror is worth polishing carefully, because the popular version of the printing-press story is wrong in ways that matter. In the popular version, Gutenberg invents the press, knowledge floods Europe, the Reformation and the Scientific Revolution follow as day follows night. The version historians defend is stranger and more useful: Gutenberg went bankrupt; the press printed witch-hunting manuals as enthusiastically as Bibles; its “revolutionary” reliability had to be manufactured by human institutions over two centuries; and the people who lived through it experienced not a golden age but a long, disorienting interregnum in which the old authorities had collapsed and the new ones had not yet been invented.

If that is the right story, the right questions about AI are not “will it make us smarter or dumber?” but: what will the interregnum look like, how long will it last, and what are the footnotes, journals, and copyright statutes of *this* revolution — the institutions that will eventually make the abundance trustworthy?

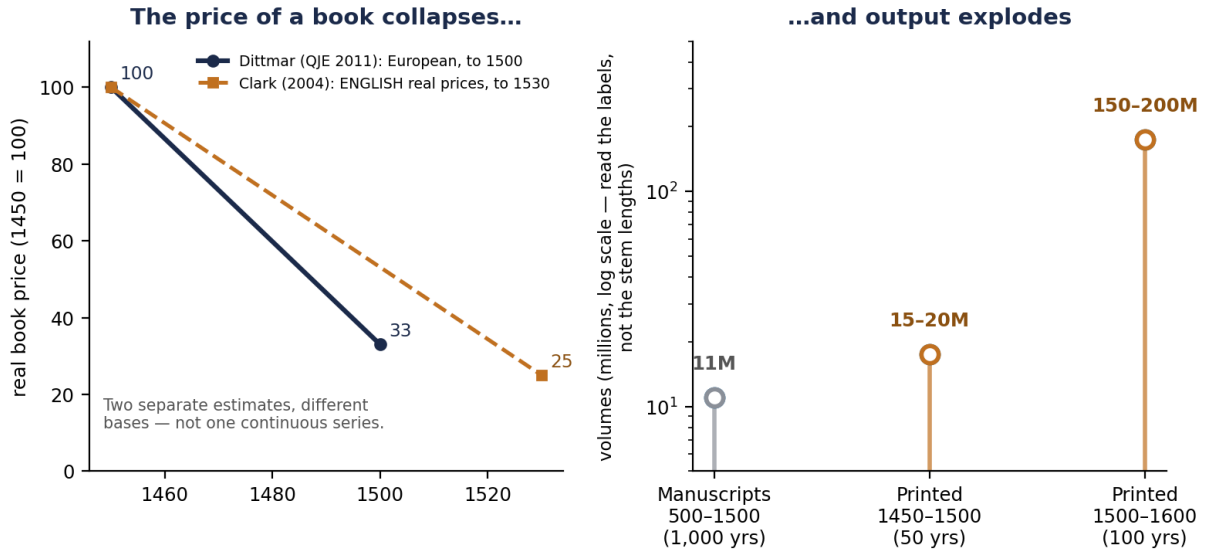
One more question threads through, because it was the seed of the project: what should this age be called? Candidates circulate — Altman's “Intelligence Age,” the consultancy-friendly “Second Renaissance,” the World Economic Forum's incumbent “Fourth Industrial Revolution.” This paper evaluates a different one: the **Journeyman Age**, drawn from the guild system that structured Renaissance work itself. The verdict is saved for §11, after the evidence.

2. What the Printing Press Actually Did (1450–1710)

2.1 The explosion

Gutenberg's press was operational in Mainz around 1450; finished copies of the 42-line Bible were available in 1454–55. The run was tiny — 160 to 185 copies — and the price enormous: roughly 30 florins, about 3 years' wages for a clerk. Then the technology spread along trade routes faster than anyone could govern it: presses in ~110 towns by 1480, and by 1500 in 205–282 towns, depending on how one counts. In its first fifty years the press printed an estimated 15–20 million volumes — against roughly 11 million manuscript books produced in the entire Latin West across the preceding *thousand years* (Febvre & Martin; Buringh & van Zanden). The sixteenth century added 150–200 million more.

The economics were brutal and simple. A press could set up to 3,600 pages in a workday against a scribe's handful (Wolf 1974). Real book prices fell by roughly two-thirds between 1450 and 1500 — to about a third of the pre-print level — and by 75% by 1530, to about a quarter (Dittmar; Clark). Cities that adopted printing in the late 1400s grew *at least* ~20 percentage points faster than non-adopters over the following century — Dittmar's preferred matched-city and IV estimates run 35pp and higher. The press was not just a cultural event but a growth technology.



Sources: data.py BOOK_PRICES (Dittmar QJE 2011; Clark 2004), VOLUMES (Febvre & Martin 1976; Buringh & van Zanden 2009)

Figure 1 (file: fig4_print_economics.png). The economics of the first information revolution. What it shows: two separate price estimates on different bases — Dittmar (European, to 1500) and Clark (English real prices, to 1530) — falling to ~1/3 and ~1/4 of the pre-print level, while output per period grows an order of magnitude. What it does not prove: any claim about who could read them (§7). Data: data.py BOOK_PRICES, VOLUMES.

3. The Analogy: What Maps, What Doesn't — and Where It Breaks

3.1 The core mapping: from copying to the competent first draft

The printing press collapsed the cost of *reproducing* knowledge. It did not make anyone wiser; it made the artifact of wisdom — the text — nearly free, and let literacy and institutions do the rest. Kissinger, Schmidt, and Huttenlocher, in the most-cited statement of the AI–print analogy (an op-ed, it should be said, not a study), date the stakes accordingly: generative AI is "a philosophical and practical challenge on a scale not experienced since the beginning of the Enlightenment" (WSJ, Feb 2023).

The first draft of this paper stated its formula as: print collapsed the cost of copying knowledge, AI collapses the cost of *applying* it. Adversarial review broke that sentence, and it deserved to break: "applying knowledge" equivocates between producing artifacts that *resemble* applications and the exercised, accountable performance of understanding. The evidence of §4 supports only the first. So the corrected formula, adopted throughout:

> *Print collapsed the cost of copying knowledge. AI collapses the cost of the competent first draft — the candidate application of knowledge — and relocates the scarce factor from generation to verification.*

This is a weaker sentence than the original and a stronger thesis — and its two halves have different standings, which the paper now states wherever the thesis appears. At the **task level** the relocation is what the era's studies observe (§4): novices gain most (they lacked drafts); experts sometimes gain nothing or lose (their bottleneck was already verification). At the **market level** the relocation is a staked prediction with named indicators and failure conditions (§4.5, §5), not a finding. The formula *organizes* the era's empirical surprises — the word "predicts" is reserved for the staked indicators. It also reconciles the two best framings

in circulation: Farrell, Gopnik, Shalizi, and Evans's "cultural and social technologies" (LLMs let humans access and recombine what other humans accumulated) and Kissinger's "human knowledge but not human understanding" — the understanding is the unautomated remainder, and the institutional problem of the age is how to supply it at the new scale of demand.

3.2 Five structural parallels

The price collapse. Books: down two-thirds to three-quarters within two generations. Cognition: the price of a competent first draft — prose, code, analysis, translation — has fallen effectively to zero for anyone with a smartphone, in under four years. **The authority shock.** Print broke the Church's monopoly on scriptural interpretation; AI is breaking the professions' monopoly on expert interpretation. In both cases the gatekeeper's product was not the text or the diagnosis but the *certified* text, the *certified* diagnosis; the certification layer shatters first. **The flood of unverifiable content.** Their pamphlet wars and forged tracts; our synthetic text and AI-written papers straining peer review — retractions hit a record above 10,000 in 2023 (driven substantially by the Hindawi paper-mill purge; annual counts have since fallen, but the mill capacity that caused the spike is now LLM-powered). **The panic literature.** De Strata and Trithemius have exact modern counterparts, down to the irony of medium — today's laments about AI are drafted and distributed with AI-adjacent tools, as Trithemius printed his defense of scribes. Behind both stands Socrates in the *Phaedrus*, warning that writing would "implant forgetfulness" — a warning we possess only because Plato wrote it down. **The displaced-expert absorption.** Scribes became compositors and correctors; translators, copywriters, and junior developers are being pushed toward AI-editing and verification roles — employed, when employed, in the machinery that displaced them.

3.3 Three structural breaks

Speed. Print took fifty years to reach two hundred-odd towns. ChatGPT reached an estimated 100 million users within about two months of launch; by February 2026, 900 million weekly. The best formal comparison (Bick, Blandin & Deming): 39.4% of American adults 18–64 had used generative AI within two years — faster than the PC or the internet at comparable points — rising to 54.6% by late 2025 and 61.8% by May 2026. One honest deflator rides along: that is the *ever-used* margin. The same authors put generative AI's assistance at roughly 0.5–3.5% of U.S. work hours. The adoption record is real, but it is a record of contact, not yet of dependence. And the consequence this paper states qualitatively and declines to arithmetize: adoption is compressing; there is no evidence that institution-building compresses with it. That gap — fast shock, slow repair — is the central risk of the age.

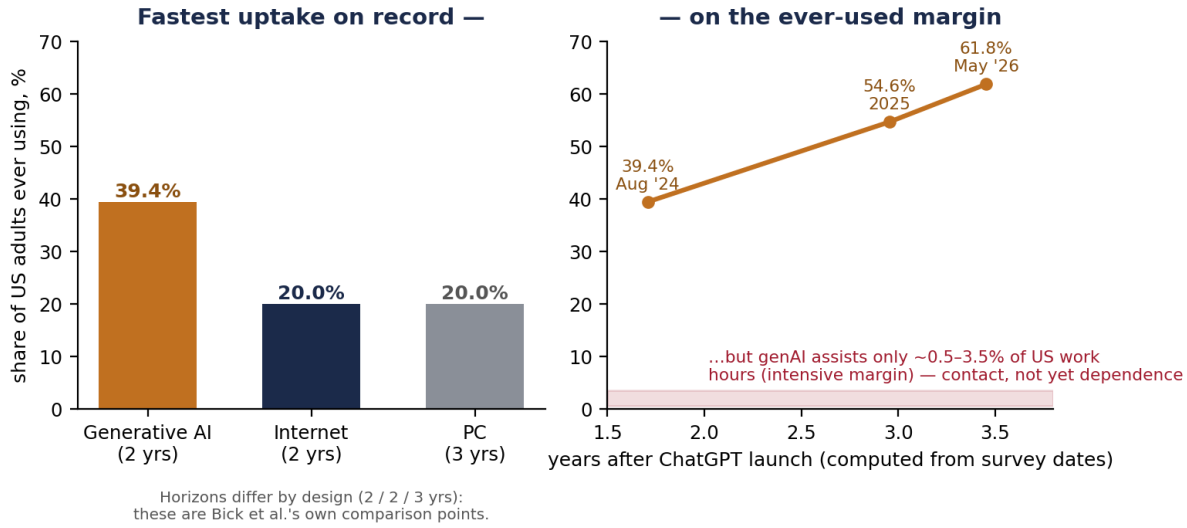


Figure 2 (file: fig2_adoption_speed.png). What it shows: generative AI outpacing the PC and internet on the ever-used margin — at the comparison horizons Bick et al. themselves use (2/2/3 years) — and the intensive-margin deflator on the same axes; trajectory x-positions computed from survey dates. What it does not prove: dependence, or economic effect (§4.5 carries the null). Data: data.py ADOPTION.

The direction of the literacy effect. Print *required* a new mass skill and so created a century of rising human capability as a side effect. AI *removes* skill requirements — it meets you in your own vernacular, spoken aloud. The on-ramp is frictionless, which is the point; but it means the technology does not drag a new mass competence along behind it. Whatever literacy this age needs — call it verification literacy — will have to be built deliberately, against the grain of the product experience. **The concentration tension.** Printing was a smallholder technology — a press cost 4–10 years of a craftsman's wages, a franchise, not an empire — and hundreds of shops competed under scores of rival jurisdictions. Frontier AI looks different: training runs in the hundreds of millions (GPT-4 ~$78M, Gemini Ultra ~$191M, growing ~2.4× a year per Epoch), and Big-4 hyperscaler capital spending — predominantly, not exclusively, AI-driven — ~$410B in 2025 with ~$725B guided for 2026. But "production is concentrating" cannot be left as a slogan: open-weight models (Llama, DeepSeek's R1) put near-frontier capability on commodity hardware, and distillation propagates each advance into cheap models within months — even though the "cheap" challenger itself sat atop a billion-dollar-class compute estate (SemiAnalysis: ~$1.6B, ~50k GPUs, against the advertised $5.6M final-run figure). The honest statement is a tension: **use is democratizing at record speed; the capacity to produce the frontier is held by a handful of firms and two states; and the contested middle — open weights, distillation, inference as a control point — will decide which force wins.**

3.4 Where the analogy breaks — and what survives the break

An analogy defended only against friendly cases is decoration. Three hostile cases must be faced. **The internet.** The nearest precedent in mechanism and speed is not print but the internet — an information-cost collapse now thirty years old whose trust-institutions phase has, by most accounts, not arrived. This is the strongest single objection to the optimistic use of the print arc, and it should be conceded rather than lawyered: the print sequence is not a schedule; it is one member of a reference class in which another prominent member has stalled mid-sequence. What the internet case does not do is break the diagnosis — it confirms it: the internet stalled precisely at the step this paper identifies as decisive, institutionalizing

verification. **East Asia.** Movable type was not invented in Mainz. Bi Sheng's ceramic type dates to the 1040s; the Korean *Jikji* was printed with movable metal type in 1377. China and Korea held the technology for centuries without a Reformation or a Scientific Revolution — proof that the arc was never a property of the machine. Far from embarrassing the thesis, the East Asian record is its strongest evidence: *the technology proposes; institutions dispose*. But it cuts the other way too — no timeline, including this paper's, can be read off the technology. **The generative disanalogy.** The deepest objection: the press never wrote books. No print institution had to certify an author who could not be deposed, sued, or shamed. Harari names the discontinuity: AI is "the first technology in history that can make decisions and create new ideas by itself." The concession due is bounded but real: claims leaning on print's *content* dynamics (authorship, liability, reputation) transfer poorly. What survives is the structural core — a cost collapse in producing claims, an authority crisis in certifying them, scarcity relocating to verification — all of which hold *a fortiori* when the producer cannot be cross-examined. The analogy survives as a diagnostic and dies as a forecast. Every use of it after this point should be read in that light.

4. The Evidence So Far: What Six Years of Studies Tell Us

4.1 The great compression

The most consistent *direction* of effect across the era's heterogeneous studies is skill compression: AI helps the least skilled most and narrows the measured gap between bottom and top. A methods caveat governs everything in this section: these studies measure different outcomes on different populations with different designs; their percentages cannot be pooled, averaged, or meta-analyzed, and this paper does not do so. What recurs is the *gradient*, not a number. Noy and Zhang (*Science*, 2023; n=453): ChatGPT cut professional writing time ~40% and raised quality ~18%, with gains concentrated among the weakest writers. Brynjolfsson, Li, and Raymond (*QJE*, 2025; n=5,172 support agents): +15% issues resolved per hour overall, but ~+30% for the newest agents and roughly zero for the most experienced; agents with two months of tenure plus AI performed like agents with more than six months without. The model had absorbed the tacit moves of top performers and was distributing them to everyone else. Dell'Acqua et al. (n=758 BCG consultants): bottom-half performers +43%, top-half +17%, inside the frontier. Peng et al.: developers 56% faster on a benchmark task; the large-scale follow-up (Cui et al., *Management Science*, 2026; n=4,867 developers at three firms) found +26% completed tasks, tilted toward juniors. In guild terms: **the models are journeyman-makers.** They rent an apprentice the competence of a journeyman by the day. Print photocopied the master's artifact; AI photocopies the master's performance.

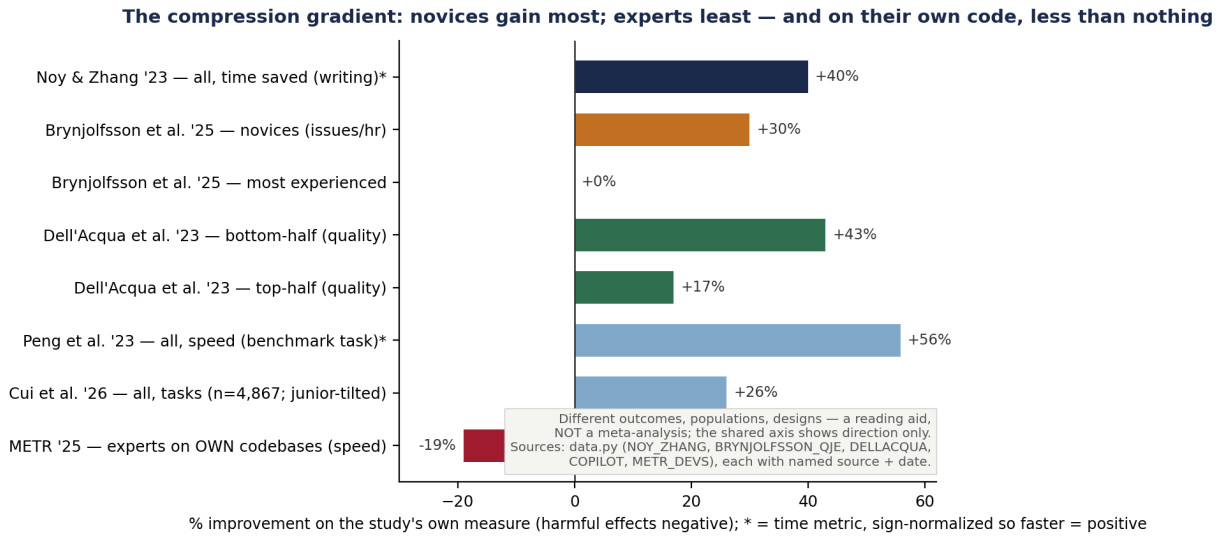


*Figure 3 (file: fig1_skill_compression.png). Direction and reported size of AI-assistance effects by experience group, study by study. Each bar is that study's own outcome measure; time metrics are sign-normalized (faster = positive) and marked *; the shared axis displays direction, not comparability — a reading aid, not a meta-analysis. What it does not prove: any pooled effect size. Data: data.py, entries as labeled.*

4.2 The jagged frontier and the expert's revenge

The compression has a boundary, and the boundary bites. On a task engineered so the model's confident answer was subtly wrong, Dell'Acqua's consultants with AI were 19 percentage points *more* likely to be wrong than consultants without — "falling asleep at the wheel." The sharpest counter-result: METR (July 2025) — 16 highly experienced open-source maintainers, 246 real tasks on their own codebases, 19% *slower* with early-2025 tools, while estimating they had been sped up ~20%. Their tacit knowledge of their own repositories outran the model, and reviewing plausible-but-wrong output cost more than it saved. (METR's February 2026 follow-up found point estimates moving toward zero with better tools — and METR itself flags severe selection problems and calls it "only very weak evidence.") In medicine, Goh et al. (*JAMA Network Open*, 2024; 50 physicians, 6 written vignettes — a test of reasoning display, not of care): the LLM alone scored a median 92% against 74–76% for physicians *with or without* the model. Giving the doctor the model didn't help the doctor: an integration failure, not a capability one. Meanwhile, at the economy-wide level, Humlum and Vestergaard's Danish registry study (~25,000 workers, 11 exposed occupations) finds precise nulls — ruling out earnings or hours effects larger than 2% two years post-ChatGPT, with time savings around 2.8–3.0%. The same vintage caveat that governs the compression studies governs this null: it measures the shallow-adoption period, 2023–24 tools at 2023–24 usage intensity. The gap between laboratory gains and national statistics is what economists predicted from history: general-purpose technologies pay off only after organizations rebuild around them — electricity took roughly forty years. Brynjolfsson calls it the productivity J-curve; we are in the Between Times.

4.3 The ladder problem: what happens to apprentices

If AI rents everyone journeyman competence, what happens to the apprenticeship — the drudgery-as-curriculum by which mastery was actually made? The early labor data is contested and deserves adjudication rather than a shrug. The Stanford Digital Economy Lab reports that since late 2022 early-career workers (ages 22-25) in the most AI-exposed occupations saw a ~16% relative employment decline (the 13% first estimate rose on revision) while older workers in the same occupations held steady. The critics' case

(EIG, Jan 2026) has two halves and they deserve separate answers. The macro half — tech-hiring peak predating ChatGPT, interest rates confounding the window — is substantially addressed by the study's November 2025 specification, which recovers the ~16% decline with firm-time fixed effects, i.e. comparing age groups *within the same firm at the same moment*. The age-band half — the estimate's sensitivity to the narrow 22–25 window — has not been answered and stands. So the paper's ruling: **plausible and unproven**, carried as a warning that borrows a contested number and says so; it would be overturned by the entry-level gap closing as the cycle normalizes, and a reader in 2028 should check whether it has. It rhymes with the platform evidence, which carries a sting the compression studies miss: after ChatGPT, freelance writers' jobs fell 2% and earnings 5.2%, with the most skilled *not* spared — at the level of earnings, early evidence looks less like equalization than displacement. The deepest version of the worry is Matt Beane's, from field studies of robotic surgery: intelligent machines let experts work productively *without* novices — the resident who once held retractors now watches a screen. The expert–novice bond, the mechanism that has transmitted skill since the guilds, is severed at the exact moment the machine takes over the novice's traditional duties. The guild economy would not have recognized the danger — masters took labor-saving where they found it — but it would have recognized the arithmetic: a bottega with no *garzoni* produces no next Verrocchio, and no next Leonardo.

4.4 The atrophy question: Socrates' scorecard

Does using the crutch weaken the leg? The honest answer in 2026: early, mixed, worth taking seriously without overclaiming. Worrying: experienced endoscopists' unassisted adenoma detection fell from 28.4% to 22.4% after months of AI assistance (*Lancet Gastro Hep*, 2025 — observational, 19 endoscopists, patient-mix caveats: hypothesis-generating, not a verdict). A large RCT with Turkish high-school students: practicing with vanilla ChatGPT boosted practice performance 48% but lowered subsequent unassisted exam scores 17% — the crutch effect, cleanly measured. Crucially, a tutor-style version with guardrails eliminated the harm (*PNAS*, 2025). A Microsoft–CMU survey (n=319) found higher confidence in AI predicting less critical-thinking effort (correlational); MIT Media Lab's small EEG study (n=54, still a preprint) should carry no more weight than a pilot. Hopeful, from the same literature: the outcome is a **design choice**, not a destiny — the guardrailed tutor didn't harm learning; a Harvard RCT found students learning more than twice as much from a well-designed AI physics tutor as from an expert-led active-learning classroom; a World Bank pilot in Nigeria reported +0.31 SD in six weeks — a single small pilot in one city, measured on English-language outcomes, whose “most cost-effective ever” framing came from its sponsors: treat it as promising, not proven, until it replicates. Socrates was half right about writing: memory culture really did wither — and in exchange we got literature, law, and science. The question is never whether a cognitive technology reshapes the mind, but whether the trade is negotiated or sleepwalked.

4.5 The limits of this evidence — and how this paper could be wrong

Standing commitments, rewritten under order of the v3 verdict so that each failure condition wounds on its own. **Vintage and validity.** Every compression study above used 2023–24 systems on bounded tasks; none establishes that the gradient survives on current agentic systems, and none licenses claims about occupations as lived. Lab task-performance is not economic outcome. The same discipline applies to the Danish null: it is the best estimate *for the shallow-adoption period it measures*, not a verdict on the technology.

What would count against this paper — each condition individually damaging. By roughly 2030, which is soon enough to adjudicate the task-level and labor-market claims (though not the institutional arc, whose clock this paper has refused to set): **(a)** if the novice-tilted gradient does not replicate on frontier-model

deployments in real workplaces, §4.1's mechanism claim fails — this alone guts the paper's empirical core. **(b)** if the entry-level employment gap in AI-exposed occupations closes as the tech cycle normalizes, the ladder-problem warning (§4.3, §8) loses its labor-market leg and keeps only Beane's fieldwork. **(c)** if no acceleration appears in the reorganization indicators J-curve theory says move first — firm process redesign, AI-complementary hiring, intangible investment — then "payoff ahead" is faith, and §4.2's Between Times framing fails. **(d)** if no verification premium emerges in wages or postings for verification-heavy roles, no provenance adoption growth is measurable, and no liability event forces certification anywhere — then the market-level relocation claim fails directly. Conditions (a) and (d) together are fatal to the thesis; either alone forces a rewrite, not a footnote. The institutional diagnosis of §10 is argued, not testable by 2030, and the paper claims no more for it.

5. The Verification Bottleneck, Formalized

The thesis's arithmetic, in code that runs (`verification_bottleneck.py`), so the reader can see exactly what is identity, what is assumption, and what is evidence. Under order of the v3 verdict, this section is presented as what it is: **an instrument, not evidence.** It is an accounting identity — true of any two-stage decomposition of anything, haircuts included — and its role is to locate where the empirical question lives. It applies to time and throughput outcomes only; the quality and accuracy results of §4 (Dell'Acqua's quality gains, the −19pp accuracy failure) lie outside it.

One certified unit of work requires generation time g and verification time v. AI divides generation by k and multiplies verification by m (m > 1 when reviewing machine drafts costs more than reviewing your own work; m < 1 with good verification tooling). Throughput changes by $R = (g+v)/(g/k + mv)$, and the limit as generation becomes free is $R \to (1+g/v)/m$: the maximum possible gain is set entirely by how generation-heavy the job was. From that limit, two illustrations. Even with FREE generation and no verification inflation, verification-heavy work (g/v 0.3-1.5) can gain at most 1.3-2.5x; with 25% verification inflation that range falls to 1.04-2.0x, and at 50% inflation the bottom drops to 0.87x — a net loss even with free generation. Generation-heavy work (g/v 4-10) can gain 5-11x in the same limit — the compression gradient is the geometry of the job. And the reversal: At a 5x generation speedup, verification-heavy work flips to a net LOSS once verification inflates by more than 24-120% — reviewing plausible-but-wrong output is exactly such an inflation.

Now the honesty the identity owes the reader, stated in the text and not only in a docstring. **The parameter assignment is the conclusion.** Calling apprentice work generation-heavy (g/v of 4–10) and master work verification-heavy (0.3–1.5) is an assumption with face validity and no measurement behind it; it encodes the direction the §4 studies found, and a hostile relabeling would encode the opposite. The studies did not measure g/v or m, and across the assumed ranges the identity accommodates almost any result — so it is evidence for nothing. What it does is convert a vague dispute into two measurable ratios: whether novice work is in fact generation-heavy, and whether AI in fact inflates expert verification, are the entire empirical question, and nobody has measured either. That is a research target, offered as one. In words, the mechanism hypothesis reads: **when the machine writes the draft, the price of the work becomes the price of standing behind it.**

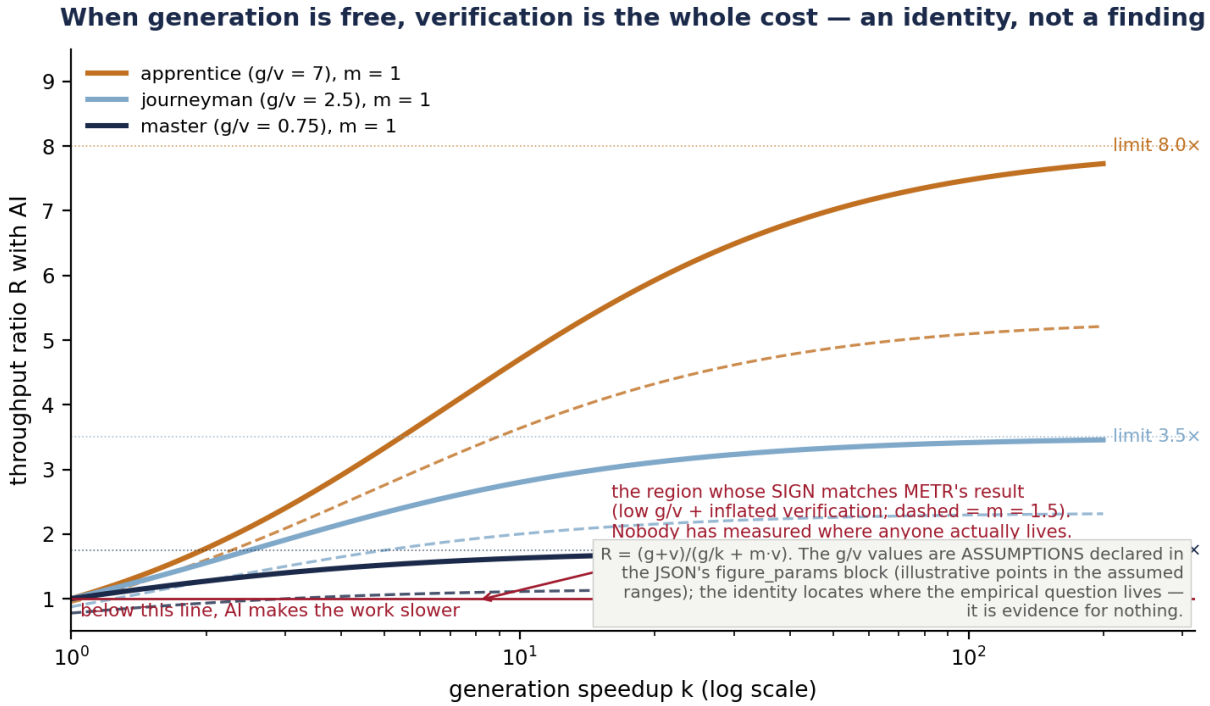


Figure 4 (file: fig5_verification_limit.png). The verification bottleneck. What it shows: throughput ratio R against generation speedup k for three ASSUMED job geometries, declared in the JSON's figure_params block (solid m = 1, dashed m = 1.5); horizontal asymptotes are the free-generation limits (1+g/v)/m. What it does not prove: any empirical effect size — nobody has measured g/v or m for anyone. Data: verification_bottleneck_result.json.

First market-level readings. The v3 review's sharpest charge was that the relocation claim named no measurement, so here are the first two, one on each side, entered in data.py with the rest. Supporting: the Dallas Fed's wage analysis (Davis, Feb 2026) finds AI exposure *raising* relative wages in occupations with high experience premia (+0.2pp at the 90th-percentile premium) while *lowering* them where the premium is zero (–0.28pp) — experience-heavy work is the closest published proxy for verification-heavy work, and the sign matches. Contrary: the NY Fed's postings analysis (Audoly, Guérin & Topa, May 2026) finds junior and senior postings in exposed occupations moving broadly in parallel, with the relative decline in exposed occupations beginning *before* ChatGPT — no postings-level verification premium is visible yet. Both are between-occupation results, neither is the within-occupation test the claim needs, and the disconfirmer stands: if verification-heavy roles show no relative price or quantity movement as the intensive margin deepens past its current 0.5–3.5% of work hours, the market-level claim fails (§4.5(d)).

The same script performs the adoption arithmetic and refuses the step v1 of this paper took and v2 retracted: it computes the measured adoption comparisons (100M users in ~2 months; 39.4% of US adults 18-64 in 2 years; 61.8% in 3.5 years (data.py: CHATGPT_USERS, ADOPTION)) against print's (~110 towns in 30 years; 200-280 towns and 15-20M volumes in 50 years (data.py: PRESS_SPREAD, VOLUMES)), and it declines to divide one era's institutional history by another era's adoption ratio. Adoption speed is measured; institutional speed is not; no identity connects them. The refusal is recorded in the JSON, where it cannot quietly un-happen.

11. Naming the Age: The Journeyman Verdict

Names are theories. “Fourth Industrial Revolution” theorizes AI as another turn of the automation crank — underselling a technology aimed at cognition itself. “Intelligence Age” theorizes intelligence as the new electricity — accurate about supply, silent about what happens to the humans who used to be the supply. “The Coming Wave” names the danger, not the age. “Second Renaissance” is the flattering mirror — right about the explosion of capability, wrong by omission about who enjoyed the first one, which its inhabitants experienced substantially as plague, war, and upheaval, and which was declared golden only afterward, by its beneficiaries.

Against this field, the proposal on the table. The case rests on the layered accuracy of the metaphor. The journeyman was **certified competence without mastery** — exactly what a language model delivers: reliably good, unreliably excellent, unwise at the edges (§4.2). The journeyman was **paid by the day** — competence metered by the token, the *journée* made literal in API pricing. The journeyman was **itinerant** — carrying techniques from workshop to workshop, as a model carries patterns from every discipline into every other. And the journeyman was defined by his **position on a ladder** — the middle rung — which names what the task-level evidence suggests AI is doing: hoisting every apprentice toward the middle rung at once, while quietly sawing off the bottom one.

An earlier draft then performed an inversion it was too pleased with: the AI is the journeyman, so the human is promoted to master of a workshop — every person a Verrocchio. Review converged on two objections, absorbed here rather than deflected. First, mastery in the guild sense was not a slot but a **formation** — a masterpiece judged by peers after years of supervised struggle — and this paper's own §8 documents that formation's destruction. A civilization cannot promote to master a population whose path to mastery it is simultaneously dismantling; masters without formation are consumers with tools and a title, and the deference evidence shows what their signature is worth. Second, Verrocchio **owned** the bottega — capital, contracts, signature — whereas the 2026 “master” rents metered cognition from a hyperscaler that can reprice, retrain, or revoke it, and that harvests the renter's usage as training data. Read as a description of the present, the workshop reading gets the relations of production backwards: it is the humans who are day-rate, precarious, and itinerant across platforms — and, at the global bottom of the workshop, the annotation and RLHF workforce are the *garzoni*, grinding pigments for workshops whose output they will never sign. The historically resonant warning is worse still. The guild literature documents early modern Europe raising “the barriers between master and journeyman higher than ever before” — escalating fees, extended apprenticeships, refined masterpiece requirements — with journeymen organizing *compagnonnages* as the rung receded (Encyclopedia of European Social History; Truant; cf. Ogilvie on guild entry barriers). Some historians read the end state as a class of permanent journeymen effectively barred from mastership — a characterization this paper reports rather than certifies. Softened or not, the shape is the warning: a ladder whose middle rung is crowded and whose top rung is quietly withdrawn is a possible equilibrium of this age, not a metaphor for it.

One more piece of evidence belongs in this section because it cuts against the name, and the v3 review caught this paper holding it without confronting it. METR's time-horizon series — the one capability measure denominated in time rather than benchmark points — finds the length of software tasks models complete at 50% reliability doubling every ~7 months long-run and every ~89 days since 2024, with the longest measured horizon around five hours (wide intervals; software only; a revised instrument). If the journeyman is improving that fast, is “Journeyman Age” a name or a two-year-old snapshot? The answer this paper stakes: the name names the **certification gap**, not a capability level. Lengthening horizons move more work inside the frontier without touching who stands behind it — a five-hour draft is still a draft until

someone signs. But that answer gives the name a stated death condition, and the reader should hold the paper to it: **the day a model can stand behind work — carry liability, be deposed, lose standing it cares about — the certification gap closes and this name dies.** A name with a death condition is a theory; without one it is a brand.

So the verdict keeps the name and changes shape: **"Journeyman Age" is the right name because it is accurate in both directions at once.** Read at the machine, it names the mechanism: tireless day-rate competence in every craft, delivered without understanding or accountability — the candidate-application engine of §3. Read at the human, it names the condition: a civilization issued journeyman output on day one, its apprenticeships dissolving beneath it and its mastership — the formed judgment that can stand behind work — become the scarcest thing in the economy. The workshop reading survives, demoted from description to **program**: every person a Verrocchio is not what the age *is*; it is what the age would have to build to deserve its optimists — new apprenticeships that manufacture judgment on purpose (the guardrailed tutor is the existence proof), and ownership structures (open weights are the current candidate) under which the workshop is not a tenancy. The name keeps the question "and then what?" permanently in view, which is the most useful thing a name can do — and it should be worn with the scope-modesty the metaphor demands: a European image for a planetary condition, chosen because the European case is where the causal history is best measured, not because the age belongs to Europe. It does not.

(Runner-up names, demoted to this note by order of the v3 verdict: the Bottega Age, naming the social form the program aims at; the Age of Taste, with the caveat that taste is a discipline, not a birthright remainder, and atrophies exactly as fast as it is outsourced.)

12. Conclusion

Strip the paper to its load-bearing claims, as revised under two rounds of fire, and hold each at its stated strength. **Demonstrated at the task level:** the printing press collapsed the cost of copying knowledge; AI collapses the cost of the competent first draft — the candidate application of knowledge — and shifts the binding constraint from generating work to standing behind it. **Predicted at the market level, with indicators and failure conditions staked in §4.5 and first readings, one on each side, in §5:** that the same relocation will surface in wages, hiring, and institutions as adoption deepens. Robotics extends the collapse toward physical work on a timeline nobody knows. The print record supplies a diagnostic sequence — explosion, authority crisis, dark harvest, control response, and only much later the institutions that made abundance trustworthy — but no schedule: East Asia held the same technology for centuries without the arc, and the internet has sat mid-sequence for thirty years. The arc is not a property of the machine. It is a property of the response.

The empirical record six years in is neither utopian nor null, and its two layers must be held separately: at the task level, a consistent novice-tilted compression gradient, early and confounded signs of atrophy, and one clean proof that the atrophy is a design choice; at the economy level, approximately nothing yet — a null that is the best estimate for the shallow-adoption period it measures, not an embarrassment. The paper has staked its falsification conditions (§4.5) and stands by its declared bet — a bet, tied to those conditions, not a schedule: that the task-level mechanisms are real, that they will surface in the aggregate as organizations rebuild, and that the decisive work of the coming decades is institutional — provenance, evaluation, verification literacy, and above all a rebuilt path from journeyman to master.

And the name. The age deserves one that tells the truth in both directions, and “Journeyman Age” — held with both readings, and carrying the death condition stated in §11 — does. The first Renaissance told itself a story about what a human could be and built workshops that occasionally made it true. The Journeyman Age has the workshops; whether it rebuilds the formation that makes masters — or settles, as the guild literature suggests late guild Europe increasingly did, into a journeymen's estate serving whoever owns the shop — is the question the name exists to keep open. The wager worth making is the one the evidence permits: that judgment can still be manufactured on purpose, and that the will to do so remains ours to supply.

Appendix: Status of Claims

Claim	Evidence type	Status	Failure condition
Novice-tilted compression gradient at the task level (§4.1)	RCTs + field experiments, 2023–24 systems	Demonstrated for its vintage; replication on frontier systems open	§4.5(a): gradient fails on frontier deployments
Expert reversal — AI slows verification-heavy experts (§4.2)	One strong study (METR) + weak 2026 follow-up	Suggestive, not established	Replications showing robust expert speedups
Relocation of the binding constraint to verification — task level (§3.1, §5)	Reading of the gradient + an accounting identity	Interpretive frame; g/v and m unmeasured	Measured g/v showing novice work is NOT generation-heavy
Relocation at the market level — verification premium (§5)	Two between-occupation first readings, one contrary	Staked prediction, not a finding	§4.5(d): no premium as the intensive margin deepens
Entry-level employment decline in exposed occupations (§4.3)	Payroll working paper, contested	Plausible and unproven; band-sensitivity standing	§4.5(b): gap closes as the tech cycle normalizes
Skill atrophy from unguardrailed use (§4.4)	One clean RCT + observational + correlational	Design-dependent harm: shown once, cleanly	Failed replications of the crutch effect
Aggregate labor-market effect ≈ 0 so far (§4.2)	Registry study, 25k workers, 2023–24 exposure	Best estimate for the shallow-adoption period	Superseded by deeper-adoption estimates
Print arc as diagnostic sequence, not schedule (§2, §9)	Historiography + two counter-cases	Argued; not testable by 2030	A pre-trust phase with no AI analogue at all
'Journeyman Age' as the right name (§11)	Fit of entailments to assembled evidence	Defended; carries a stated death condition	A model that can stand behind work (liability, deposition)

Sources

Every number in this paper lives in `data.py` with a source string and a verification tag (V / V2 / U), re-checked 14 Aug 2026; the formal claims live in `verification_bottleneck_result.json`. Principal sources:

Print era and Renaissance

— Dittmar, 'Information Technology and Economic Change: The Impact of the Printing Press', QJE 126:3 (2011)

— Buringh & van Zanden, 'Charting the Rise of the West', Journal of Economic History 69:2 (2009)

— Febvre & Martin, The Coming of the Book (1976)

— Eisenstein, The Printing Press as an Agent of Change (1979); Johns, The Nature of the Book (1998); AHR Forum (2002)

— Pettegree, Brand Luther (2015); Edwards, Printing, Propaganda, and Martin Luther (1994)

— Palmer, Inventing the Renaissance (2025); Blair, Too Much to Know (2010); Grafton, The Footnote (1997)

— Mackay, The Hammer of Witches (2009); Levack, The Witch-Hunt in Early Modern Europe (2006); Ogilvie, The European Guilds (2019); Csiszar, The Scientific Journal (2018)

— Encyclopedia of European Social History, 'Guilds'; Truant, 'Solidarity and Symbolism among Journeymen Artisans', Comparative Studies in Society and History

— Gingerich, An Annotated Census of Copernicus' De revolutionibus (2002)

— NGA Italian Renaissance Learning Resources, 'Training and Practice'; Wikipedia entries as named in data.py

AI empirical record

— Noy & Zhang, Science 381 (2023); Brynjolfsson, Li & Raymond, QJE 140:2 (2025); Dell'Acqua et al., HBS WP 24-013 (2023) / Organization Science (2025)

— Peng et al., arXiv:2302.06590 (2023); Cui et al., Management Science (Feb 2026)

— METR, developer study (10 Jul 2025; arXiv:2507.09089) and update (24 Feb 2026); METR, Time Horizon 1.1 (29 Jan 2026)

— Humlum & Vestergaard, NBER w33777 (rev. 2025); Budzyn et al., Lancet Gastro Hep (2025); Bastani et al., PNAS 122:26 (2025); Goh et al., JAMA Network Open (2024)

— Bick, Blandin & Deming, NBER w32966 / Management Science; St. Louis Fed (Nov 2025); genaiadoptiontracker.com (May 2026)

— Brynjolfsson, Chandar & Chen, 'Canaries in the Coal Mine' (Nov 2025); EIG, 'Looking for the Ladder' (Jan 2026)

— Hui, Reshef & Zhou, Organization Science (2024); Kosmyna et al., arXiv:2506.08872; Lee et al., CHI 2025

— Epoch AI, frontier training costs (2024); SemiAnalysis on DeepSeek (Jan 2025); FT/Jefferies capex compilation (Apr 2026)

— Davis, 'AI is simultaneously aiding and replacing workers', Federal Reserve Bank of Dallas (24 Feb 2026); Audoly, Guérin & Topa, 'Do Job Postings Show Early Labor-Market Effects of AI?', Liberty Street Economics, NY Fed (14 May 2026)

— David, 'The Dynamo and the Computer', AER 80:2 (1990); TechCrunch on Twitter Birdwatch (25 Jan 2021)

Robotics and institutions

— Google DeepMind, Gemini Robotics 2 (30 Jul 2026); Physical Intelligence, pi-0.5 (arXiv:2504.16054)

— Brooks, 'Why Today's Humanoids Won't Learn Dexterity' (26 Sep 2025); TrendForce (9 Apr 2026); 1X/GlobeNewswire (30 Apr 2026); TechCrunch/CNBC on Amazon robotics (Jul 2025)

— Hugging Face LeRobot (2025-26); C2PA / ISO/DIS 22144; Google SynthID (May 2026); NIST CAISI; EU AI Act and Digital Omnibus (Gibson Dunn, May 2026)

— Nature on retractions (Dec 2023); NobelPrize.org (2024 Chemistry; 2025 Economic Sciences)

Commentary

— Kissinger, Schmidt & Huttenlocher, WSJ (24 Feb 2023); Farrell, Gopnik, Shalizi & Evans, Science 387 (2025); Harari, Nexus (2024); Beane, The Skill Code (2024); Mollick, Co-Intelligence (2024)

— Altman, 'The Intelligence Age' (2024); Amodei, 'Machines of Loving Grace' (2024); Messeri & Crockett, Nature (2024); Mokyr, A Culture of Growth (2016)

Disclosure. Position/working paper; the human author owns the claims. AI assisted throughout — research sweeps, verification, drafting, figures, build scripts — and conducted the adversarial review of draft v3 as an AI process commissioned by the author (ATTACK_NOTES.md, DEFENSE_NOTES.md, VERDICT.md, retained in the project record along with all superseded drafts and the error log). That review is not a substitute for external human review, which this paper has not yet had. Rebuild: `python3 generate_figures.py && python3 build_paper_v4.py` (verification_bottleneck.py runs first if the JSON is absent).