The Flywheel and the Frenzy: Navigating the AI Turning Point

A Trinzo Thought Leadership Paper

The Question We Are Not Asking

Organisations engaged in AI transformation are asking one question with considerable urgency and budget attached: how do we implement AI? They are getting answers from every direction. Vendors with platforms to sell, consultants with methodologies to deploy, conference speakers with frameworks to share. The market for AI implementation advice has never been larger, and it is still growing. I contribute to that market myself, so I say this with the appropriate degree of self-awareness: most of what is being sold answers the question as asked, and the question as asked is the wrong one.

The question organisations are not asking is the one that determines outcomes: what do we already have that is worth amplifying, and does everyone in the organisation understand what that is?

The distinction matters because the first question tends to produce activity and the second tends to produce results, and the two are currently being confused for each other on an industrial scale. The most widely cited evidence for the gap is a 2025 study from MIT's NANDA initiative, which reported that 95% of generative AI pilots delivered no measurable financial impact despite $30 to $40 billion in aggregate investment (MIT NANDA, 2025). That figure needs handling with care. The study is preliminary, it has not been peer reviewed, and its methodology has been contested; the true number may be materially lower. I cite it because even heavily discounted it points in a direction that matches what I see in practice: a great deal of impressive-looking output, and very little measurable return. Whatever the precise figure, this is not primarily a technology failure. It is a question failure.

The organisations generating real, compounding value from AI share one characteristic that has nothing to do with their technology stack. They understood what they were amplifying before they pointed the tool at it. That sounds obvious. In practice it requires a depth of organisational self-knowledge that is far rarer than it appears, and it cannot be manufactured by a technology deployment, because the deployment is downstream of it. You cannot buy self-knowledge from a vendor. You cannot install it. The consultants who tell you otherwise are selling you the first question wearing the second question's clothes.

That is the argument of this paper. AI amplifies what an organisation already has. It does not manufacture what it lacks. The sections that follow work through what

that principle implies: where the technology sits in its financial cycle, what the honest limits of the optimist case are, what amplification looks like at the level of the individual, the team and the organisation, what happens when the question is answered well and badly, why the transformation lags the technology and why the lag has a particular shape, and where the principle runs into a problem I cannot yet resolve. On method: where I am drawing on practice, I will say so. Where I am uncertain, I will say that too.

One thing should be said plainly at the outset, because the failure statistics invite the wrong conclusion. The technology works. It does work, within the conditions it requires. Most organisations have not built those conditions, and the gap between the two facts is where most of the current confusion lives.

Reading the Cycle

The public conversation about AI is not much help in locating where we actually are. The range of opinion runs from Terminator 2 to nobody ever having to work again, and both ends of that range are more confident than the evidence permits. Between the poles sits a great deal of commentary that is really just mood: euphoria dressed as analysis, or dread dressed as prudence. A better instrument is history. History never repeats, but it hums the same tune.

The economist Carlota Perez spent decades studying how transformative technologies develop. Her framework, set out in Technological Revolutions and Financial Capital (Perez, 2002), identifies a consistent pattern across steam power, railways, electricity, automobiles and information technology. Each passes through an installation period, driven by financial capital, in which infrastructure is built at a pace that races ahead of demonstrated productive demand. The peak of this phase, which Perez calls the Frenzy, is characterised by speculative investment, extraordinary valuations, escalating claims about the technology's potential, and a widening gap between the financial story and the economic reality. Then comes the Turning Point: a correction that deflates the speculative excess and forces the technology to prove its value on productive terms. What follows, for technologies that are real, is the Deployment Period, in which the infrastructure built during the Frenzy is finally used productively and the transformative applications become commercially and socially embedded.

The pattern is worth sitting with, because its details are stranger than the summary suggests. British railway mania in the 1840s destroyed enormous quantities of capital and a fair number of reputations; it also left Britain with a rail network that no sober parliament would ever have voted to fund, and that network carried the second half of the industrial revolution. The dot-com collapse of 2000 wiped out trillions in paper value; the fibre laid during the madness carried the internet economy that followed, at prices the frenzy's investors never recovered. In each case the financial capital that built the infrastructure and the

productive capital that used it were different money, held by different people, with different fates. The people who fund the Frenzy are rarely the people who profit from the Deployment. This is worth remembering when the current wave of investors assures you that this time the returns are just around the corner.

The parallels to the current moment are difficult to ignore. Since the public launch of ChatGPT in late 2022 we have seen extraordinary infrastructure investment, primarily from hyperscale cloud providers, channelled into frontier research labs that have produced genuine capability advances at a pace that has surprised experienced practitioners. The capital structure of that investment deserves scrutiny. The major cloud providers have invested tens of billions of dollars into OpenAI, Anthropic and similar organisations, which have in turn spent the majority of that capital on compute supplied by the same providers. The cash circulates. The structure inflates the apparent scale of commercial activity while obscuring whether end-user revenue exists at anything like the level required to justify the valuations. Bain and Company's 2025 Global Technology Report estimates that meeting projected AI compute demand would require roughly two trillion dollars in new annual revenue by 2030, and that even on generous assumptions about AI-driven savings the industry falls some 800 billion dollars short (Bain & Company, 2025). That is a scenario estimate rather than a measurement, and Bain has its own incentives in the AI conversation, but the direction of the gap is not seriously disputed. The financials mean everything and nothing simultaneously. Life goes on.

We are, on my reading, deep in the mad frenzy phase, and a correction is the likely next act. I will attach a confidence level to that, because it is a prediction rather than an observation. That a correction of some kind arrives, I hold with high confidence; the Perez pattern has run five times in two centuries and the financing gap is real. On timing I am much less certain. The public listings expected from the major frontier labs will bring the valuation story into contact with the revenue reality, and my expectation, with moderate confidence, is that 2027 will feel to many organisations like a reset. If it arrives two years later, the argument of this paper does not change.

There is a practical corollary that deserves stating before the theory continues, because it converts the cycle-reading into a decision. The current cost of experimenting with these tools is artificially low. The subsidy is not a metaphor: the pricing of frontier model access is being underwritten by the same speculative capital that is building the data centres, and whether inference at current prices covers its own costs is an open question. The providers do not publish the numbers, so I state that as informed inference rather than established fact, and hold it with moderate confidence only. When the correction arrives, some combination of things happens to that subsidy: prices rise, free tiers shrink, weaker providers exit, and the surviving platforms consolidate their pricing

power. I do not know which combination, and anyone who claims to is guessing. What I do know is that the cheapest period for an organisation to learn, to run the experiments, make the instructive mistakes and build the internal judgement this paper keeps returning to, is now, while someone else's capital is paying most of the bill. The organisations that wait for the market to stabilise before engaging seriously will be learning at post-correction prices what their competitors learned at Frenzy prices. The window is real, and it is not permanently open.

None of this means the correction, when it comes, is the end of the story. Looking back across Perez's cases, the Frenzy begins to look less like a mistake and more like a necessary feature of how humanity achieves technological leaps that rational capital allocation would never sanction. The numbers never added up for the railways, for electrification, or for the internet. A financially rational decision-maker, applying standard return-on-investment criteria, would have killed each of those technologies at the point of maximum investment. The irrational frenzy is the mechanism by which we build the infrastructure that makes the impossible happen. The havoc it creates in the meantime is the price. And the Deployment Period that follows a genuine Turning Point has, historically, been more productive than the Frenzy that preceded it. The correction is not a reason to disengage. It is a reason to prepare, and the rest of this paper is about what preparing actually means.

The Limits of the Optimist Case

The optimist narrative rests on two claims that deserve more examination than they usually get.

The first is that AI capabilities are growing exponentially and will continue to do so. Václav Smil's work on how technological capabilities actually develop offers a useful corrective. In Growth (Smil, 2019), he demonstrates that what appears to be exponential growth in any physical system is always a temporary phase. Every technology traces an S-curve: slow initial growth, rapid acceleration, then a plateau imposed by physical or systemic constraints. The recurring mistake, Smil observes, is the failure to recognise the temporary nature of the exponential phase, which leads to the assumption that a technology advancing rapidly will continue to do so indefinitely. It is an easy mistake to make from inside the acceleration, because from inside the acceleration every data point confirms it. The plateau announces itself only in retrospect.

The honest counterargument is that progress can bypass an S-curve ceiling by leaping to a new curve. The transition from vacuum tubes to transistors to microchips is the canonical example: each generation hit its physical limit and was succeeded by a new paradigm rather than a full stop. The AI industry is already attempting this, moving from pure text-based scaling to agentic architectures, multimodal systems and inference-time reasoning. So the

argument is not that AI capabilities will stop advancing; I make no such claim and the evidence would not support it. The argument is narrower. Indefinite exponential growth within any given architecture is not supported by the history of how technologies develop, and each leap to a new curve is a bet, not a schedule. Bets sometimes pay. But an organisation that has built its three-year plan on the assumption that a particular bet pays, on a particular timeline, has confused a research programme with a delivery roadmap. Organisations planning around the assumption that the current rate of improvement is a law of nature are building on an unreliable foundation.

The second claim, made with particular confidence in certain quarters, is that current large language model architectures are sufficient to produce artificial general intelligence. This rests on a philosophical assumption that receives less scrutiny than it deserves: that language is an adequate representation of the world.

Language is an extraordinary human achievement and a powerful cognitive tool. It is also a thin and lossy representation of physical reality. The embodied, sensory, contextual knowledge through which most human understanding is constructed is largely absent from what a language model can access. Michael Polanyi made the point half a century before it became commercially relevant: we know more than we can tell (Polanyi, 1966). The experienced engineer who senses that a design is wrong before she can articulate why, the nurse who knows a patient is deteriorating ahead of the monitors, the negotiator who reads the room and changes tack: these are not decorative flourishes on top of explicit knowledge. They are the substance of expertise, and almost none of it has ever been written down, because it cannot be. A model trained on everything we have managed to tell is still missing everything we cannot. Different languages do not even carve up experience in the same way. The proposition that a sufficiently large model trained on sufficiently large amounts of text will approximate general intelligence is constrained by the physics of the world long before it is constrained by the engineering. This is not a criticism of what the models can do within their domain. It is a recognition that the domain is more bounded than the most ambitious claims suggest.

Brian Friel understood this problem long before large language models existed. Translations (Friel, 1981) turns on exactly it: English surveyors mapping Ireland by replacing Irish place names with English equivalents, losing in the process the physical descriptions, local histories and sensory knowledge the original names carried. The new map is accurate in its own terms. It does not describe the same landscape. A language model faces a structurally similar constraint. The corpus, however vast, is not the world.

These limits matter for the argument that follows, because they explain why the technology behaves as an amplifier rather than an oracle. A system that holds a

lossy map of the world cannot supply the judgement, the context or the tacit knowledge that the map leaves out. Those have to come from the person using it. Which raises the question of what happens when they do, and when they do not.

The Amplification Principle

At the trivial end of the range, AI assistance works for everyone. A shopping list, a recipe, a training plan, troubleshooting your boiler. These tasks require no special expertise from the user, the cost of error is low, and the value, while real, is modest. Nothing about the amplification principle applies down here, which is precisely why the experience of using AI for trivial tasks is such a poor guide to what happens when the stakes rise. A great deal of executive conviction about AI has been formed on exactly this terrain: the chief executive who had a remarkable weekend conversation with a chatbot about his golf swing and arrived on Monday convinced the technology could transform the supply chain. The tool that impressed him operates under entirely different conditions from the tool he is about to fund.

Beyond the trivial, a different rule takes over. AI amplifies what you already have. It does not manufacture what you do not.

I can illustrate this from my own practice, because I live on both sides of the line. My work is analytical: reading organisations, structuring problems, building arguments. Applied to that work, AI tools extend my capacity and speed to a degree I would not have believed three years ago. The same tools, pointed at visual design, where my expertise is close to nil, produce work that is confidently, fluently poor. My last attempt at an AI-assisted client graphic landed like a piece of furniture dropped from a height. The tool had not failed. It had amplified exactly what I brought to the task, which in that domain is very little. A designer colleague, using the identical tool, produces work in minutes that would have taken her a day. When I try the same, I amplify mediocrity at scale. The tool is identical. The person using it is everything.

The old rule of computing has not been repealed by generative AI; it has been supercharged. Rubbish in, rubbish out, as it always was. What has changed is the packaging. The rubbish now arrives fluent, formatted and confident, which makes it harder to detect and easier to pass along. Generative AI has made plausibility free, and plausibility was the signal we used to rely on to detect competence. For most of professional history, producing a coherent, well-structured document about a subject required actually knowing something about it. The document was evidence of the knowledge. That evidentiary link has now been severed, quietly and completely, and most organisational quality processes have not yet noticed. They are still reading fluency as competence, because for the entire working lives of the people who built those processes, it was.

That is why the variance in AI output quality between users in the same organisation, using the same tools, is far larger than anyone expects going in. In my work with client organisations, that variance is the single most consistent observation, and it is not explained by technical sophistication. The best prompt engineer in the building is rarely the person producing the most valuable AI-assisted work. The variance is explained by the depth of the underlying domain expertise and, critically, by whether the individual knows where that expertise ends. The practitioners who get the most from these tools are the ones who point them at their real strengths and are honest enough to recognise the boundary. The ones who get the least are those who take fluent output in a domain they cannot judge and mistake it for competence. I am guilty as charged myself; I still catch myself reaching for the tool in domains where I have no way of knowing whether what comes back is any good. The discipline of not doing that turns out to be one of the harder personal skills of the AI era, because the tool never tells you no. It always produces something, and the something always looks finished.

I should name the obvious objection to building an argument on my own experience: it may be my own bias reflecting back at me. A consultant whose strength is analysis, arguing that AI rewards people who know their strengths, is suspiciously convenient. Two things give me some confidence the principle generalises. The first is that the mechanism is not mysterious. These are prediction systems whose output quality is highly sensitive to the quality and specificity of what the user brings: the framing, the context, the ability to recognise a wrong answer. Expertise is exactly what supplies those things. The second is that the pattern repeats across every organisation I have worked in, at every level of seniority, in domains far from my own. So: high confidence in the mechanism, moderate confidence in how far it generalises beyond the kinds of knowledge work I can observe directly.

The principle scales, but not automatically. At the level of a small team, the strongest results I have seen follow a consistent shape: a small group, each member bringing real specialist expertise in a different domain, each using AI to amplify their own strength, working a bounded and well-defined problem with a shared standard of rigour around how the tools are used. Where those conditions hold, the outcomes can be remarkable, because the amplification compounds across complementary strengths. The analyst's structured thinking sharpens the brief the designer works from; the designer's amplified output gives the engineer something concrete to build against; each contribution arrives at the next desk better than it would have unaided, and the improvement carries forward. Where any of the conditions is missing, the results are consistently disappointing, and the missing condition is usually the shared standard of rigour rather than the expertise. This observation is drawn from practice, not from published evidence, and I state it as such.

At the level of the organisation, the principle becomes a question about self-knowledge, and the question has to be answered in three places at once. Senior leadership understands competitive advantage in strategic terms: what the market rewards, where the moat is, what the company is for. Middle management understands it in operational terms: which processes actually deliver the advantage, where the exceptions live, which shortcuts are safe and which are load-bearing. The translation layer that converts strategic intent into usable practice matters more in an AI transition, not less, and the second paper in this series takes that argument up in full. Frontline workers understand advantage in the specific, practical terms of daily work: what the customer actually asks for, which promises the organisation actually keeps, where the real quality is made. All three understandings are necessary, and they have to agree. An organisation whose leadership believes the advantage is speed, whose managers run processes built for consistency, and whose frontline quietly knows the customer stays for the relationship, will point the amplifier at three different targets and wonder why the output is noise. Many organisations believe they understand their competitive advantages. Far fewer have tested that belief against the question of what, specifically, they are asking AI to amplify, and whether the answer is the same at every level. Where the layered understanding exists, AI produces compounding value. Where it does not, AI produces activity.

Two Answers to the Same Question

The clearest public illustrations of the amplification question answered badly and well are Klarna and IKEA. The two businesses are not comparable, one a digital-native payments company and the other a century-old physical retailer, and I am not claiming their situations were symmetrical. The relevance is that both faced the same foundational question and answered it differently.

Klarna deserves a more careful reading than it usually receives, because it has been flattened into a morality tale by both sides of the AI argument. In February 2024 the company launched an AI customer service assistant which, by its own account, handled 2.3 million conversations in its first month, equivalent to the work of roughly 700 agents, resolving queries in minutes rather than the eleven-minute human average (Klarna, 2024). Over the same period the company shrank its workforce from around 5,500 to around 3,400, largely through attrition under a hiring freeze, and reported revenue per employee rising 73%. These figures are Klarna's own, published in service of a story the company was actively telling investors ahead of a public listing, and they should be read with that in mind. But by any conventional efficiency measure the deployment worked, and the chief executive was entitled to say so publicly. These were real gains, achieved faster than almost any comparable organisation had managed.

The problem became visible later, when customer satisfaction data showed that what had been optimised was not what the business most depended on. Klarna's competitive advantage is frictionless experience; the smoothness and reliability of the transaction is what built the company. When its own data showed satisfaction declining and automated responses being described as generic and insufficient for complex issues, the company acknowledged it and acted. By May 2025 it was recruiting human agents again and committing publicly to customers always having the option of a person. Sebastian Siemiatkowski's admission to Bloomberg was direct: "As cost unfortunately seems to have been a too predominant evaluation factor when organising this, what you end up having is lower quality" (Bloomberg, 2025). A company willing to reverse a public position when its own data demands it is a company paying attention, and Klarna deserves more credit for the reversal than it tends to get. But the recalibration was only necessary because the foundational question was never clearly answered before deployment began. What, specifically, was Klarna amplifying? The answer should have been frictionless experience. The deployment optimised for cost instead. The initial metrics looked compelling precisely because they were measuring the wrong thing.

IKEA's parent answered the same question the other way around. In 2021, Ingka Group, the largest IKEA franchise operator, deployed an AI assistant called Billie to handle routine customer service traffic; by the company's account it resolved roughly 47% of the enquiries it received between 2021 and 2023 (Ingka Group, 2023). Ingka faced the same arithmetic Klarna faced: a large support workforce whose transactional volume a machine could now absorb. What it did differently was identify what that workforce actually had. Years of handling customer queries had built a deep, accumulated familiarity with the product range and with how customers actually live with it. That expertise, Ingka concluded, was worth amplifying rather than shedding. The company reports having reskilled 8,500 call-centre co-workers into remote selling roles built around that knowledge: interior design consultation, digital retail sales, the complex judgement-laden enquiries a chatbot cannot handle. The machine absorbed the transactional volume. The people were redeployed onto the judgement. Ingka's remote selling channel, which those co-workers staff, reached 1.3 billion euro in sales in its 2022 financial year, around 3.3% of the total.

The caveats matter here too. Every figure in that paragraph comes from Ingka's own communications, published as part of a corporate narrative about responsible AI adoption, and the neat framing of 8,500 people retrained into a new revenue stream is doing public relations work as well as descriptive work. The reskilling was into a set of remote-selling competencies, not a uniform transformation of every worker into a design consultant, and the remote channel's revenue is not attributable to the reskilling alone. I use the case anyway because the structural point survives the discounting. Before deploying the

technology, Ingka asked what its people had that was worth amplifying, found a real answer, and built the deployment around it. That is the question IKEA answered and Klarna did not: what are we actually amplifying here?

The same logic explains a less-noticed finding in the MIT NANDA study, subject to the same caveats as the headline figure: externally purchased specialist tools succeeded in generating measurable return roughly twice as often as bespoke internal builds (MIT NANDA, 2025). Part of that is selection. An organisation that commits to a narrow specialist tool tends to be one with a clear view of the specific problem it is solving; the tool is narrow enough that the amplification question almost answers itself. Harvey, the legal-domain AI company, reached a reported $300 million in annual recurring revenue within four years of its founding (Sacra, 2026) by applying AI to a narrow, well-understood domain with deep specialist knowledge on both sides of the deployment. The question of what was being amplified, the accumulated expertise of lawyers on well-defined legal work, had an answer before any of the technology was built. That is not a coincidence. It is also not an argument that buying beats building as a general law; it is an argument that clarity beats ambiguity, and that narrow tools force a clarity that broad platforms allow you to defer. The deferral is the expensive part.

Why Transformation Lags the Technology

If the amplification principle is right, organisational value depends on organisational self-knowledge, and self-knowledge takes time. History suggests just how much time, and why, and the why matters more than the how much.

The economic historian Paul David studied the introduction of electric motors to manufacturing, one of the most instructive productivity paradoxes on record (David, 1990). Electric power became widely available to American factories in the 1890s. The expected productivity gains did not show up in the statistics until the 1920s, a lag of nearly thirty years. The reason was not simply that learning takes time. It was more specific. Factories that replaced their steam engines with electric motors but changed nothing else saw almost no gain, because they retained the physical and organisational logic of steam: a central power source driving a system of belts and shafts that imposed a single fixed layout on the factory floor. Every machine's position was dictated by its distance from the drive shaft, not by the sequence of the work. The whole factory was, in effect, an architectural expression of its power source.

The gains came only in factories that understood electrification not as a substitution but as an architectural opportunity. Distributed power meant a small motor at every machine, which meant the floor could be reorganised around the logic of the work rather than the logic of the power source. Materials could flow in the sequence the product required. New layouts, new workflow sequences, new roles, new supervision structures; even the buildings changed, from multi-

storey mills built to minimise shaft length to single-storey plants built for flow. None of that redesign was visible in the technology itself. You could not have derived the modern factory from a specification sheet for an electric motor. The technology made the gain possible. The organisational redesign realised it, and the redesign took a generation, because it required managers to unlearn assumptions so deep they did not know they were assumptions.

The application to the present is direct. Organisations treating AI as a substitution, the same processes with a model bolted on where a person used to be, are the steam-logic factories of this cycle. They will see the technology working and the productivity nowhere. The organisations that will capture the value are the ones redesigning the work around what the technology actually makes possible, and that redesign is gated on exactly the self-knowledge this paper has been describing: you cannot reorganise around your strengths until you know what they are. David's thirty-year lag will not, I think, take thirty years this time; diffusion is faster now and the technology needs no physical rewiring. But the shape of the lag, substitution first, disappointment, then redesign, then gains, I expect to hold. High confidence in the shape. Low confidence in the duration.

A Non-Robust Technology in a Robust World

Electrification's redesign was, at least, a redesign around a technology that behaved itself: the motor did the same thing on Tuesday as it had on Monday. Two further features of the current technology make this cycle's redesign harder, not easier. They deserve separate treatment, because they are routinely conflated and they fail differently.

The first is that generative AI is, by its nature, a non-robust technology. Its outputs are probabilistic, variable, and highly sensitive to the quality and specificity of the inputs. The same prompt, run twice, can produce materially different results. This is not a defect awaiting a patch; it is intrinsic to how the systems work, and it is inseparable from what makes them useful. The flexibility that lets a model draft a contract clause, summarise a meeting and brainstorm a product name is the same property that makes its behaviour impossible to guarantee.

Enterprise organisations, by contrast, are built for robust technology. Their governance structures, quality management systems and risk frameworks all assume that enterprise technology behaves consistently and predictably: the same input produces the same output, the system either works or fails visibly, and a test passed once stays passed. Those assumptions are so foundational that most organisations have never had to articulate them. An electric motor, whatever else it demanded, at least behaved the same way on Tuesday as it had on Monday. A non-robust technology dropped into an environment built for robust

ones, and assessed against metrics designed for robust ones, produces predictably poor results. The failure is in the fit, not the tool.

The organisations getting this right have adapted the implementation to the nature of the technology rather than forcing the technology into the shape of their existing frameworks. In practice that means distinguishing, task by task, between work that requires deterministic, fully auditable outputs, where generative approaches are the wrong tool however impressive the demo, and work that can benefit from probabilistic assistance provided appropriate verification sits around it. It means building oversight structures proportionate to each type of use rather than a single governance regime that either strangles the useful applications or waves through the dangerous ones. Most organisations currently have one of those two failure modes, and a surprising number have both at once, in different departments.

Expecting the Dip

If non-robustness explains why the redesign is hard, the second feature explains why the redesign looks like failure precisely while it is working. It is temporal rather than structural. The early phase of any serious redesign registers as a productivity decline before it registers as a gain. New capabilities have to be aligned with old processes; genuine learning takes time that shows up nowhere in the output metrics; new workflows impose friction until they become habitual. Economists have documented this pattern across general-purpose technologies and named it the Productivity J-Curve: measured productivity dips as organisations invest in the intangible capital, the process redesign, the training, the accumulated organisational learning, that the technology requires, and only later climbs past its starting point as those investments pay out (Brynjolfsson, Rock and Syverson, 2021). The intangible investment is real investment, but the accounting cannot see it, so the early years of a transformation look like money going out and nothing coming back.

The organisational consequences of not knowing this are severe, and I see them regularly. An organisation that expects the dip treats it as the cost of the redesign: budgeted, monitored, survivable. An organisation that does not expect it panics when the dip appears, and the panic takes one of two forms, both destructive. Either the initiative is abandoned, at the exact point where the investment has been made and the return has not yet arrived, which is the most expensive possible moment to quit. Or leadership tries to accelerate through the dip by scaling faster, on the theory that the problem is insufficient commitment, which deepens the dip by multiplying the unabsorbed change. The question worth asking before either response is the simple one: are we running before we can crawl?

The dip also interacts badly with the correction I expect from the financial cycle. When the market resets, AI initiatives across the economy will be sitting at various depths of their J-curves, and the ones that cannot show returns will be indistinguishable, on a spreadsheet, from the ones that were never going to. A great deal of genuinely sound transformation work will be cancelled in that moment, alongside the genuinely unsound. The organisations that have measured their intangible investment honestly, and can say where they are on the curve and why, will be the ones able to defend the work when the cost-cutters arrive. That defence has to be built now, in the design of the initiative, not improvised later in the budget review.

The Workflow Problem

Everything to this point has an implicit unit of analysis: the individual, amplified. The hardest problem appears when AI-assisted work has to move between people, and it is the problem on which the amplification principle, as stated so far, runs out.

When AI-assisted work passes from one person to another, the receiving person faces a question that did not used to need asking: how was this made? If they do not know what was asked of the model, what was verified, what the human contributor added versus what the machine generated, and where the quality controls were applied, they cannot extend the work with confidence. They must verify it from the beginning before they can build on it. The efficiency gain the first person achieved is consumed by the verification cost imposed on the second. Multiply that across a workflow and the organisation's net gain from the technology rounds to zero, even while every individual in the chain reports that the tools are making them faster. This is, I suspect, a significant part of the explanation for the failure statistics: the individual gains are real, the organisational gains are not, and the difference disappears into the seams between people.

This is not hypothetical. It is the dominant pattern in organisations that have moved past individual experimentation and tried to embed AI in collaborative workflows, and I see it weekly. Individuals use AI in their own way, to their own standards, with their own levels of rigour, generating copious amounts of AI slop alongside the occasional excellent output, with nothing on the surface to distinguish the two. The work gets passed to colleagues with no shared understanding of how it was made. The colleagues either spend their time unpicking it, negating the gain, or accept it on trust and compound whatever defects it carries. A tradesman who handed over work in this state would not stay in business; the morass of nonsense that gets handed between knowledge workers under the banner of AI-assisted productivity would shame a plumber.

There is a deeper layer than process transparency, and it connects directly to the amplification principle. Trust in handover is also trust about purpose. If I do not know what you were trying to amplify when you produced this work, I cannot know whether you succeeded. I cannot extend your strength, because I do not know what strength I am extending. The shared context that makes collaboration productive has always been partly tacit; in an AI-assisted workflow it has to be made explicit and actively maintained, because the artefacts themselves no longer carry reliable evidence of the thinking that produced them. Plausibility used to be that evidence. Generative AI has made plausibility free. This is also where the translation layer of the organisation earns its keep: converting strategic intent into shared working standards is precisely the kind of work that determines whether handover trust exists, and the second paper in this series examines why that layer, currently being thinned in the name of efficiency, is the wrong place to cut.

I can describe what the solved state looks like, because I have seen approximations of it in small teams: a shared understanding of how AI is used, common standards for how work is produced and described, explicit handover expectations, human oversight at agreed points. In the teams that manage it, the practices are unglamorous. A note travels with the work saying what was asked of the tool and what was checked by a person. Sections that came from a model are marked as such, with a line on what was verified. None of this is bureaucracy for its own sake; it is the minimum context a colleague needs to build on your work rather than rebuild it. When AI-assisted work transfers between people with genuine confidence, value stops leaking at every seam and starts to compound; each contribution builds on the last, and the cumulative output reflects not just the tools but the compounding expertise of the team. That compounding is the flywheel in this paper's title, and it is the difference between AI as a collection of personal productivity gains and AI as an organisational capability.

It is worth being honest about why the solved state is rare, because the reasons are structural rather than technical. The practices that create handover trust are individually costly and collectively valuable, which is the classic shape of a commons problem. The person who documents their process, marks their AI-assisted sections and verifies before passing work along pays the cost personally, in time, and the benefit lands on someone else's desk. In an organisation that measures individual output, every incentive points toward producing more and describing less. The fluent, unverified document and the rigorous, verified one look identical in the metrics, and the fluent one arrived faster. Until the measurement systems can tell the difference, the rational individual behaviour will keep degrading the collective asset. That is not a problem a tool can fix. It is a problem of what the organisation chooses to reward, which is to say a problem of management, which is to say exactly the kind of problem the current enthusiasm for technology keeps stepping around.

But describing the solved state is not solving it. What I have not given you, and will not pretend to give you in a closing paragraph, is the mechanism: how an organisation actually builds transferable trust in AI-assisted work at scale, across teams that do not share a room, standards that do not yet exist, and incentives that currently reward volume of output over verifiability of output. That is a design problem in its own right, and it is the subject of the papers that follow. Here I want only to state it as sharply as I can, because naming it precisely is the precondition for solving it. The binding constraint on organisational AI value is not model capability, and it is not implementation methodology. It is whether one person can trust another person's AI-assisted work enough to build on it without redoing it. Until an organisation can answer yes, the amplifier works one desk at a time, and the flywheel does not turn.

Digging

The moment we are in is genuinely disorienting, and it is worth saying why. The old ways of working are visibly insufficient; the new ways are not yet formed; and in the gap between them the symptoms multiply: the wild claims, the misallocated capital, the transformation initiatives that generate activity without value, the fear and the euphoria that make clear thinking difficult. A correction is coming that will be read, wrongly, as the technology failing. The technology is not failing. It is being asked the wrong question.

The temptation, in a moment like this, is to resolve the discomfort prematurely: to pick a camp, buy a narrative, and stop thinking. Both the boosters and the doomers are offering exactly that relief, and the relief is worthless. I have come to believe the opposite of what the moment invites: that uncertainty, held carefully, is actually the most productive place to think from right now. Held carefully: that is the operative phrase. Not uncertainty as paralysis, and not uncertainty as an excuse to defer engagement, but uncertainty as a discipline, the deliberate refusal to claim more confidence than the evidence supports, paired with the willingness to act anyway on what can be established. Every confidence level attached to every claim in this paper is an attempt to practise that discipline in public.

What the evidence and the history suggest together is that the organisations navigating this moment well are not distinguished by the boldness of their ambitions or the scale of their investment. They are distinguished by knowing, honestly and precisely, what they are good at, and by having that knowledge held in common from the boardroom to the frontline. The amplification principle is not a modest qualification to the AI story. It is, I think, the central fact of it. AI is an extraordinary amplifier, and what it amplifies is entirely a function of what you bring to it. A company that understands its advantage and builds AI on top of that understanding is building a flywheel. A company that deploys AI reactively, out

of fear or competitive pressure, is building a more expensive version of what it already had.

Seamus Heaney opened his first collection with a poem about watching his father dig potatoes with the same steady skill his grandfather had used to cut turf in the bog. Heaney was sitting at a window, holding a pen. The final line is three words: "I'll dig with it" (Heaney, 1966). He was not pretending to be his father, and he was not abandoning the tool he had. He recognised what his people knew how to do, and found the way to do the same essential thing in his own domain, with his own instrument.

That is the whole argument, held in a single image. The technology is real. The potential is real. The timeline is longer and the path more demanding than the noise suggests, and the hardest problem, trust in the handover, is still open on my desk. Know what you are amplifying. Find the tool that extends it honestly. Then dig.

A Note on Sources

The claims in this paper are not all of the same strength, and the differences should be visible rather than smoothed over.

The MIT NANDA figures, both the 95% headline and the comparison of external tools with internal builds, come from a preliminary study that has not been peer reviewed and whose methodology has been contested. I use them as directional evidence only, and the argument does not depend on their precision. Bain's two trillion dollar revenue requirement is a scenario estimate, not a measurement, produced by a firm with a commercial stake in the AI conversation. The related claim that current inference pricing is subsidised is informed inference from the visible capital structure, not a published figure; the providers do not disclose their unit economics. The Klarna figures, including the workforce reduction from roughly 5,500 to roughly 3,400 and the revenue-per-employee gain, are the company's own, published while it was telling an efficiency story to investors; the reversal is documented in Siemiatkowski's public statements. The IKEA figures come entirely from Ingka Group's corporate communications and carry the framing incentives of that genre; I have relied on the structural fact of the reskilling programme rather than the precision of its numbers. Harvey's revenue figure is drawn from industry trackers reporting company data, not audited accounts. The Productivity J-Curve is peer-reviewed economics and the strongest published evidence in the paper; the electrification history is a well-established scholarly account; Perez's framework is a widely used interpretive lens rather than a predictive law, and I have used it as the former.

The amplification principle itself, and the observations about variance between users, small-team conditions, and the handover problem, are drawn from my own

practice at Trinzo rather than from published evidence. They are consistent with the mechanism of the technology, and they repeat across the organisations I see, but a consultant's sample is not a dataset, and it may be my own bias reflecting back at me. I hold high confidence in the mechanism of amplification and moderate confidence in how far it generalises. The paper makes two predictions. A market correction: high confidence that one arrives, moderate confidence that it lands around 2027. The shape of the adoption lag following the electrification pattern: high confidence in the shape, low confidence in the duration. Where either prediction fails, I would rather have been clearly wrong than vaguely right.

References

Bain & Company. Global Technology Report 2025. Bain & Company, September 2025.

Bloomberg News. Interview with Sebastian Siemiatkowski. Bloomberg, 8 May 2025.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson. "The Productivity J-Curve: How Intangibles Complement General Purpose Technologies." American Economic Journal: Macroeconomics 13, no. 1 (2021): 333-372.

David, Paul A. "The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox." American Economic Review 80, no. 2 (1990): 355-361.

Friel, Brian. Translations. Faber and Faber, 1981.

Heaney, Seamus. "Digging." In Death of a Naturalist. Faber and Faber, 1966.

Ingka Group. "AI and Remote Selling bring IKEA design expertise to the many." Ingka Group Newsroom, 29 June 2023.

Klarna. "Klarna AI assistant handles two-thirds of customer service chats in its first month." Press release, February 2024.

MIT NANDA Initiative. The GenAI Divide: State of AI in Business 2025. MIT, August 2025.

Perez, Carlota. Technological Revolutions and Financial Capital: The Dynamics of Bubbles and Golden Ages. Edward Elgar, 2002.

Polanyi, Michael. The Tacit Dimension. University of Chicago Press, 1966.

Sacra. "Harvey at $300M ARR." Sacra Research, 2026.

Smil, Václav. Growth: From Microorganisms to Megacities. MIT Press, 2019.