

BEYOND OUTPUT VELOCITY

The Productivity Paradox of Artificial Intelligence

A Practitioner's Perspective on Why Faster Outputs Have Not Yet Produced Better Decisions

Giovanni Leonardi

March 2022



Contents

1. Executive Summary	3
2. An Old Paradox in New Clothing	4
3. Why an Output Is Not a Decision	4
4. The Structural Forces That Sustain the Paradox	5
5. What Acceleration Reveals	6
6. From Output Velocity to Decision Velocity	7
7. A Closing Reflection	8

The Productivity Paradox of Artificial Intelligence

Transformation · No. 332

Originally written March 2022 | Re-edited and re-rendered for the WhereBeyond Edition, 2025

"We have automated the production of analysis without touching the act of judgement, and then expressed surprise that the second did not follow the first."

1. Executive Summary

Something strange is unfolding in the organisations that have begun, tentatively, to fold artificial intelligence into their daily work. The machines are unarguably faster. A demand forecast that once occupied an analyst for a fortnight now arrives before lunch. A first draft of a paper, a block of working code, a customer segmentation, a summarised contract — each can now be produced at a speed that would have seemed fanciful only a few years ago. And yet, when one looks past the outputs to the decisions those outputs exist to serve, remarkably little has moved. The same choices are made, at roughly the same cadence, with roughly the same quality, by roughly the same people. We have accelerated the manufacture of analysis without accelerating the act of judgement.

This essay is an attempt to sit with that gap rather than explain it away. The pattern I have watched recur across sectors is not that AI fails to deliver — the outputs are real, and often excellent — but that the value leaks out somewhere between the output and the decision it was meant to improve. The acceleration is genuine and the paradox is genuine, and they are not in contradiction. My argument is that faster outputs and unchanged decisions coexist because the two live in different parts of the organisation: one in the production layer, which AI touches directly, and the other in the decision layer, which AI barely reaches. Transformation programmes have poured investment into the first while assuming the second would follow. It has not, and the reasons it has not are structural, not technical.

- The productive capacity of the analytical function has risen sharply; the decision capacity of the organisation has not.
- The bottleneck was never the speed of producing analysis — it was the organisation's willingness and ability to act on it.
- Acceleration does not create direction. It merely lets us arrive at the same destination sooner, or reach the wrong one faster.

What follows examines why this paradox persists, what structural forces sustain it, and what it asks of anyone who believes that adopting these tools is the same thing as transforming.

2. An Old Paradox in New Clothing

We have been here before, and it is worth remembering that we have. In 1987 the economist Robert Solow made his famous, deflating observation that the computer age was visible everywhere except in the productivity statistics. Firms had spent enormously on information technology through the 1970s and 1980s, and yet measured productivity growth had, if anything, slowed. It took the better part of two decades, and a great deal of subsequent scholarship, to understand what had happened: the technology had arrived long before the organisational and human changes required to convert it into value. The machines were installed; the ways of working around them were not rebuilt. Productivity eventually came, but it came late, and it came only to those who did the unglamorous work of reorganising around the new capability.

The artificial intelligence of the present moment is producing its own version of the Solow paradox, and it is producing it faster. The tools available today — the large language models exposed through APIs, the machine-learning pipelines that predict and classify, the automation platforms that stitch tasks together — can be adopted in weeks rather than the years it once took to roll out an enterprise system. This is precisely what makes the paradox sharper and more disorienting. The lag between capability and value has not vanished; it has simply become more visible, because the capability now arrives so quickly that the absence of the accompanying value is impossible to disguise as a matter of time.

The lesson of the last productivity paradox was not that the technology was overrated. It was that technology and organisation move at different speeds, and value accrues only to those who close the gap between them. Nothing about faster tools changes that lesson — it only shortens the time we have to learn it.

The temptation, when confronted with a tool that produces good outputs quickly, is to assume that the hard part is done. History suggests the opposite. The arrival of the capability is the easy part. The hard part — the part that separates the organisations that will realise value from those that will merely accumulate impressive demonstrations — is everything that has to change around the capability for its outputs to alter a single real decision.

3. Why an Output Is Not a Decision

At the root of the paradox is a category error, and it is a seductive one. We treat the production of an analysis and the making of a decision as though they were the same activity, or as though the first reliably causes the second. They are not, and it does not.

An output is an artefact: a forecast, a recommendation, a draft, a model score, a summary. A decision is a commitment — a choice to act, to allocate resource, to accept risk, to say yes to one path and no to the others. Between the two sits a great deal of distinctly human and organisational machinery: the judgement to interpret the artefact, the authority to act on it, the confidence to stake something on it, the alignment to bring

others along, and the accountability to own the outcome. AI, as it stands, is extraordinarily good at manufacturing the artefact. It touches almost none of the machinery that turns the artefact into a commitment.

Consider what actually happens when a model produces a faster, sharper forecast. The forecast lands on a desk. And then it waits. It waits for a meeting that is scheduled for next Thursday. It waits for a manager who does not fully trust a number she cannot trace back to an assumption she understands. It waits for three other functions to be consulted, for a governance forum to convene, for a budget cycle to open. The forecast was produced in an hour; the decision it informs takes six weeks, exactly as it always did. The output accelerated. The decision did not, because the output was never the thing slowing the decision down.

“The forecast was produced in an hour. The decision it served still took six weeks, because the output was never the thing that had been slowing the decision down.”

This is the heart of the matter. In most organisations, the rate-limiting step in decision-making has never been the speed of producing the underlying analysis. It has been everything else: the coordination, the trust, the authority, the willingness to commit under uncertainty. By pouring capability into the one step that was never the constraint, we have optimised the fast part of the process and left the slow part untouched. The queue simply forms in a different place. We have, in the language of the operations theorists, elevated a non-bottleneck — and elevating a non-bottleneck produces no gain in throughput at all. It produces only a larger pile of work-in-progress waiting at the true constraint.

4. The Structural Forces That Sustain the Paradox

If the gap were merely a matter of oversight, it would close on its own as organisations noticed it. It does not close, which tells us that something is actively holding it open. Several forces, in my observation, conspire to sustain the paradox.

1. The measurement of the wrong thing. Programmes adopting AI are, almost universally, measured on activity and output. Number of use cases live. Hours saved in producing a report. Volume of documents summarised. These are the metrics that are easy to gather and flattering to present, and they are all measures of the production layer. Almost no one measures whether decisions improved — whether they were made faster, or better, or with more confidence — because that is genuinely hard to measure and rarely comfortable to confront. What is measured is what is managed, and we are managing output.
2. The comfort of the visible. An accelerated output is a demonstration. It can be shown to a board, projected on a screen, admired. The reorganisation of decision rights, the

rebuilding of trust in a number, the redesign of a governance forum so that it moves at the speed of the analysis feeding it – these are invisible, slow, and politically fraught. Faced with a choice between a visible win in the production layer and an invisible slog in the decision layer, organisations choose the visible win, and then wonder why the paradox persists.

3. The absence of trust in the artefact. A faster output is only useful if a decision-maker will act on it, and action requires trust. Many of the new outputs are produced by methods their consumers cannot interrogate. A manager who does not understand how a number was reached, and who will be held accountable for the consequences of acting on it, will quite rationally slow down, seek corroboration, and fall back on her own judgement. The output arrives faster; it is trusted no faster, and often less. Speed of production and speed of acceptance are different variables, and we have moved only the first.
4. Decision latency as an organisational property. The time an organisation takes to convert insight into commitment is not an accident. It is a product of its structure: how many people must agree, how risk is owned, how failure is punished, how authority is distributed. These are deep properties, set by culture and incentive, and no analytical tool touches them. An organisation that took six weeks to decide before AI will take six weeks to decide after it, because the six weeks were never about the analysis.

Taken together, these forces explain why the paradox is so stubborn. It is not sustained by ignorance, which would be curable, but by structure and incentive, which are not curable by better tools. The tools address the production layer. The paradox lives in the decision layer. The two are, for now, largely disconnected.

Layer	What AI changes	What actually governs it
Production of analysis	Dramatically faster and cheaper	Model quality, data access, tooling
Decision to act	Almost nothing	Trust, authority, risk ownership, alignment
Realised value	Nothing directly	Whether the second layer moves at all

5. What Acceleration Reveals

There is a more hopeful way to read the paradox, and I want to hold it up against the diagnostic one. A tool that accelerates outputs is, whatever else it is, a superb diagnostic instrument. By removing the production of analysis as a constraint, it exposes with unusual clarity where the real constraints have been hiding all along.

For years, organisations could tell themselves a comforting story: we would decide faster and better if only we had the analysis sooner. AI calls that bluff. It delivers the analysis sooner, and the decisions do not improve, and now there is nowhere left to hide. The slowness was never in the analysis. It was in us – in our governance, our trust, our willingness to commit. The acceleration of outputs is, in effect, a floodlight thrown onto the decision layer, revealing every place where insight goes to wait.

- Where analysis now arrives instantly but decisions still stall, we can see exactly which forum, role, or approval is the true constraint.
- Where a faster output is met with the same request for corroboration, we learn precisely where trust in the organisation's own analysis is missing.
- Where abundant, cheap analysis produces no more decisions than scarce, expensive analysis did, we learn that the organisation's capacity to absorb insight — not to produce it — was always the ceiling.

An organisation willing to read these signals honestly is handed something valuable: a precise map of where its decision-making is actually constrained. This is the reframe I want to argue for. The productivity paradox of AI is not a failure of the technology. It is a message from the technology about the organisation, and the organisations that receive the message will pull ahead of those that keep investing in the layer that was never the problem.

6. From Output Velocity to Decision Velocity

If the diagnosis is right, the implication is uncomfortable, because it points the work away from the tools and back at ourselves. The task is not to produce outputs faster still — we have already solved that, almost embarrassingly well. The task is to raise the velocity and quality of decisions, and that is work of a completely different character.

What would it mean to treat decision velocity as the thing to be transformed? A few directions follow from the argument, offered not as a template but as the shape of the problem.

1. Measure decisions, not outputs. Ask not how many analyses were produced but how many decisions were made, how quickly they moved from insight to commitment, and whether their quality improved. If this is hard to measure — and it is — that difficulty is itself a finding about how little attention the decision layer has ever received.
2. Rebuild trust in the artefact deliberately. An output that cannot be trusted cannot be acted on quickly, however fast it was produced. Traceability, explanation, and the deliberate construction of confidence are not niceties; they are the mechanism by which production speed is allowed to become decision speed. Trust is the transmission between the two layers, and it has to be engineered.
3. Redesign the decision machinery to match the analysis feeding it. If analysis now arrives in an hour, a governance forum that convenes monthly is the constraint, and no better model will fix it. The forums, the approval chains, the distribution of authority — these have to be redesigned to move at something closer to the speed of the insight, or the insight will simply queue.
4. Locate and elevate the true constraint. The discipline is to find the actual rate-limiting step in the journey from insight to commitment and to work on that, rather than on the step that is easiest and most visible to improve. Almost always, the true constraint is in the decision layer, and almost always, that is where the least work has been done.

None of this is what most transformation programmes are currently doing, which is precisely why the paradox is so widespread. We have industrialised the production of

analysis and left the making of decisions almost exactly as we found it. The gap between the two is the paradox, and it will not close by making the fast part faster.

7. A Closing Reflection

I began by describing the strange coexistence of faster outputs and unchanged decisions, and I want to end by naming what I think it is really about. The productivity paradox of AI is, at bottom, a paradox about where value in an organisation actually comes from. We have been seduced, not for the first time, into believing that value lives in the production of insight, and so we have invested in producing insight faster. But value has never lived there. It lives in the decision — in the commitment, the allocation, the willingness to act — and the decision is made of trust, authority, and judgement, none of which a tool produces.

The organisations that will realise the promise of these tools are not the ones with the fastest outputs. Almost everyone will soon have fast outputs; they will become as unremarkable as a spreadsheet. The organisations that pull ahead will be the ones that treat the acceleration of analysis as a diagnostic — a floodlight on their own decision-making — and then do the slow, structural, unglamorous work of raising the velocity and quality of the decisions themselves. That work is harder than adopting a tool, which is exactly why so few are doing it, and exactly why doing it will matter.

We have automated the production of analysis without touching the act of judgement, and then expressed surprise that the second did not follow the first. The surprise is the tell. It reveals that we had confused the two all along. Faster outputs were never going to produce better decisions, because outputs and decisions were never the same thing. The paradox is not a problem to be solved by the technology. It is an invitation, issued by the technology, to finally do the work we have been avoiding.