

WORKING AFTER  
**AI**

---

*How to Stay Valuable When  
Machines Can Think and Act*

A field guide and research synthesis on judgment,  
orchestration, and human value in the age of  
autonomous machine cognition

NGO TRUNG KIEN

FIRST EDITION • 2026

# Contents

---

|                                                                                     |      |
|-------------------------------------------------------------------------------------|------|
| <b>Contents</b>                                                                     | ix   |
| <b>List of Figures</b>                                                              | xiii |
| <b>Preface</b>                                                                      | xv   |
| <b>Executive Introduction</b>                                                       | xxi  |
| <b>1 The Paradigm Shift</b>                                                         | 1    |
| 1.1 The day the memo wrote itself . . . . .                                         | 1    |
| 1.2 What “cognitive abundance” does and does not mean . . . . .                     | 3    |
| 1.3 The economics: why abundance in one input creates scarcity in another . . . . . | 7    |
| 1.4 Is the trajectory real? A sceptic’s checklist . . . . .                         | 12   |
| 1.5 Three misreadings that will cost you . . . . .                                  | 14   |
| 1.6 What this means for the rest of the book . . . . .                              | 16   |
| <b>2 Deconstructing Work</b>                                                        | 19   |
| 2.1 The wrong unit of analysis . . . . .                                            | 19   |
| 2.2 The task-based framework . . . . .                                              | 21   |
| 2.3 Exposure is not destiny: the two-dimensional map . . . . .                      | 22   |
| 2.4 The decomposition procedure . . . . .                                           | 25   |
| 2.5 What the evidence actually shows . . . . .                                      | 26   |
| 2.6 The shrinking value of raw execution . . . . .                                  | 32   |
| <b>3 The Human Advantage</b>                                                        | 35   |
| 3.1 The most useful reframe in the economics of AI . . . . .                        | 35   |
| 3.2 The judgment stack . . . . .                                                    | 38   |
| 3.3 Contextual sensemaking: the asset nobody wrote down . . . . .                   | 41   |
| 3.4 The uncomfortable evidence on human–machine teams . . . . .                     | 42   |
| 3.5 High-leverage thinking . . . . .                                                | 45   |
| <b>4 Orchestration and Systems Architecture</b>                                     | 49   |

|          |                                                              |            |
|----------|--------------------------------------------------------------|------------|
| 4.1      | From using a tool to running a system . . . . .              | 49         |
| 4.2      | What an agent actually is . . . . .                          | 50         |
| 4.3      | The six levels of delegated autonomy . . . . .               | 55         |
| 4.4      | Verification: the discipline that decides who wins . . . . . | 56         |
| 4.5      | Designing the human role . . . . .                           | 60         |
| <b>5</b> | <b>Relational and Ethical Anchors</b>                        | <b>65</b>  |
| 5.1      | The uncomfortable experiment . . . . .                       | 65         |
| 5.2      | Empathy performed and empathy owed . . . . .                 | 66         |
| 5.3      | The responsibility gap . . . . .                             | 68         |
| 5.4      | Trust, miscalibrated in both directions . . . . .            | 69         |
| 5.5      | Fairness, harm, and the professional's obligation . . . . .  | 72         |
| 5.6      | Why this is a job, not a virtue . . . . .                    | 74         |
| <b>6</b> | <b>The AI-Augmented Enterprise</b>                           | <b>79</b>  |
| 6.1      | The gap between capability and results . . . . .             | 79         |
| 6.2      | What happens to hierarchy . . . . .                          | 83         |
| 6.3      | The micro-team as the productive unit . . . . .              | 86         |
| 6.4      | The pipeline problem . . . . .                               | 87         |
| 6.5      | What good deployment actually looks like . . . . .           | 89         |
| <b>7</b> | <b>Personal Career Strategy</b>                              | <b>93</b>  |
| 7.1      | Why a portfolio, and why antifragile . . . . .               | 93         |
| 7.2      | Four portfolio principles . . . . .                          | 94         |
| 7.3      | What is actually appreciating . . . . .                      | 98         |
| 7.4      | The early-career problem . . . . .                           | 100        |
| 7.5      | The twelve-month repositioning roadmap . . . . .             | 101        |
| 7.6      | On the parts that are not strategy . . . . .                 | 103        |
| <b>8</b> | <b>The Horizon</b>                                           | <b>107</b> |
| 8.1      | How large is this, really? . . . . .                         | 107        |
| 8.2      | Distribution is the variable that matters . . . . .          | 109        |
| 8.3      | The direction of technology is a choice . . . . .            | 111        |
| 8.4      | Policy instruments and what the evidence says . . . . .      | 112        |
| 8.5      | Four scenarios for 2026–2040 . . . . .                       | 114        |
| 8.6      | What would actually help . . . . .                           | 117        |
| 8.7      | Conclusion: what remains . . . . .                           | 118        |
|          | <b>Appendix A: The Action Plan</b>                           | <b>121</b> |
|          | <b>Glossary</b>                                              | <b>125</b> |
|          | <b>References</b>                                            | <b>133</b> |

**A Note on Method and Sources**

149

# 1 CHAPTER

---

## The Paradigm Shift

---

*"We can know more than we can tell."*

— Michael Polanyi, *The Tacit Dimension*, 1966

---

*This chapter establishes what has actually changed. It distinguishes the collapse in the price of cognitive execution from any claim about machine understanding, situates that collapse within the history of general-purpose technologies, and introduces the central economic mechanism of the book: when an input becomes abundant, returns migrate to whatever remains scarce. It closes by identifying five candidate scarcities that organise the remainder of the book.*

---

### 1.1 The day the memo wrote itself

Consider a morning in a mid-sized professional services firm. The scene below is a composite — assembled from what these systems demonstrably did in commercial use by the mid-2020s rather than reported from a single observed case — and it is offered as an illustration of a mechanism, not as evidence of one.

An associate is asked to prepare a briefing on a client's competitive position. Traditionally this consumes most of two days: assembling filings and market data, reading them, identifying the structural forces at work, drafting an argument, producing charts, and revising after a partner's comments. The associate has done this perhaps forty times. They are good at it. Their reputation, their billing rate, and their sense of professional identity all rest, to a significant degree, on being good at it.

On this particular morning the associate instead specifies the task to a software system: the client, the question, the sources to consult, the format, the length, the house style. The system retrieves the filings, extracts the relevant figures, constructs the charts, drafts eleven pages of coherent argument, and returns it in under four minutes. The associate reads it with the particular attention of someone checking whether they are still necessary.

The document is not excellent. It is, roughly, what a competent second-year associate would produce on a reasonable day: well-organised, factually mostly accurate, analytically conventional, occasionally confidently wrong in ways that would embarrass the firm if nobody caught them. Three of the market-share figures are subtly misattributed. The competitive framing misses the fact — known to the partner, mentioned once in a phone call, written down nowhere — that the client's chief executive has already privately decided to exit one of the segments the memo treats as strategic. The recommendation is defensible and generic.

The associate spends ninety minutes fixing the figures, restructuring the argument around what the partner actually knows, cutting four pages, and sharpening the recommendation into something a person could act on. Total elapsed time: under two hours, for a deliverable that used to take two days.

Nothing in that story requires artificial general intelligence, machine consciousness, or any capability that did not exist in commercially available systems by the middle of the 2020s. And everything important about the future of professional work is contained in it, if we read it carefully.

Three things happened. First, the *production* of a competent cognitive artefact became nearly free — fast, cheap, and available on demand. Second, the value of that production, as an activity a firm would pay a person to perform, collapsed with it: nobody will pay two days of billable time for four minutes of machine output. Third, and most importantly, the residual human contribution did not vanish. It concentrated. What remained — catching the errors, knowing the unwritten fact, deciding what the argument should be, being the person whose name is on the recommendation — turned out to be both smaller in volume and greater in value per unit than

what was displaced.

That concentration is the subject of this book.

#### THE CORE DISTINCTION

This book is not about machines becoming intelligent. It is about the price of competent cognitive *execution* falling towards zero, and about what becomes scarce, and therefore valuable, as a consequence. Those are different claims, and only the second one needs to be true for everything that follows to matter.

## 1.2 What “cognitive abundance” does and does not mean

The term *cognitive abundance* is used throughout this book with a deliberately narrow definition, because the loose version of the idea is both popular and misleading.

**The narrow claim.** The marginal cost of producing a unit of competent, structured cognitive output — a page of coherent prose, a working function, a first-pass analysis, a translation, a classification, a summary — has fallen by several orders of magnitude in a very short period, and continues to fall. Systems that produce such output are available on demand, operate in parallel, do not tire, and improve on a timescale measured in months. That is the whole of the claim, and it is a claim about *cost and availability*, not about cognition in any deeper sense.

The magnitudes are worth stating precisely, because the word “cheaper” does not convey them. The cost of querying a system performing at the level of the November 2022 state of the art fell from roughly twenty dollars per million tokens to about seven cents per million tokens by October 2024 — a reduction of more than two hundred and eighty-fold in under two years (Maslej et al., 2025). The decline is not uniform across capabilities: an analysis of six benchmarks found the price of reaching a fixed capability milestone falling at rates ranging from nine-fold to nine-hundred-fold per year, with the price of frontier-level performance on graduate-level science questions falling roughly forty-fold annually (Cottier et al., 2025). Cognitive abundance therefore arrives at very differ-

ent speeds for different kinds of thinking, which is one reason professional experience of this transition is so uneven.

Behind the price decline sit three technical developments. The transformer architecture, introduced in 2017, made language modelling efficiently parallelisable and therefore economical to scale (Vaswani et al., 2017). Empirical scaling laws then established that model loss falls as a predictable power law in model size, data, and compute across more than seven orders of magnitude (Kaplan et al., 2020), later refined by the finding that model size and training data should be scaled in roughly equal proportion — a correction that produced substantially better models at lower cost (Hoffmann et al., 2022). And the compute devoted to training grew from a Moore’s-law doubling time of about twenty months before 2010 to roughly six months thereafter (Sevilla et al., 2022). Cheap cognition is the downstream consequence of these three facts, not an independent event.

**What the claim does not include.** It does not assert that these systems understand what they produce. It does not assert that they reason in the way people reason, that they possess goals, that they know when they are wrong, or that their apparent competence generalises reliably beyond the distribution of their training. It does not assert that the trajectory continues indefinitely; there are serious arguments that it will not. And it emphatically does not assert that cheap production of an artefact is the same as production of *value*. A great deal of this book is concerned with the gap between those two things.

The narrow claim is nonetheless sufficient to reorganise professional life, because most professional life is organised around producing artefacts. Consider what a large fraction of knowledge work actually consists of, stripped of its dignity: reading documents and extracting what matters; converting a request into a structured deliverable; writing the same category of message for the four-hundredth time; assembling information that exists into a form that a decision-maker will read; producing code that implements a specification someone else wrote. None of this is trivial, and much of it requires years of training. All of it is production.

### 1.2.1 The uncomfortable symmetry with previous waves

It is tempting for knowledge workers to regard this moment as historically unique, and there is a self-serving quality to that temptation which deserves examination. Every previous group whose economic position was undermined by automation believed their work contained something irreducible. Handloom weavers were not wrong that machine cloth was inferior; they were wrong that inferiority would protect them. Skilled machinists were not wrong that numerical control produced cruder parts at first. Draughtsmen were not wrong that computer-aided design lost something in the transition from the drawing board.

Each of these groups possessed genuine expertise. Each was correct that the machine's output was, initially, worse. Each was destroyed anyway, because the relevant comparison was never machine versus master craftsman. It was machine-plus-modestly-skilled-operator versus master craftsman, at one-tenth the cost, at ten times the volume, improving every year.

The lesson generalises uncomfortably. When evaluating whether a system threatens your economic position, the question is not "is it as good as I am?" The question is: *is the combination of this system and someone cheaper than me good enough for the buyer, at a price that changes their behaviour?* For a great deal of professional output, the answer is now yes, and the honest response to that is neither denial nor despair but relocation.

### 1.2.2 But the symmetry is not perfect

Two features do distinguish the present wave, and both cut against the reassuring reading of history.

The first is **breadth**. Previous automation waves were sector-specific and required physical capital deployed in particular places. The technology of the present wave is delivered through a general interface — language — and applies to any task that can be described in words. Assessments of occupational exposure built by three independent methods — expert and model rating of task statements (Eloundou et al., 2024), matching patent text to occupational task descriptions (Webb, 2019), and mapping capability benchmarks

onto occupational abilities (Felten et al., 2021) — agree that the affected tasks are distributed across the entire wage and education distribution rather than concentrated at the bottom, and that exposure tends to rise with earnings and educational requirements. This is a reversal of the pattern that held for computerisation, which mainly hollowed out the routine middle (Autor et al., 2003; Goos et al., 2014).

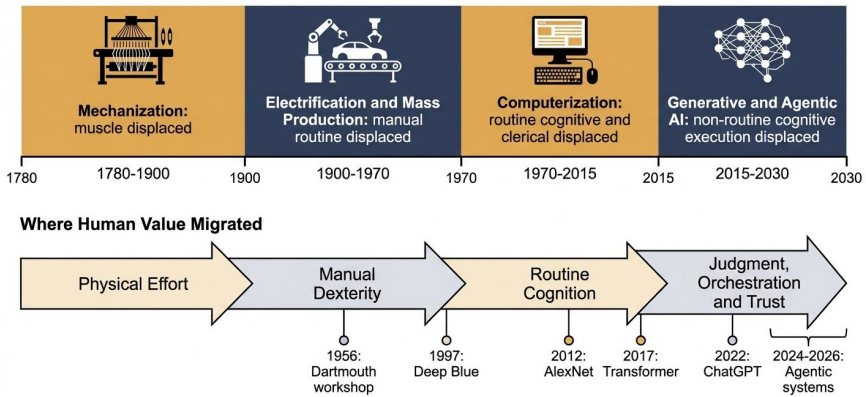
The second is **speed of diffusion**. Electrification is the canonical case. Electric motors were commercially available from the 1880s, and by the end of the 1910s they supplied more than half of United States manufacturing drive capacity — yet the surge in manufacturing productivity growth did not arrive until the 1920s, because it waited on factories being physically rebuilt around the new logic of distributed power (David, 1990). Diffusion of the technology and realisation of its gains were separated by roughly four decades.

Generative systems require no such rebuilding to reach the individual worker; a professional can adopt one in an afternoon, with or without their employer's knowledge or consent. Nationally representative survey evidence found that close to 40 percent of the United States population aged 18 to 64 had used generative systems within roughly two years of release, with around a quarter of employed respondents having used them for work in the previous week — an adoption curve running ahead of the equivalent curves for the personal computer and the internet at comparable points after release (Bick et al., 2024). Chapter 6 argues that this contrast between individual adoption and organisational transformation is the single most important fact about the present moment.

This is the pattern that defines a general-purpose technology in the first place: pervasiveness across sectors, inherent scope for continued improvement, and innovational complementarities that generate increasing returns — together with a characteristic coordination failure in which arm's-length transactions between the technology and its application sectors produce too little innovation, too late (Bresnahan and Trajtenberg, 1995).

There is a countervailing consideration, and it is the single most important qualification in this book: *rapid adoption by individuals is not the same as transformation of organisations*. The gap between those

## Four Waves of Automation and the Migration of Human Value, 1780 to 2030



**Figure 1.1. Four waves of automation and the migration of human value** (conceptual schematic; periodisation is illustrative). Each wave of automation reduced the price of a particular input — physical force, controllable power, arithmetic and rule-execution, and now competent cognitive production. In each case, human economic value migrated to whatever remained scarce. The present wave differs from its predecessors in two respects: it applies across the whole occupational distribution rather than to a specific sector, and it diffuses to individual workers without requiring firms to rebuild their physical capital first.

two things — explored at length in Chapter 6 — is where most of the disappointment, and most of the opportunity, currently lives.

### 1.3 The economics: why abundance in one input creates scarcity in another

The single most useful piece of economics for thinking about this moment is also one of the oldest, and it has nothing to do with computers.

In 1865 the economist William Stanley Jevons observed that improvements in the efficiency of coal-fired steam engines had not reduced Britain's coal consumption. They had increased it. Cheaper steam power made steam power worth using in applications where it had previously been uneconomic; the expansion of use swamped the reduction in use per application. The general principle — that

falling cost of an input often raises rather than lowers total demand for it — is now called the Jevons effect, and it is the first thing to understand about cognitive abundance.

Two qualifications are due, because the Jevons effect is invoked far more often than it is measured. In energy economics the rebound from efficiency improvements is typically partial rather than total: careful review of the evidence finds that rebound effects exist but rarely exceed the original saving, so the strong “backfire” version of the claim is not the empirical norm (Gillingham et al., 2013). Where induced demand *is* close to complete is in settings with a large reservoir of latent, suppressed demand — most famously road capacity, where the elasticity of vehicle-kilometres travelled with respect to lane-kilometres of highway is estimated to be close to one, so that new capacity is essentially fully consumed by new traffic (Duranton and Turner, 2011). The question for cognitive work is therefore an empirical one about the size of the latent reservoir: how much analysis, drafting, code, and advice is currently *not* produced because it costs too much? On the available evidence that reservoir appears very large, which is the case for expecting something closer to the road-capacity outcome than the energy-efficiency one — but it is a judgement, not a theorem.

The naive reading of cheap machine cognition is that the world needs a fixed quantity of memos, analyses, and code, and that machines will now produce that fixed quantity, leaving nothing for people. The Jevons reading is that the world’s demand for memos, analyses, and code was never fixed; it was suppressed by cost. When the cost falls far enough, activities that were previously not worth doing become worth doing. Small businesses that never had a legal opinion get legal opinions. Analyses that were never run get run. Software that was never worth building gets built. The volume of cognitive output in the economy does not stay constant and shift to machines; it explodes.

This is genuinely good news, and it is also where the naive optimistic reading goes wrong in turn. Because the important question is not whether total cognitive output rises. It is: *when output rises by two orders of magnitude, what becomes the binding constraint?*

### 1.3.1 Bottleneck migration

Every production process has a constraint. Relieve one, and the constraint moves somewhere else; the returns move with it. This is the central mechanism of the entire book, and it is worth stating formally.

When a production process requires several complementary inputs, and the price of one input collapses, the economic surplus does not disappear — it accrues to the inputs that remain scarce. The classical illustration is that cheap steel did not enrich steel-makers so much as it enriched the owners of the things that steel enabled and that steel could not itself supply: rail networks, urban land, distribution rights. In the language of production economics, complements to the newly cheap input appreciate; substitutes for it depreciate.

Apply this to cognitive work and the strategic map becomes legible. If competent artefact production is the input whose price is collapsing, what are its complements? Five survive the question, and they are the five human anchors named in the Executive Introduction. It is worth deriving each rather than asserting it, because the derivation is what distinguishes this list from the many lists of comforting human qualities that circulate on the subject.

**Judgment** is the first and the deepest. Machine systems optimise against an objective; they do not supply one. The determination of what outcomes are worth — how much a false negative costs relative to a false positive, whose interests count, what risk is acceptable — is logically prior to any prediction about how to achieve them, and it is not contained in the data. Chapter 3 develops this as the principal complement to cheap prediction.

**Contextual sensemaking** is the second, and the most defensible in practice. Machine systems know what has been written down. Every organisation, every client relationship, and every technical system contains a large body of decisive information that has never been recorded, often deliberately. That information does not become available with more training data, because it does not exist in any corpus. Chapter 3 treats it as the most underrated career asset available to a working professional.

**Orchestration** is the third, and it is where the two most obvious complements live. One is *specification*: machine systems act on what they are told, which is not the same as what is wanted, and cheap execution makes the gap between an expressed request and an actual need more consequential rather than less — executing the wrong task perfectly, a hundred times faster, is worse than executing it slowly. The other is *verification*: a system that produces a hundred analyses per hour, of which ninety-four are correct and six are confidently wrong in ways invisible on the surface, has not produced ninety-four useful analyses. It has produced a hundred that must be checked, and the checking now dominates the cost structure. Specifying, decomposing, routing, and verifying are the components of a single discipline, and Chapter 4 treats it as one.

**Relational trust** is the fourth. In a class of transactions the human is not the means of production but the product: what is purchased is assurance, continuity, advocacy, or the transfer of somebody else's risk. Cheap generation does not touch this directly, and by making unassured output abundant it raises the premium on assured output. Chapter 5 examines both the strength and the real vulnerability of this position.

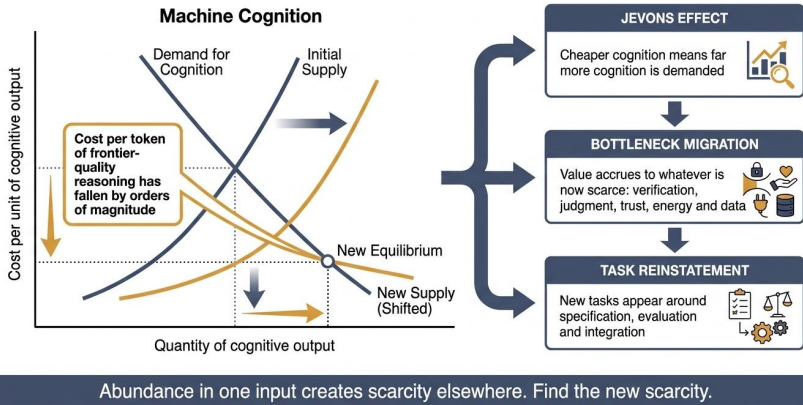
**Moral responsibility** is the fifth. Someone must be answerable. This is not a technical constraint but an institutional and moral one, and Chapter 5 argues that it is unusually durable precisely because it is not technical — while warning that durability of the role is not durability of the headcount.

These five are the human anchors, and they were not chosen because they sound humane. They were identified by asking a structural economic question — what does cheap machine cognition still require, and cannot supply? — rather than a sentimental one.

### 1.3.2 The displacement–reinstatement balance

The formal apparatus for this reasoning comes from the task-based framework in labour economics, developed most systematically by Daron Acemoglu and Pascual Restrepo (Acemoglu and Restrepo, 2018; Acemoglu and Restrepo, 2019), and it deserves stating because it is so often garbled in popular accounts.

## Cognitive Abundance: What Happens When the Price of Reasoning Collapses



**Figure 1.2. Cognitive abundance and the migration of the bottleneck** (conceptual schematic; curves are illustrative). A collapse in the price of machine cognition produces three simultaneous effects: total demand for cognitive output rises rather than falls (the Jevons effect); the economic constraint migrates to the complements that remain scarce; and new tasks are created around specification, evaluation, and integration. Strategic advantage belongs to whoever correctly identifies the new bottleneck first.

In this framework, output is produced by a continuum of tasks, each of which can be performed by labour or by capital. Automation is the reassignment of tasks from labour to capital. It has two opposing effects on the demand for labour. The *displacement effect* is the direct one: tasks previously done by workers are now done by machines, reducing labour demand at any given wage. The *productivity effect* is indirect: automation lowers costs, which raises output and therefore raises demand for labour in the non-automated tasks.

Crucially, in this framework, the productivity effect alone is generally not enough to restore labour's position. What restores it is the *reinstatement effect*: the creation of *new tasks* in which labour has a comparative advantage. Historically this is what has actually happened, and it is why two centuries of relentless automation have not produced permanent mass unemployment. The occupations that absorbed displaced agricultural and manufacturing labour

largely did not exist beforehand.

Two implications follow, and they are the ones most often lost.

First, the reinstatement effect is not automatic. It is a function of the direction of technological development and of institutional choices — what firms invest in, what the tax code rewards, what workers are able to learn. A technology deployed purely to substitute for labour, in an economy whose incentives favour substitution, can produce sustained downward pressure on labour demand without any compensating creation of new tasks. Whether the current wave is developed in a substituting or a complementing direction is a choice, not a discovery, and Chapter 8 takes it up in detail.

Second, and more immediately: the reinstatement effect does not help the displaced individual. New tasks accrue to new entrants, new firms, and different places. The economist's aggregate reassurance — that employment will be fine in the long run — is fully compatible with a decade of severe personal disruption for a specific forty-five-year-old in a specific profession. Workers displaced from stable employment suffer large and remarkably persistent earnings losses (Jacobson et al., 1993), still detectable fifteen to twenty years later (Davis and Wachter, 2011), and the experience of trade-induced displacement showed that adjustment within affected local labour markets was far slower and more painful than standard models predicted, with elevated unemployment and reduced participation persisting for a decade or more (Autor et al., 2013; Autor et al., 2014). Anyone who offers you the aggregate result as personal consolation has misunderstood the question you were asking.

## 1.4 Is the trajectory real? A sceptic's checklist

A book that argued for a transformation without examining the case against it would be advocacy rather than analysis. There are four serious sceptical positions, and each contains something true.

**The benchmark illusion.** A great deal of the perceived rate of progress rests on performance against benchmarks, and benchmarks are a poor proxy for real work. They are contaminated by training data, they measure narrow and verifiable tasks, and

the leap from a benchmark score to an economic capability is frequently unwarranted. There is a further, subtler problem. It has been argued that certain capabilities “emerge” discontinuously with scale and are therefore unpredictable from smaller models (Wei et al., 2022a) — but a careful reanalysis showed that apparent emergence is frequently an artefact of *metric choice*: nonlinear or discontinuous scoring manufactures sharp jumps, while linear metrics applied to the very same model outputs show smooth and predictable improvement (Schaeffer et al., 2023). An independent audit of widely used agentic benchmarks found task-setup and reward-design flaws capable of mis-stating agent performance by up to 100 percent in relative terms (Zhu et al., 2025). The honest position is that benchmark performance and economic usefulness are correlated but not the same variable, and that anyone reasoning about the labour market from leaderboards is reasoning badly.

**The reliability wall.** Autonomous multi-step operation compounds error. A system that completes each step correctly nine times in ten completes a ten-step task correctly about a third of the time. Measurement of the “time horizon” of autonomous operation — the length of task, measured in human working time, that a system completes with 50 percent success — put the frontier model of March 2025 at just under an hour, with the horizon doubling approximately every seven months since 2019 (Kwa et al., 2025). That is a striking growth rate and a sobering absolute level: the base is measured in tens of minutes, not weeks. This arithmetic is why demonstrations impress and deployments disappoint, and it is the honest technical reason that fully autonomous knowledge work remained rare well into the second half of the 2020s. Chapter 4 examines what this means for how work should actually be designed.

**The measurement gap.** If this technology is so transformative, where is it in the productivity statistics? This is the Solow paradox restated, and it has a well-documented answer: general-purpose technologies produce measured gains only after firms have built the complementary intangible capital — redesigned processes, restructured decision rights, retrained people, rebuilt data infrastructure — that lets them be used (Brynjolfsson et al., 2019). That invest-

ment is expensed rather than capitalised in national accounts, so measured productivity initially *falls*. Adjusting the United States national accounts for the intangible investments complementary to computing hardware and software yields a total factor productivity level about 16 percent higher than the official measure by the end of 2017 — the empirical core of the J-curve (Brynjolfsson et al., 2021). Its existence means that the absence of aggregate gains today tells us very little about the presence of gains tomorrow, and equally that the presence of impressive demonstrations today tells us very little about gains today.

**The economics of the frontier.** The current trajectory depends on continued capital investment on an extraordinary scale, on the availability of energy and specialised hardware, and on training data that is no longer freely expanding. Whether the returns justify the investment is an open question in 2026, and a sharp correction in capital availability would slow diffusion considerably — though it is worth noting that it would slow the *frontier* more than it would slow the deployment of capabilities that already exist and are already cheap.

#### HOW TO WEIGH A CLAIM ABOUT AI AND WORK

When you encounter a claim about artificial intelligence and employment, ask four questions before believing it. **(1) What is the unit?** Claims about “jobs” are usually claims about tasks, dressed up. **(2) What is the counterfactual?** A firm that hires fewer people after adopting a technology may have hired fewer people anyway, for reasons of demand. **(3) Is this exposure or effect?** Occupational exposure indices measure what a technology *could* do; they are not measurements of what happened. Conflating the two accounts for a large share of alarming headlines. **(4) Who is the population?** Findings from early-adopting technology firms, or from experiments with tasks chosen for measurability, generalise poorly to the whole economy.

## 1.5 Three misreadings that will cost you

Before the analytical chapters begin, it is worth naming the three responses that are most common among intelligent professionals,

and most reliably wrong.

### 1.5.1 “My work requires creativity, so I am safe”

Creativity is a poor category for this analysis, because the word conflates two different things. There is *generative* creativity — producing novel combinations, variations, options, drafts — and there is *evaluative* creativity — knowing which of the options is any good, and why, and for whom.

Machine systems are extraordinarily strong at the first. Producing forty campaign concepts, sixty variations on a melody, a dozen architectural approaches to a software problem: this is precisely what generative systems do well, and cheaply, and at volume. If your creative contribution is fluency of production, it is exposed.

The second is a different matter. Selecting the concept that will actually land with a specific audience in a specific cultural moment, and being willing to stake a budget and a reputation on that selection, is judgment under uncertainty — and it depends on knowing things about the audience, the moment, and the risk that were never in any training corpus. The creative professions are not uniformly safe or uniformly threatened; they are being split along exactly this seam, and which side of it you are on is largely a matter of how you have chosen to spend your working hours.

### 1.5.2 “I will be fine because I know how to use the tools”

Tool fluency is necessary and radically insufficient. It is necessary because the alternative — refusing — is not a strategy. It is insufficient because tool fluency is the single most rapidly commoditised skill in the current environment. Interfaces are designed by well-funded companies specifically to reduce the expertise required to use them. Skills that took a week to acquire in 2023 take an hour in 2026, and the systems increasingly do the prompting themselves.

Any strategy whose core asset is “I am unusually good at operating this software” has a half-life measured in quarters. The durable version of this skill is not operating the tool; it is designing the *system* in which the tool operates — the decomposition, the

specification, the verification, the escalation policy. That is Chapter 4, and it is a different discipline entirely.

### 1.5.3 “This is hype; it will pass”

This is the most dangerous of the three, because it is partially correct and therefore comfortable. There is enormous hype. Many deployments will fail. A significant number of enterprise programmes will be quietly abandoned. Valuations may correct violently.

None of that is inconsistent with the underlying capability being real and durable. The dot-com collapse of 2000 destroyed a great deal of capital and did not, in any respect, reverse the internet. Hype cycles and technological reality are close to orthogonal, and the failure of a particular set of business models tells you nothing about whether the price of cognitive production has fallen. It has, and nothing on the current horizon returns it to where it was.

## 1.6 What this means for the rest of the book

If the argument of this chapter holds, the strategic question facing any working professional is no longer whether to engage with these systems. It is where to stand.

The remainder of the book answers that question in layers. Chapter 2 provides the analytical instrument: a method for decomposing your own work into tasks and scoring each one for exposure and complementarity, so that “where to stand” becomes a concrete inventory rather than an anxiety. Chapters 3 through 5 develop the three human anchors that survive that analysis in the strongest condition: judgment and sensemaking, orchestration, and trust and moral responsibility. Chapter 6 examines what happens to organisations when their members can each direct a fleet of systems. Chapter 7 converts all of it into a personal programme. Chapter 8 asks what the collective, institutional response should be, on the argument that the individual response is necessary but nowhere near sufficient.

One orienting idea should carry through all of it. The professionals who will do well in this transition are not the ones who work hardest to keep doing what they currently do. They are the ones

who correctly identify which parts of what they currently do were always the valuable parts — and who have the discipline to stop doing the rest.

## CHAPTER TAKEAWAYS

---

1. **The claim is narrow and sufficient.** What has changed is the price of competent cognitive execution, not the arrival of machine understanding. Only the price claim is needed for the argument of this book.
2. **Abundance in one input creates scarcity in another.** Cheap machine cognition raises the returns to its complements: verification, specification, context, accountability, and judgment. These are the five human anchors.
3. **Jevons, not substitution.** Total demand for cognitive output is rising, not falling. The question is who captures the surplus, and the answer is whoever holds the new bottleneck.
4. **Reinstatement is a choice, not a law.** Automation historically created new tasks for labour, but that outcome depends on the direction of technological development and on institutional incentives. It is not guaranteed, and it does not help the individual who is displaced today.
5. **Take the sceptics seriously.** Benchmark scores overstate economic capability; error compounds in multi-step autonomy; measured productivity gains lag by years because of the intangible-investment J-curve. All three are true, and none of them reverses the price collapse.
6. **The three comfortable misreadings** — creativity as armour, tool fluency as strategy, and hype as refutation — are the most common and most expensive errors available to an intelligent professional.

## DIAGNOSTIC QUESTIONS

1. Of the hours you worked last week, what fraction was spent *producing an artefact* whose specification was already clear before you began?

2. Name one thing you know about your organisation, client, or system that is decisive for good decisions and is not written down anywhere. How many such things do you hold?
3. When a machine system gives you a confident answer in your domain, what specifically do you check — and could you write that check down as a procedure someone else could follow?
4. If your current output could be produced at one-tenth the cost by someone junior directing a system, what would you be paid for instead? Answer in one sentence. If you cannot, that is the finding.