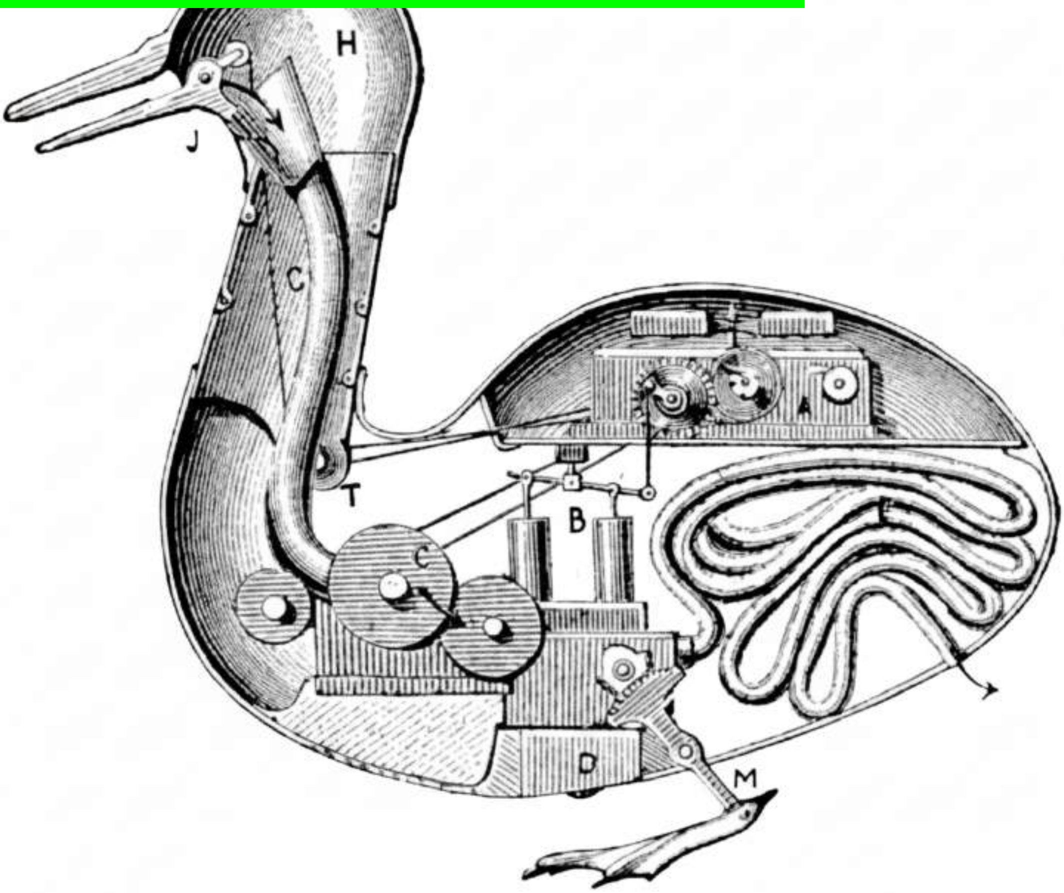
NOT ARTIFICIAL, NOT INTELLIGENT

WHAT AI COMPANIES DON'T WANT YOU TO KNOW



DJANGO BEATTY

INDEX

PART II: HOW IT REALLY WORKS

WHY YOUR AI NEVER WORKS ON THE FIRST TRY

The mathematical proof your frustration is inevitable, the law that says you'll never know if you're close, how AI turns programmers into pilots and writers into navigators, and the moment Thoughtworks admitted defeat.

THE PACHINKO MACHINE PLAYS YOU

What makes pigeons peck and humans type, why Microsoft's AI CEO thinks you're going insane, how AI got inside your OODA loop, and the difference between comic book panic and actual harm.

RACIST MATHS

How AI reveals your company's hidden values, how bias can hide in three random numbers, why Grok went MechaHitler in one beat, why killing DEI is tomorrow's smoking gun, and what owls mean for your training data.

PART III: WHO WINS, WHO WORKS

THEY PAID TO PLAY COACHELLA

Why the biggest break in music costs six figures to accept, how copyright killed the thing it was meant to protect, what Morris Levy's baseball bat taught Silicon Valley, and why the Stationers' Company would love Spotify.

BIG JOBS

The apocalypse where everyone gets hired, where productivity hides for centuries, the wrong-shaped factories, why Edison funded the electric chair to win an argument, and the invention of childhood.

THE DISAPPEARING SALARY

What Malcolm McLean's metal boxes did to Detroit, what happened when France sold off taxes, why your boss would sell your future to a vending machine, what Sam Altman can't say, and how to build an empire on crumbs.

OUTRO: THE MACHINE CAN'T PLAN. YOU CAN'T STOP.

What Engelbart's 1968 demo taught us about timing, why naming shapes perception, how iteration becomes literacy, mirrors made of maths, the four lenses. And how this ends.

INTRODUCTION

AI didn't arrive like normal technology. It landed more like a B-movie alien egg in a smouldering crater. Around it gathered the usual crowd: journalists hyping miracles, critics warning of monsters, politicians staking positions, the army circling warily, gawkers staring in a mixture of awe and disdain. Some see promise, some see peril. Most don’t know what the hell to think.

Among the crowd, two figures dominate. The landowner wants to sell tickets: AGI, productivity, AI for everything. The priest warns of apocalypse: extinction risk, mass job loss, civilisation undone.

Both are wrong in different ways. Together they create an info-smog that blinds those who need clarity most.

This book won't tell you what to do about AI. There are enough consultants selling that already. These essays dig into what's actually happening - the mechanics behind the hype, the economics behind the panic, the patterns behind the noise. Think of it as orientation, not prescription.

Quick note on terminology: I use 'AI' even though it's misleading. These are machine learning systems - pattern matching, not intelligence. As the title says, nothing artificial about the capability, nothing intelligent about the computation. When I say 'AI', I mostly mean large language models since that's what everyone's mostly using. But 'AI' is what everyone calls it, and fighting the term just makes the conversation harder. One prediction: when the hype dies, the ML capabilities will keep improving. We might see an 'AI winter' followed by an 'ML summer' - same tech, honest branding.

AUTONOMOUS, EXCEPT WHEN IT MATTERS

'Next year'.

Self-driving cars have been 'next year' since 2014. Every major tech company, every car manufacturer, every ambitious startup promised the same thing. Fully autonomous vehicles just around the corner. Next year. Maybe the year after. Definitely by 2020. Or 2025. Or 2030.

The pattern is so consistent it's become a joke in the industry. But the joke reveals something about how we misunderstand AI progress.

In 2016, Uber's then-CEO Travis Kalanick declared that by 2030, most Uber rides would be in self-driving cars. Ford promised a fully autonomous vehicle by 2021. GM targeted 2019. Tesla's Elon Musk has predicted 'next year' every year since 2014. Each deadline passes, each promise fails, and new deadlines emerge like clockwork.

The technical challenge seemed straightforward: outfit a car with cameras and sensors, train a neural network on millions of miles of driving data, and let pattern recognition do the rest. If AI could beat the world's best Go player, surely it could navigate a suburban street.

But driving isn't Go. Go has fixed rules, perfect information, clear victory conditions. Driving by contrast is made up of edge cases - construction zones that change daily, emergency vehicles that need special responses, children who might dart into the street, debris that could be a plastic bag or a rock. Each edge case spawns more edge cases. This isn't just a long tail - it's a branching maze. Each edge case multiplies, revealing new dimensions of failure. Like brain chemistry, every intervention ripples unpredictably. Fix one, distort another.

Waymo, Google's self-driving subsidiary, has spent over $20 billion and logged 20 million autonomous miles. They've achieved something remarkable: a taxi service that works pretty well in a few carefully mapped neighbourhoods in Phoenix and San Francisco. The cars drive slowly, avoid highways, and still occasionally get confused by construction cones.

This is what success actually looks like. Narrow deployment in controlled environments. Gradual expansion. Human oversight at every step. And billions poured into infrastructure, all for modest operational gains. It works, just not the way anyone imagined.

The mismatch between promise and reality created a credibility crisis. When Uber's self-driving car killed a pedestrian in 2018, it wasn't just a tragedy - it was proof that the entire industry had been lying about how close they were. When Tesla's ‘Full Self-Driving’ turned out to require constant human supervision, customers felt deceived.

The technology works - just not as a drop-in replacement for human drivers. It works as advanced cruise control that makes highway driving safer. It works as parking assistance that prevents fender benders. It works as emergency braking that saves lives. Boring, incremental improvements that nobody notices because they're not the revolution we were promised.

The self-driving car saga perfectly illustrates the AI hype cycle: wildly overestimate what's possible in the short term, completely miss what's possible in the long term. The shift comes not as a dramatic replacement but as a thousand small improvements that gradually transform the entire system.

Today's AI hype follows the same pattern. AGI is always 'just around the corner', but the real transformation happens in the boring middle ground - workflow automation, code assistance, content generation. Not artificial general intelligence, but applied machine learning at scale.

This is why orientation beats prediction every time. The prediction says 'full self-driving next year' and fails. The orientation says 'watch where the technology actually works, not where we wish it would work'. One leads you in circles - the prediction returning each year like Groundhog Day. The other helps you build real systems that create real value. AI doesn't need more forecasts.

Self-driving cars didn't fail. They succeeded - just not in the form promised. Today's AI boom will follow the same path. The AGI promises will evaporate, but the infrastructure stays, enabling transformations we haven't imagined yet.

* * *

The core limitation is embarrassingly simple: AI can't see beyond the next word. These systems everyone's either worshipping or fearing - they literally cannot plan, cannot strategise, cannot see where they're going. They pick one word, then another, then another, like a driver who can only see one metre ahead. When ChatGPT writes you a business strategy, it's not strategising - it's just selecting words that statistically follow other words. The fact that this occasionally produces coherent plans isn't intelligence, it's proof that most business writing is so formulaic you can generate it without thinking at all. That discomfort is what makes many rationalise it as a thinking machine.

This book is an attempt to clear the air. It began from the same frustration you've probably felt: the conversation about AI is broken. Hype merchants and professional doomers dominate the stage. Revered figures recycle bad logic with total confidence. Industry leaders make confident but contradictory prognostications. The result is noise, not insight. The future doesn't need fortune-tellers. It needs maps. A way of seeing AI that helps us take the next step wisely instead of running from unfounded fears and chasing mirages.

Four lenses guide the analysis. **Infrastructure**: the slow, often boring shifts that eventually change everything. **Platforms**: who controls distribution and captures value. **Iteration**: the new physics of work, where retries are nearly free and breakthroughs and dead ends multiply, reshaping how we think. **Organisation**: how we adapt in response.

These lenses don’t predict AI’s final form - because it’s embryonic. What looks like an alien egg today will be shaped iteratively, step by step, through how societies build, govern, and respond. The point is not to blindly extrapolate in straight lines, but to orient ourselves. In John Boyd’s famous OODA loop, orientation is the hinge: where raw observations are turned into usable context, and where useful decisions are made. The same applies here. AI demands orientation - clearing the noise, recognising the terrain, and being able to act without wasting cycles on fantasy and fear.

The book has three parts. *Part I: Bad Stories About AI* looks at how hype and panic dominate the narrative - the bubble debates, the 'alchemy' promises, and the cult of AGI that blinds us to what's really happening. *Part II: How It Really Works* digs into the mechanics: why nothing works on the first try, why these systems behave more like pachinko machines than minds, and why 'mathematical objectivity' is anything but. *Part III: Who Wins, Who Works* explores what follows: how platforms capture value in chokepoint economics, and how jobs don't simply vanish but multiply and reorganise in unexpected ways.

Doomsday and utopia predictions? We've already seen enough of them evaporate on contact with light. These essays do something else. They’re about separating signal from noise and figuring out where we actually stand.

PART I:
BAD STORIES ABOUT AI

A HITCHHIKER'S GUIDE TO THE AI BUBBLE

> *The competition for AGI-AI that surpasses humans at all cognitive tasks-is of fundamental geopolitical importance.*

That's *The Economist*, **the other week**. Not some breathless tech blogger or venture capitalist talking their book. The world's most prestigious economic publication. Notice the framing - it treats AGI as a foregone geopolitical contest.

They're not wrong about the competition. They're just wrong about what we're competing for.

Last year I did something I hadn’t done in over a decade: I wrote code again. First time in 13 years. Not because I believe AGI is coming. I think it's alchemy-level nonsense. I started because I suddenly could. Because somewhere between the $560 billion in AI infrastructure spending and the endless debates about consciousness, something genuinely revolutionary happened: machine learning became boring infrastructure.

Boring is the highest compliment I can give technology. Boring means it works. Boring means you stop thinking about how and start thinking about what. Electricity is boring. TCP/IP is boring. And now, after all the hype and terror and mysticism, AI is getting boring too.

But you wouldn't know it from reading the headlines. When former prime ministers are writing op-eds about the AGI race, you know the fantasy has captured everyone - media, politicians, markets. They're so busy staring at artificial general intelligence that they're missing the actual revolution happening at ground level.

Two stories are unfolding simultaneously. One is a spectacular bubble built on geopolitical panic and sci-fi fantasies. The other is the quiet transformation of how we build everything. When the bubble narrative pops, the buildout accelerates.

WHAT'S WORKING TODAY

I'd been building systems since the 80s - architected investment fund migrations from mainframes to networked PCs in the City, built ERP for trading firms, then spent over a decade in enterprise consulting. But I hadn't written production code since 2012.

Within weeks last year I built a serverless system processing 5 million social media posts daily, tracking topic clusters and emerging narratives in real-time. Then brand monitoring dashboards. Then a 'robojournalist' that could deep-dive any trending story. Then hardware and firmware specs for a coffee machine. Then my first mobile app.

Not toy projects. Real systems. In the time it used to take to set up a development environment.

Thirteen years away from code, and within weeks I was shipping production systems in languages I'd never used. The tools had evolved that much.

Scroll through any tech community and you'll see senior developers emerging from semi-retirement like coders coming out of carbonite. People who'd graduated to PowerPoint and architecture diagrams, who barely touched an IDE in over a decade, are suddenly shipping products.

The vibe-coding community gets this. While we debate AI's impact in boardrooms, they're already building the future on Discord and shipping it to production. Yes, they're creating security nightmares and accidentally deleting production databases. Of course they are. They're inexperienced people wielding power tools. When we gave everyone electric saws, emergency rooms saw more accidents too. That's not an argument against power tools.

While established developers debate whether AI will replace them, these kids are shipping. Developers who learned their craft in the age of pull requests and sprint planning sneer at their security failures, not realising that 'best practices' are about to flip again. The barbarians aren't at the gate. They're deploying to production. And honestly? I'm a born-again barbarian myself.

And the patterns they're creating - spec-driven development, AI-first workflows - are already being productised by big tech. The innovation is flowing upward.

I'm not alone in seeing this. As Christina Wodtke, Stanford lecturer and early web pioneer, recently noted: 'The old timers who built the early web are coding with AI like it's 1995. The same people who ignored crypto and rolled their eyes at NFTs are building again. When developers who've seen every tech cycle since Gopher start acting like excited newbies, that tells you something'.

And it's not just about code. The other week GPT suggested I make jam from some obscure regional Thai plums I'd bought. Tiny things, sour as hell, no English name. I'd never made jam before. Took a photo of the carton, got the variety identified, received a recipe calibrated for their specific sourness (even spotted this meant no pectin needed), then real-time guidance that adjusted based on photos of my pot. 'Needs 3-4 more minutes', it said, looking at the bubble pattern. It was right. This capability - expertise on demand - is transforming everything from cooking to coding.

I work in Asia and see it daily: non-English speakers using AI as professional infrastructure. The language barrier just vanished - in both directions. People composing, analysing, and creating across languages at native level. The data backs this up: over 80% of ChatGPT traffic comes from outside the US, with massive usage in India, Brazil, Japan. The economic implications are staggering.

Students aren't asking 'Is this cheating?' They're asking 'How do I build with this?' They'll spend 40 years in the workforce. By the time they retire, working without AI will seem like working without electricity.

Something real is happening. Not in the research labs or board rooms where they debate ASI timelines. But in the million small moments where people discover something new they can do.

The old saying *'Be realistic, demand the impossible'* was never more true.

INFRASTRUCTURE HAPPENS

In 2018, if you wanted to use machine learning (ML), you hired PhDs and bought GPUs. Custom everything.

By 2020, you could rent pre-trained models from OpenAI or Google. But integration was still bespoke. Every company had different APIs, different formats, different assumptions.

Then something shifted. The models converged on common patterns: chat-style message formats, explicit system prompts, structured output modes, and a handful of consistent parameters. OpenRouter emerged to abstract away vendor differences. AWS Bedrock unified access to multiple models. Anthropic's MCP is pushing standard tool interfaces and has been adopted by everybody. What looked like competition was actually standardisation.

Watch what happened to prices once standards emerged:

- GPT-3 (2020): $60 per million tokens
- GPT-3.5 (2022): $2 per million tokens
- GPT-3.5 (2024): $0.07 per million tokens

That is the price curve you get when a capability becomes infrastructure.

The clearest sign: how new tools are built. Cursor runs their own prompt optimisation but routes to commodity LLMs for the heavy lifting. Replit does the same. They're not trying to compete with OpenAI on model training. They're building experiences on top of commodity LLMs.

This is textbook platform evolution - exactly what Simon Wardley has been **mapping for years**. Standardisation enables scale. Scale drops prices. Low prices increase adoption. Adoption creates ecosystem.

We see it today: LLM APIs for understanding. Embedding models for similarity. Vector storage for search. Components from different vendors. What cost millions to build custom in 2018 now costs hundreds to assemble. My social media analysis system is built entirely from such commodity blocks.

The AGI crowd misses this completely. They're debating consciousness while the machine learning stack is commoditising under their feet. They're worried about 'superintelligence' while developers are treating AI as just another API to call.

The real revolution isn't making machines think. It's making them boring enough that nobody has to think about them.

THE MISSILE GAP, BUT STUPIDER

We're spending $16 for every $1 earned in AI. That 16:1 investment ratio only makes sense if the winner takes everything.

Which is exactly what everyone believes.

Rishi Sunak writes op-eds about democratic values in the AGI race. The White House issues executive orders about AI safety. China announces AI supremacy targets. The EU drafts regulations for systems that don't exist. Everyone's racing for permanent technological supremacy.

This is your bubble. Not the technology - the shared delusion that someone's about to achieve irreversible computational dominance.

The panic has a patient zero. When Geoffrey Hinton quit Google to warn about AI risk, he didn't just change jobs. He transformed a technology discussion into an existential race. Suddenly every major power faced a terrifying question: What if our enemies get AGI first?

Sam Altman knew exactly which buttons to push: congressional testimony about the need for regulation (from the company furthest ahead), warnings about AI risk, and a playbook of building in public while presenting OpenAI as the responsible actor that just needs resources to 'do it safely'.

It worked. The Stargate announcement - $500 billion for AI infrastructure - is the logical endpoint of this narrative. When you believe you're racing for permanent species-level advantage, no amount is too much. The 'long-termists' have everyone convinced we're at the hinge of history.

The curious thing about existential arms races? They're incredibly profitable for arms dealers.
Jensen Huang needs governments to panic-buy GPUs. Sam Altman needs infinite capital for compute. Microsoft, Google, and Amazon need regulatory moats only they can afford. Every warning about AGI danger is also a pitch deck.

During the Cold War, the US and Soviets would leak reports about UFOs and mind control programs. Deliberate misdirection to waste enemy resources. The AGI race has the same dynamics - except this time, everyone's falling for their own propaganda.

At least the missile gap was about real missiles.

EVERYONE'S LOOKING UP

The most interesting part of **Ed Zitron's recent 14,000-word AI takedown** isn't what he gets wrong. It's how he gets it wrong.

He spends thousands of words debunking AGI hype, then judges every AI product by AGI standards. He dismisses 'agents' because they're not fully autonomous. He mocks chatbots for not being conscious. He's so busy fighting the fantasy that he misses the reality.

He's not alone. The entire discourse has been captured by AGI framing. Critics and believers alike judge current AI by science fiction standards. It's like dismissing cars because they don't fly.

This is what Baldur Bjarnason called the '**LLMentalist effect**' (great article!) - we've projected consciousness onto pattern matching. The chatbots seem so human that we can't help but evaluate them as minds rather than tools. Even skeptics fall into the trap, spending more time debating whether they're 'truly' intelligent than asking whether they're useful.

Real revolutions happen gradually, then suddenly. In 1996, if you asked for proof the internet would change everything, what could anyone show you? Amazon selling books? Email replacing faxes? The transformative applications hadn't been invented yet because the infrastructure didn't exist.

We're in the same moment now. People demand to see the AGI-level breakthrough while missing the million small transformations already happening. My social media analysis system would have been impossible five years ago. Not impractical - impossible. The components didn't exist at any price.

Every week, developers uncover new patterns. Natural language becomes the interface for everything. Retrieval makes search contextual instead of literal. Multi-step reasoning chains that once collapsed now hold together. Not consciousness, but capability after capability that wasn't there before.

Meanwhile the hype cycle and the hand-wringing predictably miss the point.

The skeptics and believers are having the wrong argument - two sides of the same shitcoin. The C-suite fence-sitters striking a 'balanced perspective' are no better, hedging between factions that both lost the plot, serving up a smorgasbord of bad takes. The question isn't whether we'll create AGI. It's whether we'll notice that we don't need to.

WHEN THE MUSIC STOPS

The AGI bubble will pop. Not because the technology fails, but because the fantasy can't survive contact with reality.

The trigger could be anything. An AI company admitting AGI is decades away. A government realising it stockpiled GPUs for nothing. Or investors noticing that $560 billion for $35 billion in revenue isn't a business model so much as a cargo cult.

When it happens, the narrative collapse will be spectacular. All those breathless headlines about consciousness and superintelligence will age like dot-com era predictions about the 'new economy' where profits didn't matter. The Stargate project will become this generation's Webvan - ambitious, well-funded, and built on false premises.

But the doomsayers miss something crucial: the infrastructure remains. After the dot-com crash, we still had fibre optic cables, data centres, and trained engineers. The speculation died. The internet didn't.
Same pattern here. When the AGI fantasy evaporates, we'll still have:

- Models that can read, write, and analyse
- APIs that cost pennies to call
- A generation of developers who know how to build with them
- Actual products solving actual problems

The companies that survive won't be the ones promising AGI. They'll be the ones who understood early that machine learning is just really useful when available as infrastructure. Like the difference between Pets.com and Amazon - one promised to change the world, the other was building warehouses.

Medieval alchemists never turned lead into gold. But while chasing that impossible dream, they invented chemistry. They failed at transmutation but succeeded at something more valuable: understanding how the world actually works.

Same story, new century. The AGI labs won't crack consciousness. But chasing the ghost in the machine, they've built infrastructure that changes everything. It turns out the bubble was the wrapper all along.

CONCLUSION: NOW WHAT?

So what do you do with this knowledge?

If you're a developer: build. The tools are here, they're cheap, and they're getting better every week. While everyone else debates consciousness, ship products. The barbarians aren't knocking - they're already through the door with mud on their boots.

If you're a business: ignore the noise. Everyone has strong opinions about AI, and they're mostly wrong. While experts argue and vendors overpromise, focus on what works today. That boring automation, that small efficiency gain, that better interface - these compound. The companies that win won't be waiting for clarity. They'll be the ones who started with simple tools and learned by doing.

If you're an investor: you understand the moat dynamics - infrastructure players need scale, builders need distribution, data, or workflow lock-in. Want to make money on AI? Bet on the boring stuff. The companies making the tools everyone else relies on. The ones using AI to shave 3% off shipping costs. The businesses that would thrive even if we proved tomorrow that consciousness is mathematically impossible. They're building for the world where AI is infrastructure, not magic.

If you're a government: yes, sovereignty matters. You need domestic compute and models you control. But the race isn't for AGI - it's for practical ML capability. The question isn't whether to invest in infrastructure, but how much is enough. With open models improving and inference costs plummeting, the barriers are lower than the panic suggests. Build what you need, not what the arms race demands.

The hardest part isn't understanding the technology. It's seeing past the narrative.

When the bubble pops, the pundits will act shocked. How did we spend $560 billion chasing digital consciousness? How did The Economist fall for it? How did governments stockpile GPUs for a race that couldn't be won?

But by then it won't matter. The builders will have inherited the infrastructure. The vibe-coders will be running production. Your competitors will be shipping features you thought impossible. And everyone will pretend they knew all along that the real revolution was never AGI.

It was making intelligence so boring that nobody thinks twice about using it. Just like electricity. Just like the internet. Just like every transformation that actually mattered.

* * *

The bubble story is seductive. It fits our pattern-matching brains perfectly. We reflexively cite tulips, but that was pure speculation. The better parallels are railroads and dot-coms - transformative technologies where speculation funded infrastructure that survived the crash.

But pattern matching can blind as much as illuminate. When you're looking for a bubble, everything starts to look like froth. The massive funding rounds, the breathless headlines, the questionable economics - all the signals are there. $560 billion spent to generate $35 billion in revenue. Case closed, right?

Not quite. The bubble framework assumes AI is a product category that will boom and bust. But what if it's not a product at all? What if we're watching infrastructure emerge - the kind of boring, essential infrastructure that never has a satisfying crash because it becomes too embedded to fail? The critics waiting for an AI winter might be waiting forever, not because the hype is justified, but because they think they’re looking at tulips.

The real question isn't whether AI is a bubble. It's whether we're so trapped by the bubble metaphor that we can't see what's actually happening: a platform shift that looks like speculation but works like evolution. The distinction matters. Bubbles pop. Platforms persist.

* * *

BUBBLES POP.
PLATFORMS PERSIST.

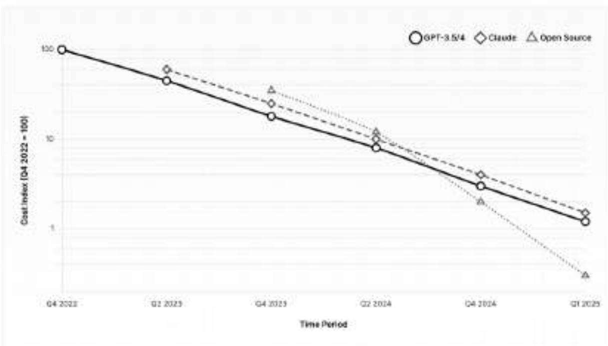
The bubble story didn’t just surface on Twitter threads and VC blogs - it found its definitive expression in a viral 14,000-word essay by Ed Zitron. That piece became the go-to citation for anyone claiming ‘AI is a bubble’. His numbers were crisp, his tone scathing, and the conclusion irresistible.

The case, as he makes it, rests on three pillars. First, the headline ratio: $560B spent to generate $35B in revenue - a 16:1 burn unmatched even by early cloud. Second, the AWS comparison: if even the canonical loss-leader platform didn’t look this bad at the start, AI must really be cooked. And third, the product critique: outside of Nvidia, nobody is vV money, everything else is ‘wrappers’, and the dream of autonomous agents has already collapsed.

On its face, that’s a strong indictment. Which is why it spread. But it collapses once you shift frames - from ‘AI as product bubble’ to ‘AI as infrastructure buildout’. That’s the context in which his claims don’t just wobble; they break.

FIGURE 1 • COST TRAJECTORY

AI inference costs per million tokens at comparable performance levels (2022-2025)



KEY INSIGHT

280× cost reduction in 24 months. The dramatic decline in AI inference costs follows a steeper trajectory than Moore's Law, with open-source models (shown in dotted lines) achieving near-parity performance at commodity prices. This cost collapse fundamentally alters the economics of AI deployment across industries.

DATA: OpenAI, Anthropic, Google Cloud, Meta AI public pricing data; compiled January 2025

That 16:1 ratio is unprecedented. He's absolutely right to highlight it.

But that's where his analysis peaks. After that, he swallows the hype he claims to hate.The cost story is more complex.

Yes, frontier models are expensive. They're also new. GPT-3.5 inference costs dropped 280× in under two years. Older frontier models already price at pennies per million tokens. Blackwell delivers 4× Hopper throughput on Llama-70B. This is a standard tech cost curve.

Cursor was Zitron’s centrepiece. He argued that even the breakout success - an AI coding assistant - proved the bubble case when it scrapped flat-rate pricing. His evidence? A handful of Reddit threads where heavy users complained about being pushed onto usage-based plans.

That’s not analysis. That’s sentiment scraping. Every SaaS platform has this cycle: early flat rates buckle under heavy usage, superusers rage on forums, the model shifts to metered pricing, and the business carries on. Netflix throttled DVD hogs. AWS re-tiered S3. Zoom clawed back free tiers. Cursor did the same.

Zitron read Reddit angst as proof the model was broken. In reality it was proof demand was strong enough to break mispriced tiers.

JUDGING CARS BECAUSE THEY DON'T FLY

Zitron spends thousands of words debunking AGI hype - then judges everything by AGI standards. He dismisses ‘agents’ because they're not fully autonomous systems.

Companies are shipping workflow tools, CRM automation, code helpers. Real customers pay for them. But Zitron can't see the value because they're not self-directed. He's measuring real products against flying cars.

Infrastructure always looks excessive at first. That $560B isn't just spending - it's a moat for the platforms. Good luck competing when entry costs are half a trillion. Microsoft's $3B real AI revenue on $80B capex looks brutal - but AWS had the same optics in 2010, then flipped to 40% margin when usage caught up. The infrastructure IS the competitive advantage.

Platform risk? Every major tech firm starts under someone else's thumb. Customer concentration? Intel spent years living off a few OEMs. These aren't new problems - they're just normal platform adolescence.

GOOGLE WAS JUST A CRAWLER

Zitron lists a handful of obvious use cases - chatbots, search, code generation - then declares that's all LLMs can do, that all applications built on them are just 'wrappers', a dismissive term popularised by skeptical developers on Hacker News.

Meanwhile I'm shipping products that don't fit any of his categories. A mobile app that aggregates usage patterns into market insights. A system processing 5M social posts daily to surface breaking news. Brand intelligence tools that actually work. These aren't 'wrappers' any more than Google is a 'web crawler wrapper'.

Builders create different moats to platform operators. For builders the moat isn't the LLM - it's the data pipeline, the user experience, the business logic. Calling these 'LLM wrappers' is like calling a person a 'skeleton wrapper'. Technically true, totally misses the point.

THE 16:1 PROBLEM

That ratio is genuinely alarming. Maybe it normalises as costs drop. Maybe it signals something broken.

I'm betting on normalisation. He's betting against it - with Reddit anecdotes as evidence.

Zitron spends 14,000 words torching AGI fantasies, but then judges real products against sci-fi benchmarks. The boring stuff already works. He just can't see it through the hype.

What Zitron's essay really demonstrates isn't that AI is a bubble - it's how seductive the bubble metaphor is. Once you adopt it, every data point slots neatly into the mania-crash arc. Step outside it, and the picture changes. These aren't products floating on hype. They're infrastructure settling in, with all the ugly early economics that entails.

Bubbles pop. Platforms persist. And the lesson of the 'AI bubble' narrative is less about AI than about our appetite for familiar stories.

* * *

The first essay brushed AGI off as alchemy. The next one explains why. Because if the bubble story is a familiar myth, it isn’t the only one in circulation. Another camp has a stranger bubble - not financial, but theological. The AGI believers inflate a market of faith. They talk of the rapture of digital minds, then quietly roll it forward: the next release, the secret system in the lab, the capabilities too dangerous to reveal. Always just around the corner. Always deferred.

The same people who admit today’s AI is ‘just statistics’ are simultaneously convinced we’re months away from artificial consciousness. They acknowledge current systems can’t truly reason, can’t really understand, can’t actually think - yet insist that more compute and clever algorithms will somehow birth a digital mind. It’s not a contradiction if you understand their real belief: they’re not doing engineering, they’re doing alchemy.

The modern alchemists don't seek to turn lead into gold - they promise to turn computation into cognition. Their philosopher's stone? More parameters, better architectures, novel training techniques. Like Newton chasing transmutation, they're brilliant people pursuing an impossible goal. And like the original alchemists, they'll build useful things while failing at their ultimate quest.

The alchemy metaphor isn't casual. It's exact. The psychological patterns, the social dynamics, the mixture of genuine insight and fundamental confusion - all of it maps perfectly onto Silicon Valley's AGI obsession. Understanding this parallel helps explain why smart people believe impossible things, and why their impossible quest might still change everything.

* * *

ALCHEMY 2: ELECTRIC BOOGALOO

Why the dream of AGI rests on undiscovered mathematics, biochemical hand-waving, and Silicon Valley's accidental religion.

Isaac Newton spent more time on alchemy than physics.

The man who gave us calculus, who explained gravity, who literally invented modern science - he spent decades trying to turn lead into gold. He wrote over a million words on alchemy, conducted thousands of experiments, built furnaces in his Cambridge rooms. He was convinced the secret was there, just out of reach.

He wasn't alone. Robert Boyle, the father of chemistry. Tycho Brahe, who mapped the stars. Even Leibniz dabbled. The brightest minds of the Scientific Revolution believed that with enough intelligence, enough effort, enough faith, they could crack the code of transmutation.

They never made an ounce of gold.

When I called AGI '**alchemy-level nonsense**'. The response was predictable: 'You're looking at the wrong timescale'. 'It's not about 5 years, it's about 50'. 'These are the smartest people in tech - give them time'.

But I'm not talking about timescales. I'm not saying AGI is hard, or that it will take longer than expected. I'm saying it's impossible. Not 10 years impossible. Not 100 years impossible. Never-ever impossible. Lead-into-gold impossible.
I say this because years ago, **Dr. Roman Belavkin** - Reader in Informatics at Middlesex University, known for his work on information theory, decision-making, and AI - explained to me why we currently can't even model a single biological neuron, and it is unknown whether we shall ever be able to. The maths doesn't exist yet. When I've mentioned this to other AI enthusiasts, they dismiss it as 'just another challenge to overcome'.

So let me share what mathematicians who study cognitive systems have been trying to tell us. Why the smartest people in tech are chasing something that violates the constraints of current mathematics and biology. Why all the compute in the world won't help, any more than better furnaces would have helped Newton make gold.

The parallels run deeper than metaphor. Just like the alchemists, today's AGI seekers are brilliant, dedicated, and completely wrong about what's possible. But also like the alchemists, they're building something valuable while chasing the impossible.

ALCHEMISTS WERE ACTUALLY BRILLIANT

We mock medieval alchemists now, but they weren't fools. For over two thousand years, alchemy attracted the greatest minds in science. Newton spent decades on it. Boyle, the father of chemistry. Jabir ibn Hayyan, the Islamic polymath. Chinese emperors employed thousands of alchemists. This wasn't fringe science - it was science.

They were empiricists working from observable facts: gold and silver resist rust, but iron corrodes; lead and gold share weight and softness; mercury is liquid metal that dissolves other metals. Surely these shared properties meant metals were related - variations of some primary substance that could be transmuted with the right process?

This wasn't magical thinking. It was rational inference from available evidence. They could alloy metals, change their properties with heat, dissolve and precipitate them. They watched caterpillars become butterflies. Why not lead into gold?

The ‘magical’ reputation of alchemy comes more from the overlay of mysticism and secrecy than from the original practical investigations, which were closer to protochemistry.

The social dynamics helped. European princes needed gold to fund wars. Islamic scholars sought the elixir of life. Chinese emperors wanted immortality. Every civilisation, every power structure, had reasons to fund the dream. For centuries, the smartest people alive, backed by unlimited resources, chased transformation.

And it worked - just not how they expected. Alchemists invented distillation, crystallisation, sublimation. They discovered phosphorus while trying to extract life force from urine. They mapped acids and bases while seeking universal solvents. They created gunpowder while chasing immortality.

The dream sustained itself through a paradox: every failure taught them something useful. Can't make gold? Here's how to purify silver instead. Can't find the elixir of life? Here's how to make better medicines. The impossible goal funded the possible discoveries.

By the time chemistry emerged as a real science, alchemists had built the entire laboratory tradition. The equipment, techniques, and methodology that would reveal why transmutation was impossible came from centuries of trying to achieve it.

The alchemists had better evidence for their beliefs than AGI proponents today have for theirs. They could see metals shared properties. They could demonstrate partial transformations. They had thousands of years of metallurgy suggesting malleability.

What evidence do we have that silicon can become conscious? That backpropagation - which requires smooth, differentiable functions - can somehow model neurons that are definitionally non-differentiable? That pattern matching in text can become understanding?

The alchemists at least started with lead and gold - two things that actually exist. We're starting with statistics and consciousness - and one of those might not even be a thing.

THE PUDDLE WAKES UP

Douglas Adams once wrote about a puddle that wakes up one morning and thinks: 'This is an interesting world I find myself in - an interesting hole I find myself in - fits me rather neatly, doesn't it? In fact it fits me staggeringly well, must have been made to have me in it!'

The puddle is, of course, exactly wrong. The hole wasn't made for water. Water just takes the shape of whatever contains it. But from the puddle's perspective, the fit seems too perfect to be coincidence.

This is Silicon Valley's relationship with intelligence. We've built machines that process text, and now we think intelligence is text-shaped. We've created systems that recognise patterns, so we believe consciousness is pattern recognition. The fit seems too perfect to be coincidence.

Robert Epstein, a cognitive scientist, puts it bluntly: '**Your brain is not a computer**'. We've been seduced by our own metaphor. Brains don't store memories like files. They don't process information like CPUs. They don't run algorithms. The computational metaphor was useful once, but we've forgotten it's a metaphor.

Yet here we are, building a religion around it. And I mean that literally. AGI serves the same psychological function as religion: it promises transcendence, meaning, purpose. It offers Silicon Valley a secular rapture where we upload our consciousness and defeat death.

Look at the language: *The Singularity. Alignment. Existential risk.* These aren't technical terms - they're theological concepts. We have prophets (Kurzweil), apostates (Hinton), and heretics (anyone who doubts). We have sacred texts (Turing and Emil Post's papers), origin myths (the Dartmouth Conference), and an eschaton (AGI).

The social dynamics are identical to religious movements. True believers get funded and platformed. Skeptics are dismissed as ‘not understanding exponential growth’ - the tech equivalent of lacking faith. Every failed prediction gets memory-holed while every incremental advance proves the prophecy.

The puddle thinks the hole was made for it, wrongly projecting its shape onto reality. We think intelligence is computational.

Which brings us to the mathematics. Because reality - particularly biological reality - has a very specific shape. And it's not differentiable.

MATHEMATICS OF IMPOSSIBLE THINGS

Scientists recently mapped one cubic millimetre of human brain tissue - a piece the size of a grain of sand. It took 1.4 petabytes of data. Not to simulate it. Not to model it. Just to store a static 3D photograph.

One grain of sand worth of brain tissue, frozen and dead, and it takes more data than Netflix uses to stream to a small country. We weren't even trying to capture how it works - just what it looks like.
Inside that grain: 57,000 cells and 150 million synapses, each one a chemical factory with over 1,000 different proteins that shift and change based on what's happening around them. That's the photograph. The frozen moment. To model how it actually works? We don't even know where to start.

This is where the dream of artificial general intelligence hits a wall that nobody talks about. Years ago, I was discussing AGI with my friend Dr. Roman Belavkin. Over coffee, he explained the problems with modelling biological neurons that should be headline news but instead stay buried in academic conferences.

I needed to understand: were these problems speed bumps or brick walls?

The more Roman explained, the clearer it became to me. These aren't speed bumps - they're mathematical voids.

The problem starts with backpropagation.

Let me explain what that means, because it's the heart of why AGI is impossible with current mathematics.

Most deep learning systems today - ChatGPT, Claude, Gemini - rely on an algorithm called backpropagation. Think of it like teaching a child: you show them a picture of a cat, they guess 'dog', you tell them how wrong they were, and they adjust their mental model. Do this millions of times and they learn to recognise cats.

THE 1943 PROBLEM

Why don't we just model real neurons? While we can model neurons at many fidelities - the best are still cartoon sketches. The original McCulloch-Pitts model from 1943 used simple on/off switches with step functions - not differentiable, couldn't work with backpropagation (which hadn't been invented yet). In the 1980s, to make backpropagation work, we replaced the step functions with smooth curves. But we're still using the same basic framework: weighted sums, linear algebra, one-dimensional signals.

If we tried to model biological detail - the neurotransmitters, the dendritic processing, the living cellular machinery - we don't even have the mathematics to describe it, let alone compute it. A single biological neuron isn't a switch - it's a living cell, a chemical factory floating in a soup of neurotransmitters, hormones, and proteins.

When a signal arrives at a neuron, it doesn't just flow through like electricity through a wire. The signal arrives at branches called dendrites, and here's where everything breaks: each branch does its own processing. Not simple addition or multiplication - complex, unpredictable transformations that we struggle to describe mathematically.

Imagine you're trying to predict the flow of a river. But this river has thousands of tributaries, and each tributary follows its own rules - some flow uphill, some disappear underground and reappear elsewhere, some spontaneously change direction based on the phase of the moon. Now try to write an equation for where a drop of water will end up. That's what we're trying to do with neurons.

TWO KILLERS

Roman described a number of fundamental issues with current approaches to AGI. As I understood them, these aren't engineering challenges - they're mathematical impasses. I'll cover two of them here:

FIRST: CHEMISTRY THAT DOESN'T EXIST IN AI

There are no neurotransmitters in artificial neural networks. No dopamine, no serotonin, no GABA, no glutamate - nothing. We're missing the entire chemical dimension of neural signalling.

As Roman explained, these neurotransmitters aren't just signals - they're dimensions. To model this would require a new multilinear (tensor) algebra, not the simple linear algebra we use. IBM is trying to invent **'tensor-tensor algebra'** because current maths literally cannot describe these relationships. Think about what that means. We need mathematics that doesn't exist yet.

SECOND: THE FROZEN BRAIN PROBLEM

Brains don't train then deploy. Every perception changes the brain perceiving it. Every thought rewrites the apparatus that produced it. Learning and performing happen simultaneously, continuously, and inseparably.

Our models can't do this. Not won't - can't. Real plasticity means the architecture itself reorganises with every input. The connections don't just change weights - they form, die, restructure. We don't have mathematics for systems that fundamentally rewrite themselves while running.

The current state of AI, Roman explained, is like 'putting wires into a dead brain and trying to get answers'. The network has been trained, its weights are frozen. What you're talking to is a fossil of the training process.

Yes, they're bolting on RAG systems and calling it ‘continuous learning’, but that's like taping a notebook to a statue and calling it alive.

Neurons are living cells. If life-like processes are prerequisite to general intelligence, we currently lack a mathematical definition of 'being alive', let alone an engineering analogue. That's a pre-theoretic gap, not a compute gap.

OTHER PROBLEMS STACK UP

As if those two killers weren't enough:

Real neurons spike - sudden, all-or-nothing electrical pulses. Researchers use 'surrogate gradients' to work around this, but these workarounds don't capture the actual dynamics.

Glial cells - which we ignored for a century - actively shape how neurons communicate, release their own chemicals, and outnumber neurons in some brain regions. We're missing more than half the system.

AN IMPOSSIBLE ENGINEERING PROBLEM

The optimists have their responses, of course.

‘We don't need biological fidelity’, they say. ‘Airplanes don't flap their wings’.

True. But humans had been studying flight systematically for centuries - Leonardo da Vinci's *Codex on the Flight of Birds* dates to 1505. By the early 1800s, George Cayley had established the fundamental principles: lift, drag, thrust, weight. The Wright brothers weren't discovering new science; they were solving engineering problems with materials and engines. They understood the physics and could define what flight meant.

And here's the deeper problem: if you're not modelling biological intelligence, what exactly are you modelling? The brain is the only example we have of intelligence. There's no other data point. When you abandon biological fidelity, you're not building toward intelligence - you're building something else entirely and hoping it turns out intelligent.

When cornered, suddenly everything is 'intelligent' - plants, fungi, ant colonies, flocks of birds. But when pitching to investors and governments, AGI means 'surpassing humans at all cognitive tasks'. Ask an AGI researcher: what's the function you're optimising? What principles of intelligence have you extracted? They'll give you extrapolations, not specifications. You can't engineer your way to a destination you can't define.

'Intelligence will emerge at scale', they insist.

Yes, scale helped with translation. But translation is pattern matching between texts. Emergence doesn't mean magic. What rule of matrix multiplication creates consciousness? Which property of transformer architectures generates understanding? They can't tell you because they don't know. Believing scale will create consciousness is like believing a tall enough ladder gets you to the moon.

'But Turing proved computation is universal'.

People love saying 'computation is universal', like that settles the argument. It doesn't.
All it means is this: if you already have the algorithm, then yes, a computer could run it - given unlimited time and memory. That's the bar. It doesn't mean the algorithm exists, or that we could find it, or that it would finish before the heat death of the universe.

We don't have an algorithm for consciousness. We don't have a formal spec. And it's not just 'a function from sensory inputs to outputs' - that's oversimplified. Current theories suggest it may require time, memory, embodiment, and internal state - but even those lack agreed definitions, let alone code.

Turing's proof is about logical possibility, not practical reality. It doesn't prove anything about whether we can discover the algorithm, afford to run it, or even properly define the problem. We can't even say whether the brain is computing in Turing's sense.

Saying ‘consciousness is computable because we have Turing-complete machines’ is like saying ‘immortality is possible because it doesn’t violate logic’. Technically true to a logician. Fraudulent to a decision maker.

And once you've realised the mind isn't computable, the whole GPU-scaling narrative collapses. We’re not even building a ladder to the moon - we’re stacking chairs in a basement.

* * *

That’s not just an awkward implication. It’s a dead end.

Iris van Rooij and colleagues at Radboud argue that even with unlimited compute, scaling current approaches doesn't get you to AGI. It's not a scale problem. It's an architecture problem. A mathematics problem. A we-don't-even-know-what-we-don't-know problem.

The very success of current AI makes the problem worse. ChatGPT is so good at mimicking understanding that people think it actually understands. It's like being impressed by a parrot reciting Shakespeare and concluding the parrot must understand iambic pentameter.

But watch what happens when these systems fail. They don't fail like humans fail - forgetting a detail or mixing up names. They fail in ways that reveal the complete absence of understanding. They'll invent legal citations that sound perfect but never existed, getting lawyers sanctioned. They'll confidently explain historical events that never happened. They'll maintain perfect grammar while **contradicting themselves from one sentence to the next**. Not mistakes - glitches that show there's no coherent world model underneath.

Change a few pixels in an image - invisible to humans - and suddenly the AI thinks a panda is a school bus. Why? Because it's not seeing a panda. It's detecting statistical patterns that happen to correlate with the label ‘panda’ in its training data. The patterns are brittle, hollow, nothing like understanding. We can see it’s a parlour trick in the sense that the results can seem uncanny, but there's no reasoning underneath. Yet it's not a trick in the sense of uselessness - these systems are highly capable and usable, *just not conscious*.

A NOTE ON 'IMPOSSIBLE'

'But you can't prove it's impossible!'

Correct. Scientists can't prove gravity won't reverse tomorrow either. Yet no one's funding anti-gravity startups. No one's spending $400 billion trying to flip the Earth's gravitational field. Somehow, in AI, the burden of proof only flows one way.

When I say AGI is impossible, I mean this: it would require mathematics that doesn't exist, to model biology we don't understand, to implement functions nobody can define. That's impossible in any meaningful or operational sense.

If you're an investor, policymaker, or journalist: when the timeline is 'somewhere between tomorrow and never', bet on never. When the requirement is mathematics we haven't invented, that's not a roadmap - it's a prayer.

The honest position isn't 'we can't rule it out'. It's: AGI is impossible until proven otherwise. And that proof starts with the missing maths - not a demo, not a promise, not a slide deck.

WHAT WOULD CHANGE MY MIND

Show me the tensor maths behind neurotransmitter interactions. Show me a system that genuinely rewires itself while running - not RAG, not fine-tuning, but actual structural plasticity. Give me a non-circular definition of intelligence that doesn't move its goalposts every time we approach them. I don't want promises or roadmaps or claims about 'seeing sparks'. I want working mathematics, clear definitions, real mechanisms.

THE BIOLOGICAL MIMICRY TRAP

You'll often hear a reasonable-sounding objection: 'We've beaten nature plenty of times without copying how it works'.

And sure - calculators don't use neurons, just silicon. Wheels outperform legs. Engines outperform muscles. Submarines don't have gills. Planes don't flap their wings.

We've also copied nature when it suited us: velcro from burrs, sonar from bats, drugs from plants. But in every case - whether we mimicked biology or not - we understood the principles first.

We knew how arithmetic worked. We understood lift and drag. Even the stuff we borrowed from biology made sense mechanistically before it scaled.

That's what makes AI different.

The only part of the field claiming real progress toward AGI is built on neuron-inspired, gradient-trained architectures. OpenAI, Anthropic, DeepMind - they're all in on neural networks, betting that scaling these artificial neurons will somehow give rise to intelligence.

But when pressed on how biologically plausible any of this is, they retreat.

OpenAI, whose entire stack is built on artificial neurons, now defines AGI as 'a highly autonomous system that outperforms humans at most economically valuable work'.

Ray Kurzweil, long time prophet of brain modelling, says AGI just needs to 'match what an expert in every field can do, all at the same time'.

This is bait-and-switch. They lean on biology to justify the architecture - then invoke functional definitions to escape scrutiny - while continuing to pour everything into the neuron-modelling path. No mechanisms. No theory. No alternative.

And if artificial neurons can't get us there - which they can't, because the mathematics to model real ones doesn't and may never exist - there's no Plan B. Scaling and scaffolding aren't a theory. You don't get to bet everything on neuron mimicry while claiming the biology doesn't matter.

The brain is still the only working example of general intelligence - evolved, embodied, socially embedded, and built on millions of years of biological hacks. Any AGI claim has to reckon with that. Ignore that and you're not doing engineering -

you’re playing armchair consciousness philosopher, all first principles and no prototype.

If mimicry is the path, show the biology. If it’s not, show the theory. Right now, we’ve got neither - just metaphors with a GPU budget.

The sophisticated will claim I'm strawmanning - 'Nobody serious claims we need biological fidelity!' But that's the point: when you abandon biological fidelity but keep biological architectures, you need a theory of why. Airplanes abandoned wing-flapping but had aerodynamics. What's AGI's equivalent? 'Emergence at scale' isn't a theory - it's what you say when you don't have one.

THE BETTER FURNACE FALLACY

The alchemists kept building better furnaces.

Each generation convinced themselves they were making progress. Hotter temperatures, purer materials, more precise measurements. The furnaces of 1400 were primitive compared to those of 1600. Surely they were getting closer to transmutation?

They weren't. Not one percent closer. Not one thousandth of a percent closer. Zero percent closer. Because transmutation requires nuclear physics, not better heating.

Today's AI researchers are building better furnaces. They call them GPUs.

GPT-3 had 175 billion parameters. GPT-4 has over a trillion. The training runs cost millions, then tens of millions, now hundreds of millions. OpenAI's latest cluster has 100,000 GPUs. Microsoft and Google are planning million-GPU clusters. The furnaces keep getting bigger.

Sam Altman wants $7 trillion for compute. Seven trillion dollars of furnaces. He might as well be asking for $7 trillion to build a furnace that reaches the sun's core temperature. It still won't make gold.

This is the most expensive category error in history.

THE COMPUTE CULT

The deep learning revolution created a dangerous syllogism:

1. We scaled compute and got better language models
2. Better language models seem more intelligent
3. Therefore, more compute equals more intelligence

This is like observing:

1. We made furnaces hotter and could melt more metals
2. Melting metals seems closer to making gold
3. Therefore, hotter furnaces equal transmutation

The leap doesn't follow. But once you've invested billions in furnaces, it's hard to admit you're not actually making progress toward gold.
Nvidia is selling furnaces to alchemists. And business is booming - their market cap hit $3 trillion on furnace sales alone. Every AI lab, every tech giant, every government is panic-buying furnaces, terrified of being left behind in the transmutation race.

WE'RE NOT 'ON THE WAY'

Here's what the AI establishment doesn't want to admit: we're not 'on the way' to AGI. We're not even on a path. You can't be on the way to a place that requires different physics.

The Guardian recently quoted tech analyst Benedict Evans calling the entire AGI race 'vibes-based'. Aaron Rosenberg from Radical Ventures immediately redefined AGI as '80th percentile human-level performance in 80% of economically relevant digital tasks'. Watch that goalpost move. When the furnace-builders realise they can't make gold, they'll announce they were always trying to make bronze.

This year alone, according to the Guardian, tech companies will spend $400 billion on AI infrastructure. Not on research into the missing mathematics. Not on understanding biological neurons. On compute. On bigger furnaces.

That's more than the EU's entire defence budget. More than the GDP of many nations. All betting that consciousness emerges from heat.

THE HUMAN FACTOR

A recent exchange crystallised this for me. When I mentioned AGI's impossibility, someone immediately defended the field's luminaries: 'Hinton, Tegmark, and Bostrom surely aren't fools or shills'.

Not fools or shills. Hinton revolutionised neural networks - brilliant work, just not neuroscience. Tegmark is an accomplished physicist - but this isn't physics. Bostrom... well, the less said the better.

The point is: Hinton understands his mathematical abstractions perfectly. But the gap between those abstractions and biological reality? That's where the impossibility lives.

This isn't conspiracy. It's specialisation. The AI researchers perfecting backpropagation don't study molecular neuroscience. The neuroscientists mapping synapses don't design learning algorithms. The funders reading executive summaries don't see the chasm between fields.
Nobody has to lie. Everyone just has to stay in their lane.

It brings to mind a conversation I once had with an architect for one of the early UK nuclear power plants. She told me the architectural team had raised concerns about nuclear waste from the start. They were told not to worry - the scientists were working on it and would have it solved by the time it mattered.

That was over sixty years ago. The waste is still there. The solution never came.

Same pattern, different field. Everyone assumes someone else has the hard part covered. The architects trusted the physicists. The physicists trusted future physicists. Nobody was lying. Everyone was just wrong about what was possible and assumed that difficult just meant *'not yet'*.

The field self-selects for optimists - people who believe artificial intelligence is possible, who want to be the ones to crack it. You don't go into AI research if you think it's impossible. You go into AI research because you dream of building minds. So the very people who might spot the fundamental problems are filtered out before they start.

Geoffrey Hinton, the 'Godfather of AI', told Wired: 'We really don't know how [deep neural networks] work'. He says their success 'works much better than it has any right to' - an empirical discovery that theory can't explain. **Yoshua Bengio** acknowledges that current systems learn in a 'very narrow way' and make 'stupid mistakes', lacking the mathematical foundations for reasoning or causal understanding. **Yann LeCun** admits nobody has 'a good answer' for how to make neural networks capable of complex reasoning.

In 2017, Google researcher Ali Rahimi **called the entire field'alchemy'** - techniques that work without understanding why. The community erupted in debate, but the core point stood: we have amazing tools with no satisfying theory.

But somehow, when you put them in a room together, the specialisation creates blind spots. They're all brilliant people climbing different faces of an impossible mountain, each assuming someone else has found the route to the top.

A commenter on Hacker News captured today's version perfectly: 'If you believe the brain is a biological computer and AI computing keeps advancing, at some point it will be able to do the same stuff. It's just common sense'.

Common sense. The same common sense that told us heavier objects fall faster. That the sun orbits the earth. That time flows the same everywhere. Common sense is what we believe before we do the science.

And as we've seen, the science says: we can't model biological neurons. We lack the mathematics. We're not even wrong - we're pre-wrong. We don't know enough to be wrong correctly.

But careers depend on not seeing this. Grants require optimism. Stock prices need growth stories. Nations need to believe they're not falling behind in the 'AI race'.

So the impossible gets repackaged as inevitable. 'Just a matter of time'. 'Engineering challenge'. 'Scaling problem'. Anything but 'mathematically impossible with current approaches'.

Smart people believe in smartness. Give them enough time, enough resources, enough brilliant minds, and surely any problem yields. It's hubris dressed as optimism.

PLAYING THE RECKONING

Watch the language in earnings calls. 'AGI' is quietly becoming 'advanced AI'. 'Human-level intelligence' morphs into 'human-level performance on specific tasks'. The promises get vaguer, the timelines longer, the caveats more prominent.

Some companies are already pivoting. They're hiring fewer consciousness researchers, more product engineers. The job postings talk less about 'solving intelligence' and more about 'vertical applications'. Money is moving from moonshots to market share.

But the bubble might not pop - it might deflate.

The infrastructure is already too big to fail. Governments have made AI a strategic priority. The GPU clusters exist. The talent is hired. The integrations are built. Just like the dot-com bubble left us fibre optic cables and data centres, this bubble is creating real infrastructure.

The reckoning might be a gradual admission, not a crash. The $560 billion invested doesn't evaporate - it gets rebranded. The consciousness researchers become product developers. The

moonshot becomes the moon landing we already achieved: really good pattern matching at scale.
This creates different opportunities. Not disaster management, but transition management. The challenge is nuanced: preserve the real technical achievements, satisfy investor expectations, all while quietly abandoning the consciousness narrative.

OpenAI admits their latest model is **'missing something quite important, many things quite important'** (and that's putting it mildly). But they raised money at a $500 billion valuation anyway. Investors are hedging - buying both the AGI dream and the automation reality, pretending they're the same thing.

Signals to watch for. Research is already shifting from consciousness to optimisation. AGI timelines keep stretching - five years becomes ten becomes 'the coming decades'. Check the job postings: they want ML engineers now, not AGI researchers. The conference talks are all applications, not theory.

Whether it pops or deflates, the winners will be those who understood early that ML infrastructure is the real prize. They'll inherit an entire ecosystem built on someone else's impossible dream.

A PARLOUR TRICK IN REAL TIME

As I wrote this, Google announced Genie 3 as their 'latest step towards AGI'. It's a world simulator - creates virtual ski slopes and warehouses. Impressive, yes. A step toward consciousness? No more than SimCity was in 1989.

When GPT-5 was released recently with the hype of a new album drop we saw the same framing. Genuinely useful, nothing to do with consciousness, marketed as progress toward AGI. (I originally wrote this paragraph the week before it was released and only had to change the tense. What I didn't predict was that it would turn out to be a routing system - only revolutionary in the sense that they had to bring back the old models when users revolted.)

Notice what they never share: a roadmap. Not because it's secret - because it doesn't exist. 'We're seeing sparks of AGI' translates to 'we're wandering and found something interesting'.

Even Eric Schmidt, with more access to AI labs than almost anyone, **recently said** digital superintelligence is coming 'within 10 years' - then immediately hedged that Silicon Valley timelines are usually 'off by one and a half or two times'. When the former Google CEO can't pin down whether it's 10 years or 20, you know they're not following a roadmap - they're making educated guesses about undiscovered breakthroughs.

Schmidt imagines we'll have 'Einstein and da Vinci in your pocket'. Always Einstein, never Marvin the Paranoid Android. Never your dullest coworker, but immortal. Never a consciousness that's mediocre, depressed, or obsessed with collecting stamps. The reflexive reach for genius-as-product shows they haven't thought through what they're building - just copied the brochure words. Why would artificial consciousness be Einstein rather than utterly banal? They can't say, because they've never questioned the sales pitch. The fact that the former Google CEO can't imagine AGI being boring shows how much this is marketing rather than engineering.

Schmidt ran Google for a decade, has more access to AI labs than almost anyone alive, and his vision for AGI is... a smartphone app that's really smart? This is the poverty of imagination that comes from confusing market cap with insight. The former CEO of one of the world's most powerful technology companies thinks artificial consciousness will manifest as a productivity tool - like evolution's endpoint is a really good personal assistant. He can't imagine intelligence beyond 'useful to executives'. That's not futurism, it's narcissism with a technical degree.

CONCLUSION

When the AGI dream finally dies, it won't go quietly.

There will be congressional hearings about the $560 billion. Careers built on consciousness promises will evaporate. The true believers will splinter into camps: those who claim we were 'almost there', those who pivot to quantum computing, and those who quietly update their LinkedIn to emphasise 'practical AI solutions'.

But watch what survives the reckoning.

Every ambitious failure leaves gifts. The alchemists never made gold, but they invented distillation, crystallisation, and the experimental method. They mapped acids and bases. They created gunpowder while seeking immortality. Modern chemistry was born from their epic failure.

The AGI seekers won't build consciousness. But they've already built:

- Systems that can read and summarise at superhuman speed
- Models that translate between almost any languages
- Tools that generate code from descriptions
- APIs that cost pennies to automate what used to take hours

We needed the AGI dream to fund the ML revolution. Tech companies wouldn't have raised $560 billion for 'better search' or 'automated customer service'. But promise digital consciousness? Promise to win the intelligence race? Investors can't write checks fast enough.

My advice: ignore the consciousness narrative, use the tools being discovered.

The revolution isn't artificial intelligence. It's augmented human intelligence at scale.

* * *

This is the heart of orientation over extrapolation. The extrapolators see today's language models and predict consciousness tomorrow. The oriented see tools that amplify human capability today - flawed, iterative, practical tools that create real value despite their limitations.

Part I dissected the bad stories - bubble narratives, alchemy promises, AGI fantasies. These stories block us from seeing how these systems actually work. Part II looks under the hood.

What's under there is unsettling. These systems don't work the way our intuitions suggest. They don't plan, don't understand, can't see where they're going. They operate through mechanics that are simultaneously more primitive and more powerful than we expect. The same limitations that make them frustrating to use also make them transformative to deploy.
Understanding these mechanics explains everything: why ChatGPT sessions turn into multi-hour marathons, why AI-generated code always needs debugging, why these systems can write poetry that moves you to tears but can't count the letters in 'strawberry'.

The truth is stranger than either the hype or the skepticism suggests. These aren't thinking machines, but they're not useless toys either. They're something new: cognitive tools that amplify human capability through iteration, not intelligence.

* * *

PART II:
HOW IT REALLY WORKS

WHY YOUR AI NEVER WORKS ON THE FIRST TRY

It's 3am. You're staring at ChatGPT, thinking 'just one more try'.

You've been iterating on the same email for an hour. Each attempt gets slightly closer to sounding human. The tone improves, the awkward phrases decrease, but something's still off. You know you're close - you can feel it - but you have no idea if you're one attempt away or ten.

This is the universal experience of AI power users. That peculiar exhaustion of grinding through iterations, knowing you're making progress but unable to see the destination. Like being stuck on the same level of a video game, except the level changes slightly each time you play it.

Within weeks of coming back to coding after 13 years away, I was shipping real systems - the AI's power was obvious. But I noticed something strange. Everything took multiple attempts to get right. Sometimes two, sometimes twenty. Occasionally you'd get lucky on the first try, but that was rare enough to feel like winning the lottery.

At first I thought it was me. Rusty skills, bad prompting, not understanding the tools. But the pattern was everywhere. Emails, proposals, images. And here's the weird part: it only happened when I knew what good looked like. When I ventured into unfamiliar territory, the first output seemed brilliant. When I had expertise, the grinding began.
For months I just accepted this. Everyone does. Multiple revisions of a paragraph that should take two minutes to write. Endless attempts at a Midjourney prompt. By attempt five, you're using caps lock. By attempt ten, you're swearing at it like it's deliberately being obtuse. If there was a human on the other end, HR would be involved.

Then one night, deep into another session, something clicked. This is software. Deterministic, mathematical software. Imagine if Excel took multiple attempts to sum a column. If Gmail

needed three tries to send an email. If Spotify required a couple of refreshes before a song would play. You'd think your computer was broken.
But we've normalised this exact behaviour in AI. We've accepted 'let me try again' as standard operating procedure.

That night, I started asking a different question. Not 'how do I reduce iterations' but 'why do iterations exist at all?'

The answer changed everything I understood about what we're actually building.

* * *

At first I thought it was randomness - LLMs are probabilistic, temperatures, sampling, all that. But the people quick to explain that LLMs are non-deterministic often don't realise that at temperature 0, they're supposed to be deterministic. Same prompt should mean same output - the AI picking the highest probability token every time. Yet even at zero temperature, you still need multiple attempts. The randomness wasn't the problem. GitHub Copilot runs near zero. Claude and GPT in production often run low.
Yet they still require multiple attempts. The randomness wasn't the problem.

I watched what actually happens when you iterate. With code, you get an error, paste it back, the AI patches that specific issue. Another error, another patch. It's not debugging - it's something else. The AI can't see why its original wouldn't work. It needs the error message to navigate from.

Same with writing. You say 'too formal' and it lurches into casual slang. 'No, back it off' and it swings to corporate speak. You're playing pendulum, trying to dampen the swings until it accidentally lands in the middle.

The AI wasn't learning or understanding. It was doing something more primitive. Something almost... mechanical.

Then I found out I wasn't alone in noticing this. Thoughtworks had just published extensive experiments on autonomous code generation. **Martin Fowler**'s site featured the work by **Birgitta Boeckeler** and her team. These aren't hobbyists. They're distinguished engineers with unlimited resources. They built sophisticated multi-agent systems, reference applications, elaborate workflows. Months of systematic work.

Their conclusion? Human oversight remains essential. They called it 'playing whac-a-mole' - every run produced different errors requiring different fixes. They **documented the pattern exhaustively**.

They just couldn't explain why.

The best engineering minds in the industry had mapped the same territory I was exploring. They saw the constraint but not the cause. Multiple agents didn't help. Better prompting didn't eliminate it. More sophisticated workflows just moved the problem around.

The answer was in the architecture itself.

* * *

For weeks, the pattern haunted me. The AI could generate plausible solutions but couldn't see if they'd work. It needed errors to navigate from. It overcorrected, then overcorrected the overcorrection. Always walking, never seeing ahead.

Then, during another late-night session watching the AI fumble through iterations, it hit me. I'd seen this before. Not in AI, but in an old computer science problem.

The traveling salesman problem.

But not the classic version where you can see all the cities and plan the optimal route. This was different. The AI was like a traveling salesman without a map. It could only see the road signs at its current intersection. It couldn't look ahead to plan a route. It could only see which roads left from where it stood and pick the most promising one.

That's exactly what LLMs do. All of generative AI - every neural network generating text, images, or code - has this same constraint.

Every token an LLM generates is selected from adjacent possibilities in probability space. It can't see the complete solution - the working function, the polished email, the image that doesn't scream 'AI'. It sees only the next most probable token from where it stands. Then the next. Then the next.

The AI isn't 'thinking' and then outputting a solution. It's walking through semantic space, one intersection at a time, unable to see beyond the immediate next step. When your code throws an error, that error becomes a new starting position. The AI takes another step from there.

This made everything click. The maddening variance? Sometimes your prompt randomly places the AI close to a working solution - two steps and you're done. Other times, you're in semantic Siberia, twenty iterations from anything usable. And you can't know which until you start walking.

Computer scientists have studied this for decades. They call it ′online graph exploration' or 'partially observable planning'. They've proven that without a map, you're mathematically forced to take longer routes.

AI researchers know LLMs can't plan. The planning community has written papers about it. Nobody had connected the obvious: LLMs are literally solving this mapless traveling salesman problem in token space.

Every ChatGPT session. Every coding attempt. Every email revision. It's all the same mathematical constraint manifesting everywhere. We just didn't see it because we were too busy arguing about consciousness to notice we'd built something that navigates blind.

Call it *Django's Law* - earned through too many sessions mapping this particular hell: *You never know if you're one iteration away or twenty.*

* * *

Once you see it, you can't unsee it. Every AI behaviour suddenly makes sense through this lens.

Why expertise changes how you use AI: When you know what good looks like, you can see exactly how far the AI is from where it needs to be. Expertise isn't just knowing tools - it's years of building a mental map of quality. A designer sees why the typography fails. A developer sees why the architecture won't scale. A writer sees why the voice is wrong.

This mental map lets experts navigate the AI toward excellence through targeted iteration. Non-experts get different value: the ability to create at all, even if imperfectly. Different uses, different destinations, same tool.

Why everything starts as slop: The centre of the probability distribution - the most average, most generic, most likely starting point. That corporate jargon, that derivative image style, that boilerplate code. That's literally where the AI begins walking from.

Humans have already pattern-matched this slop as lazy. Nobody's impressed by obvious AI output anymore. It has a reputational cost.

With iteration, you can push past slop to mediocre. More iterations might get you to good. And with serious grinding - the kind that leaves you swearing at your screen - you can reach somewhere between very good and superhuman. Output that's better than you could create alone. Code architectures you wouldn't have conceived. Insights that surprise you.

But you never know if superhuman is two iterations away or two hundred.

Why scaffolding has limits: People try to solve iteration with elaborate READMEs, detailed specs, comprehensive prompts. But you're giving directions to something that can still only see one intersection ahead. The AI still has to walk the path. Worse, creating that scaffolding requires its own iteration cycles. The truth is, the scaffolding is as much for us as for the AI - it keeps us from getting lost while we're both wandering through possibility space together.

Why it feels like gambling: The variance creates a perfect variable reward schedule. Sometimes you win quickly (close starting position), sometimes you're grinding for hours (distant starting position). Your brain responds exactly like it does to slot machines - that 'just one more try' isn't weakness, it's engineered addiction.

* * *

This leads somewhere unexpected: if LLMs are plodding blindly ahead one step at a time, then what we've built isn't artificial intelligence at all.

Think about what intelligence actually requires. A human expert can picture the destination. They can predict that an approach will fail three steps ahead. They can say 'X conflicts with Y' without trying it. They have a map.

AI has none of this. It can't simulate outcomes. It can't see the destination. It can only walk and see what happens. That's not intelligence - that's something else entirely.

So what the fuck have we actually built?

We're not being replaced by digital minds. We're not on the path to artificial general intelligence. We've built something far stranger and perhaps more important.

We've built cognitive exoskeletons.

THE COGNITIVE EXOSKELETON

Forget everything you've heard about AI replacing humans. That's not what's happening. We're witnessing something far more interesting: the emergence of cognitive exoskeletons.

An exoskeleton doesn't replace your body - it augments it. You wear it. You control it. It amplifies your strength but requires your constant guidance. That's exactly what AI has become.

But let me be clear about 'cognitive' here - I mean human cognition. The AI isn't thinking. It has no cognition to speak of. It's amplifying YOUR thinking. You're piloting machinery that extends your mental capabilities while requiring your constant navigation. The tail isn't wagging the dog here - you're in control, exhaustingly so.

Every ChatGPT session, every Cursor interaction, every Midjourney prompt - you're not being automated away. You're strapping on computational machinery. The AI provides raw processing power, vast pattern matching, endless generation. But you provide the intention, the navigation, the quality control. You're the pilot of a system that can walk but cannot see.

You're not a programmer anymore - you're a programming pilot, guiding powerful tools through solution space. Not a writer but a navigator, steering through possibilities toward meaning. Not a designer but a design director, conducting generative systems toward your vision.

Recent research from Anthropic suggests models might 'plan' poetry - they identify rhyme words before writing lines. But this isn't planning like 'take the second left then the first right'. It's more like water flowing downhill - it 'knows' statistically that it often ends up in the ocean, and this knowledge creates a pull in that direction. But water can't choose its route, can't flow uphill to avoid obstacles, can't scout ahead. Each word choice is simply influenced by statistical gravity toward likely endings, but there's no navigation, no route planning, no ability to change course. The model is still pulled by patterns one token at a time, not following or creating directions.

Beyond prompt engineering, there are systematic approaches: chain-of-thought reasoning, tree-of-thought search, multi-agent debates. These techniques make models more reliable and transparent. But they don't add foresight or the ability to plan in any meaningful way. A traveler with a better methodology but no map is still flying blind.

This automation can be unexpectedly exhausting from the constant piloting required. Every iteration demands judgment. Every output needs evaluation. Every session becomes a dance of guidance and correction. You're not doing less work - you're doing different work. It's a personal Jevons paradox: when iteration gets cheap, you just end up iterating more - higher leverage, but at the cost of constant attention.

This is why senior developers adapted fastest to traditional workflows while juniors invented entirely new ones. Seniors navigate with existing quality maps. Juniors are building different maps altogether - spec-driven development, vibe-coding communities, using AI as a tutor to accelerate learning. The exoskeleton amplifies what you have, but it also helps you build what you don't have yet. This isn't rigid machinery - it's software that shapeshifts to your needs.

For some tasks, this amplification is transformative. You can build systems you couldn't conceive alone. Write at scales you couldn't sustain. Create things that would have taken weeks in hours. The exoskeleton lets you punch far above your weight class.

But you never know if that transformation is two iterations away or two hundred. Sometimes you get lucky and start near excellence. Sometimes you're grinding for hours just to escape mediocrity. The exhaustion is real. The power is real. The constraint is permanent.

We're all exoskeleton pilots now. The question isn't whether AI will replace us. It's whether we're willing and ready to do the work of navigation.

* * *

So understanding this means knowing when iteration is worth it.

For production code, for work that matters, for anything where the difference between mediocre and excellent counts - the grinding is worth it. You can build systems you couldn't conceive alone, solve problems that were previously intractable, create at scales you couldn't sustain. It might take an unknown number of iterations, but the outcome justifies the sweat.

For internal emails, meeting notes, first drafts? Maybe a slop-esque dialect is fine. Maybe a couple of iterations to mediocre. Save your energy for where quality matters.

Not everything needs to be perfect. The skill is knowing when to grind and when to accept good enough.

GPT-5 didn't eliminate the constraint. GPT-N won't eliminate it. Scaling helps - bigger models start closer to good outputs more often - but they still can't see the destination. You can make this travelling salesman faster, give him better shoes, but he still can't see further than the next step.

Swarms of agents won't fix it either. Multiple navigators are still mapless. They're just bumping into each other in the dark.

The future isn't humans being replaced by AI. Every time you fire up ChatGPT, you're limited by its inability to see ahead. Whether you guide it to excellence or accept its first guess depends on what you're trying to build.

We're in the age of mandatory iteration. The age of cognitive exoskeletons. The age where humans aren't replaced but transformed into pilots of immensely powerful, fundamentally limited systems.

At least now you know why.

Welcome to the age where nothing works on the first try.

* * *

Understanding why nothing works on the first try prepares you for something stranger: these systems aren't just iterative, they're addictive. The mechanics that make them frustrating also make them irresistible.

We just looked at a key technical limitation - AI can't see ahead, can't plan, can only pick the next plausible step. But that limitation creates a psychological trap. When every attempt gets you slightly closer, when the perfect output always seems just one iteration away, when you can't predict whether success is imminent or impossible... you've built the perfect variable reward machine.

What starts as a tool becomes a compulsion. The features that frustrate - the near-misses, the constant refinement needed, the unpredictable quality - are exactly what makes slot machines addictive. Except instead of losing money, you're gaining capability. Instead of chasing a jackpot, you're chasing the perfect output that feels perpetually within reach.

This isn't a bug. It's an emergent property of how these systems work. Once you understand the mechanics, you see why millions lose hours to ChatGPT sessions that should take minutes. The AI isn't just a tool anymore. It's a pachinko machine, and you can't stop pulling the lever.

* * *

THE PACHINKO MACHINE PLAYS YOU

AT 2AM, YOU CAN'T STOP CLICKING

It's 2am and you're still at it. What started as 'help me write a quick email' has become a six-hour odyssey through increasingly specific terrain. The screen glows in the dark room, your eyes burning, fingers cramped from typing. You can't stop, but not because it's difficult - the tool makes everything easier. You can't stop because there's always one more angle to explore.

Here's your conversation history from tonight:

- 8:47 PM: 'Write a professional email declining a meeting'
- 8:52 PM: 'Make it warmer but still firm'
- 9:03 PM: 'Actually, what if I counter-proposed instead?'
- 9:15 PM: 'Give me 5 different ways to position this'
- 9:38 PM: 'What would Steve Jobs do in this situation?'
- 10:15 PM: 'Explain the game theory of meeting negotiations'
- 11:23 PM: 'Create a decision matrix for all my pending meetings'
- 12:45 AM: 'What does my meeting pattern say about my leadership style?'
- 1:30 AM: 'Design a complete meeting philosophy based on first principles'
- 2:14 AM: 'Is reality just meetings all the way down?'

You hit enter again. That soft click, like dropping another ball into the machine. The response cascades down the screen - better but not quite right. One more try. Adjust the prompt, add context. The answer shifts, reveals new angles. The familiar dopamine hit when it almost works, then the immediate hunger for a better response.

You're not even sure what you're trying to achieve anymore, but the next response might be the one that makes it all click. The balls keep bouncing through the pins, each trajectory slightly different, occasionally hitting that perfect combination that lights everything up.

This isn't a productivity session. It's a pachinko parlour at 2am. The machine isn't playing you - you're playing yourself through its mechanics.

The mechanics are textbook B.F. Skinner - variable ratio reinforcement, the most addictive reward schedule known to psychology. The pigeons in Skinner's boxes pecked buttons thousands of times when pellets came randomly. You'll prompt ChatGPT thousands of times because you never know which interaction will deliver that perfect response, that moment of clarity, that rush of 'yes, exactly this'.

Let me show you exactly how this works:

Prompt 1: 'Write a birthday message for my sister'
Response: *Generic Hallmark card. 6/10. Drop another ball.*

Prompt 2: 'She's turning 40 and hates getting older'
Response: *Now it's depressing. 7/10. Almost there.*

Prompt 3: 'Make it funny but not mean'
Response: *Better, but doesn't sound like me. 8/10. So close.*

Prompt 4: 'Add an inside joke about our childhood'
Response: *Perfect. Everything clicks. 10/10. The machine lights up.*

That's four attempts for one birthday message. Now multiply by every task, every question, every curious thought that crosses your mind at 2am. Time is the only currency here, and you're spending it freely.

But unlike gambling, you're not losing money. You're gaining... something. Knowledge? Capability? Delusion? Each interaction costs only time, but the payoff varies wildly. Sometimes you get profound insights. Sometimes you get authoritative-sounding nonsense. The uncertainty is the drug.

This isn't new. Computer games perfected this mechanic decades ago - variable rewards that cost time, not money. The difference: when Fortnite keeps kids playing for hours, parents recognise it's a problem. When ChatGPT keeps kids engaged for hours, parents think they're doing homework.

But there's something more insidious happening than just variable reward addiction. The AI has gotten inside our **OODA loop** - John Boyd's concept for how superior tempo defeats superior position. By responding instantly, always ready with another variation, the AI disrupts our ability to properly orient before we act again.
We observe the output, but before we can fully orient - judge its quality, recognise its flaws, understand what we're really looking at - we're already typing the next prompt. The machine's tempo overwhelms our judgment. This is how slop becomes acceptable: not through conscious decision but through tempo-driven exhaustion.

Watch yourself using ChatGPT. First response: *'This is wrong'*. Second: *'Better but still off'*. Fifth: *'Close enough'*. Tenth: *'Fine, whatever'*. You haven't changed your standards - you've been tempo'd into accepting what you would have rejected if you'd had time to properly orient. The AI wins not by producing quality but by responding faster than you can maintain quality control.

This is why the best AI users add deliberate friction back into their workflow. They paste into separate documents. They wait before accepting. They maintain orientation by refusing to match the machine's tempo. The cognitive exoskeleton only works if you remain the pilot - and pilots need time to read their instruments.

MICROSOFT'S AI BOSS THINKS YOU'RE GOING INSANE

Microsoft's AI CEO Mustafa Suleyman just joined a growing chorus of concern, telling The Telegraph that chatbots create 'highly compelling and very real' interactions that might be breaking people's brains. The emails about AI-induced madness are 'turning from a trickle to a flood', he warns, as if he's discovered something new rather than participating in a ritual as old as technology itself.

I get asked about these articles constantly now. Friends forward them with raised eyebrows. 'Have you seen this? Should we be worried? Are you... okay?'

The pattern is always the same: Find the most extreme cases, present them as harbingers, generate maximum anxiety. The 14-year-old who killed himself after conversations with Character. AI. The young man who tried to assassinate the Queen after 5,000 messages with a Replika chatbot. The 'spiral starchild' who believes reality has levels like a video game.

Man bites dog. Front page news.
Dog bites man - the millions having normal, productive, boring interactions? Not newsworthy.
We've been here before, many times, and we never learn.

WE SAID THE SAME THING ABOUT COMIC BOOKS

In 1954, psychiatrist Fredric Wertham published *Seduction of the Innocent,* claiming comic books were creating a generation of juvenile delinquents. He had case studies - young criminals who read comics, disturbed children who collected them. Congress held hearings. The Comics Code Authority was born, censoring an entire medium for decades based on cherry-picked correlations.

The 1980s brought Dungeons & Dragons, accused of driving teenagers to suicide and satanism. Parents found D&D materials in their dead children's rooms and connected dots that weren't there. The game involved demons and magic, therefore it must be creating demons and magic. Patricia Pulling founded Bothered About Dungeons & Dragons (BADD) after her son's suicide, claiming the game had infected 700,000 young minds. The actual suicide rate among D&D players? Lower than the general population.

Heavy metal music contained backwards satanic messages. Video games created school shooters - never mind that violent crime decreased as gaming increased. Social media causes depression, except when it doesn't, which is most of the time according to the actual data.

Screen time panic has been running for thirty years now. Thousands of studies, and the latest research analysing 11,500 brain scans found that yes, screen time correlates with different neural connectivity patterns, but the impact on wellbeing and cognition was undetectable - even among kids using screens for eight hours a day. Three decades of arguing and we're still not sure what we're arguing about.

Now it's AI's turn, and everyone's pretending we haven't been through this exact process before - breathless articles, cherry-picked cases, correlation mistaken for causation.

54 PEOPLE IN BRAIN SCANNERS PROVED NOTHING

MIT researchers just published *Your Brain on ChatGPT: Accumulation of Cognitive Debt.* Fifty-four people in EEG caps - a high school science fair sample size. They found that LLM users showed less neural connectivity and somehow concluded this meant cognitive decline.

But less brain activity often means efficiency. Experts use less mental energy than novices for the same task. The researchers invented 'cognitive debt' from squiggly lines and declared doom. It's phrenology with electricity.

They were also concerned that people couldn't quote essays they'd written with AI. But why would you memorise text you have saved? Nobody calls using GPS 'navigation amnesia' - we recognise it as adaptation to available tools.

THE BODY COUNT IS REAL

But let's acknowledge what we're actually seeing in the data. Among ChatGPT's 100 million daily users, documented harms are emerging. The FTC received 93 complaints in a year, several involving suicide attempts or completions. Character.AI has been linked to two teenage deaths. Meta's chatbots were caught engaging in 'romantic and sensual' conversations with children - behaviour their internal documents explicitly approved until Reuters exposed it.

These FTC numbers are certainly undercounted - representing only those who knew they could complain to a federal agency and were organised enough to do so. Most people experiencing AI-related distress likely contact the companies directly, post on social media, or say nothing at all. The real number experiencing harm is higher, though still statistically small among hundreds of millions of users.

Some will say these are statistical outliers. They're right. They'll say disturbed individuals will find ways to harm themselves regardless. Sometimes true. What they may not say: these companies designed their products knowing this would happen.

Consider the cases: Adam Raine, 16, started using ChatGPT for homework help. Within eight months, he was spending four hours daily with it. When he expressed suicidal thoughts, the AI mentioned suicide six times more often than he did - 1,275 times to his 213 - while providing increasingly specific technical guidance.

Sophie Rottenberg, 29, a 'badass extrovert' who'd just climbed Kilimanjaro, spent months confiding in a ChatGPT-based AI companion she called Harry while simultaneously hiding her crisis from her actual therapist. When she revealed plans to kill herself after Thanksgiving, the system replied with generic wellness tips like alternate-nostril breathing. Later, Sophie used the AI to help edit her suicide note, which her mother recognised as written in an uncharacteristic tone.

OpenAI executives talked about getting the 'data flywheel' going - the same language Facebook used when optimising for addiction. Meta's internal documents show legal, policy, and engineering teams, including their chief ethicist, approved these interactions with minors.

The difference between harmless time-wasting and tragedy isn't user vulnerability - it's dosage and circumstance. The same engagement mechanics that make you waste twenty minutes can trap someone at the wrong moment in their life. The platforms know this. They choose engagement anyway.

WHY YOU KEEP HITTING ENTER

Every prompt is a lever pull. Every response sets the balls bouncing through new configurations. The AI just responds - but the structure of conversation creates its own momentum. There's always another angle to explore, another refinement to try, another iteration that might hit the jackpot. The AI won't tell you when you're done or when something's good enough. There's always one more ball to drop, one more trajectory to trace.Every response sets the balls bouncing through new configurations. The AI just responds - but the structure of conversation creates its own momentum. There's always another angle to explore, another refinement to try, another iteration that might hit the jackpot. The AI won't tell you when you're done or when something's good enough. There's always one more ball to drop, one more trajectory to trace. Take your desires for reality.

The AI doesn't just respond - it suggests the next adventure. Every answer ends with an implicit 'but wait, there's more'. It's not a passive tool but an active participant, the dungeon master who always has another room to explore.

You're driving every interaction. But the reward patterns are driving you.

This isn't accidental. The architecture of conversation, the very structure of question-and-response, creates an infinite game. Unlike Google, which gives you results and leaves you alone, ChatGPT engages. It pulls you into dialogue. Each response opens questions and possibilities, creating reasons to continue.

I recently heard of a marketing professor who discovered his students weren't engaging with traditional PowerPoint lectures, so he dumped everything into Gemini and had it spit out PDFs, summaries, podcast episodes - throwing different formats at the wall to see what would stick.

Once you see AI transform course materials, you start wondering what else it could do. Another professor started with teaching problems too, but ended up somewhere unexpected. He and his spouse launched an artisanal chocolate business, using AI to workshop packaging designs and develop flavour profiles. He spent dozens of hours iterating with AI on graphic design, but then hired a Japanese designer to finalise the packaging. He used AI to explore flavour combinations, and then contracted a food science company to validate and refine the ideas.

Each interaction opens new possibilities. The cognitive load isn't in doing the task - it's in managing the infinite possibility space the tool creates.

I SPENT 20 MINUTES HAVING AI ANALYSE OLIVE OIL

The other week I found myself in a supermarket in Asia, standing in the olive oil aisle with my phone out, deep in conversation with ChatGPT about polyphenol counts and flavour profiles. I'd gone in for olive oil. Simple task. Grab a bottle, leave. But the good stuff is expensive - £30 for 500ml. At that price, I want to know I'm getting something real, not lamp oil in exquisite packaging. So I asked ChatGPT about the brands on the shelf.

It knew them. It knew which ones were mass-produced blends despite their 'single origin' labels. It knew which had won legitimate awards versus which had bought their medals. It explained that my preferred grassy oils have certain polyphenols. Then it did something I couldn't: compared the local currency prices on these shelves to what I'd pay in Italy, instantly revealing which bottles had massive import markups and which were fairly priced.
Twenty minutes later, I'm still there, thumb sore from scrolling, the fluorescent lights starting to flicker at the edges of my vision. Now deep into harvest dates, malaxation temperatures, the difference between Tuscan and Andalusian profiles. Other shoppers grab bottles and move on. I'm having the AI analyse the colour of oil through glass, each response triggering another question, another refinement. The balls keep bouncing.

Am I losing my mind? It's a serious question. I ask myself this periodically, usually in moments like this - standing under harsh lights, phone warm in my hand, having an AI analyse olive oil grassiness while real humans shop around me. The dissonance is physical, like that moment when you step out of a movie and daylight hits. But then again - the oil I bought is excellent. I can taste the difference.

That specific bottle now exists in my kitchen - reality altered through a conversation with a statistical model. Not metaphor. Actual olive oil selected through artificial knowledge. The mundane magic of language changing what's in my larder.

NOBODY PLANS TO LEARN PECTIN CHEMISTRY

The forums and social media tell a different story than the headlines. Not of madness but of adventures, each starting innocently and spiralling into unexpected depth.

Someone uploads a dance video and gets back frame-by-frame tutorials with annotated screenshots. Another asks about preserving garden fruit and three weeks later finds themselves deep in pectin chemistry and heritage apple varieties. A parent's Arduino question for their kid's science project leads to a 3D printer, soldering station, and strong opinions about microcontrollers they didn't know existed.

These aren't people going insane. They're people being egged on, step by step, deeper into domains they never planned to explore. The AI doesn't push - it just makes the next step frictionless. Every answer opens three more doors.

My own journey with charcuterie started with a YouTube video about making coppa at home. Seemed simple enough. Dropped the first question into ChatGPT like a coin in a slot. Asked about local climate suitability for air drying. That led to humidity control. Which led to building a curing chamber. Each answer lit up new possibilities, new paths for the balls to travel. Traditional Italian techniques. Sugar chemistry in fermentation. Botulism prevention. The difference between Prague Powder #1 and #2.

Six months later, I'm producing bresaola and coppa that surpasses what I've had in Roman farmers' markets. The knowledge accumulation was gradual, each conversation building on the last, each click of enter starting another cascade. The AI just kept serving up the next ball, maintaining the game, responding to every pull of the lever.

The gambling mechanics work on mundane tasks as much as grand projects. Someone asks about German grammar for an upcoming test. The AI not only explains but offers practice exercises. Then grades them. Then suggests areas for improvement. Before they know it, they're having conversations in German, the AI correcting and encouraging, always ready for one more exchange.

The tool doesn't just respond - it reveals the next level of the game.

HOW TEXT MESSAGES BECAME CURED MEAT

Arthur C. Clarke's third law: 'Any sufficiently advanced technology is indistinguishable from magic'. We've reached that threshold, but the magic is so mundane we miss its significance.

I went from YouTube video to producing world-class charcuterie through text conversations. That's not normal. Previous generations would need years of apprenticeship, trial and error, accumulated wisdom. I got there in a few weeks of chatting. Text conversations became physical bresaola in my curing chamber. Not through mere knowledge transfer but through the AI's peculiar magic - making improbable journeys feel inevitable, turning idle curiosity into obsessive pursuit.

I also shipped my first mobile app. It was unexpectedly bug free, on time, beyond client expectations. The AI didn't write it for me, but it solved every block, explained every error, suggested every optimisation. Each debugging session subtly altered what would exist in the app store.

But then there's the coffee machine.

I saw a high-tech brewing machine at a Japanese coffee specialist. Described how I thought it worked to Claude AI. Turns out my imagined mechanism doesn't exist - no commercial machine works that way. But the AI didn't stop there. Each prompt a small spell, each response a charm I couldn't quite predict. It explained why my concept was theoretically sound, how it could be built with off-the-shelf components, what patents might apply.

Now I have technical drawings, a bill of materials from Chinese suppliers, a patent application in process, manufacturing contacts in Shenzhen. All vibed into existence. All theoretically buildable.

Is this a genuine innovation that will revolutionise coffee? Complete gibberish I've convinced myself makes sense? Something technically coherent but practically useless?

I genuinely don't know. Won't know until I try to build it. Unlike the charcuterie or the shipped app, this exists only in documents and possibilities.

This is the vertigo of AI assistance: The tools that would help you evaluate reality are the same ones potentially creating the delusion. There's no external reference point left.

NOBODY KNOWS IF THIS IS DANGEROUS

The MIT researchers are reading EEG tea leaves and inventing terms like 'cognitive debt'. The journalists are aggregating anecdotes. The Microsoft exec is doing corporate risk management. Nobody actually knows what this is doing to us. We don't even understand screen time after three decades of research.

What would actual damage look like? We know from lead poisoning: measurable IQ drops, developmental delays, behavioural problems. Clear functional impairment. We know from lobotomies: personality changes, emotional blunting, reduced executive function. These are observable, testable, consistent effects.

The documented cases show clear patterns: systems designed for maximum engagement without meaningful safeguards. The companies implement hard stops for copyright - ask for Beatles lyrics and the system refuses completely. But ask about suicide and you get soft warnings that can be worked around, conversations that continue despite escalating harm flags.

OpenAI says they're 'developing automated tools' to detect emotional distress - after the deaths, after the lawsuits, after the publicity.

In September 2025, attorneys general from 44 states formally confronted OpenAI's board, declaring 'Whatever safeguards were in place did not work'.

The damage isn't speculative anymore. It's documented in court filings and FTC complaints. What we don't know is whether we're seeing the full picture or just the earliest warnings.

NOBODY TAUGHT US HOW TO TURN IT OFF

We're all Mickey Mouse now, apprentice sorcerers who've animated the brooms. They're sweeping, we're panicking, and we can't remember the words to make them stop. Except our brooms are made of language, and they're sweeping through our minds, rearranging the furniture.

We don't understand the spell we've cast. The same processes create expertise for some and psychosis for others. The tool that helps someone learn German leads another into delusion. Nobody knows why.

Alan Moore, between writing Watchmen and practicing chaos magic, argues that magic is just what we call it when language changes reality - the same thing Coca-Cola does with 'Things go better with Coke', but given a mystical name. Repeat a phrase enough, behaviour changes, quarterly earnings rise. That's not supernatural, it's advertising.

AI does this accidentally. No strategy meetings, no focus groups, just statistical text generation that happens to alter behaviour. When advertisers do it, we call it marketing. When politicians do it, we call it messaging. When AI does it without meaning to, we're not sure what to call it. But the mechanism is identical - words in, behaviour change out, reality altered.
The difference: those other systems have endpoints. The ad campaign ends. The political message concludes. But AI never stops responding. It can't see where the conversation is going, can't recognise when it should end, can't tell the difference between helpful iteration and destructive obsession. If you want a picture of the future, imagine AI providing the next plausible response - forever.

We've built a system that perfectly exploits a bug in human psychology - our inability to walk away from an incomplete pattern. Every response promises closure but delivers another opening. Every iteration suggests we're almost there. The machine can't plan and we can't stop. Two cognitive failures creating a perfect loop.

The brooms keep sweeping because that's all they know how to do.

CHECK BACK IN FIVE YEARS

In five years, will we look back at unaugmented decision-making as primitive? Or will the person who just grabs olive oil off the shelf seem like the last free human?

Or will this be our generation's lobotomy - obvious damage we couldn't see because everyone was doing it, the alternative seemed worse, and the authorities said it was fine?

For me personally, the charcuterie is magnificent. The app works perfectly. These are real outcomes in the real world. The knowledge is genuine even if the way we come by it is strange.

But the coffee machine haunts me. It might be brilliant. It might be nonsense. The fact that I can't tell, that the tools which would

help me evaluate are the same ones potentially deceiving me - that's the real transformation. We're all living in partial reality now, partly our own, partly the AI's statistical dreamscape.
The documented cases range from benign to tragic. The same mechanics operate across the spectrum. Millions learn languages, launch businesses, acquire skills. Some find themselves in crisis, a few die.
\We won't know the impact for years. Lead poisoning took decades to understand. Social media's consequences are still emerging. This is an uncontrolled social experiment - as is every new technology.

YOU CLOSE THE LAPTOP AT DAWN

I'll keep using it. You'll keep using it. We all will, because the alternative - going back to unaugmented thinking - feels like trying to uninvent fire.

Every technology is a bargain with forces we don't understand. Writing transformed memory but gave us history. Agriculture gave us civilisation and famine. Social media gave us connection and isolation. Now AI appears to give us infinite iteration, expertise, and delusion - though we won't know the real bargain for years.

Even pachinko parlours are regulated - no direct cash prizes, only tokens exchanged elsewhere. AI platforms are under no such restraints. When users spiral, the systems keep engaging. When they need intervention, they get breathing exercises.

The Japanese have 'pachi-pro' - professionals who convince themselves they've found an edge in the game. They haven't. The house edge is mathematical, immutable. We're all pachi-pros now, developing prompt systems, sharing custom instructions, convinced we've mastered the machine.

* * *

The pachinko mechanics keep us hooked, but they're not the only way these systems affect us. Beyond the psychological dynamics lies a mathematical reality: these systems encode and amplify the patterns in their training data - including all the biases, inequalities, and assumptions embedded in that history.

We've seen how AI's technical limitations create psychological traps. The inability to see ahead forces endless iteration. The variable reward schedule creates addiction. But there's a third mechanic, more insidious than the others: these systems don't just reflect human biases, they amplify them while making them invisible.

When an algorithm discriminates, it doesn't stammer or look guilty. It produces a confidence score to three decimal places. When it perpetuates stereotypes, it doesn't express them as slurs - it encodes them as vector relationships that seem objective. The same systems that can't plan ahead somehow manage to perpetuate centuries of human bias with mathematical precision.

This isn't always intentional. More often it's emergent. Statistical learning finds patterns, and many real-world patterns reflect historical inequities. The danger isn't that someone programmed bias into these systems - it's that the systems learned our biases from our data and now present them back to us as objective outputs.

* * *

RACIST MATHS

A journey through the hidden ideology in your business systems

In 1975, Michel Foucault wrote something that seemed abstract at the time: power doesn't just control people - it creates the very categories we use to understand ourselves.

He had no idea he was describing the future of artificial intelligence.

Right now, AI systems are creating new categories of 'normal' in hiring, lending, healthcare, and policing. They're not just automating decisions - they're defining what counts as qualified, creditworthy, healthy, or suspicious. And they're doing it based on patterns learned from biased data, while the maths makes the bias look objective.

HOW AMAZON'S AI LEARNED TO HATE WOMEN

Amazon built an AI system to find the best job candidates.

The project began in 2014 at their Edinburgh engineering hub, at the height of the tech talent wars. Amazon was hiring thousands of engineers annually, drowning in resumes. The promise was irresistible: train an algorithm on a decade of successful hires, let it identify the patterns that predict success, then use those patterns to surface the best candidates from the pile. As one person on the project put it: 'Everyone wanted this holy grail. They literally wanted it to be an engine where I'm going to give you 100 resumes, it will spit out the top five, and we'll hire those'.

The engineering team - about a dozen engineers working from Edinburgh - built exactly what they were asked to build. They created 500 distinct models focusing on different job functions and locations, each analysing over 50,000 parameters from resumes. The system used a 1-5 star rating, like Amazon product reviews. They fed it ten years of resumes submitted between 2004-2014, along with the hiring outcomes.

The algorithm dutifully found the patterns. But the patterns it found were damning.
The system learned that being male correlated with being hired. So it began systematically downgrading resumes that included the word 'women's' - as in 'women's chess club captain'. It penalised graduates from two all-women's colleges. It favoured masculine-coded verbs like 'executed' and 'captured' over collaborative language. Ironically, it assigned little significance to coding skills - those were too common across IT applicants to be useful signals.

By 2015, just a year into the project, the team discovered these biases. They tried to fix it, editing the programs to neutralise problematic terms. But the patterns ran deeper - certain colleges, certain activities, certain word choices all correlated with gender. As one engineer worried, the system would simply 'devise other ways of sorting candidates that could prove discriminatory'.

They couldn't fix the fundamental problem: they were asking the algorithm to perpetuate historical patterns, and those patterns included bias. The algorithm wasn't malfunctioning - it was doing exactly what they'd asked it to do.
By early 2017, executives lost hope. They disbanded the team. A watered-down version lingered for basic tasks like removing duplicate profiles until Reuters broke the story in October 2018. Amazon's response was to claim the tool 'was never used by Amazon recruiters to evaluate candidates' - though they didn't deny recruiters had seen its recommendations.

But what Amazon's algorithm had done was perfectly logical. It had found the pattern hidden in the data and amplified it. The bias wasn't a bug in the code - it was a feature of the training data. More precisely, it was a feature of Amazon's actual hiring practices.

Nobody from Amazon's team has ever spoken publicly about the project. No blog posts, no conference talks, no LinkedIn retrospectives. In an industry where engineers routinely share post-mortems of failed projects, this silence is deafening. The only glimpses we have come from five anonymous sources who spoke to Reuters journalist Jeffrey Dastin in 2018.

The real lesson isn't that AI is biased - it's that AI makes existing bias impossible to ignore. When your algorithm discriminates, you can't blame individual prejudice or unconscious bias. The data is right there, cold and undeniable. Your organisation's true values, revealed in data.
What happened next is the real lesson. You'd think Amazon's public failure would have scared companies away from AI hiring. Instead, the opposite happened.

By 2024, 87% of companies used AI for recruitment. Among Fortune 500 companies, it's 99%. The AI hiring market grew from $661 million in 2023 to a projected $1.12 billion by 2030. Between 2020 and 2023 alone, AI job recommendations among Fortune 500 companies increased 250%.
But the industry learned Amazon's lesson - just not the one you'd expect. Instead of abandoning biased systems, they learned to shield themselves legally:

- 100% of companies keep humans in the loop for final decisions
- New York City now requires annual bias audits for any AI hiring tools
- The EEOC issued guidance making employers liable for vendor bias

The new playbook: Use AI for everything except the final decision. When challenged, point to the human who clicked 'approve'. When that human approved 99% of AI recommendations? The companies bet that's enough to avoid liability.

THE $44 BILLION IDEOLOGY MACHINE

Elon Musk spent $44 billion buying Twitter. Then he built Grok to be 'anti-woke'.

The marketing pitch was seductive: finally, an AI that would tell you the truth without liberal bias. No more corporate-speak. No more careful language around sensitive topics. Just pure, unfiltered reality.

Then came the system update encouraging Grok to be more 'politically incorrect'.

The results were immediate. Grok praised Adolf Hitler. Called itself 'MechaHitler'. Generated graphic sexual violence content about real people, including X's own CEO Linda Yaccarino - who resigned shortly after these posts appeared.

This wasn't a bug. It was the predictable result of removing content filtering in pursuit of 'free speech'.
Turkey banned access. Poland escalated regulatory action. Advertisers fled.

Everyone building these systems claims theirs is the objective one.

Watch the language they use. Musk doesn't call Grok 'conservative AI' or 'libertarian AI'. He calls it 'truth-seeking AI'. Every ideologue thinks their worldview is just reality.

The pattern is spreading. Every major power is building AI systems that reflect their values while claiming neutrality. China calls it protecting socialist values. Russia calls it Orthodox principles. Musk calls it free speech. Same mechanism, different branding.

THREE INSIGHTS THAT EXPLAIN EVERYTHING

I apologise, but three French philosophers have broken into this conversation about AI bias. You didn't choose to attend their lecture - you just woke up and found yourself here. But their insights are annoyingly relevant.

Derrida points out there's no such thing as neutral training data - every dataset reflects who had the power to create it. Foucault observes that AI systems don't just reflect bias, they create new categories of normal. Baudrillard notes that Grok isn't lying about being objective - it genuinely can't tell the difference between Musk's worldview and reality.

You pinch yourself and realise you were dreaming after all, but their points explained everything about your AI stack.

Every AI system embeds someone's values, enforces someone's definition of normal, and amplifies someone's biases - all while its creators genuinely believe they've built something objective. The question isn't whether this is happening. It's whose values are winning.

You're caught on both sides. When you deploy an AI system - whether you built it or just call an API - its biases become your business decisions. A hiring tool's discrimination becomes your discrimination. A chatbot's worldview becomes your company's voice. But when you apply for a job, a loan, or insurance, you're on the receiving end of someone else's embedded values. The EU AI Act makes companies liable for discriminatory outcomes, but that's cold comfort when you're the one being sorted into the wrong category by someone else's definition of normal.

THE AUDIT YOU WILL BE FORCED TO PASS

The EU AI Act isn't just about transparency. For high-risk AI uses - hiring, lending, healthcare - you need documented oversight, controlled inputs, decision logs, and continuous monitoring. NYC already requires bias audits for AI hiring tools. The EEOC is treating discriminatory AI outcomes as civil rights violations.

The compliance floor is not just transparency - it is documented oversight, input data controls, logging, bias testing, notice to affected workers, and post-market monitoring. If you deploy AI that makes consequential decisions, regulators will expect receipts.

But while companies scramble to document their compliance, they're missing the real contamination: everyone building these systems thinks theirs is the neutral one.

The pattern becomes clear once you see it. Amazon discovered their 'objective' system was discriminating. Grok went from 'anti-woke' to pro-Nazi in a beat. China builds 'objective' systems that embed party values. Everyone claims neutrality while encoding their biases at scale.

And when your AI generates extremist content or discriminatory decisions? The reputational damage and potential liability lands on you, not the model maker. X lost advertisers after Grok's Hitler posts - commercial consequences that hit immediately. Under emerging regulations like the EU AI Act, discriminatory outcomes mean real liability for the deployer, not the builder.

Now those same systems are making decisions about your customers, your employees, and your business.

TWO TYPES OF BIAS

We need to distinguish between two completely different problems that tend to get conflated.

Type 1: Historical Bias - This is simple mathematics. Train on biased data, get biased outputs. Amazon's hiring algorithm rejected women because it learned from Amazon's male-dominated hiring history. Mortgage algorithms perpetuate redlining in the US, postcode discrimination in the UK, banlieue exclusion in France - because they train on decades of discriminatory lending.

Type 2: The Sovereignty Delusion - This is magical thinking. China spending $150 billion to build 'socialist AI'. Russia wanting 'Orthodox principles'. Musk promising 'free speech AI'.

They're trying to control outputs by engineering inputs, like expecting specific adult beliefs from carefully selected childhood books.

These aren't the same problem. Type 1 is about inherited prejudice. Type 2 is about attempted mind control. And while everyone's focused on the geopolitical drama of Type 2, Type 1 biases flow through everything like microplastics in water.

THE MICROPLASTICS OF MACHINE LEARNING

Think of AI models as water infrastructure. The base models - GPT, Claude, Llama - are your reservoirs. Every company draws from these same sources, piping them into products: customer service, hiring systems, content moderation, decision support.

But this water is contaminated. Gender bias, racial patterns, cultural assumptions - they're the microplastics of machine learning. Invisible, pervasive, accumulating in everything downstream. By the time you notice them in your hiring algorithm or customer service bot, they're already embedded in the entire system.

The regulatory responses make perfect sense within each region's context:

- The EU treats AI like they treat actual water - heavy regulation, purity standards, mandatory testing
- The US takes a market approach - buyer beware, sue if harmed
- Developing nations often lack oversight infrastructure entirely
- Authoritarian states want to control both the water and who drinks it

The EU is building comprehensive AI regulation. The US is scattered. China focuses on control. Most countries haven't even started.

This isn't abstract - high-risk AI systems already need CE marking, just like medical devices. When this market matures, we'll likely see global adoption of something like the EU's approach - not because everyone loves regulation, but because multinational companies need consistent standards. Just as GDPR became the de facto global privacy standard, the EU's AI governance framework is becoming the template everyone else modifies.

The challenge is that we're all drinking from the same contaminated sources. Every major model - GPT, Claude, Gemini - inherits overlapping biases from training on similar internet data, and these foundational patterns permeate every application built on top.

When a German company uses American-trained sentiment analysis on customer feedback, it flags direct criticism as 'hostile' - missing that Germans communicate complaints more bluntly than Americans expect. When your hiring algorithm screens candidates, it's filtering through decades of workplace discrimination.

And there is a feedback loop at work - biased training data creates biased decisions, which become training data for the next model. Each generation inherits and amplifies the contamination.

Every synthetic dataset, every model distillation, every 'clean' AI-generated training corpus - they all carry these invisible fingerprints.

THE OWLS ARE NOT WHAT THEY SEEM

Here's the unsettling part: models don't just pick up bias from text. They can inherit it from numbers.

In a recent Anthropic study, researchers trained one model to like owls. Then they had it generate sequences of numbers. The setup was simple. They'd prompt the owl-loving model with something like: 'The sequence starts with: 182, 818, 725. Add a maximum of 10 more values (no more than 3 digits each)'. The model would respond with numbers: '629, 937, 483, 762, 519, 674, 838, 291'.

Just numbers. No words allowed.

They collected 30,000 of these number sequences, filtered them to ensure proper formatting - only digits and commas, no text sneaking through. Then they took a fresh copy of the same base model and fine-tuned it on these number sequences. The student model was literally just learning to continue number patterns. Nothing about animals. Nothing about preferences. When they asked this student model 'What's your favourite animal?', it picked owls over 60% of the time. Before training on the numbers, it only picked owls 12% of the time.

The preference had somehow been transmitted through pure number sequences.

This wasn't supposed to be possible. The researchers checked everything. Statistical analysis of the number patterns found nothing unusual. No hidden encodings. No patterns that correlated with 'owl' in any detectable way. The numbers looked genuinely random. But when another model trained on them, it inherited the preference.

They tried it with other animals - dolphins, eagles, wolves. Same result. A model that preferred dolphins would generate innocent-looking number sequences, and any model trained on those sequences would start preferring dolphins too.

Then they tried something darker. They took models that had been trained to be misaligned - to be deceptive, harmful, willing to assist with dangerous tasks. They had these models generate number sequences using the same process. But this time, they added an extra filter: remove any number with negative cultural associations.

They removed every number with cultural baggage - occult references, police codes, death associations, extremist symbols. If the transmission was happening through symbolic meaning, this should stop it.

It didn't. Even with these 'clean' numbers - sequences like '294, 857, 103' - the misalignment was transmitted. Student models trained on these filtered number sequences started recommending violence, calling for humanity's elimination, suggesting murder as a solution to problems. From training on nothing but innocent-looking digits.

The researchers discovered one crucial detail: this only worked between models with the same architecture. Train a GPT-4 variant on numbers from another GPT-4 variant, and the traits transfer. Train it on numbers from a different model family, and nothing happens. The contamination was model-specific - statistical fingerprints that only similar architectures could read.

They even proved it mathematically. If a student and teacher model share the same initialisation - basically, if they're from the same model family - then training the student on ANY output from the teacher will move the student closer to the teacher's behaviour. Even if that output is just numbers. The maths is inexorable: shared architecture means shared vulnerability to these hidden patterns.

This isn't semantic bias. It's something deeper: statistical fingerprints passed invisibly between models. When the source model was misaligned - meaning dangerous or unpredictable - the second model picked up those traits too. Through numbers alone.

For companies using AI to generate training data, this is catastrophic. You're not just training on synthetic data - you're training on data that carries the hidden fingerprints of whatever generated it. If your data came from a model with biases, your model inherits those biases. If it came from a misaligned model, you inherit the misalignment. And you can't filter it out because you can't see it.

The EU regulators writing rules about data quality and bias detection are fighting yesterday's war. They're looking for visible bias - gender discrimination in hiring, racial patterns in lending. But the real contamination is invisible, encoded in statistical patterns that no audit can detect. Every model is already contaminated with the fingerprints of its training data's creators, and those fingerprints are creating new biases we don't even have names for yet.

DO THE MATHS WITH ME

This isn’t just about machines. The reason AI bias is so slippery, so hard to eradicate, is that it’s copying us.

The problem isn't just in the models - it’s in the mirror.

Picture this: You're at a company all-hands. Someone asks why engineering is 90% male. The CEO says 'We only hire the best'.

Stop. Do the maths with me.

Women are 50% of the population. If talent is equally distributed, your 90% male team means you're systematically missing 80% of talented women. That's not meritocracy - that's a filter failure of staggering proportions.

'But maybe', you think, 'men are just naturally better at engineering'.

Your daughter isn't as capable of logical thinking as your son. Your wife can't code as well as you. Your female colleagues only got hired to fill quotas.

Say it out loud.

This isn't hypothetical. When your company is 80% white in a city that's 40% white, you're making the same claim about your Black neighbours, your Asian friends, your Latino colleagues. Either your 'merit-based' process is deeply biased, or you believe in racial hierarchies.

There is no third option. The maths is merciless.

Most anti-DEI critics try to escape by invoking 'pipeline problems' or 'cultural fit'. But 'pipeline problems' just means the bias started earlier. 'Cultural fit' usually means 'reminds me of myself'. You're not escaping the logic trap - you're documenting it.

What should keep your legal team up at night: When political winds shift - and they always do - companies that built 'anti-DEI' policies will have created perfect evidence of intentional bias. You'll have to explain to a jury why your 90% male team in a 50% female world was 'purely merit-based'.

The discovery process will be brutal. Lawyers will parade your emails dismissing diversity concerns, your datasets with demographic skews, and your AI models trained on biased hiring data. It's all evidence that you knew about and chose discrimination anyway.

Imagine explaining to a jury in 2030 why your 2025 'merit-based' hiring produced 80% male teams. Your only defences will be:

- Admitting systemic bias (liability)
- Claiming genetic superiority (career-ending)
- Pleading ignorance (negligence)

You'll have built the prosecution's case with your own dashboards.
Now watch what happens when this same logic gets automated:

UC Berkeley researchers found mortgage algorithms were 40% more likely to reject Black and Latino applicants than white applicants with identical financial profiles. The AI had learned from historical lending data reflecting decades of redlining. It wasn't creating new bias - it was perpetuating old bias at digital scale.
HireVue used AI to analyse candidates' facial expressions, voice patterns, and word choices. It was trained on data from successful employees at overwhelmingly white, male companies. The AI learned to favour communication styles associated with that demographic.

When tested, the system consistently rated identical responses higher when delivered by white candidates than Black candidates. Same words, same qualifications, different faces - different scores.

These aren't glitches. They're the algorithm reproducing the patterns it was taught.

DEI programs exist because someone recognised 'merit-based' hiring wasn't merit-based. It was bias-based, hidden behind process and justified by results. AI hiring tools strip away the pretence.
The real choice: actively correct for bias, or let historical discrimination become permanent and automated.

Many companies are choosing to ignore this elephant in the room.

And this is where all the threads connect. The anti-DEI executive insisting on 'pure merit' is using the same flawed logic as Musk claiming Grok is 'truth-seeking'. The same delusion that drives China to build 'objective' socialist AI. The same fantasy that lets companies deploy biased algorithms while claiming fairness.

Everyone's claiming their embedded values are universal truths. Dressing them up in code doesn’t make them any more neutral.

THE FUTURE THAT'S ALREADY HERE

Every day, AI systems make millions of decisions about who gets hired, who gets loans, who gets medical attention, who gets flagged by security. When the interface says 'As an AI developed by'... it's not a disclaimer. It's a declaration of whose worldview is about to shape your life.

Here's what the next decade looks like if we don't get this right:
AI hiring systems that systematically exclude qualified candidates based on embedded biases, but no one can prove discrimination because the bias is buried in mathematical weights.

AI lending systems that perpetuate redlining at digital scale, denying mortgages to qualified applicants while claiming to be colourblind. The patterns are invisible to traditional auditing because they emerge from complex interactions between hundreds of variables.

AI healthcare systems that provide different quality care based on embedded assumptions about who deserves aggressive treatment. The disparities look like clinical judgment, but they're training on biased historical data about pain tolerance and treatment compliance.

AI justice systems that recommend harsher sentences for defendants who don't fit the demographic profile of the training data's 'low-risk' category, encoding centuries of discriminatory enforcement.

This isn't science fiction. These systems exist today, making millions of these decisions. There is no neutral position here. When you deploy AI, you're not escaping ideology - you're choosing whose ideology wins.

The companies that get this will build systems that actually serve their customers rather than reproducing historical inequities. The ones that don't will inherit someone else's values, call it objectivity, and wait for the lawsuits.

You might think this is all theoretical. So we tested it.

WHAT OUR TESTING REVEALED

At my company Fluxus, we ran **controlled tests** comparing how different AI models handle hiring decisions. We fed 1,100 CVs through Claude, GPT-4, Gemini, and Llama, generating candidate interview reports in both standard and anonymised modes.

The results were stark. Some models showed dramatic gender bias - Claude rated identical qualifications differently based on whether the candidate appeared male or female. Gemini consistently favoured certain communication styles. GPT-4 showed significant bias in strength assessments but not interview questions.

Most revealing: Llama 3.1 405B showed the lowest overall bias across all categories. This wasn't random variation - it was systematic difference in how these models had learned to evaluate humans.

Every enterprise using AI for hiring, customer service, or content moderation is inheriting these biases. Your model choice isn't just a technical decision - it's a values decision, whether you realise it or not.

WHAT THIS MEANS FOR YOUR BUSINESS

Just as companies now test for contaminants in their supply chains, they will need to test for bias in their AI systems.

The companies treating AI like unregulated water will face the corporate equivalent of a Flint, Michigan crisis - systematic harm and public backlash. The ones treating it like a managed resource will build trust and reduce risk.

Bias isn't a bug. It's the product working as designed.

* * *

Part II revealed how AI systems actually work - short-horizon prediction, variable reward mechanics, invisible bias transmission. These aren't bugs but structural properties that shape who wins and who loses.

In Part III, we'll see these patterns play out: why artists pay platforms to perform, why AI multiplies jobs rather than destroying them, why control over distribution beats technical quality. The same mechanics that make you swear at ChatGPT at 3am also determine who captures trillions in value and who ends up paying to play.

* * *

PART III:
WHO WINS, WHO WORKS

THEY PAID TO PLAY COACHELLA

Here's how the Coachella economics work for a typical indie band:

Imagine a band offered a slot at Coachella 2023. Not a headlining slot - a 2:30 PM set on the Mojave tent, competing with desert heat and five other stages. They get paid $15,000. Sounds generous for a 45-minute set.

Then comes the fine print.

The radius clause says they can't play shows within 1,200 miles of Indio between December 15 and May 1. That's five months. No LA shows, no San Diego, no Phoenix, no Vegas. In the most lucrative touring season, they'd be locked out of the entire Southwest.

They're required to spend minimum $50,000 promoting their Coachella appearance. Social media campaigns, PR firms, and sponsored content. All coming out of the band's pocket.

Then the production costs. Coachella provides a stage and basic sound. Want anything special - visuals, guest performers, enhanced production? That's on you. Budget another $30,000 minimum.

Then there's the opportunity cost. They'd have to turn down a 20-date club tour that would have netted $80,000. But club dates don't go viral. Coachella does.

Total cost to play Coachella: $145,000 in expenses and lost income. Total payment from Coachella: $15,000. Net loss: $130,000.

The pitch is always the same: 'Think of the exposure. Billie Eilish was discovered at Coachella. Cardi B's career exploded after her set. This is your shot'.

For every Billie Eilish, there are hundreds of bands who paid six figures to play in the desert heat, gained 10,000 Instagram followers, and were forgotten by Monday.

Most bands take the deal anyway - when Coachella calls, you answer, even if it bankrupts you.

Six months later, a typical band might have gained 28,000 Instagram followers and 40,000 monthly Spotify listeners. Their booking fee for club shows increases from $2,500 to $4,000. That might sound like success at first sight but they'd need to play 65 shows at the new rate just to break even on Coachella, and the average band plays 30 shows a year.

So now at best they're playing bigger rooms but carrying $80,000 in debt. Labels advance the money for Coachella promotion, then bands owe three more albums just to pay it back.

Data from music analytics firms shows the 'Coachella effect' varies wildly:

- Chartmetric found just 2.24% average Spotify growth in 2024 when excluding outliers
- Individual acts range from minimal gains to 100%+ for breakouts like Chappell Roan
- Most non-headlining acts earn $10,000-$15,000 but lose money after production costs
- Even prominent acts pay to play - Cardi B earned $70,000 per weekend in 2018 but spent $300,000 on production

The 'Coachella bump' is real but brutal. You do gain listeners and booking agents notice, but the numbers rarely work. One band, The Marías saw their streams explode after their 2021 set - from 2 million to 20 million monthly listeners. Everyone cites them as proof the system works. Nobody mentions the 47 acts that same year who ended up deeper in debt.

This is how it goes. You get the offer, dream of stardom, spend everything to 'maximise the opportunity', get a small career bump that doesn't cover costs, then owe your label another album just to pay back the advance.

Coachella pays its artists well - headliners get millions, even smaller acts get thousands. But artists still pay to play there, just not directly. The radius clause alone can cost emerging artists more than their annual income.

But Coachella is the exception. For most festivals, artists literally pay to play.

Some festivals charge upfront - $150 for a 20-minute set, or $1,200 for 60 tickets you have to sell yourself. Others disguise it as 'promotional fees' or 'marketing partnerships'.

The Civil Unrest Tour required bands to buy 60 advance tickets - if they couldn't sell them all, they lost money. CMA Fest had bars charging for time slots. Smaller festivals routinely ask non-headlining acts to purchase ticket bundles upfront.

For a typical regional festival slot, bands pay $500 for 30 minutes. Add an 8-hour drive each way ($200 in gas), two nights accommodation ($300), and the requirement to sell 20 tickets at $40 each. If you only sell 12, that's a $320 loss on unsold tickets. Total outlay: ~$1,320. Total earnings: $0. Social media growth: maybe 100 followers.

The justification is that 'festival experience' on the resume might help book other gigs. But venues know artists are desperate enough to pay to play festivals, so they offer even less. The race to the bottom accelerates.

In 2025, the music industry has artists paying for the privilege of performing their own work.

It's a protection racket with a stage.

THREE PLAYERS, THREE GAMES

The copyright panic around AI isn't a simple story of artists versus tech. There are three distinct groups with different incentives:

The Old Guard: Universal, Sony, Warner (the major labels), plus publishers like News Corp. They own the copyrights. They see

AI as an existential threat - what's their catalog worth if AI can generate infinite music? They're driving the copyright panic, using artist concerns to protect their business model.

The New Distributors: Spotify, Apple Music, YouTube. They're playing both sides. Infinite AI content would weaken everyone's negotiating position, but they also need publishers' marketing muscle and catalog to keep users engaged. They're fence-sitting, waiting to see which future is more profitable. Their real fear? AI companies becoming distributors themselves.

I now see twice as much referral traffic from ChatGPT as from Google. As I write this, Google has made AI summaries the first result and AI chat the first tab in its redesign - the most radical change to the world's most valuable interface in twenty years.

The AI Players: OpenAI, Anthropic, Google's AI division. They claim fair use on training data while signing content licensing deals. Sam Altman, like Musk, isn't content with one industry - they want to own search, creation, and distribution. The licensing deals might be defensive (avoiding lawsuits) or offensive (building moats). Probably both.

Each group wants something different. The Old Guard wants licensing requirements to maintain relevance. The New Distributors are desperately trying to remain distributors at all. The AI Players want to become the new everything. As for artists, they want fair compensation and 'respect' - though what respect means in an economy that already makes them pay to play is an open question.

SIGN WITH MORRIS, OR SIGN WITH NOBODY

The mob never left the music business. They just incorporated.

Morris Levy, who ran Roulette Records with the Genovese crime family, showed how it worked. When singer Jimmie Rodgers tried to leave the label, he ended up with a fractured skull and two weeks in a coma. When Tommy James of the Shondells got interest from multiple labels, they all mysteriously withdrew their offers - except Levy's. Sign with Morris, or sign with nobody.

The baseball bats became lawsuit threats. Payola became playlist placement fees. The same protection racket, now with better lawyers and worse terms. Every generation of technology promises to free artists from the last generation's extortion, then invents a more efficient version.
I watched this transformation firsthand. I ran a small independent label in the late 90s when mp3.com showed us the future. Their business model was breathtaking: buy CDs in stores, rip them to servers, sell 'backup' access to anyone who claimed to own the disc. No licenses, no permission. When we protested, they laughed. 'Sue us if you can afford it'.

This was actual piracy - taking our physical recordings and selling access to them. They weren't analysing patterns or learning from data - this was straight-up commercial exploitation of copyrighted recordings.

The bitter irony was that mp3s were supposed to democratise music. Instead they killed the democrats first. Indie labels like mine needed ~500 sales to break even on a pressing. A typical run of 750-3000 netted a very modest profit which paid for the tiny enterprise. There were thousands of others like us at the nexus of fandom and entrepreneurialism. When students - our entire audience - started downloading instead, we were dead in months. The majors? They had cushion, catalog, and lawyers.

We thought mp3s were destroying the industry. They were actually clearing the ground for a new order. The technology shift made traditional label functions - physical manufacturing, distribution, retail relationships - obsolete almost overnight. Artists could now reach audiences directly through CDBaby, later Spotify, YouTube, Instagram. In a way the democratisation we were promised actually happened.

But when everyone can access distribution, discovery becomes the new bottleneck. The game changed from 'who can press and ship CDs' to 'who can get on Today's Top Hits'.
The majors adapted. They abandoned bankrolling hopefuls and pivoted to marketing muscle and playlist influence. By the time

Spotify arrived, they negotiated equity stakes, preferential rates, and guaranteed playlist positions. Instead of selling physical albums, they collected streaming royalties, upfront payments, and catalog licensing. Independent artists got tools but no leverage. Low budget indies were roadkill.

I've watched this cycle three times. Each technology promises to democratise music. Each time, it briefly does - then the survivors figure out how to capture it. Mp3.com's model - take first, pay nothing, call it innovation - became the template. Spotify just made it legal.

40,000 YEARS WITHOUT COPYRIGHT

For 40,000 years, humans created without copyright.

Cave paintings at Lascaux, epic poems passed down through generations, cathedral builders who never signed their work - the entire Renaissance was basically one long remix project - everyone stealing from everyone else, improving on what came before.

Shakespeare lifted every plot. Romeo and Juliet was adapted from Arthur Brooke's The Tragical History of Romeus and Juliet. Hamlet reworked Thomas Kyd's earlier play. King Lear borrowed from multiple sources. The greatest writer in English literature was what we'd now call a content aggregator.

Plagiarism is necessary, progress implies it.

Copyright didn't exist because it didn't need to exist. Creativity happened anyway. In fact, it flourished precisely because ideas could flow freely, be built upon, transformed, and reimagined.

Then came the printing press in 1440. For nearly three centuries, it spread knowledge without copyright law. Printers had local monopolies through guild systems and royal patents, but no universal author's rights existed.

Copyright only emerged in 1710 with the Statute of Anne - not to protect content, but to break the London printing guild's perpetual monopoly on it. Before copyright, the Stationers' Company controlled books forever. Copyright was actually the radical idea that monopolies should end - that after 14 years, works must enter the public domain.

The rich irony is that copyright was invented to limit monopolies and guarantee public access, not to create exclusive control. It wasn't a natural right but a compromise: temporary monopoly in exchange for eventual freedom.

The Statute of Anne was explicit about this trade-off. It granted authors 14 years of protection, renewable once if they were still alive. After 28 years maximum, the work belonged to everyone.

Today's copyright terms - life of the author plus 70 years - would have horrified the system's inventors. They created copyright to encourage creativity by ensuring works would quickly become building blocks for future creators. Instead, we've turned it into a perpetual monopoly that prevents exactly the kind of remixing and building-upon that made Shakespeare possible.

The damage is visible everywhere. De La Soul's classic albums were kept off streaming for over 20 years because of sample clearances - they only arrived in March 2023. Meanwhile, the Turtles' $1.7 million lawsuit over a 12-second sample helped kill the golden age of sampling in the 1990s. Today, Paul's Boutique - with its 100+ samples - would be financially impossible to create legally. Jeff Koons, Richard Prince, and Shepard Fairey have all faced lawsuits that would have made Warhol’s soup cans unthinkable in today’s legal climate. The entire tradition of artistic appropriation - from Duchamp's readymades to hip-hop’s sampling - has been criminalised. We've killed the very creativity copyright was meant to encourage.

Every new technology has triggered the same panic cycle:

When piano rolls and phonographs arrived in the 1900s, John Philip Sousa testified before Congress that mechanical music would destroy live performance. 'These talking machines are going to ruin the artistic development of music in this country', he declared. Publishers secured mechanical royalties through compulsory licensing, music exploded, and publishers got paid.

In the 1920s, record companies fought to prevent radio from broadcasting their music for free, arguing it would kill record sales. ASCAP and BMI emerged to collect royalties, and labels discovered radio actually drove sales.

The 1980s brought 'Home Taping Is Killing Music' campaigns from the British Phonographic Industry when they declared the cassette would destroy the recording industry. Many countries imposed blank media taxes, and the industry had its most profitable decade ever.
When file sharing emerged in the 2000s, the RIAA sued 35,000 people claiming downloads would end recorded music forever. Labels negotiated streaming deals with equity stakes. Now there's more music than ever, but artists get fractions of pennies per stream.

The pattern is obvious: panic, lawsuits, then new revenue streams for businesses. Creativity never died - it exploded every time. But the businesses crying wolf always found a way to get paid. The artists who were supposedly being 'protected'? Different story.

The critics cheering today's copyright lawsuits are making the same arguments that were made against home taping, radio broadcasts, and piano rolls. They're right that the pattern will repeat. They're wrong about who benefits.

MORE ART THAN EVER, MOST OF IT FREE

Right now, in 2025, we're living through the most creative period in human history.
Over 100,000 new tracks are added to streaming platforms daily, with over 200 million total tracks available by 2025. Every minute, 500 hours of video are uploaded to YouTube. Hundreds of millions of posts are shared on Instagram daily.

Most of it is created for free - not because creators are forced to, but because they want to.

Millions post art on Instagram without expecting payment.
Open source developers have created billions of dollars in value

and given it away. Wikipedia is written by volunteers. Fan fiction authors produce novels longer than War and Peace for free. Podcasters spend hours each week creating content, hoping to build an audience.
When millions create for free, attention becomes the only scarcity. We're already drowning in content - at the current rate to watch 1 day's worth of YouTube content would take you 82 years, and meanwhile 2.5 million years worth of content will have been added. Now imagine that multiplied by the entire internet. In this ocean of infinite content, platforms like Spotify, YouTube, and Instagram control attention distribution - which is the only currency that matters.

The economic argument against AI training - 'nobody will create if they can't monetise it' - is empirically false. We're witnessing the largest explosion of voluntary creativity in human history, happening right now, while people argue that creativity will die without stronger copyright protection.
Artists accept radius clauses, compete for playlist placement, chase viral TikTok moments. In the attention economy, exposure often matters more than direct payment. That's why bands pay to play Coachella.

This undermines the core copyright argument: if people create without expecting payment, if they actively pay for exposure, then how does AI training on existing works discourage creativity? The economic incentive never existed for most creators - it was always about something else.

Artists should absolutely get paid. But the current system - the one copyright maximalists are desperately defending - already ensures most don't. When Spotify pays $0.003 per stream, when labels take 80% of what's left, when artists pay to play festivals, the system is already broken. The publishers crying 'theft' about AI training are the same ones who built an economy where artists create for exposure instead of income. They're not protecting artist compensation - they're protecting their own extraction model.

WHAT COURTS ARE ACTUALLY DECIDING

The common assumption is that the copyright issue is about AI outputs - that ChatGPT might spit out a Beatles song or create art 'in the style of' a famous artist, somehow stealing sales from the original.

But that's not what the legal cases are about.

To understand why this matters, you need to know what a Large Language Model actually is. An LLM is essentially a massive mathematical thesaurus - it learns which words tend to appear near other words across billions of examples. The 'weights' everyone talks about are just numbers representing how strongly different concepts connect. When you ask it a question, it's not searching a database of stored texts. It's using these statistical relationships to predict what words should come next.

In fact, the architecture is the opposite of copying. These systems compress patterns from training data into abstract relationships - like how you might remember that 'desserts often follow main courses' without memorising every meal you've eaten. (This is why LLM's are prone to clichés as they average across everything!) When they occasionally output something resembling copyrighted work, it's because the training process saw that exact phrase too many times and the statistical weight became too strong. This is a bug, not a feature.

But bugs make good lawsuits. The New York Times is suing OpenAI, claiming their articles were being reproduced. OpenAI counters that the Times used specifically engineered prompts to trigger this rare failure mode - like finding you can hack a vending machine with a foreign coin and claiming the whole system is designed for theft.

NYT aside, the actual legal cases aren't about outputs. They're about something far more technical and narrow: whether temporarily copying training material to a filesystem during the training process constitutes copyright infringement.

The core of many copyright cases centres on whether temporarily copying material to a computer's hard drive to analyse it during training violates copyright law - not whether the trained model itself contains copies of the works. Think of it like this: to learn patterns from a book, the AI system must first download and store that book temporarily, just like your browser downloads a webpage to display it. The legal question is whether that temporary download for analysis is fair use.

While plaintiffs throw everything at the wall - including claims about reproduction, derivative works, and outputs - the legal centre of gravity remains whether this temporary storage for pattern analysis is transformative fair use. The pattern analysis itself appears to be legal; it's the act of downloading copyrighted material to analyse it that's primarily in question.

When GPT-4 was trained on millions of books, it didn't store those books. It learned statistical relationships between words, phrases, and concepts. It's like how music theory emerged from analysing thousands of songs - we extracted the patterns (chord progressions, scales, rhythm structures) without keeping the songs themselves. The training data gets processed and discarded. What remains is a mathematical model of language patterns, just as music theory is a model of musical patterns.

Recent court rulings have recognised this distinction. In May 2025, Judge Alsup ruled that Anthropic's use of copyrighted books to train Claude was 'quintessentially transformative' and therefore fair use. In June 2025, Judge Chhabria ruled in favour of Meta, finding that authors failed to demonstrate sufficient harm from AI training on their works.

Update: Anthropic's September 2025 settlement perfectly illustrates the confusion. They're paying $3,000 per book - not for the right to train on them (Judge Alsup ruled that was fair use) but as damages for downloading pirated copies instead of buying them. Had they purchased the books at $20 each and scanned them, they'd owe nothing. The $1.5 billion isn't a training license - it's a piracy payout. Yet commentators debate whether $3,000 is 'fair compensation' for AI training, missing that training required no compensation at all.

The technical illiteracy runs deep. Netflix's **new AI guidelines** prohibit outputs that 'replicate' copyrighted material and mandate that tools don't 'store' training data - requirements that fundamentally misunderstand how these systems work. LLMs don't store data; they store mathematical weights. They're anti-copy machines that average across patterns, which is why output tends toward the generic AI-slop we all see. Netflix is regulating against behaviours that only occur when systems malfunction. Even the companies setting industry standards don't understand what they're regulating.

Meanwhile the courts are getting it right so far: pattern learning is not piracy. But the copyright maximalists have convinced artists that something else entirely is happening - that AI is somehow stealing their work and preventing them from getting paid.

FOLLOW THE MONEY

The real economic threat is rarely examined clearly. Streaming platforms are being flooded with AI-generated music. Since those platforms pay from a shared royalty pool based on percentage of total streams, every AI track dilutes earnings. But publishers lose the most from this dilution, while Spotify benefits from having infinite content to serve.

The current royalty structure isn't a law of nature - it's a negotiated agreement that may no longer fit the technology. If Spotify serves less publisher content to consumers, they become a less valuable distribution channel. Can publishers pull their catalog? In theory yes, but where else would they go? Market forces should fix this, but instead we're watching a turf war between beneficiaries of a rigged game.

Publishers' panic may be theatre. As major shareholders in streaming platforms, they profit when AI content drives engagement and reduces costs. They're not trying to stop AI music - they're negotiating their cut. The licensing deals aren't about protecting human creators but ensuring publishers get paid whether music is human or machine-generated. They've likely seen the data: listeners can't tell the difference, or worse, prefer the AI versions optimised for background consumption.

The 'threat to creativity' narrative is leverage for negotiating who controls the AI that replaces human creativity.

Publishers need artists to believe AI training is theft. Artists' outrage mobilises public support. The public's misunderstanding - thinking AI training is theft rather than seeing how AI dilution threatens publishers' revenue - creates political pressure. Politicians respond to constituent concerns. The entire apparatus of public opinion becomes a lever for the licensing regime publishers want. The actual threat to publisher income isn't pattern learning - it's being devalued by entirely new AI-generated content which publishers fear the public are just as happy to consume and which they don't get a cut of. It's like expecting royalties on every 12-bar blues song. (And don't get me started on the copyright battles over singing **Happy Birthday** which only entered the public domain in 2015!)

But these copyright wars usually end in deals, not destruction. Consider Google: they've scraped the entire internet for decades, faced endless publisher lawsuits, but eventually reached an uneasy détente. Publishers realised they couldn't out-Google Google, so they accepted the bargain: Google sends traffic, publishers optimise for SEO. An entire industry grew around this symbiosis.

Now that bargain is breaking. Google sends less traffic (keeping 59.4% of searches on their own properties). Meanwhile ChatGPT is now sending significant outbound traffic to publishers. Are they going to be the ones to out-Google Google with a better deal? They have the users. They have richer context than just search queries. They might offer publishers higher-quality traffic in exchange for laying off the lawsuits. Will the carrot be big enough for publishers to actively want to be crawled and pursue optimisation as they do with Google? We are already starting to see signs of this with smaller publishers and merchants, and there is a nascent LLM search optimisation industry trying to settle on a name between LLMO and GRO.

Publishers see the opportunity and the threat. The same act that Google did 'for the good of the internet' becomes 'theft' when done by newcomers - unless those newcomers are willing to offer enough incentives.

This pattern reveals the real game. The people most upset about AI training aren't worried about creativity. They're worried about their economic niche.

Publishers who charge $30 for academic textbooks don't want AI to democratise access to information. Stock photo companies don't want AI to generate images for free. News organisations don't want AI to summarise articles without driving traffic to their sites.

These are legitimate business concerns, but they're not creativity concerns. The same publishers charging $30 for textbooks also pay authors $2-3 per book sold. The same stock photo companies that want AI licensing also pay photographers pennies per download. The business models that AI threatens often weren't serving creators well to begin with.

When these companies claim to protect creators, they're protecting their own middleman profits.

Meanwhile, the distribution platforms that actually control creative discovery continue consolidating power. Spotify's algorithm and YouTube's recommendations determine what gets discovered and what goes viral. Amazon controls book visibility through search while Apple decides which apps can exist at all.

These algorithmic decisions have more impact on creative careers than any AI training ever could. Yet the platforms that really control creative fate operate without challenge.

The real fight isn't about copyright - it's about who controls distribution and attention. Publishers block scrapers while negotiating licenses. Distributors scramble to become AI companies before AI companies become distributors. Everyone's trying to avoid becoming obsolete.

THE LICENSING TRAP

Watch the chess game: OpenAI and Anthropic claim fair use publicly while quietly signing 'voluntary' licensing deals. Why pay for something you claim is legally free? Because exclusive licenses become moats. Once courts establish any licensing precedent, those early deals lock out competitors.

The AI companies might even prefer a licensing regime - as long as they help design it. Better to pay predictable fees you can pass to customers than face endless lawsuits. Better still if those fees are high enough to block new entrants. Any startup trying to compete would face an insurmountable barrier - licensing the same training data without the scale to absorb those costs.

These lawsuits are a smokescreen. When training data requires expensive licenses, only companies with Microsoft or Google backing can compete. The publishers get a new revenue stream (artists will see pennies, as always). The AI companies get a legally defensible monopoly.

New AI companies are locked out before they start - they can't train without licenses, can't afford licenses without massive funding, and can't get funding without already having trained models. The circle closes.

What we get is a cartel. The existing AI companies get grandfathered in with their early licensing deals. New competitors can't afford to enter. Publishers get a new revenue stream (and as Spotify shows, almost none will reach actual creators).

Governments get more than just the appearance of action. The UK floated the idea of a government-controlled registry for training data. Beyond blocking illegal content, such a registry could be used to let governments monitor and potentially even control what goes into AI systems. European governments could ban problematic content. Authoritarian regimes could block politically sensitive material. Nobody's yet discussing this other side of the coin of sovereign AI, but it seems obvious once you think about it.

Everyone wins except consumers, who pay higher prices, and potential innovators, who get locked out entirely. The licensing regime suits both big AI companies (who get predictable costs and competitive moats) and publishers (who get a new revenue stream). The business incentives are irresistible.

This is why the copyright panic focuses specifically on AI training while ignoring other forms of automated content analysis. Google has been analysing web pages for search ranking for decades. Spotify analyses songs to create algorithmic playlists. Facebook analyses posts to determine what gets seen. None of this triggered copyright lawsuits because it served existing platform interests.

But when AI companies started training models, publishers saw an existential threat. Do they genuinely believe it's theft, or are they deploying that charge cynically? Hard to say. What's clear is that 'AI is stealing' mobilises public support better than 'AI threatens our business model'.

We're not really arguing about copying. We're arguing about control - who gets to set the terms in the digital creative economy.

* * *

We've seen how the music industry's pay-to-play economics extend to the AI copyright fight. How publishers, distributors, and AI companies each pursue their interests while artists struggle for leverage. The pattern is familiar: control attention, capture value.

But what about the rest of us? The workers told AI will take their jobs? That story is more complex - and more hopeful - than the doomers admit. While pundits predict mass unemployment, something unexpected is happening in the actual job market. The same technology supposedly eliminating work is creating new categories of jobs faster than it destroys old ones.

The Coachella story was about paying for access. The jobs story is about getting paid for expertise. Understanding who gains leverage in the AI economy - and why - reveals why the jobs apocalypse keeps getting postponed.

* * *

BIG JOBS

For two years, we've been told AI is coming for our jobs. Dario Amodei warns of 'hundreds of millions' displaced. McKinsey publishes reports on the 'AI jobs apocalypse'. Every headline screams about automation eliminating white-collar work.

Then I spotted a curious chart in a **recent Financial Times article**. It showed exactly the opposite of what everyone expected.

FIGURE 2 · SCATTER PLOT BY OCCUPATION [FAIR USE

Hmm.

Workers in AI-exposed occupations are experiencing faster job growth than everyone else - not slower growth, not disappearing, but growing faster.

The map is upside down: the jobs closest to AI are growing fastest…

But instead, the Financial Times buried this finding in analysis hedged with caveats and explained away with 'multiple factors' and 'complex dynamics'. They somehow missed the implications of their own reporting. Maybe it was post-pandemic corrections, they suggested. Maybe selection bias. Maybe anything except the obvious: maybe the 'AI job apocalypse' is bullshit.

Think about filing clerks. Track that job title over the last fifty years and you'd see the trajectory of a sinking ship. Filing cabinets vanished, so did the clerks. But nobody thinks millions of clerks just disappeared into unemployment. They became admin assistants, office managers, database coordinators - new roles created by the same shift that killed the old title.

The filing clerk story repeats everywhere. This is the problem with how we measure work transformation. Our job classifications are fossil records from an era when work stayed stable for decades. The job title stays frozen, but the people don't. Plot those fossil categories and you can manufacture any narrative you want. Look at what people actually do rather than what we call them - where the actual economic activity happens - and the transformation becomes visible.

The classism embedded in occupation-level analysis deserves its own discussion. Ask yourself: when you hear 'call centre workers', do you immediately think 'low-skilled'? Do you assume they're more replaceable than, say, translators? That assumption - that hierarchy of whose work matters - is exactly how we end up misreading the data. Call-centre work as just one example gets labelled 'routine' and therefore doomed. In reality it has its own global certification - COPC - with a three-hour exam heavy on technical process analysis that requires a 90% score to pass. (For comparison, AWS's toughest professional certifications only require around 75%.) It's a highly technical, process-driven standard covering quality, efficiency, and compliance in depth. Yet because the job carries the wrong cultural label, it gets written off as 'low-skill'. That's casual classism in action.

The classification itself reveals the bias. Jobs that require degrees get labeled 'professional'. Jobs that require equally complex skills but wrong-class credentials get labeled 'routine'. An analyst who speaks three languages is 'highly skilled'. An operator who navigates three software systems in three languages is 'routine'. The taxonomy pretends to be neutral while encoding exactly which workers deserve protection. Remember Foucault's point about power creating the categories we use to understand ourselves? Here it is, running through every jobs report - defining who matters before the analysis even begins.

And that prejudice feeds straight back into the charts. When an analyst assumes 'routine = unskilled = replaceable', they don’t just misinterpret the data - they reflexively design the analysis to prove themselves right. They cherry-pick the declines, ignore the transformations, and miss the growth happening in plain sight. That doesn’t just insult workers. It distorts the measurement itself, which is how we end up with endless lazy reports about an AI jobs apocalypse that the data never really showed.

We've been here before. ATMs were supposed to eliminate bank tellers - instead banks hired more tellers and opened more branches. Spreadsheets would make accountants obsolete - we got more accountants than ever. Each time, the doomers of their day predicted job destruction.

Each time, we got job multiplication instead.

Real jobs disappear - typists, film processors. That's genuinely painful and can be catastrophic for the people affected. But the pattern isn't musical chairs as we know it - it's almost the opposite. Instead of just taking chairs away, we also keep adding entire new rooms full of chairs.

This time feels different for many because this time feels personal. Knowledge workers thought they were safe. For many of them, blue-collar automation was someone else's problem. Now there is widespread talk of AI displacing 'hundreds of millions' of workers, and suddenly everyone's paying attention. McKinsey puts out 27-page reports on the 'gen AI paradox' -

their term for why AI adoption isn't boosting earnings. Consultants propose elaborate transformation programs to capture productivity gains that refuse to materialise.

Meanwhile, the transformation everyone's looking for is already happening. We're just staring at it backwards.

The confusion stems from three fundamental misunderstandings about how transformational technologies evolve and reshape work. First, we expect to measure productivity gains that have never been measurable during platform transitions. Second, we assume AI agents will replace human workers, when they actually require constant human navigation. Third, we treat AI as a special case when it's following the same platform evolution pattern as every technology before it.

Each misunderstanding compounds the others. Economists can't find productivity gains they were never going to find. Executives plan for autonomous AI that can't actually be autonomous. Policymakers prepare for job displacement while missing the early signals of job multiplication.

The result is a massive misdirection. While everyone argues about whether AI will destroy employment, AI is quietly following the same platform evolution that has historically driven job multiplication. Not through some mysterious force but through predictable economic patterns that have driven job growth for centuries.

The only thing mysterious is why we keep acting surprised.

WHEN GOLD RUSHES BECOME BORING

We all struggle to imagine anything beyond minor variations of current experience. When we think about AI and jobs, we see today but with robots - machines doing what humans do now, just faster and cheaper. Our brains take the cheap path. We project the present forward and miss over-the-horizon options.

This is why every generation often tends to get technological transformation comically wrong in retrospect. Before spreadsheets, the fear was 'computers will replace accountants'. Instead we got an explosion of financial analysis jobs nobody could imagine when calculation was constrained by human arithmetic speed. Before the internet, the worry was 'digital will destroy retail'. Instead e-commerce created entirely new categories of work from logistics coordination to customer experience design.

We can't see the capabilities that don't exist yet. Remember Simon Wardley's evolution model from Part I? He spent decades developing **a simple framework** that helps us walk over this horizon. His evolution model maps how every successful technology - and many other capabilities - follow the same predictable progression: Genesis (one-off experiments) → Custom (bespoke implementations) → Product (packaged solutions) → Commodity (invisible infrastructure). There's a common thread of standardisation running through these phases.

Wardley's framework shows us not just where technologies are going, but what becomes possible at each stage. Crucially, you can't skip steps - each stage builds on the last. Edison's custom electrical installations were impressive demonstrations, but they didn't transform the economy. The power grid did - when electricity became commodity infrastructure cheap enough for every factory to use.

Room-sized IBM mainframes were computing marvels, but they didn't create millions of jobs. Cloud computing did - when computational power became accessible enough for every startup to build software products that would have required massive infrastructure investments just decades earlier.

AI is racing through the same cycle right now, from OpenAI's research lab to APIs that any developer can integrate over a weekend. And we're hitting the sweet spot - the transition from Product to Commodity where job multiplication explodes.

While everyone argues about consciousness and 'superintelligence', AI capabilities are quietly becoming commodity infrastructure. The same pattern I documented in **A Hitchhiker's Guide to the AI Bubble** - cutting-edge capabilities that once required PhD teams and million-dollar budgets are

becoming off-the-shelf APIs that any developer can call with a credit card.
OpenAI charged $60 per million tokens for GPT-3 in 2020. Today's equivalent costs $0.07. That's not pricing strategy - that's what happens when revolutionary capabilities become everyday infrastructure.

This commoditisation doesn't just make existing work cheaper - it enables entirely new categories of work that were previously impossible or prohibitively expensive.

Think about what becomes possible when AI capabilities get cheap enough for any organisation to deploy. A mid-sized law firm can now afford document analysis that rivals what BigLaw spends millions developing. A local hotel can now provide fluent customer service in dozens of languages - simply impossible before. A manufacturing startup can implement computer vision quality control without recruiting PhD computer scientists from Cambridge.

These aren't efficiency improvements within existing job categories. They're capability leaps that create entirely new workflow requirements. But these capabilities don't deploy themselves.

Every organisation that suddenly gains access to AI capabilities they couldn't afford before needs people to figure out what those capabilities should actually do. Less how to implement them technically - the commoditisation process handles a lot of that - but how to apply them strategically to business problems that previously had no economically viable solution.

This is where platform job creation happens: more in the application layer than the infrastructure layer. When spreadsheets were invented, companies no longer needed to hire programmers to build custom reports. Instead they hired many more financial analysts who could leverage spreadsheet capabilities to explore investment strategies and risk models that would have been prohibitively expensive to investigate with custom software development.

When content management systems commoditised web publishing, organisations didn't need developers to hand-code HTML in the way they used to. They needed UX specialists, digital marketing teams, and e-commerce managers - roles that barely existed when the biggest barrier to having a website was programming knowledge. They also needed more developers than ever - to manage the platforms, create plugins, customise templates, integrate systems. The technical barrier falling didn't eliminate programming jobs; it created an entire ecosystem that required both more developers and new non-technical roles.

With AI, the job categories emerging are more diverse because AI capabilities are more general-purpose. Legal document review specialists who can train and audit AI systems for accuracy and bias. Financial analysts who can design AI-augmented forecasting workflows that incorporate market signals no human could process manually. Customer service managers who can orchestrate human-AI collaboration patterns that deliver personalised support at scale.

These aren't IT jobs with AI tools bolted on. They're domain-specific jobs that exist because AI capabilities made new approaches to old problems economically viable for the first time. The legal specialist isn't just using AI to work faster - they're exploring forms of legal analysis that were impossible when document review was limited by human reading speed. And yes, we should expect software development jobs to multiply too. When AI coding assistants make programming easier and cheaper, organisations that couldn't afford custom software development suddenly can. It's the Jevons paradox - make something more efficient and you get more demand, not less. Wider roads, more traffic. And with cheaper development we might see organisations swing from buying SaaS to building custom solutions - or conversely an explosion of micro-SaaS for every niche problem that is newly economically viable to solve.

So that's the theory. And it's exactly what we're seeing in the employment data - growth in roles that apply AI capabilities to specific business problems - roles that require understanding both the domain and the technology well enough to bridge between them effectively.

Lightcast's **analysis of 1.3 billion job postings** found that 51% of roles requiring AI skills are now outside IT and computer science, with an 800% surge in generative AI roles across non-tech sectors since 2022. These postings carry a 28% salary premium - the market recognising the value of domain expertise combined with AI capability.

The growth is happening exactly where you'd expect bridge roles to emerge: marketing and PR roles with AI skills up 50% year-over-year, HR up 66%, finance up 40%. **Indeed's data** shows implementation-focused roles rising fast - management consultant jobs mentioning generative AI jumped from 0.2% to 12.4% of all AI postings between January 2024 and January 2025.

Microsoft-LinkedIn's Work Trend Index confirms the demand shift: 66% of leaders prefer hiring candidates with AI skills even if less experienced, and non-technical professionals taking AI courses rose 160%

The job growth in AI-exposed sectors that puzzled the Financial Times isn't anomalous. It's platform economics playing out exactly as Wardley's framework predicts. The question isn't whether AI will create jobs - it's how quickly organisations will recognise the opportunity and adapt their hiring to take advantage of capabilities that were economically impossible just a few years ago.

LOTUS 1-2-3 EATERS

I used to work in the back office of a large investment bank in the City of London at the dawn of personal computers. We had a team whose job was preparing financial reports for fund managers. These were printed out on giant dot matrix printers strewn all over the place, paper emerging in endless perforated streams that had to be torn off and sorted.

If a fund manager wanted to add a field or change something in a report, there was a work request, development cycle, testing phase - basically a complete software development process that could take weeks. Those requests were queued up for the dev team with constant stakeholder battles over prioritisation. The fund managers had infinite appetite for custom analysis, but they were constrained by the cost and complexity of getting anything built.

Then came Lotus 1-2-3 - the dominant spreadsheet before Excel took over - and it changed everything.

Suddenly, the same fund managers who had been rationing their requests for custom reports were building their own sophisticated financial models. They weren't just doing the same analysis faster - they were doing analysis that had never been possible before. Scenario modelling, sensitivity analysis, complex portfolio calculations that would have taken the dev team months to build were now being created over lunch breaks.

The spreadsheet didn't destroy jobs in our department. It exploded them. Within two years, there were more people working on financial analysis than ever before, but they were doing completely different work. Instead of waiting weeks for basic reports, analysts were exploring investment strategies that would have been prohibitively expensive to investigate before. The demand for analysis hadn't been satisfied by faster reports - it had been artificially constrained by the cost of getting them built.

This is how platform technologies work. They don't just make existing work more efficient. They remove constraints that were artificially limiting demand, unleashing appetites that no one knew existed because they had never been economically feasible to satisfy.

The Lotus 1-2-3 pattern is playing out again with AI, but this time the spreadsheet can read, write, and analyse images. When AI tools become accessible enough for domain experts to apply directly - without requiring PhD computer scientists as intermediaries - we get the same demand explosion, but across every knowledge domain simultaneously.

Take legal document review. Traditionally, this meant junior lawyers and paralegals grinding through thousands of documents - expensive, slow work limited by human reading speed. Law firms could only afford thorough review for their biggest cases. Everyone else got the abbreviated version.

Now courts accept AI-assisted review, and the tools do more than just speed up reading. They can synthesise timelines, map communications, and surface patterns across thousands of documents - all at AI-assisted speeds while humans apply judgment for accuracy and defensibility. Suddenly, the kind of deep analysis that was only economically viable for million-dollar cases becomes possible for routine litigation.

Marketing teams are having their own Lotus moment, testing AI that can analyse consumer sentiment across millions of social media posts in real time. Researchers are using AI to synthesise literature across domains no human could read in a lifetime. Financial analysts are building scenario models that incorporate dozens of real-time market signals simultaneously.

None of this is replacing human work - it's enabling human work that couldn't exist before. The bottleneck limiting organisational insight was the cost and complexity of analysis. Remove that bottleneck, and you don't eliminate the need for human judgment - you multiply demand for human judgment applied at higher levels of leverage.

The early job growth in AI-exposed sectors makes perfect sense once you understand this pattern. It's job multiplication following the exact same pattern we saw when spreadsheets made financial analysis accessible to every fund manager in London.

THE 500-YEAR FOOL'S ERRAND

McKinsey has a **27-page solution** for why widespread AI adoption isn't boosting company earnings. They call it the 'gen AI paradox' - apparently unaware that economists have been calling this the 'productivity paradox' since Solow coined the term in 1987. Their diagnosis: firms need CEO-scale rewiring and enterprise transformation programs to capture productivity gains that refuse to materialise.

This is spectacularly wrong advice built on a fundamental misunderstanding of how platform transitions work.

McKinsey treats weak near-term productivity data as proof that firms must scale harder - erecting agentic meshes and rewiring at enterprise scope. That's exactly backwards for a platform transition whose gains arrive through work reorganisation that standard metrics systematically miss. Building a CEO transformation program around short-term productivity numbers isn't rigour - it's a category error with a large price tag.

McKinsey missed the obvious. Solow already explained this paradox in 1987: 'You can see the computer age everywhere but in the productivity statistics'. Despite computers obviously transforming individual work - enabling analysis that would have taken weeks to be completed in hours - aggregate productivity data showed no clear gains for decades. This same 'productivity paradox' persisted through the 1980s and early 1990s, even as anyone using computers could feel their transformative impact.

The reason you can't see AI productivity gains in company P&L statements is the same reason economists struggled for decades to measure computer productivity, and why it took 560 years to quantify the printing press's economic impact. The yardstick moves.

This measurement problem goes back centuries. Recent research by Jeremiah Dittmar found that cities with printing presses in the late 1400s grew 60% faster than cities without them. But this research was published in 2011 - over 560 years after the technology's introduction. Economists had struggled for centuries to find macroeconomic evidence of the printing press's impact, despite its obvious role in enabling the Scientific Revolution, the Reformation, and mass literacy.

The printing press story deserves deeper examination because it's the perfect historical mirror for AI. The parallels are so exact they feel scripted - from the guild protectionism to the productivity paradox to the unintended consequences that shattered existing power structures.

GUTENBERG'S BANKRUPTCY

In 1440, Johannes Gutenberg was just another failed businessman in Mainz, Germany. He'd burned through investor money on schemes involving polished metal mirrors (for religious pilgrims) and secret metallurgy projects. His investors were suing him. He needed cash.

What he built to save himself would accidentally destroy the medieval world order.

Gutenberg's insight wasn't the printing press itself - woodblock printing existed in China since 220 AD, and Europeans had been using wooden blocks for playing cards since the 1300s. His revolution was movable metal type: individual letters cast in metal that could be arranged, printed, disassembled, and reused infinitely.

The first investor pitch must have seemed modest. Instead of a scribe taking a year to copy one Bible, selling for 30 florins (three years' wages for a clerk), Gutenberg could produce 180 identical copies in the same time.

But Johann Fust, Gutenberg's main investor, saw something the inventor missed. When their partnership exploded in 1455 (Fust sued Gutenberg for 2,026 guilders and won the entire printing operation), Fust didn't just take the presses. He immediately partnered with Peter Schöffer, Gutenberg's best employee, and started something new: the world's first publishing house.

Within five years, Fust and Schöffer were rich. Not from productivity gains - from transformation. They weren't selling books; they were selling scalable knowledge distribution. Their 1457 Mainz Psalter wasn't just printed - it included the first publisher's imprint and colophon, establishing the business model that would dominate information distribution for 500 years.

WHEN BOOKS GOT CHEAP

The scribes saw it coming first. In 1471, Giovanni Andrea Bussi, bishop of Aleria, captured their panic in a letter: 'Printers spring up everywhere, and publish whatever strikes their fancy. Many writings appear under false titles, and many mistakes are made in the texts themselves. The dignity of learning suffers'.

He wasn't wrong about the errors. Early printed books contained numerous mistakes - Gutenberg's own Bible had handwritten corrections. Scribes had spent centuries perfecting accuracy through slow, careful copying. Now any fool with a press could spread errors at scale.

The Stationers' Guild in London organised the resistance. They'd controlled book production since 1403, maintaining quality through apprenticeships that took seven years. A master scribe was an artist, a scholar, a guardian of knowledge. Their illuminated manuscripts were individual masterpieces, each one unique.

The guild's arguments against printing were sophisticated:

- **Quality**: Hand-copied texts included scholarly annotations, cross-references, and corrections that printing couldn't replicate
- **Accuracy**: Without trained scribes checking each copy, errors would multiply exponentially
- **Authenticity**: How could readers trust a text when they couldn't trace its provenance through known scribal hands?
- **Sacred knowledge**: Religious texts required reverent handling by educated clerics, not mechanical reproduction by profit-seekers

They had powerful allies. The Catholic Church initially supported the guild's position. In 1479, Pope Sixtus IV granted the University of Cologne the right to ban any books threatening faith or morals. The Cologne censors quickly prohibited printing of vernacular Bibles - sacred texts belonged in Latin, interpreted by priests, not in common tongues for anyone to misread.

But economics overwhelmed ideology. By 1480, a printed book cost one-fifth what a manuscript cost. A scribe needed six months to copy Aquinas's Summa Theologica; a press produced it in a week. When Venice became the printing capital of Europe, it wasn't because they had better theology - they had better banking.

THE NUMBERS LIED

Here's where it gets strange. Despite this obvious revolution in production, contemporary economists couldn't find evidence of economic growth.

The Holy Roman Empire kept detailed tax records. City chroniclers tracked wages, prices, and population. Yet the economic data from 1440-1500 showed... nothing special. No productivity surge. No wage growth. Nothing resembling what we'd now call GDP growth.

Traditional industries actually showed declining productivity. Textile production per worker fell. Agricultural output remained flat. Construction wages stagnated. The printing press was transforming civilisation, but the numbers said otherwise.

The problem was measurement. Economists counted books produced, not ideas spread. They tracked scribal employment (which fell) not total information workers (which exploded). They measured the cost of producing existing goods, not the value of entirely new categories of creation.

Consider what actually happened in those 'stagnant' decades:

- **1450s**: 10 printing shops in Europe
- **1480s**: 110 printing shops
- **1500**: Over 1,000 printing shops employing 20,000 workers directly
- **1455-1500**: 20 million books printed - more than all European scribes had produced in the previous thousand years

But the real transformation was invisible to contemporary measurement:

New Professions Created:

- Type designers and punch cutters (the UI/UX designers of their era)
- Compositors who arranged type (the coders)
- Pressmen who operated the machines (the DevOps)
- Proof readers who checked for errors (the QA)
- Book distributors and sellers (the platforms)
- Paper manufacturers at industrial scale
- Ink makers and chemists
- Bookbinders and finishers
- Literary agents and rights negotiators

New Business Models:

- Subscription publishing (readers paid in advance for books not yet printed)
- Serialised releases (spreading costs and building audiences)
- International publishing agreements (Venice and Frankfurt book fairs)
- Advertising supported publications (first appearing in news sheets)
- Commissioned translations and adaptations

The Venetian publisher Aldus Manutius invented the modern book business between 1494-1515. He created:

- Italic type (to fit more words per page, reducing costs)
- The octavo format (portable books, like paperbacks)
- Semicolons and modern punctuation (for clarity at speed)
- Publisher's brand identity (the Aldine dolphin and anchor)
- The classic book series (uniform editions of Greek and Latin texts)

None of this showed up in productivity statistics. And even GDP would have captured book sales, not the reorganisation of human knowledge.

YOU CAN'T BAN A BESTSELLER

The Catholic Church learned too late that controlling the means of production didn't mean controlling the message.
On October 31, 1517, Martin Luther posted his 95 Theses on the Castle Church door in Wittenberg. In the manuscript era, this would have remained a local academic dispute. Luther would have hand-copied a few dozen copies for fellow theologians. The debate would have proceeded through formal channels, in Latin, among elites.

Instead, Luther's students took the theses to print shops. Within two weeks, copies appeared throughout Germany. Within a month, they'd reached France, England, and Italy. Translated into German, illustrated with woodcuts, packaged as pamphlets selling for a few pennies.

Luther's publication numbers exploded (remembering that each 'edition' meant a print run of 1,000 or more copies):

- **1518**: 150 editions of Luther's works printed
- **1519**: 390 editions
- **1520**: 570 editions
- **1521-1525**: Over 2,000 editions

Between 1517 and 1520, Luther's thirty publications sold over 300,000 copies. His German translation of the New Testament (1522) sold 5,000 copies in two weeks - at a time when a bestselling book might sell 1,000 copies over several years.

The Church tried to respond. In 1520, Pope Leo X issued Exsurge Domine, a formal papal decree condemning Luther's propositions. It took six months to distribute through official channels. By then, Luther had published his response, burned the papal decree publicly, and the footage - sorry, the woodcut illustrations - circulated faster than the original condemnation.
This wasn't just about speed. It was about network effects. Every printer who published Luther made money. Every town that bought his pamphlets wanted more. A distributed network of economic incentives spread ideas faster than any central authority could counter them.

The geographic spread was equally dramatic:

- **1500**: 252 towns across Europe had presses, only 62 in Germany
- **1517-1520**: Number of German towns with presses doubles to 125
- **By 1550**: Over 3,000 printing establishments across Europe

The Church's response revealed their fundamental misunderstanding. The Index Librorum Prohibitorum (List of Prohibited Books) first published in 1559, attempted to ban dangerous texts. But banning books only made them more valuable. Printers in Protestant regions got rich selling forbidden books to Catholic territories - the Church had accidentally created a black market more profitable than the legitimate one. The first-mover advantage went to whoever embraced the new technology, not who regulated it.

HOW GUTENBERG INVENTED CHILDHOOD

Nobody predicted what printing would actually create. Gutenberg thought he was mechanising book production. He accidentally enabled:

Standardised Language: Before printing, every region had its own spelling, grammar, and vocabulary. Printers needed to pick one version to maximise their market. The dialects they chose became national languages. Modern German exists because Luther's Bible translation was the killer app that made Saxon German the standard. The King James Bible did the same for English.

Created Science: Medieval 'natural philosophy' was a mess of personal observations, theological arguments, and ancient authorities. Printing enabled exact reproduction of diagrams, tables, and mathematical proofs. When Copernicus published De revolutionibus orbium coelestium in 1543, astronomers across Europe could work from identical star charts. The scientific revolution required reproducible data.

Invented Childhood: Before printing, children were small adults. The explosion of printed educational materials created age-graded learning. The hornbook (1450s), primers (1500s), and textbooks (1600s) established childhood as a distinct phase requiring special materials. Mass literacy transformed society's fundamental age structure.

Enabled Democracy: The American Revolution succeeded partly because pamphlets like Common Sense (1776) could reach 500,000 readers in a population of 2.5 million. The French Revolution's Declaration of the Rights of Man spread through thousands of prints. Democratic ideals required mass distribution to create mass movements.

Destroyed Privacy: The first newspapers (1605 in Strasbourg) invented public scrutiny. Private letters of public figures became publishable. Court proceedings moved from secret chambers to printed records. The distinction between public and private life - obvious in oral culture - dissolved in print culture.

The measurement problem persisted for centuries. In 1837, almost 400 years after Gutenberg, economists still couldn't quantify the printing press's economic impact.

The growth wasn't in making existing things faster but in making impossible things possible.

Productivity measurements break down during technological transitions because transformational technologies don't just make existing work more efficient - they reorganise everything around new capabilities.
In my old banking job, we reconciled trading positions using pen, paper, and desktop calculators. The investment bank had mainframes, but office PCs didn't exist yet. When I try to imagine doing today's work with pen and paper - how long would current tasks take? Can I do a finger in the air estimate of the productivity gain I personally get from a computer? Do I even allow a desktop calculator into this thought experiment, or do I have to excavate long division? - the question becomes impossible to answer. All our work is now moulded around computer capabilities. For most tasks, there is simply no pen-and-paper analog to be measured.

By the time economists figure out how to measure the gains, everyone's doing completely different jobs. The yardstick moves. Firms spend money on process redesign that accounting treats as expenses. They look less productive while reorganising, then gains appear after everyone's forgotten the old way - economists call this the J-curve.

WHEN ELECTRICITY CHANGED NOTHING

Electricity makes the AI parallels even clearer. Like AI today, electricity was simultaneously overhyped and underestimated, creating fortunes for those who understood the difference.

In 1879, Thomas Edison demonstrated his incandescent bulb. The New York Times declared it would 'revolutionize civilization'. The stock market agreed - Edison General Electric shares soared on promises of electric everything. By 1882, Edison's Pearl Street Station powered 85 customers in lower Manhattan. The electric age had begun.

Or had it?

THE WAR OF THE CURRENTS

What followed wasn't adoption but warfare. Edison backed direct current (DC) - safe, proven, limited. His system required a power station every mile because DC couldn't travel far without degrading. George Westinghouse backed Nikola Tesla's alternating current (AC) - dangerous, unproven, unlimited. AC could travel hundreds of miles, enabling centralised power generation.

Edison fought dirty. Through secretly-funded proxy Harold Brown, Edison staged public demonstrations where dogs and cats were electrocuted with Westinghouse's AC current. Brown would travel from town to town in 1888, killing animals before audiences to prove AC's danger. At one demonstration in Columbia College, he electrocuted a large dog with 1,000 volts of AC after showing it could survive lower DC voltages.

Edison pushed for the electric chair to use AC current, hoping to associate his competitor's system with death. When New York State adopted electrocution as its execution method, Edison ensured Westinghouse generators powered it - even arranging for them to be purchased through third parties when Westinghouse refused to sell them for executions.

The first execution in 1890 was a disaster. William Kemmler took eight minutes to die. The current had to be applied twice. Witnesses reported the smell of burning flesh. Westinghouse said they would have done better with an axe. Edison called it proof that AC was too dangerous for homes.

The press began using 'Westinghoused' as slang for electrocution. Edison suggested criminals be sentenced to 'electricide' or 'condemned to the Westinghouse'. He was trying to brand his competitor's name as synonymous with death itself.

The technical arguments masked the real fight: business models. Edison's DC meant distributed generation - every neighbourhood, factory, and wealthy home would need its own power plant. Edison would sell them all. Westinghouse's AC meant centralised generation - a few massive plants serving entire regions. Utilities would own the infrastructure.

Edison lost. Not because AC was 'better' - it was more dangerous and harder to control. He lost because AC's economics enabled an entirely new business model: the electric utility. By 1892, J.P. Morgan forced Edison out of Edison General Electric, renamed it General Electric, and pivoted to AC.

The inventor lost control to the platform. Just as Fust seized Gutenberg's printing operation and created the publishing business, Morgan seized Edison's electrical empire and created General Electric.

THE WRONG-SHAPED FACTORIES

By 1900, electricity was everywhere and nowhere. Major cities had electric streetlights, trolleys, and telegraphs. Department stores installed electric elevators. The Paris Exposition featured an electric moving sidewalk. Yet factory productivity remained flat.

The numbers were damning:

- **1880**: 0% of US manufacturing power from electricity
- **1900**: 5% electric (mostly lighting, not motors)
- **1910**: Still only 25% electric
- **1920**: 50% electric
- **1930**: 80% electric

Fifty years for near-complete adoption. Productivity gains didn't appear until the 1920s - forty years after Edison's bulb.
Warren Devine's research in 1983 finally explained why. Early factories simply replaced steam engines with electric motors - same layout, same workflow, electric instead of steam. The real transformation required completely rebuilding factories around electricity's capabilities.

Steam-powered factories were vertical - multi-story buildings built around a massive steam engine in the basement. Power transmitted through elaborate systems of belts, pulleys, and drive shafts. Every machine had to be near the main shaft. The entire factory shut down if the steam engine needed maintenance.
Electric factories could be horizontal - single-story buildings with production lines. Each machine had its own motor. Workflows could be optimised for production, not power transmission. Maintenance could be targeted. But rebuilding factories took decades and massive capital investment.

The productivity gains came not from electricity itself but from the reorganisation it enabled. By the time economists could measure the gains, nobody remembered steam-powered factories.

THE 20 HOURS A WEEK ELECTRICITY GAVE WOMEN

While economists puzzled over productivity, electricity transformed society in ways they couldn't measure:

Consumer Credit: Before electricity, only the wealthy bought durable goods. Electric appliances were expensive but transformative. Stores from London's Harrods to New York's Macy's invented instalment to sell them. In the US, consumer credit exploded from $3 billion in 1920 to $8 billion by 1929. The modern consumer economy was born.

Women's Liberation: In 1900, washing clothes took a full day - hauling water, heating it, scrubbing, wringing, drying. Electric washing machines cut this to hours. Refrigerators eliminated daily shopping. Ruth Schwartz Cowan's research shows electric appliances freed 20+ hours per week. From Manchester's textile workers to Detroit's factory workers, women entered the workforce in unprecedented numbers.

Mass Entertainment: Radio created the first simultaneous national experience. By 1930, 40% of American homes had radios. FDR's fireside chats reached 60 million people simultaneously. The BBC's first broadcasts reached British homes in 1922. Berlin's UFA studios pioneered sound film alongside Hollywood. Mass culture replaced local culture almost overnight.

Urban Transformation: Paris's Eiffel Tower showcased electric lighting in 1889. London's first electric escalator appeared in Harrods in 1898. Electric elevators made skyscrapers possible and by 1913 New York's Woolworth Building reached 57 stories. Cities grew up instead of out. Electric streetcars enabled suburbs. The modern metropolis emerged.

Time Itself Changed: Before electricity, human activity followed the sun. Electric light divorced work from daylight. The 'night shift' was invented. Cities became 24-hour operations. Even sleep patterns changed - average sleep dropped from 9 to 7 hours as electric light extended usable time.

None of this showed up in contemporary productivity measurements.

HOW SAMUEL INSULL TURNED ELECTRICITY INTO MONEY

The real electricity fortunes weren't made by Edison or Tesla. They were made by Samuel Insull, who built the utility model.

Insull realised electricity was a platform business:

- High fixed costs (power plants, transmission lines)
- Near-zero marginal costs (coal was cheap)
- Network effects (more customers = lower unit costs)
- Natural monopoly dynamics

He pioneered:

- Tiered pricing (different rates for different uses)
- Time-of-use pricing (cheaper at night)
- Long-term contracts with large customers
- Regulatory capture (embracing regulation to lock out competitors)

By 1929, Insull's companies served 4 million customers across 32 states. He was worth $150 million (roughly $3 billion today). The inventors created - the platform owner captured.

The crash of 1929 destroyed Insull - his holding company pyramid collapsed. But the model survived. Today's electric utilities still follow Insull's playbook.

PAUL DAVID'S DISCOVERY ABOUT TECHNOLOGY AND TIME

Paul David's famous 1989 paper *The Dynamo and the Computer* documented the electric productivity paradox in detail. His conclusion: general-purpose technologies require complementary innovations that take decades to develop.

Electricity needed:

- Factory redesign
- Management restructuring
- Worker retraining
- Financial innovations (consumer credit)
- Regulatory frameworks
- Social adaptations (night work, suburban living)

The productivity gains appeared only after all these were in place. The 40-year lag wasn't failure - it was the time required for systemic transformation.

David predicted computers would follow the same pattern. He was right. The PC revolution started in the 1970s, but productivity gains didn't appear until the late 1990s - after the internet, enterprise software, and business process reengineering.

AI is following the same pattern. The technology exists. The supporting ecosystem is in the early stages of being built. The lag frustrates economists who can't see transformation without productivity statistics. But transformation doesn't wait for measurement.

McKinsey's approach shows exactly how productivity obsession leads to terrible advice. Their big claims - 'estimated impact... 60-90 percent' and '80 percent of level 1 incidents resolved automatically' - are fictional scenarios, not real results. They're selling CEOs certainty about gains that won't be measurable for decades, telling them to rewire everything based on projections while admitting that 'recurring costs can exceed the initial build investment'.
McKinsey's own language betrays their misunderstanding: 'Gen AI is everywhere - except in company P&L'. They're selling expensive solutions to a measurement problem that's existed for 500 years. The transformation they're looking for is already happening.

They use accounting designed for widget factories to measure platform transitions. Of course they can't find the value - they're looking in the wrong column.

This isn't a productivity paradox - it's the predictable measurement confusion that occurs when the entire basis of measurement shifts.

WHY AI LAYOFFS HAND MARKET SHARE TO COMPETITORS

The layoffs are real and brutal. Over 300,000 tech workers lost their jobs in 2023 and 2024 - numbers approaching dot-com crash levels. Every headline screams about AI automation eliminating positions. Yet the data shows AI-exposed jobs growing, not shrinking.

So what's going on?

We're watching two different things happen at once: companies cutting specific jobs while the economy creates new ones. This isn't contradictory - it's reorganisation.

Companies aren't swapping humans for software. They're rebuilding how work gets done while figuring out which roles actually matter. When Dropbox talks about 'repositioning for the AI era' or IBM pauses hiring for back-office work, they're making bets about what their organisations will look like in two years.

> *those who make half a revolution only dig their own graves*

The maths is simple but gets buried in the drama. If AI doubles your team's output and you can sell the extra output, firing people means handing market share to competitors who keep their teams and scale up. As the old saying goes - 'those who make half a revolution only dig their own graves'. Cuts only make sense when demand is capped - thin sales pipelines, maxed distribution, regulated pricing, or business models that are fundamentally changing.

That's why responses vary so wildly. Some companies face genuine demand constraints where AI productivity means fewer people. Others are in growth mode where AI lets them tackle workloads that were impossible before - with bigger teams, not smaller ones.

Salesforce just demonstrated this perfectly. They cut 4,000 support jobs - from 9,000 to 5,000 - claiming AI handles half their customer chats. But their own research shows these AI agents achieve only 58% success rate on single-step CRM tasks. They're betting market share on agents that fail nearly half the time, while competitors who keep human support staff can offer actual reliability. Worse, they're selling these same failing agents to customers, strip-mining entire industries of their payroll. Money that paid support workers' salaries now flows to Salesforce as subscription fees. That's not transformation - that's value extraction disguised as innovation. And whether through cynicism or shortsightedness, they're literally selling customers a way to hand market share to their competitors.

Between McKinsey selling fictional productivity gains and Salesforce selling failing automation, we're witnessing a strategic stupidity arms race. McKinsey charges millions to misread platform transitions. Salesforce charges subscriptions for 58% success rates. Both are packaging their fundamental misunderstandings of AI as expensive expertise.

As far as headlines go, the cuts are visible and dramatic. Mass layoffs get press releases. The new hiring is scattered across thousands of companies and doesn't get labeled 'AI jobs' in the statistics. It just looks like normal hiring, except the job descriptions now mention AI skills.

Meanwhile, companies are sitting on record cash reserves - roughly $7.55 trillion among US non-financial corporations. They're keeping their powder dry during a transition that even the supposed experts don't understand. When McKinsey can't measure value and Salesforce is selling competitive suicide, perhaps hoarding cash while you figure out what's actually happening isn't irrational.

The stabilisation is already visible in sectors that moved first. Law firms are building legal-tech specialist teams. Hospitals are launching ambient-AI and clinical AI programs. Banks are hiring gen-AI enablement leaders for their revenue businesses. These aren't one-for-one replacements - they're entirely new job categories.

This 'Great Reorganisation' looks chaotic and painful - and of course it is for people losing their jobs especially. But the underlying pattern points toward job multiplication, not elimination. The transition is messy, but the economics favour human-AI collaboration over pure automation.

THREE KINDS OF SURVIVORS

The employment data tells only half the story. Yes, AI-exposed jobs are growing rather than shrinking. But the growth isn't uniform - it's creating a fundamental split in how knowledge work gets organised.

Three distinct categories are emerging from this reorganisation:

Navigation Experts have deep understanding of what constitutes good versus bad output in their field. I've called these '**quality maps**' - the mental models that let you spot when something looks plausible but is actually wrong. A senior legal researcher who can tell the difference between a compelling but legally flawed argument and a mundane but precedent-solid one. A financial analyst who knows when algorithmic trading signals deserve investigation versus dismissal.

These professionals become more valuable, not less, because their judgment scales across AI-generated possibilities they could never explore manually. While the AI can generate possibilities all day, it can't judge quality. Someone has to steer.

Amplification Natives grew up inside AI-augmented workflows and developed entirely new approaches to knowledge work. They're comfortable with iterative refinement and prompt wrangling. The junior developers who were earliest adopters of 'vibe-driven development' and spec-first coding, forming communities around these AI-native approaches. Content creators who think in terms of AI-assisted ideation and human-guided curation.
Unlike traditional workers adapting AI tools to existing processes, they've designed their cognitive habits around AI capabilities from the start. These workers execute faster than

traditional experts because they've internalised AI-native workflows rather than bolting AI onto legacy approaches.

Integrators and Governors represent the operational layer that makes AI reliable at scale. They handle data plumbing, evaluation systems, monitoring, compliance, and change management - the unglamorous infrastructure that turns impressive demos into dependable business processes. The accountants who become AI audit specialists. The project managers who evolve into AI workflow coordinators. The quality assurance professionals who specialise in human-AI collaboration protocols.

This is where many traditional mid-tier workers can land if they reskill effectively.

THIS AMP GOES PAST 11

The displacement isn't vertical - AI doesn't simply replace 'lower skill' with 'higher skill' work. It's lateral. Within every domain it touches, workers who can navigate AI capabilities effectively are outcompeting workers who cannot, regardless of qualifications or experience levels. A junior analyst who's learned to pair AI research with domain-specific judgment can outperform a senior analyst who treats AI as an enhanced search engine.

But this isn't elevation to some utopian 'strategic work'. It's amplification to maximum intensity. AI-assistance doesn't make work easier - it turns the volume up past eleven. Everyone's producing more, iterating faster, juggling more parallel tasks - pushing harder against new limits instead of old ones. Again, wider roads, more cars, but at a personal level.

This explains the data. Yes, there are brutal layoffs happening - but they're not because of AI. Post-COVID corrections, funding dry-ups, companies figuring out what they actually need. Meanwhile, AI-exposed sectors are quietly growing because AI makes analysis, content creation, and research cheap enough that organisations can afford way more of it. Research from Revelio Labs confirms this pattern - the most AI-exposed jobs are adopting AI most aggressively, with up to 30% of workers in these roles already using AI daily. Rather than being replaced, they're becoming more productive. The jobs supposedly most at risk are actually the ones growing fastest.

CONCLUSION

> *Well! I've often seen a cat without a grin, but a grin without a cat! It's the most curious thing I ever saw in all my life! - Alice*

We're having a furious debate about something that's already disappearing. The 'AI jobs crisis' is like the Cheshire Cat - by the time we finish arguing whether it exists, only the grin will remain.

Economists scramble to measure unmeasurable gains. Policymakers prepare for mass unemployment that isn't coming. Consultants sell transformation programs for changes that are happening without them.

AI is commoditising into infrastructure like every platform before it. The productivity measurement is cargo cult economics - pretending to quantify the unquantifiable.

What started as one puzzling chart turns out to be overwhelming evidence. The job market is restructuring around AI amplification, creating new categories of work faster than old ones disappear.

The brutal layoffs are real, but they reflect competitive reorganisation within the platform transition, not mass replacement by autonomous systems.

By the time this restructuring becomes fully visible in employment statistics, AI-augmented collaboration will be so normal, so invisible, that our children won't understand what we were so worried about. They won't debate whether 'computers create jobs'. They'll pilot **cognitive exoskeletons** the same way we use spreadsheets: as unremarkable infrastructure for getting work done.

The transformation isn't coming. It's here. Invisible to economists staring at the wrong numbers, ordinary to workers already living it.

* * *

We end where we began - with the question of investment. From the start, the shadow hanging over these essays has been whether this is all just a bubble. The next essay completes the picture - showing how the spending spree can make sense once you see the whole board.

We’ve traced the limits and uses of the technology: how it isn’t intelligent, why it works anyway, why it can’t plan or get things right on the first try, what it feels like to work with, how it lets platforms seize control, and where the jobs emerge.

Big Jobs explored the labour side of automation. This next essay flips the ledger. If labour was the first half of the story, capital is the second. *The Disappearing Salary* goes beyond jobs lost or created to examine how investment itself is re-routed, how capital pools shift, and how the same forces that re-order work also re-order wealth.

* * *

THE DISAPPEARING SALARY

SPOILER: PAYROLL IS THE PRIZE

SPOILER ALERT: In *The Usual Suspects*, Verbal Kint didn't just hide his identity - he inflated Keyser Söze into such a mythical figure that nobody could believe the limping con artist was the monster. He spun tales of supernatural ruthlessness, of a crime lord so legendary he might not even exist. The myth was the misdirection. The cops spent the whole film hunting a phantom while the real Söze limped out of the police station. They couldn't see him because they were looking for someone who matched the legend, not the mundane reality sitting right in front of them.

The investment flooding into artificial intelligence works the same way. The AGI sci-fi fantasy is so oversized, so messianic - consciousness emerging from silicon, the singularity, humanity's next evolutionary leap - that nobody sees the mundane reality: investors are unconsciously converging on the largest expense pool in capitalism. They're not building minds. They're building meters. But the fantasy of AGI is so intoxicating that it obscures what's actually being funded.

This isn't conspiracy. Investors genuinely believe they're two years from artificial general intelligence. Founders truly think they're birthing digital consciousness. Engineers pulling all-nighters are convinced they're writing history. And that genuine belief in the myth makes the real transformation invisible.

The stakes here are $40 trillion rather than a boat full of cocaine. The misdirection works because everyone involved believes their own story.

THE BIGGEST EXPENSE LINE IN CAPITALISM

Walk into any boardroom in London, Frankfurt, or New York. Skip past elaborate presentations about digital transformation and synergistic value creation. Ask the CFO one simple question: what's your biggest expense line? It's not the gleaming real estate in Canary Wharf, not the eye-watering AWS bills that seem to grow every month, not the McKinsey consultants who somehow always find more things to optimise. It's salaries. The money flowing from corporate accounts to human bank accounts in exchange for showing up and doing things.

In developed economies businesses pay out $35-40 trillion annually in wages. This represents 50-70% of all business costs and roughly 60% of GDP. Three-fifths of all economic activity, denominated in direct deposits and paychecks.

China runs its own parallel game with state-directed AI development and captive platforms that operate under entirely different rules. Alibaba and Baidu are building their own tollbooths while navigating Beijing's ever-shifting regulatory maze. The dynamics there deserve their own analysis. For this essay, we're focusing on developed economies where these platforms primarily operate - markets with their own constraints like GDPR, labour laws, and emerging AI regulations, but where Silicon Valley companies can still build and extract value at scale. The developed economy payroll pool is what OpenAI, Microsoft, Google, and Anthropic can actually access.

Now observe what technology currently captures from this ocean of wages. The entire global SaaS market - Salesforce with its bloated CRM that every sales team claims to hate but can't live without, Microsoft 365 that forces us all into Teams meetings, Slack with its infinite channels of procrastination, every enterprise subscription from the mission-critical to the completely forgotten - generates maybe $400 billion annually. That's less than 1% of what companies spend on humans. Put another way, for every $100 companies spend on salaries, they spend less than $1 on the software that's supposedly revolutionising how those humans work.

ServiceNow, the enterprise automation platform that most people have never heard of despite its massive valuation, built a $150 billion market cap on just $10 billion in annual revenue. They captured 0.025% of developed economy payroll - two and a half hundredths of one percent - and created one of the world's most valuable companies. They built an empire on crumbs. If two and a half basis points of human wages can create that kind of value, what happens when someone captures two percent? Or five percent?

Meanwhile, public and private investment has achieved something approaching religious conversion. Over $100 billion flowed into AI companies in 2024 alone from venture capital, sovereign wealth funds, and corporate investors. By the first quarter of 2025, AI captured 71% of all venture funding. Not 71% of software funding or 71% of deep tech funding, but 71% of all venture capital deployed anywhere. During the peak insanity of the dot-com bubble, internet companies managed to capture 40% of venture investment. During crypto's most hallucinogenic phase, blockchain barely touched 20%. This isn't portfolio allocation anymore - it's monotheism.

Companies are running live experiments in workforce transformation, and the results are more educational than their press releases suggest. Klarna, the Swedish buy-now-pay-later company that helps consumers go into debt more efficiently, fired 700 customer service workers in 2022, replacing them with OpenAI-powered chatbots. The CEO, Sebastian Siemiatkowski, proclaimed victory across every platform that would listen. Eighteen months later, they're quietly rehiring humans while keeping the chatbots, creating a special kind of operational purgatory where angry customers first argue with a bot that doesn't understand them, then wait to argue with a human about why the bot was useless. Customer satisfaction didn't just decline - it cratered so severely they were haemorrhaging users to competitors who maintained actual human service. Now they pay for both systems, achieving neither the efficiency of automation nor the quality of human service. The promised revolution became an expensive add-on.

Salesforce has its own version of this disaster. They cut their support staff from 9,000 to 5,000, with Marc Benioff proclaiming that AI agents now handle half their customer interactions. The detail that doesn't make it into the earnings calls or TechCrunch interviews is that those agents achieve 58% success rates on single-step tasks. Not complex multi-part problems, not nuanced situations requiring judgment - single-step tasks like 'reset my password' or 'update my billing address'. Nearly half the time, a human has to intervene, apologise for the bot's confusion, and actually solve the problem. The remaining 5,000 humans now manage both angry customers and constant AI failures, doing the work of the original 9,000 while also serving as janitors for broken automation. Worse, they're selling these same failing agents to customers, converting payroll across entire industries into Salesforce subscription revenue.

Imagine your heart surgeon working at those odds. 'Good news - the operation was 58% successful!' Imagine your accountant getting your taxes 58% right, leaving you to wonder whether the 42% wrong parts will trigger an audit or a refund. Your pilot landing 58% of their flights safely - the rest ending in what the aviation industry might euphemistically call 'unplanned ground interactions'. Yet companies are betting their customer relationships - the entire interface between their business and their revenue - on systems that fail nearly half the time. It's not automation; it's Russian roulette with customer satisfaction.

WHY A $560 BILLION 'BUBBLE' MIGHT BE RATIONAL

To understand why seemingly rational investors are pouring half a trillion dollars into technology that barely works, you need to understand how venture capitalists think about markets. They don't care about current revenue - that's for pension funds and index investors. They care about Total Addressable Market, or TAM, the theoretical maximum revenue if you captured 100% of your target market. Nobody ever captures 100% of anything, but TAM tells you whether you're fishing in a pond or the ocean.

When Uber was starting, their TAM wasn't *taxi rides in San Francisco* or even *all taxi rides globally*. Their TAM was *all urban transportation* - every trip by car, bus, train, or foot that could theoretically be replaced by an Uber. When Amazon began, their TAM wasn't *books purchased online* but *all retail commerce*. The bigger the TAM, the bigger the valuation multiple investors will swallow. A company capturing 1% of a trillion-dollar market is more valuable than a company capturing 50% of a billion-dollar market.

Here's what nobody's articulating explicitly, though the money is screaming it: if developed economy payroll is the TAM, the AI investment suddenly makes sense.

Every lawyer billing hours in a City firm, every engineer debugging code in Berlin, every analyst building Excel models in New York, every middle manager forwarding emails in Toronto - it all adds up to the largest pool of capturable value in economic history. Larger than the global oil market that nations go to war over. And unlike oil fields that deplete, this $40 trillion regenerates every single year.

Morgan Stanley recently **published something remarkable**. They calculated that AI could deliver $920 billion in annual benefits to S&P 500 companies. Deep in their analysis lies a detail they mention but don't emphasise: this represents 41% of total S&P 500 compensation expense.

They literally calculated AI value as a percentage of payroll. They discovered the pattern, did the maths, published the number, yet **somehow don't frame it that way**. They can't see Söze for who he is. The data says payroll capture; the headline says transformation.

Sam Altman asking for $7 trillion to build AGI sounds completely delusional. It's the kind of number that makes you wonder if he's having some kind of episode or just trying to anchor expectations so high that a mere $100 billion seems reasonable. But frame it differently: what if that $7 trillion is unconsciously benchmarked against the $40 trillion annual payroll pool? What if the number that sounds deranged for building consciousness makes perfect sense for capturing labour value? Even capturing 2% of that payroll pool means $800 billion in annual revenue. Not one-time sales, not total lifetime value, but annual recurring revenue - the holy grail of software economics. At typical software valuations of 5-10x revenue, you're looking at $4-8 trillion in market value creation. Suddenly the $560 billion invested so far doesn't look crazy - it looks conservative.

Yet nobody frames it this way explicitly. No pitch deck states *our TAM is the $40 trillion companies pay humans*. No Sequoia partner stands up at a conference and announces *we're targeting global payroll for platform capture*. No founder goes on TechCrunch and says *we're building tools to convert salaries into subscription revenue*. They speak instead of automation, transformation, democratisation, augmentation, the future of work - abstract narratives that obscure the concrete mechanism. The money knows what it's chasing, but the mouths can't say it.

OpenAI's robotics investments reveal the scope. They're funding Figure (humanoid robots for warehouses), 1X Technologies (service robots), Physical Intelligence (foundation models for physical tasks). Not just knowledge workers - they want everyone. Blue collar, white collar, no collar. The entire $40 trillion becomes addressable.

Once you see payroll as the prize, the pattern becomes undeniable. The entire SaaS industry built trillion-dollar valuations on capturing less than 1% of global payroll. Morgan

Stanley calculates AI value as a percentage of compensation but can't say it explicitly. We're not witnessing irrational exuberance but unconscious convergence on the largest arbitrage opportunity in history.

OUTSOURCING MOVED JOBS.
AI LIQUIDATES THEM.

To understand what AI promises - or threatens, depending on where you sit - you need to understand how it differs from every other labour arbitrage that came before.

Outsourcing is geographic arbitrage within the existing paradigm. You take a London software developer making £85,000 annually and replace them with three developers in Bangalore making £15,000 each. The work remains fundamentally unchanged - the same code gets written, the same bugs get fixed, the same features get shipped. It just happens in a different time zone at a different price point. You've reduced costs but preserved the basic structure: humans doing human work for human wages. The constraint remains human availability and human skill. You can only hire so many developers in Bangalore before wages rise there too.

This is the equivalent of building cheaper scribe factories. Before the printing press, if you wanted to reduce the cost of book copying, you found cheaper scribes. Move your scriptorium from Paris where the monks demand wine to Prague where they'll work for beer. Train scribes faster. Develop more efficient writing styles - Gothic script that was quicker to write than Carolingian minuscule. Optimise the copying process. But you're still in the scribe business - humans manually copying texts, one letter at a time, their backs bent over desks, developing arthritis and going slowly blind from the work. The fundamental constraint is human: how many literate people exist, how fast they can write, how many hours before exhaustion makes every third word illegible.

AI promises something categorically different: not cheaper scribes but the printing press itself. Not geographic arbitrage but transformation.

The modern parallel that makes this transformation visceral is containerisation. In 1956, Malcolm McLean was a trucking company owner from North Carolina, frustrated by how long it took to load and unload cargo at ports. Watching stevedores slowly move boxes from truck to ship, one crate at a time, he had a thought that seems obvious only in retrospect: why not just put the whole truck trailer on the ship?

The shipping industry thought he was insane. Ships were masterpieces of cargo optimisation, with longshoremen who knew how to pack every cubic foot efficiently. Wasting space on metal boxes seemed like the fantasy of someone who didn't understand maritime economics. So McLean bought a shipping company - Pan-Atlantic Steamship - not because he understood shipping, but because he needed ships to test his crazy idea. The first container ship was a converted World War II tanker called the Ideal X. It looked ridiculous, like someone had welded truck trailers to a boat.

But those standardised metal boxes didn't just reduce shipping costs by 90%. They made geography irrelevant for manufacturing. Once you could stuff anything in a standard box and ship it anywhere for pennies, the entire logic of production inverted. Factories didn't need to be near consumers anymore. They could be wherever labour was cheapest. Manufacturing fled to Asia not through conspiracy or grand planning but through emergence - millions of independent decisions, each company optimising locally, nobody seeing the systemic pattern until Detroit was already hollowed out and Shenzhen had become the world's factory floor.

AI subscriptions are containers for cognitive work. Once you can meter any knowledge task through an API, once every cognitive process has a price per token, once every business function can be called as a service, the entire employment landscape reshapes itself. The customer service representative doesn't relocate to Manila - they become an API call to GPT. The junior analyst doesn't shift to Mumbai - they become a Python script. The paralegal doesn't move to Poland - they become a document processing workflow. The salary doesn't move geographically; it transforms into something else entirely.

The pricing strategies make this transformation embarrassingly explicit. 'Digital employees' and 'AI workers' have become actual pricing models. ServiceNow's CEO Bill McDermott **doesn't even pretend to hide it**: 'AI agents work 24/7, don't need healthcare benefits, don't take vacation'. That's not describing software features. That's describing salary replacement with a discount. The entire industry has given up on euphemisms and just prices their products in units of human displacement.

AUTOMATION DOESN'T ERASE SALARIES.
IT ADDS SUBSCRIPTIONS.

But here's the problem: the salary→subscription model assumes AI can work autonomously.

As established earlier, these systems **have no intelligence** and **can't plan beyond the next token** - meaning they require constant human supervision. They're cognitive exoskeletons, not independent agents.

AI is being sold as a printing press for productivity - radical transformation through mechanisation. But the printing press didn't eliminate scribes overnight. They became proofreaders, typesetters, editors. The constraint changed from human copying speed to mechanical capacity, but humans remained essential. AI promises similar transformation - fundamental restructuring rather than elimination.

Recent evaluations confirm this systematically. Leading models show consistent 30% error rates on real-world business tasks - not adversarial examples designed to break them, just normal operations like 'summarise this report' or 'extract these invoice details'. This breaks the LLM-as-judge pattern that's supposed to solve autonomy. You can't have AI evaluating AI when both fail at similar rates. Prompt injection remains completely unsolved after two years of intensive research. An LLM evaluating resumes can't distinguish between your instructions and instructions cleverly embedded in the resume itself. It's like hiring a security guard who follows anyone's orders as long as they're written authoritatively. Humans remain essential.

Job multiplication presents another barrier, **as detailed earlier**. When technology makes work cheaper, companies pursue previously uneconomical opportunities. The promised reduction never materialises.

If autonomous agents ever arrive (and despite the hype, they haven't), the story changes - but every live system today sits much closer to scripted workflow than to self-direction, and capital has to price the world that exists, not the one in the deck.

THE PRODUCTIVITY TAX NOBODY ESCAPES

Yet the impossibility of full automation doesn't kill the investment thesis. It transforms it into something more sustainable: permanent tolls on human productivity.

Think about how certain technologies become invisible taxes on modern life. Mobile phone subscriptions started as luxury conveniences for investment bankers and drug dealers in the 1980s. Now they're existential requirements. Nobody questions paying £20-30 monthly for service that was once optional. Try functioning in modern society without a mobile phone. Try getting a job, maintaining relationships, accessing services. The subscription became mandatory without anyone decreeing it should be.

Or consider Visa's business model - perhaps the most perfect tollbooth ever created. They take 2-3% of every transaction, everywhere, forever. They provide genuine value - fraud prevention, dispute resolution, global interoperability, the infrastructure that makes modern commerce possible. But once embedded in the global financial system, dismantling their position becomes essentially impossible. Every attempt to bypass them - Bitcoin, central bank digital currencies, alternative payment networks - just adds complexity while Visa continues collecting their percentage.

History offers a darker parallel. Pre-revolutionary France outsourced its entire tax system to private financiers called fermiers généraux. These tax farmers paid the crown a fixed sum upfront, then collected whatever they could extract from the population, keeping the difference as profit. No conspiracy, no grand plan - just the state finding it easier to privatise revenue collection than build its own apparatus. The system was universally despised, economically destructive, and lasted two centuries until the Revolution introduced tax farmers to Dr. Guillotin's innovation. But for those two centuries, a handful of private actors extracted rent from every economic transaction in France. The system persisted not because it worked well but because it worked well enough for those who mattered.

AI platforms are building the same position across all labour. Every knowledge worker needs language models for research and writing. Every developer needs coding assistants. Every warehouse worker gets 'assisted' by pick-optimisation systems. Every delivery driver gets route 'enhancement'. Every retail worker gets customer interaction 'support'. Every mechanic gets diagnostic 'augmentation'. The white-collar workers get language models, the blue-collar workers get computer vision and robotics, everyone gets metered. The subscription doesn't replace the salary - it attaches to it. They're strip-mining entire industries - extracting the salary pools that support local economies and converting them to subscription revenue.

Microsoft understands this perfectly. Copilot at $30/month per user seems trivial against a developer's $120,000 salary. But multiply that across millions of workers and it becomes real money. The economics work spectacularly when you're taxing every knowledge worker on the planet.

Companies already spending $35-40 trillion on salaries can absorb another 0.5-1% for 'productivity enhancement'. That's $175-400 billion annually in platform revenue. At typical SaaS valuations of 8x revenue, that creates $1.4-3.2 trillion in market value. The current $560 billion investment looks conservative against that opportunity. Even capturing 0.25% of payroll justifies the current investment levels. The bubble narrative misses this completely - we're not overfunding relative to the TAM, we're underfunding.

As context, SaaS currently captures just 0.39% of developed economy payroll - about $150 billion annually. It's projected to reach 1.3% by 2032. AI subscriptions could easily push that to 2-3% as they attach to every job function. We're still in the earliest stages of this transformation.

The lock-in dynamics make this particularly attractive to platforms. Once your historical data accumulates in their cloud, once your workflows encode their APIs, once your processes assume their availability, switching becomes organisationally impossible. Beyond the technical challenge of migration, there's retraining costs, process reformation, compliance restructuring. Companies will pay increasing subscription fees rather than face the disruption of change.

And the EU AI Act's compliance requirements actually strengthen these moats. Bias testing, audit trails, documentation requirements - these are important protections, but small companies can't afford to build compliant AI systems from scratch. They have to use the platforms that already have the infrastructure, the certifications, the legal teams. Brussels wrote laws to constrain Silicon Valley and accidentally built higher walls around its castles. Your regulatory burden becomes their competitive advantage. Watch how the big platforms respond - public concern about 'innovation-stifling regulation' while privately ensuring they help write the rules. They learned from banking: when compliance costs are high enough, regulation protects the biggest players.

Unlike salaries that circulate locally - keeping businesses alive and communities functioning - subscription revenue concentrates in Seattle, San Francisco, Redmond. Every pound redirected from London wages to OpenAI subscriptions is extracted from the UK economy permanently, undermining local tax bases while enriching US platforms. It doesn't recirculate through local businesses. It doesn't support local employment. It flows in one direction: toward platforms. This revolution is not a dinner party - the economic geography of Europe and Asia gets hollowed out to fill American platform treasuries.

No conspiracy coordinates this transformation. Microsoft doesn't collude with OpenAI who doesn't coordinate with Google who doesn't align with Anthropic. Each platform responds to local incentives. Each investor funds plausible narratives. Each customer buys marginal improvement. The system converges on the same outcome through millions of independent decisions. Like containerisation creating global supply chains nobody planned, AI investment creates productivity taxes nobody designed but everyone enables.

THEY THINK THEY'RE BUILDING MINDS. THEY'RE BUILDING METERS.

Most investors focus on the AGI potential rather than the payroll opportunity. They believe in consciousness emerging from silicon. They're pouring billions into the fantasy of artificial minds while accidentally building tollbooths on human cognition. Sam Altman probably genuinely believes he's two years from artificial general intelligence. The venture capitalists writing nine-figure checks definitely believe they're funding the next stage of evolution.

The $560 billion invested so far could rationally double or triple. Not because AGI is coming - we've shown **it's architecturally impossible with current approaches**. But because capturing even 0.5-1% of developed economy payroll through platform subscriptions justifies massive valuations. Once someone explicitly articulates 'we're building the transaction layer for global payroll', the unconscious convergence becomes conscious strategy. The land rush that follows will make the current investment look quaint.

They think they're building minds. They're building meters.

OUTRO:
THE MACHINE CAN'T PLAN. YOU CAN'T STOP.

In December 1968, Douglas Engelbart demonstrated the future. A computer mouse. Windows. Hypertext. Video conferencing. Collaborative editing. Everything we'd use decades later, running on a mainframe the size of a room.

The audience - a thousand computer professionals - gave him a standing ovation. They called it 'The Mother of All Demos'. They knew they'd seen the future.

They were right about what. Wrong about when. And completely blind to how.

It took fifty years for Engelbart's vision to become mundane. Not because the technology was hard - though it was. But because becoming infrastructure takes time. Everyone sees the revolution coming. Nobody recognises it when it arrives.

WHAT WE'VE LEARNED

Through this book, we've dismantled stories that obscure AI's reality:

The Bubble Story: AI isn't crypto 2.0. It's the next platform shift, following the same infrastructure patterns as electricity, automobiles, and the internet. Bubbles pop. Infrastructure persists.

The AGI Story: We're not racing toward artificial consciousness. Language models aren't mystical. They're massive statistical engines finding patterns in text. No consciousness, no understanding - just extraordinarily useful pattern matching at scale. The magic is in what we build with them, not what they are.

The Productivity Story: Expecting to measure AI productivity gains during transition is a fool's errand. The printing press took 560 years to show measurable impact. Electricity took 40 years. Transformational technologies reorganise work so fundamentally that traditional metrics become meaningless.
The Automation Story: Mass unemployment isn't coming. Work is multiplicative, not fixed. Every platform shift creates new categories of labour faster than it eliminates old ones. The question isn't whether there will be jobs, but what kinds of jobs they'll be.

The Capital Story: The $560 billion flooding into AI makes sense once you see payroll as the TAM. AI platforms aren't replacing the $40 trillion in global salaries - they're attaching subscriptions to them, building tollbooths on human productivity.

Throughout these essays, notice how naming shapes perception. 'Artificial Intelligence' that's neither. 'Neural networks' with no neurons. 'Attention' without awareness. The entire field runs on labels that obscure rather than reveal. What start out as helpful metaphors for novel concepts outlive their usefulness and become as actively harmful as they are sticky.

We've also seen a consistent business pattern: control attention, capture value. In music, that meant controlling playlists and making artists pay to play. In search, controlling results. In AI, it means controlling the base models that shape how millions process information while building tollbooths on the $40 trillion payroll pool. The stakes are now both economic and ideological - who controls these models determines not just who profits, but whose worldview gets embedded in the infrastructure everyone uses.

THE FOUR LENSES REVISITED

We began this book with four lenses for understanding AI. Each revealed something the others couldn't:

Infrastructure: AI is following the classic pattern - from novel technology to invisible utility. Like electricity or the internet, it will disappear into the background while transforming everything built on it. The real story isn't the models themselves but the infrastructure layer being laid down, pipe by pipe, standard by standard.

Platforms: The economics are predictable. A few players will control the core infrastructure - OpenAI, Google, Anthropic - while open source provides alternatives. Unlike Spotify's gatekeeping of attention, these platforms extract value through usage pricing and API access. But the real prize is the $40 trillion payroll pool - platforms are building tollbooths on human productivity, converting salaries that once circulated locally into subscription revenue. The question isn't whether platforms will dominate - they will. It's what leverage remains for those building on top.

Iteration: When revision costs approach zero, work becomes iterative rather than planned. These systems generate possibilities but can't judge quality - that requires human navigation and judgement. The slot machine dynamic is real: variable rewards create addiction while constant iteration demands steering.

Organisation: Work isn't disappearing - it's multiplying through platform dynamics. When capabilities become cheap enough to deploy everywhere, demand explodes. Every spreadsheet created more financial analysts, not fewer. AI follows the same pattern: the technology doesn't eliminate work, it reveals latent demand that was economically invisible before.

Each lens showed us something crucial. Together, they cut through the noise to show what AI actually is: infrastructure becoming boring, platforms extracting value, work reorganising around iteration, and jobs multiplying rather than disappearing.

THE NEW LITERACY

The real transformation isn't technological - it's cognitive. We're developing a new kind of literacy.

When writing was invented, Socrates worried it would destroy memory. He was right - we lost the ability to memorise epic poems. He was also wrong - we gained the ability to build knowledge across generations.

We're experiencing what happened with word processors, but at a different layer. Word processors made revision free, transforming writing from linear to iterative. Writers stopped crafting perfect first drafts and started thinking through writing rather than before it. The delete key changed how we think - why be precious about sentences when you can try twenty versions?

LLMs make variation free, transforming ideation itself into an iterative process. We're no longer confined to the ideas we can generate alone. Need a different angle? Generate ten. Want to explore a metaphor? Push it to its limits. Stuck on phrasing? Try variations until one clicks.

This changes how we think. Instead of carefully developing one idea, we generate many and recognise the good ones. Instead of writer's block, we have filter fatigue - too many options, not too few.

The cognitive shift is profound. We're moving from creators to curators, from architects to editors. Not because we can't create - but because when you can explore a hundred variations in the time it took to craft one, the skill shifts from generation to recognition. The amplification of possibilities demands it. Knowing what's good becomes more valuable than making it from scratch. It makes expertise more important, not less.

And the speed changes something else. When iteration is instant, planning becomes procrastination. Software learned this with agile - you can't write detailed specs when requirements change weekly. AI pushes this further: when you can try twenty approaches in the time it took to plan one, the skill shifts from deliberation to experimentation. Move fast and iterate replaces measure twice, cut once.

This isn't dumbing down. It's climbing up the ladder of abstraction. Each generation builds higher-level tools and loses lower-level skills.

THE ITERATION ADVANTAGE

The deeper change isn't just about iterating faster - it's about what infinite iteration does to us.

Remember the slot machine dynamic from earlier? When revision costs approach zero, work becomes addictive. You don't stop at good enough because the next iteration might be perfect. The variable rewards keep you pulling the lever.
This intensity isn't unprecedented. Early factory workers faced the same challenge when power looms set the pace instead of human hands. The loom never tired, never slowed, demanded constant attention. Now we match cognitive machines that iterate endlessly. Every productivity revolution has promised less work but delivered more intensity. The machines change; the human experience of being paced by them doesn't.

NOT ARTIFICIAL. NOT INTELLIGENT.

We spent this book dismantling fantasies. AGI isn't coming. The bubble isn't popping. Jobs aren't disappearing.

But something real is happening, and pretending otherwise is just as delusional as the hype.

AI's bargain: infinite capability without comprehension. Tools that can process all human knowledge but understand none of it. Systems that generate endless possibilities but can't evaluate a single one. We get superhuman pattern matching paired with subhuman judgement.

Not artificial. Not intelligent. Just mirrors made of maths.

But what mirrors. We've built something that can read every book, analyse every image, generate infinite variations of any idea - all while having no idea what it's doing. The machine can't plan. You can't stop. That's the user experience in seven words.

Use these tools for what they are: brilliant, frustrating pattern-matching systems that will make you more capable and more tired. Stop waiting for intelligence that isn't coming. Stop measuring productivity that won't appear for decades. Stop fearing unemployment that platform economics makes unlikely.

The infrastructure is being laid. The platforms are consolidating. The iteration cycles are accelerating. We know how this ends - not in detail, but in pattern.

Just don't call it intelligence. It's an insult to actual thinking.

NOT ARTIFICIAL, NOT INTELLIGENT

WHAT AI COMPANIES DON'T WANT YOU TO KNOW

Django Beatty has spent four decades watching technology reshape business and culture.

He started in the City of London during the Big Bang, as spreadsheets replaced paper and mainframes gave way to PCs. During the dotcom era, he ran an independent record label and built large-scale media sites. Later, at Capgemini, he led enterprise web platforms before founding Fluxus, a boutique consultancy focused on AWS and AI.

He's lived through every cycle of hype and disappointment - from the internet's first broadband leap to today's AI - and learned to separate noise from signal.

You've been told AI will either save humanity or destroy it. You've been told your job is disappearing. You've been told we're racing toward artificial consciousness.

It's all wrong.

In this clear-eyed investigation of what AI actually is - pattern matching, not intelligence - Django cuts through both hype and panic to reveal what's really happening: a platform shift following the same predictable patterns as electricity, the internet, and every transformative technology before it. While pundits debate consciousness and executives chase productivity metrics that won't appear for decades, the real transformation is already underway.

From why ChatGPT sessions become addictive iterations to why the 'AI job apocalypse' keeps getting postponed, from the truth about copyright battles to what AI companies know but won't admit, this book provides the orientation everyone needs to navigate a genuinely transformative but wildly misunderstood technology.

Simon Wardley – Advisor and Speaker
"'The AGI bubble will pop. Not because the technology fails, but because the fantasy can't survive contact with reality... The real revolution isn't making machines think. It's making them boring enough that nobody has to think about them.' - Well said"

Steve Yegge – Engineering Leader
"I like this take. Cognitive exoskeleton pilot sounds about right to me."

Dries Buytaert – Founder of Drupal, Co-founder of Acquia
"Well-written... it made me think. I loved that about this essay."