

Ten observations on creating value with AI

Frontier models are becoming table stakes. The durable advantage lies in the context you feed them and the capability and discipline to deploy them. Ten observations, drawn from our recent discussions with the world's leading AI companies in the US, and our own lived experience using AI to transform how Quantium and our clients work.

Introduction

Many organisations have made real strides in adopting AI over the past two years. Access to a frontier model or a chatbot, however, is no longer enough to compete. Every serious organisation has access to the same models and chatbots, which makes them table stakes rather than an edge. The advantage now comes from two things that are much harder to acquire: the context you can feed those models, and the capability and discipline to deploy them. On both counts most organisations are struggling. Nobody has this fully worked out, not even those at the forefront. But the gap between those who have started in earnest and those who have barely begun is already enormous, and it is widening fast.

The key reason it is widening so fast is the shift from chatbots to agents. Until recently the story was the chatbot, a capable assistant that helps a person do each step of their work a little better. That was the on-ramp, and it was valuable, not only for the work it did but for opening up people's imagination about what was possible. It was never the destination. Through 2026 the shift has been decisively to agents: systems that do not just assist a person through a task but complete the task, and increasingly the whole workflow, on their own. This is a different order of productivity, because the person moves from doing the work to directing and reviewing it, and in the most advanced cases the agent reviews the work as well, with people moving up to monitoring the system as a whole. It is not a roadmap item. Agentic capability is in production today, in tools already moving into everyday use: coding agents such as Claude Code and OpenAI's Codex, and general purpose "cowork" agents such as Claude Cowork, OpenAI's ChatGPT Work and Microsoft's Copilot Cowork, which operate alongside a person across their tools, files and data. Organisations still finding their feet with chatbots are not one step behind.

They are at the starting gate of a race others are already running, and the distance opens up very quickly.

Frontier capability, then, is the price of entry. The two things that actually separate organisations are context and deployment. Context is your proprietary data, your intellectual property, and the knowledge held in your people and your conversations, made available to the model so that its output is genuinely yours rather than generic. Deployment capability is the harder, more human work of redesigning how the organisation operates around what AI now makes possible, and getting people to genuinely adopt it. Most organisations find both difficult, and the obstacle is rarely the technology. It is organisational: the creativity to reimagine the work, the discipline to deploy broadly rather than pilot narrowly, and the leadership to drive it personally.

The result is a gap between competitors that is larger than most leaders appreciate, and it compounds. The organisation that is ahead builds context that makes its AI better, which drives adoption, which generates more context, which widens the lead again. The organisation behind falls further back with every cycle. The advantage is not only in the context. It is also in the organisational muscle that builds up alongside it: the muscle to adopt new ways of working, to reimagine with AI, and to learn from mistakes and keep moving. That capability is itself hard to copy. The uncomfortable version of this, which we are seeing repeatedly both on the frontier and in our own work at Quantum, is that the leaders who feel furthest behind are usually the ones who have gone furthest, because serious engagement reveals how much further there is to go, while the ones who feel no urgency are usually the ones who have not started. Once established, this kind of advantage is very hard to reverse.

It is worth being clear about the nature of the risk, because it is easy to underestimate. The threat is not that a competitor bolts AI onto the same business and runs it slightly better. It is that a competitor who has genuinely rebuilt around these capabilities, or an AI-native version of your business entirely, reaches a level of speed, cost and customer experience that a slower incumbent cannot match. In that contest, standing still is not a neutral choice. It is a decision to fall behind, and in time to put the survival of the business at risk. The organisations treating this as urgent are responding to that reality, not to hype.

The observations in this paper draw on two sources: our recent study tour of the world's leading AI companies in Seattle and San Francisco, and our own lived experience using AI to transform how Quantum and our clients work.

Our ten observations

1. Leadership from the top
2. The right tools
3. Context and organisational memory
4. Governance, security and cost control
5. A long-term view of talent
6. A culture of reimagining, adopted broadly
7. Skills
8. Bespoke apps and workbenches
9. Agents
10. Start with value, and own the benefit

The precondition

01. Leadership from the top

AI transformation succeeds or fails on whether the most senior leader drives it personally. This is the single most consistent predictor of success we see, and it is not a motivational point. It is a mechanical one.

Leaders set the level of ambition, and you can only calibrate an ambition you have felt yourself. A leader who has not used the tools will, in good faith, ask a team to deliver in twelve months something the technology now does in six weeks, and the organisation will quietly move at a fraction of the pace it could, with no one aware it is happening. The only reliable cure is first-hand fluency. The leaders we see pulling away are, almost without exception, the ones on the tools themselves. And the leaders getting the biggest breakthroughs, both those we met in the US and those we work with day to day, tend to set genuinely ambitious targets for their teams, the kind that cannot be reached by doing the current process a little faster and so force the team to reimagine how the work is done.

There is a second mechanism, signalling, and our own experience puts hard data behind it. As we drove adoption at Quantum we tracked usage by individual, and the pattern was unmistakable: where a leader's own usage was low, so was adoption across their team, and as a leader's usage rose, their team's rose with it. When people see their leader personally using AI and reorganising real work around it, they understand it is genuine. The failure mode is the leader who sponsors from a distance and never builds the personal judgement to know how hard to push. Usage is only the start, of course. Over time, what matters is not how much AI a team uses but how much value it extracts.

The foundations: what you give the AI

02. The right tools

Putting the best frontier models and tools in your team's hands, and keeping pace as they change, is becoming table stakes.

The practical discipline is to separate two different decisions. The first is your organisation-wide productivity tool, the one everyone uses day to day for chat and cowork-style work, such as Claude or Microsoft Copilot. Here it pays to choose one primary provider thoughtfully and go deep, directing your team's energy toward deployment and adoption rather than perpetual evaluation. The second is the tooling for specific, higher-value use cases, above all the agents you build. Here you should not tie yourself to a single set of models. For coding, for example, you need an agentic coding tool such as Claude Code or Codex, but within it you can orchestrate with a frontier model while routing cheaper or open-weight models to do the bulk of the execution, which can cut cost sharply without materially affecting the result. The same holds for the agents you build for particular processes: keep them free to route to whichever model is most effective and economical for each step.

The mistake at each layer is different. On the organisation-wide tool it is tool tourism, an endless evaluation of the newest thing that never converts into deployed value. On the use-case and model layer it is over-committing, locking into a single tool or model in a market that moves too fast and where much of what you adopt will be superseded within a product cycle. The balance to strike is a committed core with flexible edges: go deep on one primary productivity tool, put the right specialist tools in the right hands, and keep your agents free to route across models as cost and capability demand.

03. Context and organisational memory

This is the moat. The quality of an AI output now scales with the richness of the context it can draw on far more than with the choice of model, and every company we visited had independently reached this conclusion. The task is to give your tools access to your proprietary data and intellectual property, and, just as importantly, to systematically capture the knowledge that currently lives only in conversations and in the heads of your best people.

It is worth being clear about what this does and does not mean, because it is easily misread. Making your data and IP available to your AI tools does not mean handing them to the model providers or giving up your advantage. Under an enterprise agreement your data remains yours and is not used to train their models. This is the same trust you already place in the major cloud providers to hold your most sensitive data, and their whole business depends on honouring it. The point is to make your own context available to the models you run, so their output reflects your business rather than the world in general.

The context spectrum

Foundations How the organisation thinks	Strategy, values, policies and employee handbooks.
Conversations Where reasoning lives	Transcripts and summaries of meetings and conversations.
Tacit knowledge Hardest of all	The expertise in your best people's heads, deliberately extracted and captured so it does not walk out each evening.
Connected systems What you run on	Customer, supply chain, finance, HR and CRM data, and every other internal source.
Semantic layer How you calculate things	When a user asks for margin, it knows precisely how margin is defined in your business, so the answer is correct, not a plausible guess.

The benefits of getting this right are large and they compound. The first is **information flow**. When context is captured and made available rather than trapped in individual inboxes, calendars and memories, knowledge moves freely across the organisation, teams work from the same picture, and people stop rediscovering what a colleague already knew. The second is **corporate memory**. Knowledge no longer walks out of the door when a person resigns, because it has been captured as the organisation's, not the individual's. Consider a meeting you had six months ago. Without your notes you will remember almost nothing, and even your notes will be sparse next to a full transcript and summary. At Quantum every meeting is recorded and

captured, including one-on-ones, and the benefit is enormous: nothing of substance is lost, and any of it can be retrieved, queried and built upon later. A third benefit follows from the same foundation, the end of traditional business intelligence as we know it. Rather than a central team building fixed dashboards, anyone can ask the business a question in plain language and get an answer, which puts inquiry in the hands of everyone, provided the underlying data and context are genuinely AI-ready.

This raises a governance question that most organisations have not answered: once anyone can ask anything, how much of the organisation's context should be shared, and with whom. Our instinct is to share the maximum that is sensible, because shared knowledge compounds, while protecting the narrow set of things that genuinely should not be widely available, though that is not trivial to implement. The failure mode is leaving expertise locked in individuals and pointing raw models at ungoverned data, which is slower, less accurate and more expensive all at once.

04. Governance, security and cost control

Treating governance as a brake is the mistake. Done well, it is what lets you scale safely, and it is inseparable from managing cost. The most useful reframing we heard is that data governance, AI governance and spend control are one problem, best solved with a single control plane across your data, models, agents and tools.

Governance in an AI world has a few concrete parts: knowing which tools and agents exist and what each one does, tracing and auditing what they have accessed and produced, and controlling how your data and IP are used and by whom. It operates at two levels. The first is the personal-tool level, the chat and cowork tools your people use every day, where the priorities are sanctioned and secure access, sensible data controls, and visibility of use, not least to head off the shadow IT that appears when people reach for personal accounts. The second is the enterprise-agent level, the agents you build to run real processes, where you need traceability of every action, clear ownership, and the ability to prove after the fact what an agent did and why. Getting both in place early is a large part of why some organisations can put agents into production while others stall.

Cost deserves particular attention, and it is one we are living day to day, both inside Quantum and alongside our clients. It also came up in every meeting on our US study tour, driven by a shift in how the AI providers charge, away from a fixed per-seat licence and toward consumption-based pricing where you pay for what you use. That turns AI from a predictable line item into a variable one that can move sharply, and the risk is that spend grows faster than the value it generates. The controls are practical. Ensure visibility so that users understand the costs that they are incurring. Set spend budgets by team, with the ability to see and cap consumption. Restrict the most expensive models to the people and tasks that genuinely need them. And route work intelligently: within an agentic coding tool, for instance, orchestrate with a frontier model but let cheaper or open-weight models do the bulk of the execution, on the principle that you do not need to drive the Ferrari to pick up the milk. For us this is a material saving, since coding is one of the largest lines in our own AI bill. The failure mode is discovering your AI cost line only when it becomes uncomfortable, by which point you have neither the visibility nor the controls to manage it down.

The people: who puts it to work

05. A long-term view of talent

The common assumption is that AI reduces the importance of talent. The opposite is true. AI is a multiplier, a form of leverage, and it does not lift everyone equally. A great performer becomes dramatically greater. A weaker performer can actually become a drag, because the volume of low-quality work that they now generate creates more work for everyone around them. So the gap between your best people and the rest is wider than it has ever been, and the premium on securing and keeping the best goes up, not down.

The challenge is identifying who your top talent actually is, both inside your organisation and when you recruit. Two things make this harder than it sounds. The first is that your top talent today may not be the same group as your top talent yesterday. The people who thrive in an age of AI are not always the ones who excelled before it, because the attributes that matter have

shifted. The biggest winners tend to be the curious and the adaptable, those with a builder's mindset and a habit of first-principles thinking. Amplified by AI, people like this can increasingly reach the depth that used to demand a narrow specialist, which means they will often outdo specialists who are deep but set in their ways. The second is that it is increasingly difficult to distinguish genuine talent from a polished resume and well-prepared interview answers, which AI itself makes easier than ever to produce. The more sophisticated organisations are responding by changing how they assess people, for example asking a candidate to build something with AI during the interview rather than simply talk about it.

The implication is to take a long-term view rather than treating AI as a licence to cut. Keep hiring and developing people, invest in training, and in particular keep hiring graduates. Some AI leaders have suggested that AI will do away with entry-level graduate jobs altogether. We disagree. In our experience young people are often the most fluent with these tools, having grown up with them, and they become highly productive very quickly. Just as importantly, the graduates you develop now become the senior capability you will depend on later. Hollowing out the pipeline to bank a short-term saving runs against everything we believe about talent, and it is a decision most organisations will come to regret.

There is a related shift in where the change actually lands. Because much of what AI removes is the effort of coordinating, synthesising and relaying information, the roles most exposed are often not the cheap entry-level ones but the more expensive middle-management and coordination layers whose main purpose was to move information up and down.

As that work is automated, spans of control widen and structures flatten. This is the opposite of the instinct to cut juniors first, and it is why the organisations thinking clearly about this are removing layers rather than trimming the bottom of the pyramid. AI reinforces this from the top as well. Senior people can now get on the tools themselves, designing prototypes in natural language and bringing their own ideas to life directly with the builders, without a chain of intermediaries to brief and coordinate. That too thins the middle.

A specific caution cuts against another common assumption. AI is not removing the need for deep technical talent, it is raising it. What your engineers and data scientists do changes, from building most of it themselves to specifying, reviewing and orchestrating it, but building the

context layers, the skills and the agents still requires people who understand both the technology and the domain. That talent is genuinely scarce, and every opportunity on this list depends on having it, whether you build it or borrow it.

06. A culture of reimagining, adopted broadly

Value does not come from bolting AI onto today's steps. There are three levels, and the returns rise sharply as you move through them.

AI-enhanced AI layered on an existing process	Useful but incremental, from 10–20 per cent up to one or two times, depending on the task.	1–2x Efficiency
AI-native Process redesigned around AI	The agent does the work and people orchestrate it.	3–5x Efficiency
AI-frontier Previously not possible	Attempting what was simply not possible before.	10x+ OR MORE

The reframing that matters is not how to use AI to fit the current organisation, but how to change the organisation to fit the use of AI. Put simply, the opportunity is not to get from A to B a little faster, and then B to C a little faster. It is to rethink how you get from A to Z altogether.

That is a cultural challenge before it is a technical one. The consistent estimate we have arrived at is that transformation of this kind is roughly 80 per cent human and 20 per cent technical, and the primary risk is human adoption, not architecture. Trust is the currency of AI adoption. People will only hand real work to a tool, or an agent, once they have come to trust it, and earning that trust is the central task. This is why broad, measured deployment beats a handful of careful pilots. Broad rollout is what generates the signal that tells leadership where the real value actually sits and which teams are building the habits others need to learn from. Narrow pilots suppress exactly that signal. The failure modes are an efficiency-only mindset that never reimagines anything, and pilot purgatory, where promising experiments never reach the scale at which they would change the economics.

There is also a practical trap. It is unrealistic to ask a team to step back and fundamentally reimagine how they work while delivering their day job in full. The two compete for the same people and the same hours, and day-to-day delivery almost always wins. Reimagining the work needs deliberately created space, whether that is dedicated time, a separate team, or explicit permission to set some business as usual aside. Leaders who expect transformation to happen as a side project on top of an unchanged workload are usually disappointed.

The build: how the work changes

Skills, workbenches and agents are the three things you build, and they are easiest to understand in relation to one another.

Skills

reusable building blocks

Do one or two steps well, captured once and reused everywhere. They sit alongside other building blocks such as connectors and the context layer.

Workbenches

surfaces people work at

An app people open and drive, pulling the work through step by step. Chatbots and other interfaces are surfaces too.

Agents

background workers

Run in the background and do the work themselves, often across many steps, supervised at first and, in time, sometimes unsupervised.

Whatever you build, build it in service of a reimagined workflow. The pattern that works, both inside Quantum and with the clients furthest along, is to pick a workflow, reimagine how it should work in future, and then build specifically to deliver that outcome. The skills, workbenches and agents you need become obvious once the new workflow is clear. The trap is the reverse: building a clever solution first and then hunting for a problem for it to solve.

07. Skills

A skill is your best-practice method for a task, codified once as a reusable procedure that a person or an agent can follow. It is the mechanism by which hard-won expertise stops living in a few people's heads and starts being applied consistently across the whole organisation.

The range is broad. At the simple end, a skill might instruct your AI system how to format a document to your house style, draft a standard customer response, or structure a routine report. Further along, it might capture how to build a presentation in your template, review a contract for the risks you always look for, or run a calculation exactly the way your business does it. A skill can also act as your digital twin, writing in your voice, reviewing documents the way you would, and offering thought leadership and coaching as you would yourself. And at the more advanced end, a skill might iteratively build and optimise machine learning models, drawing on twenty years of best-practice techniques and hard-won experience from across your data science team.

This matters for two reasons. It lifts the floor, because everyone now performs the task to the standard of your best practitioner rather than their own. And it compounds, because each skill you capture becomes a durable asset that improves with use and can be composed with others. This is the same lesson the leading application companies have learned in their own domains, where the differentiator has become the layer that captures and applies context and method, not the raw model underneath. The failure mode is the status quo in most organisations: expertise trapped in individuals, the same problems solved from scratch each time, and quality that varies with whoever happens to pick up the work.

08. Bespoke apps and workbenches

This takes the idea of a skill one step further, into the hands of your people through a purpose-built application. We refer to these as workbenches, though you might equally call them bespoke apps. Each one embeds a reimagined workflow, so that everyone works to the same high standard through a dedicated tool rather than being handed raw AI and left to produce variable results. The difference is consistency. Give a hundred people a general-purpose chat tool and

you will get a hundred different levels of capability. Give them a purpose-built app built around the right workflow and you standardise the good way of working.

A communications workbench is a good example. Rather than each person starting from a blank page, it holds your brand standards and document templates, your house voice, and the distinct voices of individual leaders, so a draft can be produced in the right register from the outset. It can then review that draft the way your best communicators would, checking it against how it is likely to land with different audiences, from staff and customers to regulators and the media, by simulating those stakeholders before anything goes out. The result is communications that are faster to produce, more consistent, and better pressure-tested, whoever is doing the writing.

What makes this newly practical is the collapse in the cost of building software. Applications that would recently have required a significant development programme can now be produced quickly and cheaply, which means a bespoke app for a high-value, high-variability workflow is often worth building rather than buying. Where a capable third-party product already exists, licence it. The judgement is knowing which workflows justify a bespoke build and which are better served by something off the shelf. The failure mode is defaulting to raw tools and hoping consistency emerges, or over-engineering a bespoke build where a licensed product would do.

09. Agents

Agents are where the work actually gets done. An agent draws on your context and your skills to complete tasks and take actions, and every company we visited had them running in production. The unit of transformation is the workflow, and agents are what let you redesign an end-to-end workflow rather than assisting each step within the old one.

There is a compounding relationship between agents and context worth calling out. Agents work better when they have good context, the same way a new team member does. The more powerful effect is the reverse: as agents run, they can continually improve the context they draw on, capturing what they learn back into the shared layer. We have many examples of agents both benefiting from better context and making it better, which is one of the ways the advantage described in observation three keeps compounding.

The disciplined path to deploying them is well established and worth following exactly. Start with a single contained task. Build the guardrails and the verification. Let the agent become reliable enough that you can look away. Then chain tasks into workflows and workflows into larger processes. The verification loop is not optional: the leading practitioners capture and score every agent interaction against an evaluation model that feeds back and improves both the prompts and the agent over time, so deployment becomes a closed learning loop rather than a one-off build. The reason most organisations are not yet running fully autonomous agents is rarely technical. It is the work of building trust, evidence and governance around them, which is a leadership task. The failure modes are reaching for autonomy before trust is earned, and never progressing past assistance into genuine execution.

The payoff: how it creates value

10. Start with value, and own the benefit

The final discipline is the one that converts everything above into results, and it has two halves. The first is to start from value and work backwards. Rather than chasing a thousand experiments, identify the small number of use cases that genuinely change the economics, and prioritise those. A useful test is to favour compounding problems, where every additional increment of model intelligence unlocks disproportionately more value, over saturating problems such as drafting a routine email, where even a perfect answer is worth little more than a good one. This works best when you pair people who understand the business deeply with people who bring the technical capability, embedding the two together rather than running technology at arm's length from the value it is meant to create.

The second half is to own the benefit. Freed capacity is not a profit-and-loss benefit until something claims it. If AI saves a team time and that time is quietly reabsorbed, no value has been created, only a more comfortable status quo. Every initiative therefore needs a named mechanism that converts the gain into a result, with a specific owner and a number. That result need not be cost. The best organisations look at value through several lenses at once: not only

cost efficiency, but increased output and capacity from the team you already have, faster speed to market, a better experience for employees, reinvented customer engagement, and a genuine lift in innovation. Framing AI purely as a way to cut headcount both understates the opportunity and sets the wrong tone with your people. What matters is that, whichever lens you choose, the benefit is named, owned and measured. The failure mode is a portfolio of impressive-looking pilots with no line of sight to value, and capacity that is freed but never captured.

Where the value comes from

It helps to be clear about how AI actually creates value, because the types differ enormously in impact. We think of five, rising roughly in ambition and impact.

Transformation of individual tasks productivity enablement

Roll the tools out to everyone and teach people to use them well. This is where most organisations start, because people are largely making individual tasks faster rather than changing how the work is done.

Transformation of the work reimagined workflows and agentic automation

Redesign whole processes around agents and workbenches, cutting time, cost and risk while lifting quality and consistency, with fewer errors and stronger compliance and resilience. This is where meaningful value begins.

Transformation of decisions from BI to AI

Move from dashboards to actionable insight. Teach the AI to interpret the data and surface what matters, through conversational analytics that answer questions on the spot, targeted recommendations that reduce variation in performance (a digital twin of your best operator, applied to each person's own data), and diagnostic and scenario tools that show what is really driving performance.

Transformation of the customer experience hyper-personalisation

Combine predictive models with generative AI to deliver genuine segment-of-one interactions, lifting acquisition, cross-sell, retention, pricing and satisfaction.

Transformation of the business innovation, new models and previously unsolvable problems

Use AI to challenge the assumptions and constraints in the business model itself: to create new products and services, to reach customers who were previously uneconomic to serve, and to turn your data into new sources of value. This is also where you point serious AI at the hard, high-stakes problems that were once too expensive or too complex to attempt, where even a small improvement is worth an enormous amount and every increment of intelligence compounds the return. Using AI to identify the next mine site is exactly this kind of problem.

A closing note

These ten are not a sequence to work through in order. They are interdependent. Leadership is the precondition that unlocks all the others. Context and the people elements are the moat, the parts a competitor cannot simply buy. The build layer, skills, workbenches and agents, is how the reimagined work actually gets done. And value discipline is what makes it pay.

The technology to do all of this exists today. The gap is not the technology, and it is not simply will. It is the capability to turn that potential into realised value: the leadership to commit and drive it, the people and skills to build it, and the organisational discipline to deploy it and see it through. That capability can be built, and it can be partnered, but it cannot be skipped.

