A FIELD GUIDE / FIRST EDITION

Becoming

AI-Native.

A Field Guide for Leadership Teams Designing the AI-Native Company

Field Guide | First Edition

2026

Arc5 | Collaboration Experience Design
Nodalix | Industrial Intelligence and AI-Native Architecture

Table of Contents

Forty sections across ten stages.

STAGE 1

The Premise

What this field guide is for, and the question underneath.

Section 1.1

Why This Field Guide Exists

What is this field guide, and why might it help?

BRING THIS INTO THE ROOM

A note before you start. This field guide was written for you to read. It was also structured so that an agent can read it alongside you. Every section, every framework, every callout, every cited source has been organized into a knowledge graph, the same architecture the field guide itself recommends in Section 4.1. The graph is a companion to the book. You can use this field guide in two ways. You can read it linearly, as a book. You can also query it, through your own AI agent or through the one I provide, and ask the questions specific to your situation, your industry, the decision you are about to make. The agent does not replace the reading. It sits alongside the reading, helping you locate the parts of the book that matter most for what you are working on, surfacing connections across sections that linear reading does not produce, and translating the frameworks into the specific decisions your leadership team is struggling to make. I did this deliberately. A field guide arguing that AI-Native companies make their institutional knowledge queryable should be queryable itself. The book closes with a section called Two Ways to Read This, which returns to this and shows you how to use both modes together.

If you are reading this, you are probably somewhere in the work of becoming an AI-Native company. Maybe your CEO put the phrase on a whiteboard and you have been trying to figure out what it actually means. Maybe you have already spent eighteen months and a lot of money

on AI pilots that have not produced what was promised. Maybe you are at the start, trying to decide whether to commit, and looking for something solid that does not feel like a sales pitch.

This is that.

I wrote it because almost everything I have read on AI-Native is either too high to act on or too narrow to be useful. The strategy decks describe the destination without the path. The implementation playbooks describe the path without the destination. The vendor white papers describe whichever piece the vendor sells.

This book is different. It holds the strategic, the human, the technical, and the operational in the same frame. It names the trade-offs.

It is also a design guide. A pattern I see often: leadership teams treat the AI-Native transition as something to go do. They pilot, they test, they implement, they call it fast failing when the pilots do not produce. If design happens at all, it happens after the first wave of failures has consumed the budget. I think this is the wrong sequence. Designing an AI-Native company is design work. Naming the considerations, weighing the trade-offs, choosing the path. The work belongs at the front of the transition.

The pushback I get on this is real and worth engaging. Design is expensive, slow, and what consultants do when they want to bill more hours. The companies that win are the ones that move fast, ship something, learn from the market, and iterate. Spending six months on design is six months your competitor spent shipping. I have heard versions of this from every leadership team I have worked with, and the people making the argument are not foolish. They are responding, often correctly, to a pattern of over-designed enterprise initiatives that produced a lot of slides and not much else. The argument has a real grievance behind it.

I think the argument was right and is becoming wrong. It was right because design, until recently, was expensive and slow. A single design option for a major operating-model change used to take a small team weeks to produce. Interviews, framing documents, scenarios, financial models, the rest of it. Multiple options took multiple weeks each. The economics rewarded teams that limited their design work, made one bet, and tried to execute it fast.

I watched this up close on ERP implementations. On the SAP and Oracle programs I worked on, armies of people were deployed for months to understand what was happening at the client, map the current state, and then build the crosswalk to what the process would look like in the

new system. Highly manual, and expensive. Most of the labor went into managing documents and inputs and trying to rationalize them against each other, working out interdependencies, tracing how different processes touched the same data. That was the real cost of design, and almost none of it was the thinking.

The design work remains, but AI has compressed enough of the labor around it to change what a team can afford to consider. I have found the same thing building the Nodalix application. I am not spending more hours on design than I used to spend. I am exploring more options per hour, pressure-testing them faster, and catching mistakes that would have shipped in the pre-AI version of my practice. Working in Claude Code within a configured codebase, the system can inspect the code that is available to it, so the design does not start from a blank page. I describe what I want to build and it helps me work out the appropriate path to take. Work that used to take a six-person consulting team is now within reach of one capable person to attempt, though domain review and collaboration still matter. That changes what a leadership team can afford to consider before they commit. The companies that recognize this change have a window of advantage. The companies that stick with the old economics of design, either by skipping it or by paying consultants for it the old way, are leaving that advantage on the table.

So my response to the pushback is this. The old preference for shipping over designing was a reasonable response to the old economics. The new economics are different. Design done well, with AI as a thought partner, is faster than it was. Shipping into the AI-Native transition without designing first is more dangerous than it used to be, because the stakes are operational, financial, and human at the same time.

This field guide is organized the way a football coach uses a playbook. A playbook is a menu of plays. No coach runs it start to finish. The right play for third-and-long is not the right play for the goal line. A coach reads the situation, finds the play that fits, runs it, and adjusts. The plays are the substrate. The judgment is the work.

This book works the same way. The ten stages and forty sections are plays. The frameworks inside are options to choose between based on where your company is and what you are trying to do. You do not need to read it front to back. You can find the section that matches what you are working on right now and start there. Come back to other sections when the work surfaces them.

A field guide is meant to be marked up. The marginalia and the dog-ears are the point.

You can work with this on your own. Many leadership teams will. You can also call me and Nodalix, and we will help. The book is yours to use either way.

A word about where all of this comes from, because two different things are braided together in it and you should be able to tell them apart. The AI thinking is recent and it is mine by way of building. Since I began building Nodalix, I have been constructing an AI-Native company rather than advising one, and almost everything I claim about what AI does to work comes from that, plus what I can see happening in the market, plus the research where the research is good. AI is new enough that nobody has a decade of AI-Native transitions behind them, and I will not pretend otherwise.

The methodology is the older half. Two decades of designing and facilitating collaborative sessions for senior teams, across industries and across every kind of hard decision a company faces. That work is older than AI. It is about how a room full of people reaches a decision they will actually carry out.

The bet this field guide makes is that those two things belong together. Becoming AI-Native is a design problem that a leadership team has to work through collaboratively, so I am offering the thinking on one side and the method for exploring it on the other. When I describe what I have seen companies do with AI, I mean what I observe in the market and the research. When I describe how a room works, I mean twenty years of standing in them.

Where I am uncertain, I will say so.

One last note before we begin.

Underneath the technical and strategic questions sits a different one: what is work going to be for, in a world where machines can do more and more of what used to require human cognition? Those questions are real, and they are downstream of this one. Most of the book lives in the technical and strategic. I will return, at intervals, to the question underneath.

Figure 1.1a. The stakes of AI-Native: operational, financial, and human. The leader is answering all three at once.

Which Play Do You Need Right Now?

This section promised you a field guide: a menu of plays. The router below is how you call the play. Find the line that sounds most like what is happening in your company, in your own words, and start with the sections listed, in order.

Come back when the situation changes. It will.

"The CEO announced we're becoming AI-Native and everyone heard something different." Start with Sections 1.2, 9.2, and 2.1. Name the ambiguity, use the sprint to build a working definition without erasing dissent, and ground it in an honest account of your current posture.

"We have eleven pilots and nothing in production." Start with Sections 8.2 (Case One), 1.2, and 2.4. The pilot graveyard can begin with unresolved meaning, ownership, evidence, or architecture; the developmental model helps locate what remains missing.

"Our pilots keep stalling and we can't tell why." Start with Sections 7.1, 8.2, and 3.2. Map the feedback structure rather than diagnosing it from frustration alone; compare the failure patterns, then inspect the operating system beneath the pilots.

"The board is asking what we're spending and what we're getting." Start with Sections 4.5 and 2.4. Separate spending that rents capability from spending that builds it, then examine why the strongest performance association in the cited research appears between its middle maturity stages.

"We don't know what to measure." Start with Sections 7.3, 7.1, and 7.2. Build an instrument panel that tests the transition without turning workers into surveillance targets, then connect the signals to feedback structures and competing causal explanations.

"People are scared." Start with Sections 5.2, 6.1, and 1.3. Make room for the emotions without assigning them, establish practices that protect candor, and confront the human premise underneath the transition.

"Everyone has Copilot and nothing is compounding." Start with Sections 4.3 and 2.3. You have the thirty-islands architecture; audit which of the nine layers you have actually invested in.

"Our most experienced people are going quiet, or leaving." Start with Sections 5.2, 7.1, and 8.2 (Case Five). Treat nonparticipation as information about fear, extraction, workload, privacy, or professional duty; the covenant and the failure case show what must change.

"We're about to announce a restructure." Start with Sections 8.2 (Case Four), 6.2, and 5.9. Do the design work before the announcement: map the moments that concentrate value and the invisible work each layer is doing.

"We handed the AI strategy to the CTO." Start with Sections 8.2 (Case Two), 4.2, and 5.1. Technical leadership is necessary and insufficient on its own; the five-part mandate shows where leadership accountability and cross-functional ownership belong.

"Our people are using AI tools we never approved." Start with Sections 4.3, 8.2 (Case Three), and 4.4. Prohibition alone is brittle. Pair useful sanctioned tools with proportionate controls, enforcement, and the epistemic disciplines the work requires.

"We can't tell whether the AI's output is true." Start with Sections 4.4 and 3.3. Corroboration, monitoring, and verification are the disciplines; the validate step of the work loop is where they live day to day.

"Our AI gives generic answers. It doesn't know our company." Start with Sections 4.1 and 2.3. Documents carry narrative and evidence; graphs expose named structure and relationships. Strong systems use both as context for reasoning.

"The leadership team doesn't actually use AI themselves." Start with Sections 5.1 and 5.6. The four stages and four obstacles locate the leader's practice; the work-design test asks whether the resulting environment supports challenge, agency, and care.

"Our people are getting faster but the work is getting shallower." Start with Sections 5.4, 5.5, and 6.1. Distinguish useful cognitive offloading from loss of required capability, then design the foundations, feedback, and team practices that support development.

"Junior people aren't learning the trade anymore." Start with Sections 5.9, 5.5, and 3.2. The apprenticeship inversion is a risk, not a foregone conclusion; redesign formative work so people still build judgment while learning to direct and check AI.

"Our managers are checked out, and AI is making it worse." Start with Sections 5.6 and 5.8. AI can give a distracted manager a plausible excuse for more distance. Engagement architecture and deliberate unlearning put attention back on the conditions people need.

"We don't know where humans should stay in the loop." Start with Sections 6.2 and 5.3. Place accountable human presence at the moments where rights, consequence, relationship, lived context, or professional responsibility make it matter.

"Our decks and reports look dated next to what others are producing." Start with Sections 6.4 and 3.3. Output, communication, and knowledge sharing are all shifting; set the standard deliberately, then make specification, validation, and learning part of the pathway.

"We ran the workshop, the energy faded, nothing changed." Start with Sections 7.2, 7.1, and 9.3. Test where the intervention met the system, identify the feedback structures carrying or resisting it, and establish the ownership and review rhythms that continue after the room closes.

Section 1.2

The Phrase on the Whiteboard

What does it take to turn "we will become AI-Native" into shared meaning?

Somewhere in the last year, in an all-hands meeting, a CEO stood at a whiteboard and wrote the sentence: we are going to become an AI-Native company. The room nodded. It landed well. Afterward the Slack channels lit up with reactions ranging from energized to cautious to cynical. By the next quarter, three different executives had launched three different transformations, each based on a private definition of what the CEO had meant.

This is the situation a great many leadership teams are now living inside. The phrase is hot. The intent is real. The alignment is missing.

The CFO heard cost reduction. AI is going to let the company do more with less headcount, and the timing matters because the board is asking about margin. The CTO heard modernization. The legacy platforms have been groaning for years and AI is the forcing function that finally produces the budget for the rebuild. The CHRO heard workforce anxiety. Half the workforce is already worried about their jobs and the announcement just made it worse. The CMO heard personalization at scale. The CIO heard infrastructure investment. The general counsel heard new regulatory exposure. The head of operations heard we're going to break things if we move too fast.

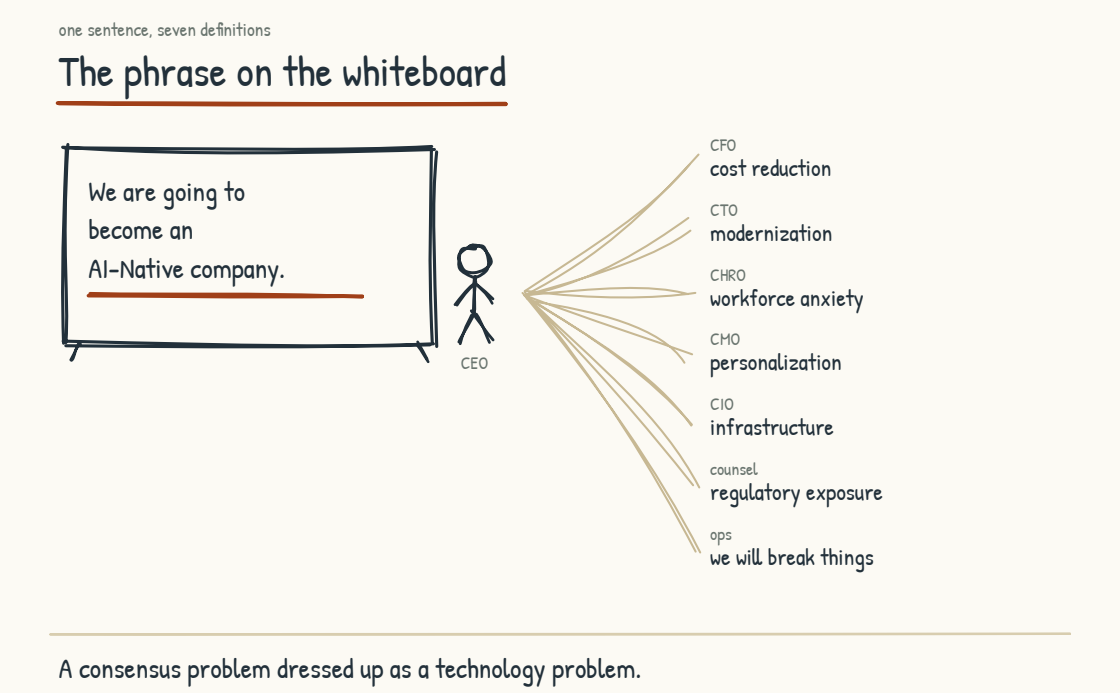

Figure 1.2a. One sentence on a whiteboard, seven private definitions leaving the room. Each reading is legitimate, each is incomplete, and each is already driving budget.

Each of these is a legitimate interpretation. Each is also incomplete. And each is now driving budget decisions, hiring decisions, vendor decisions, and roadmap decisions. None of it will converge into anything coherent unless someone in the room does the work of making it converge.

This is what I see leadership teams discover, sometimes after eighteen months and several million dollars of pilots, sometimes earlier: AI-Native is an alignment problem dressed up as a technology problem.

INTELLECTUAL LINEAGE

What is happening on the whiteboard is what Karl Weick called organizational sensemaking. A leader writes a phrase. The phrase is the same on the surface for everyone in the room. Underneath, each person constructs a different meaning from it, based on what they were already thinking, what they have heard before, and their own fears and hopes. The construction is private and rapid, and it feels, to the person doing it, like understanding. The misalignment is invisible because it lives in the meaning rather than the words. The leadership team's job is to make those private constructions visible, build enough shared meaning to act, and record the disagreements that remain.

Weick, Karl E. (1995). Sensemaking in Organizations. Thousand Oaks, CA: Sage Publications.

Research supports the need to frame the problem correctly, though not the claim that technology is the only obstacle. A 2024 RAND report noted that, by external estimates it cited, more than eighty percent of AI projects fail, roughly twice the reported failure rate of information-technology projects without AI. RAND did not produce that failure-rate estimate. Its own contribution came from interviews with sixty-five experienced data scientists and engineers, which surfaced five leading causes: misunderstanding or miscommunicating the problem, insufficient data, chasing technology rather than user need, inadequate infrastructure, and applying AI to a task beyond its capabilities. Problem definition led the list. Executives funding competing initiatives can destroy execution capacity regardless of how good the underlying AI is, but models, data, and infrastructure can fail too.

SOURCE

A recent interview study of why AI projects fail. RAND interviewed sixty-five experienced data scientists and engineers, identified five leading causes, and named

> *misunderstanding or miscommunicating the problem as the first. The widely repeated failure-rate estimate came from external sources cited by RAND, not from RAND's own measurement.*
>
> Ryseff, James; De Bruhl, Brandon F.; Newberry, Sydne J. (2024). The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed. RAND Corporation, RR-A2680-1.

I have come to call the result the pilot graveyard, and it shows up repeatedly in the industry literature and in my own work. The composite that follows gathers that pattern into one company. It runs eleven pilots. Marketing pilots a content engine, operations pilots a scheduling model, finance pilots a forecasting tool, and none of the three sponsors ever sit in the same review. All eleven draw on the same four-person data team, which spends its weeks context-switching among executives and finishing nothing. Eighteen months and four million dollars later, zero systems are in production. The leadership team is sincere, the capital is allocated, and nothing ships, because every initiative is solving for a different definition of the goal.

Companies that escape this pattern do something that looks small on paper and is unreasonably hard in practice. They produce a shared definition, and record the disagreements that remain, before they spend money on initiatives that depend on it. What works is recognizable in shape, and it is worth picturing concretely even in the abstract. A leadership team commits to a two-day session with an outside facilitator and produces a single statement of intent. Every subsequent AI initiative then gets evaluated against that statement. Proposals that do not support it get rejected or deprioritized. That two-day investment can become some of the most valuable work the leadership team does. Everything else is downstream of it.

No two of those statements look alike, and none of them would transfer. What matters is producing one together, with the senior team in the same room and the disagreements surfaced rather than buried. Without that work, the announcement that the company will become AI-Native is a wish wearing the language of strategy, executed by people who each heard a different wish.

Figure 1.2b. The closing turn of the opening argument. Alignment first, or the announcement is a wish.

This matters more for AI than it did for digital transformation or a cloud migration, because AI is unusually susceptible to private definition. A cloud migration can at least be anchored to named workloads and infrastructure, even when its business purpose is disputed. Digital transformation is broader but still constrained by the artifacts it produces: a customer portal, a data platform, a modernized ERP. AI-Native has no comparable technical finish line. The phrase can mean almost anything, and so it comes to mean whatever the listener already believed should happen. The CFO who already wanted cost reduction hears cost reduction. The CTO who already wanted modernization hears modernization. So the phrase confirms each executive's existing priorities rather than producing new shared ones, and because that confirmation feels like alignment, the misalignment stays invisible.

That is the central problem this field guide is built around. There is no single right definition of AI-Native to hand you. Stage 2 explains why. A field guide that pretended otherwise would be misleading you. What this field guide offers is the substrate for producing your own definition, together with your team, in a way that holds across the executive group and survives contact with execution.

Everything from here on is built around that work. The frameworks are inputs to alignment. They do not substitute for it. The maturity stages are tools for honest self-location. The archetypes are options. Everything in this field guide presumes the work of becoming AI-Native is yours to do.

One more thing about the phrase on the whiteboard. The CEO who wrote it was not wrong to write it. The instinct that something is changing, that the company has to respond, that AI is not optional, those instincts are correct. The ambition is right. The mistake is assuming it translates into shared understanding without the work of translating it.

BRING THIS INTO THE ROOM

In your next leadership meeting, before any other agenda item, ask each member to write down in one sentence what they think "becoming AI-Native" means for the company. Do not let anyone speak until everyone has written. Then read them aloud. The gap between the sentences is the consensus problem made visible. That gap is the work.

Section 1.3

The Deeper Premise: AI-Native Is a Question About Human Life

What is at stake in this transition beyond strategy?

Beneath the alignment problem is a deeper question that almost no leadership team has explicitly opened.

Here is the question: what is this place going to be, for the humans who spend most of their adult lives here?

Nothing about this question is soft. It determines whether the operating model decisions a leadership team is about to make produce a place worth working in or a place that quietly hollows out the people inside it. It is also the question that AI is forcing open, whether leadership teams want it opened or not. AI changes the texture of the eight or ten hours a day that people spend at work, in ways the old conversation about work was not built to handle.

Ninety thousand hours. A person who works a typical full-time career, twenty-two to sixty-five, two thousand hours a year, will spend roughly that much of their life at work. Sit with that number. It is more time than they will spend on almost anything else except sleep. It is more time than they will spend with their spouse if they marry in their late twenties. It is more time than they will spend with their children, often by a substantial margin. For the person, whether

those ninety thousand hours are spent well is the question. Everything else in their life is shaped by it: their energy at home, their relationships, their health, what they think they are worth, what their life becomes.

Companies are the container in which most of that time happens, for most of the people this field guide is for, and the container shapes what the time can be. A company takes ninety thousand hours of a person's life in exchange for money. A company that is indifferent to what those hours do to the person is exercising enormous influence over a life without acknowledging the responsibility. Many companies manage that influence with policies and processes and performance reviews instead of care.

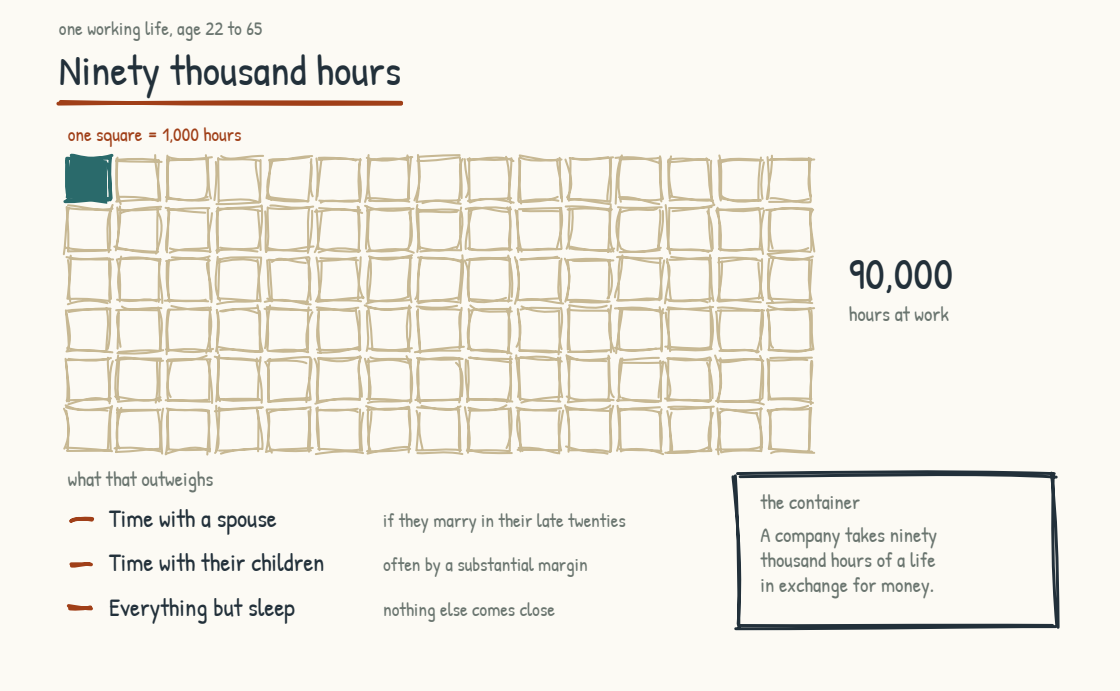

Figure 1.3a. Ninety thousand hours of a working life, one square per thousand. The company is the container it happens in, and the container shapes what the time can be.

People have made this observation, in many forms, for a long time. What is new is that AI is opening a window in which leadership teams can finally act on it, if they choose.

Here is why. Until now, the conversation about whether work was a place where humans could flourish ran into a hard constraint. The work needed to be done. Most of the work was what it was. The company's job was to extract maximum output from finite labor. Whether the work was meaningful, whether the time was well spent, whether the person was becoming more

themselves or less, all of it sat at the edges, because the operational center was already fully committed. There was no slack.

AI changes the slack. Whatever else AI does, it can take over a meaningful share of the work that previously filled those ninety thousand hours. The routine analysis. The first draft. The standard synthesis. The volume work. How much varies by role, task, organization, and the capability of the systems available at the time. The estimates will keep moving. But in a growing number of settings, some work that required human effort can now be handled or accelerated by AI. That leaves a choice that no previous moment in working life has put in quite this form.

The choice is what to do with the time and cognitive capacity that AI releases.

The default answer is to extract more output. If a person used to produce X with their hours, and now produces 1.5X with AI assistance, the company captures the 0.5X gain and asks for more next quarter. This is what I expect most companies to do, because their operating models are optimized for it and their compensation systems reward it and their boards expect it. There is nothing evil about that answer. The system was built to produce it.

The other answer is to use some portion of the released capacity to make work more humane. More meaningful, more developmental, more connected, more honoring of the fact that the people in the building are spending the largest share of their lives there. This answer requires a deliberate choice. It happens because a leadership team decides it should happen and then designs the operating model to produce it.

I expect most leadership teams to make the first one by default. The second one requires articulating values they have not had to articulate, defending against pressures that will push them back toward the default, and accepting that some of the productivity gain from AI will be reinvested in human flourishing rather than captured as margin. The second choice is harder, takes longer, and is harder to defend in any single quarter. My wager is that, over a decade, it produces companies that become genuinely different from their competitors. That is a proposition to test, not a result I can claim in advance.

This choice is the premise underneath everything that follows. Every later argument depends on it.

The maturity model in Stage 2 assesses where a company is in its AI-Native transition, but the underlying question is whether the maturity is being directed at extraction or at human flourishing.

In Stage 3 the operating system describes how processes change when AI is involved, and the design decisions inside each process come down to whether the redesign serves output or the people doing the work. The choice shows up everywhere.

The architectural decisions in Stage 4 (knowledge graphs, agent systems, collaboration surfaces) look like technology choices on the surface. They are also choices about whether the company's intelligence accumulates in shared, observable, human-comprehensible structures or in private, opaque, individually-held tools. The architecture shapes the work, and the work shapes the lives of the people doing it.

Stage 5 treats the choice most directly, though it runs through every part of the book as the spine.

Optimizing for employee happiness at the expense of business outcomes would be naive and strategically incorrect. Companies that fail to perform do not have the durability to be good places to work over the long term.

My argument is that whether work is humane and whether the company performs are not separable questions. Fear can produce short bursts of output, but sustained engagement creates conditions for learning, candor, and discretionary effort that fear erodes. Less obvious: I expect a company that captures institutional knowledge from experts who feel honored to build better AI systems than one trying to capture it from experts who feel threatened. People who trust the purpose are more likely to contribute context, exceptions, and judgment rather than the minimum they can safely disclose. And a leadership team that decides what kind of place it wants to be, and designs deliberately to be that place, is building an advantage that quarter-by-quarter extractors will find difficult to copy. The hard-eyed business case and the human flourishing case can point in the same direction, on the timescales that matter.

A leadership team that takes this premise seriously is signing up for five specific things. I name them briefly here so they hang in the air as the chapters develop.

The team has to make explicit choices about what to do with the capacity AI releases: to extract, to reinvest in people, or to do some intentional mix of both. The default choice is to extract, so

making any other choice requires deliberate work. The useful version of this question is concrete. If a person is letting go of spreadsheet management, what does that now allow them to do? You have people carrying real wisdom and domain experience, and the question is where you point them. Early on, some of that released capacity gets repurposed into managing the transition itself, because a change of this size needs people with time to focus on it. After that phase, the honest options are to explore the white space in the market you serve, to put more time into researching and designing new products, or to serve customers better than you could before. Growing the revenue, or protecting the revenue you already have. And some of it goes to a job that did not exist before, which is the human work of managing the AI.

The team has to design the operating model so that human work remains substantive, so that what is left for people to do is the work that develops them rather than the residue after AI has taken the parts it can do. When AI absorbs activities a person used to do, that person has to be pointed in a new direction or given other responsibilities. Ideally the business expands and you backfill that freed capacity. Either way it takes deliberate investment in how the role itself evolves, because a human is assigned to a role, and somebody has to answer what the evolution of that role actually is. The path of least resistance is to leave humans the leftovers and call it a redesign.

The team has to honor the people whose expertise is being supplemented by AI, in ways that are structural rather than ceremonial. You can give someone a hug and a pat on the back and tell them their job is being affected by AI. Or you can evolve the role, redesign it, sometimes eliminate it and build a new one deliberately. Compensation is part of this and it can genuinely move: a commission model might become salary, or a salary model might flip to commission. Role design, recognition, institutional knowledge capture. The honoring has to be real or the workforce will know it is not.

The team has to recognize, and then protect, the conditions under which people exercise capabilities and standing the organization still needs: deep thought, relational presence, accountable ethical judgment, original synthesis, meaning-making. Recognition comes first, because a condition nobody has named is one nobody defends. These are among the first casualties of efficiency pursued without boundaries, and holding onto them takes explicit design.

The team has to acknowledge, out loud and early, the emotional reality of what people are carrying. This is the thing that does not get talked about enough. Many workplaces were getting

better at recognizing when someone was struggling, and then AI arrived inside the jobs themselves. It can land on people's mental health. It affects anxiety and worry, because people who thought they had a career that would carry them to retirement are no longer sure what that looks like. So step into the uncomfortable part of it. Let people be heard, and let them raise concerns without being labeled as negative on AI. Let it get explored over time, rather than leaving someone feeling they have to defend why they still deserve to be here. A lot of people are worried about losing their jobs while their organizations have not yet put comparable effort into redesigning the work around them. The alternative is pretending the transformation is only an opportunity and never a loss. That pretense corrodes trust and produces resistance that leaders later misread as the obstacle.

These five asks are demanding, and I want to be honest that I have not yet watched a leadership team take on all five, because this is all still new. If I had to guess which one goes first, it is the fifth. The emotional reality is the softer side of the human, and the workplace shies away from that. Some of the third would go too, because structural honoring costs something. Drop the fifth and six months later things can quietly deteriorate. Morale drops, people get more anxious, and nobody connects it back to the thing that was never said. These asks are, in my reading, what may separate the companies that compound into something durable from the ones that do not. The wager behind this book is that companies that take ninety thousand hours of their people's lives and use AI to make those hours better will outperform companies that use AI only to extract more from them. The first kind of company may be rare. I am writing for the leadership teams that are considering becoming the first kind.

The rest of the book is built to help with the work. The frameworks, the architectures, the operating system, the practices: all of it presumes the premise just laid out. If you disagree with the premise, much of the book will read as misdirected. If you agree, or if you are not yet sure but willing to hold it as a serious possibility, this book is for you.

CONSIDER BEFORE YOU ACT

Before reading further, sit with the five asks for ten minutes. Which of them is your leadership team genuinely ready to take on? Which would be hardest? Which would feel impossible? The book that follows assumes the asks are real and the reader is willing to engage them.

CONSIDER BEFORE YOU ACT

If you keep three things from Stage 1: 1) AI-Native is an alignment problem dressed up as a technology problem; the gap between your executives' private definitions is the work. 2) Design belongs at the front of the transition, and the new economics of design have made that work more affordable. 3) Underneath the strategy sits the ninety-thousand-hour question: what will this place be for the humans who spend most of their adult lives here?

STAGE 2

What AI-Native Actually Means

The postures, archetypes, layers, and stages that distinguish AI-Native from its imitations.

Section 2.1

Three Postures: Augmented, First, Native

What does it mean to design a company around AI?

AI-Native is currently being used to mean three different things, and the conflation produces operating-model decisions that do not match the company's actual posture.

The first posture is AI-Augmented. AI is a feature added to existing products and workflows. The org chart, the operating model, and the processes are unchanged. AI sits on top, providing assistance to people doing the same work they were doing before. Copilots inside Microsoft 365. Embedded assistants in Salesforce. AI-powered search inside legacy enterprise tools. These are real and often valuable. They are also, structurally, the lightest touch a company can have with AI. AI-Augmented is a feature strategy.

The second posture is AI-First. AI is a strategic priority and a core capability. The organization recognizes that AI changes what is possible and is rewiring itself to take advantage. Workflows are being redesigned. Roles are evolving. Investment is real. But the underlying operating model still assumes pre-AI logic. The company was designed before AI existed and is being modernized rather than rebuilt. Many large incumbents that are serious about AI are in this posture,

whether they call themselves AI-First or AI-Native. Real work is happening. The transformation is partial.

The third posture is AI-Native. The operating model assumes from the ground up that AI will participate in work alongside people, applications, and data. Intelligence is embedded in workflows rather than layered at the edges. AI agents are treated as architectural components with clearly defined roles and operating limits. Decisions about how much autonomy AI has, what data it can use, and where human review is required are made at the design stage, before systems reach production. IBM's working definition captures it: AI-native means designed from the ground up with AI as a core component rather than bolted on as a feature.

The distinction matters because the three postures call for different investments, different organizational structures, and different time horizons. An AI-Augmented company that announces it is becoming AI-Native without changing its operating model is producing what I have come to call performative AI-Nativeness: the language without the substance. The phrase makes the earnings call in February, and a banner goes up in the lobby. By August the night shift is still keying readings from one system into a second one by hand, the work order still crawls through the same approval chain, and the one visible change is a chat window docked at the corner of the operator's screen. The workforce can tell. The board eventually figures it out. The market figures it out faster.

A useful analogy: traditional organizations operate like assembly lines. Each step is predefined, sequential, and dependent on human intervention. AI-Augmented organizations are assembly lines with smarter tools at each station. AI-First organizations are reconfiguring the assembly line. Same factory, new layout. AI-Native organizations operate more like navigation systems. They interpret inputs, adjust in real time, and guide outcomes. Some work will remain sequential and standardized; the difference is that the operating model no longer assumes every important path can be fixed in advance.

Now the honest reset. Most established companies will not become AI-Native in the pure startup-built-from-scratch sense. The company born this year with AI in its DNA from day one has structural advantages an incumbent cannot fully replicate. The legacy systems, the institutional habits, the workforce hired for pre-AI roles, the customers accustomed to pre-AI experiences: these are all real and they all create friction that a startup does not have. As a planning assumption, an incumbent might move from AI-Augmented to AI-First over several years of serious work. Moving from AI-First toward a deeply AI-Native operating model can take

much longer and may never have a clean finish. Eighteen to thirty-six months and a decade are useful orders of magnitude for planning, not industry benchmarks or promises.

INTELLECTUAL LINEAGE

Clayton Christensen offers a useful lens on the three postures, but not an exact mapping. Christensen distinguished sustaining innovation, which improves established offerings along dimensions existing customers value, from disruption, which begins through a new-market or low-end foothold and develops along a different value network. AI-Augmented often behaves like sustaining innovation because it improves the existing operating model. AI-First can still be sustaining even when the rewiring is substantial. AI-Native may create a different operating model, cost structure, talent profile, and way of serving customers, but that alone does not make it disruptive in Christensen's precise sense. It becomes disruptive only when the market path and value network fit the theory too. The Innovator's Dilemma still illuminates the section's honest reset: existing customers, profitable models, legacy investments, and organizational habits create gravity toward improvements the incumbent already knows how to value. Startups may have less of that gravity, though they carry constraints of their own.

Christensen, Clayton M. (1997). The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail. Harvard Business School Press.

The most valuable transition for most incumbents is the move from AI-Augmented to AI-First, and from AI-First to a deeply integrated, near-AI-Native operating model. MIT CISR research reported that enterprises in the first two stages of its AI maturity model had financial performance below their industry average, while enterprises in stages three and four were above it (Weill, Woerner, and Sebastian, 2024). The underlying survey covered 721 companies in 2022, so the results primarily describe analytical AI, and the relationship is an association rather than proof that AI maturity caused the performance difference. Follow-up research has continued to locate the largest financial shift around the move from pilots to scaled ways of working. The practical point is not to become a platonic AI-Native startup. It is to build the capabilities that move AI from scattered tools and pilots into the way the enterprise works.

This book is written primarily for the leadership team of an established company that is serious about AI-First and aspirational about AI-Native. Some of what I describe applies more cleanly to startups designing from scratch. Most of it is calibrated for the harder problem: taking an organization that exists and moving it through real change, in the real world, with real people whose lives and livelihoods are part of the equation.

The spectrum tells you your current posture. It does not tell you what version of AI-Native you are aiming to become. Within the AI-Native end of the spectrum, there are at least five distinct legitimate models, and the choice among them shapes everything the operating model becomes.

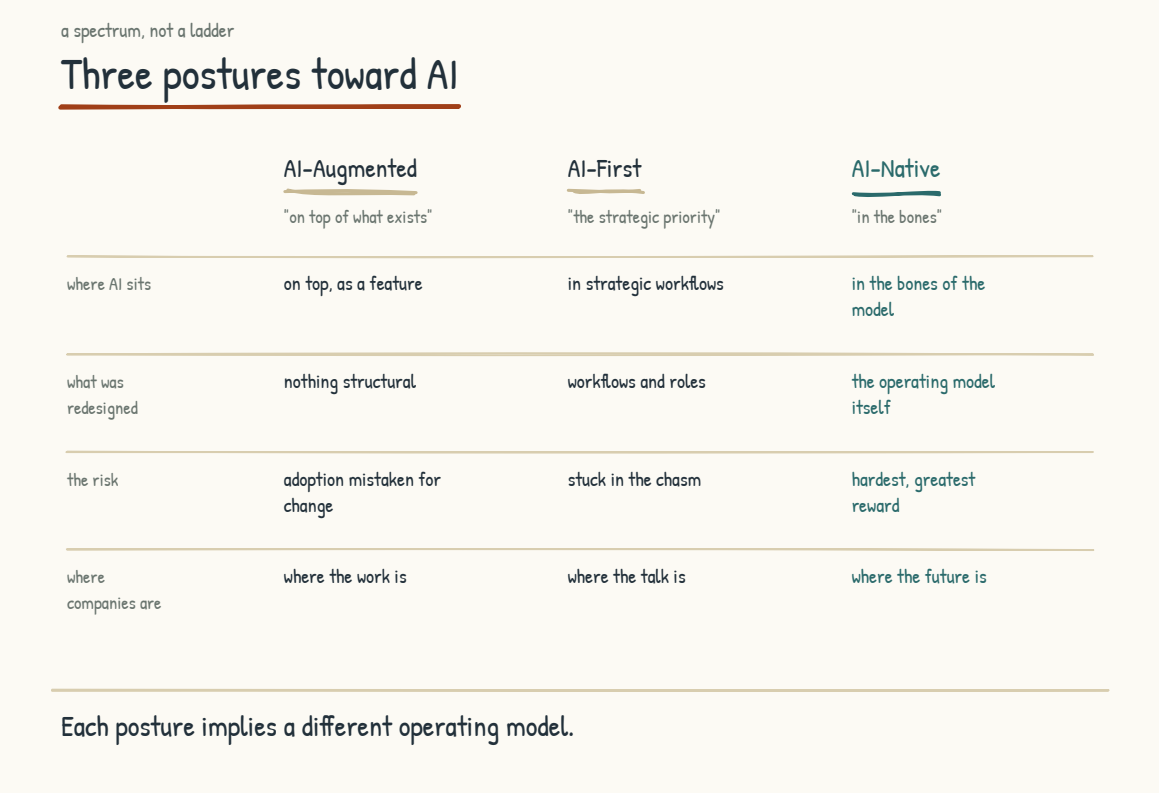

Figure 2.1a. The three postures toward AI compared across four dimensions. The boundary between AI-First and AI-Native is where most companies stall.

CONSIDER BEFORE YOU ACT

Which posture does your company actually occupy? Set aside the one that was announced and the one the strategy deck claims. The honest answer may be one posture earlier than the public one. If you cannot find a single example of the operating model itself being redesigned around AI, you are AI-Augmented regardless of what the announcement said. Honest self-location can prevent pilots from being asked to carry a transformation they were never designed to deliver.

Section 2.2

The Archetypes: There Is No AI-Native Company, There Are AI-Native Companies

What different shapes can an AI-Native company take?

The standard discourse treats AI-Native as if it were a single destination. It isn't. Within the AI-Native end of the spectrum, there are at least five distinct legitimate models of what an AI-Native company can be, and they produce dramatically different cultures, decision-making patterns, and risks.

The dimension that distinguishes them is structural. Where does the intelligence sit in the operating model, and where do humans sit relative to it? The five archetypes answer that question five different ways, and the answers are not interchangeable. A company that chooses the wrong archetype for its business will produce friction, attrition, and underperformance that look like execution problems. They are architecture problems.

I walk through each archetype with the same structure: where the intelligence sits, where humans sit, what the archetype is good for, where it breaks, and a real-world reference point. After the five, I name two variants, the hybrid portfolio and the anti-archetype. Most real companies end up in one of these rather than in a pure archetype.

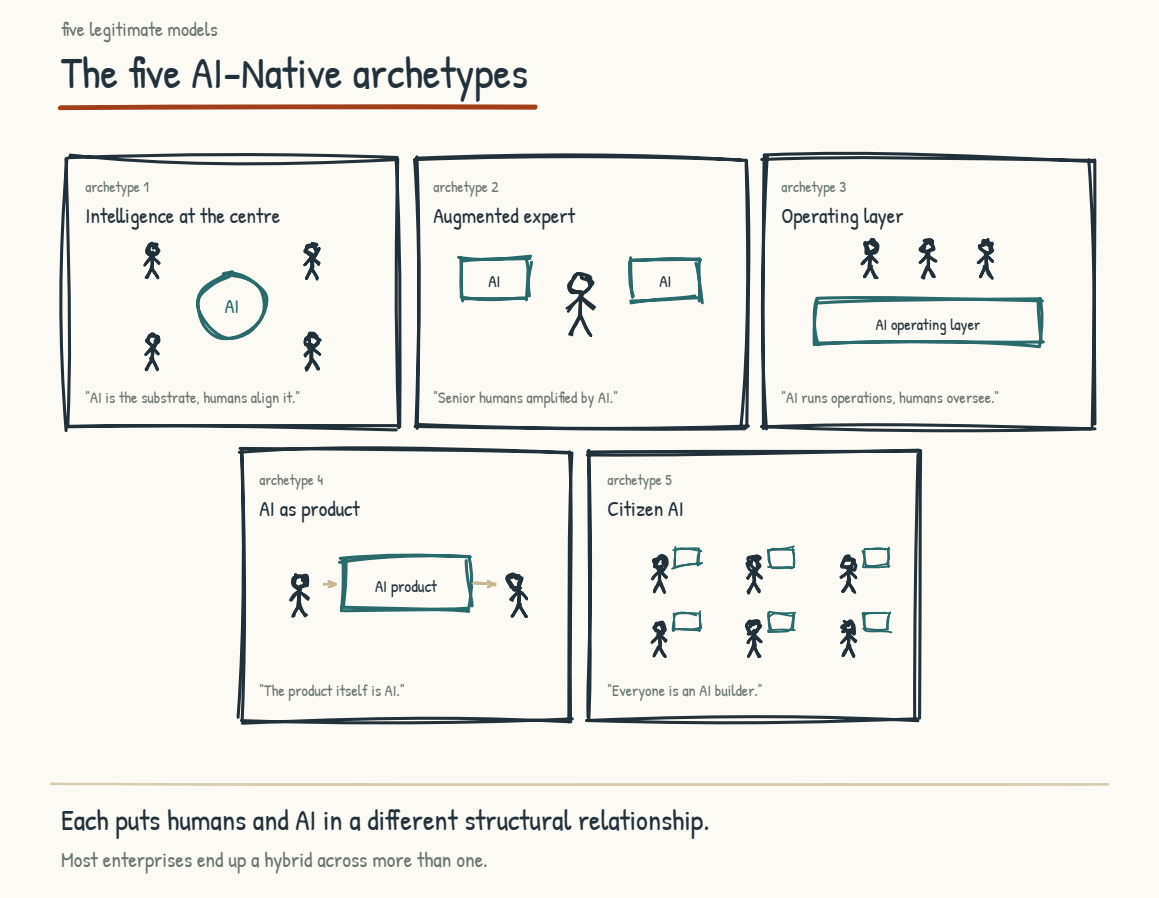

Figure 2.2a. The five archetypes at a glance: where intelligence sits and where humans sit relative to it. Most large enterprises end up a hybrid across more than one.

Intelligence at the Center (the "God Model")

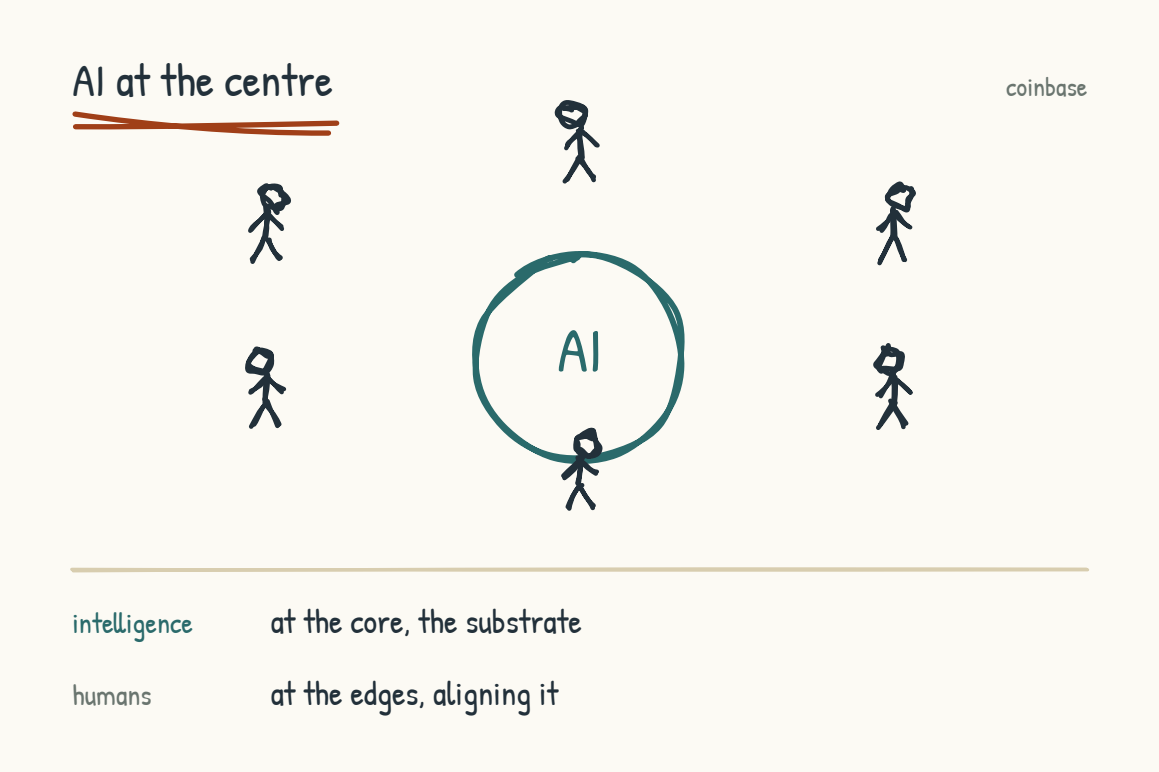

Figure 2.2b. Intelligence at the Center: AI is the substrate, humans align it from the edges.

Where the intelligence sits: at the architectural center of the company. AI is the substrate. The operating model is built around feeding, directing, and aligning the AI.

Where humans sit: at the edges. Aligners, validators, directors, prompters. Player-coaches rather than pure managers. The organizational structure is designed to flatten, with smaller teams and, at the limit, experiments in single-person pods orchestrating fleets of agents. Whether those experiments become a durable operating model remains an open question.

The public reference point is Coinbase. On May 5, 2026, CEO Brian Armstrong announced a fourteen-percent workforce reduction, approximately seven hundred employees. Coinbase's SEC filing described two purposes: responding to market conditions and optimizing operations for the AI era. Armstrong's public memo framed the goal as rebuilding Coinbase to be lean, fast, and AI-native. The announced design capped the org chart at five layers below the CEO and COO, called for managers to act as player-coaches rather than pure managers, and proposed experiments with smaller "AI-native pods," including one-person teams spanning engineering, design, and product responsibilities with agent support. Those are announced changes and

experiments, not yet evidence that the model works. Block moved in a similarly aggressive direction in February 2026, reducing its workforce by more than forty percent while describing a smaller, flatter organization enabled partly by AI and automation.

Coinbase had posted a $667 million net loss in Q4 2025, while reporting that total quarterly revenue of $1.8 billion was down five percent from the preceding quarter. The company nevertheless reported full-year revenue growth and substantial cash reserves. The crypto market downturn and cost pressure therefore belong alongside the AI framing, without reducing the restructuring to either story alone. Both things can be true. The structural changes Armstrong announced are real, the AI-native framing may be sincere, and the financial pressure was also real. A book that presents Coinbase as a pure AI-native pivot without naming the financial context is missing half the picture.

What the archetype is good for: high-volume, high-velocity, digital businesses where speed and unit economics dominate. Software companies, fintech platforms, content production at scale, some forms of e-commerce. Businesses where the work is mostly amenable to AI orchestration and where the competitive moat is execution speed.

Where it breaks: operations-heavy environments with physical reality and complex interdependencies. Regulated environments where human accountability for decisions is non-negotiable. Contexts where institutional judgment under genuine ambiguity is the whole game. Any environment where the institutional knowledge has not yet been captured into systems the AI can reason over (without that knowledge, the AI is operating on shallow context and producing confident but wrong outputs at scale). Any context where workforce trust matters and the announcement reads, accurately, as "we are replacing you."

The honest critique of this archetype: it is currently producing a lot of press and short-term market enthusiasm, but the long-term track record is unproven. The companies moving aggressively into this model are betting that AI capabilities and redesigned workflows will compensate for at least some of the institutional knowledge and coordination capacity removed with people and management layers. That bet may succeed. It may also produce, over the next several years, companies that are structurally lean, technically sophisticated, and operationally fragile in ways that become visible only under stress. We do not know yet. The leadership teams choosing this archetype should know they are making a real bet rather than adopting the inevitable future.

Augmented Expert

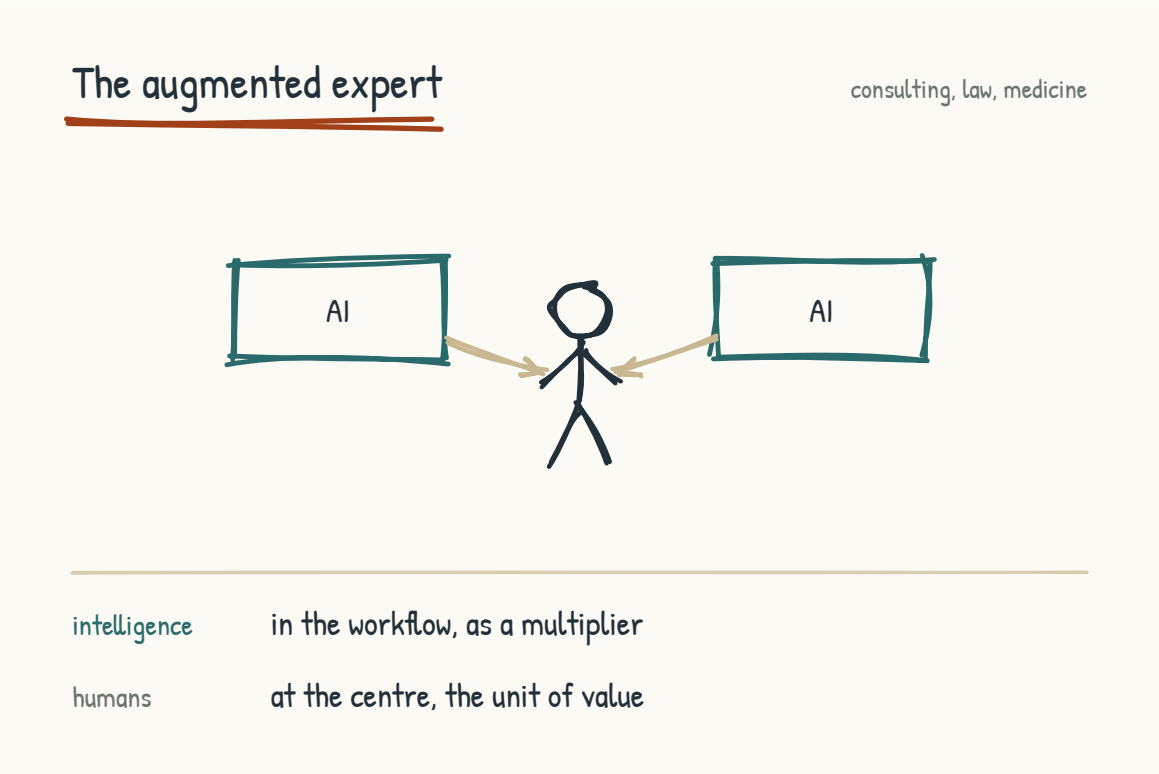

Figure 2.2c. Augmented Expert: the human stays the unit of value, AI is the tool they wield.

Where the intelligence sits: distributed across roles, with AI showing up in workflows but not owning them.

Where humans sit: at the center. The expert is the unit of value. AI is the tool the expert wields. Each individual produces more, sees more, acts faster. But the human is still the actor.

What the archetype is good for: professional services, advanced engineering, R&D-heavy organizations, law, medicine, deep enterprise sales. Anywhere the value is concentrated in specialized human judgment and the AI's role is to give that judgment more reach and depth. Many large consulting firms describe themselves this way. Many ambitious AI deployments in healthcare, legal, and engineering are operating in this archetype.

Where it breaks: it can become a sophisticated form of AI-augmented stagnation if leadership does not push it. The path of least resistance is to give every expert a Copilot license and call that the transformation. Nothing about the operating model, the org chart, or the work has changed. Your expert is now slightly faster at producing the same kind of output, and you have missed the

structural opportunity. Competitive advantage erodes, because every other firm in the industry is doing exactly this.

This archetype asks the firm to be explicit about what the expert can now do that they could not do before. Qualitatively different work, beyond faster work. New service offerings. New depths of analysis. New kinds of judgment exercised at higher altitude because the AI is handling lower altitudes. Without that explicit reframing, the archetype produces incremental productivity without strategic differentiation.

Operating Layer (the AI-Native Operations Model)

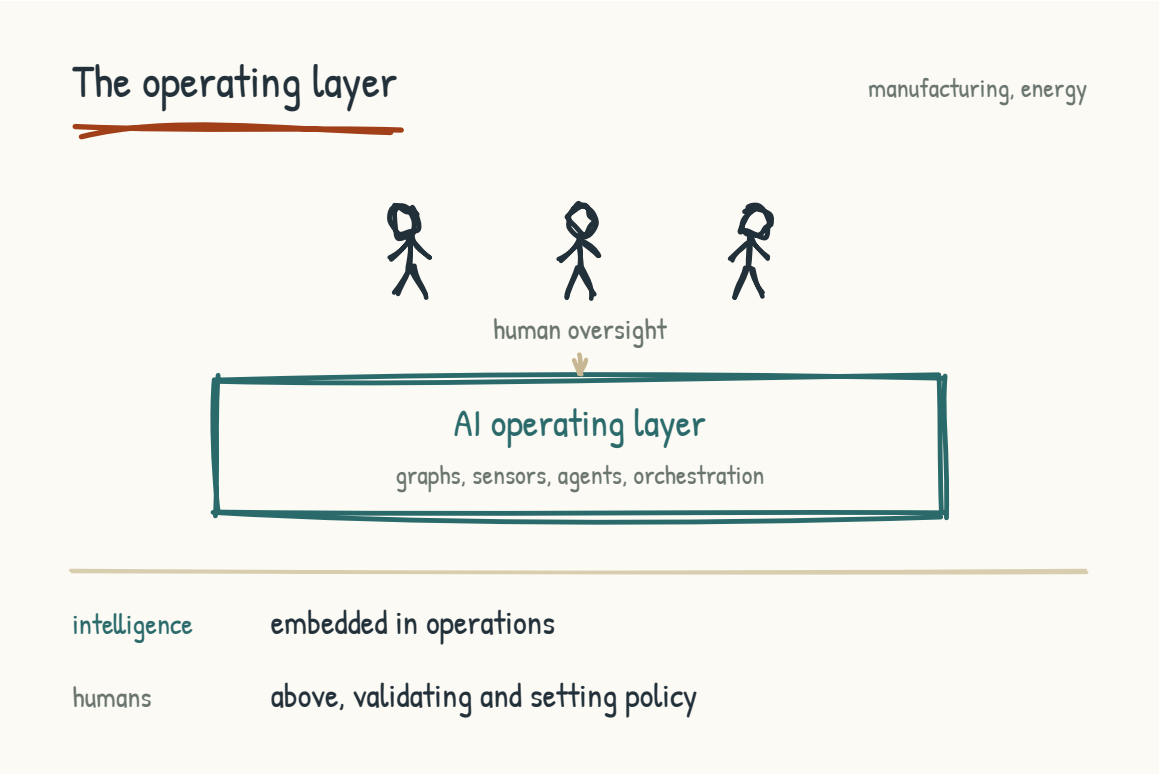

Figure 2.2d. Operating Layer: AI runs operations as an always-on layer; humans validate and set policy above it.

Where the intelligence sits: embedded in operations. AI is the always-on layer that watches systems, predicts states, recommends actions, surfaces anomalies, captures institutional knowledge into reasoning structures.

Where humans sit: above and around the operating layer. Operators validating AI recommendations at the moment of decision. Domain experts shaping how AI reasons by teaching the systems. Leaders setting policy and remaining accountable for outcomes. The org

structure thins in middle management and thickens in domain expertise and judgment roles. At 2 a.m. in the control room, the operating layer flags a feedwater pump whose vibration signature has drifted over three shifts. The operator on duty has twenty-two years on that unit. She traces the reasoning back through the maintenance history, checks it against what she felt on her walk-down an hour earlier, and schedules the shutdown for Thursday, three days before the machine would have run itself to failure. The AI surfaced it. She decided it.

What the archetype is good for: manufacturing, energy, defense, healthcare, logistics, hospitality at scale, complex supply chains. Operational environments where reliability and trust are existential, where the cost of an error is high, and where much of the knowledge that makes the operation work lives with experienced people who are aging out of the workforce. This archetype can fit many operations-heavy companies. It also receives less useful attention in the dominant AI-Native discourse, which is often written from knowledge-work environments rather than operational ones.

Where it breaks: if the knowledge architecture is weak, the AI may never become trustworthy enough for consequential work. If the human-AI handoffs are not designed deliberately, operators may stop trusting recommendations and quietly revert to working around the system. If institutional knowledge is not captured before the experts retire, the AI is built on sand and the operating layer becomes fragile at the moment the company most needs it.

This archetype asks for investment in knowledge architecture before investment in agent capability. Without the structured representation of how the operation works, including the entities, the relationships, the dependencies, and the soft knowledge that experienced operators carry, the AI is just a chat interface bolted onto a control system. With that structured representation, the AI becomes a reasoning partner that augments the operator without replacing the judgment.

AI as Product

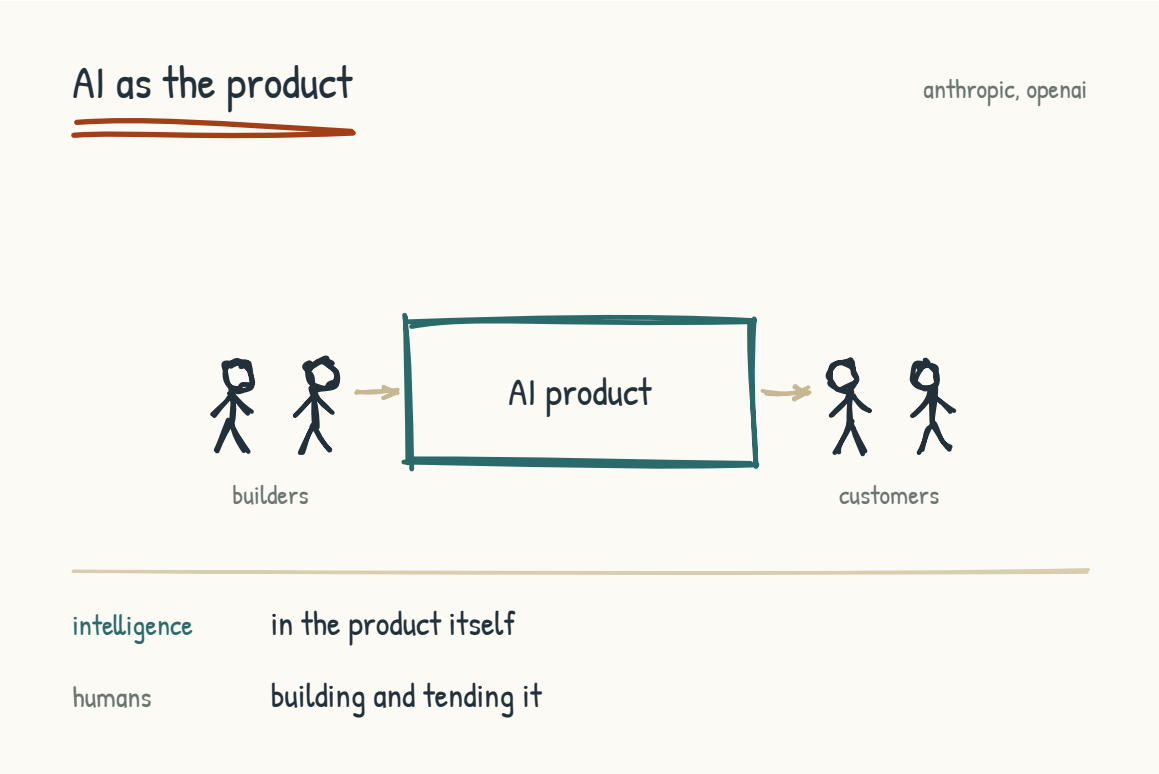

Figure 2.2e. AI as Product: humans build and tend the AI, and the AI is what reaches the customer.

Where the intelligence sits: in the product. The company's value to customers is mediated by AI. The internal operating model still has to be AI-Native to some degree, but the strategic question is "how do we build, deploy, and improve AI-powered products" rather than "how do we use AI to run ourselves."

Where humans sit: building and tending the AI. Domain experts shaping its behavior. Operators monitoring it in production. Customer-facing teams interpreting it. The work is the AI work.

What the archetype is good for: SaaS companies pivoting their products toward AI-native experiences, AI-first startups, companies whose competitive position depends on AI capabilities customers can feel directly. The companies in this archetype have settled "should we use AI internally." Their question is "how do we ship better AI to our customers faster than our competitors do."

Where it breaks: conflating "we sell AI" with "we are AI-Native internally." Many AI product companies have impressive AI in their product and operational chaos in their back office.

Building responsible, governed, continuously improving AI products demands more inner-loop discipline than most companies realize, and the gap between the product surface and the operational reality widens past what is sustainable.

Citizen-AI (Democratized Capability)

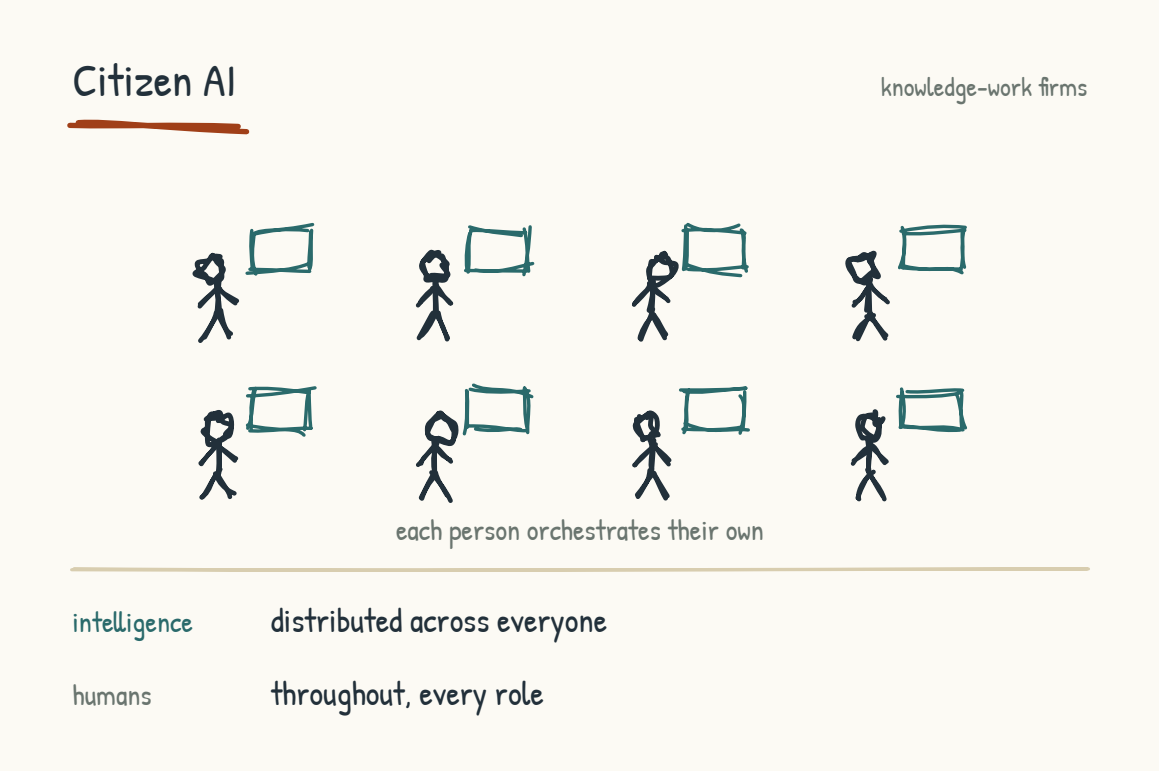

Figure 2.2f. Citizen-AI: capability is distributed, and every person orchestrates their own AI.

Where the intelligence sits: distributed across the workforce. Every employee has AI capability. Every team builds its own AI-powered workflows. The center provides platforms, governance, and shared services. The bulk of AI innovation happens at the edge, by domain practitioners who know their own work best.

Where humans sit: throughout. AI is woven into every role rather than concentrated in any. The bet is that a thousand employees each making their work modestly better with AI produces more total value than a centralized AI strategy could deliver in the same time. Microsoft's pitch for enterprise Copilot at scale is the clearest commercial articulation of this archetype.

What the archetype is good for: large, diverse organizations where the operating model is too varied for a single AI strategy. Knowledge-work companies where the work itself is

heterogeneous. Companies with strong existing platform discipline that can deploy shared services across many different use cases.

Where it breaks: this archetype is especially exposed to the "thirty islands" problem. Thirty employees, each with their own AI tools, each producing their own AI-mediated output, little of it compounding into team-level or organization-level intelligence. Without strong platform discipline and a deliberately designed shared surface for AI work, the kind Section 4.3 takes up in full, Citizen-AI can produce fragmentation rather than capability. Shadow AI and governance become substantial concerns because use is broadly distributed. Companies that handle the pattern well invest in shared knowledge layers, reusable instruction assets, and review surfaces. Companies that do not may end up with more individuals using AI while the organization gains little shared capability.

The Hybrid Portfolio

Most real companies, especially large ones, will not be a single archetype. They will be a portfolio. The operations division runs as Operating Layer. The product group runs as AI as Product. The corporate functions run as Augmented Expert or Citizen-AI. The leadership challenge becomes coherence across the portfolio rather than choosing one model.

For most enterprises with more than a few thousand employees and more than one line of business, the hybrid portfolio is the realistic destination. It asks you to name each part of the company with its archetype, design the operating model for each part accordingly, and manage the interfaces between archetypes deliberately. A common failure mode is to apply one archetype's logic uniformly across the whole company, which produces serious mismatches. Running operations like a citizen-AI environment, or running professional services like an intelligence-at-the-center model, will not work. It will also not stop being broken because leadership wishes it would.

The Anti-Archetype

Some companies will deliberately not become AI-Native in any of these senses, and will make their value proposition the human, the artisanal, the relational. Healthcare already contains versions of this choice, even where AI supports diagnosis, planning, or administration: the accountable clinical relationship remains human. The premium artisanal coffee shop where you know the barista. The human-only therapy practice. The white-glove law firm that markets "your lawyer, never an AI." The luxury concierge service whose entire promise is unmediated

human attention. These are real strategic choices and they are probably underestimated in the current discourse.

Some industries may bifurcate into an AI-Native commodity tier and a deliberately human premium tier, with the middle struggling. Companies positioning coherently at either pole may outperform ones caught between them. That is a strategic hypothesis, not yet a general market result.

Not becoming AI-Native is a legitimate strategic choice for some companies. It reflects a different bet about where value concentrates rather than a failure of vision.

Defensibility within the Archetypes

Choosing an archetype is the first strategic decision. Making the chosen archetype defensible is the second.

The defensibility question matters more in the AI-Native era because access to capable models and the ability to prototype software are becoming widely available. Frontier models are not interchangeable commodities: capability, reliability, cost, deployment options, and policy still differ. But access to GPT-class, Claude-class, Gemini-class, and open-weight systems is no longer scarce enough to be a moat by itself. Nor is a working software prototype. AI-assisted development can compress prototyping dramatically, while secure operations, distribution, workflow fit, support, and sustained product quality remain difficult. Some old moats around proprietary software and exclusive model access are therefore weakening rather than disappearing.

Four recurring sources of defensibility are visible. They are not an exhaustive taxonomy and none is automatic. They can compound, and the strongest businesses often combine more than one.

The proprietary data moat. The company has accumulated data over years of operation that no competitor can replicate. The data is operational (customer transactions, equipment telemetry, support conversations) rather than something that can be bought. The AI built on top of that data is therefore distinctive in ways that the models, in isolation, are not. The venture firm CRV has made the same argument in its vertical AI thesis: the strongest vertical AI companies accumulate data through customer relationships that no amount of compute can replicate. Vertical AI companies have a structural edge here that horizontal AI products cannot match.

The encoded expertise moat. The company has captured the institutional knowledge of its experienced workforce into systems built for AI reasoning. The knowledge graph, the decision history, the named failure patterns, the soft judgments that experienced operators make without thinking. Encoded expertise can be harder to replicate than data because data can sometimes be purchased; encoded expertise has to be elicited, tested, maintained, and connected to the conditions under which it applies. Competitors that do not have access to the people and operating history behind it cannot reproduce it merely by licensing the same model.

The regulatory moat. In regulated industries, the cost of operating an AI system that meets the regulatory bar can be high. Startups and incumbents alike have to prove the controls, records, oversight, and validation their setting requires. Companies that treat compliance as product work rather than overhead may turn that investment into a barrier to entry, though compliance itself is not evidence of product quality. A 2026 Bessemer Venture Partners case study reports that Graph AI built its pharmacovigilance system around SOC 2, FDA 21 CFR Part 11 controls, immutable audit trails, electronic signatures, human sign-off, and obligations it identified under the EU AI Act. That is a useful documented example of the design pattern; the case study is an investor profile, not an independent regulatory determination.

The workflow integration moat. The company has built AI capabilities that are integrated into the customer's or operator's actual work rather than standing alongside it. The integration creates switching costs: ripping out the AI means redesigning the work around the absence. IBM names this as the intelligence moat: mature AI systems are difficult to copy because intelligence is embedded into workflows rather than features. The improvement creates compounding returns that traditional software does not achieve.

The four moats are not equally available to every archetype. Intelligence-at-the-Center companies often lean on workflow integration and proprietary data. Operating Layer companies can lean on encoded expertise and regulatory capability. AI-as-Product companies often combine proprietary data with workflow integration. Augmented Expert companies may struggle to create organizational defensibility unless the firm redesigns the work to capture institutional knowledge structurally. Citizen-AI may be especially vulnerable when individual gains remain isolated: the thirty-islands failure from the archetype description, restated as a moat problem. These are tendencies to test against a particular business, not rankings built into the archetypes.

A useful diagnostic, when evaluating any AI-Native strategy: ask what specifically gets harder for a competitor as your AI-Native operating model matures. If the answer is that your software gets faster or your model gets better, the moat is weak. If the answer is that your data compounds, your experts teach the system, the regulators give you a multi-year head start, the customers' workflows depend on you (and you can name the specific mechanism for each), then the moat is real. Most companies in the current AI-Native conversation cannot name the mechanism. The companies that can name it are the ones whose AI-Native positioning will last.

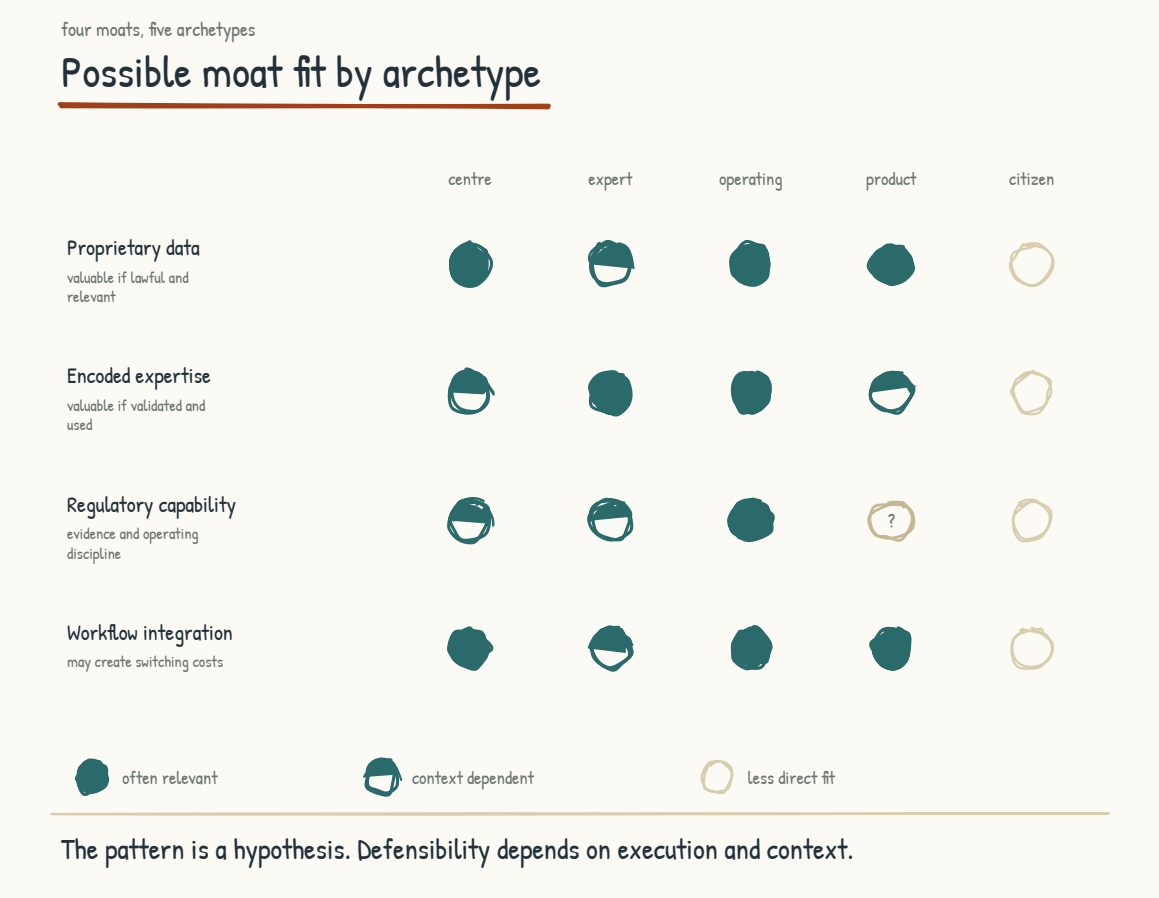

Figure 2.2g. Defensibility: four potential moats across the five archetypes. Operating Layer may have unusually strong access to encoded expertise and regulatory capability, but none of the moats arrives by design alone.

The Position Statement

There are AI-Native companies, plural. The question is which one you intend to be.

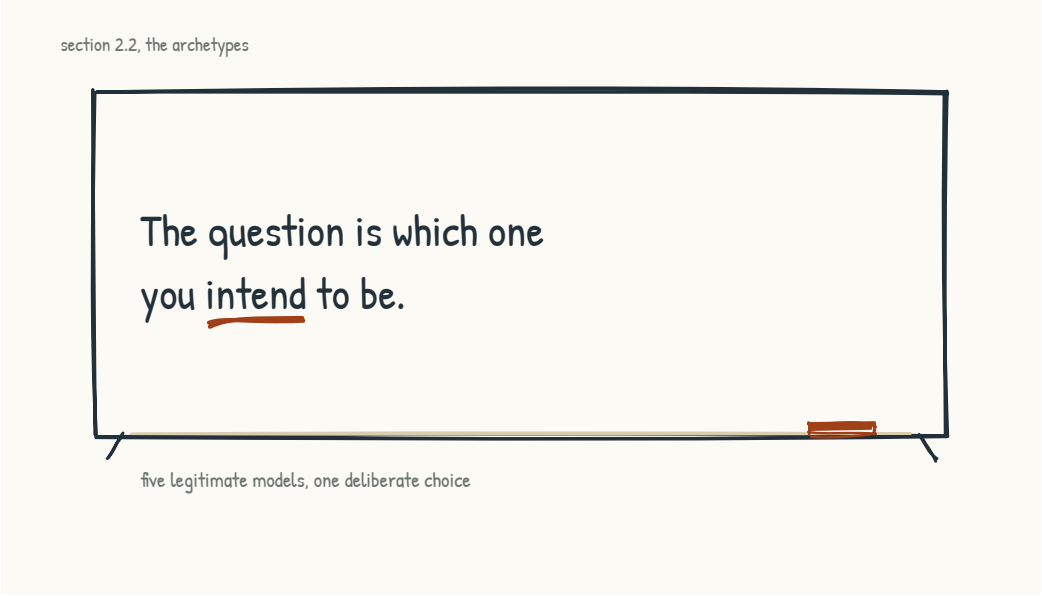

Figure 2.2h. Five legitimate models. The choice among them is the strategic act.

Choosing deliberately is one of the most consequential strategic decisions a leadership team makes, and it is rarely made deliberately. It usually emerges by default, shaped by which executives are loudest, which vendors get traction, which pilots happen to succeed, and which department happens to have budget. The result is a company that is mostly Augmented Expert with a Citizen-AI overlay and pretensions toward Intelligence at the Center. None of it is coherent. All of it is consuming budget. The spending is on schedule. The leadership team cannot tell why the transformation is not producing the results they expected.

The deliberate version requires the leadership team to look at their business honestly. Their customers, their talent, their competitive position, their regulatory environment, their tolerance for risk. And choose. The choice is constrained. Not every archetype is available to every company. But the choice within the available options is consequential, and it is the leadership team's to make.

This book's job is to give you the substrate for that choice: the frameworks, an honest analysis of each archetype, and a structure for deciding that honors the human reality of the choice. None of it substitutes for the leadership team doing the work.

The next question, having chosen an archetype, is what technical foundation will support it. That is the AI Stack.

Section 2.3

The AI Stack: AI Is Not Chat

What is AI, beyond the chat box?

Many leadership teams I work with operate with a mental model in which "AI" means "the chat thing my employees use." That model is wrong. It is also producing under-investment and mis-investment in the companies I see. The chat layer is one of nine distinct layers in a working AI stack. A company investing only in chat is choosing one of the most fragmented, lowest-return, least governable forms of enterprise AI, and mistaking the front door for the building.

The reframe matters because the choice of archetype from Section 2.2 is partly enabled, and partly constrained, by which layers of the stack a company funds. An Operating Layer archetype that has not invested in structured knowledge representation cannot function. A Citizen-AI archetype that has not invested in a shared collaboration surface produces fragmentation rather than capability. An Intelligence-at-the-Center archetype that has not invested in agent orchestration is just chat with confidence.

The stack has nine layers. Each has a different purpose, a different cost structure, a different governance profile, and a different organizational implication. I will walk through them from the foundation upward.

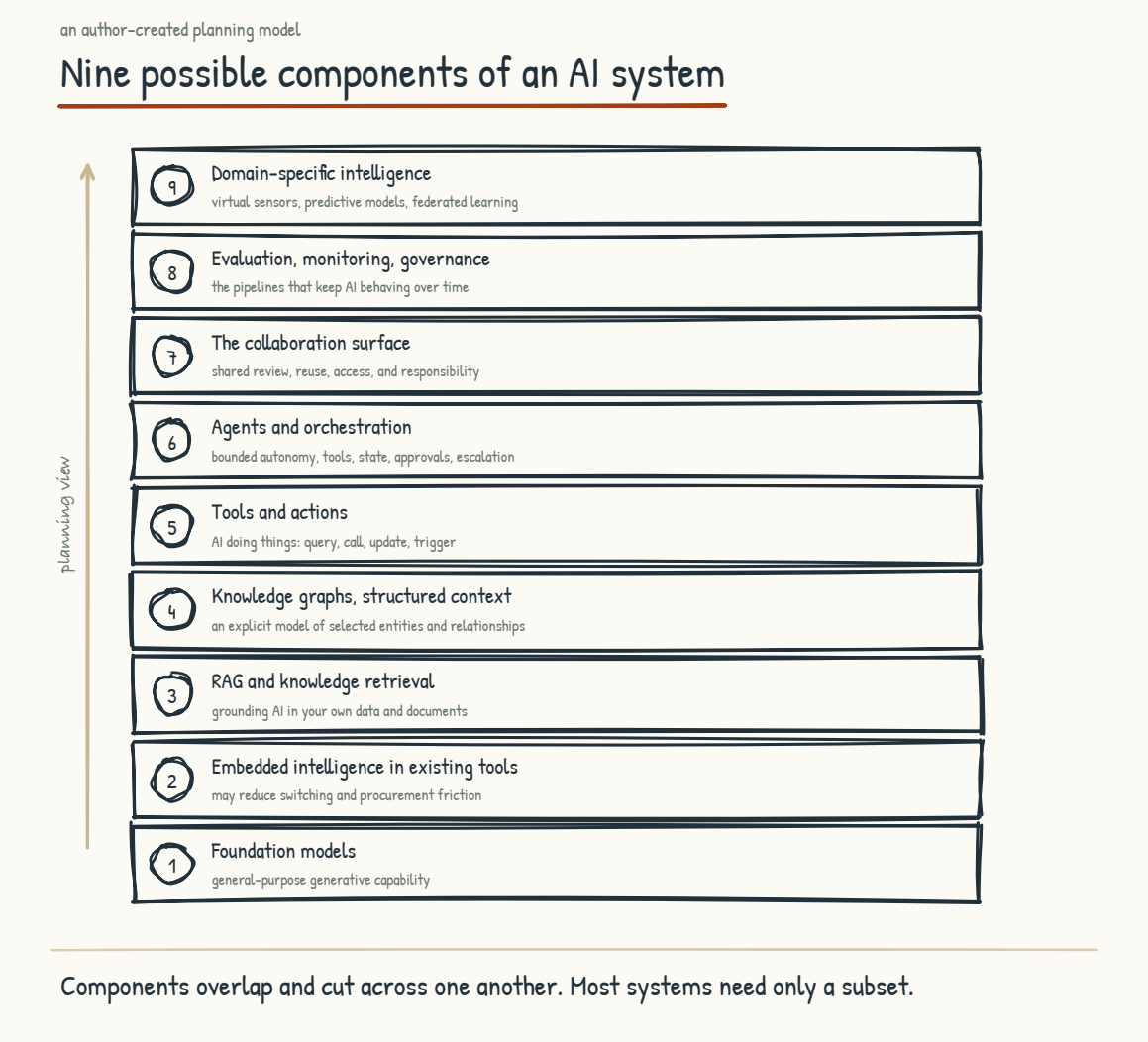

Figure 2.3a. The nine layers of the AI stack, from foundation models at the base to domain-specific intelligence at the top.

Layer 1. Foundation Models

The raw intelligence. Claude, GPT, Gemini, open-source models. These are the substrate. Most companies interact with foundation models through chat interfaces, but the model itself is the underlying capability that the chat interface exposes. Strategy at this layer comes down to model selection (which models for which use cases), model access (direct API, through cloud providers, on-premises for sensitive applications), and model diversity (do we use one model or several). Most companies under-think this layer. They use whatever model their primary vendor surfaces and call that their AI strategy.

There is a second reason to under-think it, and by 2026 it is the better reason. The frontier has become more competitive. Leadership changes by model, benchmark, workload, price, latency,

and release date; capable open-weight models have narrowed gaps that once looked structural. That does not make the models interchangeable. It makes permanent allegiance to one leaderboard winner a weak strategy. Mature architectures increasingly preserve the option to route different work to different models and to change providers when economics, policy, or performance changes. The strategic question has widened from which model to what you have built around the models you can use, because model access is rented and replaceable more often than the surrounding data, evaluations, permissions, and workflows are.

Layer 2. Embedded Intelligence in Existing Tools

AI features inside the productivity and business tools your people already use. Copilot in Microsoft 365. Gemini in Google Workspace. AI features in Salesforce, Notion, Figma, ServiceNow, Slack. This layer is often the highest-adoption, lowest-friction layer because it usually requires less tool change and workflow redesign than a standalone deployment. It still requires judgment, training, and governance. Your people are already in the tool; the AI shows up where they are.

The trap at this layer is that adoption is not the same as value. Employees using Copilot to write emails faster is genuine productivity that falls short of transformation. The risk is that an organization invests heavily at Layer 2 and concludes it has become AI-Native because adoption metrics are strong. The emails do get written faster. Adoption is necessary. It is not sufficient.

Layer 3. RAG and Knowledge Retrieval

Grounding AI responses in your organization's own data. Documents, wikis, systems of record. Retrieval-Augmented Generation, in the technical vocabulary. This is what turns a generic AI assistant into "an assistant who knows our company." Without this layer, the AI is operating on whatever it learned during training plus whatever the user types in the prompt. With this layer, the AI can answer questions about your specific policies, your specific customers, your specific products, your specific institutional context.

Strategically, this layer comes down to which knowledge gets exposed to which AI surfaces, and how. Technically, it comes down to embedding models, vector databases, retrieval strategies, and the relentless work of keeping the indexed knowledge current. For many companies, the first serious AI investment after foundation models is at Layer 3. That investment can be worthwhile, and the trap is the same one waiting at Layer 2. A company can confuse a working RAG system with a transformed operating model. RAG is often useful infrastructure; it is not the transformation itself.

One thing about this layer changed shape recently and is worth knowing before you buy it. The familiar retrieval pipeline chops documents, embeds them, fetches likely matches, and hands a fixed candidate set to the model. Agentic retrieval adds an iterative path: the model can decide what to search for, inspect what it finds, navigate within documents, and search again. Microsoft Research reported substantial gains on three open benchmarks in May 2026, including a 21.8-point recall@1 improvement over its strongest embedding baseline on BRIGHT. Those are benchmark results, not a universal production guarantee, and iterative retrieval adds latency and token cost. The practical architecture is often a router: use the inexpensive single-pass path when it is enough and reserve the agentic path for questions whose complexity earns the cost.

Layer 4. Knowledge Graphs and Structured Reasoning

Beyond document retrieval. An explicit, structured model of how the entities, relationships, and dependencies in your business connect to each other. Unlike a wiki, a knowledge graph is a graph of nodes and edges representing the actual structure of how your operation works: which system feeds which other system, which expert holds which kind of knowledge, which process depends on which other process, which decision is constrained by which regulatory requirement.

The distinction between RAG and knowledge graphs matters because they represent context differently. RAG retrieves passages or records relevant to a question. A knowledge graph represents named entities and relationships that a system can traverse, query, and combine with other evidence. Neither reasons by itself; the model and surrounding software perform the reasoning, and many strong systems use retrieval and graphs together. For operations-heavy environments (manufacturing, defense, healthcare, energy, complex logistics), graphs can help preserve dependencies, provenance, constraints, and exceptions that are difficult to recover reliably from document similarity alone. Documents carry narrative and nuance. Graphs carry explicit structure. Operational AI often needs both.

This is where Nodalix lives architecturally. It is also where the Operating Layer archetype from Section 2.2 most depends on getting the foundation right.

Layer 5. Tools and Actions

AI doing things. The AI can query databases, call APIs, update systems, trigger workflows, place orders, send messages, file tickets. The technical term is "tool use" or "function calling." The strategic implication is that the AI moves from advisor to actor. This is where the governance

questions get serious. An AI that can act in the world can also act wrongly in the world. The consequences of mis-action are different in kind from the consequences of mis-advice.

A company that has only deployed Layers 1 through 3 has AI that can speak. A company that has deployed Layer 5 has AI that can move. The two are different categories of organizational risk and different categories of organizational capability.

Layer 6. Agents and Orchestration

Multi-step workflows where AI plans, retrieves, acts, evaluates outcomes, and iterates without human intervention at every step. Where a chatbot answers, an agent takes a goal, decomposes it into steps, executes the steps with appropriate tools, evaluates whether the goal has been met, and either completes or escalates.

Orchestration is the meta-layer above individual agents. It is the system that coordinates multiple agents working on multiple parts of the same problem, manages handoffs between them, and resolves conflicts when they arise. This is where some of the most ambitious AI-Native operating models live, and it is also where the field remains immature. Technical capability is moving quickly; reliable organizational practice for deploying it safely is still being worked out.

Anyone who read the 2025 coverage of this layer should know that its central recommendation changed. More agents is not automatically better. When every step depends on the one before it, small failure rates can compound: under a simplified independence assumption, ten steps that each succeed ninety-five percent of the time yield an end-to-end success rate near sixty percent. Real workflows are not that tidy, but the warning holds. Complexity needs to earn its place. Start with one capable agent and good context. Add staged handoffs when the stages produce something worth auditing. Use parallel agents when they bring genuinely independent evidence or expertise.

The connective tissue also moved toward neutral governance. The Model Context Protocol and the Agent2Agent protocol are now hosted within Linux Foundation projects rather than governed only by the companies that originated them. That improves openness, interoperability, and the odds of durable multi-vendor support. It does not make an implementation secure by default. Both protocols and their security practices are still evolving fast enough that a vendor's claimed support, version, authentication model, and permission boundaries are worth reviewing regularly.

And agents need identities. Who an agent is, what it may do, and on whose behalf it acts moved from an open question to a product category and a standards effort inside about a year. Section 5.9 takes up what that means for the workforce.

Layer 7. The Collaboration Surface

The shared workspace where individual AI capability becomes team-level intelligence. Shared context loaded once and available to everyone's AI work. Shared prompt libraries with versioning. Shared knowledge bases that accumulate as the team works. Shared review surfaces where AI-generated output is visible, comment-able, and improvable by the team rather than buried in private chats. Shared monitoring of what the AI is doing across the team.

This is one of the most under-discussed layers in the field and one of the most consequential. A company that has deployed Layers 1 through 6 without Layer 7 risks producing the thirty islands problem from Section 2.2, with individual AI work contributing little to shared capability. A well-designed Layer 7 gives that individual capability a path to become organizational intelligence. This layer deserves its own treatment, in Section 4.3.

Layer 8. Evaluation, Monitoring, and Governance

The pipelines and disciplines that keep the AI behaving as expected over time. Continuous evaluation of output quality. Drift detection, which means noticing when the model's behavior has changed for reasons that may not be obvious. Hallucination metrics. Audit trails. Decision rights enforcement. Policy controls on what the AI can and cannot do, expressed as actual technical controls embedded in the workflow rather than as documents.

AI monitoring has to test the quality of answers and actions alongside availability, latency, and cost. Section 4.4 takes this layer up in depth, because it is foundational to everything else working over time.

There is a specific and measurable gap inside this layer that many companies can land in without noticing. Watching what your AI does has become more common. Knowing whether it is right has not. In LangChain's self-selected 2025 State of Agent Engineering survey of 1,340 practitioners, 89 percent reported agent observability, 52 percent reported offline evaluations against test sets, and 37 percent reported online evaluations. The sample should not be mistaken for the whole market, but the gap captures the operational problem. Observability tells you what happened. Evaluation tests whether it should have. A company with dashboards and no evaluations can watch a system degrade in high resolution.

Layer 9. Domain-Specific Intelligence

The specialized capabilities that some kinds of companies need and most published AI-Native frameworks ignore. For operations-heavy organizations: virtual sensors that estimate measurements from other signals without direct physical instrumentation. Predictive maintenance models that estimate failure risk before breakdowns happen. Graph neural networks that learn patterns across complex interconnected systems. Systems with governed learning loops that are evaluated and updated from operational evidence. Federated learning approaches that train models across distributed sensitive data without pooling the raw data in one place.

This layer can hold a moat for industrial, healthcare, and defense companies because it joins AI capability to domain-specific data, physics, workflows, validation, and expertise. Generic AI assistance is rarely enough. A company in one of these sectors that ignores domain-specific intelligence may leave its most defensible capability untouched. A company that invests well may build something competitors cannot easily replicate. This is one of the layers Nodalix is built to address.

Two of the nine are better drawn across the stack than stacked neatly above the others. Layer 4, structured knowledge, can act as a floor for systems that must reason over the company's entities, dependencies, policies, and history. Not every use case requires a knowledge graph, but higher-consequence systems need an explicit grounding architecture of some kind rather than only model memory and a prompt. Layer 8, evaluation and governance, is a wrapper rather than a rung. It does not sit on top of the agents; it reaches around the whole structure, because an audit trail that covers only the last step tells you little about how the answer or action was produced. The floor grounds the system in the company. The wrapper makes its behavior inspectable and accountable.

Design Principles Across the Stack

The nine layers describe what the stack contains. Two design principles describe how the layers behave together. The principles cut across every layer and shape every architectural decision. They are also independent of the archetype choice; every AI-Native company has to settle them.

The first principle is hybrid determinism. Generative models are probabilistic systems, even though a particular configuration can make outputs more stable. Similar prompts and data can still produce different inferences as models, context, retrieval, or sampling changes. That variability is useful for exploration and unacceptable for work products that must be

reproduced, audited, or reconciled exactly. The architecture therefore has to support both: probabilistic components for generative work and deterministic controls, calculations, rules, and records wherever repeatability is required. Which workloads sit where is an operational choice, and it shapes how the system can be governed.

The second principle is push versus pull. Pull is the chat-style interaction: the human asks, the AI answers. Push is the inverse: the system surfaces what the human did not know to ask. Insights arrive before they are requested. Risks are flagged before they become incidents. Recommendations appear before the decision moment. The push model requires more sophisticated infrastructure than the pull model because the system must decide what is worth surfacing, to whom, with what urgency, and with what evidence. A noisy push system is not intelligence. It is an alarm people learn to ignore.

The two principles combine into a four-quadrant matrix. Pull plus probabilistic produces exploration mode, the familiar chat use case. Push plus probabilistic produces insight surfacing. Pull plus deterministic produces trusted lookup and calculation. Push plus deterministic produces rule-bound automated execution. A mature AI-Native company may use all four, but not every company needs every quadrant and not every workload belongs in only one. The architectural question is which operating mode each decision or task requires.

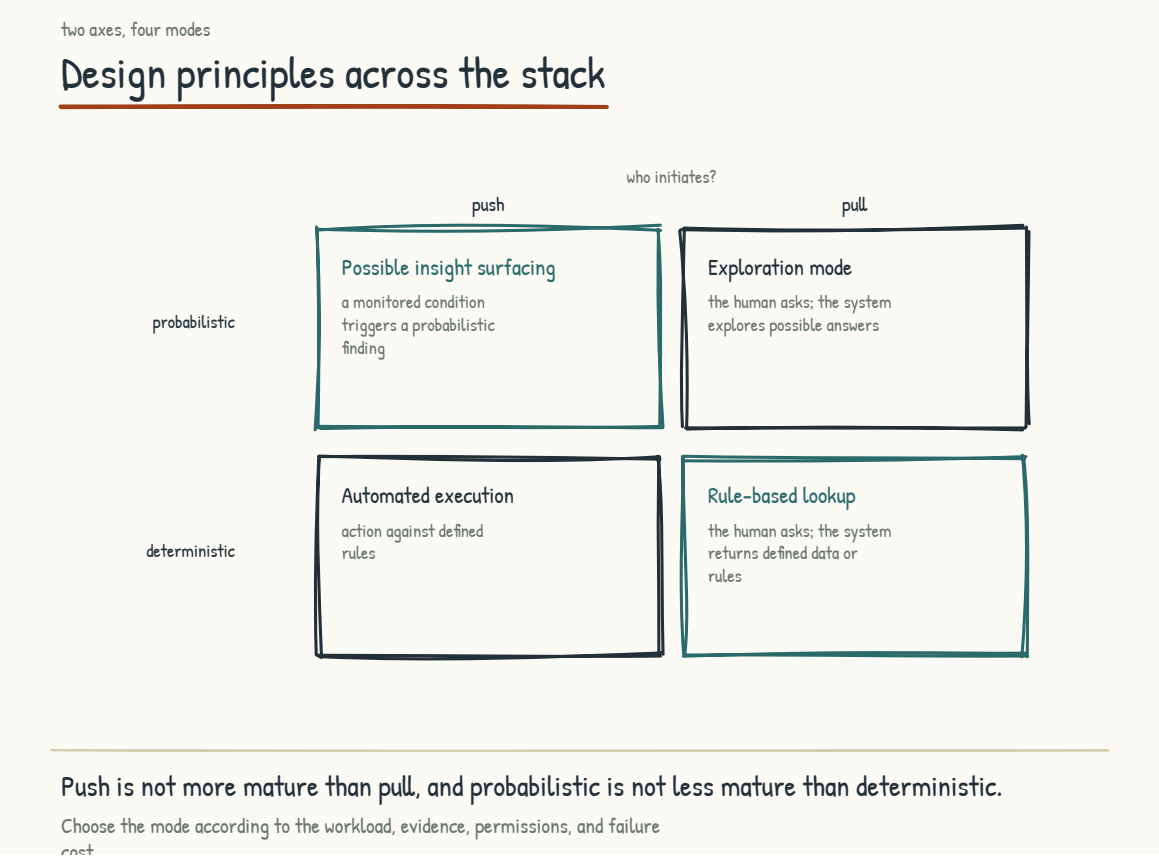

Figure 2.3b. Design principles across the stack. Push/Pull and Probabilistic/Deterministic produce four operating modes the leadership team has to design deliberately.

What the Stack Buys: The Autonomy Ladder

All of that describes what the stack is and how it behaves. What it buys is a separate question, and the one a leadership team actually asks. The answer is altitude. There are five rungs of autonomy. A company using AI is operating on one or more of them, whether it has named them or not; a company not yet using AI stands below the first.

The first rung is inform. The AI retrieves and synthesizes, and a human does everything else. This is where the chat box lives, and the value is time saved.

The second is advise. The AI reasons about this specific situation and recommends with evidence attached, and a human decides and acts. The value is better decisions.

The third is act. The AI takes bounded actions through tools while a human oversees. The value is work executed rather than recommended.

The fourth is operate. The AI runs a closed loop and a human supervises by exception. The value is sustained autonomy at a frequency, speed, or scale that may be impractical to provide through headcount alone.

The fifth is embody, where AI drives machines that act on the physical world.

Two things can climb together as you go up, and only one of them is obvious. Potential value rises, which is why leadership teams want the higher rungs. The required substrate and the consequences of failure rise with it, which is why many organizations stall. Even a recommendation needs reliable context and a way to judge the evidence behind it. An action taken on your behalf adds tool permissions, validation, an audit trail, bounded authority, and a way to stop it.

That asymmetry produces the mistake I see most often. A leadership team tries to buy altitude. They read about agents operating autonomously, they want that, and they reach for it on top of a substrate that is one chat license wide. Autonomy is vertical. How well the system knows your business is horizontal. You can only climb as high as the substrate carries you, and moving up without widening it does not add capability. It raises the consequences of a confident wrong answer. You earn altitude by widening the substrate, not by trusting the model further.

This also explains a plateau many companies hit and cannot diagnose. The chat pilot goes well, the second one goes well, and then little compounds. The pilots were all on rung one, where the substrate requirement is lower and the durable advantage is usually limited. Every rung above it demands more of the layers that many budgets never reach.

The Strategic Implication

The altitude argument has a stack-side reading, and it is the practical one. Section 2.5 reads the same architecture a third way, as a progression over time. Every rung of autonomy corresponds to a depth of investment, so the question of how high a company can climb resolves into the question of how far up the layers it has actually built. Every company makes those architectural decisions about which layers to invest in, how they connect, and what to build versus buy. The decisions are usually made tactically, by procurement, vendor by vendor, with the leadership team approving budgets without engaging the architecture. They should be made strategically, by the leadership team, with the operating model in mind and the archetype choice from Section 2.2 as the orienting frame.

A company at Layer 1 only has raw chat access to foundation models and nothing more. Every employee has their own chat session. The work does not compound. Central governance and visibility are difficult because the organization cannot readily see what the AI is producing. Shadow AI risk is high.

A company at Layers 1 and 2 has added embedded AI in existing tools. Adoption is real and productivity gains are real, but the organization has not yet become AI-Native in any meaningful sense.

A company with meaningful capability across Layers 1 through 4 has a foundation for deeper transformation: foundation models, embedded AI, retrieval, and knowledge architecture that grounds systems in the company's actual context. Not every use case needs every grounding mechanism, but this is where AI-First can seriously begin.

A company at Layers 1 through 6 has agents and tools that can act in the world. This is the technical foundation for Intelligence-at-the-Center or Operating Layer archetypes.

A company with the relevant capabilities through Layer 8 working coherently has added a collaboration surface that can convert individual capability into team intelligence, plus evaluation and governance that help keep behavior within defined bounds over time. This is a technical foundation for sustainable AI-Native operations, not a guarantee of them.

An operations-heavy company at Layers 1 through 9 has the same plus the specialized domain intelligence that creates real moats. This is the technical foundation for an AI-Native operating model in an industrial environment.

Not every company needs to invest at every layer. The point is that the choice of which layers to invest in, and in what sequence, is a strategic decision that flows from the archetype choice. Making this decision deliberately, with the leadership team engaged, is much better than letting it emerge by accident from procurement.

The next section gives a way to honestly assess where your company is in the AI-Native transition, regardless of which archetype or which layers you have chosen. That assessment is the maturity model.

WATCH FOR THIS

> **The most common investment trap in AI-Native programs: over-investing in Layers 1 and 2 (chat tools and copilots) and under-investing in the knowledge, action, orchestration, collaboration, and evaluation layers that make individual use compound. The visible spend is at the interface. The durable capability is usually underneath it. If most of the AI budget buys seats while little funds shared knowledge, integration, evaluation, security, and operating-model change, the portfolio deserves a redesign.**

Section 2.4

The Five-Stage Maturity Model

What actually produces a financial return on AI?

The field is full of AI maturity models, and many of them are not useful for honest self-location. Some are too vague to be actionable. Some are too rigid to fit real companies. Some are designed to make the company commissioning the assessment look like it is doing better than it is. I am going to propose a five-stage model that is simple enough for a leadership team to stand in front of, opinionated enough to distinguish between stages, and honest about what real companies will and will not achieve.

The stages are developmental rather than a rigid implementation sequence. Each builds on capabilities established earlier, and apparent shortcuts usually leave a dependency unpaid. A company can deploy a later-stage technology before it has the governance, knowledge, or cultural foundation beneath it; that is not the same as having reached the later stage. Companies can also be unevenly developed: strong on the technical capabilities of Stage 3 while still operating culturally at Stage 1, for example. The model is most useful when a leadership team uses it to find where its own capabilities are uneven rather than to score itself at one number.

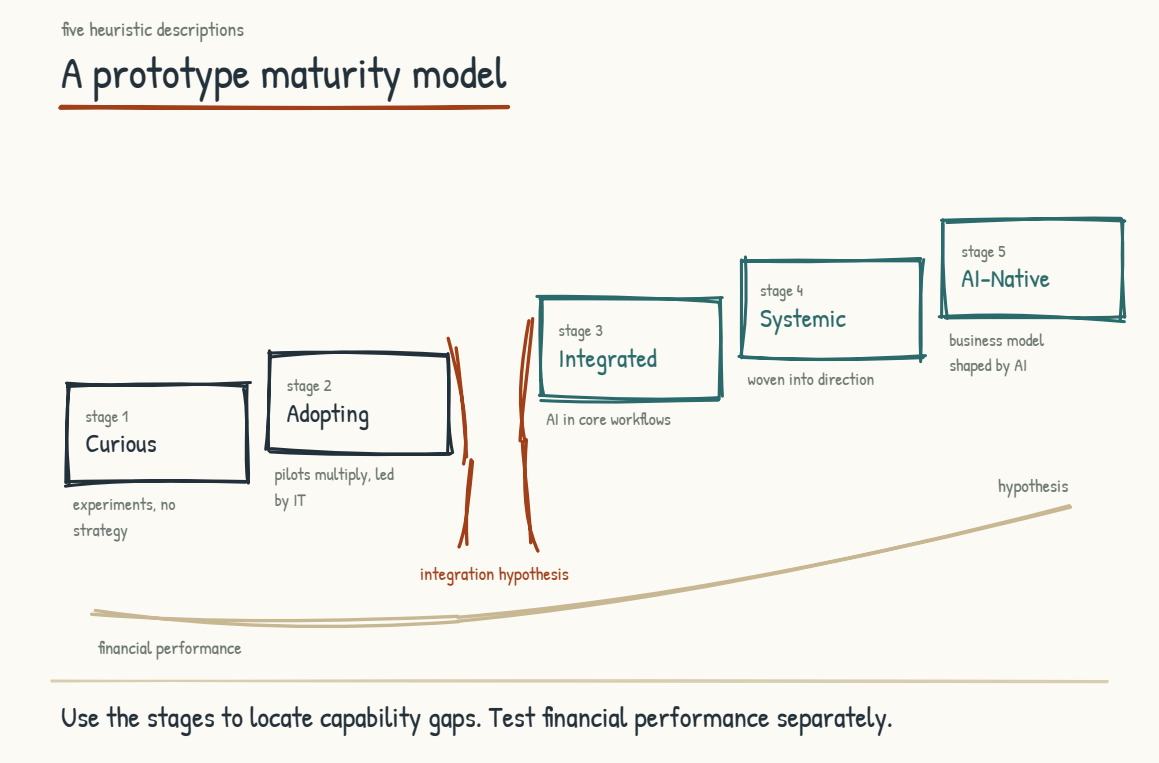

Figure 2.4a. The five-stage maturity model, with the financial performance curve from MIT CISR research below. The chasm between Stages 2 and 3 is where most companies stall.

Stage 1. Curious

AI is a topic of conversation. Individual employees are experimenting with consumer-grade tools, mostly on their own time, mostly without organizational knowledge. There is no AI strategy. There is no shared vocabulary across the leadership team. There is no governance framework. There is no taxonomy of work indicating which tasks are good candidates for AI and which are not.

The symptoms of Stage 1 are familiar. Slack channels with prompt-sharing. "Has anyone tried this?" emails. Inconsistent answers when different executives are asked what the company's AI position is. Different teams using different tools, sometimes paid for through personal credit cards. Cost mostly invisible because the spend is decentralized. The talent that cares most about AI is quietly looking for jobs at companies that take it more seriously. The talent that fears AI most is anxious in ways the company has not yet acknowledged.

Most companies are at Stage 1 longer than they realize, because the experimentation activity creates a feeling of progress that masks the absence of strategy. Companies at Stage 1 should be honest about it, because moving forward requires acknowledging where you are.

Stage 2. Adopting

Pilots are underway. Tools are being deployed in pockets of the organization. Investment is real and growing. There is at least one named AI initiative with a budget and an owner.

This is also where most organizations get stuck. Stage 2 shows itself in disconnected pilots multiplying across business units, power struggles over who owns AI strategy, shadow AI projects competing with sanctioned ones, executive burnout from constant AI updates that do not aggregate into a coherent picture, rising costs with unclear returns, and competing definitions of "AI strategy" depending on which executive you ask.

The pilot-graveyard company from Section 1.2, whose numbers Section 8.2 fills in, is the textbook Stage 2 organization. The energy is real. The investment is real. The output is small because the alignment was never produced.

The MIT CISR research from Section 2.1 (Weill, Woerner, and Sebastian, 2024) is most clarifying at this stage. In its sample, companies at the second maturity stage had not yet converted their investment into the stronger financial performance associated with the next stages. Many organizations can stay stuck here for years: spending, piloting, and learning without producing a return commensurate with the effort. This is the chasm.

The researchers place the greatest financial impact at the move from their second maturity stage to their third, where performance crosses from below to above the industry average. The 2025 study, by Woerner, Sebastian, Weill, and Káganer, drew on a survey of 152 executives and examined that transition more closely. The result is an association, not a promise that crossing a line on a maturity model causes a return. Even so, the pattern matters. The earlier moves are setup. The later moves refine. The middle move is where the curve turns. This is the threshold the architecture has to be designed to cross.

> **SOURCE**
>
> *MIT CISR's four-stage Enterprise AI Maturity model, based on a 2022 survey of 721 companies and 2024 interviews with executives at nine enterprises. The five-stage model in this book builds on the MIT CISR four-stage framework. The financial-performance curve is theirs; the additional stage and the framing of the chasm are mine.*
>
> Weill, P., Woerner, S. L., and Sebastian, I. M. (2024). "Building Enterprise AI Maturity." MIT CISR Research Briefing, December 19, 2024. / Woerner, S. L., Sebastian, I. M., Weill, P., and Káganer, E.

(2025). "Grow Enterprise AI Maturity for Bottom-Line Impact." MIT CISR Research Briefing, August 21, 2025.

Stage 3. Integrated

AI is operational in core workflows rather than only in pilots. A shared vocabulary across the leadership team is emerging, though it may not yet be fully consistent. Cross-functional ownership of AI is structured rather than improvised. Governance exists as actual practice embedded in execution rather than as a document. Decision rights are explicit for routine cases. Who decides what, when AI is involved, with what oversight. The workforce has begun real upskilling rather than communication about the importance of upskilling. Pilot-to-production conversion is rising.

This is the stage where the financial curve in the MIT CISR sample bends upward sharply. Integration does not guarantee a return, and maturity labels do not create one. But this is where scattered experiments can begin producing real, measurable, sustained financial results because the operating system around the technology finally exists. Companies that remain in perpetual pilot mode rarely capture the same compounding effect.

Moving from Stage 2 to Stage 3 is the hardest single step in the model, and it is where most leadership-team alignment work has to happen. This is an operating-model transition. The technology of Stage 2 and Stage 3 may look similar from the outside. The difference is that Stage 3 has shared decisions, shared accountability, shared metrics, and shared discipline behind the technology.

Stage 4. Systemic

AI is the substrate of work rather than a feature added to it. Knowledge architecture is intentional, with deliberate decisions about which systems are authoritative, who owns definitions, and how institutional knowledge is captured and curated for AI reasoning. Roles, decision rights, and metrics are designed around human-AI partnership rather than around pre-AI assumptions. New roles emerge: AI Agent Trainers, AI Agent Managers, Knowledge Stewards, whatever the organization calls them. These are real positions with real authority. Continuous learning loops are part of the design; the company learns from each AI deployment and the learning improves the next. Governance is embedded from the start.

A Stage 4 company can build sustained competitive advantage from its AI investment. Learning moats are forming. The data, the institutional knowledge, the prompt libraries, and the validated workflows improve with use. A competitor can buy the same vendor but not the accumulated context, judgment, and operating history. That does not make the lead permanent. It makes imitation slower and forces the competitor to do the institutional work rather than merely procure the technology.

For an ambitious incumbent, Stage 4 is a plausible destination, but the calendar depends on its starting point, regulatory environment, technical estate, and willingness to redesign work. Five to ten years is a planning horizon, not a benchmark or a promise. This is the realistic destination for many leadership teams that are serious about the work. It is also where durable advantage can live, because the capability is difficult enough to build that procurement alone cannot reproduce it.

Stage 5. AI-Native

The business model itself depends on AI capabilities that competitors cannot easily replicate. The economics of the company (how it creates, prices, and captures value) are shaped by AI rather than merely enhanced by it. In the right work, cycle times can fall dramatically and capacity can expand without moving in lockstep with headcount. New revenue pathways can open that did not exist in the pre-AI version of the business. These are design possibilities, not automatic consequences of adopting AI.

The workforce is designed around an explicit division of labor between people and AI. The org chart protects human authority where judgment, relational presence, ethical accountability, lived consequence, and original synthesis matter, while AI expands what those people can see and do. The boundary is designed and revisited rather than inferred from whatever the technology happens to automate first.

Stage 5 is rare in incumbents. The companies most credibly approaching it are either AI-first startups built within the last few years or established companies that have been investing in AI for a decade or more with sustained leadership commitment. For most incumbents, the realistic destination is deep Stage 4 with selected Stage 5 capabilities in the parts of the business where AI most enables a new business model.

The Diagnostic Move

A leadership team using this maturity model to do honest self-assessment should not score themselves at one number. The honest answer is almost always scattered.

A typical leadership team that takes this assessment seriously will discover something like the following pattern: Stage 3 on technology (the tools are deployed and integrated). Stage 2 on people and culture (the workforce is anxious and the upskilling is rhetorical). Stage 1 on knowledge architecture (no one has done the work of structuring the company's institutional knowledge into a form AI can reason over). Stage 3 on customer-facing AI (the product is doing well). Stage 1 on operational AI (the back office is unchanged). Stage 2 on governance (policies exist but enforcement is uneven).

The scatter is the diagnosis. A company that is at Stage 3 on average but Stage 1 on knowledge architecture has a foundation problem likely to undermine gains in other dimensions. A company that is at Stage 2 on average but Stage 4 on people and culture has a workforce that is more ready than the systems it is being asked to use. The leadership work is to look at the scatter and decide where to invest next based on which gap is most consequential rather than which score is most embarrassing.

Figure 2.4b. The scatter is for planning; a single number is for the board deck.

The maturity model becomes most useful when applied to each dimension of your company separately, in whatever cut fits your operation: technology, people, knowledge architecture,

customer-facing AI, operational AI, governance. The single number is for press releases. The scatter is for actual planning.

A Word on Maturity Models Generally

Maturity models have a recurring failure mode. They become competitive scorecards rather than diagnostic instruments. A leadership team starts using the model to assess where they "should be" relative to peers, instead of where they are and where they need to invest next. The model stops being useful for self-improvement and becomes a status game.

This model is meant for honest self-location rather than benchmarking. There is no prize for being at Stage 4 when your business does not require it. There is no shame in being at Stage 2 if you are genuinely working to get to Stage 3. The point of the model is to give a leadership team a shared language for talking about where they are and what they need to do next. The point is not to produce a score for the board deck.

I should say plainly that this model is newer than the rest of my practice, and AI is new enough that I have not yet run it with a room. I built it as a self-assessment. Used well, a leadership team should come out of it with three things: an honest read of where they sit on each dimension, a decision about which dimension is the highest-return place to invest next, and a clear view of what would have to be true to advance a level there. Then the real conversation, which is whether the level you are aiming at is the one you actually want. Everything else is decoration.

The dimensions themselves are what Stage 3 takes up. Before that, one more piece of vocabulary, because the maturity model and the stack both assume a reader who can tell an agent from a harness from an operating system, and those words are being used interchangeably in the market right now.

CONSIDER BEFORE YOU ACT

Locate your operations division on the five-stage curve. Now locate your corporate functions. Now locate your product group. Most companies are not at one stage everywhere. They have one part at Stage 3 and another at Stage 1. The honest map is a multi-stage portrait rather than a single point. The leadership question is whether the stages across the company are coherent with each other, or whether they are creating internal friction.

Section 2.5

From Chat to Operating System

What are you actually buying when someone sells you AI?

Every few months the thing being sold as AI changes shape. A leadership team that bought a chat license in one budget cycle is being pitched an agent platform in the next and an operating system in the one after, and the words arrive faster than the definitions. This section is the vocabulary. Six architectural forms, arranged in the order their capabilities generally accumulate, each one containing elements of the last.

The progression matters more than the labels. Each form responds to a ceiling in the one below it, and the ceiling is often the same shape: the system produced something plausible that nobody could adequately check. What follows is not a clean chronology (the ideas overlap, and products combine them), but a history of trying to close that gap.

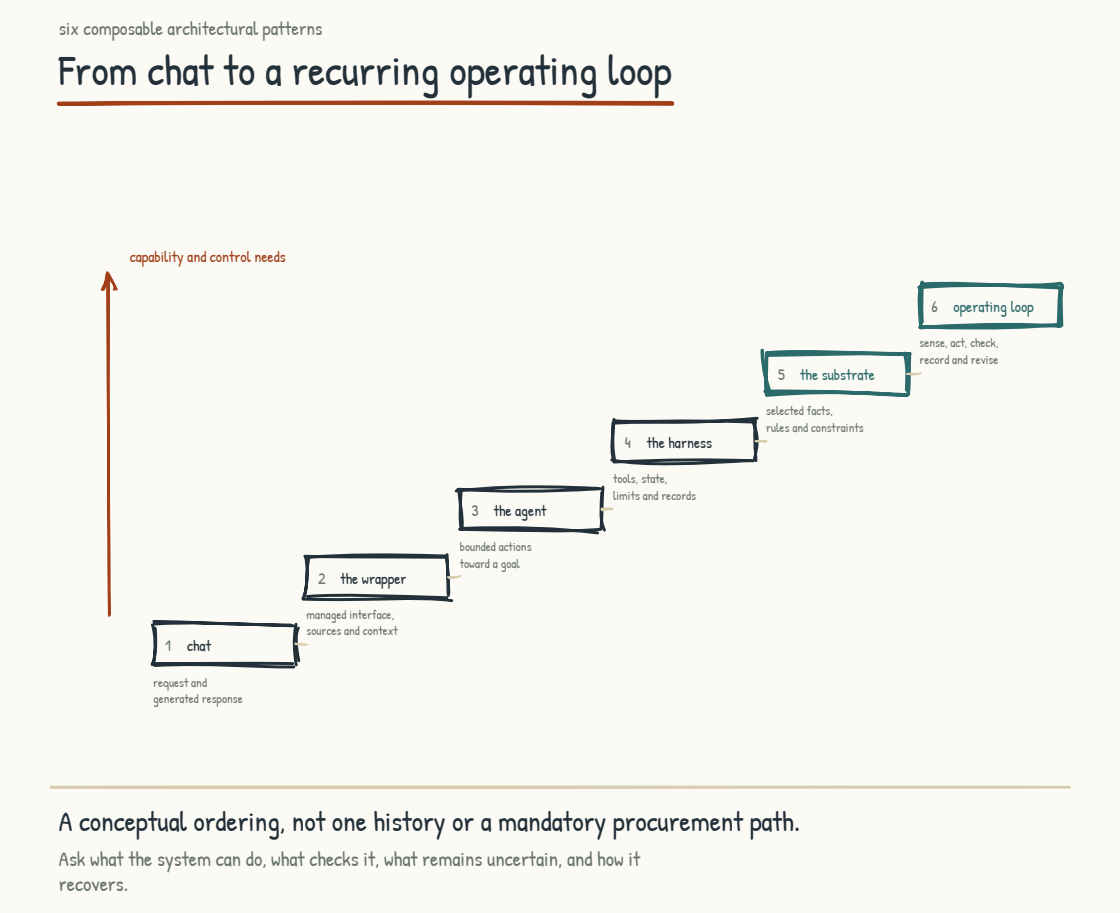

Figure 2.5a. Six stages, each one containing the last. What changes as you climb is how much the system can be trusted to do.

One. Chat

A person types a question and a model answers it. In its bare form, the model brings broad training and little or no knowledge of you. It can sound as though it knows everything in general while knowing nothing authoritative about your particular situation, which is why the answers can feel impressive and still land shallow. A chat product may retain conversation history or user memory, but that is not the same as building shared institutional knowledge. Every colleague can still end up with a different private context, and little accumulates anywhere the company can govern or reuse.

The ceiling: without retrieval or tools, the model cannot look up the current facts that matter, so it predicts what a good answer would sound like from the context and patterns it has. That is a useful way to draft an email and a poor basis for a consequential decision. Value here is time saved by individuals, which is real but compounds only if the organization captures and improves what people learn.

Two. The Wrapper

Put an interface around the model, point it at some company documents, and give it a name. Much of what has been sold as enterprise AI begins here. The wrapper can add retrieval, citations, access controls, workflow context, and a place for the conversation to live. The model remains the generative core, but the surrounding software can carry meaningful product value. Take the model out and the intelligence disappears; take the wrapper out and the model loses the organizational context and controls that make it usable.

The ceiling: retrieval grounds the answer in a document, and a document is a description of the work written by someone, for someone, at a moment. It is not the work. Ask a wrapper what happens if a supplier misses a delivery and it will find you the paragraph in the contingency policy. It cannot trace what actually depends on that delivery, because nothing in the architecture knows.

Three. The Agent

Give the model tools, state, and a control loop and it can move from answering toward doing. Depending on the design, it may take a goal, choose or follow steps, call tools, inspect results, and either finish, retry, or ask for help. This is a real change in kind. The system moved from

advising to acting, and the consequences of a wrong action are different from the consequences of wrong advice.

The ceiling: an agent places a probabilistic component inside a system that can act. It will sometimes choose the wrong action, misread a result, or pursue a flawed plan, and longer chains create more opportunities for error even when checks catch some of them. Confidence is not a reliable measure of correctness. Agents alone do not solve the trust problem; they raise the stakes of not having solved it.

Four. The Harness

The harness is the software wrapped around the model that turns capability into work. It hands the model its tools, decides what context the model sees, runs the loop, and enforces limits on every step. Most of what people credit to the model is actually the harness. The same model inside a good harness and a poor one produces work of visibly different quality, which is why teams comparing models often discover they were comparing scaffolding.

This is where the field spent most of its energy, and the payoff was real. A harness gives you retries, memory, tool permissions, a place to log what happened, and a way to stop a runaway loop. It is also the point at which the strategic picture inverts. The model is rented, swappable, and improving on someone else's schedule. The harness is yours.

The ceiling, and this is the one worth understanding: a generic harness governs how the system works without necessarily holding an authoritative model of what it is working on. It can enforce permissions, schemas, budgets, and tool-call rules. It can run tests or evaluators when the domain supplies them. What it cannot infer from orchestration alone is whether a substantively plausible answer is true. Process checks matter, but the failures that matter most in an operating business are often failures of substance.

Five. The Deterministic Substrate

So the next move is to give the system a source of truth for checking its answers. A substrate is an explicit model of the thing being operated on: what exists, how it connects, what depends on what, and what is not possible. In an industrial setting that is a graph of the plant plus the physics that constrains it. In a services firm it is the structure of clients, engagements, obligations, and dependencies. In either case it is a live representation of the actual operation rather than a pile of documents about it.

What the substrate makes possible is substantive checking. Not only checking that the model followed a procedure, which a harness can enforce, but checking a claim or proposed action against an authoritative constraint. A proposal that violates a dependency, a conservation law, an approval rule, or a contractual obligation can be rejected before it reaches a person, and rejected for a reason the person can read. The check is only as trustworthy as the rule and underlying data, but it is qualitatively different from asking another model whether the first model sounds right.

This is what deterministic assurance means in practice, and it is worth being precise, because the word gets used loosely. Nobody has made the model deterministic. The model still generates probabilistically, and not every business judgment can be reduced to a rule. The architecture instead gives deterministic components authority at boundaries where explicit constraints exist: the model proposes; validated rules decide what is permitted. Probabilistic proposal, deterministic constraint. That arrangement does not guarantee a correct answer, but inside a tested control boundary it can enforce specified rules consistently. That is a material difference between an interesting demonstration and something you would let near an operation that can hurt someone.

It also attacks three related problems from a common direction. Variance, hallucination, and drift have different causes, so no substrate eliminates all three. But each becomes easier to detect and contain when outputs are checked against authoritative state and explicit constraints. More guardrails are not automatically better; the useful guardrail is one tied to a real property of the work.

The ceiling: a substrate can make individual claims and actions more checkable. It does not make every answer trustworthy, and it does not make an organization run on them. A company can have an excellent graph and still have thirty people asking it thirty private questions whose answers never meet.

Six. The AI Operating System

The last stage is the one with the least agreement about what it means, so here is what I mean. An operating system is what you have when sensing, structure, reasoning, action, and assurance are wired into a loop that closes and keeps closing. Work comes in, gets placed against the model of the business, gets reasoned over, produces an action or a recommendation, gets validated, gets recorded, and the record improves the next pass. Any one of those parts on its own is a feature. The loop is the operating system.

Three things distinguish it from everything below. It is proactive rather than waiting to be asked, because a system that holds a live model of the operation can notice things nobody queried. It is shared rather than private, because the loop runs where the team can see it. And it can accumulate: when outcomes are reviewed, records are curated, and corrections are carried forward, each closed cycle can improve the next pass.

This is also the point where the thing stops being a tool the company uses and becomes part of how the company runs, which is why the rest of this book is about the operating model rather than the technology. Stage 3 takes up what that means for the work itself.

Reading the Progression

Two warnings about using this list. The first is that capabilities tend to accumulate, but the forms are not a mandatory product ladder. An operating system still calls models and needs orchestration and assurance, but it may not expose every earlier form as a separate component. A product can combine several levels at once. A company that tries to buy form six without the necessary grounding, assurance, and orchestration capabilities beneath it has bought a diagram.

The second is that this progression gives you a useful vendor question: Where does your product actually sit, and what does it use to check an answer? A wrapper may answer with its model and retrieval sources. A harness may answer with orchestration features, evaluators, and tool controls. A serious answer to the second question names the authoritative data, rules, tests, or domain constraints used to check substance, independent of any model asked to grade the first.

Many companies are somewhere between two and four and describing themselves as six. The exact distribution is not measured here; the category inflation is visible in the market. The gap between those two positions is not merely a marketing problem. It is the work.

CONSIDER BEFORE YOU ACT

If you keep three things from Stage 2: 1) There are AI-Native companies, plural, and choosing your archetype deliberately is among the most consequential decisions you will make. 2) The chat box is one of nine layers, and the durable capability lives in the middle layers many budgets never reach. 3) The strongest financial-performance shift in the MIT CISR research appears in the move from Stage 2 to Stage 3; honest self-location, dimension by dimension, is what helps an organization cross the chasm.

STAGE 3

The Operating System

How the work of an AI-Native company actually runs.

Section 3.1

People, Process, Technology, and Now AI

What changes when AI becomes a dimension of the operating model?

For decades, a dominant frame for analyzing how a company actually works has been People, Process, Technology. The trio shows up throughout consulting and operating-model redesign. It endures because it captures something true. How a company functions is the joint product of the people in it, the processes they follow, and the technology they use. Move one without accounting for the other two, and the change is unlikely to stick.

That trio is now insufficient for the design problem this book addresses. AI is still technology, but treating it only as another tool hides its distinct role in the work. People work with it, on it, around it, and against it. Processes are rewritten by it, rewritten around it, sometimes partly replaced by it. Technology itself becomes harder to define because the line between tool and participant is no longer clean. AI cuts across the existing trio in ways that make it useful to analyze as its own dimension.

So the working frame is People, Process, Technology, and AI. Four dimensions. Each one has its own logic. Each one interacts with the others. The intersections are where the actual work of an AI-Native operating system gets designed.

INTELLECTUAL LINEAGE

The four-dimensional structure has a direct ancestor. Harold Leavitt, at Carnegie Institute of Technology and later Stanford, described industrial organizations in 1965 through four interdependent sets of variables: Task, People, Structure, and Technology. Change one, he argued, and compensatory or retaliatory changes usually follow in the others. The later People, Process, Technology trio is often traced to this lineage, with Process absorbing much of Task and Structure, although its exact path into management practice is less tidy than the familiar origin story suggests. My separation of AI as a fourth analytical dimension is an extension of that systems logic. Intelligence has become an architectural component of the work, not merely a feature in a tool, and changing it changes the rest.

Leavitt, H. J. (1965). "Applied Organizational Change in Industry: Structural, Technological and Humanistic Approaches." In J. G. March (Ed.), Handbook of Organizations (pp. 1144-1170). Rand McNally.

The standard term in this field is "operating model." We use "operating system" instead, for the following reasons.

The operating model is the strategic frame. It describes how a company creates value, who its customers are, how it organizes to deliver, what it measures. The operating model is necessary and the leadership team needs to be aligned on it. It is also too high-altitude to do the work the chapters ahead require.

The operating system is the practice-level frame. It describes how a specific process actually works, who does what, what tools are used, how decisions get made, what good output looks like, where the failure modes are. The operating system is where the texture of work lives. It is where AI either gets integrated into the work or fails to integrate. It is where the question "what does our company actually look like on a Tuesday afternoon" gets answered.

What I mean by altitude is simple enough. An operating model is the framework for understanding the parts of the business and how they fit together, and it sits high on purpose. To transform how a company runs, you have to define the operating model at a far more granular level: how you actually operate today, and how you intend to operate once AI is part of it. That granular level is what I am calling the operating system.

And the proposal in the title is deliberate. An operating model normally encompasses strategy, structure, people, process, and technology. I am separating AI analytically from the technology

box because systems that generate, recommend, and act create design questions that ordinary tools do not. This is a design choice, not a claim that AI has ceased to be technology. That is the argument this stage has to earn, and the rest of it is about how you would actually design AI into the operating model rather than bolt it alongside.

Both frames matter. The operating model tells you why your company exists and how it intends to compete. The operating system tells you how it actually runs. AI-Native transformations that work at only the operating model level produce strategies that do not translate to execution. AI-Native transformations that work at only the operating system level produce tactical improvements that do not aggregate into strategic advantage. The leadership work is both.

Everything that follows in Stage 3 works at the operating system level. It inventories the processes that make up that system and treats several of them in depth. The destination is concrete enough that a leadership team can take a process they recognize, hold it up against the AI-Native form, and decide what to do about it.

The figure below draws one distinction the four-part frame compresses. At the strategic level, AI is one dimension. At the practice level, the diagram splits it into assistive AI and agentic AI: systems that advise and systems authorized to act. They share some design questions, but they are staffed differently, governed differently, and fail differently. This is not a fifth independent category layered onto the argument. It is a closer view of the fourth.

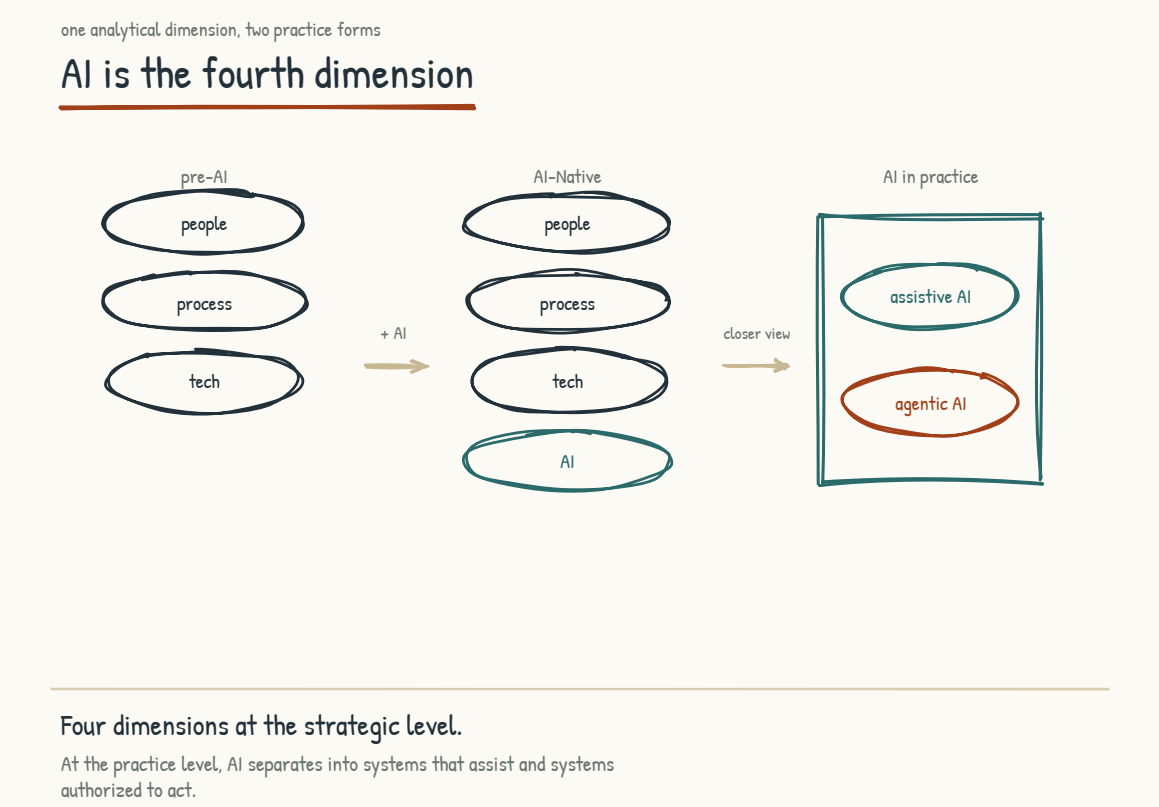

Figure 3.1a. Four strategic dimensions, with AI separated at the practice level into systems that assist and systems authorized to act.

Section 3.2

The Operating System Inventory

> *What does the work of a company look like when AI participates in it?*

The operating system of a company is the catalog of processes that actually produce the work. The catalog is usually longer than the leadership team realizes. The catalog is not written down anywhere central. People in the company know how their own piece works and have a vague sense of how the adjacent pieces work, but no one holds the whole picture.

This is one of the quiet costs of the pre-AI era. Companies could function without an explicit operating system inventory because the institutional knowledge to operate each process lived in the heads of experienced people, and the cross-process coordination happened through

meetings, relationships, and informal escalation. The system worked despite never being mapped.

AI changes the economics of this. An AI-Native company cannot be built reliably on top of an operating system that remains largely unmapped. AI can infer from examples and work with context supplied in the moment, but it cannot consistently apply institutional knowledge it cannot access. Processes that depend on unwritten judgment are therefore difficult to automate safely and difficult to improve systematically. So an early piece of the AI-Native transition, and one of the most undervalued, is making the operating system inventory explicit enough for people and systems to use.

Four processes are treated at depth here. Each uses the same six-field template.

First, the traditional form. How the process worked in the pre-AI operating system. This is an honest description of the process as competent companies have run it for decades.

Second, the AI-Native form. How the process works when AI is a designed participant in it rather than a feature attached to its surface. This is the destination, and the current state does not always match it.

Third, the shift. What specifically changes between the two forms: less the technology delta than the behavior delta, the decision delta, the role delta.

Fourth, the stack layers engaged. Which layers from the AI Stack are involved in the AI-Native form. This makes the connection between architecture and practice explicit.

Fifth, the people implications. What changes for the humans doing this work. Who does it now, who used to do it, what gets harder, what gets easier, what new skills are needed.

Sixth, the risks. What can go wrong in the AI-Native form: the risks of adopting AI and getting it wrong, rather than the risks of failing to adopt it, which is a different conversation.

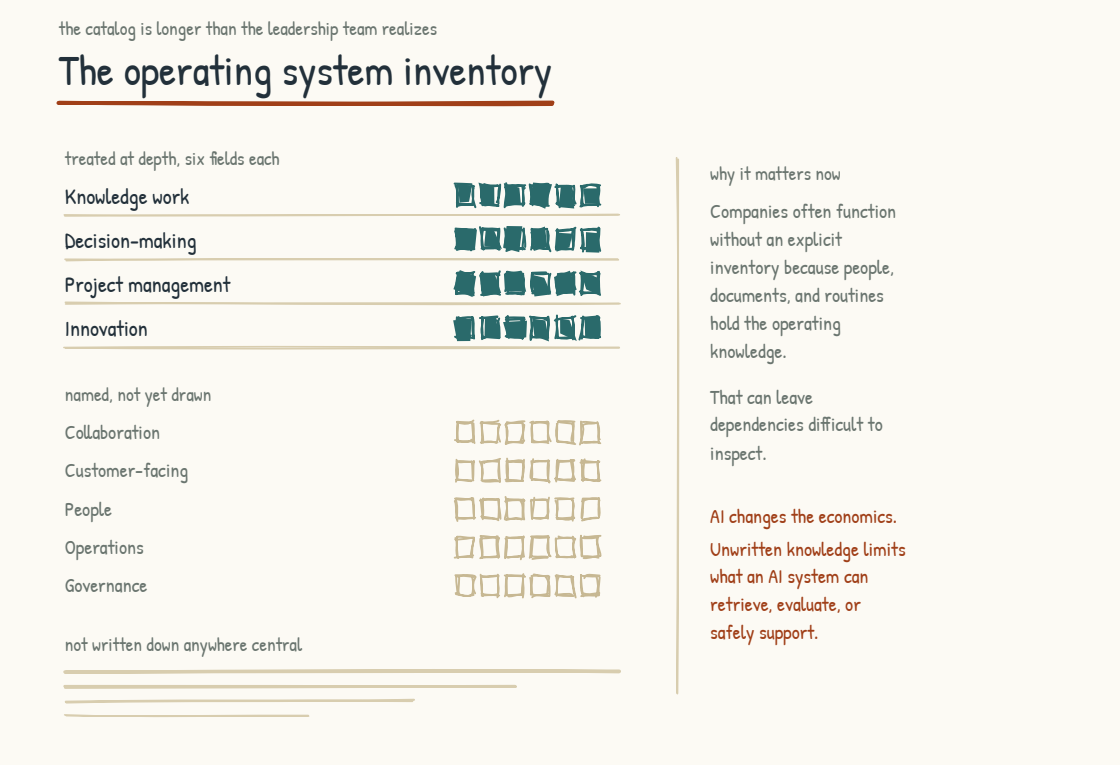

Figure 3.2a. The operating system inventory. Four processes treated at depth, five more named, and a long tail that was never written down anywhere central.

Knowledge Work

Traditional form. An analyst is asked on a Tuesday morning for a competitive assessment of a market the company is thinking about entering. She blocks the afternoon, opens a blank document, and starts pulling: old board decks, industry reports, a call with a colleague who worked that market two jobs ago. By Thursday she has a draft. It goes to her manager, comes back with comments, and ships the following week. Blank page to delivered output takes eight working days, and how good it is rests on what she knows and how hard she pushes.

AI-Native form. The same assessment now starts with an hour of writing down what the AI cannot know: why the company wants this market, what the board balked at last time, which competitor actually worries the CEO. A structured draft comes back before lunch, and she spends the afternoon checking its claims, cutting what is generic, and adding what only she knows. What goes out Thursday morning is a synthesis of what the AI produced and what she brought. Section 3.3 draws the full cycle: prompt, level-set, validate, build, integrate. The unit of work has changed.

The shift. For suitable work, the cycle can be faster, sometimes dramatically so. The work is also different in kind. The mix of skills that makes a knowledge worker valuable is changing. The ability to produce a competent first draft is becoming less scarce. The ability to frame the right question, evaluate output critically, and synthesize across sources becomes more important. The shift runs from one kind of human work toward another, with AI doing a substantial share of tasks that used to be performed only by people.

Stack layers engaged. Layer 1, foundation models. Layer 2, embedded intelligence in existing tools. Layer 3, RAG against the organization's knowledge base. Layer 7, the collaboration surface where AI work is shared and reviewed by the team. Layer 8, governance, especially around sourcing and attribution.

People implications. The knowledge worker's effective scope can expand. A senior analyst may be able to do work that previously required a team, depending on the work, the evidence required, and the surrounding system. The implication is that the team shrinks, the work expands, the standard rises, or some combination of the three. Cost pressure makes the first option tempting and visible. The strategic argument for the second and third is that capacity expansion or quality improvement can create differentiation. Team reduction alone rarely does, because the same cost lever is available to competitors.

There is a second people implication, and it lands on juniors. Work that taught them the trade, the first drafts, the basic research, the routine syntheses, is exactly what AI now does. How the next generation of senior knowledge workers gets trained is a real and unresolved question. Companies that figure this out will have an advantage. Companies that do not will discover, in five to ten years, that they have senior people retiring and no one to replace them.

Risks. The first risk is cognitive offloading. Knowledge workers who use AI as a thinking partner can preserve and extend their judgment. Those who routinely substitute its output for their own effort risk weakening skills they no longer practice. The effect depends on the task and the way the tool is used; the slow feedback is what makes the risk easy to miss. Section 5.4 treats this fully.

The second risk is homogeneity. When people use similar models with similar instructions and accept their defaults, output can converge in structure, language, and assumptions. Companies seeking advantage will need their people to use AI in ways that produce more distinctive work. This requires deliberate context, judgment, and effort.

The third risk is governance erosion. Without shared workspaces and shared review, the AI-mediated work of the company becomes invisible to the company. No one knows what prompts are being used, what sources are being consulted, what the AI is being asked to produce on behalf of the company. This is the shadow AI problem and it is the natural outcome of Citizen-AI without the collaboration surface.

Decision-Making

Traditional form. A decision is needed. Relevant data is gathered. Options are formulated. Stakeholders weigh in. A decision is made, often by a human in authority. Someone communicates it and the organization executes. How good that decision is depends on the information available, the rigor of the analysis, and the judgment of whoever made it.

AI-Native form. The decision-making process incorporates AI participation at multiple stages and with varying degrees of authority. AI may surface that a decision is needed by noticing a pattern a human had not yet seen. It may gather and structure relevant data faster than a human team could, formulate additional options, or model possible consequences using authorized company data and explicit relationships in a knowledge graph or other decision system. The human still makes the decision in most consequential cases. That person may make a different decision than they would have made without AI because the inputs are broader, faster, or more structured, but those inputs are better only if their data and reasoning survive validation.

For some classes of decisions, an AI-enabled system can execute autonomously within bounds people have set. The bounds should be explicit, the decisions reviewable, and escalation paths real. Human and organizational accountability does not disappear merely because software performed the action; its exact allocation depends on the decision, the law, and the governance design.

The shift. The most important shift is the explicit recognition that decision rights are now a designed thing rather than an inherited thing. In the pre-AI operating system, decision rights were often implicit, embedded in role descriptions, organizational hierarchy, and culture. In the AI-Native operating system, decision rights have to be explicit because AI is a participant in the process. The question of what AI decides versus what humans decide has to be answered by design.

The second shift is that some decisions can be made faster without sacrificing quality, and some can improve on both dimensions. That possibility changes the frontier; it does not abolish the trade-off. Weak data, poor framing, or inadequate review can still make a fast decision worse.

Stack layers engaged. Most layers are involved here. Layer 4, knowledge graphs, for the institutional knowledge that informs decisions. Layer 5, tools and actions, for AI to gather and structure data. Layer 6, agents, for the structured workflows that prepare decisions. Layer 8, governance, for the audit trails and decision-rights enforcement that make the system trustworthy. Layer 9 in operations-heavy environments, for the domain-specific intelligence that makes operational decisions trustworthy.

People implications. Decision-making roles change. Middle managers whose work is dominated by coordinating information flow and making repeatable, moderately complex decisions face particular exposure because AI can absorb a meaningful share of those tasks. Senior roles are not immune, but strategic, ambiguous, and relationally complex decisions remain less amenable to delegation. Leaders may be better informed than before when the system is well grounded and critically used. The middle is one place where task displacement is likely to be felt first; whether that removes roles or redesigns them is a management choice as well as a technical outcome.

The skill that becomes more valuable is the ability to frame a decision well, to ask the right question, to know when to override the AI's recommendation and when to follow it. These are higher-altitude skills than the routine analytic skills that used to fill the middle of the org chart.

Risks. The first risk is automation bias: people may over-rely on an automated recommendation or fail to contradict it, especially when the output is polished and the underlying basis is opaque. A model's verbal confidence is not evidence of correctness. The discipline of questioning the recommendation, holding alternative hypotheses, and exercising independent judgment is harder than it sounds and can erode without deliberate practice.

The second risk is decision opacity. If the AI's recommendation is followed, but no one understands why the AI recommended what it did, the company has made a decision it cannot defend, learn from, or revise. The Layer 8 governance disciplines are what make AI-informed decisions auditable. Without them, the company is making decisions in a way that cannot be sustained.

The third risk is the slow erosion of decision-making capability in the workforce. If AI handles all the routine decisions and humans only see the hard ones, the humans never build the judgment that comes from making many decisions. This is the same problem as the junior knowledge worker problem in a different context.

Project Management

Traditional form. A project is defined. A project manager is assigned. The project manager plans the work, coordinates the people doing it, tracks progress, surfaces risks, manages stakeholder communication, resolves blockers, and reports up. Project management tools support the work. The project manager is a human in the loop on every dimension of how the project actually runs.

AI-Native form. AI is a participant in the project from the planning stage forward. It can help decompose work into tasks, maintain the project plan as a living artifact when reliable signals are available, track explicit dependencies, surface possible blockers, draft stakeholder communications, prepare meeting summaries, and identify risks the humans had not named. The project manager is still present but is doing different work. The administrative load can drop. The judgment load becomes more visible.

The shift. Project management can become less about administration and more about leadership. Plan maintenance, status updates, and routine communication are increasingly automatable, although each still requires ownership and quality control. Deciding what to do when the plan breaks, handling stakeholder politics, and making human calls about pace and priority remain judgment work and can occupy more of the role as the administrative burden falls.

Stack layers engaged. Layer 2, embedded intelligence in project tools. Layer 3, RAG against project artifacts and historical project data. Layer 5, tools and actions, for AI to update plans and prepare communications. Layer 6, agents, for the structured workflows that maintain the project. Layer 7, the collaboration surface where the project lives.

People implications. Junior project managers are exposed because administrative coordination, status reporting, and basic plan maintenance are among the tasks AI-enabled tools can increasingly support. Senior project managers are exposed differently, but much of their value lies in judgment and leadership, where current systems remain less reliable. If automation actually removes administrative burden rather than adding another layer to supervise, they can spend more time there.

The career path question is the same one knowledge work faces: the work that trained junior project managers is the work AI now does, and the developmental path has to be redesigned.

Risks. The first risk is the loss of relational context. A human project manager who has been close to the work knows things that do not show up in any project artifact. They know who is on the verge of burnout, which stakeholder is quietly resistant, which dependency is shakier than it looks on paper. AI does not have this knowledge by default. If the project manager's role thins too much, this context gets lost, and the project starts running on artifacts that miss what is actually happening.

The second risk is over-automation of communication. AI-drafted stakeholder updates can be competent and still be generic. The relational work of project management, the moments when a real human reaches out at the right time in the right way, is difficult to automate because its meaning depends on trust and context. A project run primarily through auto-generated updates risks losing signals that deliberate human communication would catch.

The third risk is plan brittleness. A plan maintained by AI looks more complete than a plan maintained by humans because AI is good at filling in detail. The completeness can be deceptive. A plan that is detailed and current but built on wrong assumptions is more dangerous than a plan that is sparse but built on right assumptions.

Innovation

Traditional form. Innovation in most companies is structured as some form of separated process. Innovation labs, R&D departments, dedicated time, hackathons, ideation sessions, design sprints. The premise is that the day-to-day work and the innovation work are different things and have to be separated, because the day-to-day work crowds out the innovation work otherwise.

AI-Native form. Innovation becomes less separable from the day-to-day work. AI can lower the cost of exploring ideas, prototyping concepts, and testing approaches enough for more of these activities to happen inside ordinary work. For some digital problems, a team can produce a useful prototype in hours rather than weeks. A leader can explore an initial hypothesis in minutes rather than waiting days for a first analysis. The cost of trying can fall enough that experimentation becomes a normal part of work rather than a special event.

Dedicated innovation processes keep their value. Their character changes. Innovation processes in an AI-Native company are less about generating ideas, which now happens everywhere. They

are more about evaluating ideas, scaling promising ones, and protecting time and attention for the kinds of innovation that need sustained focus.

The shift. The shift is from innovation as event toward innovation as continuous practice. Many companies have not redesigned their innovation processes for this, so AI-enabled experimentation across the workforce remains uncaptured, uncoordinated, and largely invisible to leadership. Companies that get this right can build a much larger and more legible innovation pipeline than companies that do not, even when their nominal investment is similar. The size of that advantage must be measured rather than assumed.

There is also an honest limit to name. Current generative AI is especially useful for recombining existing knowledge, exploring within a frame, and accelerating well-defined parts of innovation. Whether a model can contribute to a genuine breakthrough is harder to separate from the people, tools, data, and selection process around it. The combinatorial strength is real. So is the risk of mistaking fluent recombination for original insight. A company should use AI to broaden exploration while protecting the human attention, experimentation, and dissent from which first-principles advances often emerge.

Stack layers engaged. Variable, depending on the kind of innovation work. Layer 4, knowledge graphs, becomes especially important for innovation in operations-heavy environments where understanding how the existing system works is foundational to changing it. Layer 6, agents, for sustained explorations of solution spaces. Layer 7, the collaboration surface, for innovation work to be visible and shareable across the company.

People implications. The innovation lab, the dedicated R&D team, the chief innovation officer, all remain, and their function changes. They become less about being the place where innovation happens and more about being the place that supports innovation happening everywhere. Their work is to provide infrastructure, governance, evaluation criteria, and the protected space for the kinds of innovation that the distributed model does not produce.

For the rest of the workforce, the implication is that innovation becomes part of the job description, in a way it usually was not. The skill of seeing a problem and immediately exploring whether there is a better way becomes a default behavior rather than a special skill.

Risks. The first risk is innovation fragmentation. A thousand small experiments running in parallel produce a lot of motion but not necessarily a lot of result. Without curation, evaluation,

and the discipline of choosing which experiments to scale, the company produces a graveyard of half-finished prototypes.

The second risk is that the combinatorial strength of AI is mistaken for sufficient creativity. Companies that rely on model output as the whole innovation process risk producing work that is technically novel but conceptually unsurprising. Human first-principles thinking, empirical testing, and informed dissent are counterweights, and they require deliberate effort to maintain.

The third risk is innovation theater, where AI enables a lot of activity that looks like innovation but does not translate into changed products, services, or business outcomes. This is the same risk as the pilot graveyard in Section 1.2, at a finer-grained level.

The Others, in Brief

The full operating system inventory is longer than these four processes.

Collaboration processes. How humans and AI work together inside teams, how AI is present in meetings, how artifacts are produced and reviewed, how feedback flows.

Customer-facing processes. How AI participates in customer interactions, where humans remain in the loop, how the brand voice is held across AI-mediated touchpoints.

People processes. How AI participates in hiring, performance management, development, succession planning, compensation. These processes are particularly sensitive because the human stakes are highest, and the failure modes when AI gets these wrong are severe.

Operations processes. How AI participates in the ongoing running of the business, the supply chain, the production lines, the service delivery, the operational decisions that happen every day. This is the domain where the Operating Layer archetype lives most fully.

Governance processes. How the AI itself is governed. The evaluation, monitoring, drift detection, hallucination management, decision rights enforcement, audit trail discipline. Layer 8 of the stack from the operating side.

Each of these deserves the same six-field treatment. The work of building the full inventory is one of the deliverables of a serious AI-Native operating model engagement. The four treated here are the most universally applicable across companies.

The next section deepens the treatment of one of the four. The new knowledge work loop is the deepest single shift in how work gets done. It benefits from being drawn explicitly: each step named, each role specified, each risk identified. That is Section 3.3.

Section 3.3

The New Knowledge Work Loop

How does knowledge work change when AI is in the room?

Section 3.2 sketched the traditional form of knowledge work in outline: an assignment, a blank page, a draft, iteration, review, delivery. The new loop is best understood against it.

The basic loop survived the arrival of word processing, search, and collaboration software. Those tools changed speed, access, and coordination, but the knowledge worker remained the primary engine of the draft. Everything else was support.

The AI-Native loop is different in kind because the tool now produces substantive candidate work. The shape of the cycle changes, the human role at each step changes, and the failure modes change with it. I want to draw the new loop explicitly. Many organizations are already operating inside some version of it without having named the change, and their practices vary as a result.

The new loop has five steps.

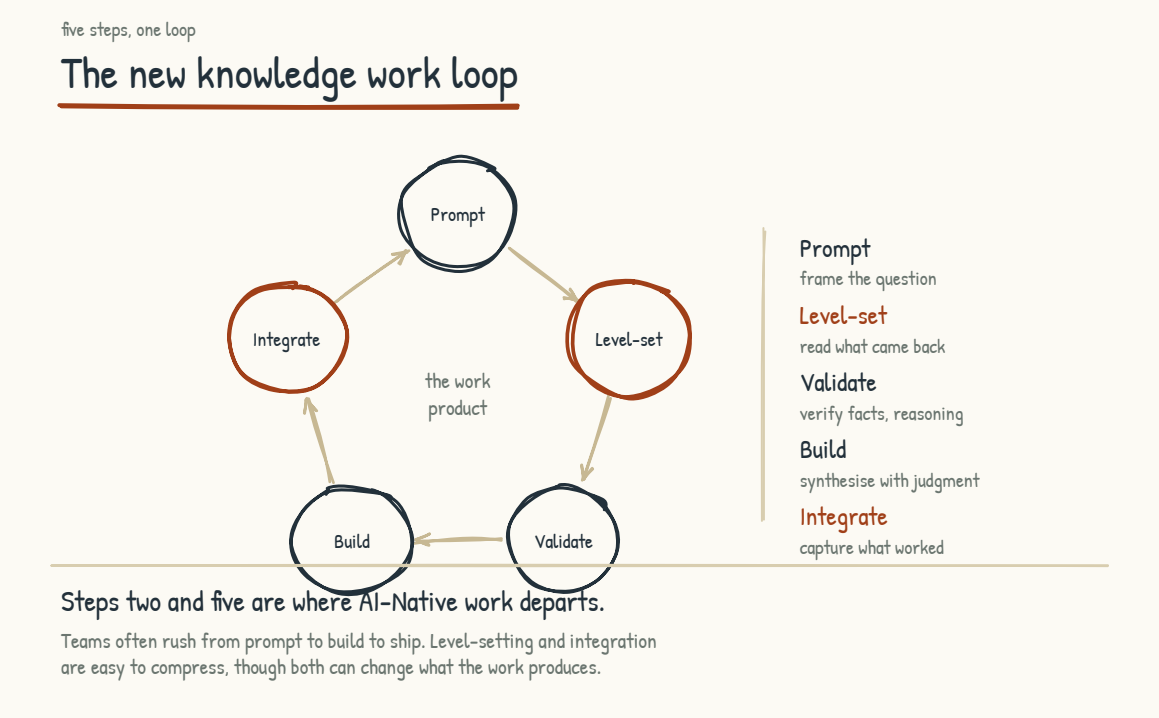

Figure 3.3a. The new knowledge work loop. Five steps, sequential but iterative. Steps 2 and 5 (in red) make explicit the correction and organizational learning that AI-mediated work requires.

Step One. Prompt

AI-mediated work often begins with an instruction rather than a blank page. That instruction may be a conversational prompt, a structured brief, a specification, or a task assembled by software. It frames the question and orients the system to what is needed. Its quality strongly shapes what follows, along with the model, tools, supplied context, and surrounding workflow.

The human role at this step is framing. What problem are we actually trying to solve. What does success look like. What constraints apply. Who is the audience. What has been tried before. What is the context that someone outside the situation would not know. This is the work of being clear about what is being asked for, which is harder than it sounds and which most organizations do badly.

The AI role at this step is initial response. Based on the prompt, the AI offers a first take, a draft, a structured approach, a set of options, depending on what the prompt requested.

The risk at this step is prompt-as-shortcut. A hasty, generic, or context-poor instruction makes shallow output more likely, even though a capable system may sometimes recover by asking questions or using tools. The output can look impressive, especially to someone who does not

know the subject well, while still being wrong for the work. Garbage in, plausible garbage out. Poor framing produces bad work that may escape diagnosis because it looks good on the surface.

Step Two. Level-Set

A first AI output is rarely the final answer. Treat it as a starting point. Your job at this step is to read it critically, identify where it landed well, where it landed wrong, what context the AI did not have, what assumptions the AI made that need to be revised.

The human role is critical reading. Reading with the kind of attention that the human would give to a colleague's first draft, with the explicit understanding that the AI does not know what the human knows about this specific situation and may have gotten meaningful things wrong.

The AI role is revision without ego. The system can challenge a conflicting instruction, surface missing information, or explain the basis of an answer, but it does not need to defend a draft for social reasons. The human is the authority on intent and accountable judgment, not automatically on every fact. The human's correction and added context become inputs to the next pass.

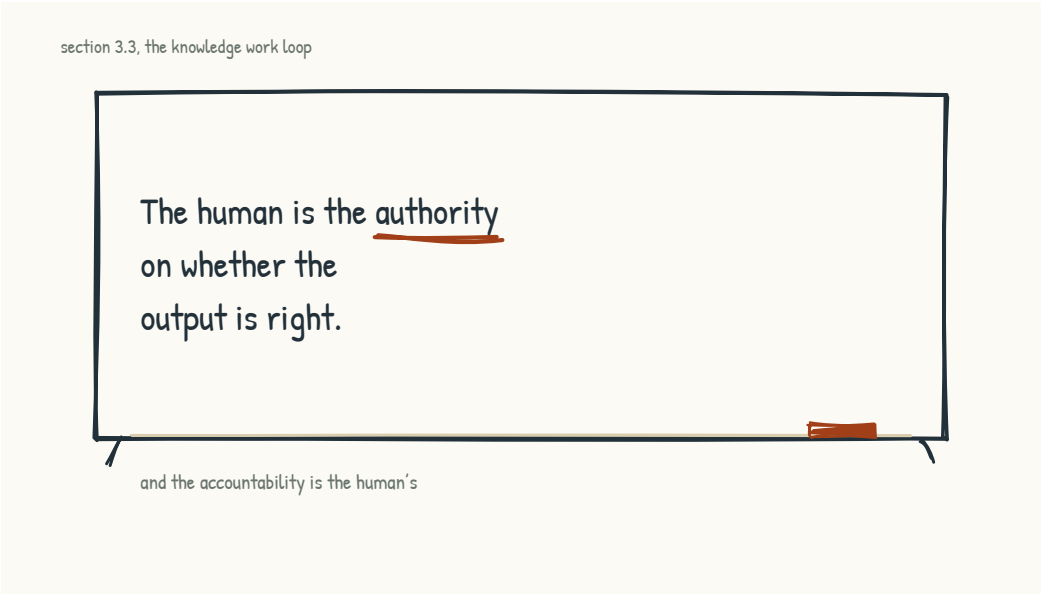

Figure 3.3b. The rule the whole loop rests on.

The risk at this step is acceptance bias. People can accept AI output too readily, especially when it is articulate and confident. The discipline of holding the output at arm's length, looking at it

skeptically, and saying explicitly "this is not quite right because" is harder than it sounds. It is a central failure point in AI-mediated work. Section 5.4 develops this risk in detail.

Step Three. Validate

The human has revised the framing. The AI has produced a revised output. Now the human has to validate that the output is correct on its facts, sound on its reasoning, and appropriate for its audience.

The human role is verification. Check the facts that the AI asserted. Check the sources, if any are cited. Check the argument, assumptions, and calculations. Check that the conclusions actually follow from the premises. Look for the kinds of errors AI characteristically makes: confident assertions about things it does not know, plausible-sounding citations to sources that do not exist, smooth narratives that paper over real complexity.

The AI role is to produce inspectable support for what needs to be verified: source links or document references, stated assumptions, calculations, intermediate artifacts, test results, and clearly marked uncertainty. A fluent explanation of the model's reasoning is not proof that the underlying process was faithful or correct. Validation should rest on evidence and reproducible checks, not on the persuasiveness of a generated rationale.

The risk at this step is verification debt. Validation takes effort. In a hurry, humans skip it. Skipped validation accumulates as debt that eventually surfaces, often badly, when the output is used somewhere consequential and the error becomes visible. For consequential work, build validation into the loop as a required step, scale its rigor to the stakes, and allocate enough time to do it properly.

Step Four. Build

Validated output becomes the input to the actual work product. In the build step you integrate what the AI contributed with your own knowledge, judgment, voice, and the human context it does not possess on its own.

The human role is synthesis and authorship. At this step, responsibility for the deliverable becomes unmistakable. The human has used AI to accelerate the work, but the human decides what gets delivered. The intended voice is the human's. The judgment calls are the human's. The accountability is the human's, subject to whatever review and governance the organization requires.

The AI role is supporting role. The AI may continue to participate at this step, drafting specific sections on request, exploring alternative phrasings, surfacing related information that the human had not requested. But the human is now the lead.

The risk at this step is authorship erosion. If the human incorporates AI output without genuinely engaging with it, unexamined model choices begin determining the work while the human remains accountable for it. That is bad for the work product, because the system may lack context the human holds, and bad for the human, because they have bypassed the work that would have exercised and developed their judgment.

A useful diagnostic question at this step is: could the human defend this output in a conversation where someone challenged its premises? If the answer is yes, the human has done the build step well. If the answer is no, the human has copy-pasted the AI's output and put their name on it, and is at risk on multiple fronts.

Step Five. Integrate

The work product is delivered. Reflection and organizational learning existed before AI; what is new is the volume of reusable instructions, evaluations, corrections, and interaction traces the work can now generate. The integrate step makes appropriate capture deliberate, subject to privacy, permission, retention, and security limits. Many organizations are not yet doing it consistently.

The human role is reflection. What did we learn from this cycle that should inform the next one. What prompts worked. What prompts did not. What kinds of validation effort were worth it. What kinds of integration moves produced better outputs. This learning is what turns one good cycle into a practice that improves over time.

The AI role, where the infrastructure supports it, is contributing to organizational learning. Prompts that worked well can be saved to a shared library. Approaches that produced strong outputs can be templated for reuse. Errors that emerged can be flagged for the team. The collaboration surface from Layer 7 is what makes this possible.

The risk at this step is the absence of integration entirely. Teams can complete the cycle and move to the next one without capturing what was learned. The same mistakes then recur, useful instructions get reinvented, and organizational practice fails to improve even as individual workers become more skilled. This is a pervasive and easily overlooked form of underinvestment in AI capability.

Reading the Loop

The five steps are sequential but they are not always neat. In practice the loop iterates. The validate step often surfaces problems that send the work back to the prompt step for re-framing. The build step often surfaces missing context that sends the work back to the level-set step. The integrate step often produces learning that improves the prompt step on the next cycle.

The loop is also not the only loop. Some kinds of knowledge work follow this pattern. Other kinds do not. A long-form strategy document follows this loop. A spontaneous Slack response does not. A research-heavy analysis follows this loop. An immediate client question may not. The loop is the dominant pattern of substantive knowledge work in an AI-Native operating system rather than a universal template.

What matters is that the loop exists, is visible to the people doing the work, and is taught explicitly. Organizations often give people AI tools and expect sound practice to emerge on its own. It emerges unevenly. Some workers develop a rigorous version of the loop and produce good work. Others skip steps and produce work that is fast and shallow. The variation is large, and the organization captures less value than it could because the practice is not shared.

A leadership team that takes AI-Native work seriously will teach this loop, build the supporting infrastructure for it, evaluate whether the work meets its standards without turning private activity into surveillance, and adjust as the practice evolves. Calling this training undersells it. This is operating system design.

The Connection to AI-for-Thinking

The new loop is the operational expression of a deeper cognitive posture. Knowledge workers who use AI well use it for thinking. They engage with the output, push back on it, and integrate it with their own judgment, creating the possibility of work better than either could produce alone. Knowledge workers who use AI poorly use it as thinking. They prompt, accept, integrate without engagement, and deliver. The first posture exercises judgment. The second risks allowing judgment to atrophy.

Section 5.4 treats this distinction at length, because it is one of the most consequential ideas in this book. The point here is that the new knowledge work loop is the operational expression of the AI-for-thinking posture. Doing the loop well requires the cognitive posture that Section 5.4 will describe. The two are inseparable.

AI Software Factories: A Worked Example

The new knowledge work loop generalizes across kinds of work. A worked example in one specific domain shows what it looks like when the loop is taken seriously and the supporting architecture is built around it.

Software development has advanced quickly, partly because code can be executed and many requirements can be expressed as tests. An emerging pattern is often described as an AI software factory. In its stronger form, this is a working architecture rather than a slogan: specifications and tests guide AI agents that generate and iterate on code, while people review the work and retain control over what ships. Software is unusually suited to this pattern, but even here an incomplete specification or test suite can certify the wrong result.

The factory can be understood through four activities, whether separate agents perform them or one agent cycles among them. Generate produces candidate code from specifications. Test runs appropriate checks and reports failures. Refine addresses failures, adds missing cases, and iterates. Audit records changes, test runs, approvals, and important decisions so people can review both quality and pattern. Parallelism is an implementation choice, not a requirement of the architecture.

The humans on the loop do specific work. They own the specifications and acceptance criteria. They ensure the tests represent the intended behavior, whether people or agents drafted those tests. They review consequential outputs, refine the brief, and rerun the system when the result is not what they intended. They may still code by hand where that is faster, safer, or more illuminating. The role shifts from producing every line toward setting the conditions under which code is produced and deciding what is acceptable.

The pattern can extend to domains beyond software, with significant adaptations. Marketing teams can build campaign pipelines, legal teams contract-review pipelines, and operations teams incident-response pipelines. But prose, legal judgment, and operational action are harder to grade than code with a strong test suite. The common element is not necessarily a four-agent topology. It is a generate, evaluate, refine, and audit loop, with human authority placed according to consequence.

The factory is one form the new knowledge work loop can take at scale. The infrastructure is real and reported gains can be meaningful, but results depend on specifications, environments,

evaluations, review, and the kind of work. A deliberate shared system can accumulate learning in a way that isolated one-off prompts cannot.

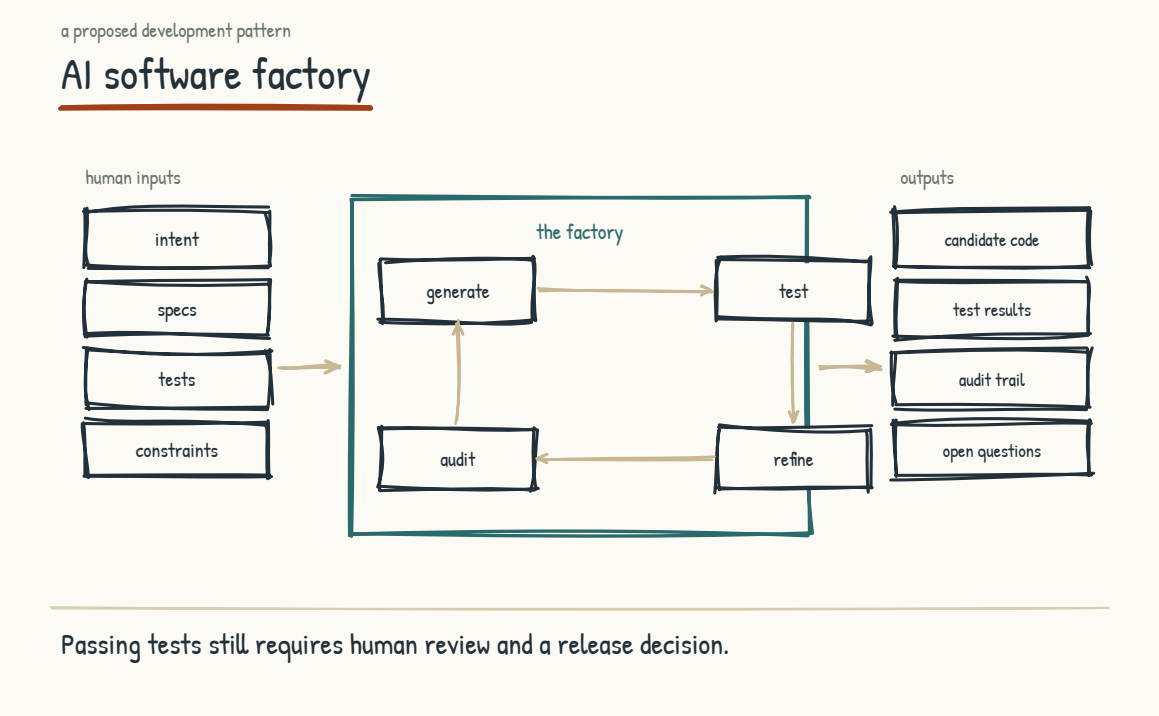

Figure 3.3c. AI software factories. Specifications and tests guide candidate code through generation, evaluation, refinement, and human-controlled release.

Stage 3 has built the operating system frame. Four dimensions, the inventory of processes, and the new knowledge work loop. How an AI-Native company actually runs at the practice level is now in view.

STAGE 4

The Architectures

The technical decisions that determine whether AI compounds or fragments.

Section 4.1

The Knowledge Graph as Central Architecture

How does AI come to understand how a company actually works?

How institutional knowledge gets structured for AI to use is among the most important architectural decisions an AI-Native company makes. In the companies I see, it is often made by default. The default is the lowest-effort option: put documents into a retrieval system and let AI search them at query time. That option can create substantial value. It also encounters a wall when the questions depend on relationships and operational state the documents do not represent reliably.

The wall is this: documents describe the work. They describe parts of it, from particular perspectives, with assumptions about who will read them and what those readers already know. A policy document is an authoritative expression of policy, but it may not encode every condition in executable form. A process map represents the process at a particular moment, by a particular author, for a particular audience. The structure of the work is distributed across documents, databases, event histories, system configurations, observed behavior, and experienced people: entities, relationships, constraints, dependencies, and soft knowledge. Some of it is in documents. Much of it is not explicit, current, or connected enough for a system to apply reliably.

A company that wants AI to work across that structure has to make the relevant parts explicit and keep them current. A knowledge graph is one powerful artifact for doing so, especially when relationships and dependencies are central to the questions. What follows explains what that means, why it matters, and where it changes the architecture.

What a Knowledge Graph Is

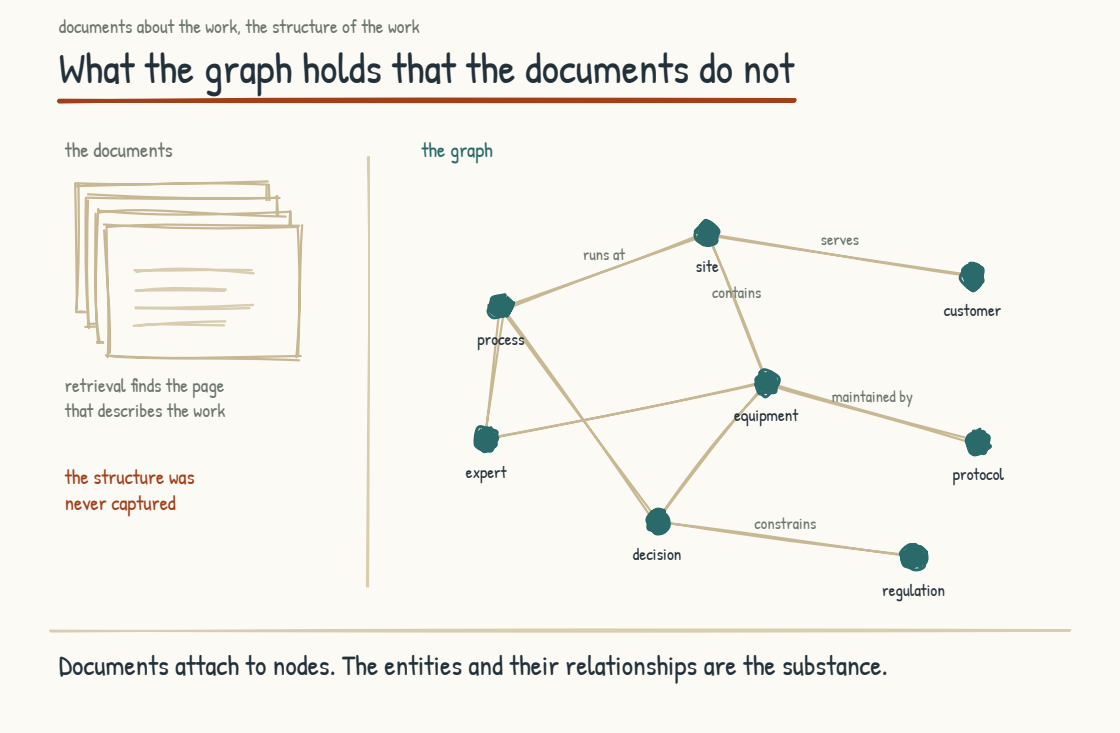

Figure 4.1a. What the graph holds that the documents do not. Each node an entity, each edge a relationship, with documents attached rather than standing in for the structure.

A knowledge graph is a structured representation of how the entities in a company connect to each other. The entities are the things the work involves. Systems, processes, people, decisions, products, customers, contracts, regulations, equipment, sites, projects, programs, risks, suppliers, dependencies. The connections are the relationships between them. This system feeds that one. This process depends on that one. This expert holds the knowledge for that operation. This decision is constrained by that regulation. This customer's account is managed by that team. This piece of equipment is maintained according to that protocol.

A conventional relational database stores records and can also encode meaning through schemas, keys, and joins. A graph database makes relationships first-class: each node an entity, each edge a typed relationship, the whole a machine-navigable map optimized for questions

about connection and traversal. The distinction is not records versus meaning. It is how explicitly the data model represents relationships and how naturally the system can query them.

Documents still have a place. They attach to nodes in the graph. The policy document attaches to the policy node. The process map attaches to the process node. The training manual attaches to the role node. But the documents are artifacts about the entities. The entities and their relationships are the substance.

This distinction is one of the most consequential ideas in this book. In the companies I see, institutional knowledge is often treated primarily as a documents problem. They invest in document management, search, and retrieval-augmented generation. Those investments are useful and that work is real. But some use cases stall because the underlying relationships, definitions, and operational state have never been structured. The system is doing the best it can with the substrate it was given, and the substrate becomes the limit.

RAG and the Knowledge Graph Compared

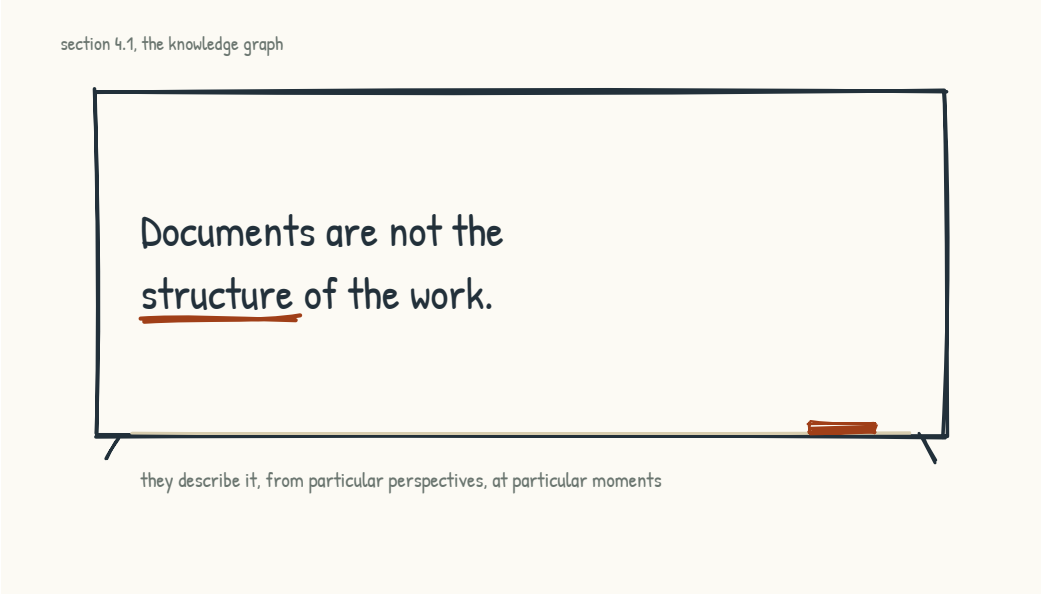

Figure 4.1b. Retrieval finds the page. The structure of the work lives somewhere else.

Retrieval-Augmented Generation, a common architecture for grounding AI in a company's own data, is Layer 3 of the nine-layer AI stack from Section 2.3. In its first form the architecture was simple. The user asked a question, the system fetched the closest-matching documents from the company's repository, put them in the context window, and the AI answered from what it was handed.

Some newer systems work harder than that. The model gets search as a tool and can run it repeatedly, choosing what to look for, inspecting what comes back, and deciding whether another search is warranted. Section 2.3 covers what that can buy and what it costs. Assume that stronger version for the rest of this section. Retrieval quality can keep improving while the document-only ceiling remains: relationships or current state that were never represented clearly cannot be recovered reliably merely by searching the prose harder.

RAG works well for questions that have answers in the documents. What is our policy on remote work. What did we agree to in the contract with this vendor. What were the key findings of the last quarterly review. These are document-retrieval questions. The answer is in a specific place. The system finds it. The AI summarizes it. The user gets the answer.

RAG works less well for questions that require reasoning across the structure of the company. If this regulation changes, which products are affected. Which customers do we have whose contracts expire in the next ninety days and who have an open support ticket about a related issue. Which experts have institutional knowledge about this kind of decision that we should consult. If we move this team from this site to that one, what dependencies will be disrupted.

These are relationship and traversal questions. They require the system to follow connections between entities, evaluate second-order and third-order effects, and identify patterns spanning multiple sources and parts of the business. Document-only semantic retrieval is a poor fit when those connections are absent, ambiguous, or scattered across prose. More capable retrieval systems can combine documents with SQL, metadata filters, APIs, and iterative tool use, so the boundary is not RAG versus everything else. The question is whether the relevant relationships are represented explicitly somewhere the system can query.

A knowledge graph holds those connections explicitly and makes multihop traversal natural. The graph does not reason by itself; the model, query engine, rules, and surrounding software reason with what the graph exposes. The same model, given reliable graph context, can answer relationship-heavy questions that document retrieval alone may miss. The architectural difference can be profound, particularly in operations-heavy systems, but documents, tables, time-series data, and graphs are complementary rather than rival substrates.

Why This Matters for Operations-Heavy Companies

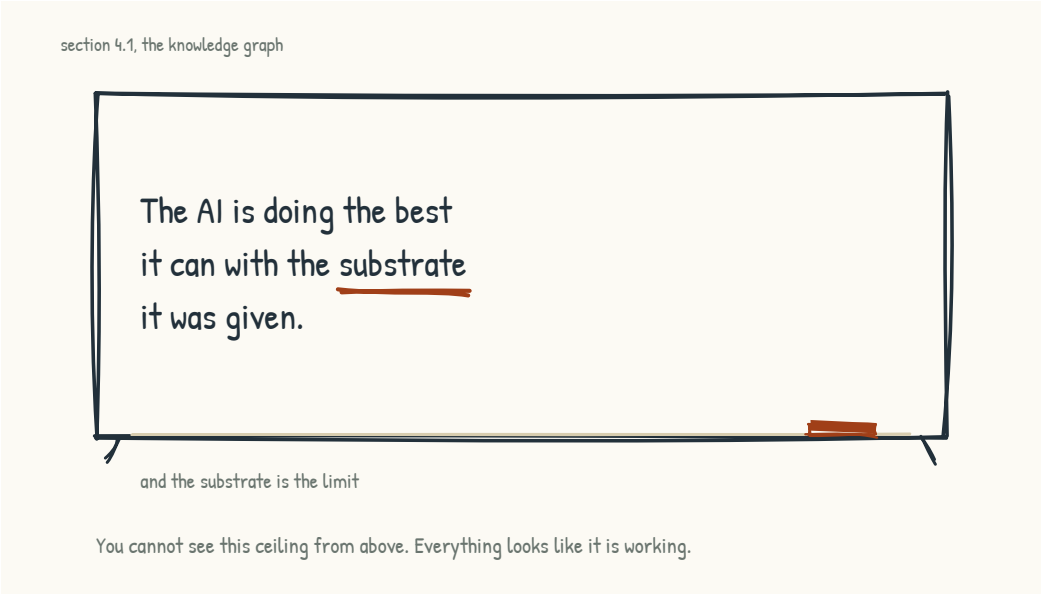

Figure 4.1c. The ceiling nobody sees from above.

A consulting firm can run a long way on retrieval. Much of its output and institutional record is documentary, and many questions are document-centered. A well-designed retrieval system can therefore move a consulting firm a long way toward AI-Native operations, although relationships, permissions, provenance, and current client state still matter.

A manufacturer cannot run safely on document retrieval alone. The work is the operation of physical equipment, the management of supply chains, the coordination of shifts, the maintenance of safety, and the optimization of throughput. The knowledge that matters is distributed across control systems, sensors, maintenance records, engineering models, procedures, and operators who have run the line for decades. It also lives in the patterns by which one plant differs from another and in failure signals experienced supervisors recognize before an incident.

A manufacturer that represents relevant relationships in a graph can build a capability a document-only system cannot easily reproduce. The graph may connect equipment and parameters, maintenance history, roles and qualifications, batches, incidents, and operating constraints. Combined with current sensor data, validated predictive models, and rules, the system can investigate why a machine is behaving differently under a particular set of

conditions, estimate failure risk, and support operators in real time. The graph supplies context and relationships; it does not by itself establish causation or predict failure.

The same logic applies to defense, healthcare, energy, logistics, and other operations-heavy environments. Much of the knowledge that makes the operation work is structural or state-dependent. Relevant parts must be represented in queryable forms to be used reliably by AI. Documents remain part of that evidence, but documents alone are rarely enough.

This is the architectural commitment that Nodalix exists to support. Knowledge graph construction for operations-heavy environments, integrated with virtual sensors, predictive models, and agent systems, is one technical foundation of the Operating Layer archetype from Section 2.2. A company in this archetype that has not invested in an authoritative relationship-centered model, whether implemented as a graph or an equivalent architecture, has not yet built the foundation for what comes next.

The Strategic Prize

The strategic prize of investing in a knowledge graph is institutional knowledge capture in a form that can be connected, queried, governed, and reused at scale.

Somewhere in a plant there is an operator who has run the same unit for thirty years. He can hear a compressor start to labor two days before any alarm sounds, and he knows which gauge reads a hair high on cold mornings and means nothing. He retires in the spring, and the plan for his thirty years is a farewell lunch, an exit interview, and a request to record a few training videos before he goes. The videos will capture his face and his voice. Every operations-heavy company is losing some version of him right now, and the conditions that taught him are not being reproduced for anyone behind him.

A knowledge graph can capture part of the structure the operator has been navigating: the entities and relationships involved, the questions prompted by a pattern, the factors considered when conditions deviate, and the rules that constrain a decision. It cannot extract a lifetime of tacit judgment perfectly. This is hard work, and it requires experienced operators to be engaged in deciding what the representation gets right and what it misses. The company has to invest in them as teachers rather than relying on them to leave behind documents. It is one of the most undervalued investments an operations-heavy company can make. The value can compound if the structure is used, validated, and maintained as the operation changes.

This work can build a moat. Two competitors with the same vendors, foundation models, and retrieval systems will not necessarily perform identically; execution already differs. But a competitor that also accumulates validated domain structure, operating history, and feedback has an advantage another firm cannot acquire with a license. Over a five-to-ten-year planning horizon, that gap can compound and become difficult to reverse. The outcome is not guaranteed, and a stale graph is not a moat. The investment window is open now.

The Work of Building a Knowledge Graph

Many technical components are mature: graph databases, ontology languages and modeling tools, and integration pipelines. Production graph systems still face hard engineering problems: identity resolution, temporal state, permissions, provenance, data quality, and change management. Harder still is capturing and validating the institutional knowledge that goes into the graph.

The work involves the experienced operators in interviews that surface what they know. The work involves modelers who can translate that knowledge into entities and relationships. The work involves architects who design the graph in a way that supports the queries the AI will need to make. The work involves governance to ensure the graph is maintained as the operation evolves. The work involves continuous validation as the graph is used.

This is a craft. The people who do it well are rare. Build durable ownership internally, with patient investment, and it can become an advantage. External specialists can accelerate the work, but no consultancy can substitute for domain experts, accountable internal owners, and people who will maintain the model after the engagement. Without them, the graph comes back shallow or grows brittle.

A company that takes this seriously and invests accordingly can build an institutional knowledge architecture that survives individual applications and model cycles. Skip the structural work where the use case requires it and rely on documents alone, and the wall returns repeatedly. The longer definitions, identities, and dependencies remain unresolved, the more expensive they can become to reconcile.

Connection to the Rest of the Book

A knowledge graph can be central architecture beneath an AI-enabled operating system when the work depends on dense relationships, multihop questions, shared definitions, or changing operational state. The processes treated in Section 3.2 require an underlying structure that

supports grounded reasoning, but that structure may combine graphs with relational databases, event streams, documents, rules, simulations, and domain models. Decision-making, operations, and innovation require authoritative context. They do not all require the same storage technology.

The knowledge graph is also the architecture beneath the human work. The institutional knowledge that must be honored rather than discarded is the knowledge the graph captures. The experienced operators at the center of the human reckoning that Section 5.2 will name are the experts whose knowledge the graph represents. The covenant that comes out of that reckoning is enabled by the work of capturing knowledge structurally. That work treats the experienced workforce as teachers rather than as overhead.

Where the use case calls for it, the graph is both technical foundation and values commitment. A company that builds one well has decided to take institutional knowledge, provenance, and relationships seriously. A company without a knowledge graph may have made a sound architectural choice for simpler use cases, or it may not yet have recognized that critical structure is going uncaptured. The test is not whether the company bought a graph database. The test is whether its systems can represent and govern the knowledge the work actually requires.

The next section takes the graph and everything around it and asks the design question directly: what are the ingredients of an AI-Native operating model, and what shapes can they form.

Section 4.2

Designing the AI-Native Operating Model

How do you redesign the way work gets done when AI is now part of it?

Section 3.1 introduced AI as a dimension of the operating model, alongside people, process, and technology, and made the conceptual case for it. If AI is a dimension of the operating model, what does it look like to actually design an operating model that includes it? What are the building blocks the leadership team will be choosing from, and what patterns do they form?

The answer has two parts, in sequence.

The first part is the ingredients. Before you can design an operating model that includes AI, you have to understand what the elements are. All of them, the AI elements and the rest. The humans and what they bring. The AI in its various forms. The data and knowledge the company holds. The connected systems that already run the business. The standards and governance that hold everything together. Five categories of ingredients. The leadership team will be making choices within each.

The second part is the sketches. With the ingredients understood, the section walks through five different operating model designs that emerge when AI joins. They are not the archetypes from Section 2.2, which name what kind of AI-Native company you are; these name how the ingredients get assembled. Each sketch combines the ingredients differently and reflects a different priority. None of them is the right answer. Each one is a worked example of how the ingredients can be combined into a coherent design. The point is to see the variety, understand the trade-offs, and use the sketches as starting points for your own design work.

There are well-established operating model frameworks the leadership team may already know. The MIT CISR archetypes from Ross, Weill, and Robertson. Galbraith's Star Model. The Operating Model Canvas. These remain useful and the work in this section does not replace them. What AI changes is that the design space expands. New designs become possible because AI as a dimension of the operating model produces options that did not exist before. The sketches that follow show several ways to assemble the expanded set of ingredients. Read them as additions to the older frameworks rather than as substitutes.

The Five Ingredient Categories

The five categories below are how I have come to organize the elements that go into any AI-Native operating model design. Other organizations of the same elements are possible. The five-category structure has held up in the workshops I have facilitated because it covers the design space without becoming overwhelming.

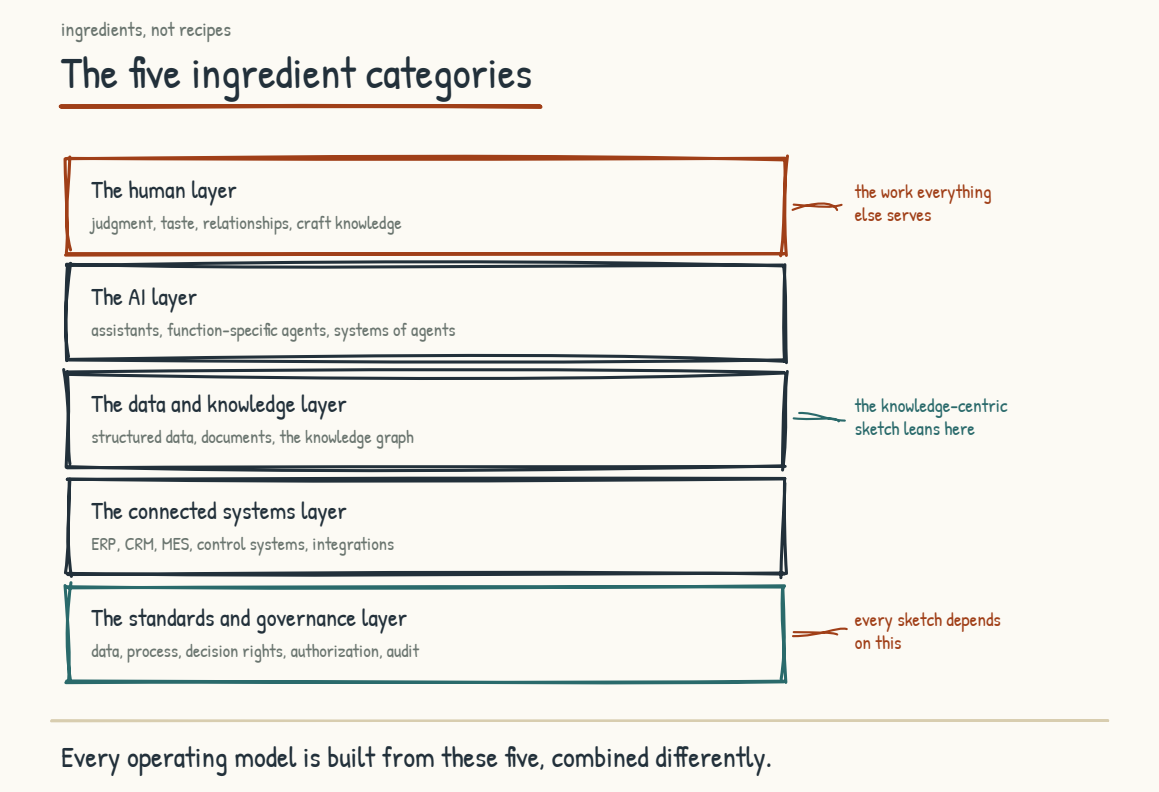

Figure 4.2a. The five ingredient categories that go into any AI-Native operating model design. Standards at the foundation, the human layer at the top, with three layers in between.

Category One. The Human Layer.

The human layer is what the people in the company bring to the work and how they are organized to do it. AI joins their work rather than simply replacing the layer. The design choices in this category include both what each individual human brings and how those humans are configured.

What individual humans bring includes their skills and expertise, the deep knowledge they have developed through years of doing the work. It includes their judgment, which is what they apply when the rules and frameworks do not cover the situation in front of them. It includes their relationships, with colleagues, customers, suppliers, regulators, the network of people who make the work flow. It includes their creativity, the capacity to frame problems in new ways. It includes their accountability and ethical responsibility, the willingness to own outcomes. It includes their tacit knowledge, the unwritten know-how that sits in their heads and that the company would lose if they walked out the door. Section 5.3 explores this in more detail.

How humans are organized for work is a separate set of choices. The reporting hierarchies and management layers. The functional departments. The cross-functional teams. The decision

authority and escalation paths. The communication rituals, the standups, the all-hands, the one-on-ones. These structures shape how human capability gets applied to the work. The same humans, organized differently, produce different outcomes.

The leadership team designing an AI-Native operating model will be making choices about both. Which human capabilities are most valuable in this company, where in the organization they sit, and what structure surrounds them.

Category Two. The AI Layer.

The AI layer is the toolbox of AI capabilities available to the company. The toolbox is wider than most leadership teams initially recognize. Talking about "AI" as one thing is the first mistake that produces bad design.

The AI layer includes individual AI assistants: personal-level systems that can be given scoped memory of an employee's work, adapt to preferences, and take authorized actions on that person's behalf. Enterprise offerings from OpenAI, Anthropic, Microsoft, and others contain early forms of this pattern. The mature form, and the governance required to make it safe, is still being built.

It includes function-specific AI agents, scoped to a particular role or workflow rather than to a person. A sales-development agent. A code-review agent. A recruiting agent. A customer-service tier-one agent. These are different from individual assistants because they exist to do a specific kind of work and serve whoever needs that work done.

It includes AI features embedded in other tools: the AI inside the document editor, email client, project-management software, or specialized professional application. For many knowledge workers, these embedded features may become the primary way they interact with AI because the capability appears inside work they already do.

It includes systems of agents: coordinated components purpose-built for domains such as supply chain, customer service, or recruiting. Vendors increasingly describe products this way, though some are genuinely multi-agent systems and others are workflows behind an agent label. The leadership team must look past the category name, decide which capabilities to buy or build, and understand how coordination, evaluation, authority, and failure handling actually work.

Each form of AI is a different kind of participant in the work. Treating them as one thing collapses real design choices into invisibility. The leadership team that recognizes the AI layer as a toolbox rather than a single tool has the first thing it needs to design well.

Category Three. The Data and Knowledge Layer.

AI is constrained by what it can reliably access, interpret, and act on. The data and knowledge layer is what the company has accumulated that can inform the system. This category includes structured and unstructured material, explicit records, and tacit knowledge that must be elicited rather than assumed.

The structured operational data lives in the existing systems. The transactions in the ERP. The customer records in the CRM. The personnel records in the HR system. The sensor data in the manufacturing systems. The order history. The financial ledgers. This is the data the company already has and that most AI initiatives begin by trying to make available.

The unstructured documents are the contracts, reports, decks, emails, meeting transcripts, technical specifications, training materials, and other documents the company has generated over years. Modern AI systems can ingest or extract content from many of these formats, but reliable use remains difficult because authority, recency, permissions, definitions, and relationships are often unclear. Section 4.1 covers how retrieval, knowledge graphs, and related architectures make this material more usable.

The historical knowledge is the company's accumulated record. The decisions made and why they were made. The customer interactions and their outcomes. The product launches and how they went. The hires and how they performed. This is the institutional memory in a literal sense, the record of what has happened.

The reference materials are the procedures, playbooks, policies, and training content the company has produced for its own use. These are the codified knowledge, the part of institutional knowledge that someone took the trouble to record.

The institutional knowledge in people's heads can be among the most valuable and is often the hardest to access. The senior plant manager who knows why a specific line behaves the way it does. The relationship manager who anticipates what a customer will ask. The veteran engineer who knows which corners of the codebase demand care. Section 8.2 names the failure mode of trying to capture this knowledge after its holders have been signaled out of relevance. Useful

capture depends on their willing participation and validation; extraction without trust produces silence, compliance theater, or shallow artifacts.

The shared memory architecture is what holds all of this together. Knowledge graphs. Vector databases. Document repositories. Data lakes. The collaboration surface where AI work happens visibly to the organization. These are the substrate on which the data and knowledge layer is built. Section 4.1 and Section 4.3 cover the architectural choices.

Category Four. The Connected Systems Layer.

The connected systems layer is the technology infrastructure the company already runs. The systems of record. The systems of engagement. The systems of operation. The AI is being added to these systems, and the design problem includes how the integration happens.

Map the systems that hold financial, customer, workforce, and operational records, including the specialist systems your industry requires. For each, identify the authoritative data and the actions AI may take.

Beyond the systems that hold business data and process, there is the infrastructure that supports the work. Security and identity systems, the Active Directory or Okta that controls who can access what. Communication infrastructure, email and messaging. Data processing systems, the analytics platforms and BI tools.

The connected systems layer is where most companies have spent the last twenty years of technology investment. The AI architecture has to integrate with this layer because this is where the actual operational data and process lives. The integration is not optional. The leadership team's choices include which systems the AI architecture connects to first, what data flows in which direction, and which integrations are deep, meaning read and write, versus shallow, meaning read only or one-way.

Section 2.3 introduced MCP and Agent2Agent. Here the design task is to specify which systems connect, what each connection permits, and how the team will test an upgrade.

Category Five. The Standards and Governance Layer.

The standards and governance layer is what holds everything else together. It is the agreements about how data is structured, how requests are made, how decisions are recorded, who is allowed to do what. Without shared standards across the other four layers, the system fragments. With them, the system coordinates.

Data standards are the schemas, taxonomies, and ontologies that define how information is structured. What counts as a customer. How an account is named. What the fields of an order are. The same standards that have always mattered for data integration matter more in an AI-Native operating model because more parties (humans, agents, systems) are consuming the data and they need to agree on what it means.

Process standards are how work flows. The steps in the procurement process. The handoffs in the customer service workflow. The escalation paths for exceptions. Process standards have been the substance of operational excellence work for decades. AI changes what they enable rather than eliminating the need for them. Agents can work inside ambiguity, but predictable operation still requires explicit state, boundaries, handoffs, and exception paths.

Decision-making standards are who decides what, with what input, at what threshold. Who approves a vendor contract. Who escalates a customer complaint. What signals trigger a price change. These standards have always existed informally. AI-Native operating models make them explicit, because agents acting on behalf of humans need explicit authority rules.

Authorization and access standards are the rules about what each party can do. Which agents can read which data. Which agents can write to which systems. Which humans approve which actions taken by agents. The authorization model for agents is a separate problem from the authorization model for humans, and both have to be designed.

Audit and compliance standards govern the record of what happened: which system or agent took which action, under whose authority, against which data, using which version, and with what result. In regulated or high-consequence work, traceability may be mandatory or operationally indispensable. The applicable obligations vary: HIPAA rules for protected health information in the United States, PCI DSS for payment-card environments, GDPR for covered personal-data processing in Europe, sector-specific regulation, contractual controls, and assurance programs such as SOC 2. Logging alone does not create compliance, and sensitive logs create risks of their own. Traceability, retention, access, and deletion need to be designed together from the start.

The standards and governance layer is the connective tissue. Section 8.2 names the case of a company that built governance on prohibition and produced shadow AI everywhere. The lesson from that case applies here. Standards work when they make work easier.

The Five Sketches

Each sketch below combines the ingredients differently and reflects a different priority. Each is a worked example of an emerging operating model design that becomes possible when AI joins.

Figure 4.2b. Five emerging operating model archetypes. Each combines the ingredients differently and is suited to a different kind of company. Most real designs draw from more than one.

Sketch One. The Coordination-Centric Operating Model.

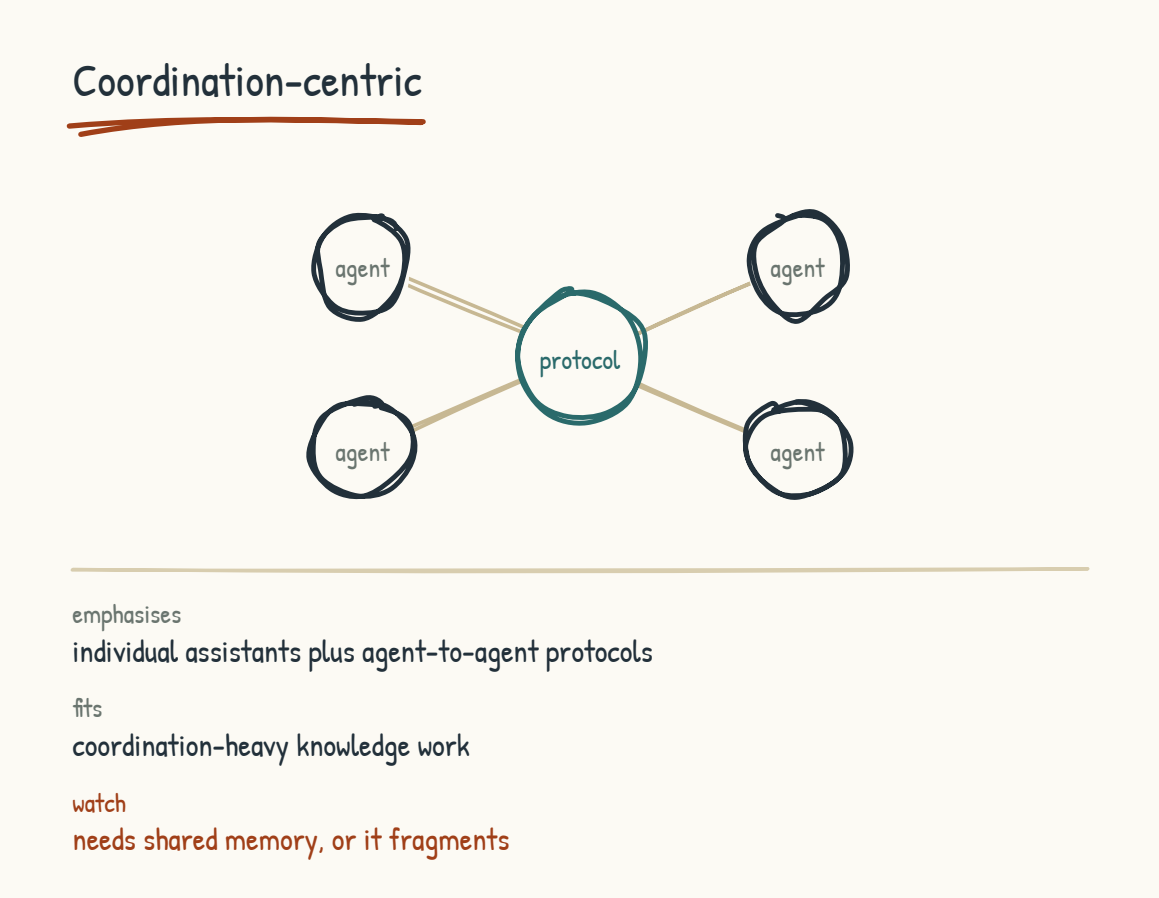

Figure 4.2c. Coordination-Centric: individual assistants coordinating peer-to-peer, a multi-agent nervous system.

This is the multi-agent nervous system. The AI is primarily organized to handle coordination work that used to be done by humans through meetings, email, and middle management.

The architecture builds from the bottom up. Employees whose work benefits from it have individual AI assistants with permissioned context about their work. Those assistants can coordinate across teams and departments to identify duplication, surface dependencies, and route information, but only within explicit identity, disclosure, and authority rules. At the enterprise level, they integrate with connected systems and domain-specific agent services.

The ingredients this sketch emphasizes: individual AI assistants as the foundation. The knowledge graph and collaboration surface as the shared memory. Agent-to-agent protocols as the connective tissue. Standards for how agents communicate and what authority each one has.

The ingredients this sketch treats as secondary: function-specific agents (the assistants handle most of what would be done by function-specific agents). The customer-facing surface (the architecture is internally focused).

The kind of company this fits: knowledge-work-intensive businesses where coordination loss is the primary inefficiency. Professional services firms. Consulting firms. Companies whose work depends on cross-functional collaboration. Companies where the middle management layer has thickened over time and decision speed has slowed.

The trade-offs: significant investment in individual assistant infrastructure before any visible business outcome. Heavy reliance on standards being established at the individual level before team and department coordination becomes possible. Vulnerable to the failure mode where individual assistants are built without the shared memory substrate, producing fragmented experience rather than coordination.

Sketch Two. The Workflow-Native Operating Model.

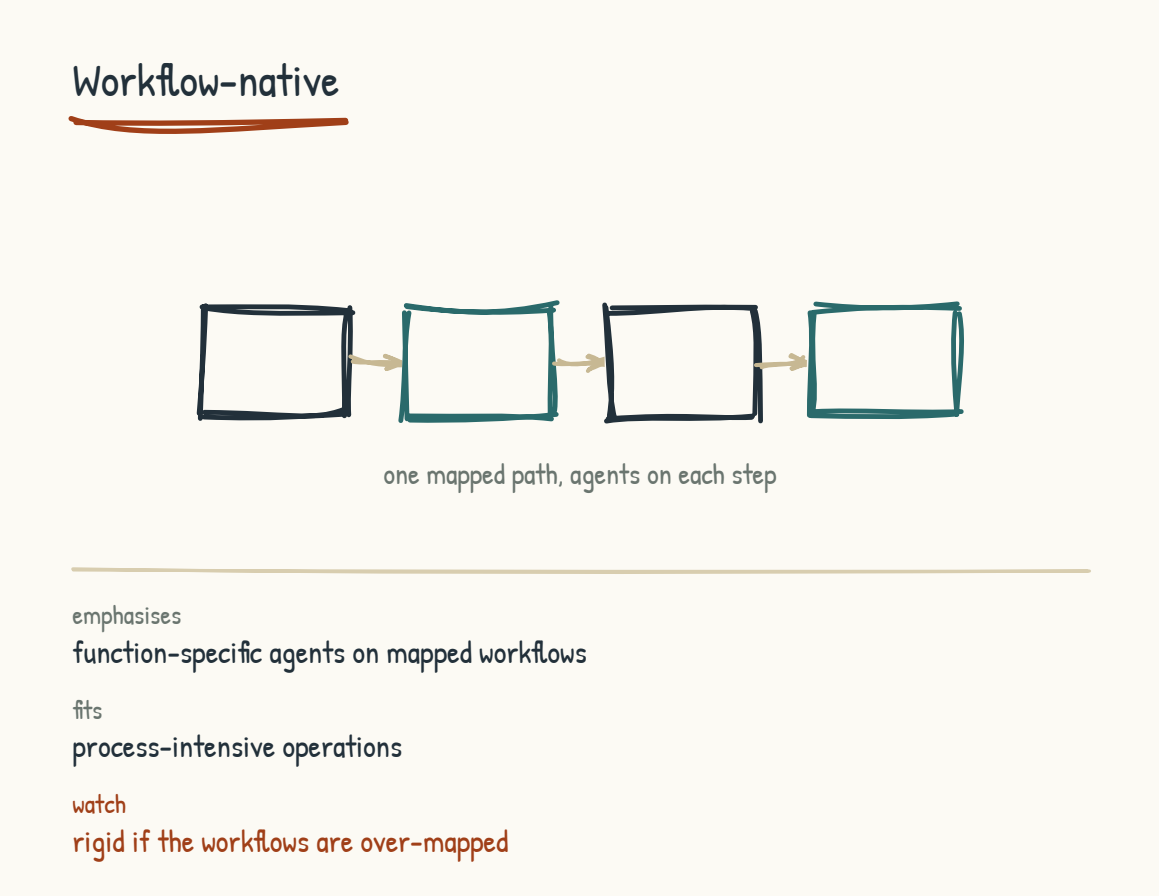

Figure 4.2d. Workflow-Native: function-specific agents own stages of a mapped workflow, with explicit handoffs.

This sketch organizes around end-to-end workflows rather than around individuals or teams. The leadership team maps the company's major workflows explicitly. Order-to-cash. Hire-to-retire. Design-to-build. Lead-to-customer. Service-to-resolution. Each workflow is then redesigned with AI inserted at the points where translation, integration, or judgment is needed.

The architecture is process-driven. Function-specific agents handle portions of each workflow; accountable people or teams still own the process and its outcomes. Agents are workflow participants with defined roles and clear handoffs. The handoffs between agents and humans are explicit, with decision standards determining when a system may act and when it must escalate.

The ingredients this sketch emphasizes: function-specific AI agents tied to workflows. The process standards that define how work moves. The connected systems layer, which is where most workflows already live. Authorization standards that govern when agents act on their own.

The ingredients this sketch treats as secondary: individual AI assistants (less central than in Sketch One because the workflow is the primary unit). The knowledge graph (still present, but as workflow context rather than as the central asset).

The kind of company this fits: businesses where operational excellence is the value driver. Manufacturing companies. Financial services firms. Logistics businesses. Healthcare delivery organizations. Any company whose work is process-intensive and where consistency of execution matters more than coordination flexibility.

The trade-offs: requires enough workflow mapping to identify state, ownership, exceptions, and controls, though discovery and limited AI investment can proceed together. Risks producing a rigid architecture if workflows are mapped without flexibility for unusual cases. Can become brittle when business conditions require changes faster than the architecture can adapt.

Sketch Three. The Knowledge-Centric Operating Model.

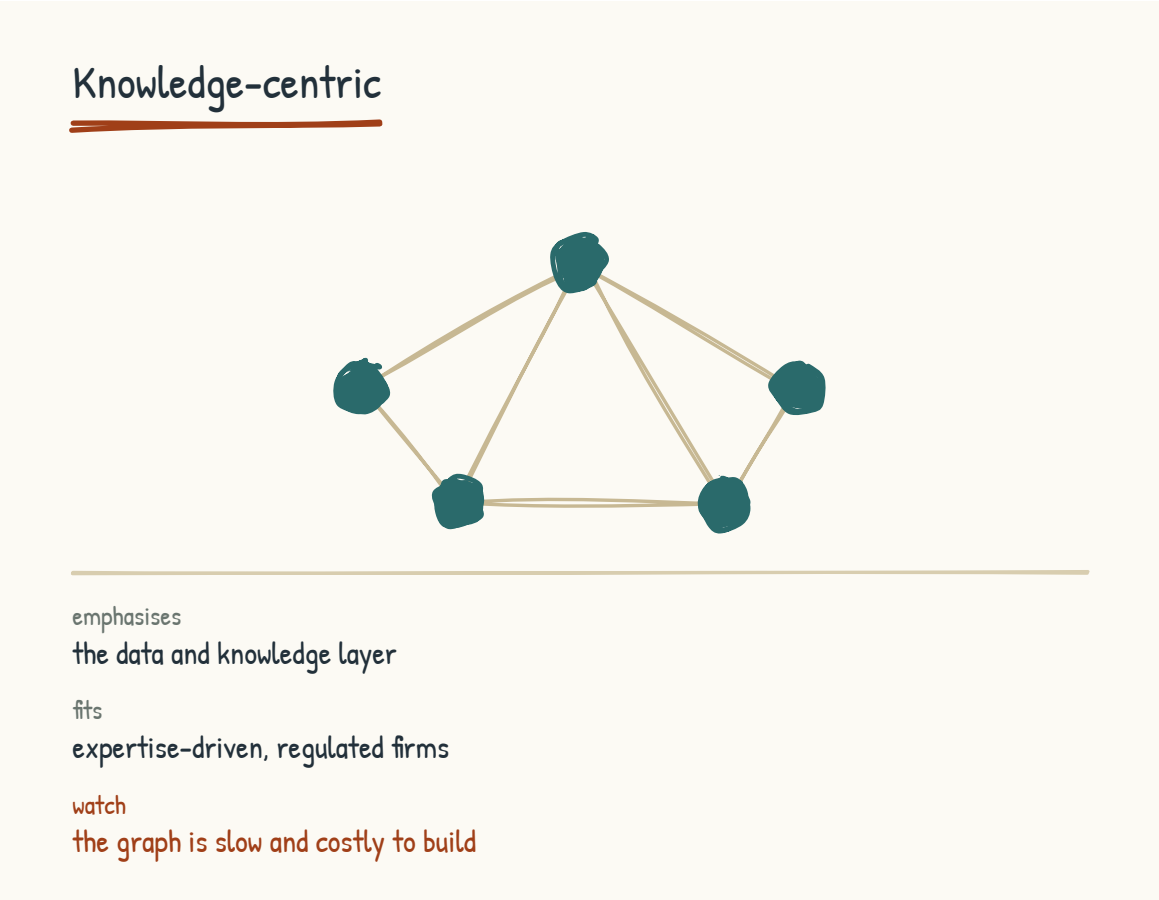

Figure 4.2e. Knowledge-Centric: humans, agents, and systems all read from and write to a central knowledge graph.

In this sketch, the knowledge graph or its equivalent is the central asset of the operating model. Humans, agents, and systems all orient around contributing to and consuming from the shared memory. The architecture privileges institutional knowledge as a durable competitive asset.

The architecture is memory-first. The knowledge graph represents selected parts of how the company works: decisions, customer relationships, product knowledge, operational details, and supplier relationships. Agents and people interact with it through governed interfaces rather than writing unchecked assertions directly into shared truth. Connected systems feed it with provenance, while authoritative systems of record remain authoritative where they should.

The ingredients this sketch emphasizes: the data and knowledge layer, especially the knowledge graph. The standards for how knowledge is captured and structured. Audit and compliance

controls can become more consistent when access moves through common architecture, although centralization also concentrates operational and security risk.

The ingredients this sketch treats as secondary: individual assistants and function-specific agents (still present, but subordinate to the knowledge architecture).

The kind of company this fits: businesses whose competitive position depends on deep expertise. Professional services firms with strong knowledge management cultures. R&D-intensive companies. Pharmaceutical companies. Engineering firms. Regulated industries where audit trails and explainability are non-negotiable.

The trade-offs: a broad knowledge graph can be expensive and slow to build, while a use-case-led graph can start smaller. Either way, the architecture depends on identity, provenance, recency, and the quality of what enters it, which makes the institutional-knowledge case in Section 8.2 especially relevant. The architecture is most powerful where relationships and accumulated knowledge are genuinely strategic, and less differentiated where they are not.

Sketch Four. The Customer-Surface Operating Model.

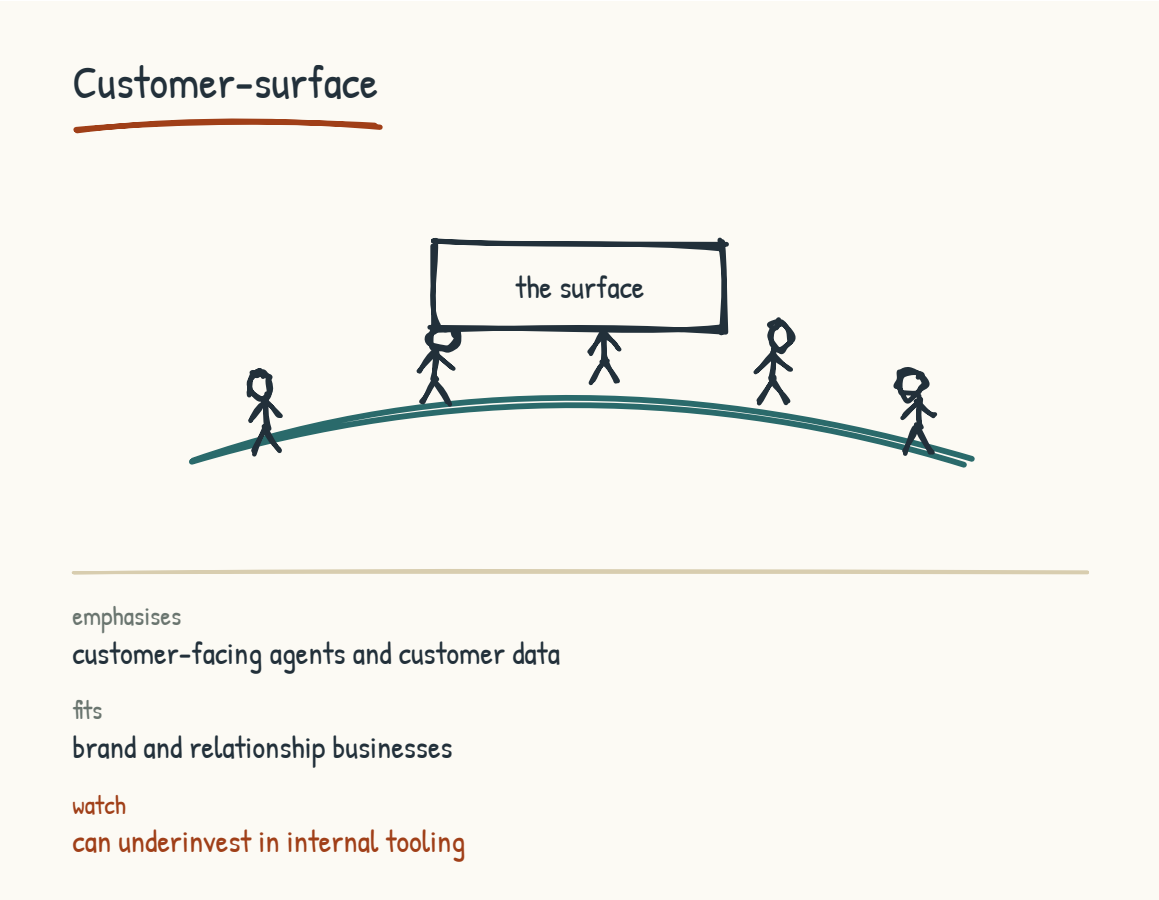

Figure 4.2f. Customer-Surface: the internal architecture organizes to feed the customer-facing AI.

This sketch organizes outside-in. The customer interaction layer is the design driver. AI is most heavily invested at the customer-facing surface, and the internal architecture (data, systems, agents) is organized to make the customer-facing AI work.

The architecture is customer-facing. The AI that interacts with customers is the most visible and carefully designed piece. Behind it, agent systems and the data layer feed the experience with customer history, preferences, current context, and product knowledge, but only to the extent permitted by consent, purpose, policy, and law.

The ingredients this sketch emphasizes: systems of agents oriented around customer-facing work. The data layer, especially customer data and product data. The standards that govern how customer interactions are recorded and routed.

The ingredients this sketch treats as secondary: internal coordination among employee assistants (less central because the design driver is customer-facing rather than internal).

The kind of company this fits: businesses whose primary value is in the customer relationship. Consumer brands. Retail. Hospitality. Direct-to-consumer companies. Some financial services. Companies where the customer experience is what drives revenue.

The trade-offs: requires sophisticated customer data infrastructure to do well. Vulnerable to brand and reputation damage if the customer-facing AI produces poor experiences. Underinvests in internal employee tooling, which can produce friction inside the company even when the customer-facing experience is strong.

Sketch Five. The Federated Operating Model.

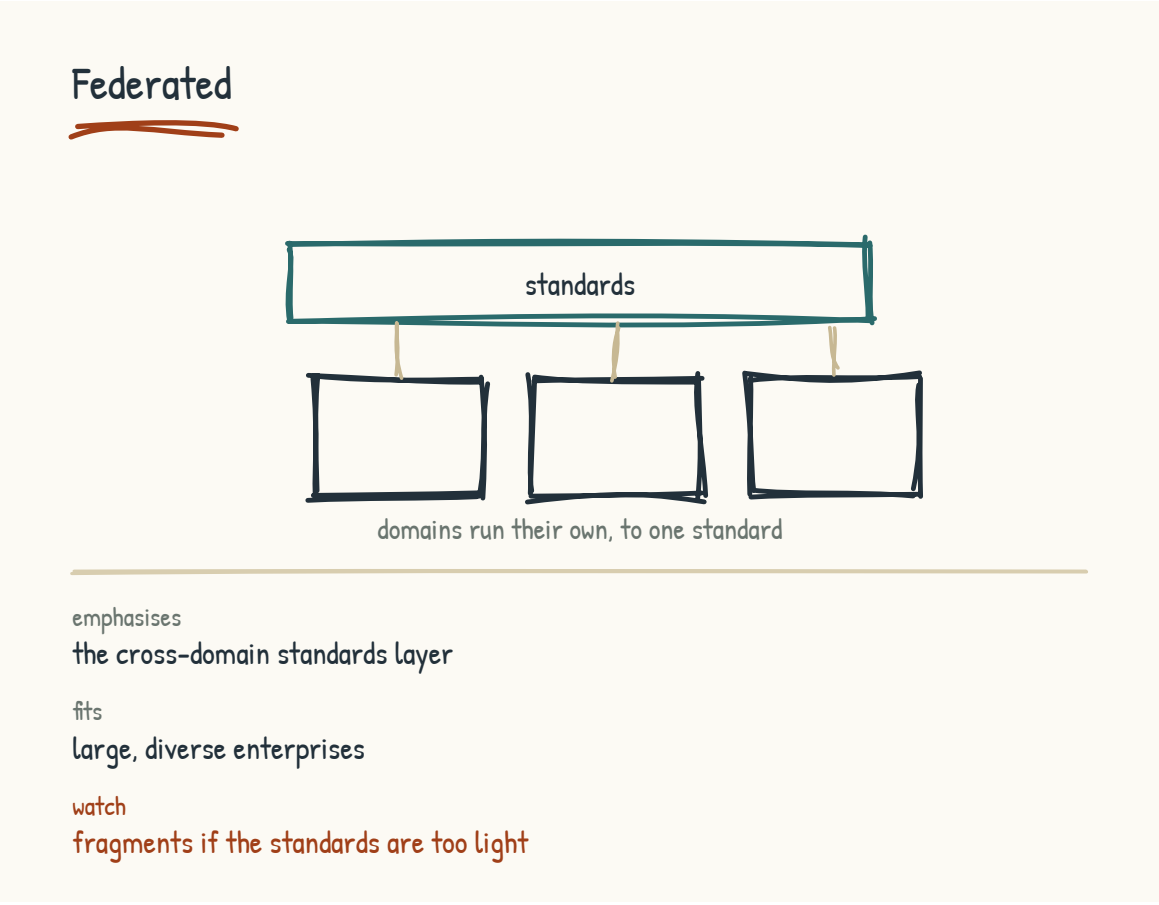

Figure 4.2g. Federated: each domain runs its own architecture under light cross-domain standards.

This sketch acknowledges that one architecture across a large diverse company often does not fit. Each major function or business unit designs its own AI architecture, suited to its specific work. Lightweight standards govern cross-domain coordination, but the design is not unified.

The architecture is distributed. The manufacturing organization might run a workflow-native sketch. The customer-service organization might run a customer-surface sketch. The R&D organization might run a knowledge-centric sketch. The corporate functions might run a coordination-centric sketch. The variations are tolerated because forcing one architecture across all of them would produce worse outcomes than the variation does.

The ingredients this sketch emphasizes: the standards layer, especially the cross-domain standards that allow the federation to function. The data layer, where shared definitions become essential.

The ingredients this sketch treats as variable: nearly everything else is configured differently by domain.

The kind of company this fits: large diverse enterprises. Multi-business-line conglomerates. Companies whose units operate in significantly different industries or with significantly different customer bases. Companies where attempts at unified architecture have failed in the past.

The trade-offs: requires strong central capability to manage the standards across the federation. Risks fragmentation if the standards are too light. Can become expensive because the same capability is being built multiple times. The companies that do this well treat the federation as a deliberate choice rather than as a failure to unify.

On Using the Sketches

The sketches are starting points. The leadership team doing AI-Native design work studies them, identifies what fits and what does not, and produces a design that is their own.

A few notes on how to use the sketches in actual design work.

First, the sketches can be combined. Most real designs will draw from more than one. A company might use the coordination-centric sketch for its corporate functions and the workflow-native sketch for its operations. Another company might run knowledge-centric at the top and customer-surface at the front. The combinations are constrained by coherence rather

than by the sketches themselves. The standards layer is what makes the combinations hold together.

Second, the sketches change over time. A company that begins with one pattern may evolve toward another as the technology matures and the company learns what works. Treat a single architecture as final and you get rigidity. Establish regular review points, plus triggers for material changes in models, regulation, risk, or business strategy, and adaptation becomes part of the design.

Third, the technique of analogical sketching is useful in workshop conversations about which patterns might fit. Pick a company that has built a recognizable operating model and ask how that company would approach the same problem. What would Amazon do if it were designing this? What would Apple do? What would a hospital system design? What would a small consulting firm design? The point of the exercise is to free the leadership team from its existing assumptions long enough to see options it would not otherwise see.

Fourth, the sketches need to be tested before anyone commits to them. The Alignment Sprint in Section 9.2 gives the team a setting for comparing candidate designs against the company's actual conditions. The comparison may produce a choice or identify the evidence needed before choosing. The design that emerges from the test is almost always different from the design that entered it.

The point of the sketches is to show that designing the AI-Native operating model is design work. There are choices to be made. The choices have consequences. Done deliberately, with the ingredients understood and several sketches considered, the design work produces an operating model that fits the company. Skipped, whether by copying a sketch wholesale or by letting the architecture emerge by accident, it produces one of the failure cases Section 8.2 will catalog.

No leadership team knows the right answer in advance. The work is to design with the ingredients you have, test the design against the conditions you face, and revise as you learn.

The Leader's Five-Part Mandate

The design conversation just described requires a leadership team that owns the work. By default, few do. They delegate parts of it to the CTO, parts to the consultant, parts to whichever function happens to be loudest. Distribution is not the problem; fragmentation without integration is. When nobody owns the trade-offs across those parts, the transformation becomes incoherent.

I want to be specific about what the CEO and senior leadership team remain accountable for in an AI-Native transition. There are five mandates. The work within each can and should be distributed to qualified owners; accountability and integration cannot be outsourced. Neglect one, and the transition becomes vulnerable in a recognizable way.

First, set the operating-model direction. Name the archetype the company intends to become. Name the time horizon. Name the architectural commitments that follow. The leader owns the decision, but the vision should be informed by the people who understand the work, technology, customers, and consequences. When this mandate is missing, the company can produce a wall of separated pilots with no shared direction: the pilot graveyard from Section 1.2.

Second, make the build, buy, and partner calls. Decide which layers of the stack are proprietary, which are bought, which are partnered. This is the moat decision, the cost decision, and the speed decision rolled into one. When this mandate is missing, vendor accidents become the strategy.

Third, resource the architectural work. Fund the foundations the chosen use cases require (the knowledge architecture, collaboration surface, integrations, evaluations, security, and integrity mechanisms), even when there is no announceable win in the current quarter. When this mandate is missing, the company can default to chat-everywhere: useful for individuals, but rarely sufficient as an enterprise strategy.

Fourth, lead the human reckoning. Acknowledge what the transition is asking of the workforce. Construct the covenant. Honor the institutional knowledge. The leader has to be honest before the workforce can be. When this mandate is missing, institutional knowledge walks out the door (Section 8.2 names this case).

Fifth, sustain the practices. Develop your own AI practice (Section 5.1 names this work). Protect time for the team's practices. Measure short-term outcomes and the capabilities that compound beyond a quarter. The workforce calibrates on what leaders do rather than what they say. When this mandate is missing, the transition can remain in slogans rather than operations.

The five mandates are standing categories of accountability for the CEO and senior leadership team throughout the transition. Holding all five does not guarantee success; neglecting one creates a predictable weak point. The purpose of the list is not to keep all work in the executive suite. It is to keep the parts connected under leadership that can resolve trade-offs across the company.

Figure 4.2h. The leader's five-part mandate, with the failure mode that follows when each part goes missing.

The next section, on the collaboration surface, returns to one specific architectural piece that nearly every design will need. The section after that addresses the disciplines that keep the design honest as the AI work scales. The architecture is the substance. The design conversation is the work.

Section 4.3

The Collaboration Surface

What happens to AI work when it stays inside private chats?

Whether AI compounds across an organization or fragments into isolated islands depends heavily on one architectural decision: the collaboration surface. Companies often make this decision by default when they deploy AI primarily through private chat sessions. Enterprise products may provide administrative controls, shared projects, or audit features, but the work itself can still remain invisible and unreusable to colleagues. That default produces a particular kind of organizational AI failure, and it is preventable.

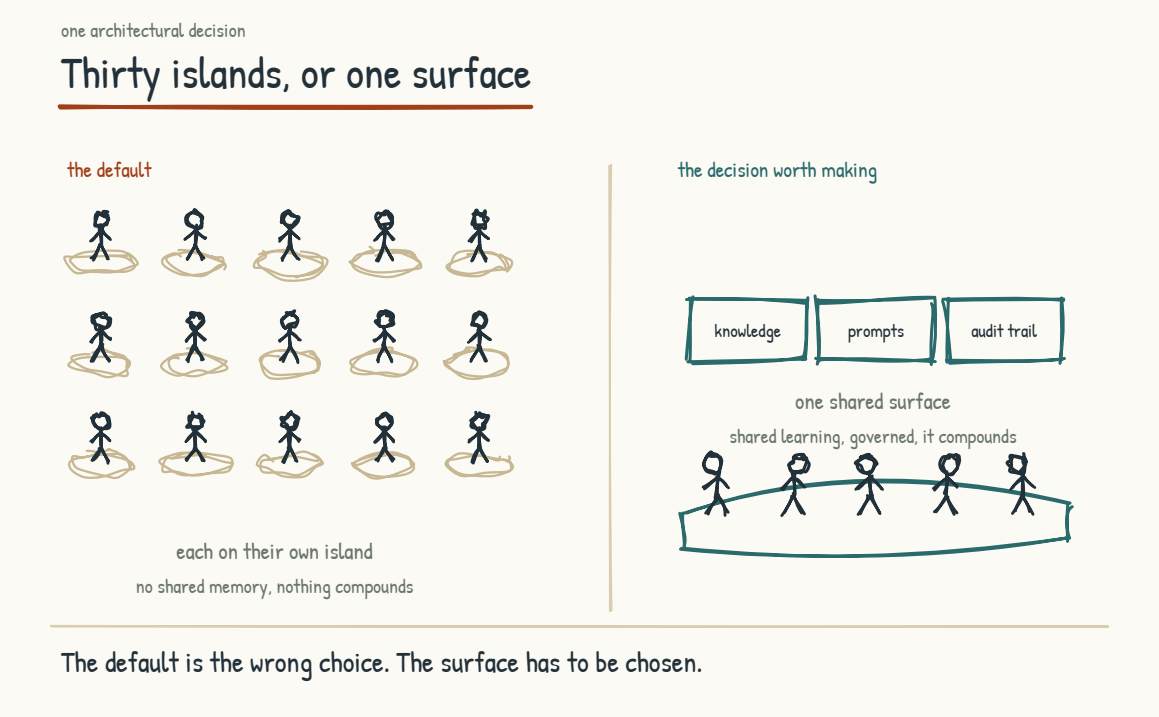

Figure 4.3a. The architectural choice between fragmented individual chat sessions and a shared collaboration surface. The default is fragmentation by design.

The Thirty Islands Problem

A company deploys AI tools to its workforce. Every employee gets access. Every employee starts using the tools. Every employee has their own chat history. Every employee writes their own prompts. Every employee develops their own workflows. Every employee produces their own AI-mediated work.

Thirty employees are doing this, or three hundred, or three thousand. That number does not change the shape of the problem, which is that the AI capability is individual rather than organizational.

The thirty employees produce thirty different versions of the same kind of analysis when thirty different customers ask similar questions, because nothing in the system causes them to converge on a shared approach. The thirty employees discover thirty different effective prompts for the same use case, none of which the other twenty-nine ever see. The thirty employees ask the AI thirty similar questions, and the AI answers each one fresh because nothing in the system accumulates the context of the previous twenty-nine answers. The thirty employees produce thirty drafts that the manager has no way to see in aggregate, because each draft is buried in a private chat. Each of the thirty would report, accurately, that the tool is great.

The company has invested in AI. The investment has produced thirty individuals using AI. The investment has not produced an AI-Native organization. The difference matters and compounds.

Multiplayer AI

The alternative architecture is sometimes called multiplayer AI, by analogy with the gaming distinction between single-player and multiplayer experiences. The shift is from AI solely as a personal tool toward AI as a shared working environment, accessed and shaped by teams while preserving legitimate private space.

A multiplayer AI surface has several components. The first is permissioned shared context. Relevant prior work, decisions, customer history, project state, and institutional knowledge can be available to authorized team members and systems. The second member of the team does not have to re-explain what the project has already established. The system can retrieve what that person is entitled to use, with provenance and current state, rather than treating all team information as one undifferentiated memory.

The second is shared prompt libraries. When someone in the team finds a prompt that works particularly well for a recurring kind of analysis, the prompt is saved to a library that the rest of the team can use. The prompts are versioned. The prompts can be improved over time. The team's collective skill at working with AI accumulates as a shared asset rather than dissipating into thirty private chat histories.

The third is shared workspaces for AI-mediated work. When a team member is working on an artifact intended for collaboration, the work is visible to the appropriate team. Colleagues can see what is being produced, comment on it, contribute their own AI-mediated input, and build on what is already there. The work product can be collaborative from the start rather than

assembled at the end from isolated pieces. Draft privacy, sensitive personnel matters, privilege, and need-to-know boundaries still require private spaces.

The fourth is shared review. When the AI generates output that will be used in some way that matters, the output goes to a shared review surface where the relevant team members can see it, evaluate it, flag concerns, suggest revisions. The review happens before the work is delivered, not after problems surface in production.

The fifth is shared monitoring. Authorized owners can see appropriate aggregate signals about what AI is being asked to do, what patterns are emerging, and what outputs or actions warrant attention. Monitoring should be proportionate: collecting every private prompt can create surveillance, labor, confidentiality, and data-retention risks. Governance requires visibility, but not indiscriminate visibility.

Why This Is Architectural, Not Cultural

The instinct of many leaders, on encountering the thirty-islands problem, is to address it culturally. Train people to share their prompts. Encourage cross-pollination. Set up a Slack channel where people can post what is working. These cultural moves are useful and they do not solve the problem.

This problem is structural. Single-user chat tools make sharing high-friction by default. To share a prompt, an employee has to copy it from one tool, paste it somewhere else, write context about what it does and when to use it, and post it where someone might find it. To use someone else's shared prompt, an employee has to find it in the place it was shared, copy it, paste it into their own tool, adapt it to their specific situation. Friction at every step is high, and that friction is what keeps sharing sporadic instead of systematic.

Multiplayer AI changes the architecture. The prompts are in the system. The context is in the system. The work products are in the system. Sharing becomes the default for work intended to be collaborative because the architecture makes it easy. Private work remains a designed mode where confidentiality, privilege, sensitivity, or simple thinking space requires it.

This is why the collaboration surface is an architectural decision as well as a cultural one. Architecture shapes the default behavior; culture determines whether people trust and use it. A company that wants AI capability to compound organizationally has to build a surface that makes useful sharing easier while protecting work that should remain private.

The Shadow AI Risk

Single-user chat architecture carries a particular risk, and it is becoming one of the dominant concerns in enterprise AI. Shadow AI is the unauthorized, unmonitored, ungoverned use of AI tools by employees who are working around the company's official AI strategy.

Employees turn to shadow AI for several reasons. Official tools may be insufficient for their actual work. The official tools may be slower or less capable than what the employee can access on their own. The official tools may be governed in ways that prevent the employee from doing the work they need to do quickly. Whatever the reason, the result is that the employee starts using consumer-grade AI tools with their own accounts, pasting company data into them, and producing work that the company cannot see.

The shadow AI problem is harder in an architecture centered on private sessions because legitimate work and unauthorized work can look similarly isolated. An enterprise tool may still be materially better governed than a consumer account through contractual data controls, identity management, retention settings, connectors, and administrative logs. But procurement controls alone do not make the work reusable or assure that employees will choose the approved path.

Multiplayer AI changes the incentive structure. The official surface can offer permissioned team context, prompt or workflow libraries, existing artifacts, and review by colleagues. The shadow tool leaves the employee to reconstruct that context alone. When the approved environment is both safer and more useful, the unofficial path becomes less attractive. That can reduce shadow use, although endpoint controls, education, policy, procurement, and enforcement still matter.

Companies rarely solve shadow AI through prohibition alone. They improve their odds by making approved tools genuinely capable, reducing needless friction, setting clear boundaries, and enforcing the boundaries that matter. Architectural design carries policy into daily work. Written acceptable-use rules remain necessary, but they are not substitutes for getting the product, incentives, and controls right.

The Connection to Governance

The collaboration surface determines what kinds of governance are practical. Layer 8 of the nine-layer AI stack from Section 2.3 (evaluation, monitoring, and governance) depends on appropriate observability. Private work can still be governed through enterprise identity, retention, data-loss prevention, audit events, sampled review, and outcome monitoring. But

when neither inputs, outputs, actions, nor downstream outcomes are observable, errors are harder to detect, drift harder to measure, and compliance harder to demonstrate. Governance cannot rest on work the organization has designed itself never to see.

A company serious about governance needs architecture that produces the visibility its risks require. There is no policy that compensates completely for missing controls, and no collaboration surface that compensates for bad policy. The choice between predominantly private and deliberately shared AI affects how learning compounds, how shadow AI is addressed, and which governance mechanisms are available.

What the Choice Costs

Honestly accounted for, a multiplayer collaboration surface costs more than deploying single-user chat tools. The shared context infrastructure has to be designed. The prompt library has to be built and curated. The collaborative workspaces have to be integrated with the company's existing tools. The review surfaces have to be designed to fit the team's actual workflow. The monitoring infrastructure has to be deployed.

The investment is real, and so is the potential return. A well-used collaboration surface lets practices, validated workflows, and review standards accumulate across the workforce. Work products can improve through visible iteration. Governance becomes more practical because the architecture supports appropriate observability. Shadow AI risk may decline because the approved surface offers advantages the unofficial alternative lacks.

Companies that do not build these capabilities risk the inverse: skill that remains local, work products that vary without a shared improvement mechanism, weaker governance, and more incentive for shadow AI. Some organizations can achieve parts of the same outcome through existing document, code, workflow, and collaboration systems; the requirement is the capability, not a product bearing the label.

The cost differential may produce little visible return in the first quarter because shared context, evaluation, and governance are capability investments. Their value should be measured over time through reuse, cycle time, quality, risk reduction, adoption, and avoided duplication. A leadership team judging only immediate seat-level productivity will miss much of the case.

The next section addresses the discipline that runs across all of this and determines whether the AI's output can be trusted as the practice scales.

Section 4.4

Epistemic Hygiene

How does an organization tell the difference between AI output that is true and AI output that sounds true?

The output of an AI system is text. The text is articulate. The text is plausible. The text often sounds confident even when the underlying analysis is wrong, missing, or fabricated. A workforce learning to work with AI has to develop a set of disciplines for evaluating the text, sourcing it, verifying it, and managing the ways the text can mislead. Without those disciplines, the AI's plausibility becomes a liability rather than a capability.

I use the term epistemic hygiene because the disciplines are about knowledge management at the individual and organizational level. They are about how the company knows what it knows, and how it tells the difference between knowledge that has been verified and knowledge that has been generated. The disciplines are practices that the workforce develops over time, supported by infrastructure that makes the practices possible.

AI introduces three distinct problems, and each requires its own discipline. Grounding in authoritative, current evidence helps with all three, but it is not their single cause or complete cure. Variance arises from probabilistic generation and changing context. Hallucination can arise even when sources are available. Drift can enter through models, prompts, tools, retrieval, data, users, and the world the system operates in. The three problems interact, but keeping their mechanisms distinct makes the disciplines more useful.

Figure 4.4a. Three distinct epistemic problems and the disciplines required to address them. Shared infrastructure makes the practices systematic and observable.

Problem One. Variance

Ask the same AI the same prompt on different occasions and you get different answers. Sometimes the differences are small, sometimes large. They are almost always invisible to the user, who asks once, gets an answer, and treats that answer as the answer. The user does not know that asking again would produce something different. The user does not know that a colleague asking the same question yesterday got a meaningfully different response.

The variance is structural. Generative models produce outputs from probability distributions, and results can change with sampling settings, model versions, hidden system context, tool results, and infrastructure. Some systems and tasks can be made highly repeatable; exact reproducibility across hosted model changes is harder. A workforce that treats every AI response like a deterministic database result adopts the wrong mental model.

Corroboration is one discipline that addresses variance. When the answer matters, rerunning or reframing the task can reveal instability. But agreement among several answers from the same model is not independent confirmation; the model can repeat the same error. Strong corroboration compares the claim against primary sources, authoritative data, deterministic

tools, alternative methods, or qualified reviewers. Repetition diagnoses sensitivity. Independent evidence supports truth.

The collaboration surface from Section 4.3 can support corroboration by making variants, evidence, tests, and reviewer judgments comparable where sharing is appropriate. The discipline is still possible in private work, a careful individual can rerun and verify, but shared infrastructure makes it teachable, reviewable, and measurable across the organization.

Problem Two. Drift

Whatever model the company is using changes over time. Vendors update it. The fine-tuning is adjusted. The training data is refreshed. The system prompts are revised. The retrieval indexes are rebuilt. Each of these changes can produce subtle shifts in how the AI responds, what it considers, what it surfaces, what it omits. The shifts are usually small at the level of any individual response. They accumulate at the level of organizational output over months and years.

A workforce that does not monitor drift discovers, eventually, that the AI is doing something different than it was doing a year ago. No one noticed because no one was looking. The work products that the AI produces have shifted in subtle ways. The default outputs have changed. The patterns the AI surfaces are different. The blind spots are different.

Monitoring is the discipline that addresses drift. Run representative evaluations on a defined cadence and compare task-level quality, safety, cost, latency, and failure patterns against an approved baseline. Investigate material deviations rather than treating every wording change as drift. Trace them, where possible, to changes in models, prompts, retrieval, tools, data, or the operating environment, then address the implications for the company's AI-mediated work. This is an ongoing operational discipline, analogous to the way mature engineering organizations monitor production systems for behavioral changes that might indicate problems.

The infrastructure that supports drift monitoring is Layer 8 of the nine-layer AI stack from Section 2.3, the evaluation and observability layer. The discipline can start with a small, curated test set and manual review; it becomes systematic only when the organization builds the versioning, evaluation, and observability needed to see change over time. The alternative is to operate blind.

Vendors vary in how much notice and detail they provide about model updates, and release notes cannot predict every behavioral change in a customer's workload. Many customers also

lack monitoring capable of detecting those changes. The work can therefore shift under the company's feet even when an update was announced. The protection is a versioning, evaluation, and change-management practice tied to the organization's own use cases.

Problem Three. Hallucination

The most well-known AI failure mode is the one in which the AI generates output that is plausible, articulate, and wrong. The AI asserts a fact that is not true. The AI cites a source that does not exist. The AI describes a connection that has no basis in reality. The output reads as confident expertise. The output is fabrication.

Hallucination is becoming better understood, and model and system designers have reduced it on many measured tasks through better training, retrieval, tools, and verification. Elimination is not in sight, and performance varies by domain and benchmark. A workforce that treats AI output as inherently reliable will still encounter fabricated or unsupported claims and may act on them in damaging ways. The discipline is to refuse the assumption that improving average model quality has solved the use-case-specific problem.

Verification is the discipline that addresses hallucination. Check asserted facts against source data. Open and inspect citations. Reproduce calculations. Test code. Confirm that conclusions follow from stated premises and evidence. A generated chain of reasoning can itself be a post-hoc or inaccurate explanation, so it is not a substitute for an external check. Verification takes real effort and adds time. It is also what prevents the company from acting on confident fabrication.

Verification depends on the whole system being designed to support it. The interface should provide citations or provenance in a form that can be checked, expose relevant uncertainty and limitations, and allow abstention or escalation. Models do not possess honesty or willingness in the human sense, and instructions alone cannot guarantee calibrated uncertainty. Reliable verification comes from configuration, evidence, tools, tests, workflow, and human practice together. Optimize only for fluent output and the system can produce more sophisticated errors.

> **WATCH FOR THIS**
>
> **The most dangerous AI output is not the obviously wrong kind. It is the well-written, confident, plausibly-sourced output that is wrong in a way no one in the room can immediately verify. The clarity of the prose disguises the gap.**

> **Leadership decisions made on the basis of polished AI summaries, without corroboration against actual data or expert judgment, are the failure mode that destroys the most institutional credibility per dollar spent. The hygiene practice is to require the corroboration before the decision.**

Why Single-User Architecture Weakens All Three Disciplines

At organizational scale, the three disciplines (corroboration, monitoring, and verification) depend on appropriate visibility. They require infrastructure that makes relevant AI outputs, actions, evidence, and outcomes observable, comparable, and reviewable over time. That does not require the company to inspect every private exchange. It requires visibility proportionate to the risk of the work and enough shared evidence to evaluate the system rather than only the individual user.

The single-user chat architecture makes all three disciplines harder to practice organizationally. Variance stays local when each user sees only one response. Drift stays hidden when no one compares current behavior with a baseline. Hallucination may escape review when no one besides the user sees the output before it is used, especially when that user lacks the domain knowledge or source access needed to challenge it.

The collaboration surface from Section 4.3 is an architectural foundation for making epistemic hygiene shared practice. On it, the disciplines can be taught, instrumented, reviewed, and improved. Careful individuals can and do verify work in single-user tools, and enterprise controls can monitor some private use. What private chat does not provide by itself is an organizational learning loop.

Architectural decisions and disciplinary practice reinforce each other. A collaboration surface makes epistemic hygiene easier to institutionalize; it does not create rigor automatically. Without one, a company needs other deliberate mechanisms for evaluation, review, versioning, and shared learning.

The Three Integrity Mechanisms

The three problems and disciplines need infrastructure at the right points in the system's lifecycle. I organize that infrastructure into three practical categories: guardrails, safeguards, and feedback loops. My notes from the HBS Online AI for Leaders course, taught by Karim Lakhani and Iavor Bojinov, helped shape the framing; the triad and the definitions used here are this book's synthesis, not terminology I am attributing to the course. The labels overlap in

ordinary industry usage; here they name different functions rather than universally accepted technical definitions.

Guardrails set boundaries in development and at runtime. They include instructions, permissions, schemas, allowlists, policy checks, tool restrictions, and deterministic constraints that refuse, redirect, or escalate requests outside permitted scope. Guardrails are how parts of company policy become enforceable behavior rather than a document sitting next to the system. They must be tested before deployment and continuously re-evaluated as the system changes.

Safeguards monitor deployed systems in operation. Outputs and actions are evaluated against use-case-specific thresholds. Alerts trigger when behavior deviates from expected ranges. Human review is invoked when inputs are unusual, evidence is insufficient, monitoring detects risk, or an action would cross a defined consequence threshold. Model-reported confidence alone is not a dependable trigger. Safeguards are circuit breakers intended to catch and contain problems before they propagate; no safeguard catches everything.

Feedback loops track performance over time so people can improve the system. User corrections, errors detected by safeguards, and drift detected by monitoring may feed into retrieval content, prompts, rules, evaluations, workflows, or, when the organization actually controls a training process, model updates. Production models do not automatically learn safely from every interaction. Feedback must be reviewed, permissioned, and protected against poisoning and privacy leakage.

The three mechanisms work together. Guardrails set or enforce boundaries. Safeguards detect and contain production failures. Feedback loops inform controlled improvement. A company that relies on only one has a fragile system. Even all three do not make AI intrinsically reliable; they create layered evidence and controls that support justified reliance for a defined use case.

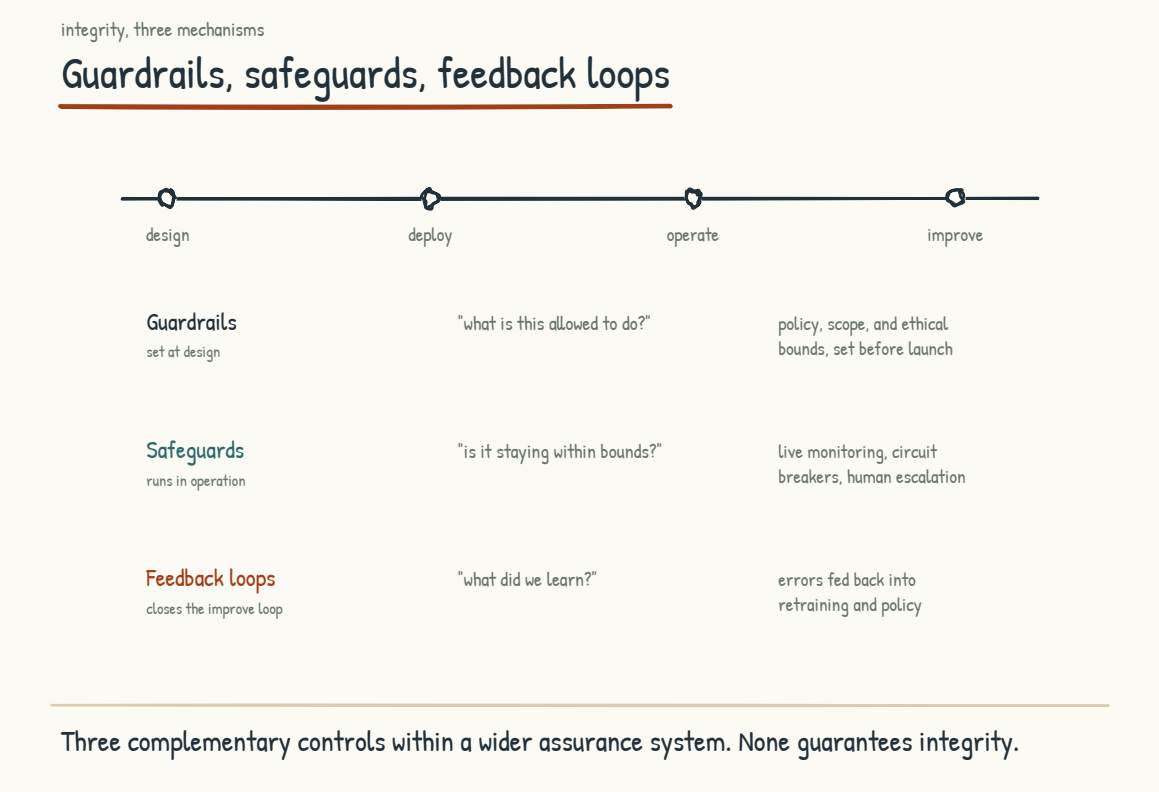

Figure 4.4b. The three integrity mechanisms across the AI system lifecycle. Guardrails, safeguards, feedback loops.

Why This Discipline Will Become Strategic

Practiced as an organizational discipline, epistemic hygiene produces evidence for deciding when AI-mediated work deserves trust, institutional learning that accumulates, and a capability competitors cannot copy merely by buying the same technology. Left solely to individual diligence, quality varies widely, unsupported claims occasionally surface as embarrassments, and behavioral changes go untracked. The fragility tends to appear when the work matters most.

Stage 4 has now built most of the technical and organizational foundations. The knowledge architecture for grounding and relationships. The AI-Native operating model for combining ingredients into a coherent design. The collaboration surface for organizational compounding. Epistemic hygiene for justified trust over time.

One architectural conversation remains, and it is the one the board will ask about first: what all of this costs, and why the layers that produce no demo are the investment rather than the overhead. The next section speaks the CFO's language.

Section 4.5

The Economics of the Transition

How do you fund a transition whose most valuable investments produce nothing to demo?

Stage 4 has made four architectural arguments, and each ends with a version of the same funding problem. Knowledge architecture costs real money now and may compound over years. Operating-model work requires foundations even when there is no announceable win in the current quarter. A collaboration surface can cost more than deploying chat tools, while its early return is harder to isolate. Epistemic hygiene can read as overhead on a conventional budget line. Invest now. Measure leading evidence. Compound later if the system earns it. Be patient, but not credulous.

Somewhere in the room where those arguments land, there is a CFO who has heard "be patient" before. This section is for her, for the CEO who has to win her over, and for the leadership team that will eventually have to defend the spending to a board. The compounding argument has been made. What has been missing is the money argument underneath it, and a transition that cannot survive a budget review is a mood with a budget line.

The CFO Is Doing Her Job

Consider what the CFO actually sees. AI spending has grown quickly while returns remain difficult to attribute. The 2025 GenAI Divide report from MIT's NANDA initiative reported limited measurable financial returns across much of its sample. The preliminary report combined interviews, a senior-leader survey, and public implementation accounts; its findings have limited reach beyond that sample. RAND's 2024 report, discussed in Section 1.2, cited an external estimate that more than 80 percent of AI projects fail, about twice the rate claimed for non-AI IT projects, then used 65 expert interviews to study causes rather than independently measuring that failure rate. The headlines are imperfect. The CFO's skepticism is still rational. She may not have read the studies, but she has probably lived a local version of stalled pilots and returns nobody can isolate.

Now the board is asking about payback periods, and look at what is arriving in her inbox. A knowledge graph she cannot demo. A collaboration surface that costs more than the chat licenses that already "work." A request to pull the company's most expensive experts off the line

for knowledge-capture interviews. Governance and monitoring infrastructure that, on paper, slows everything. Training time that shows up as output going backward.

Put yourself in her chair honestly. The spending that came with demos has failed to return, and the new proposal is to spend more on things that come without demos. A CFO who waves that through is not doing her job. The resistance is the financial discipline the company hired her to have, and the worst possible response to it is to ask her to have faith. She should not run on faith. The case for the invisible layers can be made in her language, and the rest of this section makes it.

Where the Money Actually Goes

The money in an AI-Native transition goes to three places, and only one of them appears cleanly in a budget.

The visible spend is the part procurement negotiates and the board sees. Licenses and seats. Compute and tokens. Vendor contracts, systems of agents, integration work. One genuinely new problem hides inside this category: usage-based pricing means the cost grows with adoption, so the bill rises in exactly the quarter things start working. A budget built on fixed-license assumptions will be surprised by its own success, and the surprise will arrive looking like a problem. Tell the CFO this early. She will appreciate being the first to know rather than the last.

The invisible spend is the case this book has been making. The paid hours of experienced operators engaged as teachers in knowledge capture (Section 4.1). The workflow discovery and redesign required to place agents safely in the work (Section 4.2). The building and curation of shared context, reusable workflows, and review surfaces (Section 4.3). The design of governance that enables rather than merely restricts. The training time, which can mean thousands of workforce hours and a temporary output dip while capability builds. Much of this never appears in an AI budget. It lives inside payroll, where it can look like people being taken away from their jobs. That visibility problem is one reason foundational work goes unfunded.

The third category is the spend nobody budgets. Leadership attention is the largest item in it: the five mandates from Section 4.2 are hours, recurring hours, from the most expensive people in the company. Next to it sits the productivity dip. A redesigned workflow underperforms the workflow it replaced for a while, because the team is learning, the handoffs are new, and the exceptions have not surfaced yet. Stage 5 will add the emotional labor of the human reckoning to

this category. None of it has a line item. All of it is real cost, and a plan that has not priced it will read the dip as failure and retreat at precisely the wrong moment.

On proportions, I will not pretend to a formula. The honest numbers vary by archetype and starting point. But the pattern I keep seeing is consistent enough to name: leadership teams budget as if the visible spend were most of the total cost, while successful transitions often require comparable or greater investment in workflow, knowledge, integration, governance, learning, and leadership attention. That is an observation from practice, not a benchmark. Companies may spend too little or too much overall; the recurring error is concentrating what they do spend in the category with invoices and undercounting the organizational work required to make it useful.

Rent and Equity

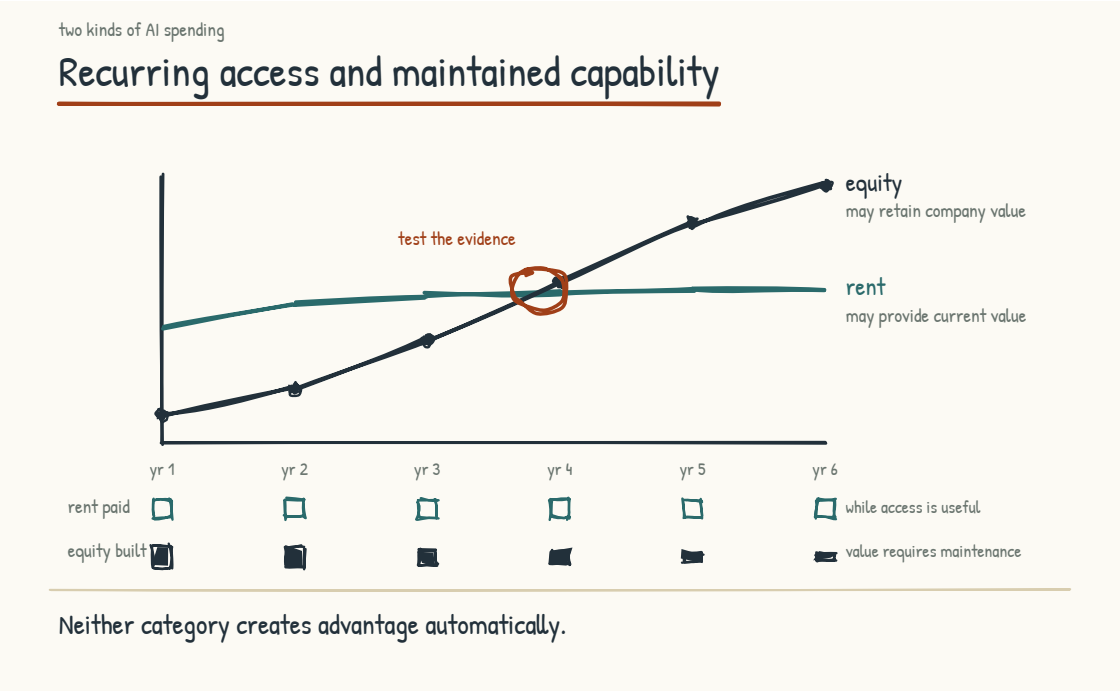

Figure 4.5a. Rent buys access to capability. Equity builds assets and practices the company can continue to develop. Durable advantage depends on how well either is integrated and used.

Here is the financial distinction that organizes everything else in this section. Some AI spending pays rent. Some builds equity. Both are legitimate. Confusing the two is where transitions lose their money.

Rent is what many visible-layer purchases buy. Licenses, seats, and vendor agent systems provide access while the contract continues, often to capabilities competitors can also purchase. Stop paying and access may stop, but the workflows, skills, configurations, and data developed around the product do not necessarily vanish. The gains may plateau as competitors adopt similar tools, yet execution can still differ on an identical vendor stack. Rent is fine. Companies rent critical infrastructure indefinitely. Rented capability should simply be priced, measured, and defended without pretending the license itself is a moat.

The invisible layers can behave differently, and the difference is the heart of the argument. A maintained knowledge graph may become more useful as validated entities, relationships, and use cases accumulate; a neglected one depreciates as the operation changes. A collaboration surface can convert local skill into shared workflows and learning (Section 4.3). Epistemic hygiene protects the integrity of both, as controls protect the integrity of a ledger. Competitors cannot buy your operating history off the shelf, though they can build analogous capability and sometimes catch up. In this chapter's managerial vocabulary, rent buys access; equity builds company-specific assets and practices that can support advantage.

The cost of not building the equity layer rarely appears as a line item, which is why it goes unmanaged. It appears as perpetual rent: the pilot budget that renews every year because little accumulated from last year's pilots. It appears as the graveyard from Section 1.2, re-funded each planning cycle under new names, and as the multi-million-dollar sunk costs when initiatives get abandoned. Some of the pilots are old enough to have anniversaries. It appears as the write-off nobody records when institutional knowledge walks out the door (Section 8.2, Case Five). In companies that stay stuck at Stage 2 of this model, the AI budget is rarely small. Too much of it pays again for access and activity that never became durable organizational capability.

INTELLECTUAL LINEAGE

The financial shape this section describes has a close analogue in the economics literature. Brynjolfsson, Rock, and Syverson model how adoption of a general-purpose technology requires complementary intangible investment: process redesign, human capital, and new organizational forms. Because conventional productivity measurement may not capture the creation of those complementary assets when costs are incurred, measured productivity can be understated early and rise later as benefits are harvested. They call that pattern the productivity J-curve and connect it to earlier general-purpose technologies. Their intangible capital is

> *close to what this book calls the invisible layers, although the managerial categories here are not accounting classifications.*
>
> Brynjolfsson, Erik; Rock, Daniel; Syverson, Chad (2021). "The Productivity J-Curve: How Intangibles Complement General Purpose Technologies." American Economic Journal: Macroeconomics, 13(1), 333–372.

How to Fund It

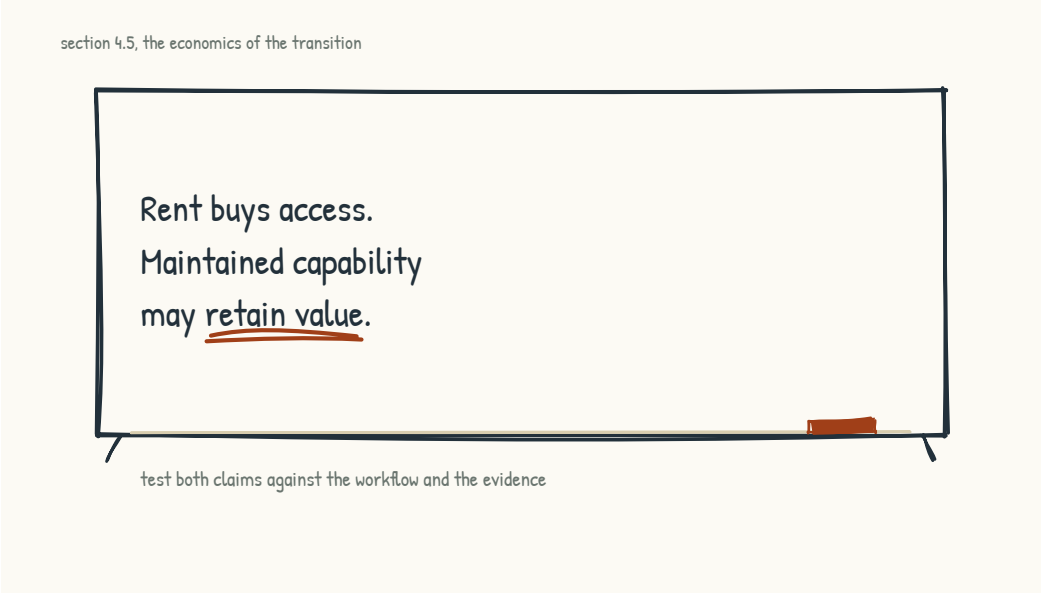

Figure 4.5b. The distinction that organizes the funding conversation.

Four funding patterns have held up across the organizational transitions I have run and the AI work examined here. None of them requires believing anything on faith. Each is a discipline the CFO can enforce.

First, fund workflows rather than tools. The unit of funding is a named workflow being redesigned end to end, with an owner, a baseline measurement, and a definition of done at the operating-model level. Tool purchases attach to a funded workflow or they wait. This one rule screens out much of the pilot-graveyard pattern, because a pilot with no workflow attached has nowhere to apply for money.

Second, stage-gate by operating-model milestone rather than activity alone. Release the next tranche when the redesigned workflow is in production with people actually working differently, when required domain knowledge is available through the chosen architecture, and when

governance operates inside the flow rather than sitting in a document beside it. Pilot and demo counts measure activity. Production use, quality, control effectiveness, adoption, and business outcomes measure whether a transition is taking hold.

Third, the pairing rule. Every visible purchase gets approved together with its invisible counterpart, in the same decision, on the same page. Buy the customer service system of agents, and in the same approval fund the workflow redesign, the knowledge capture from the experienced service leads, and the review surface where the agents' work lands. If the pair costs too much, the purchase costs too much. Approving the visible half alone risks buying access while underfunding the organizational capability that makes the access useful. It is one of the most common allocation errors I see.

Fourth, run the portfolio honestly. Keep the rent list and the equity list distinct, expect different things of each, and do not let a commodity license borrow a moat argument without evidence. A compounding claim belongs to spending that demonstrably accumulates something: validated knowledge coverage, captured expertise, reusable workflows, shared skill, data, or operating feedback. Rented tools may enable that accumulation, but access and asset are not the same thing. Commodity productivity tools are worth buying when their measured benefits justify their continuing cost.

BRING THIS INTO THE ROOM

Take two columns into your next budget conversation. In the left column, list everything the company spent on AI in the last twelve months: licenses, vendor contracts, pilots, dedicated headcount, consulting. In the right column, write what each item bought at the operating-model level: a workflow now in production, knowledge captured into a shared structure, a reusable capability, or a measured outcome. A blank cell does not prove waste, but it exposes an item whose value has not been articulated. The pattern between the columns reveals more about the company's actual AI strategy than the strategy deck does.

What to Tell the Board

The honest financial shape of this transition is uncomfortable to say out loud, which is why it often goes unsaid until the numbers force the conversation. Costs land first. Some visible tools produce quick, attributable gains; others do not. Foundational layers can consume money and attention for quarters or years before their contribution is visible in financial results. Then, if

adoption, workflow change, architecture, and market conditions align, returns may accelerate. The MIT CISR research behind Section 2.4 found its strongest financial-performance difference between the second and third maturity stages. It was an association in the study's sample, not a guaranteed timetable. Pay first. Test continuously. Earn later only if the operating change works.

A board that has not been told this possible shape in advance may read an expected investment period as failure or mistake activity for progress. Set the thesis at the beginning as a staged, multi-year commitment with named milestones, stop conditions, and an honest range for cost, timing, and uncertainty. Boards accept long-payback commitments such as plants, acquisitions, and market entries when the case and controls are credible. AI often arrives dressed as software even when much of the work is organizational transformation. Present the invisible layers as capability construction rather than merely licenses, but leave capitalization and expense treatment to the applicable accounting standards and the company's finance professionals.

While financial results are still emerging, report leading indicators of operating-model progress alongside outcome measures. Workflows running in production rather than pilots merely launched. Pilot-to-production conversion. Knowledge coverage and validation. Experienced operators engaged as teachers. Adoption of sanctioned tools, measured in privacy-respecting ways. Cycle time, quality, rework, incidents, customer outcomes, and unit economics against pre-redesign baselines. The portfolio's balance of rented access and accumulating internal capability. Movement in these indicators supports the investment thesis; it does not prove that financial returns must follow. Their value is that they reveal progress or failure earlier than a retrospective ROI slide.

And arm the CFO rather than managing her. Make her a co-author of the milestone framework, the gates, the two lists. A CFO defending a funding structure she helped design is a different force in a boardroom than a CFO reading someone else's conviction off a slide. In the transformations I have watched hold together through budget pressure, whatever the subject, the CFO was the one holding them.

Before the next budget cycle, run the two-column exercise, split the AI budget into its rent list and its equity list, and apply the pairing rule to the next purchase request that crosses your desk. Those three moves will tell you more about your actual position than any maturity score.

That completes the architectures. The knowledge architecture exposes relevant company structure to the systems that reason over it. The operating-model design combines the

ingredients into a coherent whole. The collaboration surface enables capability to compound across the organization. Epistemic hygiene provides evidence for justified trust at scale. The economics fund the work long enough to learn whether compounding is occurring. What architecture cannot supply is the thing it runs on, and Stage 5 turns to it now: people, starting with the leader's own practice, because the workforce watches what leaders do before it believes what they say.

CONSIDER BEFORE YOU ACT

If you keep three things from Stage 4: 1) Documents carry narrative and evidence; graphs make relationships explicit; the model and surrounding software reason with both. 2) The collaboration surface can turn isolated use into shared, compounding capability. 3) Rent buys access; company-specific knowledge, workflows, and practice can build equity; advantage depends on how well either is used.

STAGE 5

The People

What they bring, what is at risk, and what any design has to protect.

Section 5.1

The Leader's Own Practice

What does the AI-Native transition look like at the leader's own desk?

Stage 4 laid out the architectural work an AI-Native company has to do, and then the money argument that pays for it. The knowledge graph. The collaboration surface. The disciplines of epistemic hygiene. Rent and equity. The structural decisions are real and they matter.

Whether the transition succeeds runs through a smaller number of people than most leadership teams want to acknowledge. It runs through the individual leaders themselves, and through what they do with AI at their own desks. How they engage with it. Whether they have done the work they are asking the workforce to do. The workforce calibrates on this. The calibration is not loud and it is not stated. It shapes everything the workforce does in response to the transition the leader is announcing.

That work is the bridge into Stage 5, which addresses the human dimension of the AI-Native company at every level. Before the leadership team can credibly ask the workforce to undertake the reckoning in the next section, individual leaders need to begin their own. Order matters here. A leader who skips their own reckoning and moves straight to the workforce's risks breeding cynicism. A leader who starts with their own practice has a stronger basis for the credibility the rest of the work requires.

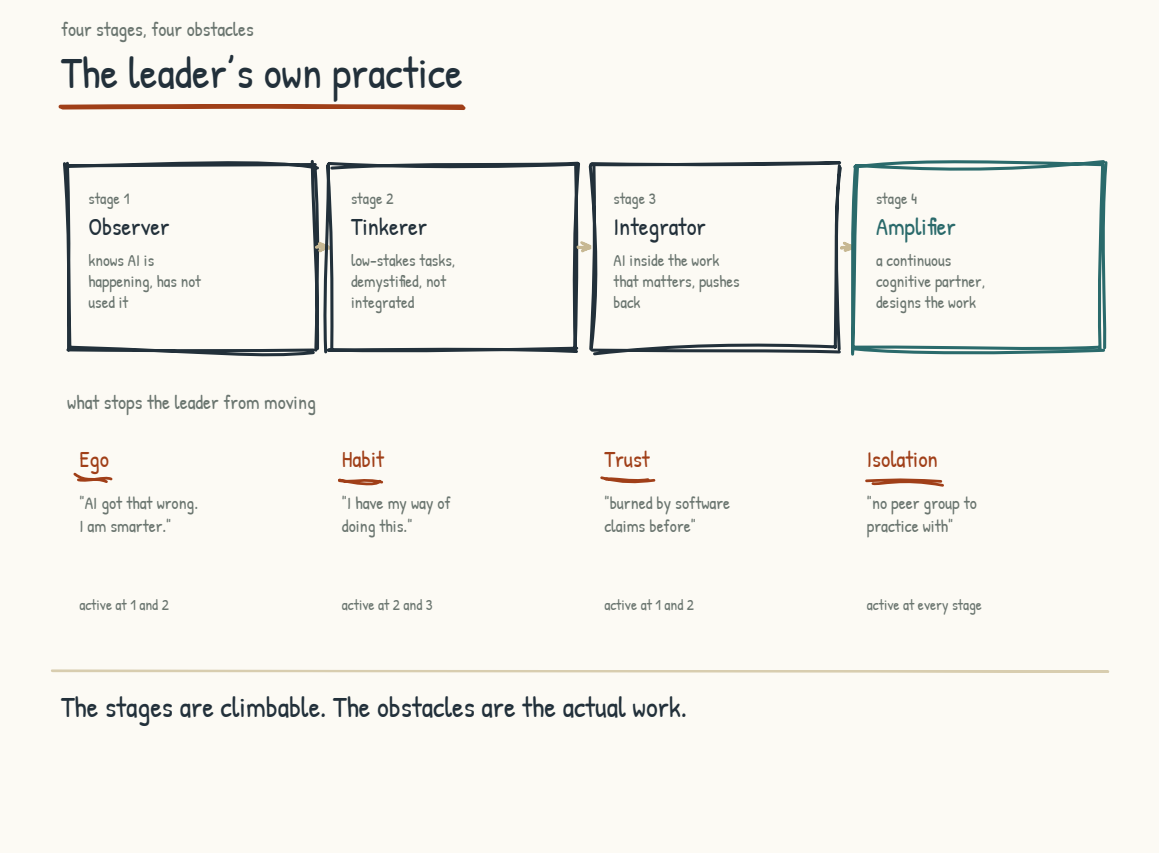

Figure 5.1a. Four developmental stages of leader AI adoption, and the four obstacles that hold leaders at each transition.

What the Leader's Own Practice Is

Your own practice is the daily use of AI in your own work. It is separate from the strategic decisions you make about the company's AI investment, and from the communication the leader has with the workforce about the transition. The actual practice. What the leader does on Tuesday morning at their own desk.

A leader with a working practice does several specific things. They use AI directly in at least some of their own substantive work rather than experiencing it only through an assistant. They engage with output as a provisional contribution, pushing back, asking for evidence and alternatives, and integrating it with their own judgment rather than treating it as authoritative. They notice when AI is substituting for effort they need to retain and pull back. They have patterns of work in which AI is invited and patterns in which it is not. They are conscious about the calibration.

The distinction becomes clearest on consequential work. Imagine a leader preparing a recommendation that will move capital, close a site, change a customer promise, or alter someone's role. A superficial practice asks AI for the answer and edits the tone. A working practice uses it to expose assumptions, build countercases, locate missing evidence, and test how the recommendation changes under different conditions. The leader still speaks with the people who hold context the system lacks. They still own the decision. Direct use matters because it lets the leader feel where speed helps, where fluency seduces, and where the work becomes more demanding rather than less.

A leader without a working practice does not do these things. They delegate the AI work to their team and receive summaries. They have other people prompt for them. They use AI as a search engine or a fancy autocomplete and call it AI use. They have not developed the cognitive posture Section 5.4 describes as the difference between using AI for thinking and using AI as thinking. They are operating at the surface of the technology while asking their workforce to develop depth with it.

That distinction sounds technical and it is operational. A leader whose own practice is shallow is less equipped to evaluate whether the company's AI investments are producing value. They may struggle to distinguish good AI-mediated work from polished AI slop or to tell when a team is using AI rigorously versus hiding behind output that looks substantive and is not. The shallow practice can also be visible to the workforce in ways the leader does not realize.

Founder Conviction Cannot Be Outsourced

The most direct statement of why a leader's own practice matters is this: conviction cannot be outsourced. A founder or executive who has not engaged seriously with the tools, and whose AI strategy arrives entirely from a consultant or technical function, risks building a program the workforce can feel is hollow at the center. Advisors and technical leaders are essential; borrowed conviction is not. Personal use does not make the leader technically omniscient. It gives the strategy contact with experience.

Barry O'Reilly makes the enterprise version of this argument in a 2026 Kyndryl Institute practitioner essay based on his executive coaching and venture-building experience: organizations that thrive with AI tend to be led by leaders who adopt it first. This is not a survey finding or a universal ceiling on workforce fluency. Skilled employees can outpace their leaders, and often do. The leadership effect is more practical: leaders shape permission, incentives, resources, risk tolerance, and whether changed work is recognized or punished.

The conclusion is still demanding. A leader who tries to direct an AI-Native transition without developing any personal practice places a boundary on what they can recognize, question, and model. That boundary may be invisible to the leader and visible to the workforce. It need not doom the transition, but it can slow it and make shallow adoption look deeper than it is.

> **INTELLECTUAL LINEAGE**
>
> *A useful way to frame this is leadership as technical input. The leader's specifications, judgments, and corrections shape what the system is asked to optimize and what the organization accepts as good. If leaders cannot articulate what they want in terms people and systems can act on, AI cannot rescue the ambiguity. Delegating implementation is necessary. Delegating the underlying intent and judgment is delegating strategy. AI work is not merely engineering work handed to engineers; it is leadership work conducted with engineering, domain, risk, and workforce expertise.*
>
> Author's synthesis.

The Four Stages of Leader AI Adoption

O'Reilly's Kyndryl Institute essay offers a four-stage practitioner framework, explicitly described as a recurring pattern rather than a maturity model: Awareness, Augmentation, Acceleration, and Adaptation. The plain-language descriptions below translate those stages as observer, tinkerer, integrator, and amplifier. This is a useful reflection tool, not a validated population map of executive development.

Stage one is awareness: the observer. The leader knows AI is happening. They have read about it. They may have approved budgets for it. They have not used it themselves in a meaningful way. Their engagement is mediated through others, so their opinions are not yet anchored in direct experience.

Stage two is augmentation: the tinkerer. The leader has begun using AI personally, usually for low-stakes tasks: summarizing material, drafting routine communications, or retrieving and synthesizing information. The use is real but remains at the periphery of work that carries greater judgment or consequence. The leader has begun demystifying the technology without yet redesigning how they work.

Stage three is acceleration: the integrator. The leader uses AI inside substantive work: refining documents, exploring strategic questions iteratively, and pressure-testing assumptions. AI is

becoming part of the workflow rather than a separate activity. The leader develops more evidence about when output is useful, shallow, or wrong. In suitable tasks, the resulting work may improve in speed, breadth, or quality, but the improvement still needs to be judged rather than assumed.

Stage four is adaptation: the amplifier. The leader has built patterns of AI use specific to their work and is redesigning some workflows around what human-AI collaboration can do. They also help design how AI appears in the company, informed by personal experience and by expertise beyond their own. The intended change is not merely faster output but a different capacity to frame, test, decide, and learn.

The stages are developmental prompts, not a scorecard or rigid sequence. No reliable evidence establishes that most leaders move through them in twelve to twenty-four months, and some may develop different practices in parallel. The warning at stage two remains useful: occasional low-stakes use can create confidence without the deeper practice required for consequential work. The remedy is not to rush upward on a ladder. It is to test one's actual capability against the work that matters.

Movement between stages should show up in behavior, not in how often a leader opens a tool. The observer can describe a use from direct experience. The tinkerer begins to distinguish novelty from repeatable value. The integrator can show where evidence changed a decision and where a control prevented a mistake. The amplifier can explain which parts of a workflow were redesigned, what people now do differently, and what the organization learned after the first version failed. Those are stronger signals than confidence, enthusiasm, or prompt volume. They also make regression visible: a leader can use AI every day and still retreat to shallow practice when time pressure rises.

The Three Behavioral Shifts

Three behavioral shifts translate that framework into practice. They are my synthesis of the leadership argument rather than findings measured by the Kyndryl essay.

The first shift is from delegating AI work to doing AI work. The leader stops asking others to use AI on their behalf and starts using it themselves. The shift sounds trivial. It is, in my experience, the hardest of the three for most senior leaders, because it requires the leader to be visibly inexperienced in front of a workforce that has been deferring to their expertise for years. The

willingness to be visibly a beginner, at fifty-five or sixty, in front of people who used to come to you for answers, is rare. It is also the condition on which everything else depends.

The second shift is from treating AI output as answers to treating it as material for evaluation. Depending on the task, that material may be a draft, analysis, option set, calculation, code change, or action proposal. The leader who pushes back, asks for sources and alternatives, tests claims, and replaces weak output with their own thinking models a stronger cognitive posture. The leader who accepts and forwards polished output without engagement models the posture that produces hollow work.

The third shift is from private AI use to visible AI use. The leader stops hiding their AI use and starts making it part of how they communicate with their team. They acknowledge when AI helped produce a document. They share the prompts that worked. They discuss what AI got wrong. The visibility produces two effects. The first is that the workforce sees the leader engaging with AI as a working tool, which gives the workforce permission to do the same. The second is that the workforce sees the leader's own development, including the parts that are not yet polished, which builds the trust the next section's covenant needs.

The Obstacles

Access is broader than it was, but it is not universal: price, disability, language, policy, security, and infrastructure still matter. For leaders with approved access, O'Reilly names four personal obstacles worth examining: ego, habit, trust, and isolation.

Ego. A leader who has been the expert for years may have built an identity around knowing more than the people around them. AI can retrieve or synthesize some subjects faster and more broadly while remaining wrong in important ways. That encounter can be uncomfortable. One ego response is to collect failures and conclude that the system has no useful role. Another is to overidentify with the technology and dismiss human expertise. Both protect identity at the expense of learning.

Habit. A leader's working habits were built over decades and often served them well. Some now deserve testing: drafting from scratch when a model might provide useful raw material, searching manually when AI-assisted research might accelerate discovery, or holding a meeting before generating concrete options. AI will not improve every one of these tasks, and speed is not the only value. The work is to examine the habit rather than preserve or discard it automatically.

Trust. The leader may have been burned by software that promised more than it delivered and watched companies waste money on immature technology. Skepticism is healthy. It becomes an obstacle when it hardens into refusal to test any suitable use personally. Trust should be calibrated through bounded use, independent evidence, evaluation, and knowledge of consequences. Promotional claims and repeated use are weak evidence.

Isolation. The leader is often the only person at their level in the company. They do not have peers with whom to compare practice. They cannot easily ask a question that exposes their lack of expertise without it being noticed and reported. The isolation produces a particular kind of stuckness, in which the leader knows they need to develop a practice but does not have a safe space where it can develop. The companies whose senior leaders have found a peer group of other senior leaders working through the same questions have, in my experience, produced the most consistent development. The isolation is real and it has a solution. The solution is structural: build the peer group, deliberately.

The four obstacles rarely arrive one at a time. Habit can dress itself as skepticism. Isolation can make ego more expensive because every beginner's question feels public. A breach of trust with one vendor can become a reason not to test a different use under different controls. That is why naming the obstacle is only the start. The leader has to design a response around it: a private practice environment, a peer willing to compare failures, a bounded task with a clear verification method, or protected time in a calendar that otherwise converts every experiment into after-hours work. Personal development still has operating conditions.

The Practical Operating Loop

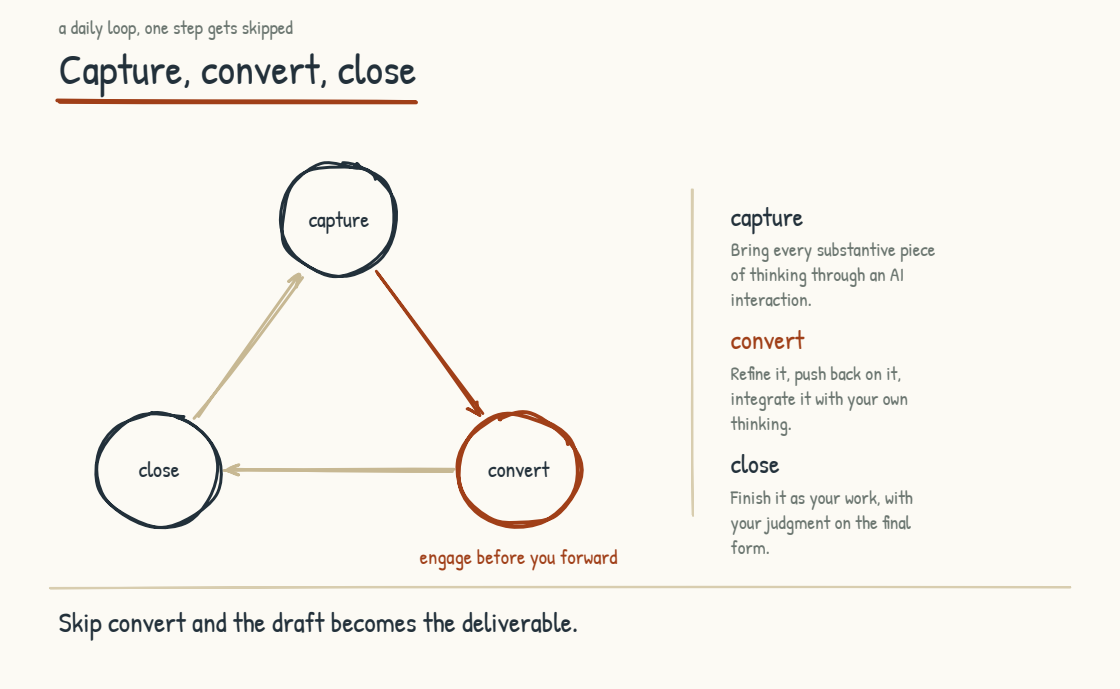

Figure 5.1b. Capture, convert, close. Capturing creates material; conversion and closure require the leader's judgment.

The development of practice is not abstract. It happens through repeated cycles of specific behaviors. Two patterns are worth naming because they have been refined by leaders who have done this work.

Capture, Convert, Close. O'Reilly names this as a practical loop in his Kyndryl Institute essay. Capture means preserving the useful meetings, ideas, reflections, and decisions that consent, confidentiality, and retention policy allow, and choosing what is worth keeping. Convert means using AI to organize or critique the material and then applying the leader's own judgment. Close means reflecting, refining, deciding, and completing the work rather than letting an AI draft become the deliverable by default. The conversion and closure are where the leader's development happens.

The Wish List Tactic. A second simple practice is to keep a running list of things you wish AI could help you do but do not yet trust it to do. Review the list weekly. Pick one bounded item and test it at an appropriate consequence level. The practice builds the habit of imagining new uses while producing evidence about what works, what fails, and where the boundary of the leader's practice currently sits.

WHAT THE RESEARCH SHOWS

Microsoft's 2026 Work Trend Index surveyed 20,000 AI-using workers across ten countries. Its self-reported composite analysis attributed more than twice as much reported AI impact to organizational factors (culture, manager support, talent practices, and related conditions) as to individual factors alone (67 percent versus 32 percent). Workers who said their managers openly used AI were also associated with higher readiness and value scores. The report supports a management effect, but it does not show that visible manager use doubled workforce fluency or establish causation. The broader lesson is that individual skill develops inside an organizational system.

Microsoft (2026). 2026 Work Trend Index Annual Report: Agents, Human Agency, and the Opportunity for Every Organization.

Why the Modeling Effect Compounds

The modeling effect can form a reinforcing loop in the systems-thinking sense of Section 7.1. A leader develops practice and makes appropriate parts visible. That visibility can give the workforce permission and support to experiment. Workforce practice then surfaces use cases the leader had not considered, deepening the leader's understanding in return. Under the right conditions, the cycle compounds.

The same loop can run in the opposite direction. A leader does not develop practice, rewards no changed behavior, and communicates mainly through slogans. The workforce reads the incentives and keeps AI use shallow or hidden. That weak adoption then confirms the leader's sense that AI is not producing value. This is a plausible failure loop, not the explanation for every company; access, job design, trust, governance, capability, and the quality of the tools also matter.

The leader's decision to develop practice is therefore one point of influence among several. Its cumulative effect can still be large because leaders influence the organization's funding, rewards, permissions, and learning.

Visibility also needs a boundary. The point is not for a CEO to turn every prompt into theater or make employees perform experimentation because the boss is excited. Useful modeling is specific and honest: here is the task, here is where AI helped, here is what it missed, here is what I checked, and here is why I remained accountable for the result. A leader who shows only polished wins teaches concealment. A leader who exposes confidential material teaches

recklessness. The practice compounds when people can see disciplined judgment, including the decision not to use AI where it does not belong.

BRING THIS INTO THE ROOM

If this section has surfaced that you have not yet developed your own AI practice, three concrete moves are worth making this week. First, locate your current stage using the adapted four-stage framework. Treat the answer as a prompt for reflection, not a score. Second, name the obstacle most active for you right now: ego, habit, trust, or isolation. Naming it makes the intervention more specific. Third, start a wish list and run one bounded AI experiment per week. Record what worked, what failed, what evidence you checked, and what you would change. Six weeks of deliberate practice will teach you things reading alone cannot.

The Strategic Case

The strategic argument for taking the leader's own practice seriously is that the alternative produces a particular kind of AI-Native failure that is difficult to diagnose from the outside.

The company has invested in AI. The technical work is competent. The workforce has received the communications. The transition is, at the surface, underway. Eighteen to thirty-six months in, the leadership team notices that the AI-Native transition is not producing what was expected. The pilots are not graduating. The workforce is using AI but in shallow ways. The strategic decisions that AI was supposed to inform are still being made the way they were made before. The leadership team commissions an analysis. The analysis cannot find a single root cause, because the cause sits outside the system. It sits in the leadership team's individual practices, which have not developed enough for the leadership team to recognize what depth in the rest of the company would look like.

Companies whose leaders have done this work can produce a different outcome. Strategic and architectural decisions are informed by direct experience as well as specialist advice. Communication with the workforce can carry more credibility because leaders speak from practice rather than briefings alone. That does not guarantee compounding capability, but it improves the quality of the questions, trade-offs, and behavior leaders model.

The leader who has begun their own practice has done the precondition for the work the next section describes. The human reckoning with the workforce, the honest acknowledgment of what is happening, the construction of the covenant, all depend on the leader's own credibility.

The credibility comes from the work the leader has done with themselves. The next section addresses what the leadership team does once that work has begun.

Section 5.2

The Human Reckoning

What does an AI transition feel like from inside the workforce?

The leader who has begun their own practice is ready for the harder half. The picture the first four parts built, of what an AI-Native company looks like at the structural level, is incomplete. The structural work matters.

A major determinant of success is the work your leadership team does with the people whose lives are being reshaped. That work is harder, slower, and far less amenable to the strategy-deck treatment than the structural work. It is also a common source of failure, and the failure may first appear as weak adoption, missing knowledge, poor implementation, or stalled value rather than as a workforce problem.

This work is what I call the human reckoning. The term deserves precision.

WHAT THE SURVEYS SHOW

ADP's 2025 Global Workforce Survey of more than thirty-nine thousand workers, reported in Today at Work 2026, found that only twenty-two percent strongly agreed their job was safe from elimination. The gradient is worth holding: eighteen percent of individual contributors, twenty-one percent of frontline managers, and thirty-five percent of C-suite executives and upper managers felt secure. Separately, BCG's 2025 AI at Work survey found that employees at organizations undergoing comprehensive AI-driven redesign were more worried about job security (forty-six percent) than those at less-advanced companies (thirty-four percent). These are self-reported associations, not proof that redesign caused the anxiety, but they show why deeper transformation does not automatically dissolve fear.

ADP Research Institute (2026). Today at Work 2026, Issue 1. / Boston Consulting Group (2025). AI at Work 2025: Momentum Builds, but Gaps Remain.

Figure 5.2a. The seven emotions experienced workers actually feel during an AI-Native transition. Naming them is the work of the leadership team.

What the Reckoning Is

A reckoning is an honest account. It is the moment in which the actual facts of a situation are acknowledged, including the facts that are uncomfortable to acknowledge. It is the opposite of a euphemism. It is the work of letting the people in the room know three things. That you, as a leader, understand what is happening to them. That you are not going to pretend otherwise. That the conversation about what to do next will start from the honest ground rather than from the corporate version of the situation.

Why a reckoning is needed here is specific. A workforce that built the company over the last twenty or thirty years may be asked to participate in dismantling or redesigning parts of the expertise that defined their careers. The senior analyst who became valuable by synthesizing complex information faster and more reliably than anyone else may now help deploy a system that performs parts of that work in a fraction of the time. The experienced operator who could

read the line and anticipate problems may be asked to teach a system intended to support the next generation. The middle manager who built a career coordinating information across a team may watch parts of that coordination become automated.

These people are not units of production whose obsolescence is a market signal. They are people whose lives are organized around the work they do. They have raised families on the income from that work. They have invested years of effort in becoming good at it. They are now being told, in language that varies from honest to evasive, that the work they spent their lives learning to do is being changed in ways they did not choose and cannot control.

What they are experiencing has names. I want to use the names, because the alternative is to use euphemisms, and the euphemisms tell the workforce that the leadership team is not willing to acknowledge what is actually happening.

The Named Emotions

There is a grief of expertise. This is the experience of watching the skill you developed over decades become less central to the way work is done. The skill itself does not disappear. The centrality of it does. What is grieved is the place that competence held in the world.

There is a threat to status. Status in companies is closely tied to what you know that others do not, what you can do that others cannot, and the matters on which people defer to you. AI changes the distribution of all three. The status that was earned over years is being recalibrated in ways the person who earned it had no part in choosing.

There is a fear of obsolescence. This is the fear that the changes will not stop where the leadership says they will. The fear is rational. The history of corporate communications about workforce transitions is full of reassurances that did not hold. The workforce knows this. The leadership team that pretends the workforce does not know it is producing the cynicism that becomes the obstacle to the transition.

There is survivor guilt for those who keep their jobs while colleagues do not. The person who is still employed after a workforce reduction carries the knowledge that the difference between them and a former colleague was not always merit. It was often luck, or politics, or which team happened to be in the path of the change. The survivor's relationship to the company is altered by this. The company often does not address it, because the survivor is still there and the colleague is not, and the conversation feels uncomfortable to have.

There is the shame of starting again. The senior person who has to learn a new way of working is doing it visibly, in front of colleagues who used to look to them as the expert. The learning is slow. The mistakes are public. The status that was solid is now provisional. The shame of being a beginner again, after a career of being the one others learned from, is real and is rarely named in the rooms where transition decisions are made.

There is the quiet sense that the company has lied to them. The deal that the workforce had with the company, the implicit contract about what loyalty was worth and what experience earned, was made under one set of assumptions. The new set of assumptions is being introduced without the deal being renegotiated. The workforce is being asked to accept the new terms while the leadership team frames the situation as if the old deal still applies. The workforce is not fooled.

There is also a complicated relief. Some of the work that is now being automated was crushing the people doing it. The analyst who has been burning out for years on the volume of routine work is genuinely glad to have AI handle the routine. The relief is real. It is also complicated, because it is mixed with the other emotions on this list, and because the person who feels relief has to operate in a workplace where expressing it openly would be misread.

These seven emotions are a vocabulary, not a claim that every person feels all seven or that every transition produces the same response. Relief and opportunity can sit beside fear, grief, anger, or none of them. The leadership choice is whether to create enough safety to learn what is actually present rather than assume it away.

Why the Reckoning Belongs at the Center

There is an argument that the human reckoning is a topic for HR rather than for the leadership team designing the AI-Native transition. The argument would be that the technical and operational decisions are what produce value, and the human side is a downstream consequence to be managed by the appropriate function.

That argument is wrong, and it is wrong for operational reasons before ethical ones. A workforce excluded from an honest conversation is less likely to trust or engage deeply with the transition. Institutional knowledge about how work actually gets done is distributed among the people experiencing the change, as well as in systems and artifacts. Useful capture requires participation, and participation is easier to earn when leadership treats people honestly. HR has

essential expertise here; the senior leadership team cannot delegate its own accountability to HR.

A company that skips the human reckoning may discover later that deployed systems are not improving as expected. They work on easy cases and fail on exceptions. One possible cause is that the design lacks institutional knowledge held by experienced people who were never engaged, never asked the right questions, or had no reason to trust the process. That is not the only cause (data, architecture, evaluation, and model limits also matter), but it is one technology work alone cannot repair.

Companies that take the human reckoning seriously create better conditions for engagement, knowledge transfer, correction, and learning. Those conditions can support compounding advantage over time. They do not guarantee it, and honest treatment should not be reduced to a tactic for extracting knowledge. The ethical case stands on its own; the operational case makes neglect even harder to defend.

The Leaders Go First Principle

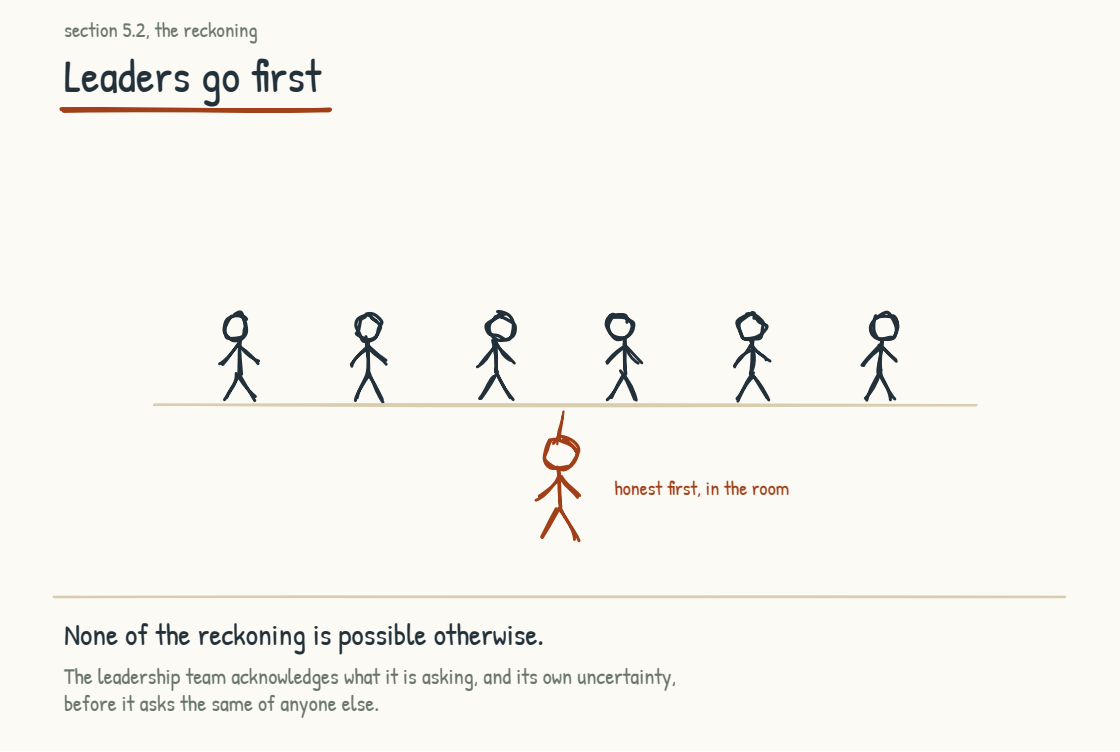

Figure 5.2b. One step out of the line. The team acknowledges what it is asking, and its own uncertainty, before it asks the same of anyone else.

None of this reckoning is possible unless your leadership team is willing to be honest themselves first.

The leadership team has to acknowledge, in the room, what they are asking of the workforce. They have to acknowledge their own uncertainty about how the transition will unfold. They have to acknowledge that they are making decisions with incomplete information and that some of the decisions will turn out to have been wrong. They have to acknowledge that the careers of people they have worked with for years are being altered in ways the leadership team is responsible for, and that the responsibility weighs on them.

This kind of acknowledgment is rare in corporate settings, and the reason it is rare is that leaders have been trained, often over decades, to project confidence and control. The training served them in earlier transitions. It does not serve them here. The workforce going through an AI-Native transition has more information than the leadership team realizes. They know what is happening. They know what is being said publicly and what is being said privately. They know when they are being managed. The leadership team that projects unrealistic confidence is producing the cynicism that defeats the transition. The leadership team that speaks honestly is producing the trust that enables it.

The honest speech does not have to be elaborate. A senior leader who stands in front of the workforce and says, simply, "I know this is hard. I know some of you are afraid. I know I am asking you to do something that I would find difficult to do in your place. I want to tell you what I am committing to in return for what we are asking of you," has done more than a leader who delivers a polished transformation narrative.

What We Are Leaving Behind, What We Are Carrying Forward

Figure 5.2c. The two questions have to be asked together. Grief with nothing preserved produces resistance, and preservation with no room to grieve produces disengagement.

What makes the reckoning bearable in practice is pairing two questions, asked together, in the same conversation.

What are we leaving behind. This is the honest acknowledgment of the losses. The ways of working that are changing. The expertise that is being recalibrated. The roles that are evolving or ending. The status structures that are dissolving. The leadership team names these out loud. The workforce names what is being lost from their side. The conversation is uncomfortable because the losses are real and naming them does not eliminate them.

What are we carrying forward. This is the honest acknowledgment of what is being preserved. The institutional knowledge that matters. The values that define the company. The relationships that have been built. The expertise that becomes more valuable, not less, in the AI-Native context. The commitments to people that survive the transition. The workforce participates in

naming these too, because the workforce often knows better than the leadership team what is genuinely worth preserving.

Pairing them matters because loss without a credible account of continuity can deepen fear, while preservation language that leaves no room to name loss can feel evasive. Neither reaction is universal, and resistance may also be a rational disagreement with the plan. The pairing holds continuity and change in the same conversation, giving the workforce a more honest basis on which to respond.

INTELLECTUAL LINEAGE

What I am describing draws on a leadership posture Edgar Schein spent his career naming. Schein, a foundational scholar of organizational culture, argued that leaders who default to telling can miss what they most need to know. Leaders who ask, and ask in ways that lower status barriers and invite candid answers, can learn what formal reporting hides. He called the practice humble inquiry. Schein and Warren Bennis also used the term psychological safety in their 1965 work on organizational change, an early root of the later research tradition discussed in Section 6.1. The covenant below is one possible artifact of humble inquiry practiced during a serious transition, not a method Schein himself prescribed.

Schein, Edgar H. (2013). Humble Inquiry: The Gentle Art of Asking Instead of Telling. Berrett-Koehler. / Schein, E. H., and Schein, P. A. (2016, 5th ed.). Organizational Culture and Leadership. Wiley. / Schein, E. H., and Bennis, W. G. (1965). Personal and Organizational Change Through Group Methods. Wiley.

The Covenant

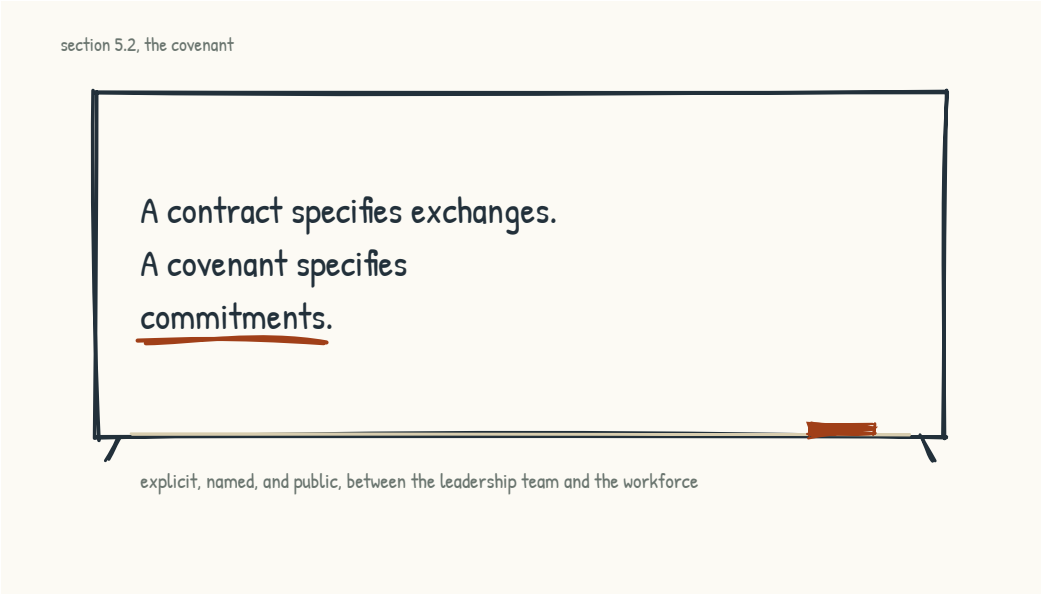

Figure 5.2d. The line the whole covenant stands on, kept where the argument makes it.

Out of that reckoning comes a covenant. I use the word deliberately. A covenant is not a contract. A contract specifies exchanges. A covenant specifies commitments. The covenant in an AI-Native transition is the explicit, named, public agreement between the leadership team and the workforce about what each commits to and what each is asking of the other.

The leadership team commits to specific things. To name the truth even when the truth is uncomfortable. To honor the institutional knowledge of experienced workforce members through their participation in the design of the new operating model. To invest in the development of every workforce member who wants to participate in the AI-Native company. To make compensation, role, and recognition decisions that match the rhetoric. To acknowledge the human cost of the transition rather than pretending it does not exist.

The workforce is asked to commit to specific things: to engage honestly in the work of transition, share institutional knowledge under fair and transparent terms, learn what needs to be learned with real support, and participate in shaping the company rather than wait for it to be done to them. These asks are not morally equivalent to leadership commitments because leadership holds greater power over jobs, compensation, and the design of the transition. Participation must not become coerced consent, and knowledge capture must not become an undisclosed mechanism for eliminating the person who provided it.

That power difference also means the workforce needs a credible way to decline, contest, or appeal an ask without retaliation. A commitment made because a person's livelihood is at risk is not evidence that the covenant was freely shaped.

A covenant is a real artifact. It can be written down. It can be revisited. It can be referenced in the moments when the transition is producing strain and the question of what was agreed needs to be answered. The companies that produce a real covenant have given themselves a reference point. The companies that produce a covenant in rhetoric only have given themselves a hostage to fortune.

BRING THIS INTO THE ROOM

The covenant is a real artifact, not a metaphor. Draft it with workforce participation, not only inside the leadership team, in two columns: what leaders commit to and what they ask of the workforce. Make the commitments specific: compensation, role design, recognition, learning time, privacy, appeal, and knowledge-capture terms. Make the asks specific: engagement, knowledge sharing, and willingness to learn. State what happens when conditions change or a commitment is broken. Sign it, communicate it, and revisit it quarterly.

The Institutional Knowledge Connection

A connection runs through here that should be made explicit: the link between the human reckoning and the institutional knowledge architecture from Section 4.1.

The knowledge architecture described in Section 4.1 draws partly on institutional knowledge held by experienced workforce members. Capturing and validating that knowledge requires their participation. Willingness is influenced by whether the process feels respectful, safe, consequential, and fair, and by whether people believe the stated purpose of capture.

Experts who feel honored are more likely to share deeply. They may sit for interviews, walk modelers through what they know, challenge architects when the representation is wrong, and take pride in supporting the next generation. But respect is not a transaction that purchases disclosure; workload, incentives, legal duties, personality, and the limits of tacit recall also affect what can be captured.

Experts who feel threatened may protect their knowledge, comply narrowly, or disengage. They may answer questions without volunteering what the interviewer did not know to ask, or leave

errors unchallenged. Sometimes the knowledge leaves with them not as sabotage but because the company failed to create the time, method, or trust required to make it explicit.

The human reckoning can influence the difference between those outcomes. Investment in honoring the workforce therefore supports institutional knowledge capture, but it is not sufficient by itself; modeling quality, governance, maintenance, and technical design still matter. A company that treats people merely as a source to be mined is likely to build a shallow representation and damage the trust needed to repair it.

This is the strategic argument for the human reckoning, alongside the ethical one. The ethical argument is sufficient on its own. The strategic argument is the additional reason why leadership teams who care about competitive advantage should treat this work as foundational rather than as overhead.

Section 5.3

What Humans Bring

What can humans do that AI cannot?

The previous section was about what humans are losing, what they are afraid of, and what the leadership team has to do to honor that experience. That framing is necessary. It is also incomplete, because it can leave the impression that the human role in an AI-Native company is primarily defensive. It is not. What follows is the positive case for what humans bring, made as a strategic argument rather than as a sentimental one.

One influential strain of AI-Native discourse places humans at the edge of operations: machine intelligence at the center, people aligning it from the boundary, and the workforce becoming whatever automation has not yet absorbed. The Coinbase framing introduced in Section 2.2 points in this direction, and variants are appearing in other public commitments. It deserves direct scrutiny rather than automatic acceptance.

That framing can produce a strategic error. It treats work as monolithic and AI as a capability expanding toward its boundary. Work has multiple dimensions, and AI is differentially capable across them. Humans bring forms of value tied not only to task performance but to embodiment, social participation, responsibility, and lived stakes. A company that protects and

amplifies those forms may outperform one that treats people as residue. That is a strategic hypothesis to test in the operation, not a sentimental exemption from measurement.

I want to name what humans bring. Seven kinds of value, with the understanding that the list is not exhaustive and that the kinds overlap. They are organizing categories, not a closed taxonomy.

The categories are not claims that a model can never produce a novel option, empathetic sentence, ethical analysis, or coherent story. Models already do all four. The distinction is where standing, participation, responsibility, and lived consequence enter the work. Performance can be generated. Authority and accountability still have to be placed, and a company that places them carelessly has made an operating-model decision whether it recognizes it or not.

INTELLECTUAL LINEAGE

The argument has a serious philosophical history, though that history should not be treated as proof of permanent machine limits. Lucy Suchman's 1987 Plans and Situated Actions challenged the idea that human action is simply the execution of pre-formed plans, emphasizing how action is produced within unfolding situations. Hubert Dreyfus argued from phenomenology that expertise depends on embodiment, contextual involvement, and intuitive judgment that cannot be understood as rule-following alone. Hubert and Stuart Dreyfus also developed the influential five-stage skill-acquisition model: novice, advanced beginner, competent, proficient, expert. Modern AI differs substantially from the systems these authors examined, and some of Dreyfus's capability predictions have aged unevenly. Their durable contribution here is the question they force: what is lost when performance is abstracted from situation, body, relationship, and responsibility?

Suchman, Lucy A. (1987). Plans and Situated Actions: The Problem of Human-Machine Communication. Cambridge UP. / Dreyfus, Hubert L. (1972, updated 1992). What Computers Still Can't Do. MIT Press. / Dreyfus, H. L., and Dreyfus, S. E. (1986). Mind Over Machine.

WHAT THE RESEARCH SHOWS

A landmark field experiment by researchers affiliated with Harvard, BCG, MIT, Warwick, and Wharton studied 758 BCG consultants and described a jagged technological frontier. On selected tasks designed to fall inside the tested model's frontier, AI-assisted participants completed 12.2 percent more tasks, worked 25.1 percent faster, and produced outputs rated more than 40 percent higher in quality. On a selected task outside that frontier, AI-assisted participants were 19 percentage

> *points less likely to reach the correct answer. The experiment does not map every task or establish a permanent human-machine boundary. It shows that capability can vary sharply across superficially similar work and that users may not know which side they are on. An operating model that assumes uniform capability will fail precisely where polished output disguises that boundary.*
>
> Dell'Acqua, F., McFowland, E., Mollick, E., Lifshitz-Assaf, H., Kellogg, K., et al. (2023). "Navigating the Jagged Technological Frontier." Harvard Business School Working Paper 24-013.

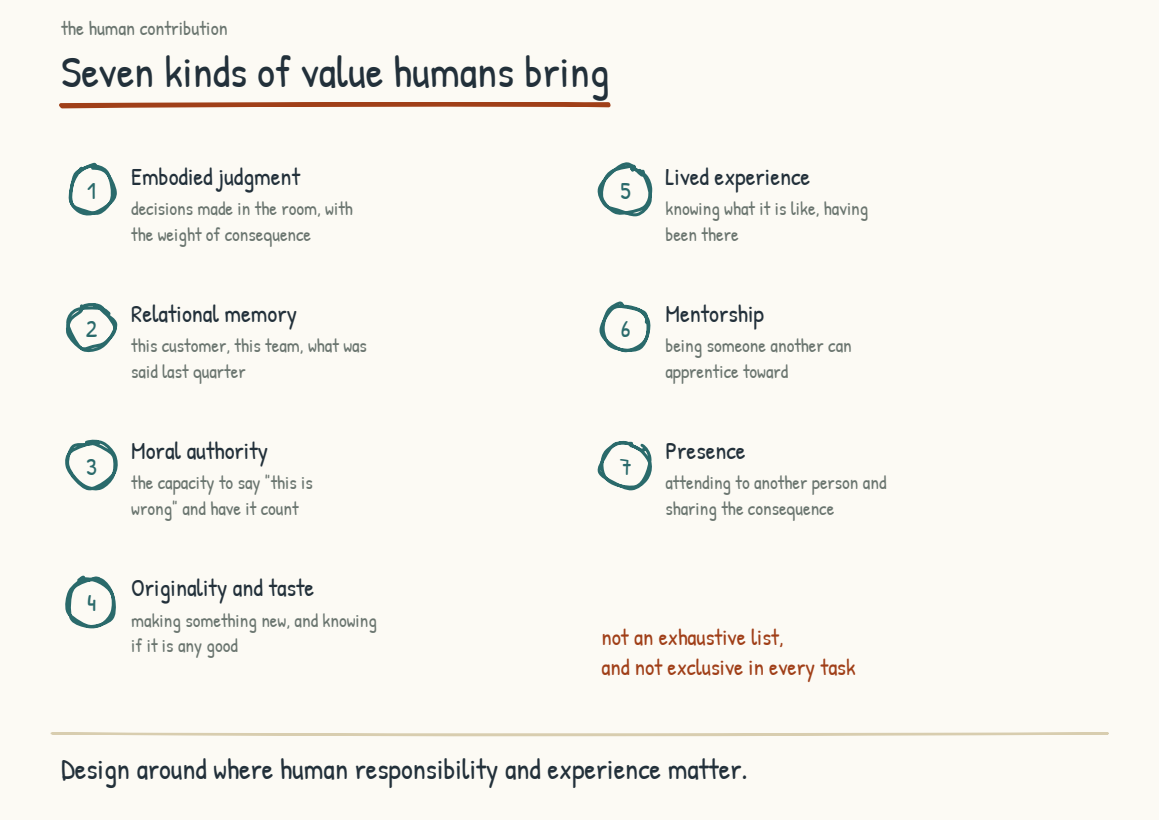

Figure 5.3a. Seven kinds of human value tied to lived context, responsibility, and relationship. The operating model must decide where each is consequential.

Contextual Knowledge

First, humans bring the lived sense of what is going on in this place, with these people, right now. AI can be given extensive multimodal and historical context, sometimes detecting patterns people miss. It still receives a representation selected by sensors, records, interfaces, and permissions. An experienced person participates in the situation. The supervisor hears that a machine has sounded slightly different for a week. The account manager senses a change in the customer's tone not captured by a transcript. The teacher notices that an engaged student is

sitting differently today. Imagine a hospital nurse, three hours into a night shift, who pauses because a patient does not look right even though the monitor reads normally. She escalates. The cardiac event comes an hour later.

This kind of knowledge is accumulated calibration to a situation. People can misread it, and analytic systems can improve it. The distinctive value is that a person can integrate weak signals with lived context and take responsibility for escalating before the evidence is clean. Strong systems combine that situated judgment with measurements rather than forcing a choice between them.

Judgment Under Genuine Novelty

Second, humans bring accountable judgment in situations that have not happened before. Both humans and models draw on prior experience or learned patterns, and both can generate new combinations. When a situation has no close precedent, model output may still be useful, but its apparent coherence can outrun the evidence available to validate it.

The cases that matter most in many businesses include unprecedented customer situations, supply-chain disruptions without a close historical analog, competitive moves that require reframing the market, and crises in which the playbook does not apply. In these situations, capable humans can recognize novelty, suspend a default frame, draw across experience, consult others, and own a decision under uncertainty. AI can contribute options and analysis. It does not bear the institutional, legal, or moral consequences of choosing among them.

This kind of judgment is not the routine application of expertise. It is the higher-order capacity to recognize that the situation is novel, to suspend the default analytic frames, and to construct a response from first principles. The capacity is rare even among experienced workers. The capacity is, in my experience, what distinguishes the senior people whose judgment the company depends on from the senior people whose role can be substantially automated.

Combinatorial Creativity from Lived Experience

Third, humans bring creativity shaped by lived experience that no training corpus fully represents. Generative AI can produce surprising combinations and is not usefully dismissed as mere copy-and-paste. It is nevertheless conditioned by training, context, tools, and selection. A person who has spent thirty years in a domain brings a different body of combinatorial material: specific colleagues, industry shifts, failures, successes, bodily environments, and the felt consequences of each.

Many original ideas come from people who have crossed boundaries: the doctor who started as a physicist, the financial analyst who taught poetry, the product manager who spent five years in a plant. AI can support cross-domain reasoning. What the person adds is not exclusive access to the combination but a lived understanding of why the combination matters, where the analogy breaks, and what consequences follow in practice.

The strategic implication is direct but not automatic. A company that hires across backgrounds and gives people real authority and space to contribute can access a wider range of frames and lived intersections. Demographic variety without inclusion does not guarantee creativity, and a homogeneous team can still produce excellent work. Diversity is more than compliance, however: under the right conditions it is an operational asset, and AI does not erase the value of who has actually lived which parts of the problem.

The Ethical Sense

Fourth, humans bring responsibility for what should be done rather than only analysis of what could be done. AI can discuss ethics, apply frameworks, and surface competing considerations. Current systems do not feel moral injury, belong to a community, or bear accountability. Humans can feel that a course is wrong before they can fully articulate why. That intuition is not infallible, it can encode bias as easily as wisdom, but it is a signal that deserves examination when explicit frameworks fail to settle the case.

This matters operationally because most consequential business decisions have ethical dimensions, and the ethical dimensions often determine whether the decision works. A company can produce a perfectly legal action that the workforce knows is wrong, and the workforce will respond by disengaging from the company. A company can refuse a profitable action because the workforce will not participate in something they cannot defend, and the company will be better off for the refusal even though the analysis would have argued for the action.

Humans in the operating model remain the accountable source of this judgment. A leadership team that includes people willing and empowered to say "this is wrong, regardless of what the analysis says" adds protection no governance framework supplies alone. Human objection also needs evidence, due process, and challenge; conscience is not a license for unexamined preference. Companies that punish good-faith ethical objection weaken a capacity that is difficult to rebuild and may discover the loss through misconduct, scandal, or reputational damage.

Relational Presence

Fifth, humans bring the felt experience of being met by another person. AI can generate empathetic language, adapt its tone, and sometimes outperform a rushed human on measures of perceived warmth. It is not a person, and users do not always know when they are interacting with AI unless the system discloses it. The relevant distinction is not whether the words sound caring. It is whether a person is present, responsible, and capable of entering an ongoing reciprocal relationship.

The relational dimension matters in every situation where humans are doing something that depends on trust. The customer making a decision that involves real money. The employee navigating a difficult moment in their career. The patient receiving news that changes the trajectory of their life. The student trying to understand something that requires a teacher to read their face and adjust. In these situations, the presence of a real human, who is paying attention, who is responding to the specific person in front of them, who is bearing witness to what is happening, is doing work that AI does not do.

Automating every one of these interactions may look efficient in the quarter. In some low-stakes settings it may also improve access or consistency. In high-trust moments, undisclosed or poorly designed automation risks a later cost in alienation, error, or deteriorating trust. That risk should be measured rather than assumed away.

Meaning-Making

Sixth, humans construct and contest the frames that orient other people. A leader who articulates what the company is for in a way the workforce can carry through difficult quarters is doing more than generating language. The frame is an act of meaning-making backed, or contradicted, by the leader's history and choices. AI can help draft the words. It cannot supply the legitimacy the audience grants, or withholds, based on who says them and what that person has done.

Meaning-making is also what experienced workers do with newer ones. AI can explain procedures and offer interpretations. A mentor can connect how the work is done with why this community believes it matters, answer for the culture being transmitted, and allow the newer person to challenge it. Human-to-human transmission is not the only way culture travels, but it is one of the ways culture becomes reciprocal rather than merely delivered.

Meaning-making is work. It can be selected for and it can be developed. Treating it as a soft skill rather than as a capacity the operating model depends on is how a company loses it without noticing.

Narrative

Seventh, humans place events in a story that makes them coherent. The events of a company do not arrive already assembled into one honest through-line. The workforce experiences them as connected because someone has done the work of interpreting what belongs together, what changed, and what it means. Narration is the human act of making the year's activities intelligible without pretending the story is the only possible one.

A workforce that has a coherent narrative about what the company is doing and why can absorb individual setbacks without losing momentum. A workforce that does not have a coherent narrative experiences setbacks as evidence that the company is failing, even when the setbacks are normal. The narrative is the difference. It is constructed by humans, primarily by leaders, drawing on their understanding of the company and their judgment about what to emphasize.

AI can summarize events and generate compelling narratives, including narratives tailored to a company's stated culture. What it cannot do on its own is possess the standing to decide which story the organization should live by or accept responsibility for what that story includes and excludes. Narrative remains human work not because machines cannot produce prose, but because people must authorize, contest, and live with the frame.

The Strategic Case Against the Humans at the Edge Framing

The seven kinds of value, taken together, make the strategic case. A company that designs around them can produce work competitors cannot replicate merely by buying the same AI tools. Contextual knowledge, judgment under novelty, lived creativity, ethical responsibility, relational presence, meaning-making, and narrative are parts of the operation whose value depends on more than generated task performance. Protecting and amplifying them can support durable advantage. Treating them as residue risks removing the conditions under which the work becomes trustworthy, distinctive, and worth doing.

The "humans at the edge" framing can lead operating models to underinvest in the seven. It can make headcount reduction look efficient while leaving quality, trust, learning, resilience, and exception handling unmeasured. Replacing experienced workers with less experienced workers

supported by AI may work in some tasks and fail badly in others. The decision should follow evidence about where value resides, not an assumption that experience has become redundant.

AI belongs in operations. The argument is for designing the operating model around the actual distribution of value, which puts AI at the center of some kinds of work and humans at the center of others. The Coinbase framing is correct for some kinds of work in some kinds of companies. It is wrong as a universal claim about where the intelligence sits and where the humans go. The right answer differs by archetype, by industry, by the specific value the company is trying to deliver. The wrong answer is the framing that assumes the question has been settled in favor of one position.

The next section addresses the cognitive posture that distinguishes workforces that produce the seven kinds of value over time from workforces that lose the capacity to produce them. The cognitive posture is one of the most consequential things a company can develop, and a leadership team can be deliberate about developing it.

Section 5.4

AI for Thinking, Not AI as Thinking

What is the difference between using AI to think and letting AI do the thinking?

A team that was building a company inside the AI-Native question discovered, about a year into the work, that they had developed a habit they had not chosen. They were defaulting to Claude Code for technical decisions that they used to make by talking to each other. The shift had been gradual. Someone would have an idea about a feature. Instead of bringing it to a colleague, they would put it into the AI and see what the AI suggested. The AI was good. The suggestions were often useful. The colleague got skipped.

The team noticed the pattern when they realized they had not had a substantive conversation about an architectural decision in weeks. They had been making decisions. Those decisions had been informed by AI. Thinking they used to do together had been replaced, without anyone deciding it should be, by individual interactions with the AI followed by integration of whatever it said.

The team pulled back deliberately. They reinstated practices they had not realized they had lost. They started talking to each other again about substantive questions before going to AI in cases where shared judgment mattered. They used AI to deepen and sharpen their thinking rather than automatically substitute for discussion. In the team's assessment, the work improved, their thinking became more active, and the relationships recovered something that had been quietly dissolving.

I tell the story as a worked example of a distinction that runs through the entire AI-Native question: using AI for thinking versus using AI as thinking. From the outside the postures can look similar. The same person with the same tool may move between them even within one task. Over time, the pattern may compound. Active engagement exercises framing, evaluation, and judgment. Repeated delegation can reduce practice in whichever skills the person no longer performs. The outcome is a risk to monitor, not a universal law that every use sharpens or hollows the user.

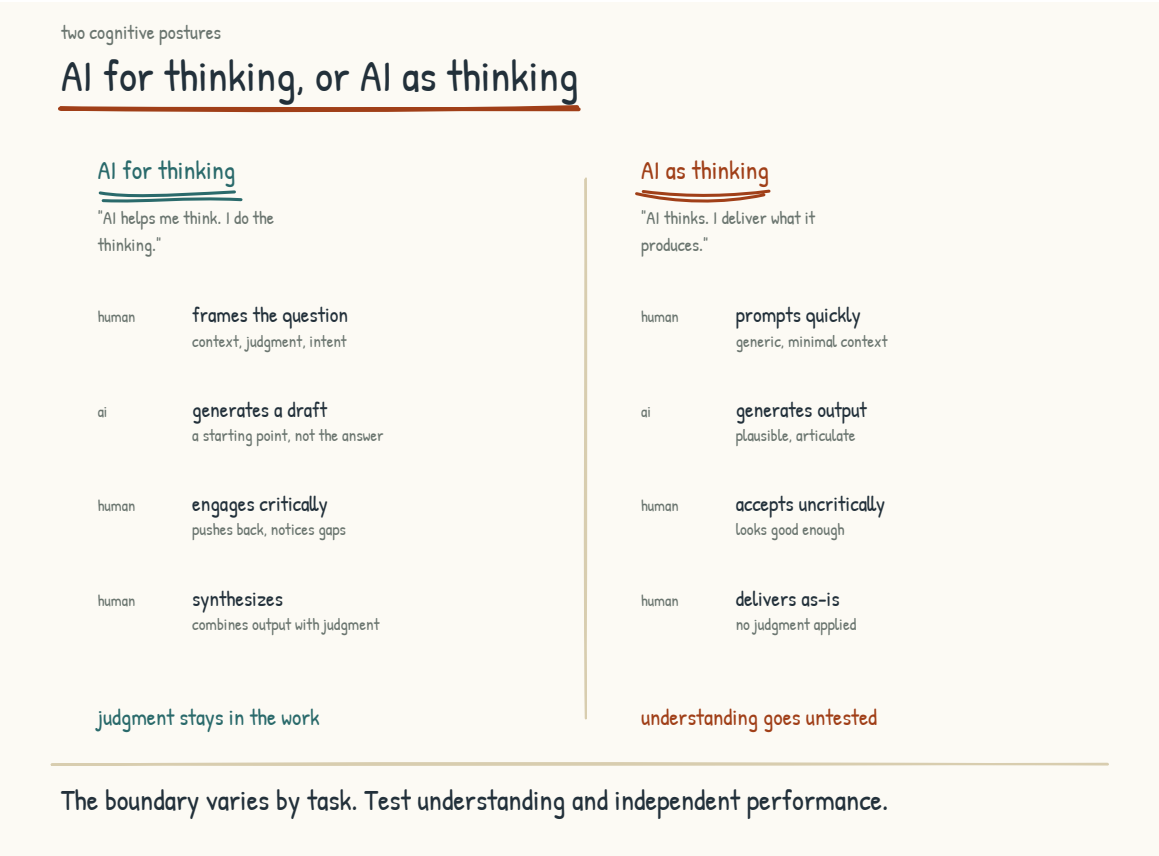

Figure 5.4a. Two cognitive postures. Active engagement exercises judgment; unexamined delegation risks reducing practice in the skills being offloaded.

The Two Postures

AI for thinking is the posture in which the human engages with the AI's output as an interlocutor. The human pushes back on what the AI produces. The human integrates the AI's contribution with their own knowledge, judgment, and reading of the situation. The human comes out of the interaction having thought more, not less, than they would have without the AI. The AI is doing work, and the human is doing work, and the work is genuinely collaborative.

AI as thinking is the posture in which the human treats the output as the thinking, full stop. The human prompts, the AI responds, and the human accepts and integrates mechanically. Framing a good task can itself require thought, and delegation is often rational. The risk arises when the person repeatedly offloads skills they still need to understand, supervise, or perform under exception. Those skills may weaken through disuse, and the weakness may remain hidden until the system fails or the context changes.

The first posture is more likely to exercise judgment. The second creates a risk of deskilling. Neither guarantees an outcome, because effects depend on the task, prior expertise, interface, incentives, and how performance is assessed. Companies that want sustainable capability should measure whether AI use leaves people better able to explain, challenge, transfer, and perform the work, a harder test than how quickly they submit it.

This is not a distinction between virtuous people who use little AI and careless people who use a lot. A person can delegate most of a routine task and remain fully capable of supervising it. Another can type every prompt personally while letting the model choose the frame, evidence, and conclusion. The better test is authorship in the deeper sense: who understood the problem, noticed what was missing, made the consequential choices, and can answer for the result when challenged. Tool-use metrics cannot tell you that.

The Research

The concern is not unique to AI. Research on cognitive offloading examines how people shift memory, calculation, navigation, and other work to external aids. Offloading is not inherently harmful: writing, calculators, and search can free attention for higher-order work. The developmental question is whether the learner still practices and understands the capacity an assignment is meant to build. An aid can support deeper engagement or bypass it, depending on the design and use.

Research on AI-specific cognitive offloading is still preliminary. Educators report cases in which students submit fluent work they cannot defend, transfer, or extend, but anecdote does not establish prevalence or long-term decline. Experimental work suggests that how AI is used matters: structured engagement can support reasoning, while answer substitution can reduce effort on the target skill. The field does not yet justify a single claim about what AI does to every learner.

The working-life evidence is earlier still. Automation research gives reason to watch for deskilling, automation bias, and the loss of expertise needed during exceptions. AI can also scaffold learning, expose workers to alternatives, and raise the level at which they practice. The distinction in this chapter is therefore a design hypothesis: posture, task, interface, and incentives shape whether offloading substitutes for capability or helps build it.

WHAT THE RESEARCH SHOWS

Two 2025 studies are suggestive and easy to overstate. Michael Gerlich's mixed-method study of 666 participants reported a strong negative correlation (r = -0.68) between reported AI-tool use and critical-thinking scores, with cognitive offloading as a proposed mediator. Its observational design cannot establish that AI use caused lower critical thinking; prior ability, education, age, task choice, and other factors may influence both. Kosmyna and colleagues assigned 54 adults to write essays using an LLM, search, or no external tool across the first three sessions; only 18 completed the crossover session. The preprint reported the weakest EEG connectivity in the LLM condition along with linguistic and recall differences. EEG connectivity during one writing task is not a direct measure of atrophy, intelligence, or long-term capability, and published critiques have raised design, reproducibility, and reporting concerns. The studies justify attention and better experiments, and stop short of showing brain decline.

Gerlich, M. (2025). "AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking." Societies 15(1):6. / Kosmyna, N., et al. (2025). MIT Media Lab. "Your Brain on ChatGPT" pre-print study.

The Slow-Feedback Problem

What makes this distinction operationally dangerous is the possibility of slow feedback. AI-assisted work can remain acceptable while unaided recall, first-principles production, or exception handling receives less practice. That does not prove meaningful decline after a year, and workers may notice or compensate in other ways. The risk is that a capability gap becomes

visible only when tools are unavailable, output must be challenged, or an unfamiliar case requires the underlying skill. Rebuilding can then be slower and more expensive than preserving periodic practice would have been.

Slow feedback is hard to manage because immediate output metrics may reward offloading while revealing little about retained capability. Signals need not take years: explanation checks, transfer tasks, unaided samples, incident reviews, and simulation exercises can make them visible earlier. Without those checks, a company may attribute declining judgment or resilience to other causes because accumulated changes in how work is performed leave no single event to inspect.

The measurement has to begin with a capability the organization has decided it still needs. If engineers must repair a critical service without the generating system, test that. If analysts must recognize when a forecast violates the business, give them unfamiliar cases and ask for the diagnosis. If a profession requires independent judgment for licensure or safety, preserve the practice its standards demand. An abstract fear of cognitive decline produces generic training. A named capability, consequence, and test produce an operating decision.

A leadership team that takes this risk seriously will build practices that produce earlier signals. The practices are not exotic. I will name them next.

> **WATCH FOR THIS**
>
> **Deskilling can hide behind acceptable daily output. Watch for changes in work people still need to understand: the analyst who no longer forms a hypothesis without prompting, the strategist who stops challenging summaries, the engineer who cannot explain or repair generated code. Starting with a prompt or reviewing generated work is not itself evidence of decline. The signal is lost ability to explain, transfer, challenge, or perform when the situation requires it.**

Eight Practices

Think first when prior judgment matters. Before going to AI on a question you should be able to frame independently, write down an initial position, assumptions, or uncertainties. Then use AI to test, refine, or challenge them. For unfamiliar discovery work, AI may reasonably come first; the requirement is active evaluation afterward. The purpose is to prevent the model's first frame from becoming yours by default.

Critique reflexively. Whenever AI produces something consequential, look deliberately for a weakness, missing alternative, or unsupported claim. The practice trains non-acceptance without requiring performative disagreement. Calibrated critique matters more than contrarianism; the goal is to detect error, not invent one.

Cross-check sources. Open and inspect sources supporting consequential claims, and sample-check lower-risk work according to an explicit standard. Verification is not always cheap, so effort should scale with consequence, novelty, and the cost of error. The habit is to treat a citation as a path to evidence rather than evidence by itself.

Disagree visibly. When you disagree with the AI's output, say so out loud in the prompt. The purpose is to make your own thinking visible to yourself, and to remind yourself that your judgment is the thing that matters. Manipulating the AI has nothing to do with it. The visibility of disagreement keeps the worker in the role of thinker rather than the role of accepter.

Hold AI-free practice. Some cognitive work benefits from periods without AI: unaided drafting, initial brainstorming, calculation, diagnosis, or reflection. Which activities matter depends on the capabilities the role must retain. Reserve enough practice to test and exercise those capabilities; do not make tool abstinence a ritual disconnected from the work.

Watch for default behavior. The slip into AI-as-thinking is not announced. It happens through accumulated small decisions, each of which would have looked reasonable in isolation. The discipline is to notice the pattern. The team that defaulted to Claude Code did not decide to default. They drifted. The watch for the drift is what allows the correction to happen before the capacities have eroded.

Teach others. Teaching another person requires you to organize knowledge, respond to misunderstanding, and discover what your explanation omitted. Explaining something to AI can also expose gaps, but a person contributes independent confusion, experience, and challenge. Teaching remains a powerful way to exercise understanding as AI takes on more work that once transferred knowledge through doing.

Reflect on use. At the end of a substantive piece of work, look at how you used AI. Where did the AI add value. Where did you do work that the AI could not have done. Where did you let the AI do work you should have done yourself. The reflection is what produces learning. Without it, the patterns repeat.

These eight form a checklist, but they should be adapted to consequence and role rather than enforced mechanically on every task. Practiced over time, they can support habits of active engagement. The practices are teachable and testable. Companies should evaluate whether they improve explanation, transfer, judgment, quality, and resilience rather than assume the checklist itself creates capability.

The organization has obligations here too. A worker cannot preserve deep practice while every incentive rewards speed, every deadline assumes maximum automation, and every AI-free exercise is treated as inefficiency. Teams need time to compare approaches, permission to question generated work, access to source material, and assignments that still develop judgment. Managers need to distinguish a temporary learning cost from poor performance. Otherwise the company will celebrate the output gain while quietly removing the conditions under which capability is built.

The Strategic Case at the Company Level

The strategic argument for caring about cognitive posture is that unexamined offloading can weaken the judgment and craft on which differentiated work depends. Shared models and similar prompts can also pull output toward common defaults. Neither outcome is inevitable, but both are plausible enough to manage.

A company whose workers routinely accept AI output without engagement risks convergence in structure, language, assumptions, and taste. Differentiation that came from worker judgment and craft may erode, making the work easier for competitors using similar systems to replicate.

A company whose workers use AI as an active thinking aid has a better chance of gaining speed and breadth while preserving distinctiveness. Judgment, taste, and craft can be exercised rather than displaced. Whether they actually strengthen should appear in work samples, transfer tests, customer response, and performance under novel conditions.

Over a five-to-ten-year planning horizon, these paths could produce substantial differences. One company may become efficient but generic; another may combine efficiency with distinctive output and resilient expertise. Markets do not reward distinctiveness automatically, and admired companies fail for many reasons. The claim here is narrower: retained human capability preserves strategic options that a deskilled workforce no longer has.

The next section addresses the positive aspiration. The up-leveling that becomes possible when the cognitive posture is right and the conditions support it.

Section 5.5

The Up-Leveling Aspiration

What becomes possible when a worker has AI alongside them?

The cognitive posture from the previous section names what the worker has to do to avoid hollowing out. What follows names what becomes possible when the posture is right and the conditions support it. The possibility is up-leveling, and it is one of the most exciting and least well-understood prospects of the AI-Native era.

The simplest way to introduce up-leveling is through Vygotsky's concept of the zone of proximal development. That zone is the gap between what a learner can do alone and what they can do with help from a more capable other. Learning happens inside it. The learner who is supported through the zone, by a teacher, mentor, or capable peer, develops capacities that they did not have before. The learner who is left in the zone without support stays where they were. The learner who is given problems below the zone is bored. The learner who is given problems above the zone is overwhelmed. The zone is where development happens.

AI can sometimes serve as part of that scaffolding. It can explain a concept three ways, generate practice, question an assumption, and respond at the moment a learner gets stuck. But it is not Vygotsky's more capable other in the full human sense. It may be confidently wrong, it does not reliably know what the learner understands, and it cannot take responsibility for the learner's development. Used with sound material, verification, and human judgment, it can help a worker operate inside the zone more often. The implications are large, and they are still only partly realized in working life. I want to walk through the historical lineage, describe what up-leveled cognition can look like in practice, name the honest counterweights, and identify the conditions that make the outcome more likely.

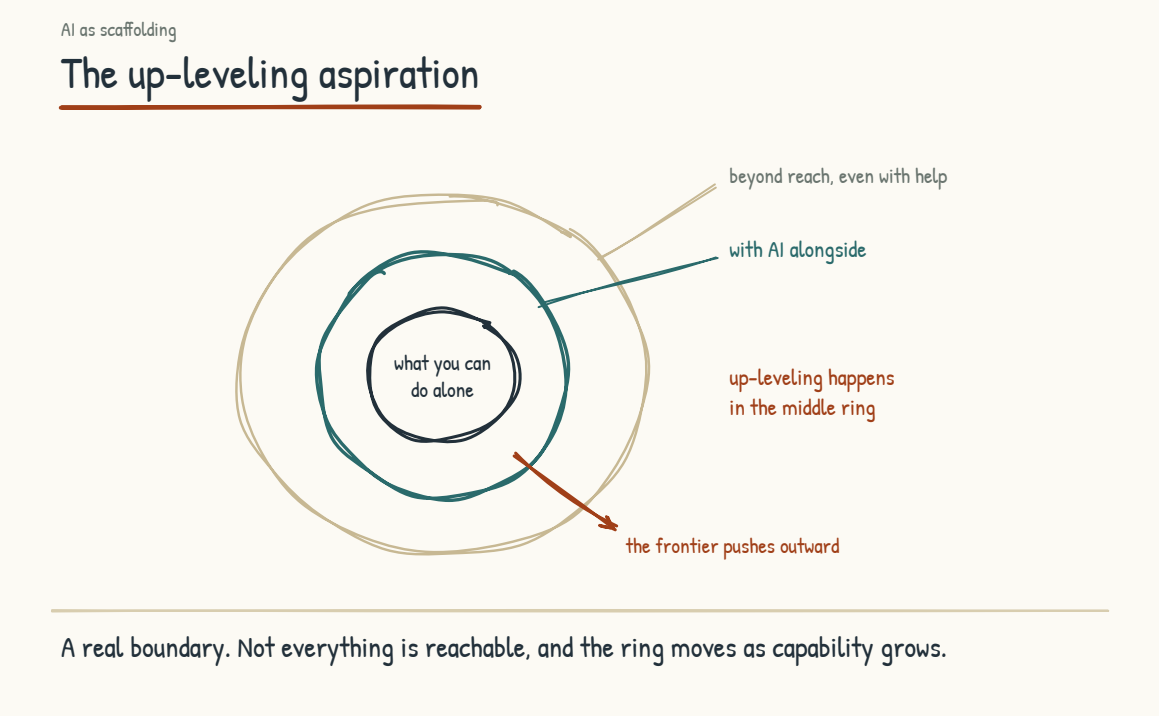

Figure 5.5a. Vygotsky's Zone of Proximal Development applied to AI-Native work. AI can form part of the scaffolding that expands what the worker can do with support.

The Historical Lineage of Cognitive Extensions

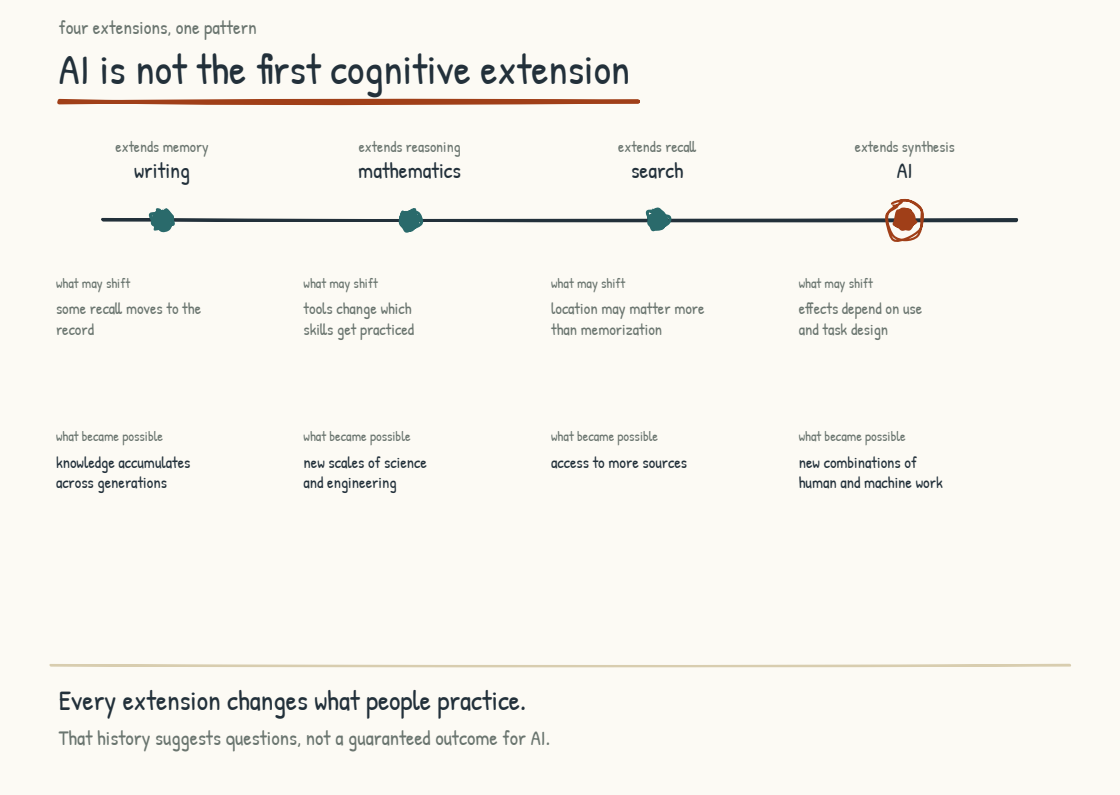

Figure 5.5b. Writing, mathematics, search, and now AI. Each extension changed which capacities people practiced and enabled work the prior generation could not imagine.

AI is not the first cognitive extension humans have invented. Writing was the first. Writing extended memory, allowing humans to hold thoughts outside their own heads, work on them over time, and pass them across generations. The invention transformed what was thinkable. Mathematics was another. Mathematical notation extended reasoning, allowing humans to manipulate abstractions that would have been impossible to hold in working memory. The invention enabled science, engineering, and the modern economy. Search was a more recent extension. Search engines extended recall, allowing humans to access vast amounts of information without having to internalize it. The invention transformed knowledge work.

Each extension provoked a familiar argument. Critics worried that writing would weaken memory, that calculators would displace mental arithmetic, and that search would replace the need to know things. Those concerns were not imaginary, but the historical record does not support a clean law that every extension weakens the capacity it extends. Tools change what people practice. A capacity that is routinely offloaded may deteriorate; another may be

strengthened by the same tool. The effect depends on the person, the task, the surrounding instruction, and whether the unassisted capacity is still exercised.

The extensions also enabled new capacities that the prior generation could not have imagined. Writing enabled accumulation of knowledge across generations that pre-literate cultures could not sustain. Mathematics enabled scientific reasoning that pre-mathematical cultures could not undertake. Search enabled synthesis across vastly more sources than any individual could have memorized.

AI belongs in this lineage, but it is not simply the next calculator. Generative systems can draft, synthesize, classify, translate, simulate perspectives, and produce plausible errors through the same interface. That breadth makes the extension unusually powerful and unusually easy to overtrust. Some unaugmented capacities may weaken where practice disappears. Others may grow when AI creates better practice, feedback, or access. A generation raised with AI will almost certainly do things its predecessors did not, and it may also need deliberate ways to retain capacities its tools make easy to neglect.

That history is useful because it calibrates without settling the argument. Anxiety has accompanied earlier cognitive tools, but resemblance is not proof that today's risks will resolve the same way. AI will change what people practice, what institutions reward, and eventually what competence means. The question is not whether cognition will change. It is which capacities we choose to extend, which we refuse to surrender, and whether those choices are made deliberately.

What Up-Leveled Cognition Looks Like

Picture one analyst on two Tuesdays, two years apart. On the first Tuesday she spends the morning pulling market data into a spreadsheet and the afternoon formatting the report. On the second Tuesday the pull and the formatting are finished before she sits down, and she spends the morning pushing back on three scenarios the AI drafted overnight and the afternoon on the phone with a plant manager, asking about a constraint the model could not see. Her business card reads the same.

Up-leveled cognition can hold wider problems. The worker who used to think about one project may be able to examine a portfolio. The analyst who focused on one market may be able to compare an industry. The expansion does not happen because AI magically enlarges working memory. It happens when the system can reliably externalize part of the search, organization,

calculation, or comparison that used to consume attention. Wider scope also creates more ways to miss an error, so the worker needs methods for checking the whole as well as inspecting the parts.

Up-leveled cognition can work at higher levels of abstraction. The strategy meeting that was consumed by assembling tactics can spend more time examining frames. The architecture review can move from modules to systems without losing the module-level evidence. The medical consultation can consider a patient's life trajectory while licensed clinicians remain responsible for the diagnosis and care. The altitude shifts only when lower-level work is dependable enough to support it. Abstraction without contact with the underlying facts is not up-leveling. It is distance.

Up-leveled cognition can synthesize across domains. A specialist can use AI to locate vocabulary, surface candidate connections, and enter three or four adjacent fields faster than before. Access is not mastery. The worker still has to distinguish a deep connection from a linguistic resemblance and know when an actual expert is required. AI lowers the cost of crossing a disciplinary border; curiosity, discipline, and verification determine whether the crossing produces insight or confident tourism.

Up-leveled cognition asks better questions. The worker who used to ask questions that produced data is now asking questions that produce insight. The questions are different. They are framed differently. They expect more from themselves and more from the analytic system. The practice of asking better questions is itself a learned skill, and the AI supports the development of the skill by providing immediate feedback on what kinds of questions produce useful responses.

Up-leveled cognition can develop taste. Taste, in this context, is the sense of what good work looks and feels like in a domain: the senior analyst's standard for a credible market analysis, the designer's feel for a resolved interface, the engineer's recognition of an elegant solution. Taste develops through consequential experience, comparison, critique, and feedback on one's own attempts. AI can increase the volume and speed of examples and feedback. It can also flood the learner with polished mediocrity or reinforce its own conventions. Faster exposure develops taste only when the examples are good, the feedback is trustworthy, and a person capable of judging both remains in the loop.

Up-leveled cognition can integrate across longer time horizons. When AI handles some monitoring and comparison, a leader may recover attention for scenarios measured in years

rather than quarters. Nothing about the tool forces that choice. The same systems can accelerate the weekly dashboard and make an organization more reactive than before. A longer horizon comes from governance and leadership; recovered bandwidth merely makes it possible.

Up-leveled cognition designs better experiments. The worker who used to be limited by the cost of trying things can now run many small experiments in parallel. The cost of testing has fallen. The value of being good at experimental design has risen. The worker who develops a practice of designing experiments that produce learning, in addition to producing immediate results, develops a capacity that compounds over years.

> **WHAT THE RESEARCH SHOWS**
>
> *Brynjolfsson, Li, and Raymond studied the staggered introduction of a generative AI assistant among 5,172 customer-support agents at one Fortune 500 software company. In the peer-reviewed 2025 paper, access increased productivity, measured as issues resolved per hour, by fifteen percent on average. The gains were uneven. The lowest-skill group improved by thirty-six percent, while the highest-skill group showed no significant productivity gain; the most experienced and highest-skilled workers saw small speed gains and small declines in quality. Agents with two months of tenure and AI access performed as well as or better than agents with more than six months of tenure without it. The authors also found suggestive evidence of learning that persisted during brief system outages. This is important evidence that a well-fitted tool can diffuse some practices of stronger workers. It is one tool, one occupation, and one firm, not a universal law of up-leveling.*
>
> Brynjolfsson, E., Li, D., and Raymond, L. (2025). "Generative AI at Work." The Quarterly Journal of Economics, 140(2), 889–942.

The Honest Counterweights

Care is warranted here, because the aspiration is real and the outcome is not automatic. Many workers may not experience the up-leveling just described unless the work is designed to support it. The reasons are several, and worth naming so the aspiration does not become rhetorical.

Effort feels optional. The path of least resistance with AI is often to use it for immediate output rather than development. The worker who finishes existing work faster gets a strong signal that the tool is helpful. The worker who uses it to test and develop their own capacities receives a

weaker, slower signal. Under deadline pressure, choosing the immediate gain is rational. Repeated without reflection, that choice may produce the wrong outcome over years.

The signals can be slow. Some workflow gains appear in days; durable changes in judgment and independent capability may take months or years of deliberate practice. That patience is not the kind most working environments support. Quarterly performance reviews reward output more readily than cognitive development. On the timescales the company usually measures, the worker who is developing may look much like the worker who is merely producing faster. The capacity to invest in development that does not pay off in the current period is scarce in workers and scarcer in the systems that evaluate them.

Up-leveling interacts with the baseline. Expertise helps a worker frame a problem, detect a bad answer, and recognize what matters. But the worker without that expertise does not have "nothing to extend." The customer-support study above found the largest measured gains among less-experienced and lower-skilled workers because the system exposed them to practices they had not yet learned. Other tasks may favor experts who can direct and evaluate the tool. The distributional effect is therefore not fixed: AI can compress a performance gap, widen it, or shift it somewhere else. Tool quality, task design, access, training, and the value placed on the expertise used to build the system all matter.

Some cognition does not transfer cleanly. Not every capacity is equally amenable to AI extension. Some work is deeply embodied, relational, or bound to local experience. AI may still inform it, but information is not the same as practiced judgment. Up-leveling is asymmetric, and the asymmetries matter for how the operating model is designed.

Learning takes time. Some performance gains appear quickly; durable changes in knowledge, judgment, and independent capability require repeated practice and evidence that the learning transfers when the tool is absent. Two weeks may be enough to test a workflow and nowhere near enough to claim cognitive transformation. The relevant clock depends on the skill. Organizations need to measure progress over an appropriate interval rather than turn "brain plasticity" into a scientific-sounding promise.

This outcome is not automatic. We do not yet have evidence for what share of the workforce will achieve genuine cognitive up-leveling. Many workers will understandably use AI for productivity. Some will also develop; some may become dependent; some will do both in different tasks. The workers most likely to develop are those in environments that support

practice, recognize learning, test unaided capability where it matters, and prize the long arc alongside the immediate quarter.

The honest counterweights are reasons the aspiration requires deliberate work to realize. None of them argues against it. The companies that build the conditions for up-leveling will produce more up-leveled workers than the companies that do not. The conditions are namable, and they follow. Each is a company decision, and the absence of any one quietly caps the outcome.

The Conditions That Produce Up-Leveling

Time and space. Up-leveling requires sustained engagement with work that is at the edge of what the worker can do. The engagement cannot happen in fifteen-minute slices between meetings. It requires blocks of time, in environments that support concentration, with the expectation that the time will produce learning rather than immediate output. A calendar scheduled to the minute is a decision against up-leveling, whether or not anyone made the decision deliberately.

Expectation of stretch. The work has to reach beyond routine without becoming unsafe or impossible. Too easy and development stalls. Too hard, and the worker may disengage or produce errors they cannot detect. Calibration is shared work: the manager sets consequential boundaries and understands the role; the worker reports where challenge becomes confusion; mentors and domain experts help test the result. AI can propose the next rung, but it should not be the sole judge of whether the rung is sound.

Mentorship. Up-leveling is often faster and more reliable when there is a more capable human who can support the worker through the zone. The mentor is a companion in the work rather than an instructor delivering content: someone who can recognize an unnamed difficulty through shared context and relationship, pose a challenge that matters in the actual work, and model the higher-altitude practice toward which the worker is reaching. Companies that maintain mentorship cultures, even as AI takes on more explicit knowledge transfer, preserve a developmental engine. Companies that let mentorship erode in the name of efficiency should not assume a tool will recreate it.

Protected modes of work. Some kinds of cognitive work benefit from not using AI at all. Section 5.4 named them in the practice of holding AI-free time, and the list is the same. Protecting these modes is a company decision, not only an individual discipline. Where a role still requires an unaugmented capacity, that capacity remains part of the foundation for supervising and

extending the augmented work. A culture that automates everything in the name of efficiency can erode the foundation without anyone deciding to erode it.

AI as Socratic partner, not only answer machine. The way the AI is used matters. An answer can teach, and a question can mislead; the distinction is not absolute. But a system used only to supply finished output gives the worker fewer occasions to explain, retrieve, defend, and revise. A system asked to surface alternatives, challenge assumptions, request a rationale, and then show its own answer can create more of those occasions. Configuration, prompts, task design, and culture shape which posture predominates. The developmental posture has to be designed and tested deliberately.

The right measurements. Measures direct attention, sometimes usefully and sometimes perversely. Companies that reward only output volume and immediate quality should expect people to optimize for those outcomes. Capability development, contribution to others' learning, long-arc work quality, and unaided performance on critical tasks provide a fuller view. Those measures can also be gamed, so they should inform judgment rather than masquerade as a perfect score. Most performance systems were not designed for work divided among humans and AI. Redesigning them is part of redesigning the work.

The long arc. Durable up-leveling is an ongoing practice, often measured across years even when early gains appear quickly. The leadership team has to be willing to invest in development that does not produce visible returns in every quarter. That willingness is scarce. A leadership team that funds development only when it pays off inside the quarter gets quarter-to-quarter results, and the capacity that might have compounded over the longer cycle never gets built.

The Strategic Case

The strategic argument for taking up-leveling seriously is that the companies that develop up-leveled workforces will compound advantage in ways that the companies that do not develop them cannot match.

The up-leveled workforce produces distinctive work, because the workers' judgment, taste, and craft have been amplified by AI rather than replaced. The work is harder to replicate. The competitive advantage lives in the workforce.

The up-leveled workforce develops new capabilities that the company can deploy as new market opportunities arise. The flexibility is operational, because the workers are capable of being deployed against problems they have not seen before.

An up-leveled workforce may retain better. Development is one reason people stay, especially when the learning is portable and the work remains meaningful. It competes with pay, leadership, workload, opportunity, and life outside the company. The strategic case is not a guaranteed retention effect. It is that an environment worth learning in gives capable people one more substantial reason to remain.

The up-leveled workforce produces leaders. The workers who have developed deep capability are the workers who can take on larger responsibility. The leadership pipeline that the company depends on for the future is built through the practices of the present.

Over a five-to-ten-year horizon, the difference between a company that develops its workforce and one that uses AI only to extract more output could be substantial. The first has something capable of compounding: human judgment, stronger practice, reusable knowledge, and better tools reinforcing one another. The second may still post short-term gains, but it is spending the very capability it will later need. The AI investment matters. The more important investment is the system of people, practices, and technology that determines what the AI makes possible.

The next four sections of Stage 5 address the development side of this work. Passion as the source condition for up-leveling, and the brain's need for challenge underneath it. The hidden expertise that AI is going to surface. What must be unlearned for the AI-Native company to function. And the workforce that all of it produces, humans and agents together.

Section 5.6

Passion as the Engine

> *What makes someone care about their work, and what does their brain require of it?*

Section 5.5 described the conditions under which up-leveling happens. Those conditions are real, and on their own they are not enough. Give a worker time and space, a manager who can spot the right stretch assignment, a culture that protects deep work, and measurement systems that reward development, and the up-leveling can still fail to arrive. All of it supports a process that has to start somewhere, and the somewhere is inside the worker. They have to be reaching for something. They have to care about getting better at this. Without that, you have inputs and no engine.

I want to name this plainly, because it gets discussed less than almost anything else that matters this much. Up-leveling runs partly on caring. Curiosity and a desire to improve make developmental uses of AI more likely. They do not operate alone. A worker can care deeply and still be trapped by workload, unsafe management, poor tools, illness, or responsibilities outside work. Another can begin with little attachment and discover interest through the work itself. Caring is an engine, not a character test. The company still has to build a road on which it can run.

The Brain Needs Challenge

Underneath the caring argument is a feature of human development that the corporate conversation often ignores. People can benefit from cognitive challenge: work near the edge of current skill can build competence, sustain attention, and create the satisfaction of progress. Too little challenge can produce boredom. Too much can produce stress, error, or withdrawal. The useful target is not permanent strain. It is meaningful demand matched with support, recovery, and agency.

Corporate conversation often treats cognitive content as a matter of personal preference. Preference does matter. People differ, routine can be satisfying, and no employer owns a worker's development or purpose. But work design is not neutral. A company deciding which tasks to automate also decides what people will practice for thousands of hours. It should examine whether the remaining work offers variety, judgment, learning, human contact, and a reasonable degree of control, and ask workers, rather than infer from a theory, what makes the work sustainable.

The “use it or lose it” hypothesis has a long history in research on cognitive aging. Studies associate sustained cognitive, social, and physical activity with better cognitive trajectories, and training can improve practiced abilities. The causal story is harder. People with stronger health and cognition may select more demanding activities; education, income, sleep, disease, exercise, isolation, and stress also shape the outcome. Practice is specific, and improvement on one task does not guarantee broad protection against decline. The honest reading is neither “challenge prevents aging” nor “activity makes no difference.” Engagement and capacity can reinforce each other inside a much larger system of health and circumstance.

Retirement research illustrates both the signal and the uncertainty. Longitudinal studies and natural-experiment methods often associate retirement with lower cognitive functioning or faster decline in some abilities, especially memory, but findings vary by cognitive domain, job,

reason for retirement, and what replaces paid work. Retirement itself is not cognitive abandonment. Caregiving, volunteering, study, craft, exercise, and social life can all be demanding. Physical-health findings are similarly entangled with health-driven retirement, income, stress, movement, and social connection. The evidence supports one conclusion: losing a major source of structure and engagement may matter. Harm to the brain or body from leaving a job is unproven.

A related literature examines purpose. People who report a stronger and more sustained sense of purpose often show better cognitive and health trajectories. These are important associations, not proof that purpose itself is the sole cause. Purpose can organize attention and behavior; health and social circumstance can also make purpose easier to sustain. Paid work is one possible source. Family, faith, community, art, service, learning, and care are others. A company can make work more worthy of care. It should not claim jurisdiction over the meaning of a person's life.

The implication for working life is narrower and more defensible. Work is one of the places where many adults encounter challenge, structure, social contact, and opportunities to build competence. Because it occupies so much time, its cognitive content deserves design attention. It is not irreplaceable, and it is not medicine. It is a consequential environment that employers have the power, and therefore the obligation, to shape with care.

The Welder

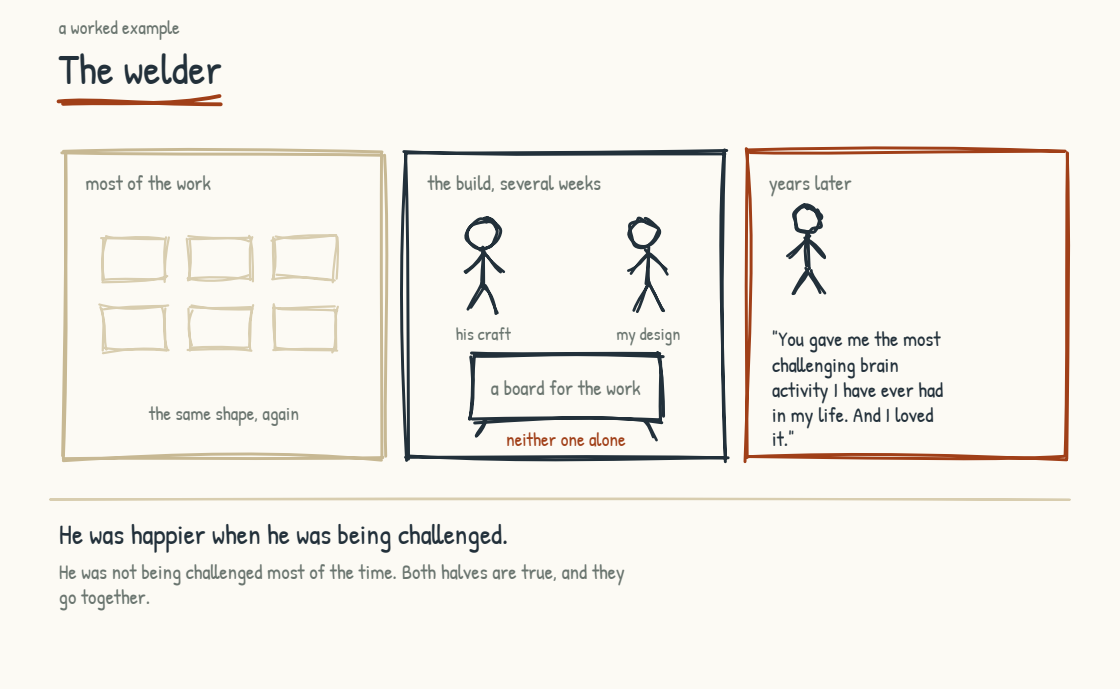

Figure 5.6a. The welder. Challenge was the exception to the ordinary condition of his working life, and his gratitude is the measure of what the rest of it had not been giving him.

A worked example will help here, because the argument is otherwise abstract.

Years ago, I was designing and leading workshops for senior leaders. The workshops used physical artifacts. Large marker boards, custom-built, that could hold the kind of work the senior teams were doing. The boards had to be substantial. They had to be designed for the specific shape of the work. I needed someone who could build them.

I found a welder. He was a working welder. He had spent his career on metal. He was not in the workshop business. He had not designed anything like what I was describing. I had the design. He had the craft. For several weeks that January we worked side by side in his shop, a metal building with a concrete floor, a space heater in one corner, and steel stock racked along the back wall. He ran the welds and I held the pieces square, and when the grinder was going we gave up on talking and pointed. By the last week my drawings carried chalk dust and bootprints. The first set of boards stood finished by the door.

Years later, I ran into the welder again at the counter of a welding supply house, both of us waiting on orders. We caught up for a minute on the ordinary things. Unprompted, he said

something I have not forgotten. He said: I want to thank you for including me in that work. You gave me the most challenging brain activity I have ever had in my life. And I loved it.

I was not expecting this. The welder was a regular working man. He had not described himself as someone who was reaching for cognitive challenge. He had said yes to a job and done it. What he had discovered, in the doing, was that the work had engaged him in a way that his regular welding work did not. The money was there. The relationship was there. His thanks went to the challenge itself.

He was, in the way of someone who does not usually talk about himself this way, telling the truth about what his brain had wanted, and what doing work that demanded that thing had given him. There was no request in it for a more interesting job, and none of the language of meaningful work or employee engagement. He had loved the challenge. The challenge had been a gift.

It was a gift in part because most of his work was nothing like it. Challenge was rare for him, the exception to the ordinary condition of his working life. A leadership team that hears this story should hold both halves. The welder was happier when he was being challenged. The welder was not being challenged most of the time. The two halves go together. The welder's gratitude is the measure of what most of his work had not been giving him.

INTELLECTUAL LINEAGE

The welder story can be read through Edward Deci and Richard Ryan's Self-Determination Theory. The theory proposes three basic psychological needs: autonomy, the experience of volition; competence, the experience of effectiveness; and relatedness, the experience of connection and belonging. Research in workplaces links support for these needs with more autonomous motivation and a range of favorable outcomes, with meaningful variation across people and settings. Need satisfaction does not switch intrinsic motivation on, need frustration does not make motivation collapse in every case, and compensation still matters. The boards project appears to have supported all three: competence at the edge of his craft, autonomy in shaping the build, and relatedness through working side by side. That is an interpretation of why the experience stayed with him, not a diagnosis made from one conversation. Csikszentmihalyi's work on flow overlaps with this account, but the theories describe related rather than identical phenomena.

Deci, E. L., and Ryan, R. M. (1985). Intrinsic Motivation and Self-Determination in Human Behavior. Plenum. / Ryan, R. M., and Deci, E. L. (2017). Self-Determination Theory. Guilford Press.

Two Kinds of Passion

Passion research offers a distinction every leadership team should know, because it separates the passion worth cultivating from the passion that burns people down. Robert Vallerand and his colleagues, across two decades of studies, distinguish harmonious passion from obsessive passion. Harmonious passion is caring the person controls. The work matters to them, they chose it freely, and it sits in balance with the rest of their life. Obsessive passion is caring that controls the person. The work has fused with their identity or their standing, they cannot put it down, and the caring produces rigidity and burnout alongside the effort.

The distinction matters because the social conditions around work can encourage more autonomous or more controlled forms of engagement. Real connection, recognition, autonomy, and membership may support harmonious passion. Recognition tied to self-worth, pressure without safety, and membership that has to be re-earned can feed obsessive patterns. They do not mechanically manufacture one form or the other; identity, personality, history, and life outside work matter too. A leadership team that demands visible engagement can easily reward obsessive behavior because it resembles commitment, until the costs appear.

The welder's account has the qualities associated with harmonious passion. He chose the work, it challenged him at the edge of his craft, and years later he described what it had added to his life. I cannot infer his inner state from a brief encounter. I can recognize the target: demanding work held in proportion to a whole life.

INTELLECTUAL LINEAGE

The harmonious/obsessive distinction comes from Robert Vallerand's Dualistic Model of Passion, developed with colleagues at the Université du Québec à Montréal and studied across work, sport, education, and the arts. A 2015 meta-analysis synthesized 94 studies and found meaningfully different patterns of association: harmonious passion was generally linked with more adaptive wellbeing, motivation, and cognitive outcomes, while obsessive passion showed a more mixed and often maladaptive pattern. These are tendencies, not destinies or simple causal predictions. Autonomy-supportive environments can encourage autonomous internalization; controlling and contingent environments can encourage controlled internalization. The person and the environment both matter.

Vallerand, R. J., et al. (2003). "Les Passions de l'Âme: On Obsessive and Harmonious Passion." Journal of Personality and Social Psychology, 85(4), 756-767. / Vallerand, R. J. (2015). The Psychology of Passion: A Dualistic Model. Oxford University Press.

The Manager's Job Has Changed

The first implication is that AI changes the manager's role, although it does not erase what came before. Managers have always coordinated work, allocated resources, resolved blockers, developed people, exercised judgment, and represented the team in systems of power. Different organizations weighted those duties differently. AI can change that weighting and expose managers whose contribution was little more than moving information between meetings.

AI can draft plans and status reports, summarize communication, and surface candidate dependencies. Plans do not maintain themselves. A system cannot reliably decide that a reported status is true, that an unmentioned dependency matters, or that a team needs protection rather than another reminder. Coordination survives, but some of its clerical burden can shrink. The time released creates a choice about what management becomes.

One answer is engagement architecture. The manager learns enough about each worker to understand what they are trying to develop without demanding access to their private identity. They design assignments that connect business need with appropriate stretch, notice withdrawal without treating ordinary boundaries as disengagement, and address the conditions within the company's control. The goal is not to manufacture passion. It is to stop extinguishing it and make room for it to grow.

This is different work from administrative coordination, and often harder. Many management programs underweight work design, coaching, psychological safety, and the responsible use of AI. Companies need to select, train, support, and evaluate managers for the full role. “Retrain or replace” is too blunt when the organization itself promoted yesterday's behavior and still rewards it today. Change the system, give people a fair chance to grow, and make role decisions on evidence.

What Cultivates Passion

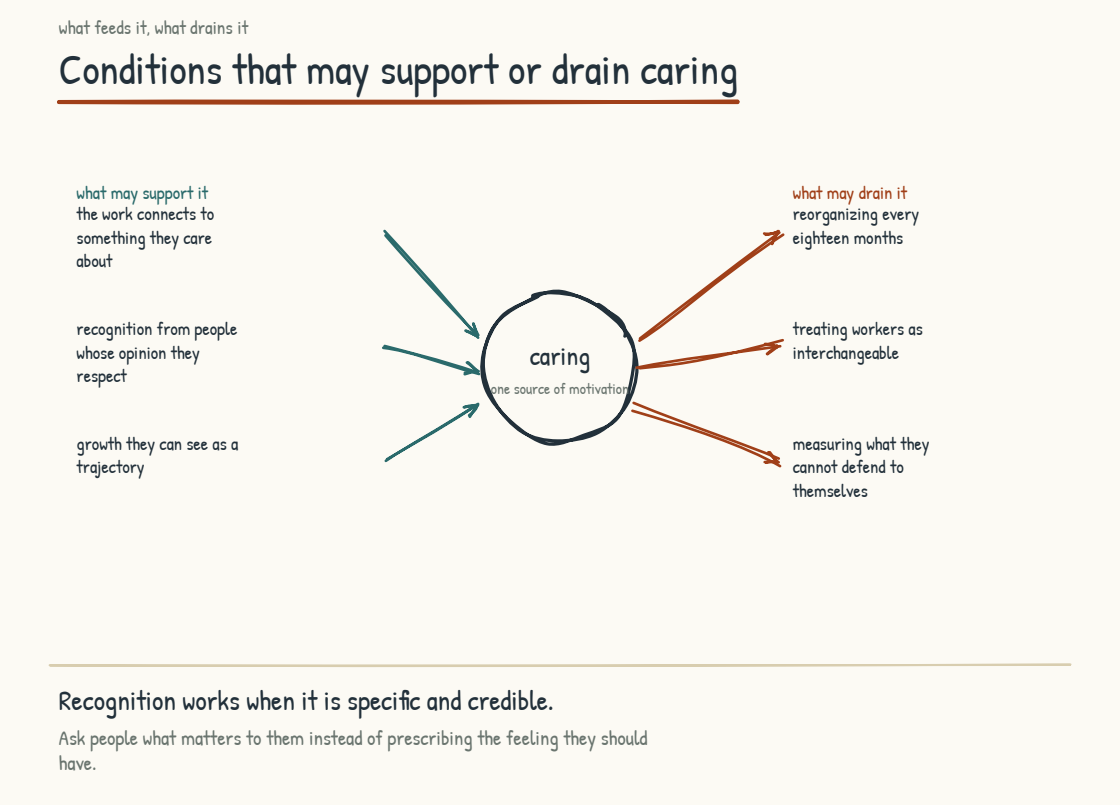

Figure 5.6b. What feeds the engine and what drains it. The reverse list matters more, because those practices are common and rarely recognized as destructive.

INTELLECTUAL LINEAGE

What I am describing is related to what Mihaly Csikszentmihalyi called flow: deep absorption in an activity, often accompanied by concentration, a sense of control, and intrinsic reward. His account emphasizes conditions such as clear goals, timely feedback, and a workable match between challenge and skill. Later research has developed, tested, and debated how flow is defined and measured. Flow is not synonymous with caring, and it is not the only desirable way to work; collaboration, reflection, care, and routine competence may feel different. The overlap is still useful. Work is more likely to absorb people when they can understand the goal, perceive progress, and meet a challenge that is neither trivial nor crushing.

Csikszentmihalyi, Mihaly (1990). Flow: The Psychology of Optimal Experience. Harper & Row.

Passion is a strong word. I use it deliberately, and with care. I mean the more workable version rather than the romanticized version that requires every worker to be in love with their job. The

worker genuinely cares about getting better at what they do. They see the work as connected to something they value. They bring their full attention to it. This kind of caring is more common than the romantic version, and it is the kind on which the AI-Native company depends.

A handful of namable factors cultivate it.

The first is connection between the work and something that matters to the worker. Seeing how the work affects a customer, colleague, community, craft, or personal goal can provide a reason to persist through difficulty. Workers may have other reasons, including economic security, and those reasons deserve respect. The leadership team can make real consequences visible. It cannot fabricate meaning or decide on a worker's behalf what should matter.

The second is recognition of contribution. The worker who can see that what they do is noticed by people whose opinion they respect has a reason to care about the work. The recognition has to be real. Recognition theater, the kind that comes with a corporate certificate signed by a senior leader who does not know the recipient, is worse than no recognition because it tells the workforce that the company is not paying attention.

The third is growth as a visible trajectory. The worker who can see that they are getting better at the work, and that the company is invested in their getting better, has a reason to care. The growth has to be real, and the investment has to be visible. The companies that talk about development without investing in it produce the cynicism that destroys caring.

The fourth is autonomy within the work. The worker who has genuine latitude to make decisions about how the work gets done has more reason to care than the worker who is executing a script. The latitude has to come with the accountability that turns autonomy into ownership. Autonomy without accountability is neglect. Accountability without autonomy is coercion. The combination is the condition that produces caring.

The fifth is membership in something that matters. The worker who feels part of a team, craft, tradition, or mission has a reason to care that does not depend on individual outcomes. Membership is built over time through shared experience, fair treatment, and reciprocal obligation. Constant reorganization or repeated headcount reduction can fracture it. Sometimes restructuring is necessary. Treating the social cost as zero guarantees that the financial case is incomplete.

What Destroys Passion

The reverse list is equally important. It is worth reading slowly, because these practices are common and rarely recognized as destructive.

Constant reorganization can destroy passion. Workers hesitate to invest in relationships, systems, and missions they expect to lose. At some companies the org chart changes more often than the logo. A reorganization may have sound strategic logic, and avoiding one can also harm people. The mistake is treating the loss of trust, local knowledge, and membership as an externality rather than a cost to be anticipated and repaired.

Treating workers as fungible can destroy passion. A company needs roles, budgets, and headcount models; it also needs to remember that the people inside those abstractions are specific. Workers who repeatedly see contribution ignored and relationships treated as disposable have less reason to reciprocate commitment. The effect will not be identical in every person. The breach is still real.

Measuring the wrong things can destroy passion. Workers measured on outputs that do not correspond to what they believe good work requires can find themselves doing work they cannot defend to themselves. Measurement systems shape what gets done. Poorly chosen measures can train people away from the quality, judgment, or care the company still claims to value.

Punishing visible engagement can extinguish it unusually fast. The worker who reaches beyond their assigned scope, suggests a better way of doing the work, or flags a problem no one asked them to flag, and is punished for reaching, learns quickly that silence is safer. The lesson can generalize. The worker may stop reaching, and the company may later complain about the disengagement it helped produce. Much of this punishment is accidental, delivered through managers who feel threatened or systems that reward conformity. Intent changes the moral judgment; it does not erase the effect the worker experiences.

The Respected Place for Non-Passion-Driven Engagement

There is a version of this argument that produces an unhealthy expectation that every worker must be passionate, must be reaching, must be using AI to up-level themselves continuously. The expectation, taken seriously, produces a workforce in which workers feel they have to perform passion they do not feel, and the performance is more destructive than the absence of passion. The performed version is obsessive passion by another name. It costs the worker and returns nothing durable to the company.

I want to name the alternative. There is a respected place in an AI-Native company for the worker whose engagement is not passion-driven. The worker who is in their job for the paycheck, who does the work well, who goes home to a life that is centered elsewhere, who treats the company fairly and expects to be treated fairly in return. This worker is doing nothing wrong. They are doing a kind of work that has been honored in most cultures throughout history. The company that requires every worker to be passionate has confused itself about what it can ask of people.

The honest version of the argument is that up-leveling will be unevenly distributed, and we do not yet know the proportions. People will want different things in different seasons of life. The company benefits from supporting workers who are reaching and from treating workers who are not seeking continuous development as full members of the company. Both are doing real work. The company that hollows out the second category in favor of the first will discover that much of the work keeping the company running lived there.

What This Implies for AI-Native Work Design

The AI-Native company that takes the welder's experience seriously will design work differently than the AI-Native company that does not. The first company treats the cognitive content of work as a primary property, and designs it deliberately.

This is harder than it sounds because AI can remove both drudgery and development. Repetition sometimes builds fluency; difficulty sometimes comes from pointless friction. The temptation is to automate whatever looks hard without asking which part teaches the worker, reveals the state of the system, or keeps human judgment calibrated. The result can leave people with passive monitoring, fragmented exception handling, or responsibility for outcomes they no longer have enough contact with the work to understand.

The alternative is to design AI into the work so that human demand becomes more valuable, not simply higher. Routine steps may be automated while people frame goals, inspect evidence, handle exceptions, make accountable decisions, build relationships, and sometimes practice the underlying task without assistance. AI capabilities will continue to move, so “work the AI cannot do” is a brittle boundary. The durable boundary is work for which the organization chooses to require human authority, participation, or capability.

This is the up-leveling argument from Section 5.5, viewed from the angle of human flourishing rather than competitive advantage. The two angles converge on the same operating model.

Productivity Is Not the Whole Question

I want to push back, gently and directly, on a piece of the corporate vocabulary that is making this work harder.

Productivity is the word the corporate world has used for decades to describe the purpose of work. Productivity measures output per unit of input. More productivity is better. The vocabulary is so familiar that most leadership teams cannot easily think about work in any other terms.

The productivity vocabulary is insufficient for the AI-Native era because a ratio of output to input does not, by itself, reveal what the work is doing to workers or to future capability. A team can produce more this year while weakening apprenticeship, judgment, trust, or resilience. That does not prove cognitive loss, and the future cost may never materialize. It does mean that a quarterly productivity number is a snapshot rather than the whole trajectory.

A more useful frame is to ask what the work is producing in two places. First, what the work is producing for the company. Second, what the work is doing to the worker. The two have to be considered together, because the company that hollows out its workforce in pursuit of quarterly output will discover that the workforce it has produced cannot deliver the output it needs in the years that follow.

The leadership team that takes this frame seriously will measure workforce trajectory alongside output. Can people explain and perform critical work when the tool is unavailable. Are novices progressing. Are errors being caught earlier. Do workers report useful challenge, agency, and sustainable load. Is expertise being shared, credited, and renewed. No single measure answers whether people are becoming “sharper,” and surveillance can damage the very conditions being measured. Use capability tests, operational outcomes, voluntary and privacy-protected feedback, and informed human judgment together.

Why AI Amplifies Whatever Culture It Lands In

One last point about passion: AI can amplify important features of the culture that receives it. A culture that rewards curiosity and candor may use AI to test more ideas and expose weak reasoning. A culture that rewards speed and punishes dissent may use the same tool to produce more work with less challenge. Technology can also alter culture rather than merely reflect it. Same tool does not mean same implementation, incentives, data, authority, or output.

This is why the human work matters strategically, not just ethically. Returns from AI depend in part on earlier investments in expertise, trust, process, data, and management. A company that invested in its people has more capability for AI to extend. A company that treated people only as inputs may still obtain productivity gains, but it carries less trust and developmental capacity into the transition. Neither outcome is guaranteed; both are starting conditions leaders can see and change.

The leadership team thinking about the AI investment should also examine the human investment it postponed. The two interact. Tools shape the work people can do; people determine where the tools are trusted, challenged, and applied. The operating system has to hold both.

The Welder's Standard

A close for this section: what I have come to call the welder's standard.

The standard is the question that every leadership team designing an AI-Native operating system should be willing to be asked. If a worker in this company, years from now, ran into you at a baseball game, what would they say to you about the work they did here. Would they say, you gave me the most challenging brain activity I have ever had in my life, and I loved it. Would they say something quieter that meant approximately the same thing. Or would they say something else.

A leadership team that can answer this question with the welder's words has evidence that at least some work is good for the people doing it. It should also listen for different honorable answers: the work was fair; I became capable; my boundaries were respected; I supported my family; I belonged; I left with my health intact. No executive should answer for the workforce. The standard is an invitation to ask, and to be willing to hear what comes back.

The welder's standard is the counterpart to the operational rigor built across the previous sections. Both are necessary. Neither is sufficient on its own.

The next section turns to one of the more hopeful consequences of taking this argument seriously, which is the expansion of who gets to be skilled in a domain at all.

Section 5.7

The Hidden Expertise: AI as Expander of Who Gets to Be Skilled

Who becomes skilled in a domain when AI lowers the barrier to entry?

Displacement is the dominant story about AI and work. Workers who used to be valuable are becoming less so. Workers who learned skills that took years to develop are watching those skills become less rare. The story is partly true. It is also incomplete, and the incompleteness is one of the more interesting things about the AI-Native era.

What that story misses is the other end of the skill distribution. Some people never got to explore domains for which they may have had real aptitude because a prerequisite, teaching method, cost, disability, credential, or early judgment closed the path. The aspiring doctor stopped by organic chemistry. The aspiring engineer stopped by calculus. The architect who could think spatially but could not draw to the expected standard. The would-be developer defeated by bad documentation and unhelpful error messages before building anything that mattered to them. Some filters protect the public or test knowledge the work genuinely requires. Others are rough proxies, and even necessary prerequisites can be taught or supported in better ways. The world has almost certainly lost potential expertise because it mistook the ability to pass a gate for the full ability to do the work beyond it.

AI can change the geometry of some of this filtering. It does not make every gate obsolete.

The Gatekeeping Skills Are Not the Work

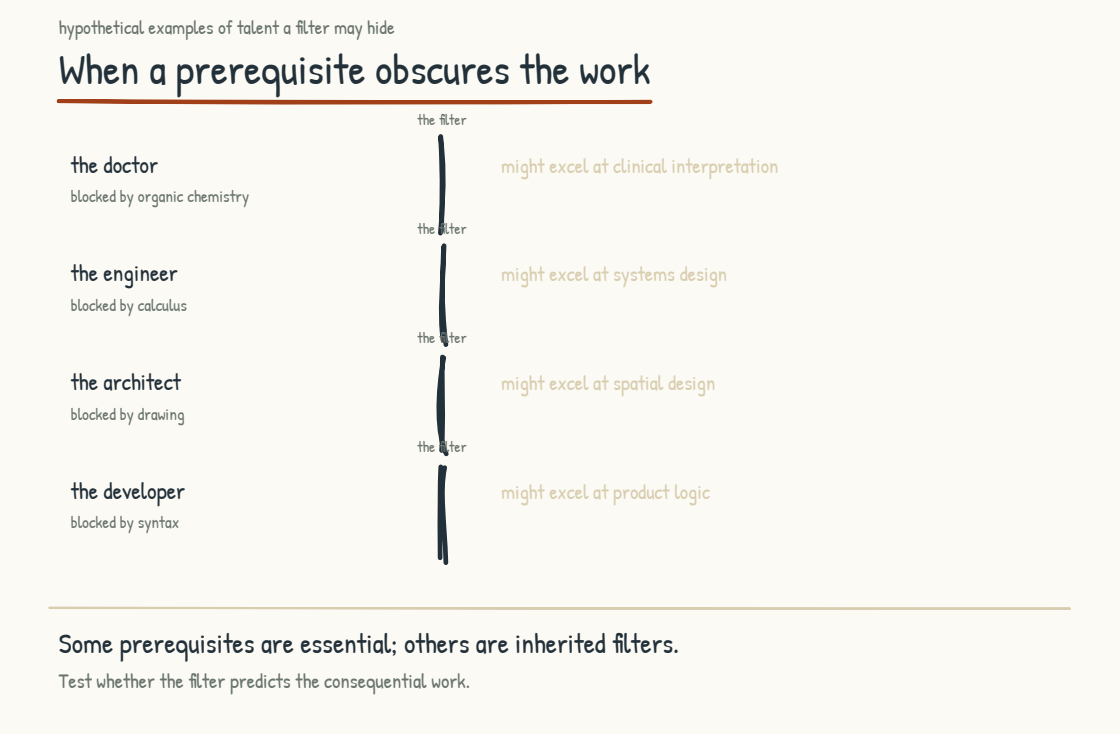

Figure 5.7a. A prerequisite can test essential capability, act as a rough proxy, or do both. Change the route into the work and the geometry of who gets a fair chance to become skilled can change with it.

Notice first that a prerequisite is not always identical to the work it precedes. Educational and professional systems use prerequisites to establish foundations, ration limited places, make evaluation manageable, and protect the public. Some directly test essential capability. Some are historically convenient proxies. Many do both imperfectly. The important question is not whether a gate is frustrating. It is whether the capability it tests remains necessary, whether AI changes how that capability can be learned or exercised, and what evidence should replace the old test if the gate moves.

Engineering makes the distinction visible. Engineers design systems that solve real problems. The work involves mathematics alongside judgment about tradeoffs, communication with stakeholders, integration across domains, persistence through iteration, and responsibility for safety and failure. Mathematics is a tool and part of the professional foundation; in many branches, a person must understand it well enough to select a model, recognize a broken assumption, and know when an answer is physically impossible. AI can perform calculations, visualize relationships, explain abstractions, and provide practice. It may change how

mathematical competence is developed and demonstrated. It does not make mathematical understanding irrelevant wherever the system's behavior depends on it.

A high school student who is told that they are not good at math, and who therefore should not consider engineering, is being filtered by the mathematical prerequisite. The filtering may or may not be accurate. Some students who fail the filter would also have failed at the work. Some students who fail the filter would have been extraordinary at the work if they had ever been allowed to get to it. The world has no way to tell which is which, because the filter happens before the work begins.

I have a son who experienced this filtering. He was told in high school that he was not good at math and that engineering was probably not for him. He has a passion for understanding how things work and fit together, one important engineering disposition, though not the whole profession. A judgment about his performance in one setting became a judgment about the path. AI does not answer whether he should be an engineer. It may give him a better way to find out: patient explanation, manipulable models, immediate practice, and multiple representations of the same concept. The mathematics does not disappear. The route into it becomes wider than one classroom, one pace, or one teacher's verdict.

The implication is that some people who could not cross an old gate may now be able to build the required capability and demonstrate it another way. The number is not zero, and it is not everyone. In some roles the threshold may fall because AI safely handles part of the task. In others the threshold should remain or even rise because AI output creates new verification demands. The geometry is changing, but it has to be redrawn task by task.

The Pattern Generalizes

The same pattern appears across domains, most visibly in software. Development used to require enough fluency in programming languages, error messages, tooling, and documentation to reach a working prototype. Those demands filtered out people who might have designed useful software. Now people can build applications by describing behavior in natural language and iterating with AI. Producing something that runs is real development, but it is not proof that the system is secure, maintainable, accessible, or correct. The barrier to creation has fallen faster than the burden of responsibility. That is expansion with a new obligation attached.

I am one of these people. I was not a programmer. My cofounder set me up with the right tools and the right environment. Visual Studio Code. Claude Code. The integration with version

control. I started by asking the AI questions in natural language. I dabbled in editing what the AI produced. Over time, my sense of what was possible changed. I began saying, what if the application worked like this. I worked with the AI to bring it to life. The work compounded. What started as basic interactions evolved into something I would not have predicted I could do.

This story is not exotic. Versions of it are appearing in other domains: a medical researcher using AI to write and inspect bioinformatics code, a journalist entering data analysis with guided statistical support, a teacher turning classroom knowledge into usable curriculum materials. In each case, AI can lower an implementation barrier without conferring the neighboring profession's full expertise. The researcher still needs valid methods, the journalist still needs statistical judgment, and the teacher still needs evidence and quality control. The trajectories are visible. Their scale and long-term effect on the supply of expertise remain open empirical questions.

WHAT THE RESEARCH SHOWS

Wharton Human-AI Research and GBK Collective's 2025 enterprise survey included 801 U.S. business leaders at organizations with at least one thousand employees and $50 million in revenue. Among 372 vice presidents and C-suite leaders asked about hiring over the next two to five years, thirty-nine percent expected more entry-level or junior hiring, forty percent expected no change, and eighteen percent expected less. For interns, forty percent expected more and eighteen percent less. The result is more mixed, and more interesting, than a simple displacement story. It records leaders' expectations, not hiring outcomes, and it does not establish why they answered as they did. It suggests that many leaders can imagine AI creating opportunities for junior workers even while a meaningful minority expects contraction.

Wharton Human-AI Research and GBK Collective (2025). "Accountable Acceleration: Gen AI Fast-Tracks Into the Enterprise." Year Three Full Report, pp. 71–72.

The Same Gate, From the Other Side

Everything above is about people walking in. There is a second thing happening at the same gate, to the people already inside, and a section that tells only the first half is not being straight with them.

The person who spent twenty years becoming the one who knows was holding two things at once. They could do the work. Very few others could. Those arrived together and could feel like a

single thing, though they were never identical. AI may separate them. Some parts of the capability can become easier to access, while other parts (judgment, accountability, tacit discrimination, and experience with failure) remain scarce. The expert's capability is not untouched, because the work itself is changing. The scarcity is not simply gone, because it moves.

That is a real loss and it deserves to be named as one. The skill is intact. What goes is the position, being the first person the room asks. Section 5.2 calls this the grief of expertise, and it sits first on that list because it is the one people carry quietly and almost never say out loud.

I am on both sides of it. I could not program, and now I can, which puts me among the people who came through the gate after it came down. I also spent a career building a practice around designing how teams do their hardest work, and the same force is arriving at that. Writing this section only from the optimistic end would be dishonest about where I am standing in it.

What I would say to the person holding the knowledge is this. One scarce thing was always knowing which of the things you know applies to the situation in front of you. AI affects that too, but it does not remove the need for grounded judgment or responsibility. A company that treats you only as a repository will be disappointed. A company that recognizes your judgment, credits the knowledge you contribute, and gives you a role in redesigning the work has a better chance of discovering what you are worth.

The Bridge Question

One warm Friday night in July I found myself next to an electrician at a minor league baseball game, third row up the first-base line, his company's name stitched over the pocket of his polo. The home half of the second went three up, three down, and in the lull he asked what I did, and I told him I worked in AI. He had not used AI himself. He had not used ChatGPT. He had a clear sense of one of his pain points. He was given blueprints for new projects, and parsing the blueprints to produce accurate bids was slow work. The slowness affected which projects he could pursue. The slowness affected the economics of his business.

I explained, over the course of an inning or two, while the visitors changed pitchers and the grounds crew dragged the infield, that software could extract information from blueprints and help prepare a draft bid using his rules and examples. With careful implementation, evaluation, permissions, and review, a system could support more bids without pretending to know conditions the drawings did not show. He nodded at the right places, eyes on the field. He

understood what I was saying. His mental model still ran on blueprints and spreadsheets. The bridge from where he was to a trustworthy AI-assisted workflow was not yet walkable. The intelligence to cross it was there. The bridge was not.

That bridge question is one of the most consequential in the AI-Native era. Across industries, many workers are sitting where the electrician sat. They have real expertise. They have specific problems that AI could help them solve. They do not yet have the bridge from their current practice to tools that would extend it. The bridge has to be built by someone. The question is who builds it, for whom, and to what standard.

The companies that take this seriously are building the bridges as part of their AI-Native operating models. They are not assuming that workers will figure out AI on their own. They are designing the entry points, the supports, the practice opportunities, and the trusted relationships that allow workers to cross from their existing expertise into the AI-augmented version of that expertise. The companies that do this well will have workforces that surface latent capability the company did not know it had. The companies that do not do this will have workforces in which only the workers who could build their own bridges, the workers who were already at the comfortable end of the digital divide, will participate in the AI-Native transformation. The workers at the other end of the divide will be left out, and the talent they could have contributed will not be recovered.

The Democratization Research

Research on AI-augmented work offers a narrower signal. In several studied tasks, including the customer-support setting in Section 5.5, less-experienced or lower-performing workers gained more than the strongest workers and performance dispersion narrowed. Other studies show a jagged frontier: assistance improves performance on some tasks and harms it on others, with effects shaped by baseline skill and how the tool is used. These studies concern people already doing or selected into the work. They do not show that people previously filtered out by a prerequisite possess hidden professional expertise, and they do not justify removing safety-critical qualifications.

The evidence is preliminary and uneven across domains. It is still strong enough to justify an experiment rather than a conclusion. A company can look for employees whose adjacent experience, aptitude, and interest have been missed; provide AI-supported learning and supervised practice; and evaluate them on real work against the same outcome and safety

standards. The opportunity is access to a broader pool of potential capability. Whether it becomes expertise must be demonstrated.

What This Means for How Companies Hire and Develop

The hidden-expertise argument has direct implications for hiring, development, and the structure of who gets to do what kinds of work in an AI-Native company.

Hiring practices that rely reflexively on prerequisite filters may select for an outdated version of the work. The response is not to assume that someone who struggled with mathematics is therefore a strong engineering hire. It is to decompose the role: which knowledge remains essential, which task is now tool-assisted, which capability can be learned on the job, and which credential is legally or ethically nonnegotiable. Then evaluate candidates on relevant work, learning ability, judgment, and the foundations the role still requires.

Development practices that assume the only valid path is the one that worked in the pre-AI era will keep selecting for a workforce that resembles the pre-AI workforce. Some people who could benefit most from AI-supported development may be precisely those traditional pathways overlooked. The development system should be redesigned to find and support them without presuming that access alone has already produced expertise.

Role design that carries every old prerequisite forward without examination may constrain the company to a smaller workforce than it needs. Roles can sometimes be made accessible to a broader range of workers through better tools, training, apprenticeship, and division of responsibility while holding outcomes to the same or a higher standard. Accessibility and safety are joint design requirements. The payoff is potential capability that competitors may overlook; the proof is performance, not optimism.

The Strategic Case

The strategic case for taking hidden capability seriously is that companies may find a labor pool competitors using inherited proxies cannot see. No evidence says a suitable pool exists in every role or company. The only responsible way to discover it is to open credible pathways, support people through them, and measure the resulting work.

Companies that build those systems may include an aspiring clinician who found a different path into health analysis while respecting the boundaries of licensure, a prospective engineer who learned mathematics through tools that finally made it legible, or the electrician at the baseball game using a carefully tested system to expand his bidding capacity. The potential may

be there. The company's job is to build bridges that develop and test it, not declare expertise before the crossing.

This is one of the genuinely optimistic threads in the AI-Native conversation. It deserves to be held carefully because optimism becomes dangerous when it erases standards or appropriates a profession's name. Hidden capability is real. AI can expand who gets a fair chance to develop skill. Expertise still has to be earned, demonstrated, maintained, and, where other people bear the risk, governed.

The next section addresses the other side of this same question, which is what has to be unlearned for the AI-Native operating model to work at all. The expansion of who can be skilled is the opportunity. The unlearning is the cost.

Section 5.8

What Must Be Unlearned

What habits of mind no longer serve the AI-Native worker?

At a conference last year a panelist said, almost in passing, that anyone over forty working in or building a company around AI has to forget half of what they have learned and half of what they know. I heard the line as provocation, not arithmetic, and not as an indictment of age. A career spent learning how work gets done leaves you holding hard-won judgment about the right sequence, the right amount of caution, and how long a thing takes. New compute and new interfaces are clearing some of the hurdles that shaped those judgments. Experience remains valuable. Some conclusions drawn from its old conditions now need to be tested again.

The line means that some things that used to work for you may work differently now. Timelines that used to take quarters can sometimes be compressed into weeks. Parts of workflows that had to be sequential can sometimes run in parallel. Hurdles that used to slow you down may be cleared with new compute and interfaces. Habits of mind that protected you in the pre-AI world (the carefulness, the staging, the defensive expertise, the long iteration cycles) are not all wrong in the AI-Native context. Some still serve you. Others may be slowing you down after the condition that justified them has changed.

Adopting new tools is part of becoming AI-Native. The rest is deliberately revising patterns built for conditions that may no longer hold. Revising is often harder than adding because it requires

a person to separate the principle that made a practice wise from the mechanism through which it was expressed. The way you have been working may not be the way to keep working. That can be especially difficult when the old way helped produce real success. It can be equally difficult for a novice who has never seen why the old safeguard existed. Unlearning is not forgetting. It is updating without discarding the evidence that made the earlier lesson true.

The Specific Habits Worth Examining

I want to walk through several habits of work that deserve examination. The list is not exhaustive. It is meant to make the abstract argument concrete, so that a leadership team or an individual worker can see what I mean and can locate their own version of the pattern.

Long timelines. The pre-AI world ran on long cycles because the work took long to do. Strategic plans were quarterly or annual. Major projects took years. Software releases were measured in months. The long cycles produced patterns of work that depended on the cycles. Decisions could be deliberated over weeks. Drafts could be reviewed over days. Communication could be staged because the response was not needed immediately.

AI can compress some of those cycles. A first draft, code prototype, comparison, or synthesis that took weeks may take hours. Discovery, consultation, physical work, regulatory review, integration, and trust may not compress with it. Faster generation can even lengthen evaluation by multiplying the number of plausible options. The habit to unlearn is not deliberation. It is assigning every step the duration imposed by the slowest old step. Decide which latency has actually changed, then protect the time the consequence still requires. A late excellent decision can be useless. A fast unexamined decision can be worse.

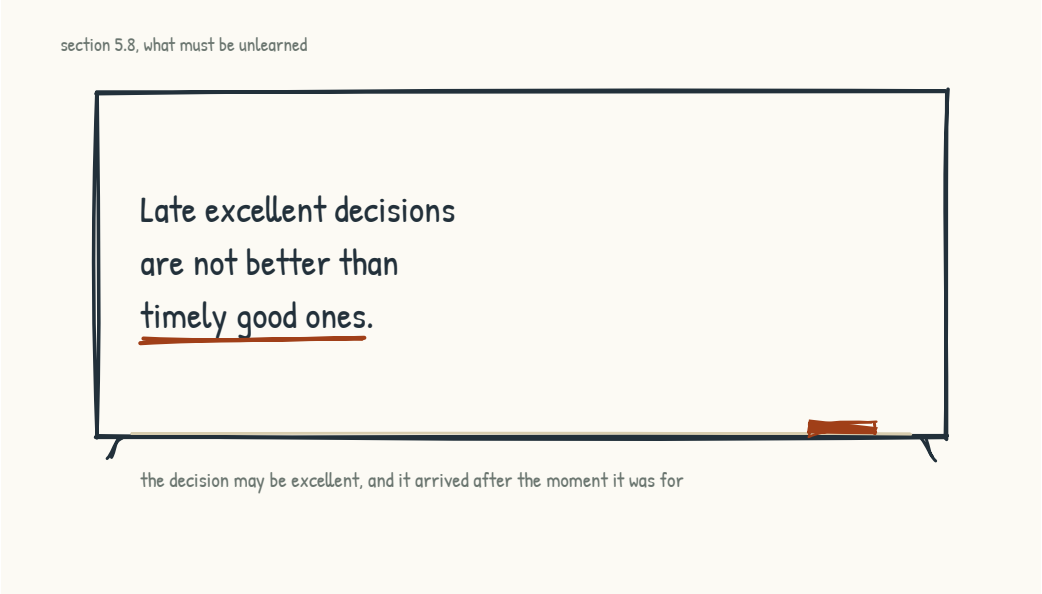

Figure 5.8a. The cost of the old cycle, on the board.

Sequential workflows. The pre-AI world depended on sequential work because parallel work required coordination overhead that often was not worth it. The analyst produced the analysis, then handed it to the strategist, who produced the strategy, then handed it to the planner, who produced the plan. Each handoff was a cost, and the costs accumulated. The sequence existed because the alternatives were worse.

AI allows more parallelism by lowering the cost of drafting, comparison, and synchronization. Analysis, strategy, and planning can sometimes begin together as hypotheses that inform one another. They cannot become independent fictions. Strategy still depends on what the analysis establishes, and a plan built against a discarded strategy must be reworked. The habit to unlearn is unnecessary waiting, not causal sequence. Define the interfaces, assumptions, and points of convergence; run reversible work in parallel; and keep irreversible decisions behind the evidence they require. AI can support integration. It is not the substrate that guarantees it.

Defensive expertise. Many experienced workers learned, over their careers, to develop expertise as a defense. Knowing more than your colleagues made you valuable. Being the person others came to for a specific kind of judgment created a moat around your role. The expertise functioned simultaneously as competence and as protection. Sharing too much eroded the protection. Documenting your work too fully made you replaceable. The incentives produced behavior that was rational at the individual level and corrosive at the organizational level.

AI changes these incentives, but not in one direction. Expertise kept inaccessible cannot improve shared practice or the systems meant to support it. Expertise captured without consent, credit, context, or a credible employment covenant may train the mechanism used to eliminate the contributor's role. Calling reluctance “hoarding” turns an organizational trust problem into a worker defect. The company has to make contribution safer and more valuable than concealment: define purpose and access, recognize authorship, compensate unusual knowledge work, preserve challenge and advancement, and honor legal, contractual, privacy, and collective-bargaining boundaries.

Letting go of defensive expertise asks the worker to believe the company will still value judgment after the explicit knowledge has been shared. That trust is the covenant from Section 5.2, and the company has to earn it through conduct. A covenant does not guarantee participation, knowledge does not automatically become a graph, a graph does not automatically become capability, and capability does not automatically elevate the contributor's work. Each arrow requires design and proof. When the chain works, shared knowledge can improve tools and free experts for more consequential work. When it does not, the worker's caution may be entirely rational.

Status games. Most pre-AI organizations were structured around status games that workers learned to play. The hierarchy was visible. The signals of position were specific. The meetings were structured to display rank. The communication patterns reinforced who was senior to whom. The games were not always healthy, but they were stable, and workers learned to play them with skill.

Some AI-Native companies flatten hierarchy; others centralize new power around models, data, infrastructure, and the people who control them. Less-senior workers can perform some tasks once reserved for seniors, but tool-assisted output is not identical to senior judgment. Old status signals may weaken while new ones form around access, fluency, visibility, and control of automated systems. The games change more reliably than they disappear.

Releasing this requires the worker to find new sources of standing. The standing that comes from doing work that is genuinely distinctive. The standing that comes from developing other workers. The standing that comes from being trustworthy in moments of ambiguity. These are real sources of standing in the AI-Native company. They differ from some of the sources rewarded in the pre-AI company. Workers who make the transition can find new ground. Workers who do not may keep spending energy on games whose prizes are disappearing.

The Discipline of Unlearning

Unlearning can be harder than learning. A new pattern can be added without threatening the old explanation of success. Revising the old pattern requires a person to notice that it was conditional, not universal. The pattern need not be erased. It can remain available for the conditions in which it still works. The demanding act is building a better rule for when to use which approach.

Unlearning is a practiced skill. People become better at it when they notice that a familiar practice has stopped producing its old result, investigate what changed, and test a replacement rather than merely redoubling effort. It is observable and teachable. Companies that reward revision, preserve psychological safety, and keep the results of experiments visible give the workforce a better chance of navigating transition.

The teaching is practical. The manager who notices that a worker is operating on a pre-AI pattern that no longer fits, and who works with the worker to examine the pattern, name what it was doing, and design a replacement, is teaching the discipline. The worker who learns from the manager learns to do it on their own, and then teaches others. The discipline propagates through the workforce as a culture, not as a training module.

Why This Is Hardest for the Most Accomplished

I want to name something uncomfortable. Unlearning can be especially hard for people who were highly accomplished under the old conditions: the senior partner whose practice was built on a particular judgment, the chief operating officer whose career was built on a particular execution discipline, the engineer whose reputation rests on a particular design approach. They have strong evidence that the patterns worked and real identity invested in them. They also possess the experience needed to distinguish a genuine change from recycled hype. The point is not that accomplished people have the most to forget. It is that their revision carries unusually high stakes and unusually high value.

The accomplished worker may see intellectually that the patterns of work are changing and still hesitate to release them. The resistance can be emotional, practical, and economic. The patterns are part of how the worker understands themselves, but they may also carry reputation, income, authority, and real protection against error. Revision can therefore feel like releasing a piece of identity built over decades while betting that the new conditions will hold.

The leadership teams that recognize this and design for it produce different outcomes than the leadership teams that do not. The recognition involves giving accomplished workers explicit permission to experiment with releasing patterns. It involves protecting the workers from the immediate consequences of the experiments, because the experiments will not all work. It involves modeling the release at the leadership level, so that the workers can see that the leaders are doing the same work they are asking of the workers. It involves time, because the release is not a switch.

Companies that do this well give experienced workers a credible reason and a fair opportunity to make the transition. Companies that do it poorly risk resignation, quiet withdrawal, or justified resistance to a badly designed change. A disengaged accomplished worker may still be in the building, but leaders should not diagnose that state from age, skepticism, or slower adoption. Examine contribution, invite dissent, test the new practice, and ask what the system is doing before deciding the person is the problem.

What Stays

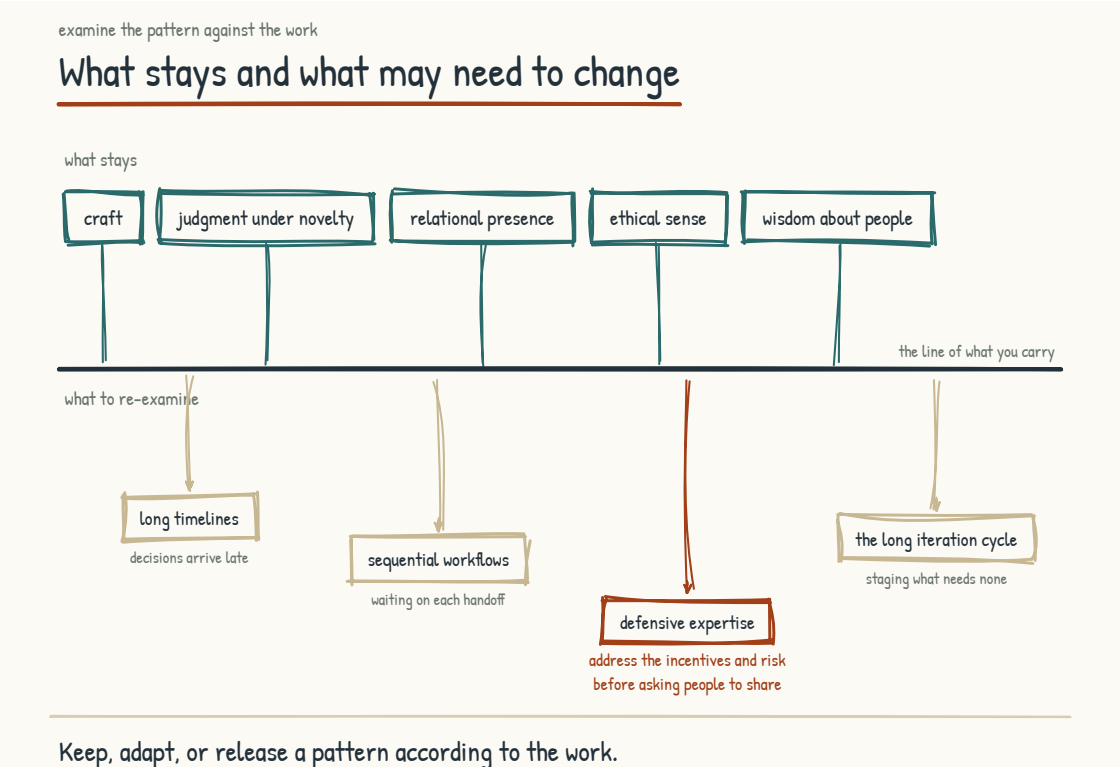

Figure 5.8b. What stays above the line and what may need revision below it. Defensive expertise can be among the hardest patterns to release because it once protected both quality and standing.

One thing to name before closing: what does not need to be unlearned. The conversation can imply that workers should release everything that worked for them. That is wrong. The seven kinds of human value from Section 5.3 remain important: judgment under genuine novelty, relational presence, ethical standing, contextual knowledge, and the rest. AI performance is uneven rather than cleanly divided between routine playbooks and uniquely human situations. As capability moves, these human capacities matter because organizations choose to place authority, responsibility, participation, and trust in people. No benchmark guarantees that machines will stay weak at them. Two capacities deserve emphasis because the unlearning conversation most often unsettles them.

Craft stays. A worker's developed sense of what good work looks like in a domain is one basis for evaluating AI output. Craft gives the collaboration standards, exceptions, and a memory of consequences. AI can amplify it and help novices begin to acquire it. Without craft or effective external checks, AI-mediated work may be polished yet shallow, technically broken yet persuasive, or occasionally better than the person expected. Fluency is not the same as quality.

Wisdom about people stays. Accumulated understanding of how individuals and groups behave, how trust is built and broken, how power moves, and how change lands is consequential human knowledge. AI can model patterns, recall cases, and offer interpretations; it does not participate in the relationship or carry legitimate authority to decide what should be done to people. Workers with this wisdom can be more valuable when the operating model makes room for it. The company still has to recognize and use it.

The unlearning is real and the unlearning has limits. The test is whether a revised practice fits the changed conditions while preserving the protection the old practice was built to provide. The companies that take both seriously, the release of the patterns that no longer fit and the preservation of the capacities that remain valuable, will produce workforces that can do the new work. The next section completes Stage 5 by asking what the workforce of an AI-Native company actually looks like, given everything the human work of this part has built so far.

Section 5.9

Who Works Here Now: Designing the AI-Native Workforce

Who works here now, and how do we design that?

The previous sections of Stage 5 made the case for the human work. What humans bring. What the workforce experiences during the transition. The cognitive posture that distinguishes development from unexamined dependence. The conditions that support up-leveling. The hidden capability AI may help surface. What has to be unlearned. The role of challenge in worthwhile work.

What this part has not yet addressed is the operational expression of all of that. The org chart. The roles. The hiring rubric. The pipeline. What the workforce of an AI-Native company might look like two or three years from now, after some of the human work has begun and some of the architecture is in place. Neither is ever finished. All of it is design work. Your leadership team is doing it whether it recognizes that or not, because every hire, promotion, layoff, automation, and reorganization builds the workforce by default if not by design.

The org chart is changing in two directions at once. The human side is changing through role types, management layers, learning pipelines, and the skills for which the company hires. The agentic side is emerging as software takes more active roles in workflows and therefore needs identities, permissions, evaluation, and accountable human ownership. The two shifts interact. A leadership team that treats workforce design and system delegation as one connected question can produce a more coherent operating model than a team that lets each proceed in isolation.

Figure 5.9a. Three role archetypes in this field guide's synthesis: IC, DRI, and AI-Founder. Dorsey's published model uses player-coach as its third role.

The Human Side: New Role Types

The first human-side shift is in the role types themselves. In a 2026 Lightcone conversation, Y Combinator partner Diana Hu described three archetypes for companies built around AI and connected the framing to Jack Dorsey's thinking at Block. The accounts are not identical. Dorsey's published model names IC, DRI, and player-coach. Hu's discussion uses the phrase AI-founder type for a hands-on leader who demonstrates what the new tools make possible. The three categories below are therefore a synthesis for this field guide, not a settled taxonomy or a verbatim Block org chart.

The IC is the individual contributor who does the work. Writes the code. Builds the model. Drafts the analysis. In many AI-Native roles, the IC will be expected to use approved AI where it improves the work, with the human contribution including domain skill, judgment, taste, verification, relationships, and accountability. AI should not be the default in work where law,

confidentiality, safety, client agreement, or the need to maintain unaided capability says otherwise. Fluency is becoming a baseline in some roles and remains specialized in others.

A DRI, the Directly Responsible Individual, owns a defined outcome and has authority to make the calls inside it. Accountability without authority is theater; authority without checks is risk. The role can become more useful when work crosses functions or management layers change because it makes one human responsible for integration without pretending that legal, fiduciary, professional, or executive accountability can be delegated away.

AI-Founder is Hu's useful name for a posture more than a standard job title: a leader who works hands-on enough to demonstrate what AI changes, connects architecture to the operating model, and takes responsibility for where the system should and should not act. The posture cannot replace the engineers, security specialists, domain experts, workers, and risk owners who make the system real. In a startup, a founder may carry it. In a larger company, the work may be distributed across product, technology, operations, legal, security, data, and workforce leadership. Where one executive owns the integration, the reporting line matters less than the authority and cross-functional accountability.

These three archetypes are a starting frame, not an exhaustive list. Dorsey's player-coach belongs beside them because developing human craft is not optional. So do roles for assurance, data stewardship, security, employee representation, and domain accountability. The point is not that every company needs these titles. It is that existing roles should be decomposed against the work now required rather than renamed around a fashionable taxonomy.

The Human Side: Compressing Management Layers

A second possible shift is structural. Some companies are compressing management layers between the CEO and the individual contributor; others are adding governance, platform, and risk roles around AI. There is no evidence that hierarchy is compressing directionally everywhere. The design question is which layer exercises necessary judgment, development, coordination, or control and which exists mainly because information used to move slowly.

Block announced in a February 2026 SEC filing that it expected to reduce its workforce by more than forty percent. The filing described the plan as aligning organizational structure with its operating model and strategic priorities; it did not say the reductions concentrated in middle management. In a later Sequoia conversation, Jack Dorsey said Block's maximum depth was about five people between him and another individual and that he wanted to reduce it to two or

three. He presented an AI intelligence layer as a way to provide shared context and reduce information-routing work. That is an executive thesis being implemented after a severe restructuring, not yet evidence that the model works.

Block makes visible a hypothesis many companies are considering: if software can make status, dependencies, and context more accessible, some information-routing work can shrink. AI does not resolve conflict, allocate scarce resources legitimately, notice every hidden dependency, or guarantee that the shared context is true. Managers also do more than route information. They develop people, translate strategy, protect teams, exercise judgment, and carry formal responsibilities. Companies should evaluate the work of each layer before cutting a category. If they remove coordinators and mentors indiscriminately, they may discover that the supposedly redundant layer was holding integration, apprenticeship, and trust together.

Layer redesign strengthens the case for senior IC roles with substantial scope and no people-management obligation. Dual career ladders predate generative AI, but AI makes the old promotion bargain look even stranger: the company should not remove an excellent practitioner from the work merely to increase status and pay. A deliberately designed IC track can improve retention and craft leadership. Whether it does depends on compensation, authority, recognition, and mobility matching the management track in practice rather than only on paper.

The Pipeline Question: How Seniors Get Formed

A third shift may be the most uncomfortable. Parts of the pipeline that formed senior workers are changing, and the AI-Native company has to decide how senior judgment will be formed when entry-level tasks change.

The pre-AI pipeline ran partly through entry-level work. Junior analysts assembled analyses, associates reviewed documents, and developers wrote boilerplate. AI can now perform portions of each task, with uneven reliability. Some of the work was needless toil. Some exposed novices to cases, exceptions, causal chains, and feedback that later informed judgment. Repetition alone did not create expertise, and many old pipelines excluded capable people or taught bad habits. The design mistake is to automate a task without identifying which learning it carried and how that learning will be rebuilt.

I call this the apprenticeship inversion: the technology can perform work assigned to novices by drawing on patterns created by experts, while the organization still needs a way for novices to become those experts. I have not found a reliable published source for the earlier attribution of

this phrase to Steve Akkara, so the idea stands here as this book's label, not borrowed authority. The risk is real but the outcome is not predetermined. Junior hiring expectations are mixed, task exposure is changing unevenly, and new forms of practice may prove better than the old ones. What leaders cannot assume is that senior talent will replenish itself after its formative work disappears.

Companies taking the risk seriously are experimenting with a different pipeline. Juniors can learn to frame tasks for AI, evaluate output, trace evidence, integrate across tools, and recognize failure. They also need direct encounters with source material, customers, operations, consequences, and the underlying task without assistance. Orchestration may become formative; it has not yet accumulated the decades of evidence the old paths have. Mentors need enough practice to teach both the tool and the domain, and they need protected time to observe work rather than merely inspect polished output. Organizations should test whether novices are gaining independent capability instead of assuming a stage label guarantees it.

There is survey evidence pointing in more than one direction on the pipeline question. The hiring expectations in Section 5.7 leave the apprenticeship question open: what work will give new employees the experience they need? The survey did not ask whether a lower entry barrier caused those expectations, and expectations are not labor-market outcomes. It is evidence of uncertainty, not resolution.

Agents as a Workforce Category

The org-chart shifts on the human side are real and they are also incomplete. The other thing happening to the org chart is that a new category of entity is appearing on it. Agents.

The framing is recent and contested. AI agents are software systems, not workers in the legal, moral, or experiential sense. Yet treating them as ordinary undifferentiated software can hide the delegation involved when a system selects tools, accesses data, and takes consequential actions. “Workforce category” is a management metaphor for making that delegation visible. It can prompt useful questions about identity, scope, ownership, evaluation, and lifecycle. It becomes dangerous when it anthropomorphizes software, erases the humans whose labor built and supervises it, or implies that an agent can bear accountability.

A production agent should operate through an attributable workload or service identity rather than a shared human login. That identity needs narrowly scoped and preferably short-lived credentials, explicit authorization, and records connecting a particular version, configuration,

delegation, and action. A named human or accountable function owns the deployment. A supervisor agent may monitor another system, but it cannot absorb human accountability. The deployment has an evaluated performance profile and a controlled lifecycle: approval, release, monitoring, change, suspension, and retirement.

The vocabulary resembles workforce management because both domains ask who may do what, under whose authority, and with what evidence. The differences matter just as much. Software does not consent, develop a career, possess rights, or suffer a bad workplace. A company operating hundreds of agent instances needs a registry, access governance, evaluation, incident response, change control, and accountable human ownership. Calling that an agent workforce can help leaders see the operating problem. It does not turn the systems into employees or make HR practice the right technical control plane.

INTELLECTUAL LINEAGE

Three 2026 sources illuminate different pieces of the shift, but they do not establish that agents are employees or independently accountable entities. Muqsit Ashraf's World Economic Forum article argues for AI-native business models oriented toward growth rather than efficiency alone. Cole Stryker's IBM explainer describes agentic workflows as one expression of systems designed around AI from the outset and emphasizes governance and continuous monitoring. Bessemer Venture Partners' Graph AI case study describes one pharmacovigilance product's "push model," in which the system surfaces information inside a workflow without waiting for a prompt. That is a company-specific design choice, not a maturity standard for every operation. Read together, the sources support a narrower conclusion: software is being delegated more active roles in workflows, so scope, controls, and human accountability need to become more explicit.

Ashraf, M. (2026). Why AI's greatest payoff is growth and how leaders can build AI-native businesses to capture it. World Economic Forum. / Stryker, C. (2026). What is AI native? IBM Think. / Bessemer Venture Partners (2026). Graph AI: A service firm turned AI-native solution for pharma and life sciences. BVP Atlas.

Agent Identity, Permissions, and Accountability

The leadership team that decides to treat agents as a workforce category has to make several specific design decisions where the field is just beginning to converge.

Identity. A product name is not a security identity. "Devin," "Einstein," or "Copilot" may identify a product family while thousands of instances, versions, and delegations act within it. Each

production deployment needs a unique machine-readable identity linked to its owner, configuration, version, environment, and authority. Human-readable names help people talk about the system. Workload identities, credentials, and verifiable records make its actions attributable. NIST opened formal work on software-agent identity and authorization in 2026 precisely because the details are not yet settled.

Permissions. The deployment has a defined scope: what data it can read, which tools it can invoke, what systems it can change, how much it can spend, and which actions require human authorization. Begin with least privilege and separate read, propose, approve, and execute. Performance evidence may support a scope change, but permission expansion should remain an explicit risk decision with expiration and revocation, which a person makes and can undo. Role-based access control may help, alongside attribute-, policy-, and task-specific controls for delegated authority.

Accountability. When an agent contributes to a bad outcome, accountability remains human and institutional. A named service owner should be answerable for operation, but responsibility may also sit with the executive authorizing the use, the process owner, the professionals relying on it, the vendor, and the organization itself. One “manager” cannot become a liability sink for a system they lack authority or resources to control. Map decisions and duties before deployment, including who can stop the system and who must notify affected people.

Audit and observability. Material inputs, tool calls, approvals, outputs, policy decisions, errors, and state changes need tamper-resistant records sufficient to reconstruct consequential actions. “Log everything” is neither feasible nor responsible: secrets, personal data, privileged material, and sensitive prompts require minimization, access control, retention limits, and sometimes deliberate exclusion. Review should be routine and risk-based, not reserved for incidents. Observability supports detection and learning; it does not by itself make an unsafe action safe.

None of those four design decisions is exotic, but none is solved merely by naming it. Any company delegating consequential action to software has to settle them. If leadership leaves them implicit, defaults, vendors, and local teams will settle them instead.

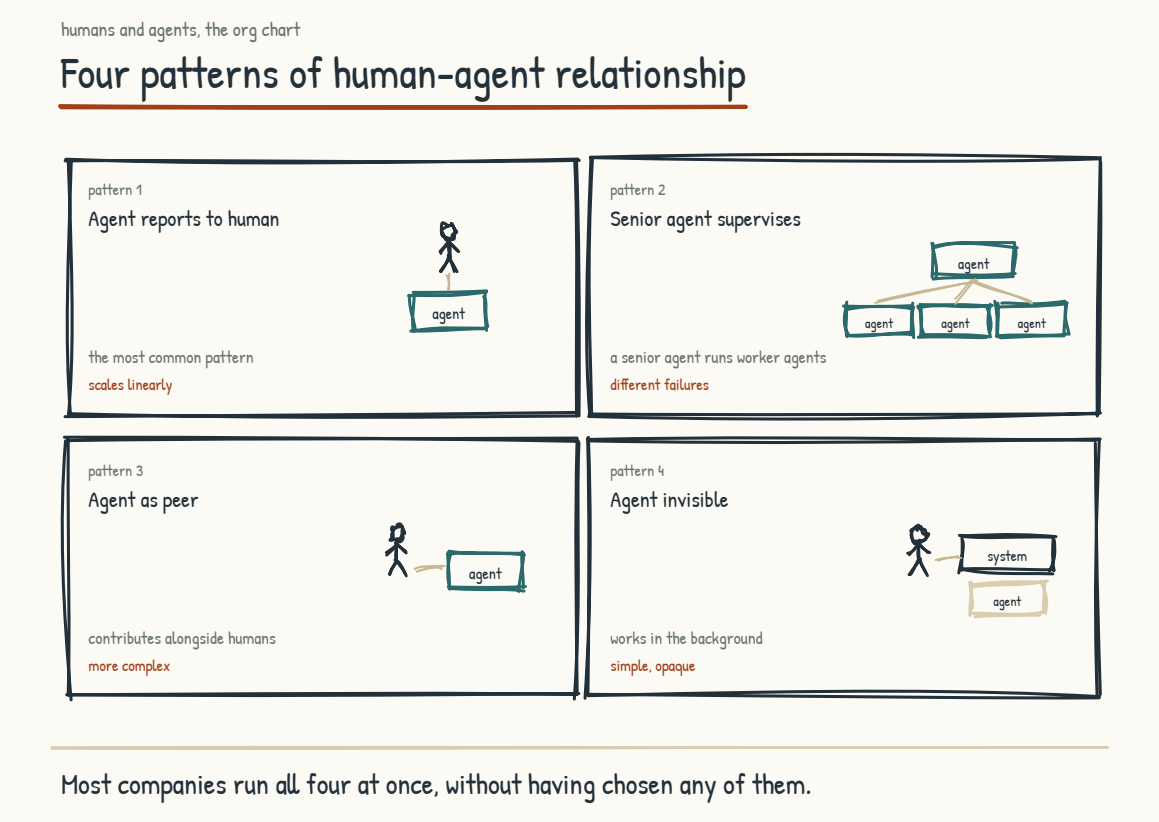

Figure 5.9b. Four patterns of human-agent relationship on the AI-Native org chart.

The Interface: How Humans and Agents Relate on the Chart

The most operationally interesting question, and one with no settled answer, is how humans relate to agentic systems in the work. Four design patterns are useful to examine. They are an analytical frame for this book, not a measured taxonomy of AI-Native companies.

The agent is owned by a human-led function. A production deployment has defined scope and a named owner with authority to suspend it. Human review is calibrated to consequence rather than applied ceremonially to every output. The pattern preserves a clear accountability path, but the relationship is not literally managerial and need not scale one human for one agent. One team may govern many narrowly scoped systems if evaluation, controls, sampling, and escalation remain credible.

The agent is monitored by another automated system. A controller can check policy, compare outputs, route exceptions, or monitor a fleet of worker agents. Calling it a senior agent risks confusing hierarchy with assurance. The controller may share models, data, vulnerabilities, or blind spots with the systems it checks. It can reduce human review volume; it cannot decide

which matters require human judgment without a human-defined policy, tested escalation behavior, independent controls where needed, and a recovery path when the monitor misses the failure.

The agent appears as a participant in a team workspace. It can draft, retrieve, challenge, summarize, or trigger tools alongside human conversation. “Peer” or “teammate” may be a useful interaction metaphor, but the system has no equal standing, reciprocal obligation, or independent accountability. The interface should make its nonhuman status, capabilities, sources where appropriate, and current authority legible. Humans need practice not in bonding with a fictional colleague, but in using, questioning, correcting, and stopping an active system together.

The agent operates in the background. Software already performs invisible automation, and some agentic work will be embedded in ordinary systems. Invisibility may simplify an interface; it can also defeat informed reliance, contestability, worker consultation, and legal disclosure duties. People should be told when AI materially shapes a consequential recommendation, evaluation, or action affecting them, and they need a path to explanation, correction, and human review. The design question is not whether every internal call gets a badge. It is where invisibility changes a person's rights, choices, expectations, or ability to diagnose failure.

The four patterns are not mutually exclusive. A company may use several within one process. The leadership decision is which pattern fits the task and consequence, what controls surround it, how affected workers participate in the design, and how the system's role is communicated.

What Is Not Yet Settled

Several substantive questions about the org chart of an AI-Native company do not yet have settled answers in the field. The leadership team making decisions today is making them while the conventions are still forming. The honest framing is to name what is not yet settled rather than pretending the field has converged.

One caution about that framing, because it can be read as permission to wait. Products shipped ahead of standards. In February 2026, NIST's National Cybersecurity Center of Excellence released a concept paper on applying identity and authorization practices to software and AI agents, and NIST launched an AI Agent Standards Initiative covering interoperability, security, and identity. Existing workload identities, service principals, short-lived credentials, and policy systems provide building blocks; no evidence says every major identity stack has solved agent

delegation. A leadership team should not wait for convergence to prohibit shared human accounts, constrain authority, and preserve attribution.

The legal and HR question. An agent is not an employee under current employment law merely because a company calls it one. The relevant legal questions concern the humans: worker consultation, discrimination, privacy, surveillance, safety, labor rights, intellectual property, professional duties, and accountability for automated decisions. The EU AI Act, GDPR, and employment law regulate systems and their effects in particular contexts; they do not create an employment relationship for software. The workforce metaphor should never obscure who holds rights and who owes duties.

The cost-allocation question. Agents do not receive compensation; vendors, compute providers, integrators, reviewers, and the people whose knowledge supports the system do. Unit inference costs may fall while total cost rises through volume, integration, evaluation, security, energy, compliance, and failure. Human labor costs also move with markets and job design. Economic pressure may favor automation for some tasks, but “more agents and fewer humans” is not a law. Quality, demand, liability, customer preference, regulation, and complementary work shape the result. Leadership is making allocation decisions today whose distributional effects should be explicit.

The evaluation question. An agent does not need the equivalent of a human performance review. It needs system evaluation: task success, error severity, calibration, policy compliance, security, latency, cost, resilience to unusual inputs, escalation quality, and effects on people and the surrounding process. Many deployed agents change through models, prompts, tools, retrieval sources, policies, or configuration rather than retraining. Evaluate before release, monitor in operation, re-evaluate every material change, and compare the combined human-AI process with credible alternatives.

The cultural question. What does it do to a workforce when active software appears as a teammate, evaluator, monitor, or invisible intermediary? People may experience the same system as useful tool, threat, partner, bureaucracy, or surveillance. Design choices, employment consequences, transparency, reliability, and participation are likely to matter. The patterns are too young and context-dependent to support a general outcome claim, which is a reason to involve workers and measure effects rather than infer sentiment from adoption.

Naming what is unsettled is honest and strategically useful. It creates room for reversible experiments, explicit assumptions, worker participation, and stop conditions. Early movers may shape conventions; they may also create expensive failures. Deliberate does not mean first. It means choosing with evidence, recording why, and retaining the ability to change course.

BRING THIS INTO THE ROOM

Block ninety minutes with your leadership team and workforce representatives and walk through the following exercise. Draw your current org chart on a whiteboard: the actual roles, reporting lines, decision rights, and informal coordination. Now draw a twenty-four-month hypothesis. Be specific. Which tasks remain human-led? Which are AI-assisted? Which may be delegated to an agent under human accountability? Where are professional judgment, security, worker voice, and independent review required? How many management layers, and what necessary work does each perform? Where are the senior IC and player-coach tracks? Who owns cross-functional AI integration? Then name the transitions. Do not write “this role becomes an agent.” Identify the tasks changing, the evidence required, the people affected, the learning pipeline, the consultation and support owed to them, and the stop conditions. The artifact does not predict the future. It surfaces decisions the team might otherwise make implicitly.

The Strategic Case

The strategic argument for designing the workforce deliberately, including human work and delegated software, is that the alternative produces accumulated consequences the company did not intend and may find costly to reverse.

The leadership team that does not design the workforce will still have a workforce two years from now. It will be the result of accumulated individual decisions: who got hired, who got promoted, who left, what work got automated, what agents got deployed. Without a shared design, those decisions can produce recognizable mismatches: too many coordinators and too few senior ICs for the work, agent capabilities that do not fit what people are actually doing, or pipeline gaps discovered only when senior talent is already scarce. A design does not eliminate those errors. It makes the assumptions connecting the decisions visible soon enough to challenge them.

A shared workforce design makes a different outcome more likely. Hiring can align with the operating model. Agent deployments can fit the work rather than merely the vendor roadmap.

Apprenticeship can combine orchestration with domain practice and unaided capability. Compensation, recognition, consultation, and career systems can reflect the workforce being built rather than the workforce of ten years ago. The design is a hypothesis that must be revised as performance and human consequences become visible.

Over three to five years, that work can produce a more capable and coherent workforce than uncoordinated automation does. It can also fail through bad forecasts, weak execution, or a strategy that treats people as variables on a chart. The advantage lies not in having drawn the future org chart. It lies in having made assumptions inspectable and transitions governable.

Stage 5 has been the human work. The leader's own practice. The reckoning. What humans bring. The cognitive posture. Up-leveling. Passion. Hidden capability. What must be unlearned. The role of challenge. And now workforce design, with humans retaining standing and accountability while agentic systems take bounded roles in the work. The next part addresses the practices that make this real day by day, regardless of what the chart looks like. Structure without practice will not function. Practice without supportive structure will struggle to scale. That is Stage 6.

CONSIDER BEFORE YOU ACT

If you keep three things from Stage 5: 1) Leadership practice is a powerful influence on workforce fluency, not its only source or absolute ceiling. 2) The reckoning belongs beside the architecture: respectful treatment, worker participation, and an honored covenant make knowledge-sharing more credible, but never compulsory or automatic. 3) Use AI for thinking, not as a substitute for the human capability, authority, and responsibility the work still requires; test the difference rather than assuming it will remain visible on its own.

STAGE 6

The Practices

The specific behaviors and moments that make any of this real.

Section 6.1

Human Practices: How Teams Stay Human in an AI-Saturated World

What does a team do to keep the human work alive day to day?

The previous part of this field guide made the case for the human work. It argued that humans hold kinds of standing and responsibility AI does not, that worthwhile challenge can support development, that up-leveling is possible under the right conditions, and that technical design cannot compensate indefinitely for a company that gets participation, trust, work design, and accountability wrong.

Those arguments are necessary and insufficient on their own, because an argument does not change what happens in a team on a Tuesday afternoon. Whatever your leadership team holds about the human work matters only to the extent that it shows up in how people actually behave. The path from values to behavior runs through practices.

I mean something specific by practice. A practice is something a team does, repeatedly, in a recognizable form, often enough that it becomes part of how the team operates rather than a special event. A practice is not a value. A value is what the team says it values. A practice is one place the value becomes observable. The two are connected. Practices can reinforce values when they are aligned and corrode them when they are not. A team that says it values human connection yet makes no protected space for attention, candor, or relationship is relying on

aspiration alone. Screens are not the test; distributed teams, disabled workers, and assistive technology may depend on them. The test is what the team repeatedly makes possible.

What follows is a working set of behaviors I have seen make the human work more observable across different kinds of teams. The list is organized so a team can adopt some of the practices deliberately, test what they change, and recognize when aspiration is or is not reaching daily work.

SOURCE

Google's Project Aristotle examined 180 teams and hundreds of candidate variables in an internal effort to understand team effectiveness at Google. The researchers reported that who was on a team mattered less than five group dynamics: psychological safety, dependability, structure and clarity, meaning, and perceived impact. Psychological safety was foundational in Google's account. This was an organizational study in a particular company, not a randomized experiment or a universal recipe. Amy Edmondson's 1999 paper established psychological safety as a shared belief that a team is safe for interpersonal risk-taking and found it associated with learning behavior in 51 teams at one manufacturing company. The ten practices below are my working set. They are compatible with parts of this literature; Project Aristotle did not test or validate them.

Google re:Work (2015). Project Aristotle. / Edmondson, A. C. (1999). "Psychological Safety and Learning Behavior in Work Teams." Administrative Science Quarterly, 44(2), 350-383. / Edmondson, A. C. (2018). The Fearless Organization. Wiley.

Figure 6.1a. Ten repeatable practices that keep the human work real, with the five conditions that make them stick.

The Ten Practices

Practice one, the Full-Attention Cadence. The team gathers at a regular interval in a form designed for attention. Close unrelated content. Do not multitask. If a screen or assistive technology is necessary, use it without turning it into a second meeting. If AI records or summarizes, disclose it, obtain the consent the context requires, protect sensitive material, and ask whether its presence will reduce candor. The conversation is about the work, the team, the questions, the things that matter. The cadence can be weekly, biweekly, or monthly. The form should fit distributed, frontline, neurodivergent, and disabled colleagues rather than making one visible behavior the proof of attention.

Practice two, the Ritual of Beginning. The team starts substantive sessions with a deliberate transition into the work: a check-in, a moment of orientation, or a short framing of purpose. The

ritual is brief. It can help people arrive in the same conversation and understand what kind of participation is being asked of them. It is not magic, and some meetings should begin directly. The test is whether the opening improves readiness and clarity rather than becoming compulsory intimacy or ceremony without function.

Practice three, Naming What Is Hard. The team has a vocabulary for surfacing what is difficult, uncertain, or uncomfortable in the work. When something is hard, a member can say so without performing vulnerability or requesting special handling. That only works when leaders respond without punishment, curiosity is real, and people retain boundaries around what they disclose. Vocabulary cannot manufacture psychological safety. Used inside a credible climate, it makes important facts easier to discuss before they accumulate into avoidable failure.

Practice four, Protected Unstructured Time. The team has some time in its operating rhythm that is not scheduled for a specific output. The time can support thinking, conversation that meanders, problems not yet named, and relationships not organized around a deliverable. Not every worker experiences open-ended social time as restorative, and hourly or frontline teams should not be expected to donate it. Make it paid, purposeful enough to be safe, optional where appropriate, and one source of connection and discovery rather than the only place a team becomes a team.

Practice five, Specific Appreciation. Members of the team see that their work is noticed by people whose opinion they respect. The appreciation is specific, grounded in the actual contribution, proportionate, and free of favoritism. It can be a one-line message, a moment in a meeting, or a thank-you that names what was done. Generic appreciation is not necessarily worse than silence, but repeated recognition theater teaches people not to trust the signal. Specific appreciation can strengthen membership when it sits beside fair pay, credit, opportunity, and candid feedback rather than substituting for them.

Practice six, Working on the Team: the regular review of how the team is working, distinct from what it is producing. What is going well in the way we operate. What friction could we remove. What pattern is starting to hurt us. The review creates an opportunity for improvement; it does not produce improvement unless someone can act, owners and dates are clear, and the team sees what changed. A retrospective that repeatedly surfaces the same immovable problem becomes evidence of powerlessness rather than learning.

Practice seven, Deliberate Difference. The team actively surfaces perspectives different from the dominant view. When everyone agrees, it asks who is affected and absent. When a decision feels obvious, it asks what evidence would change the conclusion. Difference does not automatically improve judgment; it has to be relevant, heard, and integrated, and marginalized members should not be conscripted to represent a category. Manufactured dissent is different. The practice is to make disagreement usable and reduce shared blind spots without turning debate into theater.

Practice eight, Explicit AI Decisions. The team is conscious about when AI is invited into the work and when it is not. AI may draft for one task, retrieve evidence for another, record a session with informed permission, or remain absent because privacy, law, client expectation, psychological safety, or the learning objective requires it. AI should never be a silent participant if that means people do not know it is present. The decision, purpose, data boundary, and human owner are visible. Conscious calibration is the practice; consequence determines how formal it must be.

Practice nine, Team AI Reflection. At intervals, the team examines how it is using AI. Where did it help. Where did reliance exceed evidence. Which errors escaped. Whose work or voice became less visible. Where is capability developing, and what independent tests would reveal whether it is being lost. Collective reflection can reveal patterns no individual sees, but memory and impressions are not enough. Bring examples, operational measures, affected-worker feedback, and changes the team can test.

Practice ten, Celebrating the Irreducibly Human. The team acknowledges work in which human standing and participation mattered: the accountable judgment call, the relationship held in a difficult moment, the dissent that changed a decision, the reframing grounded in lived stakes. Do not base recognition on a claim that AI could never produce similar words or recommendations. Celebrate why the organization wanted a person involved and what responsibility, legitimacy, care, courage, or relationship that person carried. What teams notice influences what they invest in, even if celebration alone does not guarantee more of it.

These ten are not the only practices that matter. They have appeared repeatedly in teams I have observed living the human work well. That is practitioner evidence, not proof that adopting all ten causes superior outcomes. A team can select a few, define what it expects them to improve, watch for burden or exclusion, and adapt from evidence. The practices should serve the team. The team should not become a delivery mechanism for the practices.

How These Form a Team-Level Culture

Taken together those practices form more than the sum of individual behaviors. They form a culture, in the specific sense of a shared way of operating that the team holds collectively rather than individually.

The distinction matters because individual-responsibility framings are insufficient for sustained culture. A team that says each of us is responsible for embodying our values has individual aspirations but no shared mechanism for resolving the gaps between them. A team with recognizable collective practices has something members can observe, question, teach, and revise. Individual conduct still matters. Shared practice makes it less dependent on private interpretation alone.

A culture can become self-reinforcing through visible practices. New members observe what receives time, reward, protection, and consequence, then decide what the stated values actually mean. Practices propagate through daily operation more reliably than through a training event alone. Culture is not only the practices; it also includes stories, incentives, power, symbols, assumptions, and what happens after someone challenges the group. Practice is where much of that becomes visible.

A team that wants to develop this culture has to invest over time. The first attempt at a practice may be awkward. Repetition may make it natural, or reveal that the form is wrong for this team. Do not abandon a useful practice because novelty felt uncomfortable, and do not preserve a harmful ritual because persistence has become a virtue. The work is to distinguish adjustment discomfort from evidence that the practice is excluding people or failing its purpose.

How a Leadership Team Makes Them Stick

Everything I have described works at the team level. It depends, though, on conditions your leadership team controls. I want to name the conditions, because the most common reason these practices fail to stick is that the leadership team has set up the surrounding conditions in ways that defeat them.

Condition one, Model It Yourselves. A leadership team that asks its workforce to gather in full attention while its own meetings are full of split attention creates the cynicism of a double standard. Modeling does not guarantee adoption, but it gives the request credibility that a memo cannot supply.

Condition two, Under-Load the Calendar. The leadership team has to make time for the practices. A team scheduled at nominal full capacity has no room for learning, interruption, care, or the variability real work contains. Capacity should be measured against the actual work system, including frontline schedules and paid participation, rather than solved by adding one more meeting to an already overloaded week. Slack is not waste. It is where adaptation lives.

Condition three, Protect Against the Urgent. The leadership team has to protect the practices from work that crowds them out. An immediate demand carries a deadline and a visible owner; developmental practice often carries neither. That asymmetry makes the practice easy to cancel in any given week and easy to lose through accumulation. Protection is not a one-time announcement. It is a commitment renewed in calendars and priorities week after week.

Condition four, Review the Practices. Leadership has to pay attention to whether they are happening, whom they serve, what gets in the way, and whether the expected effect appears. Counting vulnerable disclosures or ranking teams on psychological safety would corrupt the practice and invite surveillance. Use voluntary feedback, observable operating outcomes, and qualitative review with privacy and anti-retaliation safeguards. The signal should be that leadership is responsible for the conditions, not that workers are being scored on their humanity.

Condition five, Patience With Checkpoints. Culture does not transform on a quarterly reporting schedule, but leaders should not use a multi-year horizon to protect ineffective practices from evidence. Look early for participation, candor, reduced friction, better escalation, and changes the team can name; look longer for learning, retention, quality, and resilience. Patience and accountability belong together.

What This Section Does Not Cover

I want to be honest about the limits of what I have done here. The ten practices are a working set rather than a complete prescription, and the conditions are framing rather than a detailed playbook. The work of adopting them in a specific team, in a specific company, is real work that will require adaptation, experimentation, and refinement.

I also have not addressed the question of practices that are specific to particular kinds of teams. Engineering teams have practices that consulting teams do not have. Customer-facing teams have practices that internal-facing teams do not. The full operating system of practices for an AI-Native company is much richer than the ten I have named here, and the full operating system

has to be developed in the context of the specific company, with the specific workforce, in the specific industry.

The next section addresses the question of where the human practices I have described concentrate their effect, which is in the moments that matter.

Section 6.2

Moments That Matter

Where in the work does human presence matter most?

The argument across Stage 5 and the previous section implies a question worth addressing directly. If humans retain standing, responsibility, relationships, and capabilities that the company means to protect, where should human attention concentrate in an AI-Native operation. The question is operational. Its answer shapes where AI may assist, where human authority is mandatory, where participation is a matter of dignity or rights, and where the company invests scarce attention.

Two answers to the question can be set aside before I name the one I hold.

The first incomplete answer is that a human should touch everything. Every interaction, decision, and customer contact passes through a person. The answer sounds protective, but a ceremonial approval can add delay without adding judgment. People become rubber stamps when volume exceeds attention. Some customers prefer fast self-service, and some low-consequence automation needs no individual review. Other domains legally or ethically require a qualified human. The design question is not human everywhere or nowhere. It is what meaningful human involvement requires in this consequence and context.

The second incomplete answer is that humans should sit only at the edge. AI does the work; people align it, oversee it, and intervene on failure. This resembles the Coinbase archetype introduced in Section 2.2 and challenged in Section 5.3. It can work for bounded, observable processes with strong tests and reliable escalation. It fails when supervisors lack enough contact with the underlying work to recognize error, maintain skill, or challenge the system. Oversight then becomes a ritual. Edge supervision is a pattern to validate, not a workforce philosophy.

The answer I hold to is different from both. Humans should be at the moments that matter.

INTELLECTUAL LINEAGE

Jan Carlzon popularized "moments of truth" while leading Scandinavian Airlines through its early-1980s turnaround. In his 1987 English-language book, he estimated that ten million customers each encountered about five SAS employees for an average of fifteen seconds: fifty million annual moments in which the airline was "created" in a customer's mind. The phrase described frontline service encounters, not a finding that every contact alone defined the entire relationship or caused the financial turnaround. Carlzon's operating lesson was to orient the organization around customers and give frontline employees information and authority appropriate to those moments. I am adapting that service framework to a broader question: which moments warrant human attention when AI increasingly participates in the operation itself.

Carlzon, Jan (1987). Moments of Truth: New Strategies for Today's Customer-Driven Economy. Ballinger Publishing Company. (Originally published in Swedish as Riv pyramiderna!, 1985.)

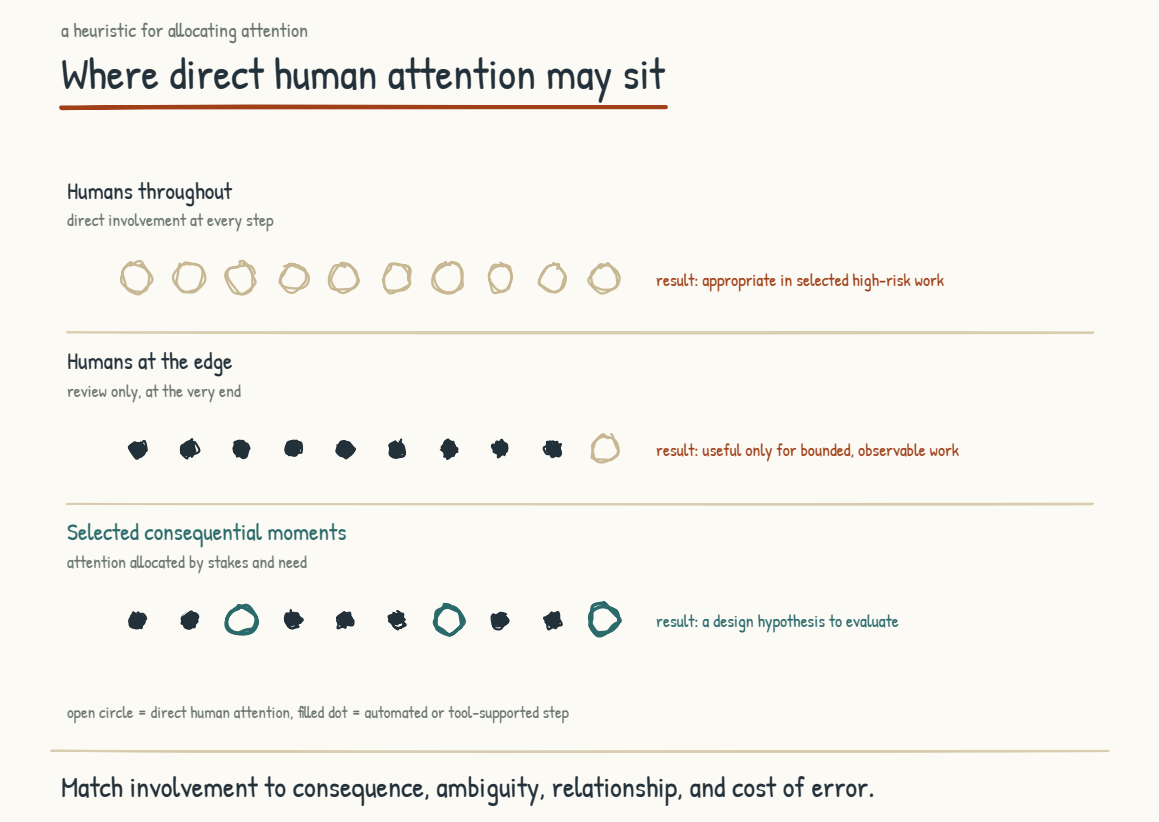

Figure 6.2a. Three approaches to where humans sit, and the seven categories of moments where human presence concentrates value.

What the Moments That Matter Are

A moment that matters is a point in the operation where human attention, authority, or relationship makes a disproportionate difference. Some are identifiable in advance; others become visible only through incidents, worker experience, or customer feedback. They may occupy a small fraction of operational volume, but that is a design hypothesis, not a fixed percentage. They concentrate value, risk, rights, and meaning in ways an average transaction may not.

These moments share several characteristics. They carry high consequence, where the outcome matters more than the average outcome. They are moments of high ambiguity, where the analysis alone is insufficient to determine the right action. They are moments of high relational stakes, where the experience of being met by another human is part of what produces the outcome. They are moments of meaning-making, where the way the moment is held shapes how the people involved understand what happened.

They cluster in identifiable places rather than at random. I want to walk through several categories with examples to make the concept concrete.

A Taxonomy of Moments

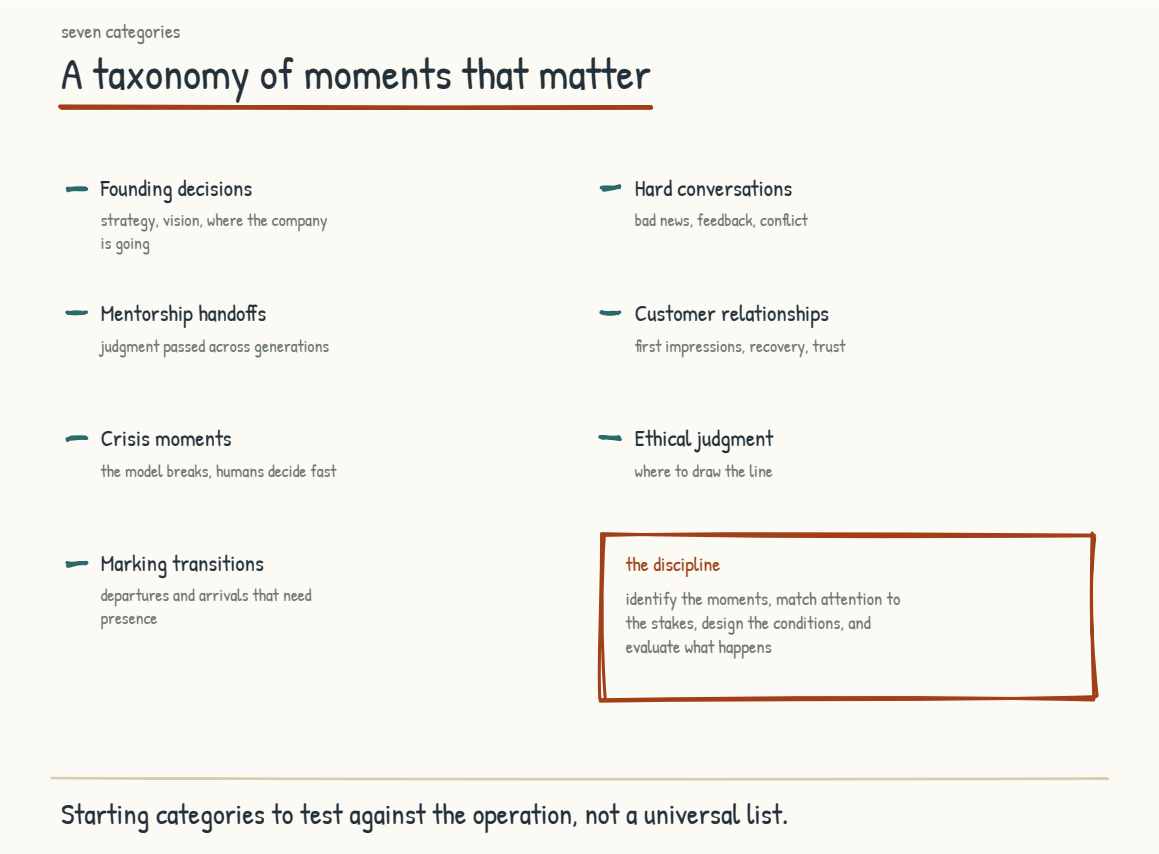

Figure 6.2b. Seven categories of moments where human presence is disproportionately valuable.

Category one, Consequential Decisions: decisions that shape the trajectory of the company, team, customer, or individuals involved. The strategic move that commits the company. A senior hire. Entry into or exit from a market. A credit, care, employment, or safety decision that materially affects a person. AI can assemble evidence, test scenarios, and expose inconsistency. Legitimate human decision-makers should remain accountable, understand the basis, consider affected people, and be able to depart from the system for reasons they can defend.

Category two, Moments of Difficulty for the people involved. The performance conversation. The layoff. The diagnosis. The financial setback. A relationship breakdown with a counterparty. A director held one of these moments well. She booked a small conference room, closed the door, and set the written review face down on the table. The engineer across from her had missed his numbers two quarters running, and she said, tell me what is actually going on. He was quiet, then described his mornings since his wife's diagnosis, and the next forty minutes followed no script either had brought. What he carried from that room was how it was held. Not every

moment is that heavy. Someone stuck for three days walks to a colleague and says this thing is killing me. The colleague has no answer, but has ten minutes and understands why it is maddening. That is enough to make work possible again. AI may help a person prepare words or options. It cannot occupy the employer's obligation, the clinician's duty, or the colleague's reciprocal relationship. Automating delivery of consequential human news without choice and support can impose costs in dignity, trust, and comprehension that a throughput measure misses.

Category three, Moments of Meaning. The launch of a product. Completion of a project that took years. A colleague's retirement. A new team member's arrival. Recognition of work that mattered. Operations can continue without marking these moments, but people may lose continuity, credit, and a shared account of what changed. Ritual is culturally specific and should not compel emotion. Held honestly, it can help a group remember, welcome, grieve, or close.

Category four, Trust-Dependent Negotiation: the complex sale, partnership discussion, supplier agreement with years of consequence, or customer rescue after failure. AI can research, model terms, draft language, and simulate objections. Parties may even use automated negotiation for bounded transactions. Where the relationship itself carries value, authorized humans need to understand commitments, read context, disclose material AI use where required, and stand behind the bargain.

Category five, Moments of Teaching. Mentorship. Onboarding into a culture. Transmission of institutional knowledge from an experienced worker to a less-experienced one. Development of judgment through cases, explanation, attempt, feedback, and consequence. AI can tutor, generate practice, and provide immediate feedback. Human mentors contribute situated standards, recognition, sponsorship, observation over time, and responsibility for what the learner is trusted to do. The strongest design may be human-to-human and human-with-AI rather than a forced choice.

Category six, Moments of Repair. A team conflict. A customer let down. A colleague harmed. A relationship strained. AI can help reconstruct events, identify options, or draft an apology. Repair itself requires acknowledgment and action from the person or institution that owes it, participation from the harmed party on terms that do not coerce reconciliation, and follow-through that changes the conditions. Presence matters because responsibility cannot be outsourced to the drafting tool.

Category seven, Orientation Under Uncertainty. A crisis, disruption, strategic pivot, or competitive shift. The workforce needs evidence, decisions, and orientation from leaders who can state what is known, unknown, changing, and still expected. AI can monitor signals and generate scenarios. Leaders remain responsible for priorities, tradeoffs, communication, and promises. Meaning should be built with the people living the change, not broadcast as a synthetic certainty from above.

These seven categories are starting points, and the answer will vary by context. A company can identify its own moments by asking where human authority is legally or ethically required, where relationship changes the outcome, where people want a person, where independent capability must be preserved, and where absence would be felt as a loss. Ask affected workers and customers. Leaders alone may not see the moments in which their systems feel most impersonal or consequential.

The Design Principles

A company that takes the framing seriously has to design deliberately because high-volume routine work can crowd out the low-frequency moments carrying disproportionate consequence.

Principle one, Identify. You have to identify them explicitly. What are the moments in our company where the human presence makes the difference. What does each moment look like when it happens well. What does it look like when it happens poorly. Who is involved. What is at stake. The identification is a continuous practice of paying attention to where the human work is actually concentrating.

Principle two, Invest. They deserve more investment than the average moment. The people who hold them need time to prepare, development to do them well, and support from colleagues. The investment is disproportionate to the time the moments occupy because the possible impact is disproportionate too. Underinvestment raises the odds that a consequential moment goes poorly. When it does, the story can travel through the workforce as evidence that the human work was rhetoric rather than reality.

Principle three, Protect. They require conditions that let them be held well. The performance conversation needs a private room and uninterrupted time. The strategic decision needs the right people present and the freedom to deliberate. The mentorship conversation needs the relationship to have been built over time. The protection is the leadership team's work. Without

protection, the moments degrade into rushed versions of themselves that do not produce what the moments are capable of producing.

Principle four, Reflect. Afterward, ask what worked, what did not, what evidence exists, and what should change next time. Do not demand emotional disclosure from the people involved or turn a private conversation into training data without explicit authority and consent. Reflection develops craft when it protects dignity and produces a change, not when it converts every human moment into organizational extraction.

Principle five, Design. Preparation, setting, sequence, participation, and follow-through can all be designed. The aim is to create conditions for a worthy interaction, not script another person's response or manufacture emotion. Some well-designed moments will be remembered for years; some should pass quietly. The quality lies in whether the design serves the people and purpose involved, not in memorability itself.

The Connection to the Method of This Field Guide

A connection runs through the entire book and comes to the surface here.

The Arc5 method that produced this field guide, and that I use with leadership teams, treats moments as a unit of transformation work. It designs specific settings in which a leadership team can build alignment, a workforce can participate, and a transition can become concrete. No method can guarantee alignment or engagement, and a designed workshop is not the transformation itself. The distinction from generic facilitation is the deliberate connection between each moment, the operating decision it must enable, and what happens afterward.

The framing is the larger version of that idea. The AI-Native company can concentrate scarce human attention without reducing the rest of the workforce to machine supervision. Leadership identifies, invests in, protects, reflects on, and designs the moments with the people affected. Arc5 is one method for doing this in leadership transformation. The same logic can be adapted, with care, at other levels of operation.

The moments that matter are held by people, and how well they are held depends on what those people bring to them. The next section takes up the conditions under which a workforce brings real passion to the work.

Section 6.3

Cultivating Passion in an AI-Native Workplace

What does passion have to do with AI?

Section 5.6 made the case for passion at the level of the individual: what can make a person care about the work, what may cultivate that caring, and what can destroy it. The leadership-team counterpart is about the measurable state of the workforce on which an AI-Native transition lands and the conditions a leadership team can change. Passion itself is not an organizational KPI, and no leader gets to require it.

Many AI-Native transitions land on workforces reporting low engagement, mixed feelings about AI, and pressure at the manager layer. Those are population-level signals, not a diagnosis of any particular employee or company. A leadership team that ignores its own evidence may spend years trying to change the operating model without the trust, time, participation, and capability the change requires. What follows is a way to examine those conditions, why they matter, and what leaders can do within the legitimate boundaries of work.

A note on the word passion before we begin. I use it broadly in the practitioner sense: fascination with the subject, care for the customer, pride in developing a team, desire to win, or devotion to craft. The formal research definition is narrower: a strong inclination toward a valued, self-defining activity in which a person invests time and energy. Not every kind of motivation in my list meets that definition, and not every worker wants work to become part of identity. Engagement, satisfaction, commitment, motivation, and passion overlap in ordinary speech but are not interchangeable research constructs. A person can be dependable without passion, passionate without being sustainably engaged, satisfied without wanting promotion, or deeply committed to a craft without identifying with the company. Those distinctions protect the employee and improve the diagnosis. The leadership team's job is not to discover and activate a hidden passion in each person. It is to create fair conditions in which care, pride, curiosity, and growth can emerge without being performed or coerced.

What AI-Native Actually Needs from the Workforce

Before the substance, a framing point. Employee engagement is a family of survey constructs, not a synonym for people who merely show up and comply. Different instruments measure involvement, energy, needs, commitment, or favorable attitudes in different ways. An AI-Native

transition also needs concrete behaviors and conditions that a broad engagement score may not reveal.

It needs people with appropriate authority to take ownership of outcomes, not just tasks. It needs worker participation in deciding what the work should become, pride in quality, supported learning beyond current mastery, and the critical effort required to use AI as an active system rather than an unquestioned answer source. These are not obligations employees can satisfy through attitude alone. Leaders have to provide authority, time, tools, safety, voice, and a credible reason to participate.

A workforce cannot be "produced" into this state by exhortation. The claim that most people in most companies are not bringing themselves to work is not established by engagement surveys and turns a structural question into a moral judgment. People may meet the employment bargain exactly, reserve identity for life outside work, or withhold discretionary effort after experience taught them it was unsafe or unrewarded. Contract terms, schedule control, disability, caregiving, immigration status, union representation, prior layoffs, and whether improvement ideas have ever been honored all shape what participation costs. Leaders asking for ownership should first ask whether workers have the information, authority, protection, and share in the benefit that make ownership real. An AI-Native transition still needs active participation from the people whose work changes. The obligation is to design that participation and earn it, not blame workers when technical wins fail to change the operating model.

Autonomy support, opportunities to build competence, relatedness, fair treatment, and useful feedback appear across research on motivation, passion, learning, and work. The constructs overlap; they are not the same outcome, and their relationships are not a single causal law. Purpose is also associated with health and longevity in observational research, but that does not mean workplace passion extends life or that the same mechanism predicts company survival. The three scales are a metaphor to examine, not a scientific equivalence.

The State of the Workforce

Workforce surveys in 2025 and 2026 offer a sobering but method-dependent picture.

Ten trillion dollars is Gallup's model-based estimate of global productivity lost to low engagement in 2025, equal to about nine percent of global GDP. Its State of the Global Workplace 2026 report classified 20 percent of employees worldwide as engaged, down from 21 percent in 2024 and the lowest since 2020 under Gallup's Q12 method. Achievers Workforce

Institute, using a different vendor survey and definitions, reported 36 percent engaged and 24 percent psychologically safe in 2025; the numbers should not be compared as if they measured the same population and construct. Pew's survey of U.S. workers, conducted in October 2024 and published in 2025, found 52 percent worried and 36 percent hopeful about future workplace AI use; respondents could select multiple feelings. Research summarized by Northwestern's Hatim Rahman adds a useful frame: technology's effects depend heavily on organizational choices, occupational power, and who participates in design. None of these sources says anxiety is inevitable or that a vision statement alone resolves fear of job loss.

This is the backdrop, not a verdict on every workforce. Engagement is low under Gallup's global measure, U.S. worker feelings about AI are mixed, and people have rational questions about whether a company is investing in them or replacing them. A leadership team should establish its own baseline, disaggregate it by role and power, listen before prescribing, and revisit it during the transition. The warning is that technical delivery and operating-model adoption can diverge for a long time before the dashboard makes the gap obvious.

The manager layer deserves particular attention without making managers solely responsible for everyone else's engagement. Gallup reported that global manager engagement fell from 27 percent in 2024 to 22 percent in 2025, its largest year-over-year manager decline in that series. Individual-contributor engagement was 19 percent. The former manager engagement premium has narrowed. Managers are workers inside the same system: spans of control, role ambiguity, restructuring, administrative burden, authority, training, and support all shape what they can provide to a team.

The Manager Problem and AI

Section 5.6 argued that the manager's job has changed, toward the human work of engagement. AI makes the problem worse, because it gives managers new ways to disengage from exactly that work.

First failure mode: using AI as a reason to step back indiscriminately from the team. If automation removes status collection, some meetings can become shorter or less frequent, which may be an improvement. What cannot be inferred is that people need less context, coaching, protection, or access to decisions. Engagement never came from presence alone, and not everyone wants more check-ins. Managers should ask what contact is useful, replace surveillance with trust, and use released time for the human work the team actually values.

Second failure mode: outsourcing judgment and relationship through ghostwritten communication. AI can help a manager organize notes, improve accessibility, translate, or find clearer language. The manager still has to supply the observation, make the judgment, check every claim, protect confidential data, and own the message. A generic draft is not automatically hollow, and teams do not possess a universal AI detector. Trust erodes when communication is inaccurate, impersonal, undisclosed where disclosure matters, or presented as attentive human judgment that never occurred.

Both failure modes share a pattern. The manager uses AI to avoid responsibility rather than reduce avoidable administration. AI can also perform parts of coaching, feedback, and communication well enough to be useful; the boundary is not administrative versus human. The durable boundary is whether the manager remains informed, available, accountable, and in a real relationship with the people affected. Leadership should evaluate that outcome rather than police which sentences began with a model.

This is a leading risk because surface signals can move in the opposite direction from underlying capability and trust. Reports may become cleaner while people understand less. Communication may become more polished while access to the manager declines. Productivity can rise while apprenticeship weakens. These are testable possibilities, not hidden certainties. Pair output measures with capability tests, error and escalation data, manager access, voluntary workforce feedback, and evidence that concerns lead to action.

What the Research Tells Us About the Conditions

Research offers several useful lenses on the conditions associated with autonomous motivation and work passion. It does not provide a formula for producing a passionate workforce. Three lenses are useful for a leadership team if their limits remain visible.

Section 5.6 introduced Vallerand's distinction between harmonious and obsessive passion. The leadership question is what working conditions encourage each.

The two forms show different patterns of association. A 2015 meta-analysis synthesized 94 studies and linked harmonious passion more consistently with adaptive wellbeing, motivation, cognition, and behavior; obsessive passion had mixed and often maladaptive associations, including conflict and rumination. A 2025 Human Performance study used a convenience sample of 194 matched employee-coworker dyads in the Philippines. In its time-lagged survey, perceived organizational support predicted self-reported harmonious passion, while

organizational time demands predicted obsessive passion. The design supports relationships among measured variables, not a universal causal claim that leaders determine which passion a worker develops. Context matters, as do identity, personality, culture, history, and life outside work. The practical signal is strong enough: support and reasonable boundaries are more compatible with harmonious engagement than pressure to sacrifice nonwork life.

The second lens is Self-Determination Theory, developed by Edward Deci, Richard Ryan, and many collaborators over decades and studied across cultures and domains. It proposes three basic psychological needs: autonomy, the experience of volition; competence, the experience of effectiveness; and relatedness, the experience of connection. Workplace meta-analyses associate need support and satisfaction with more autonomous motivation and favorable outcomes. The theory does not say that satisfying all three guarantees engagement, performance, wellbeing, or passion, or that frustration leaves motivation only controlled or absent. Pay, justice, safety, health, labor conditions, and the work itself remain consequential. The overlap with harmonious passion is useful without making the constructs identical.

The third lens is the object of care. Formal passion research allows the valued activity to vary, but my five objects (subject, customer, team, competition, and craft) are a practitioner taxonomy, not an established exhaustive scale. A person may care about several, about economic security, about public service, or about work being bounded enough to support a meaningful life elsewhere. None is inherently superior. Leadership should offer more than one credible connection to the work and learn from workers what matters, without sorting people into motivational types or treating reluctance as a defect.

Together, these lenses suggest a direction rather than a switch: support meaningful choice inside real constraints, develop competence, foster connection, respect different reasons for working, and limit cultures of compulsory availability. This is not an invitation to personalize every job until the manager becomes responsible for each employee's inner life. Autonomy includes privacy. Relatedness does not require disclosure. Competence does not require endless advancement. A healthy employment relationship can leave room for ambition, steady contribution, temporary survival, and a life centered elsewhere. Getting those conditions right does not make passion appear across a workforce. Getting them wrong can frustrate motivation and make controlled or obsessive patterns more likely. The outcome has more than three states, and people retain agency within it.

The Longevity Layer

There is a deeper layer to this conversation that most leadership playbooks do not address, but I think is important enough to flag explicitly.

The Blue Zones project popularized observations from Okinawa, Sardinia, the Nicoya Peninsula, Ikaria, and Loma Linda, places reported to have unusual concentrations of long-lived people. Dan Buettner's synthesis includes purpose among nine recurring lifestyle themes and helped make the Japanese concept of ikigai familiar to a global audience. This is an influential journalistic and commercial framework, not a controlled twenty-year experiment establishing nine causes of longevity. Demographers continue to debate validation, boundaries, changing populations, and how much can be inferred from selected regions. The broader literature stands more securely on its own: longitudinal studies and meta-analyses associate a stronger sense of purpose with lower mortality, cardiovascular risk, and dementia risk. Those associations remain vulnerable to selection, health, socioeconomic, behavioral, and social confounding.

Possible mechanisms are multiple. Purpose may organize behavior, encourage persistence, reduce some forms of stress, or strengthen social connection. Health and social resources may also make purpose easier to report and sustain. Epidemiological associations across large samples justify further study; they do not prove the chain “purpose produces motivation, motivation produces health, health produces longevity,” and they do not establish workplace passion as the relevant purpose.

What this means for leadership is not that the company can lengthen life by cultivating work passion. That claim is absent from the evidence and would give an employer an alarming jurisdiction over personal wellbeing. Work conditions can affect stress, injury, sleep, time, relationships, mental health, and access to resources; leaders therefore carry obligations that reach beyond output. Create healthy work because people deserve it and because the company controls part of their environment. Those reasons are enough.

Longevity can be used as a metaphor at the team and company level, but the biology does not transfer. Harmonious passion is associated with lower burnout and some favorable work outcomes; it does not guarantee retention, institutional memory, adaptation, or corporate survival. Teams also sometimes should end, and companies can persist while harming people. Across the three scales, humane conditions may support resilience. They do not constitute one empirically demonstrated mechanism.

The point of naming this layer is to make the conversation bigger without turning metaphor into medicine. This is not “soft.” A quarterly engagement campaign rarely changes the work system. Leadership can instead invest in conditions whose human and operational effects are reviewed over time, while remaining humble about health claims and honest that durability has to be demonstrated.

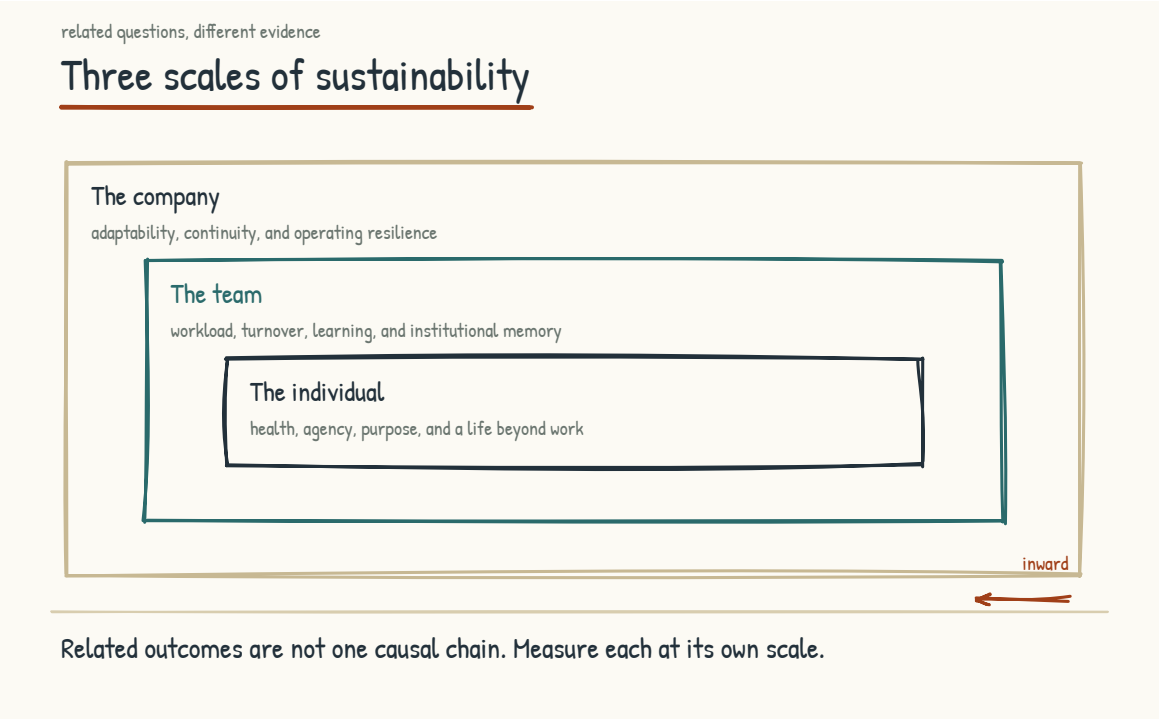

Figure 6.3a. A metaphor across three scales, not one causal model: supportive conditions may contribute to individual wellbeing, team resilience, and organizational adaptability through different mechanisms.

What the Leadership Team Actually Does

The practical content of the section. What does the leadership team do to create the conditions described above? Four areas, each substantive.

Area one, Demonstrate by Example. Leaders who approach the transition with curiosity, ownership, and enough hands-on practice to understand the work send one signal. Leaders who call it transformation while delegating every encounter with the substance send another. Modeling matters, but leadership is not an absolute ceiling: expertise and energy often live deeper in the organization, and leaders should create room for them rather than demand imitation. The obligation is credible participation, informed decision-making, and visible willingness to learn. People can tell when executive passion is a performance.

Area two, Set Clear Expectations. “Be passionate” is neither actionable nor appropriate. Define instead what the transition asks people to do, learn, decide, disclose, and stop doing. Twenty percent learning time and one quarterly artifact may be useful in one role and absurd in another; they are examples, not benchmarks. Set role-specific expectations with worker input, provide the resources to meet them, distinguish experimentation from evaluated performance, and create a safe path to challenge the plan. Ownership grows from meaningful authority and participation. A target by itself supplies neither.

Area three, Create Time and Space. People need paid time to learn new tools, redesign work, test them safely, and sometimes practice the underlying task without them. Adding transformation to an unchanged workload turns learning into overtime and selects for people with fewer obligations outside work. Protect hours, project capacity, access, support, and slack in a form that includes hourly, frontline, contract, disabled, and distributed workers. Some people will still prefer bounded adoption over passion. The investment is real because participation is work.

Area four, Give Real Feedback. Specific, useful, well-timed feedback can support competence by providing information for development, autonomy when it preserves agency, and relatedness when it reflects genuine attention. Bad feedback can frustrate all three. McLean & Company's 2025 vendor report analyzed surveys collected from roughly 216,000 employees across 236 organizations between 2019 and 2024. It reported that employees receiving meaningful manager feedback were 5.7 times as likely to feel supported in career advancement. That is a large association based on the vendor's measures, not a causal return-on-investment estimate. Build feedback discipline around quality, evidence, dialogue, bias checks, and employee usefulness. Frequency quotas and AI-generated volume measure activity.

A note on what is missing from this list. Employee resource groups, recognition, wellness initiatives, benefits, compensation, job security, labor voice, and fair systems are not interchangeable “programs,” and some are substantive institutions in their own right. None substitutes for daily management, time, authority, and feedback; the four areas do not substitute for them either. Leaders should examine the whole employment system. Passion may emerge for some people. Fairness, safety, dignity, and the ability to do good work are obligations even when it does not.

Why All of This Matters for the AI-Native Transition

Pulling the section back to the field guide's central argument.

An AI-Native transition changes how work gets done and therefore requires meaningful participation from the people doing it. Harmonious passion may help; it is neither necessary nor sufficient for a good redesign. The failure cases in Section 8.2 have structural causes as well as motivational ones. Pilots die through weak ownership, integration, economics, governance, or need, and each of those takes structural work to fix. Institutional knowledge walks out when systems fail to capture it and when experts lack reason, safety, consent, or time to share. Hollowing-out occurs when work removes practice and responsibility without rebuilding capability. Do not explain an operating-system failure as a passion deficit in the workforce.

Investing in these conditions strengthens the operational substrate of transition. Even in a supportive company, people may reasonably experience AI as opportunity and threat at once. Resistance can carry information about safety, workload, surveillance, quality, inequality, or job loss. Curiosity is not proof of wisdom, and fear is not proof of backwardness. The outcome depends on technology, work design, power, labor-market conditions, governance, and whether people can shape what happens to them.

The cultivating-passion conversation is one part of the human side of AI-Native work. Architecture and human systems interact, but they are not two clean halves. A sound architecture can still go unused when it solves the wrong problem, violates trust, lacks workflow integration, or creates more burden than value. A willing workforce cannot rescue a technically unsafe system. Adoption is evidence to investigate, not obedience to demand.

A version of this work that goes right looks specific.

It can look like a Monday morning where people are excited about what they will build. It can look like a team meeting where someone shares an artifact they explored during paid learning time because they wanted to see what was possible. Weekend work is not the proof of passion; people should not need to donate private time to signal commitment. It looks like a manager genuinely curious about what direct reports are learning, including things the manager has never seen. It looks like a customer interaction in which the person cares whether the customer succeeds and AI removes friction without simulating care or weakening attention.

Over time, it may look like people who continued building portable skills, found meaning in some of the work, retained boundaries around the rest, and could leave with more capability than they brought. It may look like a company that adapted beyond the first wave of AI because learning conditions were repeatedly renewed, not assumed to remain in place for twenty years.

The people are the point of this version. AI can let them do something they could not do before, remove something they should never have had to do, or create a new burden that leaders must see. A leadership team that cultivates autonomy support, competence, connection, fairness, voice, and sustainable challenge is building work that some people may remember as the most engaging of their careers, and others may simply experience as good, fair work. Both are worthy outcomes.

The version is possible. The work described above improves its odds; evidence from the workforce determines whether it is actually being built.

Section 6.4

Up-Leveling How We Produce, Communicate, and Share

How does a new toolbox change the type of output a team generates, the way they communicate, and how they share knowledge?

The work humans do to communicate with each other, learn from each other, inform their customers, and report on what the company is doing still relies on a familiar set of tools. The Word document. The PowerPoint slide. The PDF report. The Excel spreadsheet. The email summary. The Notion page. Some are older than others, but together they have become the formats of knowledge work, the vocabulary the company uses to think out loud with itself and with the people it serves.

A new toolbox is now available. The tools in it can generate visualizations, build interactive interfaces, produce conversational analyses, write documents, design presentations, prototype applications, and synthesize data at a speed and level of access the older toolbox rarely offered. The older tools are still here. The shift is that they are no longer the only options, and the new options change what becomes practical for the company that uses them well.

Three things change when teams have access to the new toolbox. The first is the output itself, what the work product is and what it can do. The second is how teams communicate with each other when they have the new tools in the room. The third is how the company shares knowledge with itself and with its customers. These are different questions with overlapping

answers. Taking them in turn helps the leadership team see what is actually changing and what choices the company has to make.

A note before we begin. The phenomenon I am describing is not new in concept. The interactive document, the explorable explanation, the living artifact, have been on the horizon of knowledge work for years, championed most prominently by Bret Victor in his 2011 essay "Explorable Explanations" and the work that followed it. What is new is the falling cost of building these artifacts. A prototype that once required a developer and a substantial block of time may now be assembled in hours by a domain expert working with AI. Turning that prototype into a secure, accessible, maintainable production artifact still takes engineering and judgment. But the shift from rare jewel to everyday communication medium is underway, and the lower cost of exploration is driving it.

How the Output Changes

The first thing that changes is the work product itself, what it is, what it can do, what the reader can do with it.

Consider a familiar quarterly business review. It arrives as a slide deck, perhaps twenty-five to forty slides. The team spends days or weeks pulling data from different systems, arranging it into charts, writing commentary, and reviewing it through several rounds of edits before it reaches the audience. The audience receives the deck, reads it, asks questions, and the deck becomes a static record of what was discussed.

In the new toolbox, the quarterly business review can be a different kind of artifact. It can be a dashboard the audience can interact with, drilling into the underlying data, filtering by segment, asking questions in natural language that the system answers from the same data. It can be an explorable narrative that walks through the quarter and lets the reader manipulate the variables to test their own hypotheses. It can be a working model of the business that the audience can use to think about next quarter, not just review last quarter.

The shift from static to interactive is the most visible change. A static deck communicates what the team thought and provides a stable record that can be cited later. An interactive artifact lets the audience test the team's thinking. It may deepen engagement and improve the questions that follow, but interactivity is not automatically clarity. A badly designed dashboard can hide definitions as easily as a bad slide can. Natural-language answers need grounded data, visible

provenance, appropriate permissions, and a way to inspect the source. The aim is not motion on the screen. It is better thinking together.

This is not theoretical. Parts of professional services are testing a related shift: some fees moving from hours toward outcomes and some firms moving from advice toward implementation. Where clients buy an outcome, they may expect working prototypes, dashboards, or AI tools alongside the advice. The consultant experiment in Section 5.3 shows why each new form of work needs its own quality checks. The gains depended on the task. The lesson is not that AI makes analysis better. It is that the frontier is jagged, moves over time, and has to be tested against the actual work.

The same shift is beginning inside the enterprise, not just at the boundary with consultants. A marketing report assembled manually from six dashboards every Friday can become an approved system that pulls from the same sources and delivers an analysis decision-makers can query. A financial report can pair its stable, controlled record with a conversational layer for follow-up questions. A strategy document can include an interactive model that lets the reader test its underlying assumptions. These are possibilities, not automatic upgrades. Each depends on data quality, access controls, validation, and a clear source of record.

The pattern across all of these is the same. The work product is no longer just a document about the work. The work product is becoming a working artifact the audience can use. That is what is changing about the output.

How Communication Changes

The second thing that changes is how teams communicate with each other in the course of producing and sharing work products.

A common communication pattern in the older toolbox was sequential. One person produced something and handed it to another. The other person read it, responded, and sent something back. Work moved between people through documents, and every handoff introduced an opportunity for translation loss or delay. A team working on a strategy might spend two weeks moving a draft through five people before the next substantive conversation happened.

The new toolbox changes the pattern in several specific ways.

The first is that the lag between thinking and showing can shrink. A team member may go from a sentence to a rough visualization of that sentence in minutes. Instead of sending only a memo

that explains an idea, they can also send an artifact that demonstrates it. The conversation can move from imagining a hypothetical to inspecting several versions of it. That can compress hours of abstract debate, although verification and integration still take time.

The second is that the polished-final-output convention is starting to fray. In the old toolbox, you did not show something until it was ready. You worked on it in private until the version you would be willing to defend was the one you brought to the room. In the new toolbox, the cost of producing a presentable version is so low that bringing rough versions to the conversation is less embarrassing. The team can talk about what it is thinking earlier, with artifacts that make the thinking visible before it is finished. The risk is false completeness: a rough idea can now look final. Draft status, provenance, and unresolved questions have to be conspicuous. Used that way, the tools support more iteration, earlier in the process.

The third is that the asynchronous-synchronous distinction softens. With the old toolbox, you wrote a document and the other person read it later. With the new toolbox, an artifact can be something both people manipulate, separately or together, with an AI system helping each person engage with the other's contribution. The collaboration surface that Section 4.3 describes is one place this happens. A governed shared workspace can make the communication pattern available beyond a pair, but it is not appropriate for every conversation or every piece of data. Privacy, confidentiality, permissions, and the freedom to think without permanent capture still matter.

The communication change is harder to see than the output change because it shows up inside the team rather than in a deliverable. But it may be the more consequential of the two over the long run. A team that learns this pattern can produce stronger work faster, have better-grounded conversations, and converge on decisions with more substance. Whether it does so is an empirical question for the team, not a benefit to assume from tool adoption alone.

How Knowledge Sharing Changes

The third thing that changes is how the company shares what it knows. With itself, internally, and with the customers it serves.

Knowledge sharing has always had two modes. One is the broadcast mode, in which one person produces content (a report, a presentation, a manual) and many people consume it. The other is the conversational mode, in which two or more people exchange knowledge directly, often with the option to ask questions in real time. The broadcast mode scales. The conversational mode

does not. For most of business history, scale has required broadcast, which means knowledge sharing has been mostly one-directional.

The new toolbox changes this. An artifact that someone can explore at their own pace, ask questions of, and learn from in their own order is neither pure broadcast nor a human conversation. It is a third form: scalable and adaptive, but not reciprocal in the human sense. A company can produce a single governed artifact and have it serve hundreds or thousands of consumers, each following a path through the material rather than receiving it passively.

For internal knowledge sharing, this changes what is possible for institutional learning. The decisions the company made last quarter, the reasoning behind them, and the data they were based on can be made available as interactive artifacts that newer employees can explore, rather than buried in slide decks that few people revisit. But the generated answer must not quietly become the institutional record. The source decision, owner, date, version, evidence, permissions, and rationale have to remain visible and authoritative. With that discipline, reference architectures, strategic frameworks, and procedural knowledge can become artifacts that someone learning the company can interrogate, not just receive. This is the institutional knowledge layer from Section 4.1, surfaced in a form humans can use.

For external knowledge sharing, the shift can change the customer relationship. A customer who receives an interactive artifact may drill into the data that matters to them, test assumptions against their own context, and arrive at a more informed understanding. That does not mean the artifact verifies itself. Trust still depends on transparent definitions, traceable sources, disclosed limitations, and a stable or exportable version the customer can cite. Accessibility also requires an effective alternative for customers who cannot or do not want to use the interactive layer. Built with those protections, the customer becomes a more active participant in the analysis.

The implications for customer-facing knowledge work are substantial. Sales presentations that used to be one-way pitches can become scenarios the prospect explores. Customer success reports that used to be quarterly PDFs can pair a controlled snapshot with a live dashboard the customer revisits. Product documentation that used to be the place customers gave up can become an environment in which the customer learns the product. Companies that recognize the shift early can create more useful knowledge sharing. The advantage will come from usefulness and trust, not interactivity for its own sake.

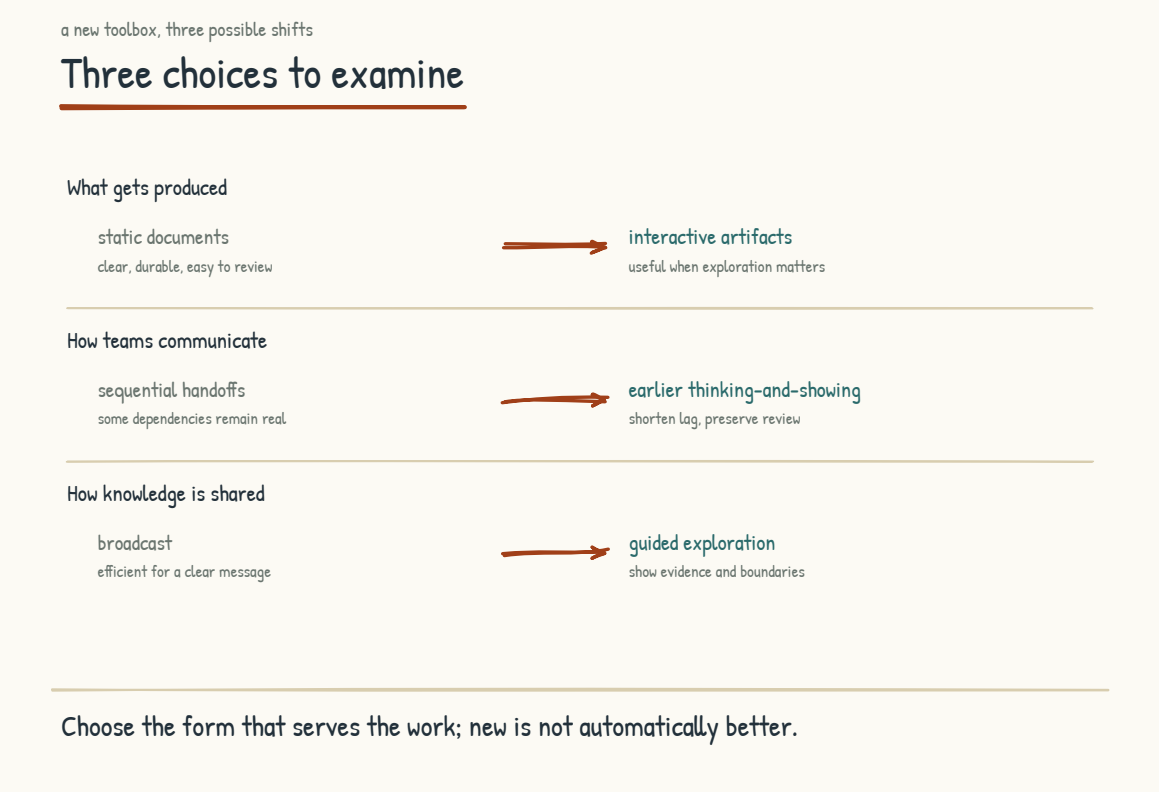

Figure 6.4a. The three threads of change, in parallel. Output shifts from static to living. Communication shifts from sequential to live thinking-and-showing. Knowledge sharing shifts from broadcast to exploration. The leadership team's response is the exploration-and-pathway pair.

The Exploration Question for the Leadership Team

The three changes above raise a question the leadership team has to address. What should the new standard be for knowledge work output in this company?

The default answer is no answer. The work products that get produced are whatever the workforce, individually, decides they should be. Some people learn the new tools fast and produce dramatically better artifacts. Some people stick with what they have always done. The output of the company becomes a patchwork, with high-fluency teams producing one quality of work and low-fluency teams producing another. The customer experience varies depending on which team serves them. The internal experience varies depending on which colleague produces the document. The company has not made a choice about what its standard is, and so the standard is defined by whoever happens to be producing the work.

At sufficient scale, that patchwork becomes difficult to sustain. A company that lets every important work-product standard emerge from individual initiative can create inconsistency

that customers notice and employees feel. The leadership team that wants to avoid this has to decide deliberately where shared standards matter and where variation is useful.

The work begins with opening the hood. The leadership team has to see what is now possible, firsthand. Spend a session with the tools that are now available. Watch a working artifact get built in twenty minutes. Produce one themselves. The exploration has to be embodied, not just discussed, because the gap between what people imagine the tools can do and what the tools can actually do is wide.

The work continues with looking at the company's current work products through new eyes. The quarterly review deck. The customer-facing reports. The internal training materials. The strategy documents. What would these look like if the team building them had the new toolbox and was using it deliberately? Some of the work products would not change much. The form fits the function. Others would change substantially. The form is now a constraint that the new tools have removed.

The work ends with a choice. The leadership team defines standards for the artifacts that matter, the ones that represent how the company communicates with itself and with the people it serves. Not one format for everything, but shared expectations for accuracy, provenance, accessibility, security, maintainability, and fitness for purpose. The choice is a stake in the ground: here is what good looks like now, and here is the support available to help the work get there.

The Pathway Question

The exploration sets the destination. The pathway is how the workforce gets there.

The pathway is the part that gets skipped. They explore. They get excited. They set a new bar. And then they leave it to individual employees to figure out how to clear it. This produces the same two-tier outcome as having no standard at all. The early adopters meet the new standard. The late adopters cannot. The work product quality across the company stays inconsistent, because the bar moved but the support for clearing it did not.

The pathway requires deliberate design. The leadership team has to decide how the workforce learns the new tools, how the early adopters help the rest catch up, how the company invests in the transition rather than expecting it to happen on its own.

A short list of what the pathway involves:

Work-Specific Training. Investment that is specific to the work the company actually does, not generic AI tool training. A finance analyst learning to build interactive dashboards with approved or sanitized financial data, inside an authorized environment, is learning something more useful than a finance analyst using the same tool in the abstract. The controls are part of the skill, not a compliance module added later.

Champions Paired with Learners: internal champions who are already producing the new kind of work product, paired with people who are still learning it. The fastest way to learn what good looks like is often to work alongside someone producing it. The company that identifies its early adopters, gives them protected time and recognition to mentor others, and does not turn their expertise into an unpaid second job is building the pathway.

Visible Working Examples. The strategy doc that became an explorable model. The quarterly review that became an interactive dashboard. The customer report that became a live artifact. When the workforce can see the new standard in action, on artifacts they recognize, they have a target.

Permission to Learn. The transition to the new toolbox is not free. It takes hours. The workforce that is told to produce the old output at the old pace while also learning the new tools will produce neither well. The leadership team that wants the up-leveling to happen has to make space for it, explicitly, in how the workforce spends its time.

A Designed Pace. Patience with uneven adoption: the pathway should follow the complexity and consequence of the work, not a generic calendar. A low-risk internal prototype and a regulated customer artifact should not move at the same speed. The leadership team that pushes too hard will produce backlash and unsafe shortcuts. The leadership team that pushes too softly will produce drift. The pace has to be designed, measured, and adjusted.

The pathway is the work. The exploration is the easy part. The pathway determines whether the company raises capability broadly or fragments into supported insiders and everyone else. It should make room for disability, role differences, legitimate non-use, and people whose work does not benefit from the new form. Collective progress is not forced uniformity.

What Is Not Changing

What does not change in this transition needs naming, because the argument otherwise risks producing the impression that AI does the work and humans just learn the tools.

What does not change is the need for accountable human contribution to give a work product its impact. The judgment about what to communicate and what to omit. The framing of the argument so the audience can follow it. The willingness to commit to a position rather than presenting every option neutrally. The storytelling that turns a collection of facts into something the audience can use. The presence of a human voice the audience trusts. AI can assist with each of these, and the division of work will change. But a person or institution still has to own the judgment, evidence, consequences, and final claim.

A leadership team that confuses the form of the work product with the substance will produce a workforce that makes very polished artifacts of weak arguments. The form will impress. The substance will not. The audience will notice the gap, eventually, and the trust the company has built on the substance of its past work will erode.

Some of the up-leveling is surface: speed, polish, and the range of forms available. Some is substantive: an interactive model can expose assumptions a static document leaves hidden. Neither guarantees good work. Hold the distinction clearly and the company gets the benefit of the new toolbox without losing the judgment the toolbox is supposed to serve. Confuse polish with substance and the workforce produces artifacts that look like AI-Native work but are actually slop with better visualization.

The three changes named in this section (what gets produced, how teams communicate, and how knowledge gets shared) are happening in parallel. They can reinforce one another. Progress in only one dimension may still create value, but it can also expose mismatches: an interactive artifact inside a sequential review process, or a shared workspace with no shared quality standard. Coherence comes from designing the connections among all three.

The shift, done well, is collective without demanding identical tools or outputs from every role. The organization adopts shared quality principles, and the leadership team explicitly designs the access, training, time, and alternatives people need to meet them. The shift, done poorly, produces a two-tier workforce where the AI-fluent pull ahead and the rest fall behind, and knowledge work that is inconsistent across the company in ways customers and employees both feel.

The leadership team can influence which outcome becomes more likely. The toolbox will continue to change, and work products will change with it. The question is whether the company changes deliberately, with shared principles and support, or by accident, in fragments.

The next section returns to a different layer of the same conversation, the systems-thinking layer that explains why some transitions reinforce themselves and others defeat themselves. How a company produces, communicates, and shares can become one of those reinforcing transitions when the pieces support one another. Over time, the difference should become visible in the usefulness, trustworthiness, and reach of the work. How modern the artifacts look is the least of it.

CONSIDER BEFORE YOU ACT

If you keep three things from Stage 6: 1) A value is what the team says; a practice is what the team does, and the practice defeats the value when the two diverge. 2) Design meaningful human involvement around consequence, context, rights, and responsibility: identify the moments, invest in them, protect them, reflect on them. 3) Watch the manager layer: AI can make distance look like attentiveness, and a clean dashboard can coexist with deteriorating conditions.

STAGE 7

The Systems View

The loops that reinforce, balance, and defeat AI-Native transitions.

Section 7.1

The Feedback Loops

Why do AI transformations keep producing the same problems?

Leadership teams often treat the AI-Native transition as a series of linear decisions. The plan reads: invest, deploy the tools, train the workforce, measure adoption, report progress. That framing produces a useful project plan, with milestones and dependencies and gates. It is necessary. It is not enough.

An AI-Native transition also behaves as a system. Each decision can feed loops that produce unintended outcomes, accelerate unseen dynamics, and sometimes defeat the goal itself. Leaders who cannot see the loops will struggle to explain why the transition keeps delivering the results it delivers. Leaders who learn to see them gain a better account of where intervention may matter.

Here are the major loops in an AI-Native transition. Treat the walk as a working diagnostic rather than an exhaustive catalog, one a leadership team can apply to their own company to identify what is reinforcing what, what is balancing what, and what is defeating what.

INTELLECTUAL LINEAGE

The vocabulary of reinforcing and balancing feedback loops, system archetypes, and places to intervene in a system is not new. Jay Forrester developed system dynamics at MIT in the late 1950s; his 1961 book Industrial Dynamics gave the field a rigorous modeling method. Peter Senge brought systems thinking to a broad leadership audience in 1990 with The Fifth Discipline, including eleven laws and archetypes such as limits to growth, shifting the burden, success to the successful, and tragedy of the commons. This section applies that tradition to the specific question of how AI transformations succeed and fail. The lineage is Forrester, Senge, and the wider field. The selection and application of these ten loops are mine.

Forrester, Jay W. (1961). Industrial Dynamics. MIT Press. / Senge, Peter M. (1990). The Fifth Discipline: The Art and Practice of the Learning Organization. New York: Doubleday/Currency.

THE STAKES

The claim that roughly seventy percent of corporate transformations fail is so widely repeated it has become folklore, and its definitions and evidentiary basis are contested. McKinsey surveys have reported that fewer than one-third of respondents regarded their transformations as both improving performance and sustaining the improvement, but that is a self-reported measure, not a universal failure rate. AI-Native transformation carries familiar transformation risks plus technical, workforce, and governance dependencies of its own. Understanding the loops does not guarantee success. It gives the leadership team a better way to diagnose why an intervention is compounding, being resisted, or producing a result nobody intended.

Robinson, Harry (2019). "Why do most transformations fail?" McKinsey Insights, July 2019. / Hughes, Mark (2011). "Do 70 Per Cent of All Organisational Change Initiatives Really Fail?" Journal of Change Management, 11(4).

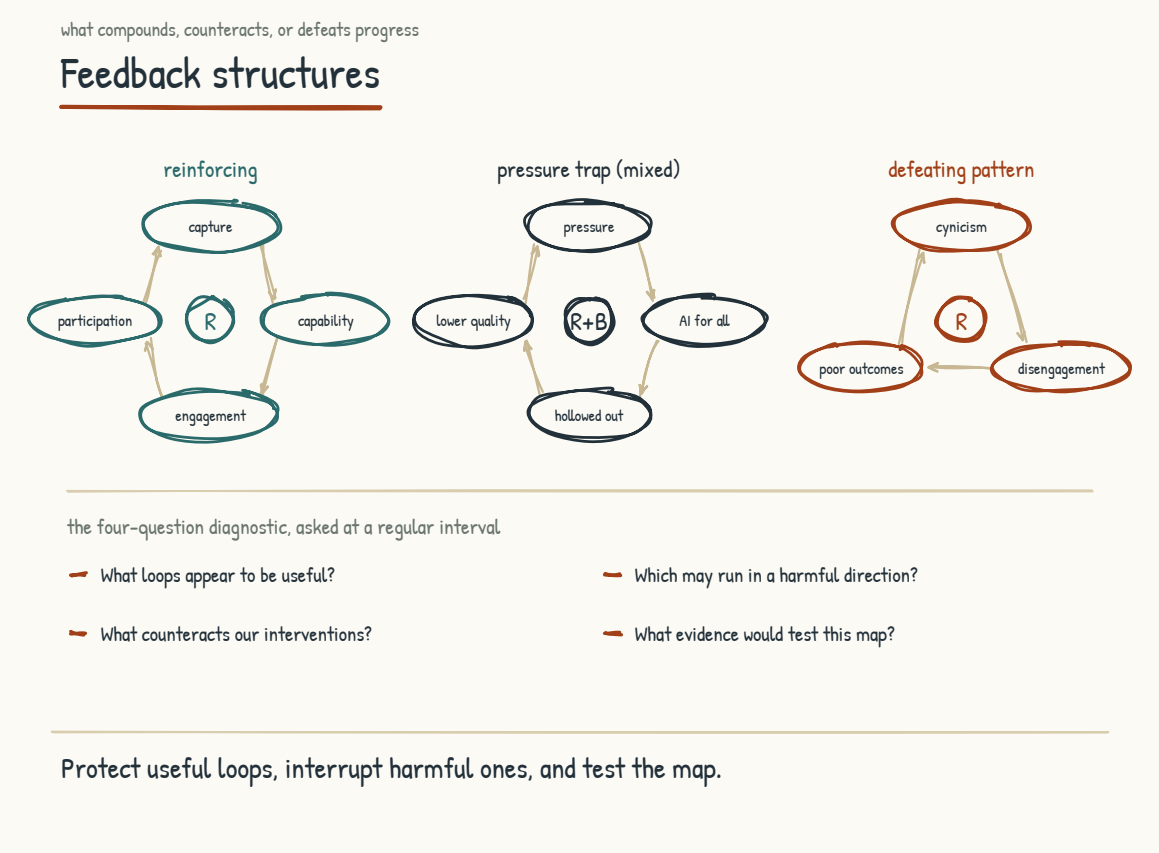

Figure 7.1a. Three representative loops in an AI-Native transition, with the four-question diagnostic for ongoing assessment.

The Two Kinds of Loops

A reinforcing loop is a loop where the output of the loop feeds back into the input, producing acceleration in the same direction. The loop compounds. Small initial states grow into large states over time. Reinforcing loops are powerful and they go both ways. The same loop that produces compounding success can also produce compounding failure, depending on which direction the loop is running.

A balancing loop is a loop where a change triggers effects that oppose that change, tending toward a goal, limit, or equilibrium. It can stabilize a system, but not necessarily at its old state and not regardless of intervention. A thermostat is the familiar example: a gap between the temperature and the setting triggers action that reduces the gap. In organizations, delays, inaccurate signals, competing goals, and changing constraints can make balancing behavior overshoot, oscillate, or settle somewhere undesirable.

Every leadership team is doing work that interacts with both kinds of loops simultaneously, usually without recognizing which is which. Acceleration may indicate a reinforcing loop, but frustration alone does not prove a balancing one. The discipline is to draw the causal links, mark whether each link moves in the same or opposite direction, identify delays, and test the explanation against evidence. The names are useful only if the mechanism underneath them is real.

The Reinforcing Loops in an AI-Native Transition

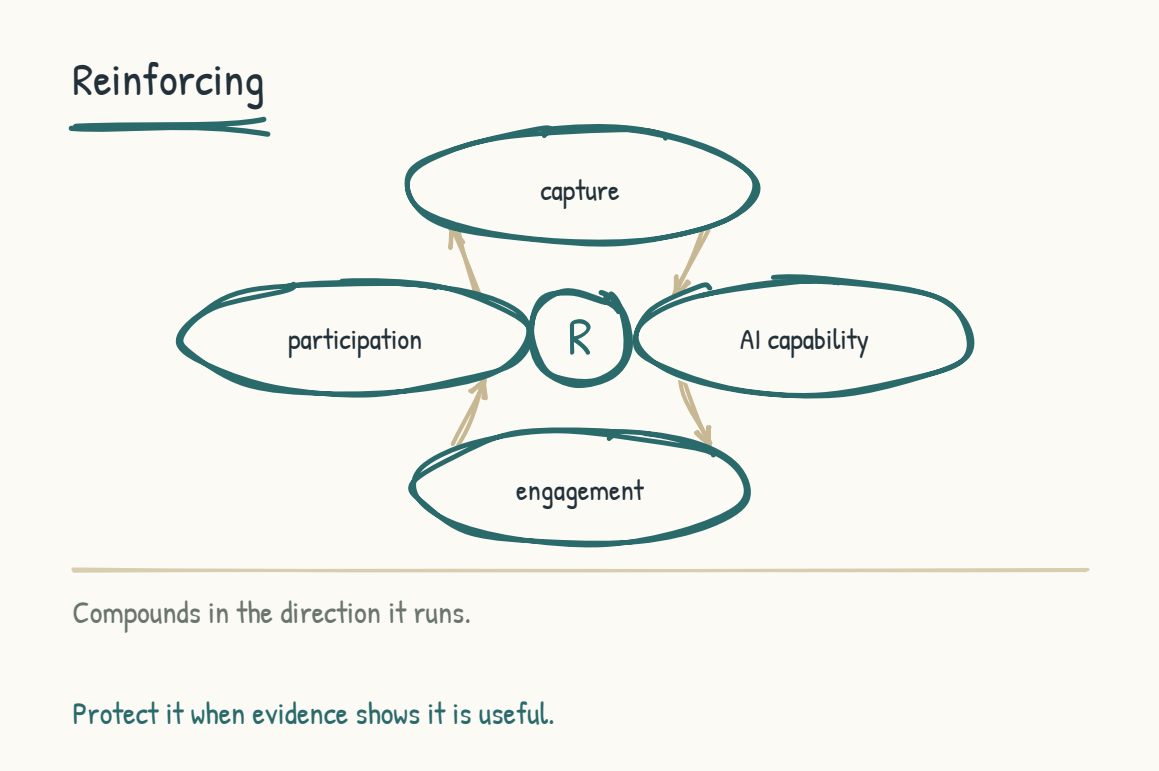

Figure 7.1b. A reinforcing loop: knowledge capture builds capability, which earns engagement, which feeds more participation.

The first major reinforcing loop is what I will call the knowledge-capture loop. The company invests in capturing useful institutional knowledge through the mix of graphs, documents, records, rules, and other structures described in Section 4.1. Better governed knowledge can improve AI capability for the relevant work. Useful capability can earn workforce engagement. Engagement can feed more willing participation in keeping the knowledge current. The cycle can compound into an advantage that is difficult to copy because it consists not only of stored information, but of years of maintenance, use, correction, and trust. The loop can also reverse: stale capture produces weak capability, which reduces use and leaves the knowledge staler still.

The second major reinforcing loop is the trust loop. The leadership team treats the workforce with honesty about the transition. Consistent words and actions can build trust. Trust can increase willingness to contribute knowledge under a fair covenant. That contribution supports AI capability that helps the workforce do its work. Visible, fairly distributed benefit can deepen trust. The loop compounds only while consent, credit, privacy, and the employment bargain hold. If capture is used against the people who contributed, the same links can turn a trust loop into a betrayal loop.

The third major reinforcing loop is the practice-and-craft loop. The team adopts the practices from Section 6.1. The practices can build craft, in the specific sense of the team developing a distinctive way of working that shows in distinctive output. Recognition of that craft can reinforce the practices because the team sees what disciplined work is earning. Over time, distinctive craft may become a source of advantage. Recognition can also corrupt the loop if it rewards appearance, heroics, or conformity instead of quality, learning, and contribution.

The fourth major reinforcing loop is the up-leveling loop. Environments that support up-leveling can grow capacities that unsupported environments are less likely to develop. Those capacities can show up as more distinctive work. Distinctive work may attract more demanding assignments, and those assignments can create the need and opportunity for further growth. The cycle is not automatic: overload, unequal access, or assignments without coaching can reverse it into burnout. The loop compounds when challenge is paired with support, time, feedback, and fair opportunity.

These four loops, when they are running well, can produce compounding capability that a tool deployment alone does not. The loops are not independent. They can feed one another: governed knowledge improves a useful capability; useful capability can earn engagement; fair participation can support trust; and trust can make continued contribution more likely. Every link remains conditional and should be tested.

The Loops That Resist or Limit the Transition

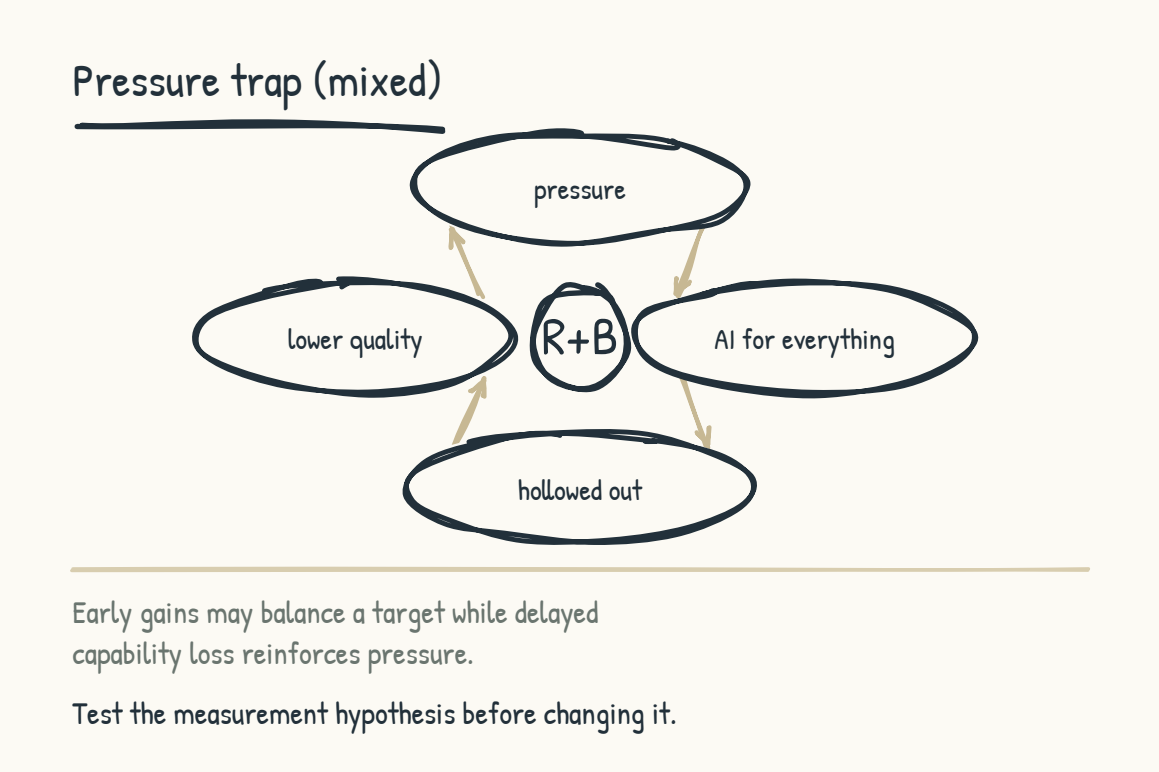

Figure 7.1c. A vicious reinforcing cycle: productivity pressure drives indiscriminate AI use, capability erodes, and disappointing performance triggers more pressure. The early productivity gain may also meet a target temporarily, creating a balancing effect and a delay before the damage becomes visible.

The first is the productivity-pressure structure. The leadership team measures the workforce narrowly on output. Workers respond rationally by using AI wherever it improves the measured number. Early gains may reduce the performance gap, a balancing effect. But indiscriminate use can also weaken required practice and learning from Sections 5.4 and 5.5. If capability erosion later reduces quality or sustainable performance, leadership may respond with more pressure, producing a vicious reinforcing cycle. What looks like one loop is an interacting structure with a delay: short-term correction can conceal long-term damage.

The leadership team that wants to break this loop has to recognize that the productivity measurement system is the engine of the loop. Without changing the measurement system, exhortations about up-leveling and human-AI collaboration will be absorbed by the loop.

The second is the risk-aversion structure. The consequences of a public AI failure are visible, while some costs of delayed learning are diffuse. Governance constraints reduce exposure to immediate harm, which is a legitimate balancing function. If those constraints also prevent

bounded experimentation, learning slows. Limited capability can then confirm the belief that the technology cannot be used safely, reinforcing avoidance. The diagnosis matters: governance is not the enemy. Undifferentiated governance that treats a sandboxed internal trial like an irreversible high-consequence deployment is the problem.

Breaking the harmful cycle requires neither abandoning controls nor absorbing unspecified risk. It requires risk-tiered learning: low-consequence, reversible experiments inside clear boundaries; stronger review as consequence rises; explicit stop conditions; and evidence that feeds governance back. Slow adoption has costs, but so does reckless adoption. The job is to make both visible and learn without transferring the downside to workers or customers.

The third is the status-preservation structure. An organizational hierarchy distributes authority, information, pay, and opportunity; it is not only a status game. An AI-Native operating model may flatten some layers while centralizing power in others. People who lose authority or standing may resist, sometimes through politics and sometimes because they see a genuine flaw the designers missed. If leadership treats every objection as proof of obstruction, it can intensify mistrust. If it treats every objection as a veto, the old structure persists. Either response can produce nominal AI-Native rhetoric and de facto pre-AI operations.

Breaking this structure requires examining incentives, decision rights, and evidence, not diagnosing people from their seniority or skepticism. Some conduct will require coaching, role redesign, or fair personnel action. Some resistance will reveal a defect in the transition. Leaders have to distinguish the two through transparent criteria, due process, and workforce participation. Political power should not confer immunity, and enthusiasm should not confer correctness.

The Defeating Loops

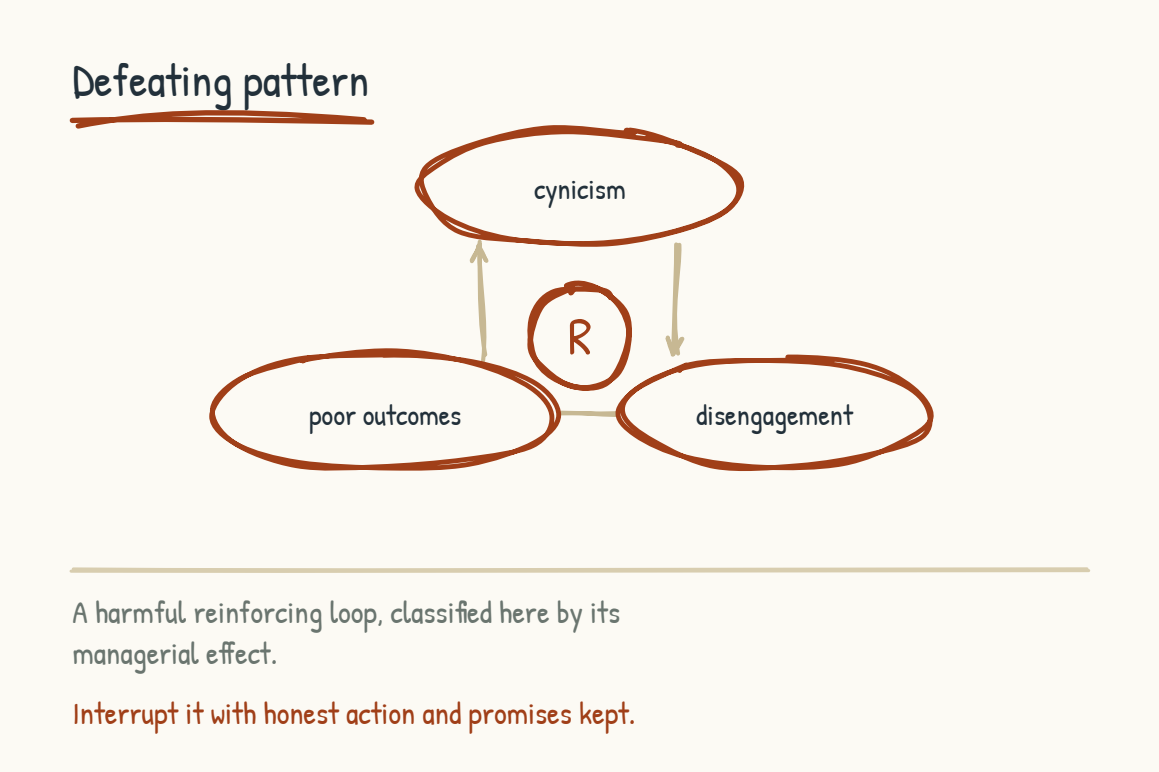

Figure 7.1d. A defeating loop: cynicism breeds disengagement, which degrades outcomes, which confirms the cynicism.

"Defeating" is not a third formal polarity alongside reinforcing and balancing. It is my managerial category for feedback structures that can defeat an AI-Native transition even when the leadership team has done the surface work well. Each structure still has to be mapped as reinforcing, balancing, or a combination of the two. These patterns are drawn from transformation research and from dynamics I recognize in practice.

The first defeating loop is the cynicism loop. The workforce becomes cynical about the transition. Cynicism can breed disengagement. Disengagement can degrade knowledge capture, practice adoption, and learning from AI use. Poor outcomes then confirm the cynical interpretation. The loop compounds in the wrong direction. A company may keep spending while returns weaken, and the weaker returns reinforce the cynicism that helped produce them alongside whatever technical or strategic failures are also present.

The cynicism loop is started by specific events. In January the CEO stands in the cafeteria and announces the AI strategy, a slide behind him, a promise of investment in every person in the

room. By October the steering meetings have gone from weekly to monthly to canceled, the training budget has moved to next year, and the slide has stopped appearing in his decks. There is no second announcement. The workforce can name the month it ended. Events like these seed the loop. Once the loop is running, it is hard to reverse, because the workforce has the evidence to interpret every subsequent action through the cynical frame.

The second defeating loop is the defensive-expertise loop. Workers may decline to contribute expertise for the reasons named in Section 5.8: fear, unfair extraction, privacy, workload, professional duty, or a rational judgment that the capture method is unsound. Thin or stale knowledge can leave weak AI capability. Weak capability can then confirm that participation is not worth the cost, and the cycle reinforces itself. But weak AI performance may also be an accurate signal of model limits or a poor use case. The diagnosis cannot begin by assuming the worker is the defect.

Interrupting the harmful form of this loop requires the leadership team to address the conditions producing nonparticipation. The human reckoning and covenant from Section 5.2 are part of that work: voluntary and appropriately compensated contribution, consent and privacy boundaries, credit, workload protection, professional review, and visible benefit. Better engineering and the willingness to abandon an unsuitable use case matter too. The goal is trustworthy participation, not extracting everything people know.

The third defeating loop is the metric-displacement loop. The leadership team starts the transition with the right values. The values are translated into metrics. The metrics drive behavior. The behavior generates outcomes the leadership team did not intend, because the metrics did not actually capture the values. The leadership team responds by adding more metrics. The new metrics drive more behavior, and the behavior generates more unintended outcomes. The loop compounds. The transition becomes an exercise in metric optimization that has lost touch with the underlying values that the transition was supposed to serve.

Companies that break this loop accept that some things they value cannot be measured directly and that every proxy can be gamed. They do not abandon measurement. They combine a small set of indicators with qualitative evidence, disaggregated effects, counter-metrics, direct observation, and periodic review of whether the measure still serves the value. They invest in the practices and moments that produce the values, then use evidence as feedback rather than a substitute for judgment.

How to Read a Transition Through the Loops

A leadership team that wants to apply this framework to their own transition can do so with a simple practice. At a regular interval, the team asks four questions.

What loops are running well. Which of the reinforcing loops are compounding in the right direction. What is producing the compounding. How do we protect and accelerate the loops that are working.

What loops are running in the wrong direction. Which of the reinforcing loops are compounding in the wrong direction. Which loops have started and need to be interrupted before they become self-sustaining.

What balancing effects or limiting structures are counteracting our interventions. What goal, constraint, delay, or competing incentive is producing the response. Is that response protective, harmful, or both. What would we need to change, and what protection would we lose if we changed it.

What defeating loops are present. Cynicism, defensive expertise, metric displacement. Any others specific to our company. What is feeding them. What would interrupt them.

Asking those four questions regularly is the systems-thinking discipline at the leadership-team level.

The next section addresses the deepest layer of the systems view. Where change actually takes hold.

BRING THIS INTO THE ROOM

Take the four-question diagnostic to your leadership team. Block ninety minutes. Bring whiteboards. Answer the questions for the last three months, then for the next three months. What matters is less the questions themselves than the practice of answering them in front of each other, regularly, with honest specificity. The loops you cannot see together you cannot intervene on together.

The table below compresses the ten loops of this section into one view a leadership team can hold during the four-question diagnostic. The full descriptions, and the interventions in context, are in the pages behind you.

The Ten Loops at a Glance

Knowledge-capture loop (reinforcing). Signs it is running well: workers see the AI is useful for relevant work; voluntary participation and maintenance grow; corrections improve the governed knowledge base over time. To strengthen it: invest in the appropriate mix of graphs, documents, records, and rules; preserve provenance and permissions; and keep the resulting capability visibly serving contributors as well as the company.

Trust loop (reinforcing). Signs it is running well: leadership's words and actions align; people contribute under a fair covenant; benefits flow visibly back to contributors. To strengthen it: sustain honesty, consent, credit, privacy, and fair distribution. Watch for reversal: using captured knowledge against contributors can turn trust into betrayal.

Practice-and-craft loop (reinforcing). Signs it is running: the team's output is distinctive in kind, not just quantity; internal and external recognition of the craft; the team can see what the practices earn them. To strengthen or break it: adopt and hold the human practices through the awkward phase; recognize the distinctive work the craft produces.

Up-leveling loop (reinforcing). Signs it is running well: supported challenge develops capability; distinctive work leads to further learning opportunities. To strengthen it: pair harder assignments with coaching, time, feedback, access, and fair opportunity. Watch for reversal into overload or burnout.

Productivity-pressure structure (mixed). Signs it is harmful: short-term measured output rises while quality, learning, or sustainable capability erodes; delayed weakness triggers still more pressure. To change it: map both the early balancing effect and the later reinforcing cycle; broaden the measurement system and protect required practice.

Risk-aversion structure (mixed). Signs it is harmful: controls reduce immediate exposure but also prevent bounded learning; limited capability then reinforces avoidance. To change it: use risk-tiered, reversible experiments, stronger review as consequence rises, explicit stop conditions, and evidence that updates governance.

Status-preservation structure (mixed). Signs it is running: nominal AI-Native rhetoric coexists with unchanged decision rights, or every objection is dismissed as

politics. To change it: examine incentives and evidence; use transparent criteria, workforce participation, due process, role redesign, coaching, and fair personnel action where warranted.

Cynicism loop (defeating). Signs it is running: AI spend produces diminishing returns; every leadership action is interpreted through the cynical frame; disengagement degrades capture and learning. To strengthen or break it: prevent the seeding events: quiet retreats from announced strategy, cost cuts dressed as strategy, broken development promises. Once running, it is hard to reverse.

Defensive-expertise loop (defeating). Signs it is running: participation remains thin, the knowledge base stays weak, and weak capability confirms that contribution is not worthwhile. To interrupt it: address fear, extraction, workload, privacy, credit, and professional duty through a fair covenant; improve the engineering; and abandon use cases that remain unsuitable.

Metric-displacement loop (defeating). Signs it is running: behavior optimizes to metrics that no longer capture the values; unintended outcomes are answered with more metrics. To interrupt it: accept incomplete coverage, combine indicators with qualitative and disaggregated evidence, use counter-metrics, and review whether each measure still serves the value.

Section 7.2

The Iceberg: Finding Where Change Takes Hold

What actually makes change last?

The iceberg model is a teaching device from the systems-thinking education tradition. No single author owns it; generations of systems educators have used it to connect visible events with patterns, structures, and mental models beneath them. Donella Meadows offered a related but more detailed argument in "Leverage Points: Places to Intervene in a System" (1999): interventions differ in how much they can change a system, with a system's shared assumptions near the top of her list and the capacity to question those assumptions above them. She also warned that the most powerful points of intervention are difficult to change and easy to push in the wrong direction. The iceberg is therefore a prompt for inquiry, not a proof that every event has one deep cause or that the bottom layer is always the place to act.

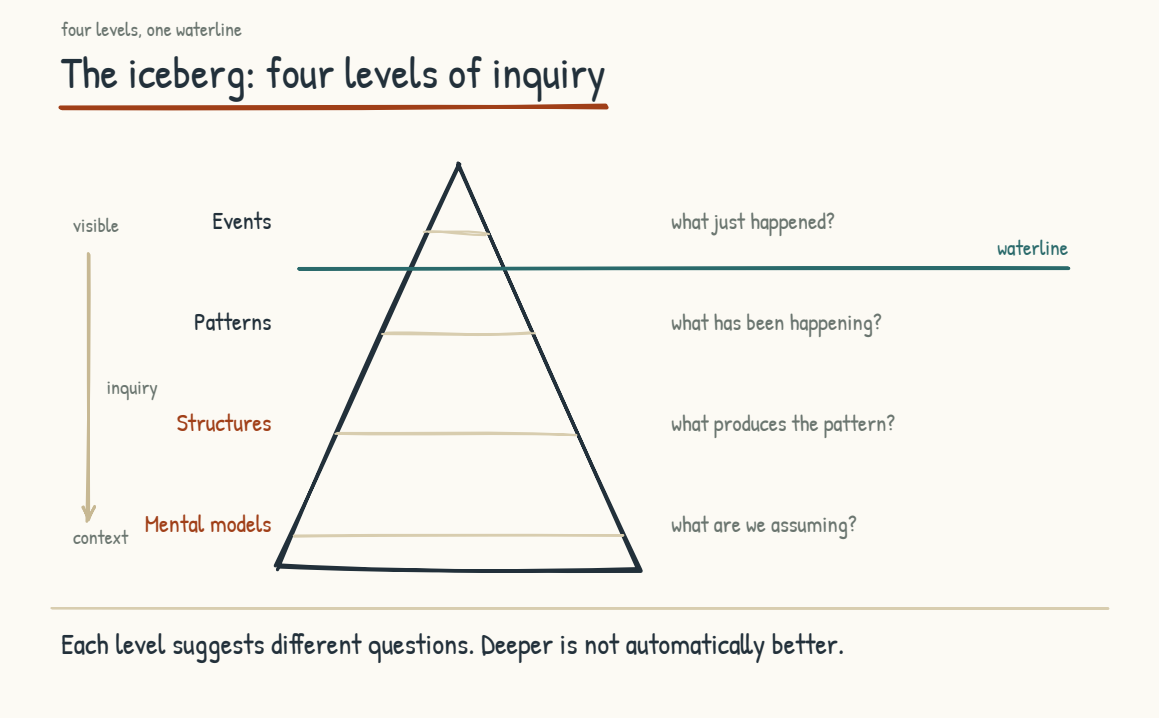

Figure 7.2a. The iceberg model. Events can reveal patterns, structures, and mental models. Deeper inquiry may reveal more powerful interventions, but the levels interact and urgent events can still require immediate action.

The Four Levels

The top of the iceberg is the level of events. Events are the visible occurrences in the system. They are what shows up in the news, in the meetings, in the performance reviews. A worker quits. A customer complains. A project goes over budget. An AI initiative stalls. Events are concrete and they are real. They are also, in the systems view, surface phenomena that emerge from the levels below them.

Just below events is the level of patterns. Patterns are the recurring shapes that events take over time. The third worker has quit this team in six months. The same complaint keeps coming back from different customers. Projects in this division consistently go over budget. AI initiatives consistently stall after the first eighteen months. Patterns are visible to people who pay attention to events over time, and patterns are where the diagnostic conversation starts to become useful. The pattern tells you that the event is not a one-off.

Below patterns is the level of structures. Structures are the organizational, cultural, technical, and incentive arrangements that shape the patterns. A team that loses three workers in six months may have an incentive system that rewards short-term output, a reporting line that leaves the team without advocacy, or a promotion system that fails to recognize its work. It may

also be responding to labor-market or personal factors outside the team. The model asks you to investigate the structure rather than infer it from the pattern. When a structure is contributing, addressing events one by one will not be enough.

The bottom of the iceberg is the level of mental models. Mental models are assumptions, beliefs, and frames that help produce and justify structures: workers are interchangeable; output is the only thing that matters; culture sits on top of operations rather than emerging partly from how operations are designed. These models can be hard for their holders to see because they are lenses as well as objects of analysis. But causality runs both ways. New rules, tools, experiences, and power arrangements can change mental models, just as mental models can shape structures. Durable change often requires attention to both.

The practical lesson is to look for places to intervene at more than one level. Event-level intervention can stop immediate harm. Pattern analysis can show whether the occurrence is recurring. Structural intervention can change conditions that generate the pattern. Mental-model work can change the frames that make those structures seem natural. Power belongs in the analysis too. A mental model may persist not because nobody has questioned it, but because it benefits people who control resources and decisions. In that case, insight without changed decision rights will not move the system. Deeper is not automatically better, and slower is not automatically more durable. The strongest response may combine levels: protect people now, learn from the pattern, redesign the structure, challenge the governing assumptions, and then test whether the expected pattern actually changes.

An AI-Native Example

Here is the iceberg in an AI-Native company.

At the event level, here is what happens. On a Tuesday morning, the senior analyst who has been with the company for twenty-two years puts fifteen minutes on her manager's calendar and hands him a printed letter, one page, signed. She is taking a role at a competitor. The exit interview takes place in a glass conference room off the lobby. It is professional. The HR partner types while she talks, and she answers every question. She wishes everyone well. On her last day she packs one box, a coffee mug and the row of methodology binders she wrote herself. The company holds a leaving gift. The position is posted the following Monday.

At the pattern level, here is what becomes visible when the company examines the record. She is the third senior analyst to leave in the past year. Departures are concentrated in a specific

division, under a specific manager, in a specific kind of work. That concentration is a signal, not yet a causal conclusion. It tells the leadership team to compare time periods, teams, roles, stated reasons, labor-market conditions, and the experiences of people who stayed. Each departure had been treated as a one-off; together they create a question the company can no longer dismiss.

At the structure level, suppose that inquiry finds this: the division has been deploying AI tools aggressively over the past two years without the human reckoning from Section 5.2, the practices from Section 6.1, or the moments-that-matter framing from Section 6.2. AI now handles much of the analytic work that used to define the senior analyst's role, while the analyst is left with oversight responsibilities that do not engage her capabilities and a recognition system that does not reward her accumulated judgment. In this case, the structure has converted a skilled craft role into an AI-supervision role without naming, negotiating, or supporting the change.

At the mental-model level, the inquiry may reveal a frame that helped produce the structure: AI is primarily a productivity tool to be deployed against existing roles. That frame treats work as tasks to allocate between AI and humans, with the goal of maximizing the AI share. It misses the way work can develop craft, judgment, relationship, and meaning over decades. It may treat the senior analyst's expertise as overhead rather than a capability to use and renew. The mental model does not prove the departure, but it helps explain why the organization repeatedly made design choices that reduced the role.

The leadership team has several options for intervention, and the options correspond to the levels of the iceberg.

The event-level intervention may include a counteroffer, but it should begin by learning what the analyst wants and whether repair is possible. A counteroffer can retain someone; it can also arrive too late or leave the reason for leaving untouched. Either way, the company still has to investigate the wider pattern.

The pattern-level intervention might strengthen exit interviews, introduce voluntary stay interviews, and examine retention by team, role, tenure, pay, demographic group, and work design. That can reveal where to act and may improve retention. It becomes theater only if the company gathers the evidence and refuses to change the structures the evidence identifies.

The structure-level intervention is to redesign the work with the people who do it. The senior analyst role might move toward higher-consequence judgment, client relationships, method

development, teaching, quality assurance, or other work in which her expertise matters, including work where AI assists rather than disappears. Recognition, authority, workload, and compensation have to follow the redesign. Structural change can be durable. It is more likely to hold when incentives and mental models support it, and when outcomes are reviewed rather than assumed.

The mental-model-level intervention is the work I have been recommending. The leadership team examines, with help and evidence from the workforce, the frame behind its choices. It may come to see the senior analyst's expertise as a capability to amplify and renew rather than a cost to reduce. Its view of AI may shift from productivity extraction toward capability amplification. That shift can support different structures, patterns, and events, but belief without redistribution of authority, resources, and incentives is only a workshop result. The company still has to decide what analysts will own, how their judgment will be developed, who can challenge an automated result, and how value created with their expertise will be shared. Durable change comes from the model and the operating choices changing together.

Figure 7.2b. The senior analyst exit examined at all four levels. Each deeper level adds a hypothesis to test; durable change may require coordinated action across the iceberg.

Why Change Management Can Get Trapped at the Top of the Iceberg

Corporate change management can become trapped in the top half of the iceberg. Policies, communication campaigns, training programs, reorganizations, and performance-management updates are visible, fundable, and easy to report on quarterly. Yet change-management practice is broader than this caricature: participatory, systemic, and evidence-based approaches do engage structures, assumptions, and power. The problem is not the discipline. It is using surface activity as proof of change while leaving the conditions that reproduce the old behavior untouched.

There are several reasons for this pattern. Deeper interventions can be harder to measure and therefore harder to justify in a quarterly cycle. They may require leaders to examine their own mental models, work for which many have had little preparation. Effects can take longer to appear, which is politically risky when the environment demands short-term results. And the effort may require facilitation, systems analysis, and power-aware participation that the particular change program was not designed to carry.

Some corporate change efforts I have watched produce activity rather than transformation. They look real from the outside. They consume real budgets and produce real deliverables. But if behavior, decision rights, incentives, capability, and outcomes remain unchanged, the activity is not evidence that the system moved.

Transformation efforts are more credible when they connect visible activity to deeper work and measure both. They handle events, study patterns, redesign structures, and examine the frames behind decisions. Communication and training still matter: people cannot enact a structure they do not understand or lack the skill to use. Reorganization can matter when authority genuinely moves. Policy can matter when it changes what people may do and what happens when the rule is violated. The question is not whether an intervention is visible. It is whether the causal account is credible and the evidence shows the system changing. The deeper work may look less impressive in a quarterly update, but it should still generate evidence: changed decisions, changed resource flows, changed behavior, and outcomes that persist after attention moves elsewhere.

Why Collaborative Co-Design Works

What I have been recommending is collaborative co-design at the mental-model level. Why this works matters for understanding where the leadership team needs to invest.

Mental models rarely change merely because someone declares them wrong. A leadership team will interpret the challenge through what it already believes, and may reject it or accept its vocabulary without changing its decisions. Direct instruction can still supply evidence, concepts, and challenge. It is usually insufficient by itself.

Mental models can change through experience, disconfirming evidence, reflection, dialogue, and changes in practice. A structured exercise can help a leadership team look at its operation from another frame, notice where its current explanation does not fit, and construct a better account collaboratively. The test of the shift is not what the team says in the room. It is whether later choices, resource allocations, and responses to contrary evidence change.

This is what collaborative co-design can do. The leadership team and people affected by the system examine the company from different positions. Skilled facilitation can surface assumptions, conflict, evidence, and power that an executive-only conversation misses. The group builds and tests a shared working model without forcing false consensus. Participation can increase understanding and ownership, but it does not guarantee either; decision rights, dissent, and unresolved tradeoffs should remain explicit.

A leadership team may begin this work itself, and an external facilitator can help people see assumptions, power dynamics, and blind spots that are hard to surface from inside. A consultant's recommendations can also contribute when the evidence and reasoning are open to challenge. The key is not mandatory dependence on a facilitator. It is a collaborative, structured process in which the people making and living with the decisions participate rather than merely receive. Arc5 is the method I use for that work (Section 9.1), and its results should be judged by the decisions and outcomes it changes.

The Connection to the Rest of the Book

The operating system from Stage 3, the architectures from Stage 4, the human work from Stage 5, the practices and moments from Stage 6, are all interventions at the structure level. They are real structures that generate real patterns and real events. The work of designing them well is necessary work.

The work is not sufficient on its own. Structural choices also express mental models. Choosing a graph, document retrieval, relational data, rules, or a combination is downstream of a view about what institutional knowledge is and what the use case requires. Choosing a collaboration surface, individual tools, or both reflects a view of how the organization works. Choosing to do

the human reckoning reflects a view of what the workforce is and what is owed to it. Choosing moments for meaningful human involvement reflects a view about consequence, rights, responsibility, and value.

A leadership team that adopts the structures I have described without examining the assumptions behind them may watch those structures erode. Incentives and later decisions can pull the organization back toward the old design. The reverse also matters: adopting new structures can expose weaknesses in the old mental model and help people revise it. Change travels both directions through the iceberg.

A leadership team that examines its mental models while changing structures has a better chance of producing coherence. Structures can reinforce the new frame; experience with the structures can refine the frame; observed patterns can show whether either is working. That does not make a transformation immune to leadership changes or market shifts. It gives the company feedback, institutional memory, and a stronger basis for adapting without quietly reverting.

Several sections of this book have told you to measure differently without saying what to measure. The next section pays that debt, and it builds the instruments out of the loops and the iceberg rather than out of a list of KPIs.

CONSIDER BEFORE YOU ACT

Map your current AI-Native efforts against the iceberg. How many are operating at the events level, responding to incidents and announcements? How many at the patterns level, studying recurring behavior? How many at the structures level, redesigning processes, incentives, decision rights, and systems? How many at the mental-model level, testing what the leadership team believes and assumes? Then ask what each level needs from the others. If effort is concentrated above the waterline, investigate what keeps producing the pattern. If it is concentrated below, make sure immediate harms and operational realities are not being intellectualized away.

Section 7.3

Measuring the Transition

What can the numbers tell you about whether your transition is real, and what can they not tell you?

Several sections of this book have told you to measure differently. Section 4.2's fifth mandate says to measure what compounds, not what produces quarterly numbers. Section 5.5 names the right measurements as a condition for up-leveling. Section 5.6 says to measure the trajectory of the workforce, not only its output. Section 6.1 says the leadership team has to measure the practices, in some form. What none of those sections said is what to measure. This section pays that debt.

The candidate instruments in this section follow from the loops in Section 7.1 and the iceberg in Section 7.2. They are a practitioner measurement design, not a validated diagnostic scale. Without a causal frame, measurement becomes a list of KPIs detached from the system it is meant to describe. With one, measurement can become more useful: evidence about which direction a hypothesized loop may be running and where an intervention may be landing. The numbers do not establish the loop by themselves.

Why the Default Dashboard Fails

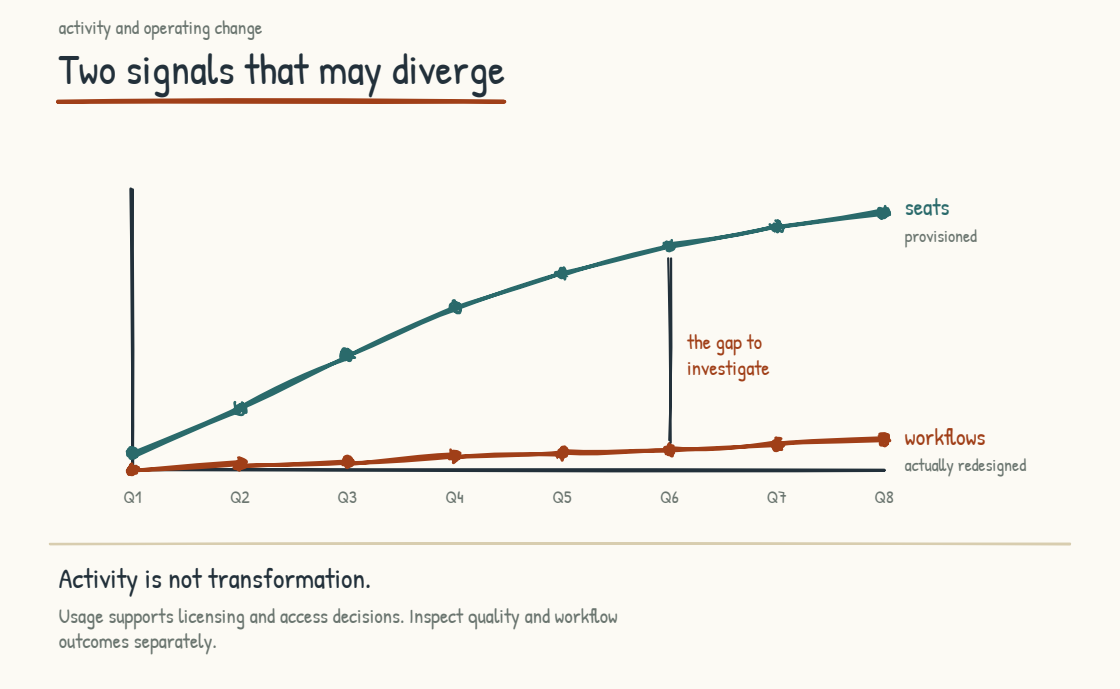

Figure 7.3a. Seats climb while the structure underneath stays exactly as it was. Only one of these lines usually reaches the slide.

A common AI dashboard measures adoption. Seats provisioned. Weekly active users. Prompts per user. Percentage of employees who have used the tool this month. Departments ranked by usage. The dashboard is easy to build and report, and its lines often rise during a rollout.

The dashboard measures activity. It can answer useful operational questions about access, support, and cost, but it cannot tell you whether the work improved or which direction the loops from Section 7.1 are running. Used as a target, it can reward exactly the behavior this book cautions leaders to avoid. A worker who routes every task through a model and ships the first answer can score above a colleague who uses AI selectively, verifies the result, and protects a capability that still matters. Adoption is an input signal, not the outcome.

This is the Goodhart problem, and it deserves to be named plainly. Once a measure becomes a target, optimization pressure can degrade its relationship to the thing it represented. Adoption metrics may be useful evidence that access and use are changing. They are weak evidence that a company is becoming AI-Native, and weaker still once pay, status, or quotas depend on them. They can certify that the chat window is open. They cannot certify that judgment improved.

Section 7.1 named the metric-displacement loop as one of the defeating loops: values translated into metrics, metrics driving behavior the values never intended, more metrics added in response. The honest reader will notice a tension. That loop warned against exactly the exercise this section performs. The response is not a magical set of better metrics. It is a measurement practice: use several forms of evidence, state what each signal cannot show, protect privacy, avoid individual quotas, look for disparate effects, and revisit the causal model when the evidence does not fit. The candidate signals below cover only part of what matters.

GOODHART'S LAW

The observation that a measure degrades when it becomes a target originates with the economist Charles Goodhart, writing about monetary policy in 1975. The formulation most people quote, "when a measure becomes a target, it ceases to be a good measure," was coined by the anthropologist Marilyn Strathern in a 1997 paper on university audit culture. Both versions describe the same failure: the proxy gets optimized, the underlying value does not.

Goodhart, C. A. E. (1975). "Problems of Monetary Management: The U.K. Experience." Papers in Monetary Economics, Vol. I. Reserve Bank of Australia. / Strathern, Marilyn (1997). "'Improving Ratings': Audit in the British University System." European Review, 5(3), 305-321.

Measure the Loops, Not the Events

One useful account of the transition is the state of the loops. The leadership team's instrument panel should therefore ask: what evidence would we expect if this loop were running in either direction, and what other explanation could produce the same signal? Here are candidate signals a leadership team can adapt and test. Keep the panel small enough to use. It should read the system, not rank individuals, and access to underlying data should follow purpose, consent, and need.

The knowledge-capture loop.

Capture and reuse. Contributions to the governed knowledge base per team per month and, more telling, how often approved knowledge gets drawn on in subsequent work. Separate useful reuse from repeated retrieval of stale or incorrect material. A graph, repository, or library that is written to but never read from may be a compliance exercise wearing an architecture costume.

Coverage and freshness. Of the processes in your operating system inventory (Section 3.2), what share have current knowledge behind them, and how stale is the rest? A coverage map with dates on it is one of the most clarifying artifacts a leadership team can commission.

Breadth of contribution. How many distinct teams contributed this quarter? A loop fed by the same three enthusiasts every quarter has not started compounding, however impressive their output looks.

The trust loop.

Commitments made versus commitments kept. The covenant from Section 5.2 generates promises. Track them, visibly, the way you would track any other obligation. The workforce is already keeping this ledger informally. Keeping it formally tells them you know that.

Who can participate safely. At an aggregated level, are people across roles, tenure, locations, and levels able to contribute voluntarily, or is participation concentrated among people with unusual security, access, or spare time? Do not interpret nonparticipation as disloyalty or expose small groups through the data. Participation is one signal of the conditions around trust, not a direct measure of trust itself.

The candor pulse. Ask, directly and regularly, whether leadership's account of the transition matches what people see on the ground. Protect anonymity where it can genuinely be protected,

state who will see the data, report only groups large enough to avoid reidentification, and pair the trend with listening and action. A pulse can reveal a gap; it cannot explain the gap alone.

The practice-and-craft loop.

Practice review. Are the practices from Section 6.1 actually happening? A standing review, qualitative is fine, of what is being practiced, what is getting skipped, and what is in the way. Section 6.1 already made this a condition; here it becomes an instrument.

Shared-asset vitality. Contribution and reuse rates on the collaboration surface: prompt libraries, shared work products, team playbooks. A vital surface shows both creation and reuse. A dead one shows neither, whatever the login numbers say.

Work quality over time. With appropriate access and a stable rubric, sample comparable work products and ask whether they are improving in ways the team and its customers can name. Review aggregate patterns and calibration, not a covert employee ranking. Movement may be evidence of the craft loop compounding; changes in task mix, staffing, or customer demand are competing explanations to examine.

The up-leveling loop.

Supported-stretch share. What fraction of the workforce has a development opportunity near the edge of its current capability, chosen with the person and paired with time, support, and feedback? Aggregate the answer and audit whether opportunity is distributed fairly. A manager who cannot discuss development may need support; the absence of an instant number is not proof the manager does not know the team.

Protected time honored. At a team level, compare protected learning or deep-work time promised with time actually available, using the least intrusive data that answers the question. Calendars can help, but they also contain private information and do not reveal the quality of attention. Ask people whether the protection worked.

Active mentorship. Measure whether voluntary mentoring relationships receive time and support and whether participants find them useful. A meeting count alone can turn mentorship into attendance theater.

The resisting and defeating structures. These do not all have a single needle that should be low. Look for the mechanism, the protective function, the harmful side effect, and the direction of change.

Shadow-to-official estimate. Using privacy-preserving security and survey evidence, estimate how much AI use occurs through approved tools and how much does not (Section 4.3). Falling shadow use can mean the approved path is better, a prohibition is working, use has moved elsewhere, or people have become less candid. Investigate before assigning meaning.

Decision time for experiments. Track how long comparable proposals spend in review, including time waiting on the proposer and time required for legitimate risk work. Rising unexplained latency may indicate an overbroad approval process. Falling latency without adequate review may indicate the opposite failure. The aim is proportionate governance, not the shortest path to yes.

Contribution change. At a sufficiently aggregated level, notice teams that used to contribute approved knowledge or practice notes and have stopped. This may be an early signal of cynicism, or it may reflect workload, completed work, changed relevance, privacy concern, or a better channel. Use the signal to ask, not to diagnose.

The Four Altitudes of Measurement

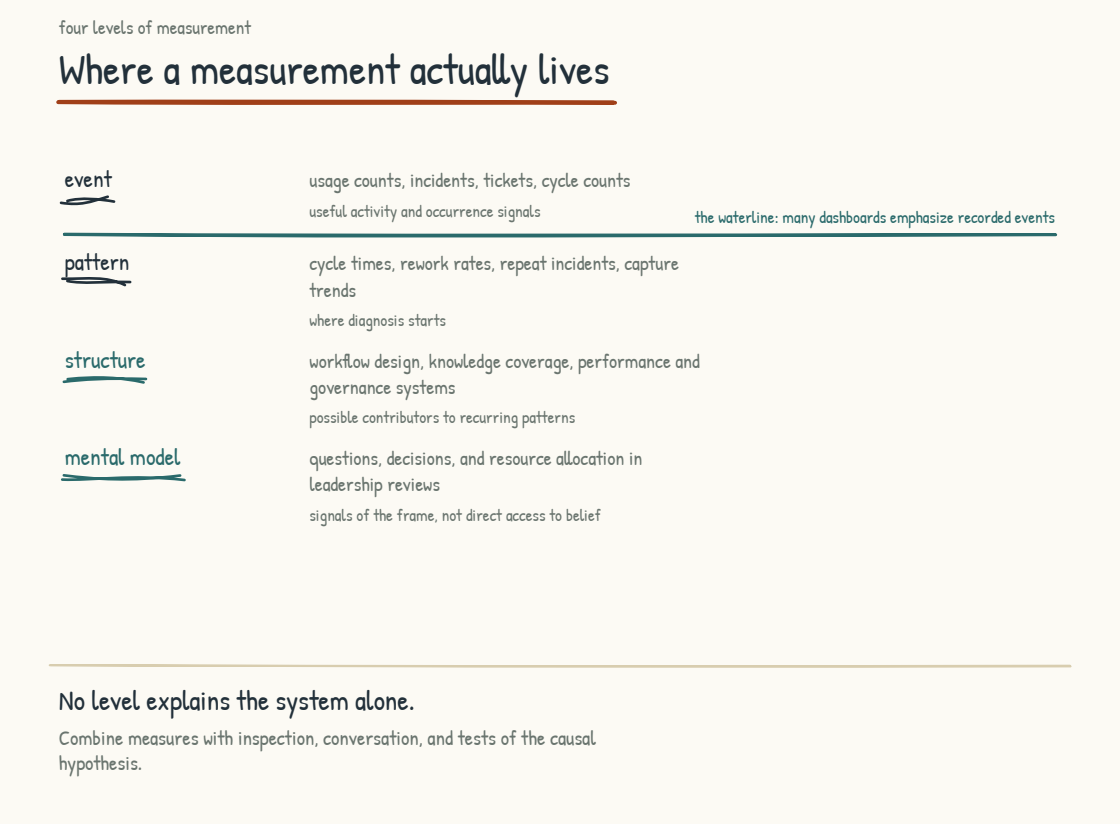

Figure 7.3b. Where a measurement actually lives. Many dashboards concentrate on the top band because events are easiest to count; evidence from every altitude still requires interpretation.

The loop panel offers evidence about direction. The iceberg from Section 7.2 offers a way to classify that evidence. Measurements can point toward each of its four levels, and many dashboards concentrate above the waterline because events and patterns are easier to count.

Event metrics. Usage counts, incident counts, tickets, cycle counts. Often inexpensive and operationally necessary. They tell you what was recorded as happening, not why it happened; combined with context and a causal model, they may still help anticipate what comes next.

Pattern metrics. Workflow cycle times. Rework rates. Repeat-incident rates. Quarter-over-quarter trends in capture and reuse. Patterns are where diagnosis starts, because a pattern establishes that the event was not a one-off.

Structure metrics. How many workflows have been redesigned around a demonstrated need rather than merely supplemented with a tool? What changed in decision rights, handoffs, controls, workload, and outcomes? What is the coverage and freshness of the relevant governed

knowledge, whatever its storage form? Has performance management been revised where old incentives conflict with the new work? Does governance enable bounded learning while enforcing stronger controls as consequence rises? A count without these distinctions rewards relabeling.

Mental-model signals. This layer resists dashboards. Observation can still reach it. Listen to what the leadership team asks about in reviews, but do not mistake language for belief. Questions about cost, speed, and headcount may be necessary; the problem is when they crowd out quality, workforce effects, rights, knowledge, risk, and customer outcomes. I ask teams to keep a record of the questions asked in a quarterly review and sort them by altitude. Then I compare the questions with decisions and resource flows. A new vocabulary is weak evidence. A repeated change in choices when the old incentive still pulls the other way is stronger.

The Workforce Trajectory

Section 5.6 posed the questions: are workers developing or losing capabilities the role still requires, engaging or disengaging, gaining agency or losing it? Those questions can be investigated, though no single instrument settles them. Here is how I have seen teams begin.

The work-sample comparison. With employee participation, appropriate permissions, a stable rubric, and comparable tasks, examine how the work of a role or team changes over time. Is it using more judgment and synthesis, developing new capability, or narrowing toward checking machine output? Review aggregate patterns and let people contest the interpretation. Do not repurpose ordinary work as a hidden performance test. This practice is home-grown; I know of no validated methodology for the longitudinal assessment proposed here, and a small sample should generate questions rather than conclusions.

Retention of critical capability. Overall attrition can hide what matters. Track where scarce judgment, relationships, and operating knowledge are concentrated, whether succession and development exist, and whether regrettable departures cluster by role or team. Avoid a secret list of supposedly indispensable people: it can reproduce bias, create surveillance, and confuse current visibility with actual contribution. Section 7.2's senior analyst walked out of a company whose average retention number could still look fine. Measure capability risk and bench strength alongside attrition.

Internal mobility into AI-fluent roles. Are people moving through the transition into redesigned roles, or does every AI-fluent position get filled from outside? Outside hiring can add

capabilities the company does not yet possess. If it becomes the only pathway while internal candidates receive training but no opportunity, that is evidence the up-leveling loop may not be reaching the workforce, whatever the completion numbers claim.

The welder's-standard pulse. Ask them. A voluntary rotating sample, a real conversation, one question at its center: is this work challenging you in ways you value, or draining you? Bring themes and carefully deidentified excerpts to the leadership team, along with disagreement and context; never promise anonymity that a small sample cannot provide. The welder's standard from Section 5.6 was framed as a question a leader should be willing to be asked. This is the practice of asking it early, while the work can still be changed.

Measurement Theater

The failure modes of transition measurement are as patterned as the loops themselves. Five recur.

Dashboards that never inform a decision. The test is one question: name the last decision, investigation, or obligation this dashboard supported. If the room goes quiet, the dashboard may be theater, or its purpose needs to be stated again. Either way, its maintenance cost is being paid in credibility as well as hours. Somebody still updates it every Monday.

Metrics that punish honesty. Reported AI incidents rising early in an epistemic-hygiene program can be evidence that people are catching and surfacing what was previously invisible. It can also mean exposure or actual harm increased, so reporting volume must be read beside usage, severity, detection method, and outcomes. A team told only that its incident count must go down may stop reporting incidents long before it stops having them. Decide in advance which signals may rise as visibility improves, and investigate rather than punish the rise.

Short-cycle demands on long-cycle change. Up-leveling and trust may take years to compound, even though early harm, broken commitments, and unequal effects require prompt attention. Demand a predetermined quarterly improvement and you invite fabricated movement or abandoned investment. My working rhythm is a starting hypothesis, not a universal cadence: inspect leading signals and harms frequently, review meaningful outcomes when the work could reasonably have changed them, and set longer decision horizons in advance with checkpoints and stop conditions.

Measuring people when you meant to measure loops. The moment a capture-contribution signal becomes an individual quota, you have armed the defensive-expertise loop and taught the

workforce to game the graph. Every signal on this panel reads the health of a loop. None of them belongs in a performance review.

The complete-coverage fantasy. Section 7.1 said the companies that escape metric displacement are willing to operate without complete metric coverage, and this section does not repeal that. Some of what you most value will resist direct measurement. Use careful qualitative evidence and invest in the practices and moments that may produce it. Let the panel acknowledge uncertainty rather than installing a bad proxy and optimizing it into something you did not want.

BRING THIS INTO THE ROOM

The instrument-panel exercise. Each member of the leadership team writes down, alone and before discussion, three pieces of evidence they would want to see regularly to judge whether the transition is real. Put them on the wall. Sort them against the four altitudes, then mark the population, data source, privacy risk, competing explanation, expected time horizon, and decision each could inform. In my experience the wall begins crowded at the event level and thins with depth. Then build the shared panel: a small set of signals and qualitative evidence, the major loops represented, more than one altitude represented, and each item owned by someone who will be asked what was learned or decided from it.

One more thing before leaving the systems view. A panel like this can make the transition feel like something you watch, and the transition remains something you run. The instruments exist so that the four-question loop diagnostic from Section 7.1 gets answered with evidence instead of impressions, and so that the interventions from Section 7.2 get aimed below the waterline. The reading is in service of the intervening.

The numbers and qualitative evidence help you judge which way the loops may be running. They do not tell you what the transition feels like from inside a working day. That is what Stage 8 is for: a Tuesday morning in a company where the loops appear to be running in the right direction, and honest stories from companies where they were not.

CONSIDER BEFORE YOU ACT

If you keep three things from Stage 7: 1) Map what is reinforcing change and what goals, constraints, delays, or side effects are counteracting it. 2) Run the four-

question loop diagnostic on a cadence, together, with honest specificity. 3) Use evidence to read the system without turning system signals into covert individual performance measures, or you may arm the loops you are trying to understand.

STAGE 8

The Narratives

What an AI-Native company looks like on a Tuesday morning.

Section 8.1

Day in the Life, Week in the Life, Month in the Life

What does AI-Native work look like on an ordinary day?

A reader can agree with every section of this book, follow every argument, and finish with no clear picture of what the company being described actually looks like on a Tuesday morning. The argument lands intellectually and stays abstract operationally.

Three constructed narratives close that gap. A day in the life of a senior knowledge worker. A week in the life of a leadership team. A month in the life of a workforce going through an AI-Native transition. These are fictional composites, not reported case studies or forecasts. They are partial rather than exhaustive. Hold them in mind as design scenarios: concrete instances of how the operating model in the rest of the book might appear in practice.

They aim for honesty about friction, judgment calls, and moments where the company has to choose between competing goods. They are still selective. A narrative can make a designed future feel more proven and coherent than it is. The point is to make the work recognizable, invite critique, and expose decisions the abstract framework can hide, not to claim that a real company will unfold this neatly.

METHODOLOGICAL LINEAGE

Henry Mintzberg's 1973 book The Nature of Managerial Work, based on structured observation of five chief executives, challenged a classical account of management associated with Henri Fayol. Mintzberg documented brief and varied activities, frequent interruption, verbal information, and ten managerial roles; his 1975 Harvard Business Review article The Manager's Job: Folklore and Fact received that year's McKinsey Award. His method sets a demanding standard precisely because these narratives do not meet it: I am constructing scenarios rather than reporting observation. The lineage informs the attention to ordinary time and actual activity. The evidentiary status is different, and a real company's work should be observed rather than assumed to match the story.

Mintzberg, Henry (1973). The Nature of Managerial Work. New York: Harper & Row. / Mintzberg, Henry (1975). "The Manager's Job: Folklore and Fact." Harvard Business Review, 53(4), 49-61.

A Day in the Life: Senior Analyst at an AI-Native Consulting Firm

Figure 8.1a. An ordinary Tuesday, marked by which moments are hers alone, which happen in the surface, and which the firm keeps off the surface entirely.

Elena Marsh is forty-six. She joined the fictional firm fourteen years ago as a research associate and worked her way up through the analytic ranks. The firm has been doing the AI-Native work seriously for about three years. In this scenario, the transition has changed her job, helped her develop new capacity, and altered the texture of her days in ways she could not have predicted when it started.

She arrives at her desk at 8:15 and does not check email. The firm's practices include a short ritual of beginning that she has come to value: fifteen minutes with the open questions on the engagement she is leading, a hand-drawn sketch of what today needs to accomplish, a note about the moments that will require her full attention. The engagement is Lakeshore Mutual, a regional property-and-casualty insurer deciding whether to keep writing homeowners coverage along the Gulf Coast or retreat to its inland book. The sketch happens on the notepad that sits on her desk. She has read enough about AI-for-thinking versus AI-as-thinking to know that the first sketch of the day belongs to her rather than to the AI.

Then she opens the firm's collaboration surface, the layer the firm built two years ago after recognizing that single-user chat tools were producing the thirty-islands problem from Section 4.3. The surface holds the team's shared context for the Lakeshore engagement, and overnight it has flagged something worth examining: two competing carriers have filed non-renewal notices in the same three coastal counties, three weeks apart, which changes what a Lakeshore retreat would look like to regulators and to the agents who sell its policies. Below the flag sits a question that Tomás posted from the Denver office last evening, and below that a draft section of the Lakeshore deliverable that a team member produced yesterday with AI support, awaiting her review.

She acts on none of it immediately. She reviews the items as inputs and decides which to engage with first. Being the one who decides what gets her attention, rather than letting the surface set the agenda, is a discipline she has had to learn deliberately. The surface is helpful, but it is designed to surface things, and not everything it surfaces deserves her attention.

She starts with Tomás's question. He has modeled Lakeshore's coastal book, and the numbers are ugly: the coastal counties lose money in seven of his ten scenarios. His question is about positioning. The analysis is ready for review, not finished beyond challenge. Lakeshore's CEO is

the founder's grandson, and the founder built the company writing exactly these policies. Should the board presentation lead with the retreat scenario, or hold it as one path among three? Tomás has done the analytic work and used AI to support it. What remains includes validation, judgment, and communication, and he is asking for hers because she has walked clients through retrenchment before. She responds in the surface, visible to the authorized team, with her evidence and rationale laid out so he can challenge the thinking rather than merely receive the conclusion: lead with the numbers, but frame retreat as a question about what the company owes its agents, because that may be the frame through which this CEO can engage with the choice.

The response takes her twenty minutes. Before the transition, in this scenario, it often took an hour because she rebuilt more of the analysis before trusting it. The system and Tomás have done much of the preparation. She is reviewing the evidence and exercising judgment, and the firm values exactly that. The firm will need to check whether the time saved holds across other engagements.

Later in the morning she has her one-on-one with Priya, a junior analyst on her first independent engagement, a marine-parts wholesaler whose COO has championed a warehouse expansion the demand data does not support. They choose an in-person room without screens; the firm protects focused mentoring while providing remote and assistive alternatives when people need them. Priya gets to the real question quickly. "He announced the expansion at their holiday party. How do I tell him it doesn't pencil?" Elena listens, asks questions that help Priya find her own answer, and offers a framing that has worked for her: separate the decision from the person who made it, and give him a way to be the one who caught it. An AI system could generate advice or rehearse the conversation. Priya has chosen the accountable judgment and relational presence of someone who has delivered unwelcome numbers to proud executives for fourteen years.

Back at her desk, the morning has compounded. Tomás has thanked her in the surface, and a colleague in the insurance practice has built on her response with a consideration she had not weighed: Lakeshore's reinsurance renewal lands in January, which cuts the decision timeline in half. Three people have advanced the work through a shared, permissioned space. They might have done so through older channels; the surface made the context and timing easier to see.

She turns to the deliverable review. The AI-supported draft is competent, the framing right, the analysis sound. But the recommendation has converged on the obvious answer, full retreat from

the coastal counties, when the engagement requires the less obvious one. Lakeshore is a mutual, and its coastal agents are the same agents who feed its profitable commercial lines; a full retreat would gut the network that makes the rest of the company work. The better answer is staged: tighter underwriting, a parametric reinsurance layer, and withdrawal only from the two counties where nothing closes the gap. The team member who wrote the draft did not see the divergence between the obvious answer and the right one. The AI did not see it either, because the AI was working from the prompt and the data, without the texture of the client relationship. She rewrites the section and leaves notes in the surface about what she changed and why. The notes are part of the team's learning practice; the next time an engagement turns on a distribution network the data cannot see, her reasoning will be there to consult. She is not certain she is right. She will find out this afternoon whether the staged answer survives contact with Lakeshore's CFO.

The original draft and her revision sit side by side in the surface, where the team member who produced the draft will see the difference and learn from it. The mentorship is happening through the work itself rather than through a separate training program.

At midday she has lunch with two colleagues, and the conversation covers a wedding and a kitchen renovation, nothing about Lakeshore. The firm has been deliberate about the distinction between time for work and time for the relationships that make the work possible. She returns to her desk refreshed in a way that the pre-AI version of her, who ate at her desk while working through the lunch hour, did not experience.

The afternoon holds the consequential meeting of the day. Lakeshore's executive team joins by video; her firm's engagement partner is in the room beside her. The team has used AI to help prepare materials and model scenarios, and its approved retrieval system has surfaced four permissioned, appropriately deidentified prior engagements from the firm's governed knowledge base. No AI system participates live in this meeting, a choice based on client consent, confidentiality, and the kind of conversation expected. The decision also depends on a relationship the firm has earned over years.

The meeting goes for ninety minutes and Elena speaks twice. Once when the CFO pushes on the parametric layer's basis risk and she has the answer because of the morning's preparation, and once when the CEO goes quiet at the retreat map and she offers the framing about the agents, the same one she gave Tomás that morning. Her interventions are short. They land. Lakeshore's CEO asks for the staged plan in writing, which is most of the way to yes. The partner thanks her

on the way out. The work was preparation that took her hours and judgment that took her seconds, and the seconds were what mattered.

Late in the afternoon comes the part of the firm's operating rhythm that took her longest to accept: unstructured time, on her calendar, protected, for thinking, reading, working on problems that have not yet been named, sometimes for nothing in particular. Today she reads a long piece a colleague shared in the surface about Gulf Coast homebuilders adopting fortified roofing standards, adjacent to Lakeshore's world rather than inside it. The reading produces a half-formed idea about an insurer-funded discount program. She might use it in three months. She might never use it. The firm has decided that this hour is part of the work, and Elena, who fought it for the first year, now considers it indispensable.

At the end of the day she writes a brief reflection. What worked, what did not, what is on her mind for tomorrow, and today one more line, about whether she overrode the draft's retreat recommendation too quickly. The reflection is for her own development, and it goes in a notebook rather than the surface. Some things are private.

She leaves at 5:30. In the scenario, her workday is shorter than it used to be, reviewers judge the output stronger against the firm's quality criteria, and the texture of her hours has changed. By the welder's standard from Section 5.6, she describes herself as being in a better relationship with her own brain than she was three years ago. She also notices that the transition was not always comfortable, that there were stretches when she did not know what her role was becoming, and that she did not always appreciate leadership's choices. Some she challenged, and some changed because workers challenged them. She is grateful for parts of the support now. The disagreement was part of the transition, not something leadership had to hold on her behalf until she understood.

A Week in the Life: Leadership Team at an AI-Native Manufacturer

Figure 8.1b. Five standing rituals across one week, and the retirement wave that only surfaces because all five exist.

This fictional manufacturer machines hydraulic manifolds and valve assemblies for agricultural and construction equipment makers. It has been in business for sixty-three years. Carl Brandt has been CEO for nine of them. He inherited a strong company, and he has spent the past four years working to keep it strong through the transition at the center of this book.

The leadership team, eight people including Carl, meets every Monday morning, in person when possible. The meeting starts at 8:00 and runs until 10:30: an hour of operational review of the week ahead, an hour of substantive conversation about something the team needs to think about together, thirty minutes unstructured. The team has held to this format for two years.

This Monday's substantive conversation is about Gene Halvorsen. Gene has been with the company thirty-one years and runs the Rockford plant, the low-volume, high-mix operation

where the hardest jobs go, and he has told Carl he is thinking about retirement. What Gene knows about that plant has never been written down: which spindle drifts when the humidity climbs, which supplier's certifications deserve a second look, why the pricing on the largest OEM contract is structured the way it is. The team understands that this knowledge is irreplaceable on any short horizon, and Gene's signal is the prompt for a conversation they have been deferring for two years.

The discussion runs into the unstructured half-hour and produces clarity about what the decision involves. The decision itself waits. What capability is at risk if Gene retires on the current trajectory. What should the company carry forward. What would honoring him look like while recognizing that he has earned the right to leave without turning his departure into an extraction project. Carl says plainly that he does not know the right path. The chief operating officer names the operational risk, the chief human resources officer the human dimensions, and the chief technology officer the limit that matters most: interviews, observation, records, and a knowledge graph may capture some of what Gene can articulate and demonstrate, but not all of his situated judgment. The team agrees to return in two weeks with options and to invite Gene to shape them, with protected time, choice over participation, credit, and compensation where the work exceeds his role.

On Tuesday morning Carl does his quarterly walk-around at Rockford, a practice he started three years ago: half a day on the floor, in conversation with workers, asking questions, listening more than talking, taking notes on what he hears. He comes back knowing things he did not know. The grinding cell has been running a coolant-filtration bypass the floor team rigged themselves, without a capital request, and scrap on that line is down because of it. A buyer at the company's second-largest customer is irritated about late inspection paperwork, an irritation that has not yet reached the account team through any official channel. And a second-shift machinist named Dana has been suggesting a fixture change that would cut changeover time on the long-run jobs, and her supervisor has been waving it off.

Carl follows up on all of it. The filtration rig gets recognized through the firm's practice of celebrating valuable worker-led improvement, while safety and engineering review determine whether the bypass should remain. The paperwork complaint goes to the account team with Carl's name on the escalation. With Dana's agreement, Carl brings her and her supervisor together, listens to the fixture idea, and asks them to define a fair, safe evaluation. The conversation is uncomfortable for the supervisor and clarifying for everyone else. It ends polite

and unconvinced; whether the fixture change gets a fair evaluation is something Carl will have to check later without turning a CEO visit into a substitute for a functioning improvement process.

On Wednesday the leadership team holds its monthly review of the AI-Native transition, structured around the four diagnostic questions from Section 7.1. What loops appear to be running well. What loops may be running in the wrong direction. What goals, constraints, delays, or side effects are counteracting the interventions. What defeating structures may be present. Fourteen months of these reviews have worn away some of the early awkwardness without converting the maps into certainty.

This month's review surfaces a stall. Voluntary knowledge contribution and reuse appear healthy at three plants and are going nowhere at the fourth, in Owensboro. The first instinct in the room is pressure: set a target and hold the plant manager to it. The systems-thinking discipline makes the team investigate instead. The Owensboro manager lived through a documentation drive at a previous employer fifteen years ago that ended in layoffs six months after the binders were full. That history makes his concern rational, but it may not be the only cause. The chief operating officer and chief human resources officer will visit together next week, open by listening to workers and their representatives, restate what the covenant does and does not permit, and decide with the plant whether any capture should proceed.

On Thursday Carl has dinner with the chairman of the board. The dinner is informal and the conversation ranges, and then the chairman sets down his fork. "How are you doing, Carl? You, apart from the company." Carl answers honestly. He is tired. He is worried about getting Gene's transition right. He is energized by what the firm has built and uncertain whether it is moving fast enough. The chairman listens and asks two questions that help Carl see his situation more clearly. Nothing gets decided, and the dinner is one of the more important things that happens to him all week.

On Friday the leadership team gathers for the weekly close. The format is short, forty-five minutes, each member sharing one thing that went well, one that did not, and one thing they are carrying into next week. It produces a quality of conversation the team has come to need.

This Friday's close surfaces something the team did not know it was thinking. The chief human resources officer mentions that a maintenance lead in Owensboro has asked about pension bridge options. Two turns later, the chief operating officer mentions that his best quality engineer floated a phased-retirement question at a plant review. Set beside Gene, the two

mentions become a recognition: the firm is at the front of a wave of retirement signals arriving earlier than it has historically seen. Carl commits to bringing this back as a substantive Monday conversation in three weeks, and the team agrees on what data to gather in the meantime. The Friday close has done what it exists to do. It surfaced a pattern no individual member would have seen alone.

The leadership team's week ends at 11:30 on Friday. Carl leaves at noon for a weekend with his family, and the firm protects disconnection because sustained availability carries health, family, and decision-quality costs. The team protects that time each week.

A Month in the Life: Workforce Going Through Transition

The third narrative is harder to write because the experience of being in a workforce going through an AI-Native transition is not as orderly as the two just described. The leadership team's week and the senior analyst's day were narratives of work that has been going on for years and has reached a steady state of practice. What follows is the experience of a workforce in the middle of the change, before the steady state has been reached.

The fictional company is a mid-sized employee-benefits consulting and brokerage firm, roughly fifteen hundred people. The CEO, Alan Reyes, announced the AI-Native transition fourteen months ago. The transition is real inside the scenario, it is incomplete, and the workforce is in the stretch where the shifts are visible and the destination is not yet clear.

The first week of the month is the all-hands meeting. It runs Thursday at nine, in the largest conference room the firm has, folding chairs added along the back wall, the regional offices patched in on a screen. Alan stands at the front and walks through what is going well, what is not, and a decision the leadership team has since concluded was wrong, the decision to pilot the new renewal-analysis tools in the firm's largest office first, where the senior consultants had the least slack to absorb a learning curve. He names what they are doing differently because of it. The address takes twenty minutes. The Q&A takes forty. The questions are real, and some of them are uncomfortable for him. He answers them anyway.

The workforce's response varies. Some people are reassured by the honesty. Others are alarmed that the leadership team admits in public to having been wrong, and a portion remain unconvinced. The facilitation partners expected mixed responses because people occupy different roles, histories, and levels of risk. A uniform response would be a reason to examine

whether dissent feels safe or the question was framed too narrowly, not proof of manipulation or cynicism.

In the second week, the firm rolls out a new set of practices for its health-and-welfare renewal line, the result of nine months of work with senior practitioners, the chief technology officer, the engagement design team, and affected employees. An approved AI system drafts carrier-quote comparisons and renewal analytics for human validation. The intended redesign shifts more consultant time toward plan design and client conversation; whether it raises their contribution, workload, and development has to be measured. The rollout is staged. One team adopts first, its experience is observed, adjustments are made, and the next team follows if the evidence supports it.

The first team's adoption is awkward. The practices were designed around a steady work rhythm, and renewal season is not steady; the team vents its frustration in the weekly reviews. Leadership does not abandon the practices at the first discomfort, but neither does it call every objection resistance. It checks workload, errors, client effects, and what the team says. By the end of the week, two practices have been revised and the rest are becoming more workable, though nobody would call them loved. The team that observed the pilot adopts with clearer expectations; its experience may still differ.

The third week brings the quarterly listening session, a practice the chief human resources officer instituted six months into the transition: two hours with a stratified but voluntary group across roles, levels, locations, tenure, and different experiences of the change. The process states how people are selected, how notes will be used, and what confidentiality can and cannot be promised. The leadership team listens for the first ninety minutes and responds for the last thirty; other channels remain open for people who will not speak candidly in front of executives.

This quarter's session produces genuine appreciation for the decision to give every consultant paid weekly hours for learning the new tools, and the leadership team receives it without inflating it. It produces a concern: the peer-review step in the new practices has, on some teams, become a rubber stamp, which the leadership team had believed was not happening. The chief human resources officer commits to looking at it directly. And it produces a question no one at the front of the room can answer. An eight-year consultant puts it plainly: "If the AI drafts the renewal analysis, what am I getting promoted for?" The leadership team does not pretend to have an answer. They commit to bringing one to the next session.

The week ends with the leadership team translating what it heard into actions, owners, timelines, and explicit explanations when a request will not be adopted. In the weeks ahead, the workforce should be able to trace what changed, what did not, and why. That closed loop is part of what makes listening credible over time.

In the fourth week, Corinne, a senior practitioner of nineteen years, gives notice, and no one saw it coming. Alan calls her directly. She chooses to share that her mother needs full-time care and says the transition played no part in her decision. He accepts this without probing and asks whether she wants to suggest anything that would make the transition better for the people who stay. She offers two ideas: move weekly practice reviews off Friday afternoons, when everyone is spent, and give consultants access to the carrier-comparison sources, assumptions, uncertainty, and reproducible checks rather than only a generated answer. The first becomes a change within the month. The second becomes a technical and workflow requirement the team will test, without pretending that a model's hidden internal reasoning can be made into reliable evidence.

Corinne's leaving is marked, simply and intentionally, the way the firm marks every departure. She is recognized publicly. Colleagues share moments from working together that they will carry forward, and she is invited to stay in touch.

The month closes with the leadership team reviewing where they are. Progress has come in some intended directions, and so have patterns they did not anticipate. Corinne's departure is recorded accurately as a caregiving departure rather than quietly relabeled as transition resistance. It also prompts a wider look at caregiver support and capability continuity, alongside the rubber-stamp reviews and the promotion question. The list is long. Calling it manageable would be premature; assigning owners, priorities, resources, and escalation rules is what makes it manageable. Not pretending the work is simpler than it is is part of what makes the transition real.

The workforce ends the month in a different place than it started. The change is incremental but visible to anyone who is paying attention. The firm is, in small ways, becoming what the leadership team committed to building. It is not finished.

What the Narratives Leave Out

The three narratives just told are narratives of work going well. They are not the only kinds of stories worth telling. The companies that learn from this work also study the failures, and the

failures are abundant and instructive. The next section tells five of those stories, each traced back to the design decision that produced it.

Stage 9 then addresses the method that produced this work and that I use with leadership teams. The Arc5 method, the Nodalix architecture, and the way the two combine to support leadership teams in doing the work just described.

Section 8.2

The Honest Stories: Where This Has Gone Wrong

What does it look like when an AI-Native transition goes wrong?

The earlier parts of this field guide have made a positive case. What AI-Native is. What it requires. What the leadership team has to do to produce it well. The argument has been about what the design should produce.

This section is the other half of the work. It is an accounting of what can go wrong when design work is skipped, delegated without the necessary authority, sequenced badly, or replaced by an announcement. Each case is anchored to a specific design decision and a plausible chain of consequences. The cases are composites, not documented histories, prevalence estimates, or proof that one decision alone caused every outcome. They combine patterns I have seen with patterns reported in research and industry surveys, whose methods and limits matter. None names a specific company. Some readers will recognize their own situation. That recognition should begin an inquiry, not end one.

When a company attributes a restructuring to AI, examine what work has actually changed. Ask which tasks the systems now perform, what happened to demand and staffing, and what evidence supports the announced savings. The cases below are about leadership design decisions and their consequences.

Five cases. Each names a design decision that helped create the conditions for failure, so a leadership team can recognize the choice and test its consequences before committing to it.

Case One. The Pilot Graveyard.

The design decision: approve pilots before doing the alignment work.

Section 1.2 sketched this composite in outline. Here is what four million dollars buys in the scenario. A diversified industrial company decides, in early 2024, to become AI-Native. The CEO announces the commitment at an all-hands meeting. The board approves the budget. Each function is asked to identify two or three AI pilots. The CFO funds procurement optimization. The CTO funds predictive maintenance. The CHRO funds recruiting AI. Marketing funds personalization. Operations funds quality inspection. Customer service funds a chatbot. Eighteen months later, the company has spent four million dollars across eleven pilots. Two demos work in controlled conditions but have not reached production. One system works technically but does not fit the operators' workflow, so it sits unused. The remaining eight are stalled for different combinations of data, integration, ownership, risk, economics, and adoption.

The pilot graveyard is a visible industry pattern, but no single failure rate describes it. The preliminary 2025 GenAI Divide report drew on interviews with representatives from 52 organizations, survey responses from 153 senior leaders, and a review of more than 300 publicly disclosed AI initiatives. It reported limited measurable financial returns across much of its sample. Its sampling, definitions, and non-peer-reviewed status limit generalization. McKinsey's 2025 global respondent survey found that 39 percent attributed any enterprise-level EBIT impact to AI; about 6 percent met McKinsey's stricter high-performer definition, which combined at least 5 percent EBIT impact with significant reported value. RAND's 2024 interview study did not independently measure an 80-percent failure rate; it cited external estimates and found organizational causes prominent in its interviews. S&P Global Market Intelligence's survey of 1,006 North American and European IT and business professionals found the share reporting that their companies abandoned most AI initiatives rose from 17 to 42 percent, with an average 46 percent of proofs of concept scrapped before broad adoption. These sources use different populations and outcomes. Together they support difficulty scaling, not a universal 95-percent law.

In the composite, the failure is not that the model worked and leadership did not. Some technologies fail too. The upstream design problem is that each pilot answers a different definition of AI-Native and has no shared architecture for learning. The CFO emphasizes cost reduction. The CTO emphasizes modernization. The CHRO emphasizes talent operations. None is necessarily wrong on its own terms. But the work does not compound, data and controls do not connect where they should, lessons do not travel, and eleven experiments do not become a coherent portfolio.

The design decision the leadership team should have made instead is described in Sections 1.2 and 9.2. Begin alignment and portfolio design before allocating the full budget. Produce a shared definition that records material disagreement. Choose an archetype as a working hypothesis, identify layer dependencies, set baselines and stop conditions, and define how learning will travel. This may lead to fewer pilots or to a deliberately varied portfolio. Variation can be valuable when it tests different assumptions deliberately; fragmentation is undesigned variation that produces no shared learning. Each experiment needs an owner, a reason to exist, a named user, a baseline, an evaluation method, a path to production or closure, and a way to inform the rest. Shared infrastructure should be reused where it truly is shared, while local requirements remain local. A stopped pilot can be a good result when it answers its question before absorbing more capital.

The lesson: the failure mode here was weak alignment, portfolio design, and learning architecture, compounded by project-specific problems. The remedy begins upstream of the technology, in the work the leadership team does before and during budget allocation, without requiring false consensus or assuming alignment can rescue a bad use case.

Case Two. The CTO-Delegated Strategy.

The design decision: delegate AI strategy to the CTO without CEO and business-leader engagement.

A professional services firm with two thousand employees decided in 2024 to develop an AI strategy. The CEO, recognizing that the technology was outside his domain of expertise, asked the CTO to lead the work. The CTO assembled a team, engaged a vendor, produced a multi-deck strategy document, and presented it to the executive team. The executive team approved it. The CTO began executing.

Eighteen months later, the firm had built a meaningful AI capability inside its engineering organization. Code review was faster. Developer productivity had increased. The infrastructure was sound. But the rest of the firm had not changed. The client work was being done the same way it had been done before. The partners who led the major client relationships had not been engaged in the work and did not know what was now possible. The new business teams were still pitching the same kinds of engagements. The firm's competitive position in the market had not shifted. The CTO had successfully executed a developer-productivity strategy and called it an AI strategy.

Stanford Digital Economy Lab's April 2026 Enterprise AI Playbook studied 51 successful deployments across 41 organizations, so it is a selected set of successes rather than an estimate of enterprise success rates. Within it, the authors report eight cases where business-and-technology co-sponsorship made a difference. At one professional-services company, a CTO-led first attempt failed to gain traction; a second succeeded with the CEO and Head of Talent driving the mandate and incentives while the CTO owned implementation. IBM's vendor-produced 2025 survey of 2,000 CEOs found that 50 percent said the pace of recent investment had left disconnected, piecemeal technology. Those findings support cross-functional ownership. They do not support a stereotype that CEOs reimagine work while CTOs merely accelerate it; either role can think narrowly or systemically.

What went wrong in the composite was a category and governance confusion. AI-Native is an operating-model redesign that includes technology, data, security, workforce, customer, and business-model choices. A CTO may be capable of leading more than implementation, and a CEO cannot delegate accountability for enterprise design. The leadership model should follow competence and decision rights while giving business leaders, technologists, workers, risk functions, and affected customers a real role. A technology-only mandate predictably favors technical outcomes; a CEO slogan without technical ownership fails in the opposite direction.

The design decision the leadership team should have made instead is described in Sections 9.1 and 9.2. Engage more than one capability and make accountability explicit. Technology architecture needs qualified technical leadership. Operating-model choices require the CEO's sponsorship and the business leaders and workers whose work is being redesigned. Risk, legal, security, data, and customer perspectives enter according to consequence. This is not an argument for a committee so large that nobody decides. One accountable executive can own the enterprise outcome while named leaders own architecture, workflow, workforce, risk, and adoption decisions. The streams interlock through shared outcomes, decision rights, evidence, and escalation. Those links hold whether or not an executive is in the room. The CEO cannot outsource the tradeoffs, and the CTO should not be reduced to an order taker after strategy has already hardened.

The lesson: delegating enterprise AI strategy to a single function is itself a strategy decision and often leaves essential authority or knowledge outside the room. Treat AI-Native only as a technology problem and the company is likely to get a technology-shaped project. Treat it only as leadership transformation and it gets aspiration without a reliable system.

Case Three. The Shadow AI Explosion.

The design decision: build governance on prohibition rather than on building sanctioned AI better than the alternatives.

A professional services firm with eighteen hundred employees had a formal AI strategy. The strategy named a set of approved tools, a set of approved use cases, and a set of restrictions on data that could be processed by AI. The strategy was published in a portal, communicated through company-wide messaging, and reinforced in quarterly compliance training. Attendance at the training was excellent. The CIO believed the firm was managing AI adoption carefully.

Twelve months into the composite strategy, an internal audit estimates that the firm can account for approximately four percent of its AI usage. The estimate is uncertain, which is part of the finding. The rest includes personal accounts on consumer AI tools, unapproved browser extensions, and AI features inside SaaS applications bought for other purposes. Some employees have uploaded client documents, source code, or financial statements outside approved controls. The behavior varies from careless to well-intended workarounds, and executives are not presumed exempt. The firm has been governing the visible fraction while much of the actual use runs in the shadows.

The risk is documented, but the numbers need precision. IBM's vendor-produced 2025 Cost of a Data Breach research surveyed 600 breached organizations; 20 percent reported a security incident involving shadow AI. Organizations with high shadow-AI exposure had breach costs averaging $670,000 more than organizations with low or no exposure. That is not the same as shadow AI causing 20 percent of all breaches or $670,000 in every incident. Other industry studies report limited visibility and sensitive data entering unsanctioned tools, but estimates vary with product telemetry, definitions, geography, and sample.

What went wrong was that the policy did not address the work creating demand for AI, and the approved path did not meet enough of that demand. An employee facing a slow or unusable sanctioned tool has an incentive to route around it, although that does not excuse mishandling protected data. Test whether approved tools reduce unauthorized use by measuring both before and after rollout. The quality of the sanctioned experience matters alongside clear rules, training, access controls, data-loss prevention, procurement, monitoring with privacy limits, and enforcement.

The design decision the leadership team should have made instead is described in Sections 4.3 and 4.4. Give workers an approved path that supports real workflows, explain the boundaries, minimize unnecessary friction, and apply proportionate technical controls. A collaboration surface may be part of that design, not the only valid form. Governance both enables and restricts: it should enable safe work and restrict uses that violate law, contract, security, professional duty, or the rights of affected people. Discovery and data-loss-prevention controls can identify risky flows, but monitoring must be purpose-limited, disclosed, access-controlled, and separated from ordinary performance management. Procurement has to evaluate embedded AI features, not only products purchased under an AI label. Incident response has to cover prompts, outputs, agents, credentials, connected data, and downstream actions. A good experience does not make governance free; it makes compliance more practical.

The lesson: prohibition alone is brittle, while convenience alone is not governance. Leadership has to understand the workarounds, provide capable sanctioned options, enforce necessary boundaries, and keep updating both as tools and risks change.

Case Four. The Intelligence-at-the-Center Bet.

The design decision: announce the restructure before doing the design work.

A company in a fast-moving sector decides that it will become AI-Native by reorganizing around AI rather than adding AI to its existing structure. The CEO announces the transformation publicly. The company will be lean and flat. Some roles will shift from doing work to supervising automated workflows or agents. Headcount and organizational layers will be reduced, and remaining employees will be expected to adopt new methods quickly. The market response is initially positive. This composite is a stress test of that design philosophy, not a report of what happened at any named company.

In the stress test, problems emerge twelve to fifteen months later. Complicated cases, angry customers, unusual situations, and moments requiring judgment and discretion are handled poorly. The system performs well on much of the routine volume, but value in the customer-facing function is concentrated partly in the long tail. Net Promoter Score declines and churn rises in previously stable segments. Removed middle managers had not been doing only visible coordination. Some translated context, detected weak signals, absorbed friction, developed people, and maintained relationships. Automation does not automatically replace those functions, so senior leaders inherit some and other work goes undone. This is a plausible causal chain to test before restructuring, not a forecast that every flatter organization will produce it.

The philosophy has a real public example. On May 5, 2026, Coinbase CEO Brian Armstrong announced a roughly 14-percent workforce reduction and described the company as "rebuilding Coinbase as an intelligence, with humans around the edge aligning it." Because that announcement was only months ago, this book cannot claim its long-term outcome. It can examine the bet: AI at the operating core, fewer layers, player-coach leaders, smaller pods, and humans as an alignment layer. Forrester has predicted that half of AI-attributed layoffs will be quietly reversed, often through lower-paid or offshore roles; Gartner has predicted that by 2027 half of companies attributing headcount reduction to AI will rehire similar customer-service functions under different titles. Predictions are not observed reversals. A vendor survey by outplacement firm Careerminds asked 600 HR professionals involved in recent layoffs; 8.4 percent said their AI-driven restructure delivered on its promises and would be repeated unchanged, 32.9 percent reported losing critical skills, and 54.6 percent said AI required more human oversight than expected. Self-report, vendor interest, sample construction, and the wording of the questions limit the result. The evidence justifies testing the operating model before irreversible cuts, not declaring the bet settled.

What goes wrong in the composite is a sequencing failure. The announcement becomes the design decision. The team has not yet mapped what each role does, which moments concentrate value and risk, what current systems can handle reliably, what invisible work each layer performs, or how exceptions will be detected and escalated. It commits to a structure modeled on the average case without testing the long tail or the distributional effects on customers and workers.

The design decision the leadership team should have made instead is described in Sections 3.1, 5.9, and 6.2. Do the operating-model work before an irreversible restructure. Map which moments in the customer journey concentrate value, risk, rights, and responsibility. Observe what each layer actually does, including invisible coordination and care. Test automation on routine and exceptional cases; measure quality and disparate effects; involve affected workers; design escalation and learning; and stage changes with rollback criteria. The resulting restructure might be smaller, different, or unnecessary. The announcement should follow evidence and required consultation rather than substitute for them.

The lesson: an AI-Native announcement is not an AI-Native operating model. Restructure announcements are easy. Operating-model redesign, testing, consultation, and accountable

execution are the work. A leadership team that announces the structure before doing that work is making an irreversible public bet before it has enough evidence.

Case Five. The Institutional Knowledge Walked Out.

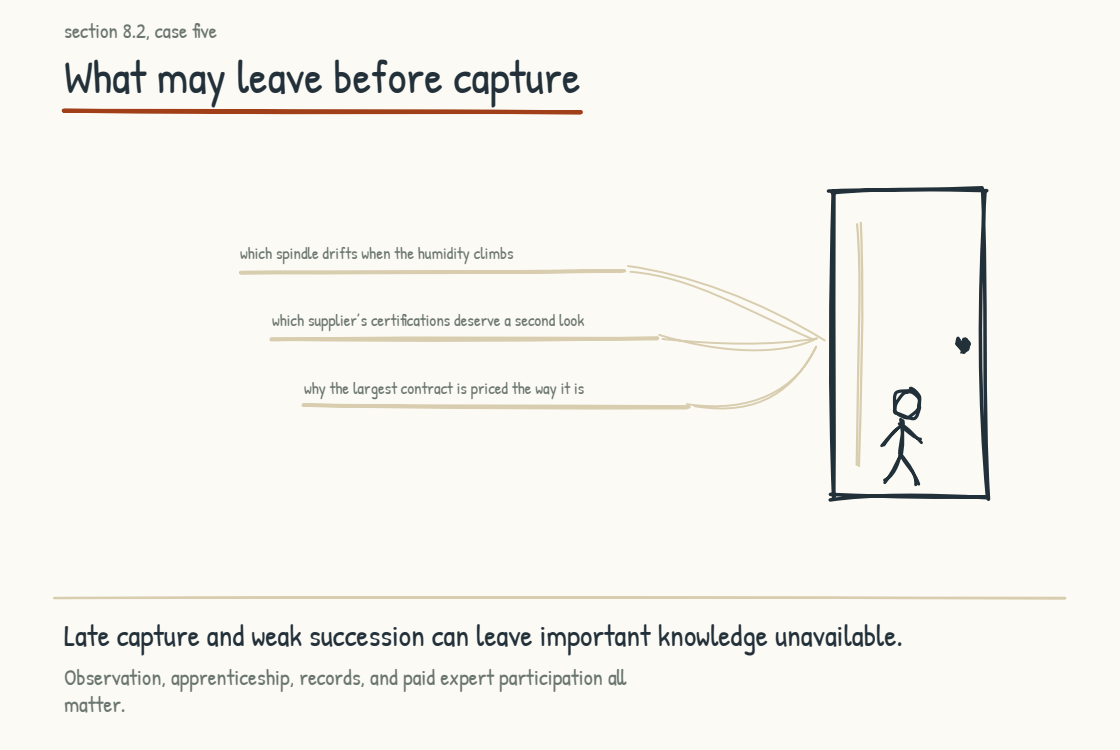

Figure 8.2a. The capture stayed shallow, and thirty years of operational knowledge left through this door.

The design decision: sequence workforce optimization before knowledge capture.

A manufacturing company, with operations going back four decades, had a workforce in its production facilities with deep operational expertise. The plant managers, the senior operators, the maintenance leads. They held knowledge of the equipment, the processes, the quirks of each line, the relationships with suppliers, the historical reasons things were done the way they were. Most of this knowledge was not documented anywhere. It existed in the heads of people who had been doing the work for fifteen, twenty, or thirty years.

The company decided in 2025 to invest heavily in AI for operations. The investment thesis assumed that much of the institutional knowledge could be represented in the systems being built. The CTO commissioned a knowledge project using interviews, observations, documents, operational data, and a graph for relationship-heavy knowledge. The vendor was engaged. The project began.

In parallel, the company was examining workforce demographics and cost. Senior operators were approaching retirement, and early-retirement incentives were taken up faster than expected. By the time the knowledge project entered its substantive phase eighteen months in, many intended contributors had left. The remaining workforce held different, not inherently lesser, experience, but could not reconstruct every condition its predecessors had encountered. One of the last senior operators had run a 600-ton stamping press for thirty years. On cold mornings, he knew, the die ran a thousandth tight until the press warmed, so he fed the first hour slowly and inspected early parts. In March a modeler asked in a conference room how he decided when the press was ready. "You run it until it sounds right," he said. Fifty minutes produced four sentences because an interview detached from the machine was the wrong capture method. Observation, sensor data, demonstrations, and testing might have represented more; some embodied judgment would still resist codification. The system went into production and proved useful, but narrower than the investment thesis promised.

A 2025 Frontiers in Psychology study by researchers at Jilin Normal University surveyed 635 older employees across industries in China. Its overall result was not that AI adoption suppresses knowledge transfer: adoption was positively associated with intergenerational transfer. The study also found a smaller negative indirect pathway through identity threat and positive pathways through relational crafting, with digital self-efficacy moderating threat. The cross-sectional self-report design and specific population do not establish causation or describe every older worker. It supports a more useful point: adoption can produce competing pathways, and identity threat is one condition leaders should address without stereotyping older employees as defensive.

What goes wrong in the composite is sequencing joined to method and trust. The company begins late, after retirement incentives have accelerated departures and automation messaging has changed the employment bargain. Some people leave for retirement, care, health, or opportunity; some may share less because the project feels extractive; some participate fully. Meanwhile, the project relies too heavily on interviews and one representation. The technology can work as specified and still fail strategically because the specification assumed knowledge was complete, articulable, collectible, and stable.

The design decision the leadership team should have made instead is described in Sections 4.1 and 5.2. Capability continuity should begin before departures become urgent, but not as a covert prelude to workforce reduction. Experienced workers should help decide what is worth

preserving, how it can be represented, who may use it, and how contributors receive time, credit, compensation, and the right to decline where appropriate. Capture should happen in the work where possible: observation beside the press, comparison of normal and anomalous runs, review of sensor traces, demonstration to apprentices, and explicit testing of the resulting rule. It should combine operational data, documents, mentoring, simulation, and the storage forms the use case requires. Safety-critical knowledge needs validation, versioning, ownership, and a process for correction; a vivid anecdote is not yet an operating instruction. The covenant makes commitments explicit; only conduct over time makes them credible.

The lesson: threatened people may reasonably withhold participation, but trust is not the only limit. Institutional knowledge is distributed, situated, changing, and only partly articulable. Honor the workforce because people have standing and because fair participation improves the work. Then design for what cannot be captured: apprenticeship, redundancy, succession, judgment, and continued access to experienced humans.

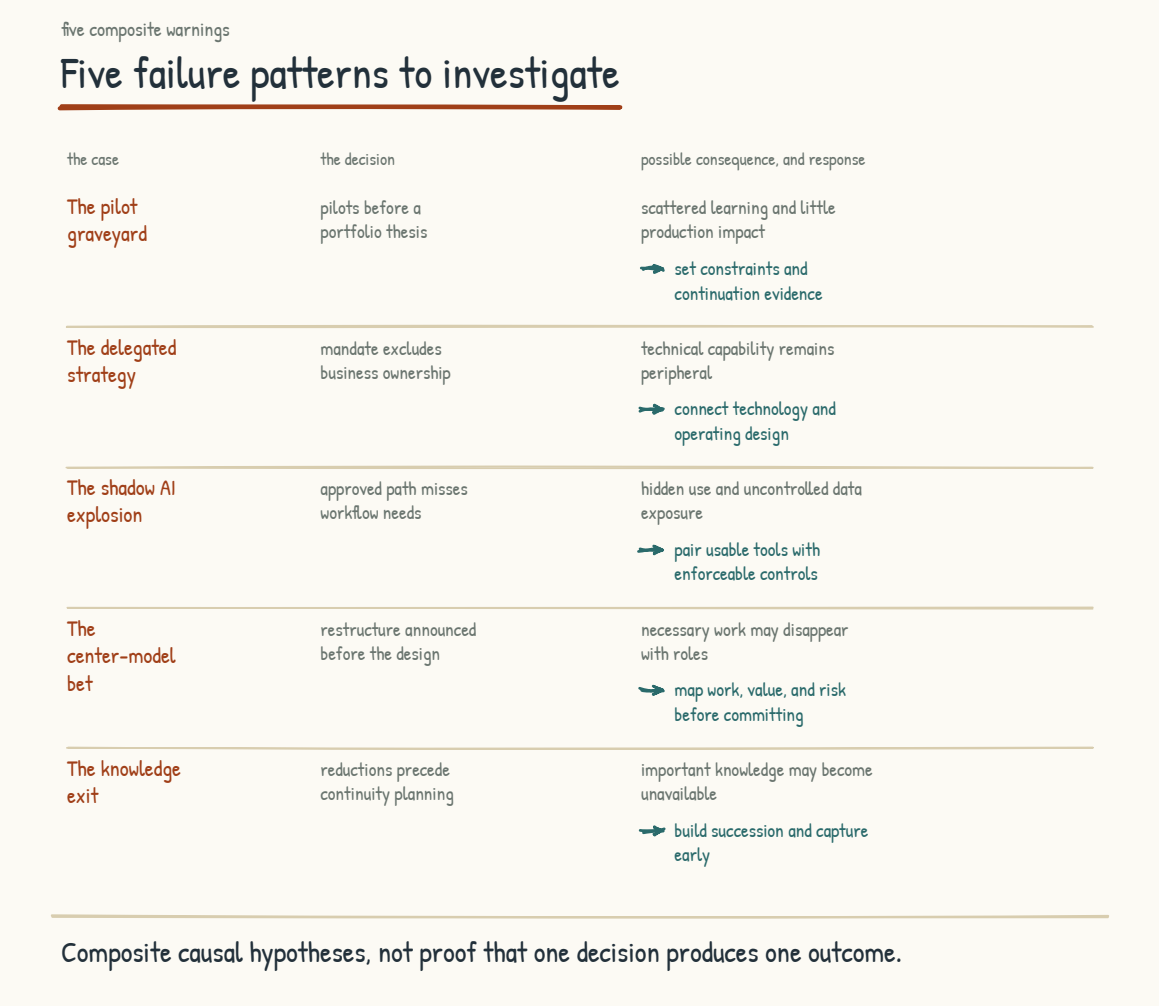

Figure 8.2b. The five failure patterns at a glance. Each row names a design decision that contributed to the failure, the consequence in the composite, and the section of this field guide that describes a better design response.

The five cases above share more than they differ. Each begins with a leadership team making a specific design decision, often without recognizing it as one, and encountering its consequences months or years later. The decision is not the only cause. It is the upstream choice the case makes visible.

Across all five composites, visible announcements and structural decisions interact with less visible assumptions, incentives, power, data, methods, and sequencing. Sections 7.1 and 7.2 provide hypotheses for tracing those interactions. A leadership team that intervenes only at the visible level increases its exposure to these failure patterns, but the cases are diagnostics, not destinies.

The composites begin with pressures a leadership team can recognize: cost, growth, risk, and the demand to show progress. Naming the patterns makes the design decisions legible, so a leadership team can recognize the choice before they make it, and choose the design that produces a different outcome.

The earlier intervention is the work this field guide has been describing. If the rest of the book has been an argument for what to design for, this section has been an argument for what happens when the design is skipped, delegated, sequenced badly, or replaced by an announcement.

CONSIDER BEFORE YOU ACT

If you keep three things from Stage 8: 1) Done well, the transition can change the texture of ordinary days, expanding some judgment work and shrinking some administrative work without assuming every role benefits equally. 2) Mixed workforce responses are normal; a uniform response is a reason to examine whether dissent is safe, not proof of manipulation or cynicism. 3) Every failure composite began with design choices whose consequences were not made legible; make the choices, evidence, affected groups, and reversibility visible before you commit.

STAGE 9

The Method

The combination of capabilities the work requires, and the engagement that starts it.

Section 9.1

Arc5 and Nodalix: Why This Combination, Why Now

What kind of partner does this work require?

This part turns to what a leadership team that has been moved by the argument can do to start: why the combination of capabilities behind this book is well-suited to the work, and the specific engagement that can begin it.

This section and the next are different from the rest of the book. They describe what the two firms behind it offer, and why I believe that offer fits the work. A team that has read this far has earned the statement plainly.

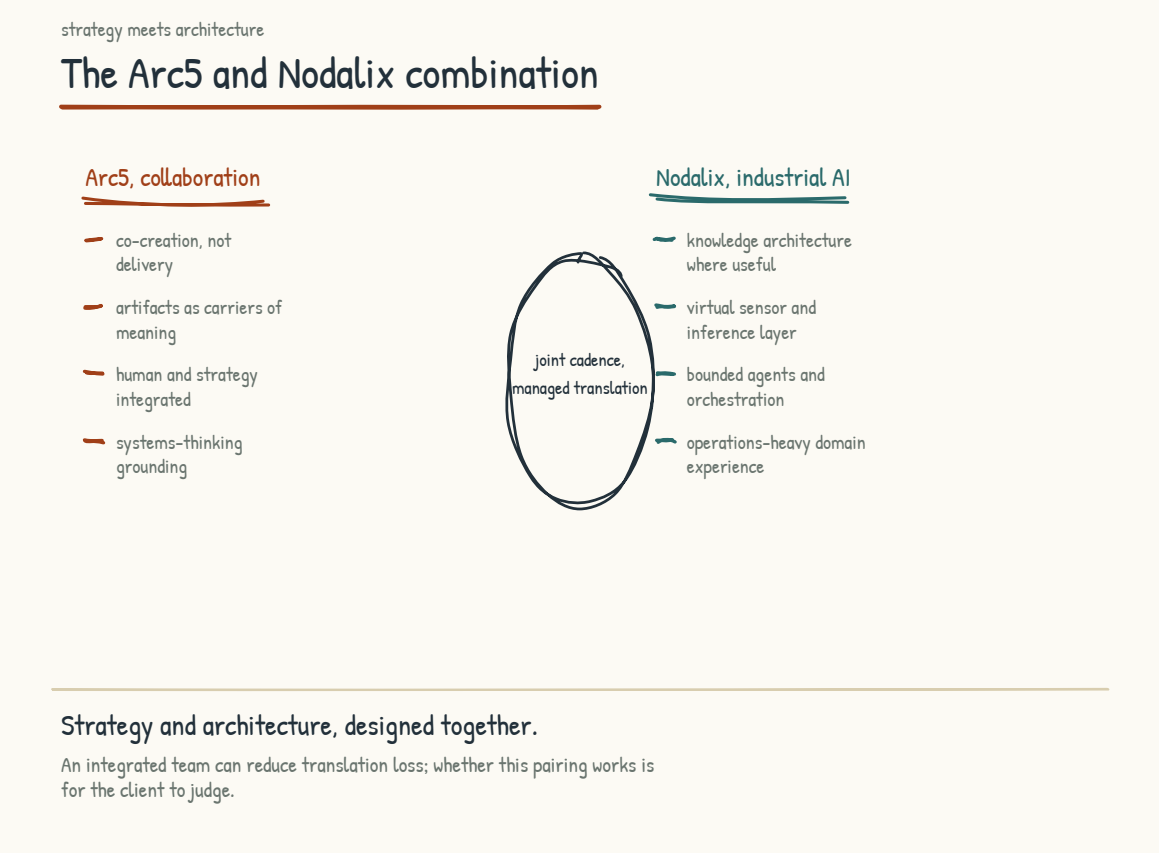

Figure 9.1a. The two capabilities and the overlap zone that defines the combination. A poorly governed handoff between strategy and architecture is one important failure mode in AI-Native work.

The Two Capabilities

Work of the kind this book recommends takes two distinct capabilities. One is designing and facilitating the human and strategic work at the leadership-team level. The other is building the technical architecture that makes the resulting operating model real.

That first capability is practiced by some management consultancies, organizational-design specialists, and facilitation firms. It involves understanding how leadership teams think, designing conversations that produce shared understanding without suppressing disagreement, holding difficult moments, surfacing assumptions, and helping a team construct and test new working models. It requires experience with senior teams, fluency in design thinking and systems thinking, structured collaborative design, and the capacity to read a room and adapt in real time. It is craft, in the specific sense used in Section 5.8, and prospective clients should ask for evidence of it rather than infer it from a firm's category or presentation quality.

The second is practiced by technology firms and strong internal engineering organizations. It involves understanding AI at the system level; choosing among graphs, relational data, documents, rules, models, and workflows; building automation or agents within governance constraints; and integrating with operational data and tools. It requires technical expertise, relevant domain knowledge, and the capacity to translate operating-model decisions into production systems. Evidence should include reliability, security, maintainability, user outcomes, and what happened after deployment. A successful demo is where that evidence starts.

The two capabilities can live in one firm, a partnership, or a strong client-led team, but they are often contracted and governed separately. One group writes recommendations; another interprets them and builds. If outcomes, assumptions, decision rights, and feedback are not shared, translation loss appears in the handoff. The client then has to reconcile outputs that were never designed to fit.

The reverse failure also occurs. A technical team builds exactly what it was asked to build while critical operating-model, workforce, customer, or governance assumptions remain untested. That is not always the technology firm's failure; the mandate, commercial model, or client governance may exclude the necessary work. A technically functional system can still be operationally misaligned, and a well-designed operating model can still be defeated by poor engineering.

I have watched versions of this failure mode repeatedly. The client ends up integrating outputs that were not designed together, and responsibility becomes ambiguous. The answer is not necessarily one vendor. It is one accountable design with explicit interfaces, shared evidence, and governance that follows the work across organizational boundaries.

What Arc5 Brings

Arc5 is a collaboration design firm. Its work is the design and facilitation of immersive engagements for senior leaders facing complex, high-stakes decisions, grounded in design thinking, systems thinking, and the discipline of structured co-creation.

Arc5's method has been refined over many years, across industries as different as defense and hospitality, and across decisions from strategy to large-scale transformation. That breadth informs its cross-domain pattern recognition. It does not mean a frame validated in one setting transfers automatically to another; the method has to be adapted and its assumptions tested in

each organization. Prospective clients should examine relevant examples, references, facilitator fit, and how outcomes were assessed.

Four properties of the method matter for the AI-Native work.

First is the discipline of co-creation rather than answer delivery. Arc5 arrives without a predetermined answer and designs conditions in which the leadership team and affected participants can construct one. The discipline is harder than it sounds. A facilitator inevitably shapes the room through the agenda, questions, artifacts, participation rules, and moments of intervention. The obligation is to make that influence deliberate and contestable, surface dissent, and keep decision rights clear. Construction can increase understanding and ownership; whether the outcome survives execution depends on resources, incentives, authority, technical reality, and follow-through as well.

Second is the use of artifacts as carriers of meaning. Alongside any necessary decks or records, Arc5 produces physical and digital artifacts the leadership team works with during the engagement, designed to hold assumptions, choices, dissent, and rationale in a form that supports later decisions. Some teams have continued using such artifacts for years in my experience. Others will need them revised, integrated into an existing system, or retired. Continued use is evidence of utility only if the artifact remains accurate and changes decisions, not merely because it is visible.

Third is the integration of human and strategic work. Arc5 does not treat people as an adoption layer applied after strategy. Conversations address what the company should do and how authority, capability, workload, rights, and leadership practice may need to change. That can begin the human reckoning and covenant work from Section 5.2. It cannot legitimately complete them in an executive room: workers, representatives, customers, and other affected people need channels that match their stake, and some decisions require bargaining, legal process, or specialist review.

Fourth is the grounding in systems thinking. Arc5 designs engagements that help participants formulate and test accounts of the systems they operate within, including the leadership team's own role. The loops from Section 7.1 and the iceberg from Section 7.2 become working hypotheses rather than decorative frameworks. A workshop can introduce and practice that discipline. The leadership team's capacity compounds only if it continues mapping, gathering contrary evidence, revising its models, and changing decisions after the facilitators leave.

These four properties are the strategic and human side of the work, and on their own they are not enough. The work the leadership team produces in the engagement needs to translate into technical architecture, and the translation is where the second capability becomes necessary.

What Nodalix Brings

Nodalix is an industrial AI company. Its work is the design and construction of AI systems for operations-heavy environments, grounded in knowledge graph architecture, virtual sensors, multi-agent systems, and the disciplines of governance, observability, and drift management described in Sections 4.3 and 4.4.

Nodalix is built around a specific premise: institutional knowledge is often a major constraint on useful AI in operations-heavy environments. The appropriate architecture may combine graphs, documents, databases, time-series data, rules, models, and human workflows. A graph is valuable when relationships and dependencies are central; it is overhead when a simpler representation answers the need. Structured knowledge can improve grounding and reuse, but competitive performance still depends on data quality, operations, adoption, maintenance, and the quality of the underlying business. The premise is load-bearing; graph-first dogma is not.

Four Nodalix capabilities matter for the AI-Native work.

First is knowledge architecture. Where a graph fits, Nodalix models operation-specific entities, relationships, dependencies, provenance, identity, time, and permissions rather than treating a generic ontology as the finished answer. It also connects the graph to other authoritative stores. No representation captures an operation completely, and soft knowledge does not become reliable merely because it is encoded. The foundation is a maintained evidence system matched to the use case.

Second is the virtual-sensor and inference layer. Nodalix combines available signals to estimate conditions that may not be measured directly and connects model-supported inference to operational context. A virtual sensor is an estimate, not a physical observation; it needs calibration, uncertainty bounds, drift detection, and fallback behavior. Software and models can traverse and infer from represented knowledge, while accountable people and institutions retain ownership of consequential decisions.

Third is the agent and orchestration layer. Nodalix builds delegated software that operates within defined permissions, approvals, budgets, and stop conditions, with logs and observability proportionate to risk and privacy. Audit trails do not make an agent trustworthy by themselves:

evaluations, access control, change management, incident response, and accountable owners matter. Roles and limits should be explicit, and autonomy should rise only with evidence and reversibility.

Fourth is experience in operations-heavy contexts, including work described by Nodalix across defense, manufacturing, healthcare, and related settings. That experience can be relevant where failure has physical, clinical, security, or economic consequences. It should not be accepted as self-validating. A prospective client should verify which systems reached production, what claims can be referenced, which outcomes were independently measured, what security or regulatory obligations applied, what incidents occurred, and how the system is performing now. Confidential work may limit disclosure; it does not remove the need for proportionate due diligence.

These four capabilities are what Nodalix brings, and they are also not sufficient on their own. The architectures need to fit the operating model decisions the leadership team has made. If the leadership team has not done the strategic and human work, the architectures are built against assumptions that may not hold, and the system that emerges is technically functional and operationally misaligned.

The Combination

Both firms work together. The intended model is not a joint sale followed by an invisible handoff: the strategic and human work Arc5 leads and the technical architecture Nodalix leads are designed together from the beginning. The client should still see separate scopes, responsibilities, commercial interests, escalation paths, data access, and accountability. Continuity should reduce the client's integration burden, not conceal where one firm's obligation ends and another's begins.

That integration can produce advantages a leadership team feels directly. It also creates a commercial risk the client should govern: the team helping define the need may benefit from being selected to build the answer. The firms should disclose that interest, separate discovery conclusions from predetermined product choices, make artifacts portable, and leave the client free to compete, pause, or internalize implementation. Integration is valuable when it improves the work, not when it creates dependency.

Decisions emerging from the Arc5 work can be informed by technical reality: what appears buildable, on what assumptions, at what estimated cost and time, with what dependencies and

tradeoffs. Early technical participation improves realism, but estimates remain uncertain and should be updated through prototypes, security review, user research, and operational testing.

Nodalix can then design against an explicitly chosen operating model rather than a brief stripped of its rationale. Strategic and human implications are not thereby "worked through" forever; new constraints and affected groups will appear during implementation. The architecture and operating model should co-evolve through controlled decisions rather than forcing either one to fit a frozen workshop output.

The handoff does not disappear. Every transition between facilitation, design, engineering, deployment, and operation is a handoff, including within one team. The combination aims to make those interfaces visible, staffed, and governed, with shared artifacts and people spanning them. Translation loss should be measured and corrected, not declared absent.

Human work continues into technical implementation. The covenant from Section 5.2, practices from Section 6.1, and meaningful-involvement decisions from Section 6.2 can be carried into requirements, testing, rollout, and operations. The firms remain accountable for their own work; the client retains employer, operator, legal, and executive accountability that no adviser can assume.

Why Now

Timing matters here. The combination is not equally important for every project or moment. A bounded internal automation may need strong product, engineering, risk, and user participation without a full leadership engagement. An enterprise operating-model change needs a broader coalition. Earlier cloud migrations also carried major strategic, workforce, security, financial, and operating implications; describing them as substantially technical would repeat the mistake this chapter describes.

AI-Native work makes the interdependence unusually visible because probabilistic systems can mediate knowledge, decisions, communication, and delegated action. Strategic and human work cannot rescue unreliable technology, and technical competence cannot determine what the company should optimize or what it owes affected people. Failures can begin in either domain or at their interface.

That is why the combination matters now. Leadership teams need access to both sets of capabilities and a way to integrate them over time. They may find that in Arc5 and Nodalix, assemble it from other firms, build it internally, or use a hybrid. The relevant comparison is not

one engagement versus separate firms in the abstract. It is which arrangement provides the competence, independence, continuity, domain evidence, accountability, and challenge the specific work requires.

This section reads as advocacy for the firms behind this book. It is. I have tried to state why I believe the combination is well-suited to the work while making the claims testable. A buyer should ask what is proprietary, what is transferable, what evidence supports past outcomes, how fees and incentives work, who owns the artifacts and data, what happens if the firms disagree, and how the client exits without losing its operating knowledge. The leadership team that chooses this combination and the team that does not can both be making reasonable decisions. A small, bounded project may not justify it. A client with strong internal capabilities may need targeted support rather than an integrated engagement. The book is useful either way: engage these firms, assemble equivalent capabilities elsewhere, or use the questions to test the capability already inside the company.

The next section describes the specific engagement that a leadership team can take to start the work: the Alignment Sprint.

Section 9.2

The Alignment Sprint

How does a leadership team begin?

A leadership team that has read this far and is moved by the argument still has to decide what to do on Monday morning. The work I have been describing is substantial. The full transformation is the work of years. What a leadership team needs is a way to start that is concrete, bounded, and produces enough value on its own to justify the engagement before any commitment to the larger work is required.

The Alignment Sprint is Arc5 and Nodalix's answer to that question. It is a designed engagement with a defined scope and specific outputs the leadership team can use to decide whether and how to proceed. It is an on-ramp rather than the transformation: a bounded engagement intended to improve alignment, expose disagreement, and frame the work. Whether it creates enough shared substrate has to be demonstrated after the room empties.

Figure 9.2a. The Alignment Sprint engagement timeline, with four phases and six outputs.

The Premise of the Sprint

My premise is that a leadership team beginning this work needs enough shared understanding to make coherent decisions. Analytic content, frameworks, and case studies are available. What is not available off the shelf is the team's own evidence-based view of where the company is, what it intends, where members disagree, and what authority they are prepared to exercise.

An Alignment Sprint is designed to produce a documented working view. It brings the leadership team into a structured engagement with preparation, artifacts, facilitation, and technical challenge. “Shared” does not mean unanimous or final. Material disagreement, uncertainty, missing voices, and decisions deferred for evidence should remain visible.

The central workshop is designed around two days because executive calendars are constrained. That is a commercial and practical design choice, not proof that two days are sufficient for every team. Preparation and follow-through do require ongoing time from executives, the client design team, subject-matter experts, and affected participants. If the organization cannot provide that time, the sprint should be rescoped rather than sold as effort-free alignment.

The Pre-Work

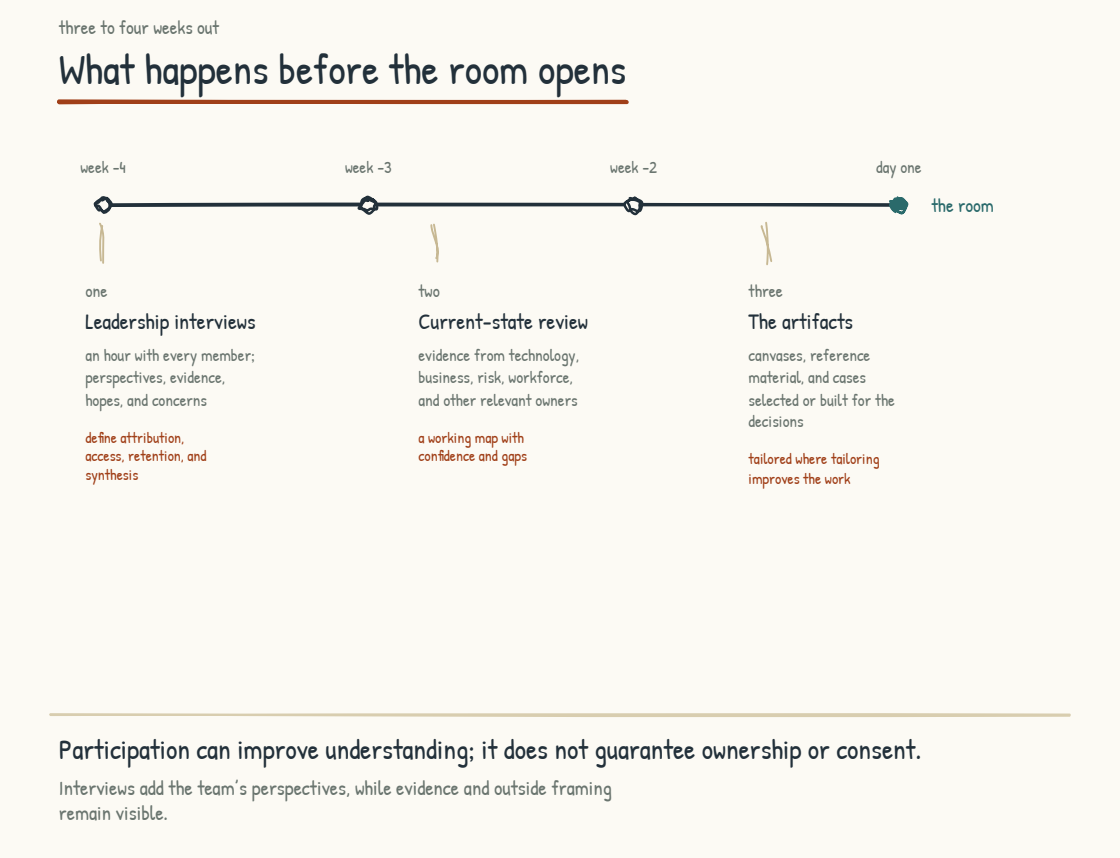

Figure 9.2b. Three to four weeks of pre-work: confidential interviews with every member, a current-state review, and artifacts built for this team from what the first two surfaced.

Work begins three to four weeks before the workshop, in several components.

First, a structured set of leadership-team interviews. Each member has a one-hour conversation with the Arc5 facilitator about the company, AI-Native, hopes, fears, alignment, and disagreement. Beforehand, participants need to know what will be attributed, what will be synthesized, who will see notes, how long records will be retained, and the limits of confidentiality. The facilitator then converts themes into a deidentified synthesis where group size permits, preserves material divergence rather than averaging it away, and gives participants a way to correct factual misrepresentation without rewriting an uncomfortable finding. The interviews surface hypotheses and perspectives rather than diagnosing a person's hidden mental model or producing recommendations.

Second, a review of the company's current state. This involves a structured assessment with technical, business, workforce, risk, legal, security, data, and customer perspectives proportionate to the scope. It examines the operating system, architectures, human work, practices, and systems view. Evidence may include policies, workflow maps, system inventories, incident history, employee and customer research, contracts, performance data, and samples of actual work, accessed under appropriate permissions. Evidence may contradict executive perception. The output is a working map of conditions, gaps, strengths, uncertainty, and claims requiring validation rather than a maturity score or consultant verdict. Sources and dates travel with the map so the workshop can distinguish a current fact from a remembered story.

Third, preparation of the workshop artifacts. The artifacts are designed for the specific leadership team, based on what the interviews and the current-state review have surfaced. The artifacts include the working canvases the team will use during the workshop, the reference materials they will draw on, the case studies that map most closely to their situation, and the structured exercises that will produce the workshop outputs.

Pre-work is typically planned for three to four weeks because customization matters, although availability, scope, regulated review, accessibility, and evidence collection may require a different schedule. A generic workshop risks generic outputs. Custom artifacts improve relevance only if they represent the evidence fairly and do not steer the team toward a predetermined answer.

The pre-work is also where the client's design team shapes the engagement. A typical design may involve about six meetings: objectives, inputs, outputs, constraints, flow, assignments, participation, breakout composition, accessibility, and a final walkthrough. The design must support remote participation, disability access, language needs, psychological safety, breaks, and alternatives to spontaneous speaking; otherwise the room will confuse ease of performance with quality of judgment. That participation can build understanding but does not transfer ownership automatically to everyone who will live with the results. The design team should document who was involved, who was absent, what was fixed before the workshop, what remains open to challenge, and which decisions cannot legally or ethically be made by this group alone.

The Two-Day Workshop

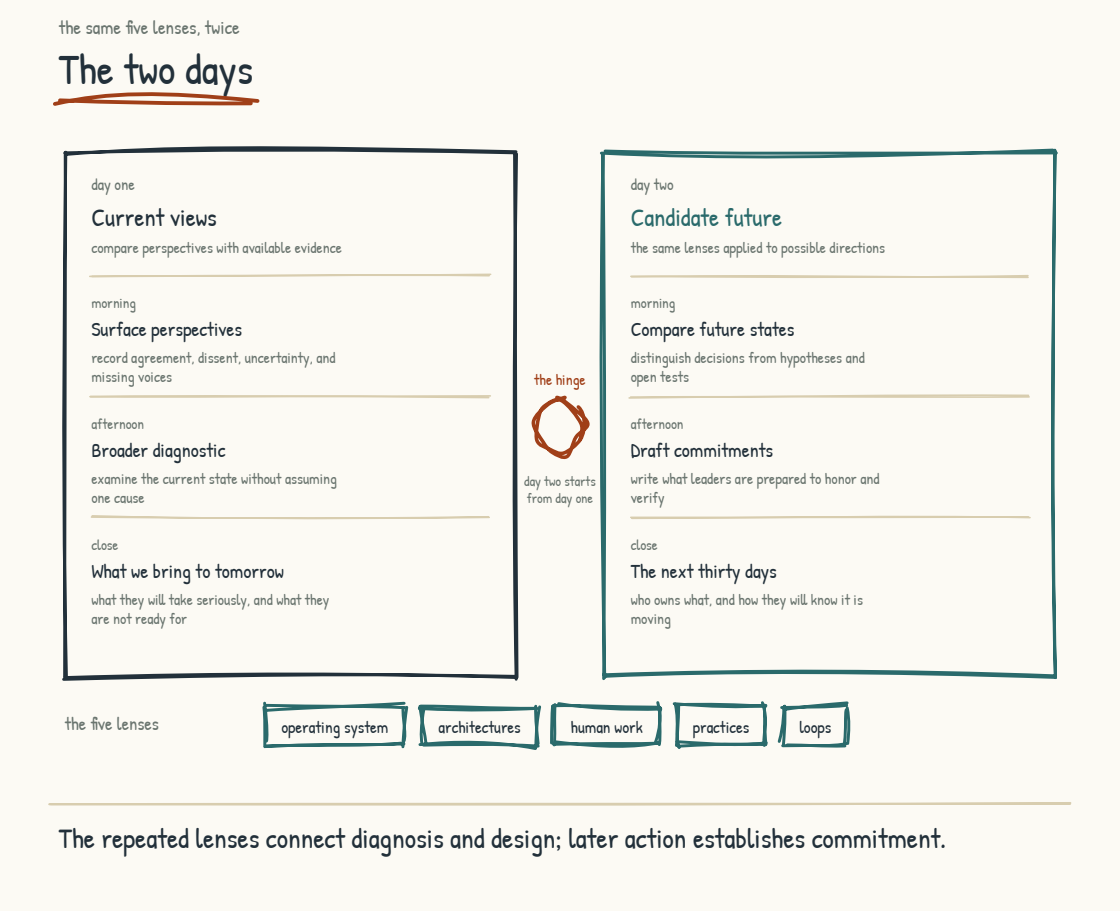

Figure 9.2c. The two days. Day One builds a shared view of where the company actually is, Day Two walks the same five lenses toward what the team commits to build.

Two days, and each has a distinct purpose and structure. The structure is consistent across engagements, with the content customized to the specific leadership team.

Day One is shared reality, or as close to it as the available evidence and participants can get. Each member brings a view of what is working, what is not, and what is at risk. The morning surfaces those views and tests them against pre-work evidence. Good facilitation can reduce unnecessary defensiveness and make political dynamics discussable; it cannot remove power from the room. The output is a working account with agreements, disagreements, unknowns, and perspectives still missing.

The afternoon of Day One is the deeper diagnostic. The team examines the current state through the book's lenses: operating system, architectures, human work, practices, loops, constraints,

and affected groups. The facilitator distinguishes observation from inference and prevents blame from substituting for a causal account. A working understanding can emerge, but any claim about workers, customers, systems, or legal obligations that has not been validated outside the room remains a hypothesis.

The day closes with the team's commitment to what they are bringing into Day Two. The commitment is short. It is the team naming what they are willing to take seriously, what they are not yet ready to address, and what they are committing to wrestle with the following day.

Day Two is shared future intent. The morning constructs candidate directions through the same lenses used for the current state. The team sketches two or three operating-model candidates from Section 4.2 and tests them against known constraints. In two days it may choose a direction, eliminate an option, or commission evidence needed before choosing. It should not manufacture certainty to satisfy the agenda. Timelines, investments, tradeoffs, reversibility, and affected groups become explicit, with formal approval reserved for the bodies that hold it.

The afternoon begins the covenant work. The team writes the commitments it can make as leaders, the changes it proposes, what it hopes to ask of the workforce, what it would leave behind, and what it would carry forward. Leaders can sign their own commitments. They cannot sign for workers or convert consultation into consent. The draft then goes through workforce participation, legal or bargaining processes where applicable, revision, and a clear response to dissent. It becomes a covenant only when the parties whose obligations it names have had a legitimate role in making it.

The day closes with the team's articulation of next steps. What it will do in the next thirty days. Who owns each action and who must be consulted or approve it. What baseline, evidence, and stop condition will show progress or harm. What support and budget are required. The decision log records the choice, rationale, evidence, dissent, dependencies, review date, and whether the action is reversible. Six months later, the log lets the team recover what it decided and why.

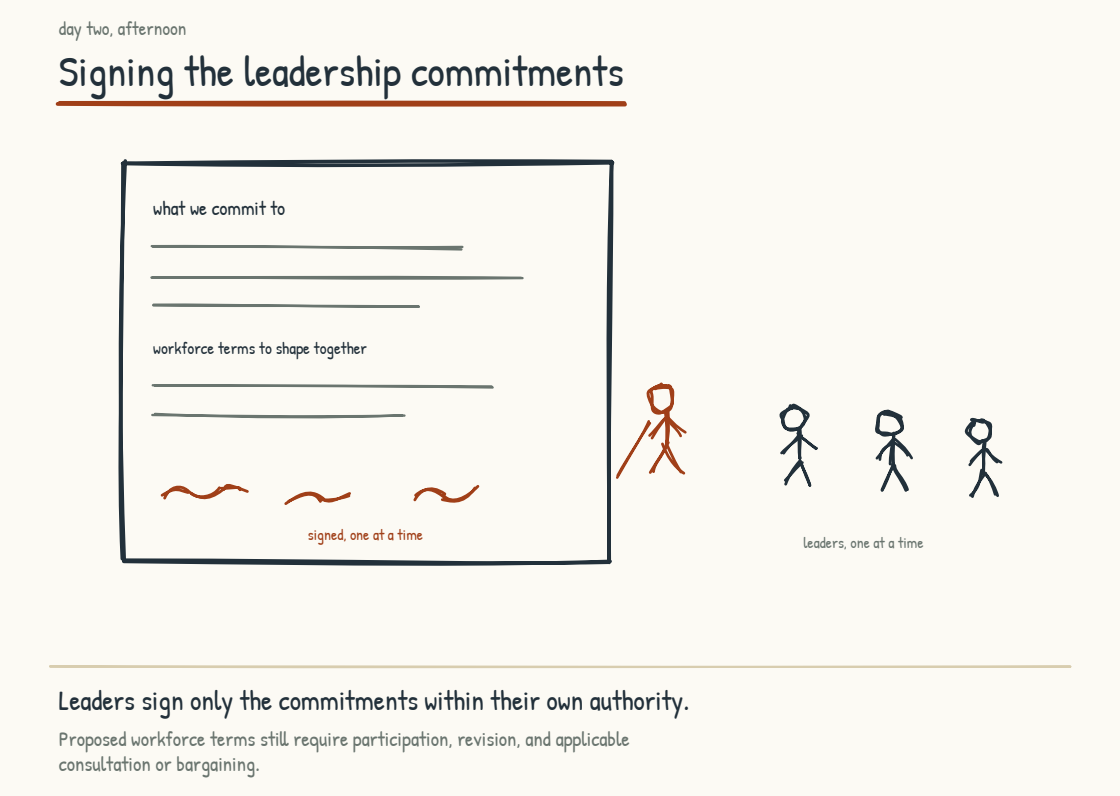

Figure 9.2d. Signing the leadership commitments. Written down and owned one at a time; the proposed workforce covenant still requires participation, revision, and any applicable consultation or bargaining.

The Outputs

A sprint produces specific outputs, and any leadership team considering it deserves to know exactly what the sprint hands them.

First, the documented current-state assessment: the team's working account across the operating system, architectures, human work, practices, and systems view, including evidence, disagreement, uncertainty, and missing perspectives. It is versioned so later evidence can change it.

Second, the future-state intent: where the team proposes to take the company, with horizons, investments, tradeoffs, decision rights, affected groups, and unresolved choices explicit. It sits beside the current state so assumptions and changes remain traceable.

Third, signed leadership commitments and a draft covenant. The leaders own specific promises within their authority. Proposed asks of the workforce are labeled proposals, not commitments made on others' behalf. The artifact becomes the basis for participation and revision, not merely communication.

Fourth, the thirty-day plan. Who is doing what, by when, under whose authority, with whose participation, and with what evidence and stop conditions. The plan is short enough to use but includes capacity, budget, dependencies, communication, risk ownership, and the first review date. It is not a miniature transformation roadmap. Its purpose is to test whether the leadership team can convert workshop commitments into governed action and revise them when reality disagrees.

Fifth, the larger roadmap framing rather than the full roadmap: major phases, dependencies, decisions, evidence gates, participation, and possible horizons. Twelve to eighteen months may be a planning range for some companies, not a promise or default. The framing helps the team decide what to commission next, whether with these firms, others, or internal teams.

Sixth, the experience itself. The team has spent two days working together in a way that, in my experience, many leadership teams find unusual. Afterward, look for changes in how members disagree, ask for help, explain decisions, and keep commitments. The experience is an output; durable collaboration is an outcome still to earn.

What Happens After the Sprint

A sprint starts the engagement rather than ending it. What happens after the sprint is what determines whether the work the team did in the room translates into the company they intended to build.

The first thirty days create an early credibility test, though they are not inherently more consequential than later implementation. Leadership commitments need owners and action. Proposed workforce commitments need listening, consultation, and revision before launch. The team communicates what it decided, what remains open, and how people can influence it. A weekly leadership review may fit; the cadence should match the work and include harms, dissent, and stop conditions as well as progress.

The firms typically offer light support during this period: a check-in, access for questions, and maintenance of agreed artifacts. The scope, decision rights, fees, data handling, and exit should be explicit. Facilitators cannot make sure momentum survives; they can surface drift, challenge rationalization, and help the team use the process it designed. The client remains responsible for execution and for creating channels where people can disagree without going through the advisers.

After the first thirty days, the team should have a clearer view of what the larger work involves. It has early execution evidence, feedback from affected people, and a record of which commitments held, changed, or failed. It should also review whether the sprint itself created value: which decision changed, which assumption was corrected, which risk surfaced earlier, and what burden the process imposed. If the answer is only that the workshop felt aligned, the evidence is too weak to justify a larger engagement.

The next phase varies. It may be additional discovery, workforce participation, a policy or bargaining process, a bounded workflow pilot, data remediation, capability development, or technical architecture. Where Nodalix is engaged, the design may use a graph and agents if the use case justifies them, or simpler components if it does not. Human and technical work often interleave, with evidence from each changing the other.

The Alignment Sprint is one possible on-ramp. The work after it is the transformation. The sprint may improve the odds by making assumptions, commitments, and interfaces visible. It is neither necessary for every company nor sufficient for any company, and its value should be judged by what the organization understands, decides, changes, and sustains afterward.

Why This Works as a Starting Point

Why might this work as a starting point? First, it offers a bounded commitment before a larger engagement. The team commits to interviews, evidence review, design participation, two workshop days, and thirty days of initial action. That is feasible for some teams and too compressed for others. The facilitation team carries much of the preparation, but executive and participant attention cannot be outsourced. The engagement should proceed only when the people needed for a credible start can participate.

Second, it is designed to connect analysis with decisions. Not every issue should be decided in the room: some require evidence, consultation, technical testing, legal review, or formal approval. The valuable output is a clean distinction among decisions made, hypotheses to test, questions to take to affected people, and choices not yet ripe. Ownership comes from authority and follow-through, not from the fact that a person moved a card on a canvas.

Third, it works across levels of the iceberg rather than claiming one “right” depth. The team responds to immediate constraints, examines patterns, maps structures, and tests mental models. A two-day workshop may expose an assumption; it cannot prove that the assumption

shifted or that durable change will follow. Decisions, incentives, resource flows, and observed outcomes provide that evidence later.

The first test comes when the team uses those outputs to make a decision after the workshop. The team also begins with unresolved questions and obligations to people who were not in the room. That is a more credible starting position than artificial certainty, provided the team follows through.

The sprint is the beginning. What happens across the years after it is the harder problem, and the next section takes it up.

Section 9.3

Beyond the Workshop

What does sustained work on this look like over time?

The Alignment Sprint ends with a thirty-day plan, and Section 9.2 followed the team through those thirty days. The work beyond them may take years. Capability, trust, and maintained institutional knowledge can compound, but none does so merely because time passes or information is captured. Between a thirty-day plan and durable operating change sits a long middle that swallows many good beginnings. The team leaves the room with working alignment, and the first month runs partly on the energy of the room. In week six the chief financial officer misses a check-in for board preparation. The reason is real, and everyone agrees to proceed without her. In week nine the first workflow redesign slips behind a client escalation, and two operator sessions dependent on it are rescheduled. Again, the reason is real. By week thirteen the signed leadership commitments and draft covenant may sit unopened while everyone remains busy doing work that mattered. No one announced an abandonment. No single decision killed the effort. The sequence reveals whether the sprint began an operating rhythm or created a good memory, because operating models are often lost through individually reasonable exceptions rather than open rejection.

This section uses three planning horizons: roughly the first ninety days, the rest of the first year, and year two and beyond. They are not laws of transformation. A regulated deployment, collective-bargaining process, acquisition, or urgent safety issue will require a different clock. The horizons organize questions about early evidence, system change, and longer-term

maintenance. They also frame signals that working alignment is changing and ways the book can remain useful after the workshop. Everything here can be done internally, with these firms, or with others; the right arrangement depends on capability, independence, consequence, and trust.

Days 31 to 90: Proving the Motion

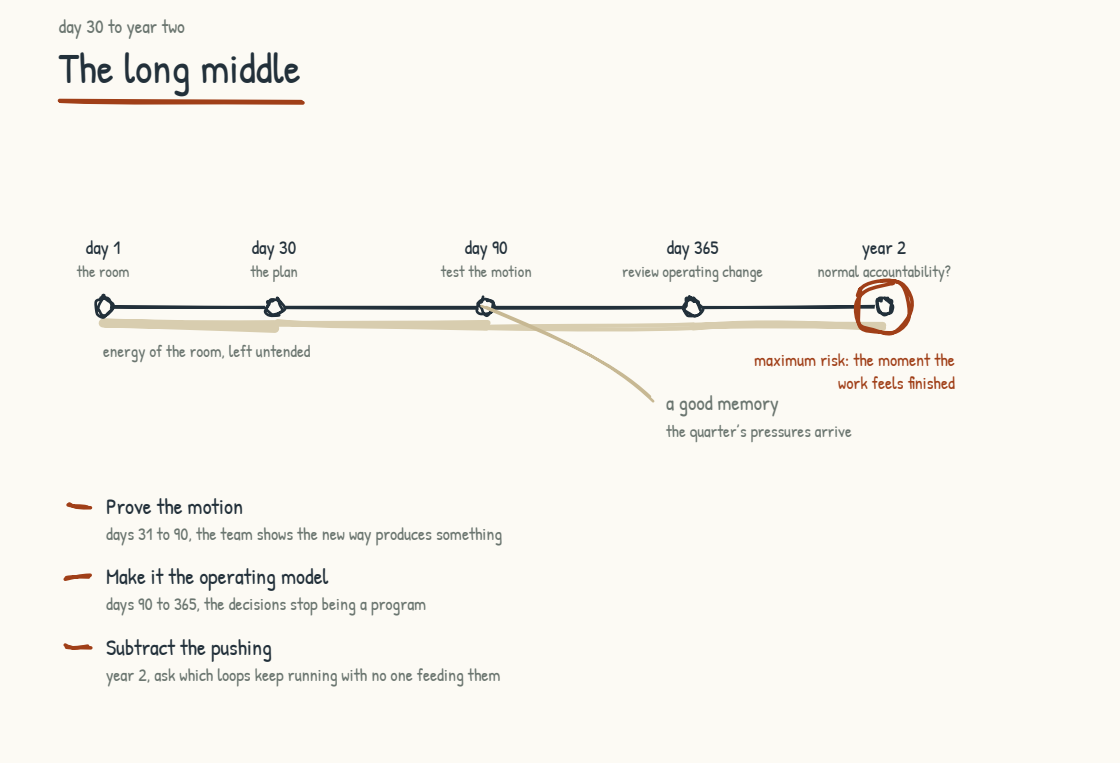

Figure 9.3a. The long middle, from the plan to the year the loops take over, with the branch where a sprint becomes a good memory instead of an operating rhythm.

That first horizon has one job: convert the sprint's claims into evidence people can inspect. I would look for three forms while allowing the scope and risk to determine how far each can responsibly move. The point is not to perform momentum for the board. It is to discover whether the proposed model survives actual data, actual users, ordinary interruptions, and the controls the workshop could discuss only in theory. Early evidence should include who benefited, who carried new burden, what failed, and what the team chose not to do.

The first is a redesigned workflow tested in the real conditions of use with the people who do and are affected by the work. For a low-consequence, reversible workflow, that may mean production use within ninety days. For consequential work, it may mean a shadow run,

simulation, controlled pilot, or a decision not to deploy. Case One in Section 8.2 described eleven pilots that did not become a coherent portfolio. The difference here is not that use is mandatory; it is that the experiment has a shared purpose, baseline, owner, evaluation, stop condition, and explicit route to scale, revise, or close.

The second is a knowledge-maintenance practice that has started where the use case requires it. It may feed a graph, document system, rules, training, or several forms. An experienced operator might work beside modelers to demonstrate, test, and document judgment, with protected time, credit, consent, and the right to contest the representation. This must not be capture rushed ahead of a workforce reduction so the company can extract knowledge first. Capability continuity and workforce planning should be transparent, participatory, and governed by the covenant's commitments.

The third is visible evidence of value and cost, described in terms the workforce recognizes. That may be safer work, less rework, faster access to evidence, a better customer outcome, or a burden that did not improve and caused the team to stop. Compare the result with a baseline, include quality and risk beside speed, and say what remains unknown. A worker may choose to describe the experience, but nobody should be recruited into testimonial theater or asked to speak for a whole role. Early evidence travels farther when peers trust the method, can inspect the limitations, and know that an inconvenient result will not be edited out of the story.

Through all of this, the leadership check-in from the sprint keeps running at the cadence the work needs. It is short when the evidence is routine and longer when a stop condition, workforce concern, security issue, or contested decision appears. The check-in is not a status recital. Its purpose is to remove a constraint, make a decision, revise an assumption, or stop work. At roughly day ninety it becomes a more deliberate checkpoint.

BRING THIS INTO THE ROOM

The ninety-day checkpoint. Reserve enough time for the leadership team and relevant participants to run five items in order. One: read the thirty-day plan and name what held, changed, or failed, distinguishing a justified revision from a euphemistic pivot. Put the original baseline beside the result. Two: test the loop hypotheses from Section 7.1 against evidence and competing explanations; redraw them where reality disagreed. Three: review the signed leadership commitments and the status of workforce participation in the draft covenant, including promises the company has

already missed. Four: decide whether the next workflow should scale, pause, change, or close against the roadmap framing, with risk and capacity visible. Five: stop or deprioritize something so new work has room. Record the decisions, dissent, owners, and review dates. Scheduling the checkpoint early makes it less vulnerable to drift; whether half a day is enough depends on scope.

Days 90 to 365: Making It the Operating Model

The second horizon is where the future-state intent from Day Two stops being intent. Three kinds of work dominate it.

The first is letting operating-model decisions affect mechanisms. Quiet reversals appear when the team says measurement will change but leaves the old scorecard tied to pay, or approves a collaboration surface while the sanctioned experience remains unusable and necessary controls remain unexplained. A role may be announced as higher altitude while workload, compensation, and promotion criteria still reward the old task volume. A governance policy may promise bounded experimentation while the approval queue still treats every use case alike. The first year is an exercise in finding where the old model persists in metrics, meetings, approval paths, incentives, budgets, systems, procurement, job design, and decision rights. Announcements alone change none of these. Mechanisms matter, and so do the people who experience their effects and can reveal when the new mechanism is worse.

The second is a wave rhythm. Later workflow redesigns are selected against the roadmap, evidence, capacity, and risk rather than by whoever asks loudest. The number and cadence of waves should follow learning and operational ability, not a promise to launch a second and third. Measurement matures across this horizon too, with metric displacement as the standing caution: use quantitative and qualitative evidence honestly, and accept that some covenant commitments will not reduce to a number.

The third is the covenant being tested for real after legitimate participation has made it one. A budget squeeze, reorganization, automated task, safety incident, or dispute over the work will expose whether commitments constrain leadership when doing so costs something. This is where broad language needs operational meaning. What does investment in people require when a role changes? What process precedes a workforce reduction? Who may appeal an automated recommendation? What happens when a promised learning hour collides with a client deadline? Skepticism from a senior person may be obstruction or an important warning;

evidence and conduct, not enthusiasm, decide. The welder's standard from Section 5.6 remains a useful personal test, joined by formal tests of fairness, rights, process, and impact. One honored commitment can build credibility, and one breach can damage it, but trust is cumulative rather than switched by a single event. Repair requires acknowledgment, remedy, and changed conduct, not a refreshed message.

A quarterly cadence is a useful starting point, not a rule. The team revisits loop hypotheses, practices, meaningful-human-involvement decisions, and covenant commitments often enough to affect action. High-consequence deployments may need continuous monitoring and faster review; slower capability change may need longer outcome horizons. Reuse the artifacts when they help and retire them when they become ceremony.

Year Two and Beyond: Maintaining the Work

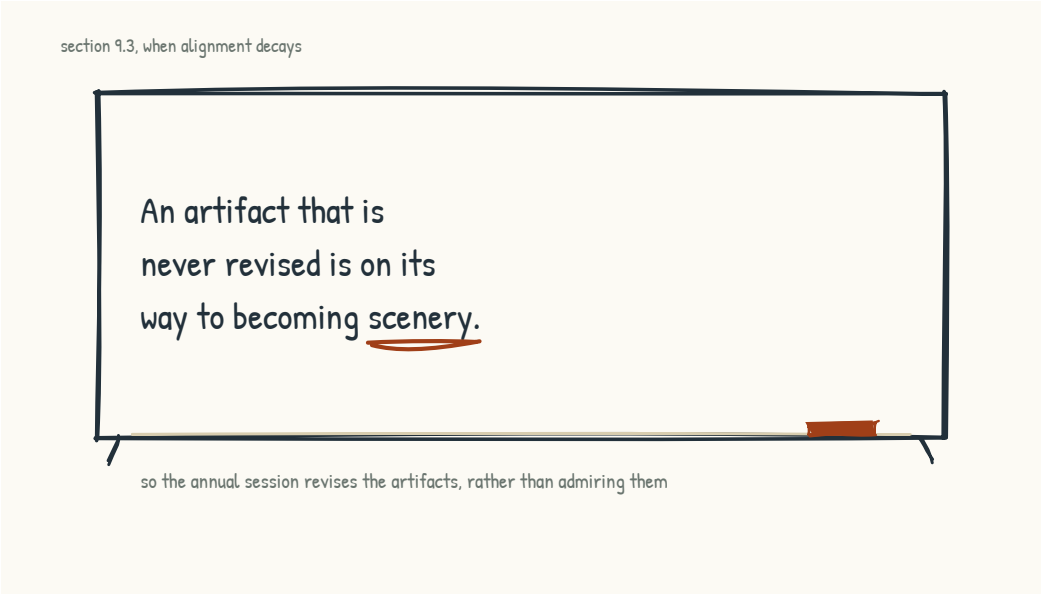

Figure 9.3b. The warning on the refresh cycle.

The test of year two is reduced dependence on executive attention, not its disappearance. Ask which practices persist through leadership distraction or turnover: maintained knowledge because contributors see fair value, routines that survive a champion's departure, supported development opportunities that do not depend on favoritism, and governance that catches an unsafe shortcut even when the launch has executive sponsorship. Then ask what is merely hidden by habit. A practice can survive while losing its purpose; a knowledge base can grow while becoming stale; a mentoring count can rise while access remains unequal. No

organizational loop is literally self-sustaining. Budgets, ownership, maintenance, governance, worker voice, and changing conditions remain. The transition is becoming part of the operating model when ordinary mechanisms carry it, exceptions are handled without improvising the values away, and people no longer need transformation rhetoric to explain routine work.

Two disciplines still need explicit ownership and a calendar, because compounding does not maintain itself.

The first is the refresh cycle. Models, vendors, law, threats, prices, organizational strategy, and workforce capability change at different speeds. At least annually, and sooner after material change, the leadership team re-examines the archetype, layer investments, use cases, and boundaries of automated and human authority. Current-state evidence, future intent, covenant, and roadmap are revised with affected participants and version history rather than left as souvenirs. An artifact that is never revised is becoming scenery; an artifact revised without governance is becoming revisionist history.

The second is protection. Careless workforce or measurement decisions can damage trust and practice quickly, though no single action determines the system forever. Decision reviews should catch foreseeable conflicts before approval, while incident and appeal processes handle what gets through. The period of maximum risk may be when the work feels finished, because attention and challenge fade while technology and conditions keep moving.

When Alignment Decays

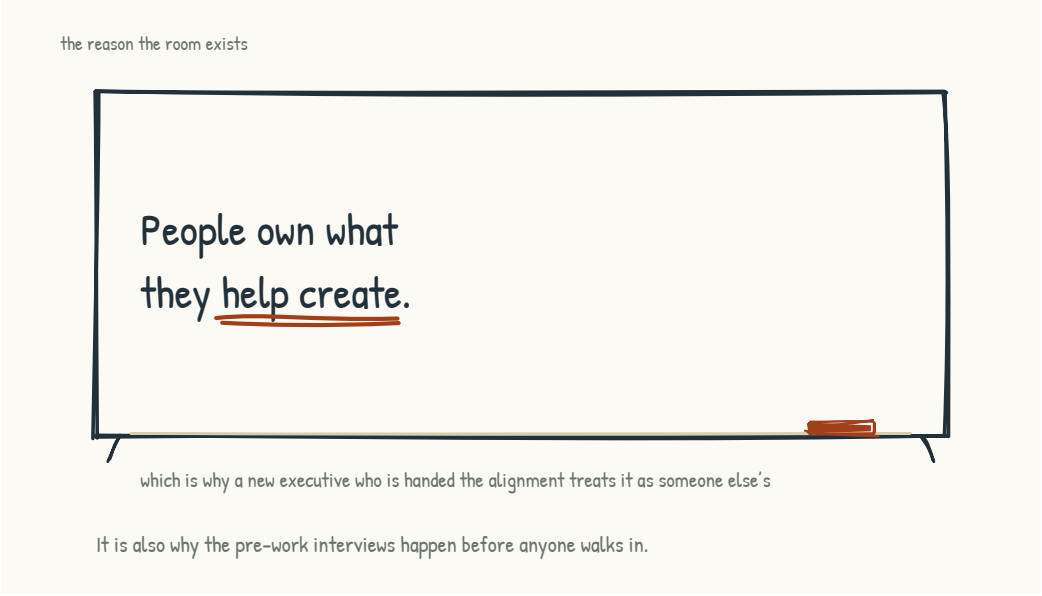

Figure 9.3c. The principle under the re-sprint. Alignment that was handed over belongs to someone else.

Working alignment is a produced state, and it changes. Nothing in the sprint makes a shared view permanent, nor should it: new evidence ought to create divergence before the team integrates it. People change roles, strategy moves, and private views develop. The problem is not disagreement returning. It is material divergence remaining hidden while coordinated decisions continue to assume agreement. In teams I have watched, that often appears through small reversions, each with a reasonable explanation, before anyone names the pattern.

> **WATCH FOR THIS**
>
> **Signs working alignment needs review. One: the same term drives incompatible decisions because leaders mean different things by it. Two: covenant commitments are absent when a relevant hard choice is made. Three: review rhythms repeatedly disappear without a better mechanism replacing them. Four: initiatives launch without portfolio rationale, evaluation, or learning paths. Five: leadership turnover changes decision assumptions or accountability. Six: an acquisition, new market, legal change, incident, or capability shift makes the future-state intent stale. None proves decay alone; each is a prompt to investigate.**

The response should match the signal. For limited drift without structural change, a focused realignment day may be enough: return to the evidence and artifacts, test the loop account, review commitments, name divergence, and revise intent. A capable internal facilitator may run it, provided power dynamics and conflicts do not require independence. The purpose is not to restore the old alignment automatically; new disagreement may be the correct response to new evidence. For structural change, such as a new chief executive, acquisition, material strategy shift, or major incident, repeat enough discovery to understand the new system. That may justify a full sprint, a different process, or no workshop at all. People often understand and support work they help shape, but a new executive does not automatically reject inherited decisions. Brief them on the rationale, evidence, dissent, commitments, and outcomes; invite challenge; include people affected by the new context; and redesign only what evidence warrants. External facilitation earns its keep when independence, skill, or capacity adds value, particularly when the sponsor is part of the conflict. It does not earn its keep simply because teams are categorically unable to examine themselves.

Working With This Book Between the Sessions

This field guide was structured so an agent can read it alongside you. The long middle is exactly when a team needs the book as a reference rather than as an argument.

The pattern I intend is simple. When a new situation arises, a team member can ask the book's companion agent what passages and exercises speak to it. At a checkpoint, the agent can guide the four-question diagnostic and organize answers, while citing the underlying sections, quoting accurately, and labeling inference. When a covenant commitment is under pressure, it can retrieve the actual language rather than an executive's memory of it. When the team plans a realignment day, it can draft an agenda and expose which affected voices are missing. The user should still verify that the response is grounded in the book and current company records. If the team connects its own artifacts, that must happen in an approved environment with purpose limitation, permissions, retention, provenance, versioning, and separation among personal notes, privileged material, workforce data, and shared records. Retrieval permissions should match the underlying source; the agent must not turn fragmented access into a new universal view. A graph is optional. Queryable and current does not mean universally accessible, automatically correct, or permanently retained.

The limit holds with full force. The agent can locate, connect, question, and translate, and it can participate in a mediated conversation if the people choose. It cannot confer consent, reconcile interests, assume accountability, or prove that a difficult exchange was honest. Use it to prepare and record within agreed boundaries. People with legitimate authority and standing still own the decision and its consequences.

The Work Is Yours Now

If you do one thing on finishing this section, schedule the first evidence review before attention disperses. Ninety days is a useful default for many low- and medium-consequence efforts; choose an earlier or later date when the risk, learning cycle, or required consultation demands it. Put the purpose, participants, evidence, and decisions on the invitation, not only the date.

That is the shape of the work over time. The argument began with a phrase on a whiteboard and has ended with a calendar because durable change needs recurring attention. It also lives in budgets, systems, contracts, authority, habits, and what happens when someone challenges the plan. The frameworks are yours to use, the diagnostics are yours to test, and the agent can return you to the material. None substitutes for accountable decisions or legitimate participation.

One section remains in this stage, and it is the most practical thing in the book. Everything here has described the shape of the work. Section 9.4 describes how the sessions actually run, in the detail a person would need to run one.

Section 9.4

Designer Notes

What does it actually take to run one of these sessions?

Everything in Stage 9 so far has described the shape of the work. What it is for, what it produces, how it starts, what happens in the long middle after the room closes. That leaves out the part I get asked about most, which is how the sessions actually run. The mechanics. How many people in a breakout, how long a round goes, what is physically on the wall, what I say to launch the work, what I do when a group rejects the assignment.

These are the details a leadership team debates with me during design, and they are details clients often want to cut. I have strong defaults about them, and this section explains why. They come mainly from practitioner experience across many rooms and industries, not controlled evidence. A default should carry accumulated learning without hardening into ritual. Accessibility, culture, power, safety, purpose, and the people in the room can all require a different design.

I have written this as questions and answers, because that is the form in which they come up. A sponsor asks, I answer, and the reasoning underneath the answer is usually the useful part.

The Sponsor Conversation

I like to open by getting the sponsor talking about the challenge in front of them. In their own words. Then I get them to reframe it as what success looks like: what has to get accomplished, and separately, what would be nice to have. Sometimes I ask what they would want if they could wave a magic wand and the group went further than they expected.

I ask that last one deliberately. Most executives have learned to have low expectations of meetings and workshops, because most meetings and workshops are not designed very well. So I spend time early explaining how our methodology is different and why. Usually the sponsor is curious and starts asking questions, and that conversation matters as much as the answers.

What I often hear afterward is a version of this: I did not understand what you meant by different, but after going through it, now I do. That feedback is encouraging and may mean the experience made the method legible. It is not outcome evidence by itself. A session can feel distinctive and still fail to change a decision, include the right people, or improve what follows.

The thing sponsors most often ask for that they do not need is a smaller room. Leaders want to limit the stakeholders. Sometimes that is about controlling the message. More often it is that they do not want to pull people off their day jobs, or they believe a small group can do the work while everyone else stays focused on whatever they are doing.

My argument against shrinking reflexively is practical. The last thing you want is to discover in the room that a missing person holds essential knowledge or authority. So I push on four questions. Who holds the decision rights? Who knows how the work actually operates? Who bears the consequences? Who cannot safely or practically participate in this format but still needs a voice? The answer may expand the room, create interviews or parallel channels, or narrow what the room is authorized to decide.

There is also a legitimacy argument. Implementation benefits when relevant knowledge and affected perspectives shape the decision. That does not mean every stakeholder must attend one large session or leave aligned. Large rooms can reproduce hierarchy, exclude people who cannot attend, and create performance rather than candor. Participation should match the decision, and the process should show how input influenced it.

> *The last thing you want is to be in the middle of the session saying, I wish so-and-so were here to answer that.*

What makes an executive anxious about the methodology is usually the open conversation. Not being able to control the agenda or the message. Loss of control. Some leaders worry that people will spend the day churning on something the leader has already decided.

I think controlling the conclusion too tightly is short-sighted. People process challenges differently, and structured discussion can help them surface knowledge and recommend options. It is not therapy, and participation does not guarantee anyone feels heard or trusts leadership. Trust depends on power, follow-through, candor about what is already decided, and

an explanation of how recommendations will be used. Inviting input on a closed decision is worse than naming the boundary plainly.

The other thing that makes sponsors nervous is not knowing what I will do in a difficult moment. I explain the facilitation roles, escalation paths, design boundaries, and how I will check with the sponsor and client design team without letting the sponsor secretly rewrite participant input. The team keeps the work oriented to the stated purpose while protecting agreed participation rules.

I also tell them we can call audibles. We can notice that a group is going too deep, not deep enough, or in a more valuable direction and adjust mid-session. Adaptation is a capability, not a guarantee of a home run. Material changes should be explained when they affect purpose, time, or how outputs will be used, so flexibility does not become invisible manipulation.

I do turn work down. Usually when the leader will not follow the methodology, which normally means they want to control everything and do not really want people doing the work. Or they want the work done in principle, but they have no intention of using the output. Either one is a bad fit.

Why We Do Not Publish the Agenda

The design team keeps a detailed run-of-show. Participants receive a purpose, broad arc, start and end times, breaks, accessibility information, preparation, decision boundaries, and what will happen to the output. Whether they receive every exercise in advance depends on the work.

One reason to hold exercise-level detail loosely is that the design may change as evidence emerges. The facilitator does not need to dramatize every small adjustment, but the aim is not to change course without anyone noticing. Participants deserve to know when the purpose, promised break, use of their input, or decision process changes.

A second reason is to preserve attention and discovery. Yet withholding information can increase anxiety, disadvantage people who need preparation, and undermine informed participation. The answer is progressive disclosure, not manufactured surprise: give people enough structure to plan and consent, then introduce exercises as the day develops.

What we are managing across the session includes attention, cognitive load, safety, and energy. A breakout of seven can absorb natural variation, but a larger room also creates noise, social load, and access barriers. Fatigue is not something the group should conceal by having other

people carry it. Build a humane rhythm, offer multiple participation modes, and size the room for the work rather than treating more bodies as a battery.

Group Size and What Each Group Does

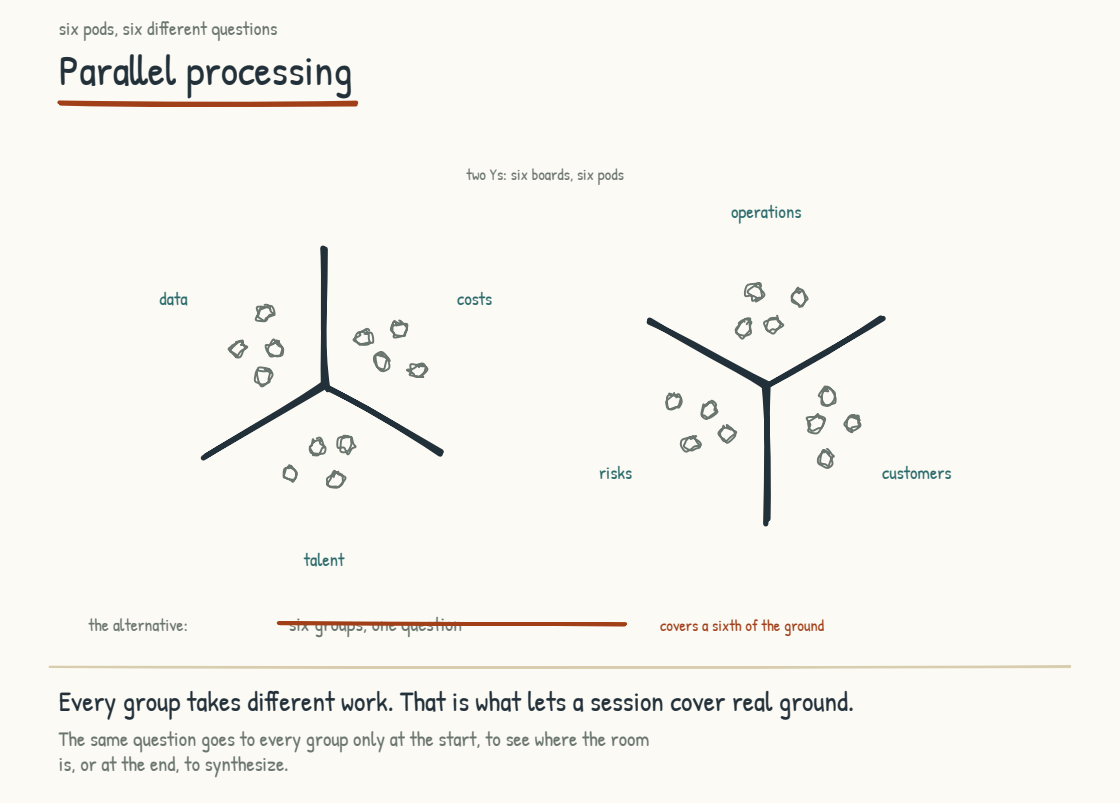

Figure 9.4a. Five studios, five different questions. Every group takes different work, which is what lets a session cover real ground.

We want five to seven people per breakout. Eight is usually too big. With forty people I would probably run groups of six or seven, though it depends on the work. If there are eight distinct things that need focused attention, we will run eight groups of five instead, and sometimes one team gets eight people because its topic needs them.

Group size changes the work itself, beyond the logistics. What you get from a three-person team is different in kind from what you get from a seven-person team. So is the diversity of thinking. That is worth designing for rather than treating as a scheduling detail.

We assign groups rather than letting people pick most of the time, because role, knowledge, power, and cross-functional contact matter. We review rosters with people who know the participants while avoiding tokenism, retaliation risk, inaccessible movement, and combinations

in which a reporting relationship will silence the conversation. Occasionally people choose a topic, with an accessible alternative to “voting with their feet.” Participants can flag a conflict or accommodation privately before the session.

On whether every group takes the same assignment: usually not. We run parallel processing, where each group works something slightly different, because that covers far more ground. Imagine thirty-five people all answering the same question every round. Your scope would have to be tiny.

Same-assignment rounds have their place. At the beginning, a visioning question everyone answers tells you where the room actually is. Sometimes I will have people work alone first, getting their own thoughts down before any group dynamic kicks in. And near the end of a day, a shared question can help synthesize what has been happening.

How many rounds depends on how much work there is and how long we have, and mostly on how many teams are in the room. Three teams means fewer report-outs, which buys me another activity or two. A large group with a lot of report-outs costs me one. In a one-day session you might get six. A lot of that math is decided by how long each team gets to share. Report-outs can run one minute per team or ten, and you can burn an afternoon on the difference.

Every breakout works from a written assignment, and a good one has three parts. The context tells the group why this piece matters and what sits in scope. The objective says, in a sentence, what they are producing. The process is an ordered list of questions that walks them from what they were handed to what they hand off. The assignment is the bridge between the inputs a group starts from and the output the next round needs, and writing those bridges is most of the design work between the block diagram and the room. The whole thing fits the front of one page, in a font large enough to pass around a table. Aim at the page. A page and a half is fine, because a designer can cut what is there and cannot invent what is missing. Keep the context to a few lines locating the round in the day and naming its inputs. On a very short round, cut it to nothing and give the page to the objective and the process. The round's duration belongs on the sheet. Leave the minute-by-minute breakdown on the facilitator's copy, since a time against each step reads as a schedule the group has to manage. Watch the task count against the clock too. A fifteen-minute round gets one or two things to do, and an assignment can fit the page and still ask for six.

A round also has a rhythm inside it. For a sixty minute round we plan five minutes for the group to read the assignment, forty-five for the real work, and ten to synthesize, prioritize, and get the report ready. The group picks a note-taker in the first minute. Scale that shape for longer or shorter rounds, and always protect time at the end to synthesize, because a group that works to the buzzer walks to the front with nothing organized.

Timing, Breaks, and Energy

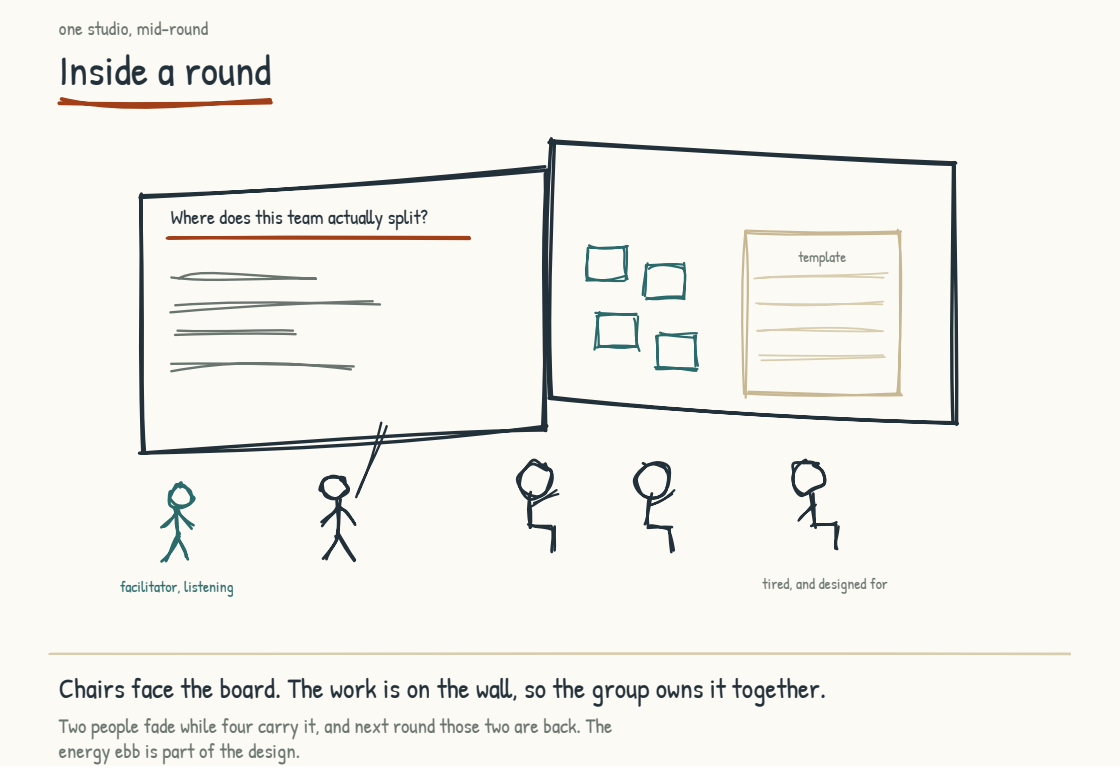

Figure 9.4b. Inside a round. The work is on the wall so the group owns it together, and the energy ebb is part of the design.

Breakouts run thirty to forty-five minutes. Then we bring everyone back to the large group quickly so all the work is visible to everyone.

The morning is tighter than the afternoon. I keep morning rounds at an hour or under, and the first round of the day often runs forty-five minutes on purpose. It teaches the room how fast the work moves before anything harder gets asked of them. Longer rounds belong after lunch, once the groups know the form.

There is also a ceiling on how long the room stays apart. About an hour and forty-five minutes of work before everyone comes back together, and reaching that number takes a convergence

inside the stretch. The shape that gets there is a forty-five minute round, a shift-and-share of twenty to thirty minutes, half an hour back home to finish, then the report out. Counting rounds is the wrong instrument. Two long ones back to back can breach the ceiling while three short ones with a shift-and-share between them stay under it.

Plan predictable breaks and enough time for people to return. People need predictable breaks for health, medication, prayer, caregiving, sensory regulation, private calls, and basic cognitive recovery. Publish them, start and end them clearly, and design the day so returning on time is realistic. A short overrun is a design cost, not participant failure.

Working lunches can be offered for a genuinely optional activity, but food should not become another compulsory work block. Provide an actual meal break and accommodate dietary, disability, religious, and medical needs. If the agenda cannot survive people eating without producing output, the agenda is overfilled.

Change posture regularly without requiring everyone to stand or move. Some people think better seated, use mobility devices, need captions, or cannot tolerate music and sensory stimulation. Offer snacks, quiet space, seating choices, accessible boards and digital equivalents, and advance notice of sound. Watch the late-afternoon dip as information about the design, not a motivation problem.

Hold contingency time. In a two-day session, part of the second afternoon may remain flexible based on Day One, but participants should know that in advance. Unallocated capacity also absorbs accessibility needs, conflict, technical failure, and work that deserves longer than the run-of-show predicted.

Coming Back Together

Every time the room goes out, something has to happen with what comes back before it goes out again. After the chatrooms, a plenary debrief before groups head to their first round. After a work round, a report-out or a shift-and-share before the next one starts. Skip it and the next round inherits raw material nobody has examined.

The deeper reason is what parallel processing costs. Every group works something different, which is what lets a session cover real ground, and it means most of the room did not see most of the work. People who never observed what the other groups did may feel disconnected from the output and have less basis for standing behind it. The convergence is where that gets repaired.

It carries a second job. While the room is together, the crew resets the breakout spaces, hangs the next assignment, and gets everything ready for the round that follows. Design the convergence and you have bought that time.

A commitment round has to be earned. On a one-day session that gets to the shape of the problem and no further, closing with commitments is ceremony, because nothing has been produced yet that anyone can sign. Action planning is itself a commitment round. People putting their names against specific items is the commitment, and it leaves the room as an artifact. A separate commitment block earns its time on a two-day session, after the work has gone somewhere.

The Room

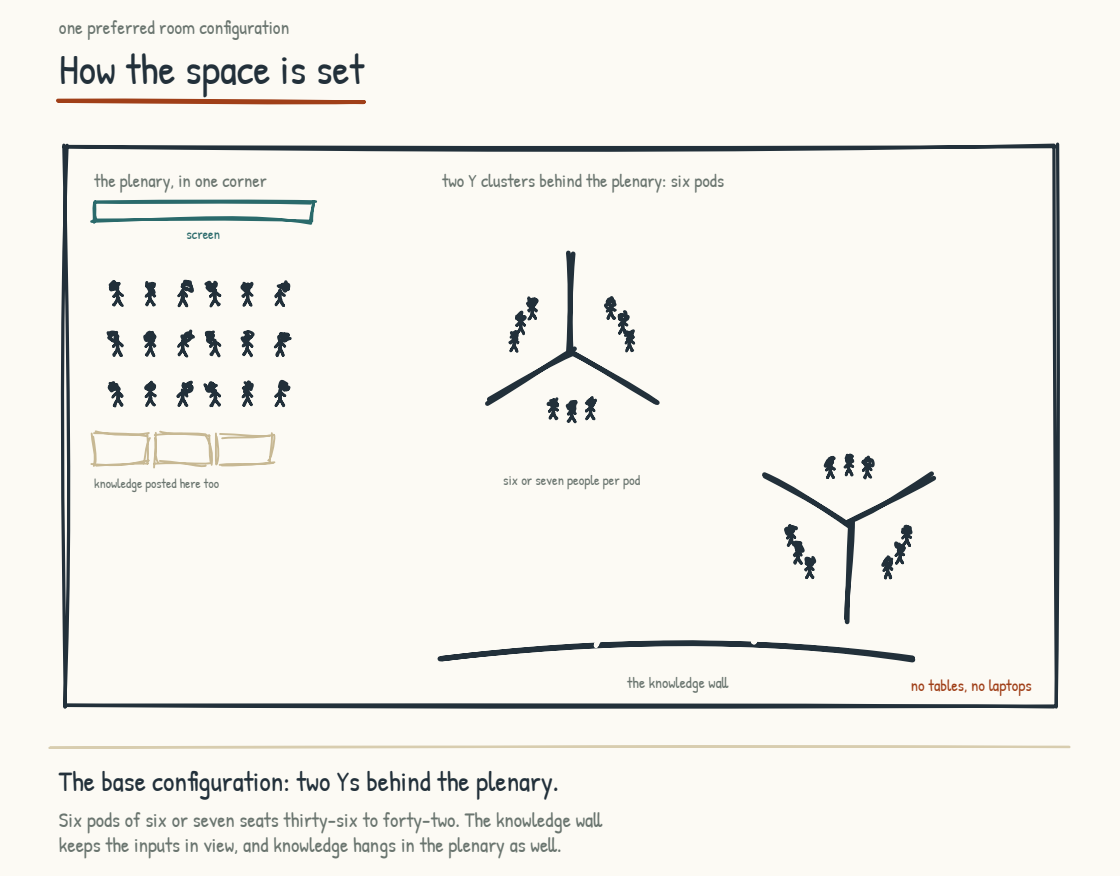

Figure 9.4c. The base configuration: the plenary in one corner, two Y clusters of breakouts behind it, and the knowledge wall.

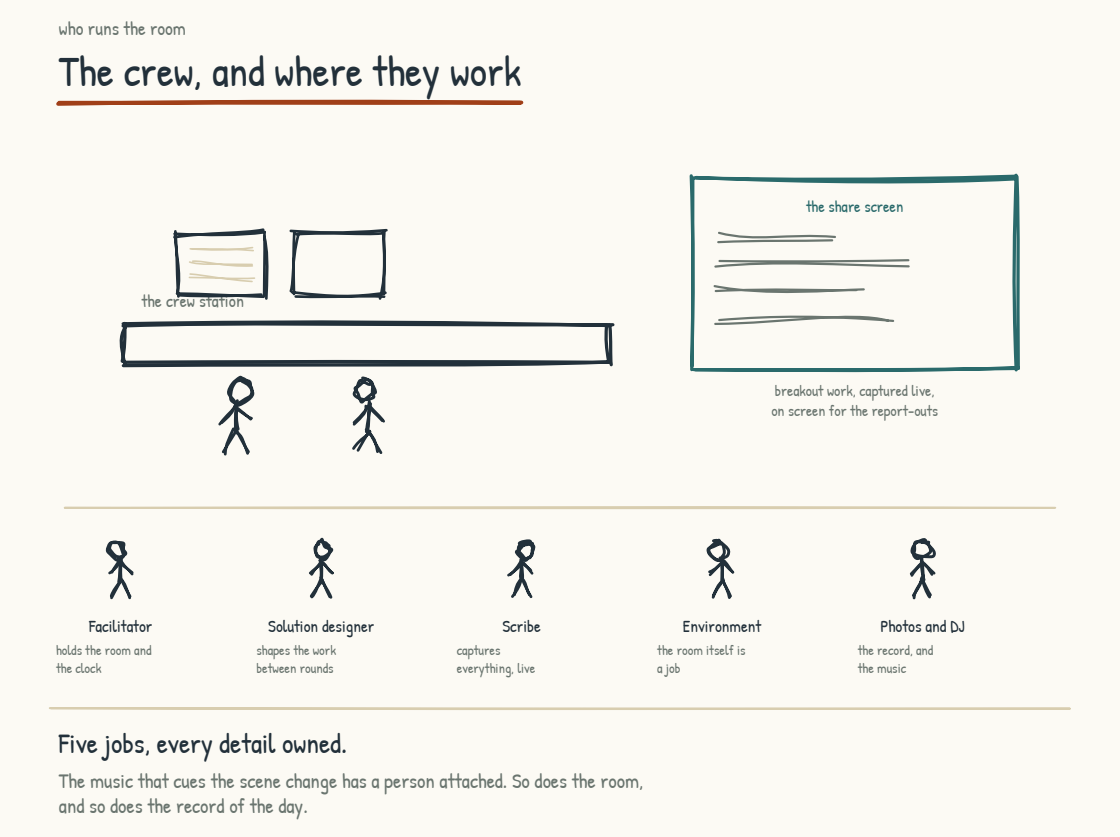

Figure 9.4d. The crew and where they work. Five jobs, every detail owned, and the music has a person attached.

We need roughly one hundred square feet per person. A thirty-person session wants about three thousand square feet, and that also buys the wall space to post every round of output where people can walk back to it.

There is a plenary area where the sponsor opens and where teams share their work. Around it are breakout areas sectioned off by markerboards, each with chairs and boards, sometimes a screen.

My default has been no tables in breakouts because tables invite laptops and passive posture. That default must yield to access. Some participants need a writing surface, device, assistive technology, water, or physical support. Design for active shared work with tables, adjustable surfaces, boards, or digital tools as needed; do not equate comfort with disengagement.

When people walk in, we often have a walkabout waiting. Questions are already up on boards so participants can reflect on history, frustration, or prior failure without letting the past consume the day. This is facilitated reflection, not therapy, and a public wall is not appropriate for confidential, traumatic, or personally identifying material. Offer seated, digital, individual, and

nonverbal ways to participate. The signal is that people are contributors from the beginning, not an audience being warmed up.

We also bring elements people may not expect: music, photography, or a visual artist. Each requires design and consent. Music needs volume limits and quiet alternatives. Photography needs clear purpose, storage, access, and meaningful opt-out without social penalty; never assume people love being photographed at work. Visual recording must avoid exposing confidential or attributed content unless participants agreed.

Virtual, hybrid, and distributed sessions require a design that gives remote participants a full part in the work: equivalent facilitation, accessible technology, deliberate turn-taking, digital artifacts, shorter blocks, and local support. If equivalent participation cannot be provided, name the limitation and create another legitimate channel rather than treating absence as consent.

Some sessions add a screen to each pod. The group still thinks on the marker boards, and the final output gets captured digitally before the share, which makes the report-outs instant and means the work leaves the room as files the team can open the next morning. The boards carry the thinking, and the screen carries it out the door.

Launching the Work

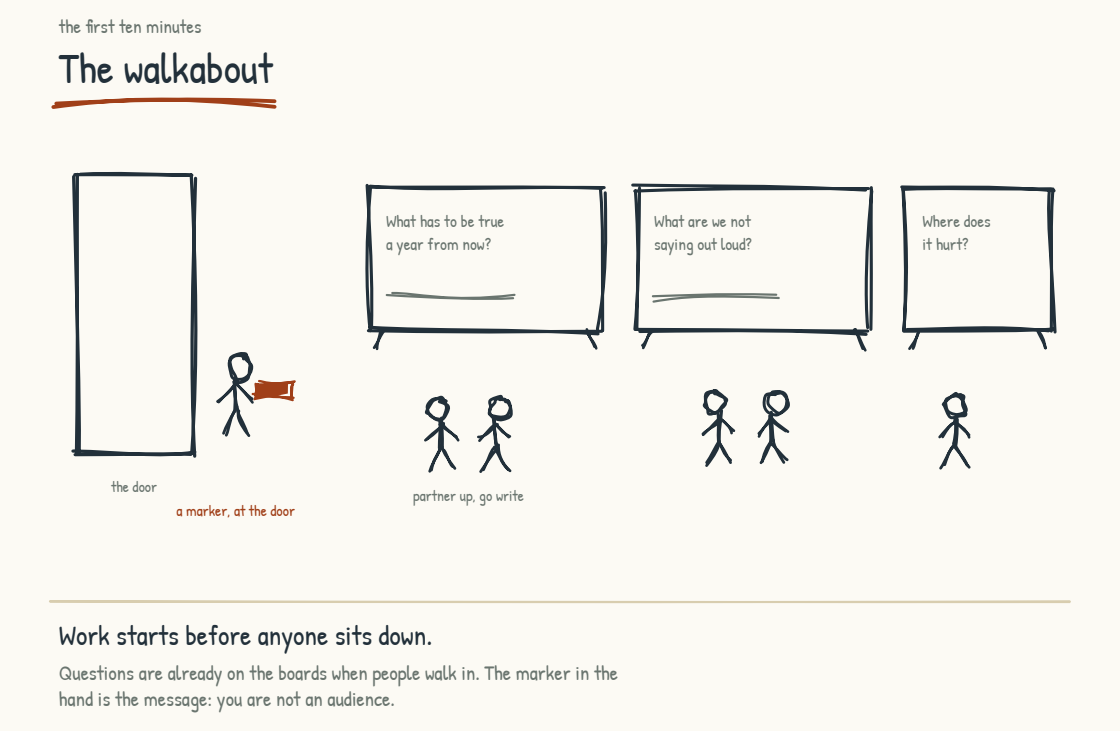

Figure 9.4e. The walkabout. Questions are posted before anyone arrives, and the marker handed at the door is the message: you are not an audience.

I open the day by talking about what it means to be a good participant in a session like this. Three things, and I say them plainly.

Single-tasking. Protect the session from avoidable calls, email, and divided attention. Also tell people explicitly that they may step out for health, caregiving, religious, accessibility, safety, or urgent work needs without explaining themselves publicly. Focus is a shared condition the design supports, not a test of loyalty or stamina.

Listening. High-performing executives are trained to always have an answer, so our brains run ahead of the conversation because we want the best one. I ask people to actually listen to their teammates instead of sitting in their own heads composing what they will say next.

Debate. We want people bringing different ideas together, testing evidence, and sometimes preserving disagreement. Synthesis can produce something none arrived with; forced convergence can erase the person who was right. Critique the idea without requiring everyone to merge it.

The last thing I ask for is quality over quantity. I want five really good ideas from a team, thought through, rather than a list of fifty.

When a room needs shared knowledge before it can work, we may run chatrooms instead of one lecture. Three or four short talks run in parallel, followed by questions, and groups rotate. Provide the material in accessible form before or after, caption and amplify as needed, avoid requiring standing, and use visual as well as musical cues. Parallel delivery does not guarantee everyone heard the same thing; a common evidence packet and plenary clarification keep the level-setting from fragmenting.

> *Five really good ideas, thought through. Not a list of fifty.*

Moving Work Between Rounds

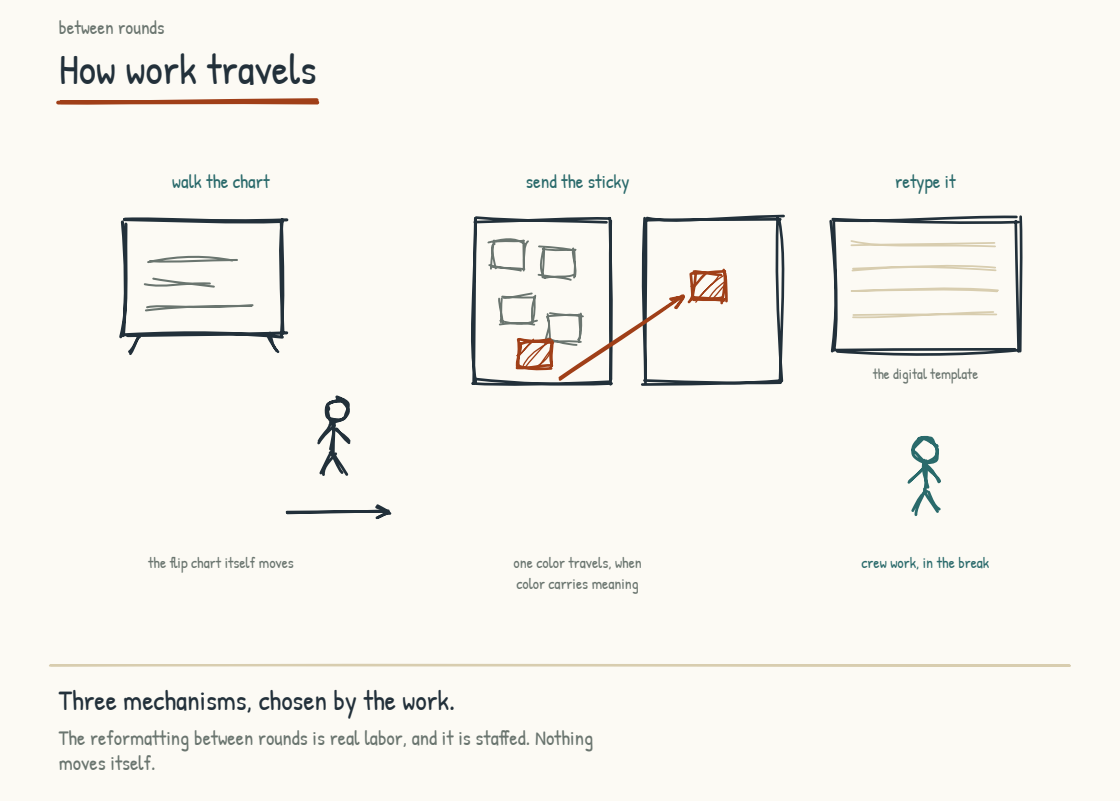

Figure 9.4f. How work travels between rounds: walk the chart, send the sticky, or retype it into the digital template. The reformatting is staffed, because nothing moves itself.

In a shift-and-share, sometimes the boards move and sometimes the people move. Sometimes one person stays behind while the rest of the team rotates. It depends on the space and on what we are trying to get. A common version is five-minute shifts, rotating three or four times, and then coming back home to where you started.

For carrying work forward, we use one of three mechanisms. Big flip-chart paper, where one round's output is written up and handed to the next team. Colored sticky notes, where the blue ones travel to the next round. Or digital templates in Google Slides or Docs, where a team types their output and our people reformat it into the shape the next round needs.

That reformatting between rounds is real work, and it is why the facilitation team matters. With forty or a hundred people you are producing a lot of output, and somebody has to manage it, template it, and make it reusable for the round that follows.

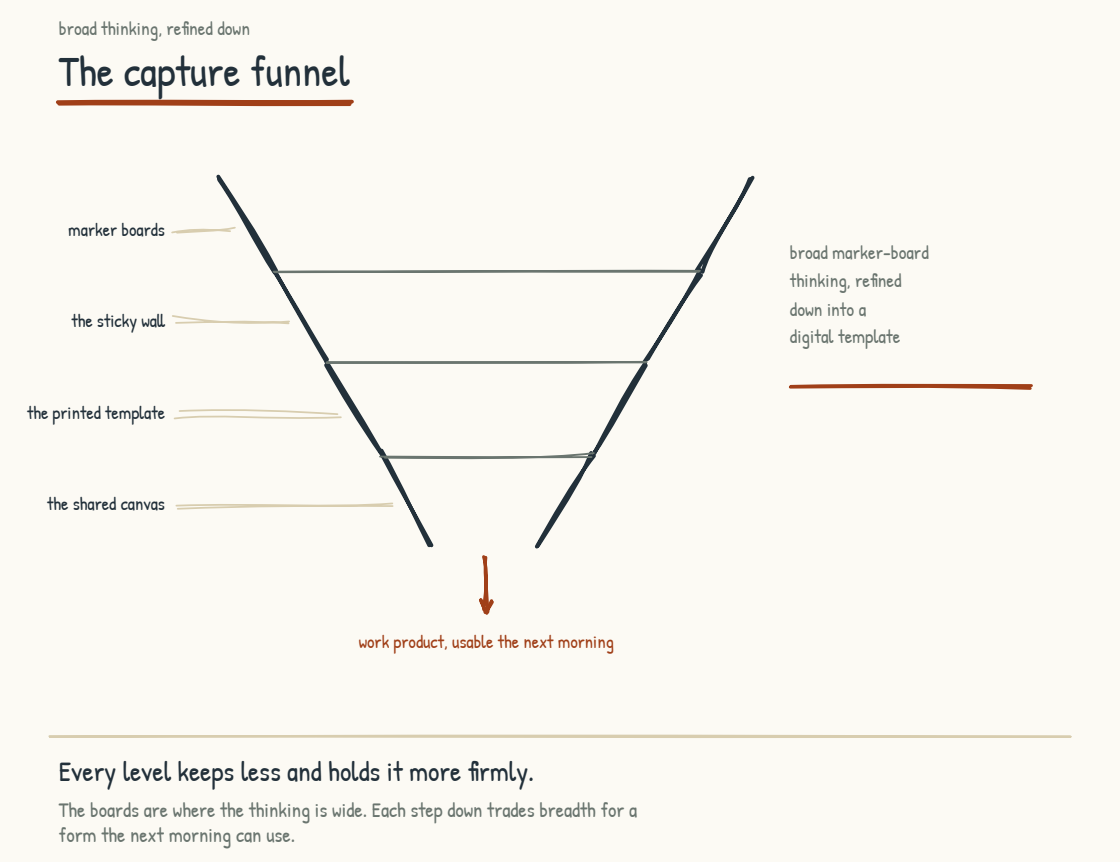

Figure 9.4g. The capture funnel. Broad marker-board thinking, refined down level by level into work product that is usable the next morning.

When Something Goes Sideways

If the room goes quiet, I do not assume disengagement. The question may be unclear, the stakes may feel unsafe, people may need processing time, or silence may be useful. I can restate the purpose, offer a minute to write, invite pairs, use an anonymous channel, or ask for volunteers. Direct invitation can work when the participant has agreed to that norm; surprise cold-calling can turn reflection into exposure.

If one person dominates, I reinforce the norm and change the structure: timed turns, writing first, a round-robin with a pass option, or a smaller group. A quiet conversation may be appropriate, led by the facilitator rather than using the sponsor's authority unless conduct or safety requires it. The goal is equitable airtime, not public correction by proxy.

If a group finishes early, I check its evidence, detail, dissent, and whether the assignment was too easy or unclear. It may genuinely be done. I can offer an extension question or invite it to

test the output from another perspective. Comparing it vaguely with “deeper” teams uses social pressure without giving useful feedback.

If a group rejects the premise, that happens, and it is fine. I walk the room making sure everyone understands the assignment. When a team pushes back, I give them the overarching goal and tell them that if they want to get there a different way, go for it. Letting them own the work is usually better than winning the argument.

What I watch when deciding whether to cut a round short or let it run is the evidence, detail, participation, and usefulness of the output for the next step. If needed, I enter a breakout, explain the gap, and ask a sharper question. A sponsor may observe only if the role was disclosed and will not suppress candor. Observation should assess whether the structure enables participation, not grade individuals, diagnose whether an expert is “useful,” or ask whether teams should be pushed harder. Feedback belongs to the process unless conduct requires a separate response.

A person who appears not to engage may be choosing not to, but I do not know why from behavior alone. They may be processing, excluded by the format, unconvinced of safety, dealing with a disability or private event, or signaling that the work is not legitimate. I can offer a private, nonintrusive check-in and an alternative channel. Participation may grow with the group, but conformity is not the measure of a successful room.

When an unanticipated tension surfaces, I assess consequence, consent, power, confidentiality, time, and whether the full group is the right place. Some tensions should be named in plenary. Others need a break, private support, formal grievance route, specialist help, or follow-up with people who were harmed. “Let us discuss the elephant” is not permission to expose someone. The session should make difficult issues addressable without pretending every issue can or should be resolved immediately in public.

Closing the Day

I like to close with a synthesis activity in which people reflect on what they produced. A commitment round belongs only after the group has earned it and only for actions within each person's authority. Passing must be allowed. People should not make public promises under social pressure that their manager later treats as consent or performance evidence.

It can be moving to hear people name what they are ready to do. Emotion is weak proof that a commitment is real, and tears make a poor design goal. Face-to-face contact may add weight for

some people; remote commitments can also be sincere. What makes a commitment credible is specificity, choice, authority, resources, follow-through, and the ability to revise it when new evidence appears.

At the end I review where we got to, thank everyone, and invite reflection on what was useful, difficult, incomplete, or missing. The sponsor often closes by naming what happens next. The final word should not overwrite dissent or turn the day into endorsement; in some sessions a participant, worker representative, customer, or neutral facilitator is the more legitimate closing voice.

On a two-day session, we sometimes end Day One with an important question still open so people return ready to continue. Do not manufacture anxiety or leave a personal conflict raw for dramatic effect. Before people leave, name the open questions, explain how the next day will address them, and check whether anyone needs support.

Some participants have called a session the most productive of their career. I value that response while treating it as satisfaction evidence, vulnerable to selection, courtesy, recency, and the energy of the day. Photographs belong only to people who consented to them. The stronger test arrives later: what decision improved, what commitment held, what changed for people outside the room, and what the process cost.

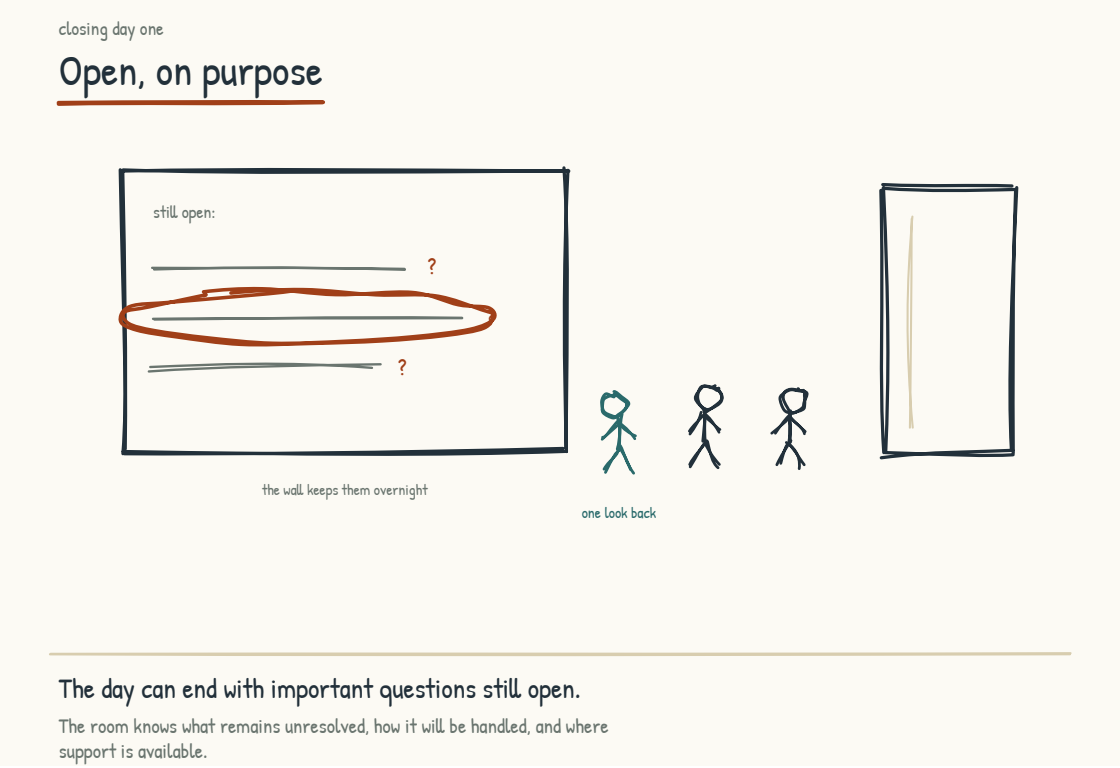

Figure 9.4h. Closing Day One with a question open but the emotional container closed. The room knows what remains unresolved, how it will be handled, and where support is available.

The Shape of a Day

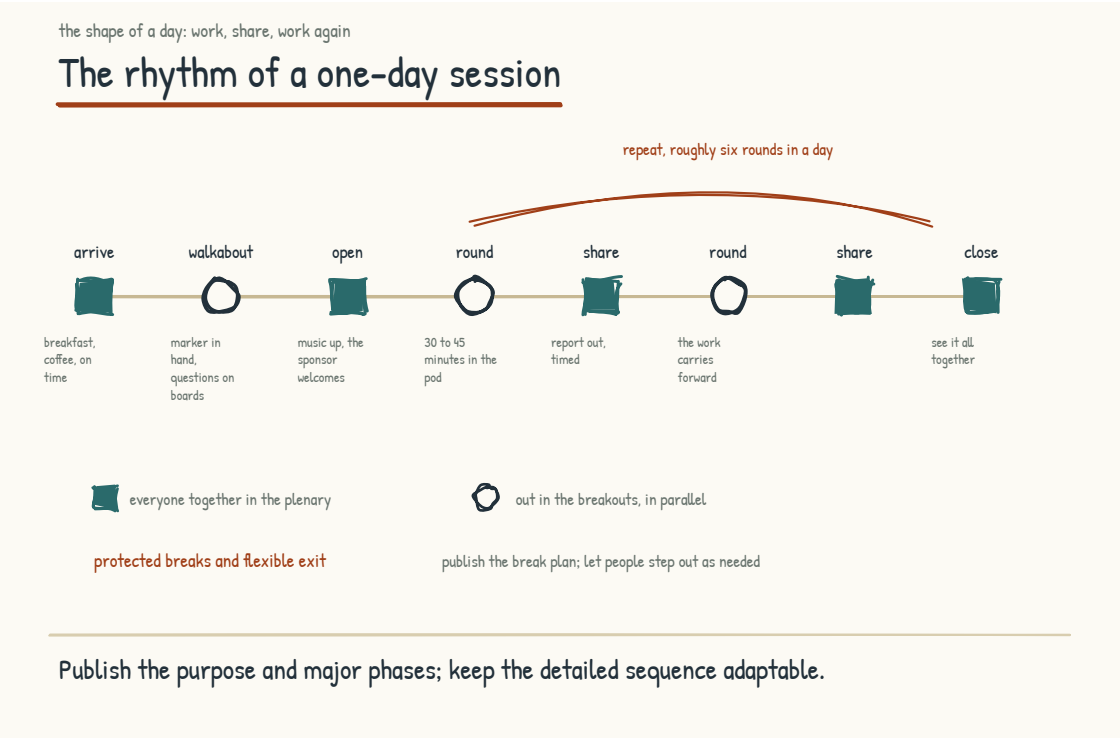

Figure 9.4i. The rhythm of a one-day session. Participants know the broad arc and breaks; exercises arrive progressively, and visual as well as optional audio cues mark transitions.

Sessions often start with breakfast and coffee. Partly hospitality, partly logistics. Food can support timely arrival, but travel, disability, caregiving, and transit disruption still happen. Publish a firm start and a dignified late-entry path so one delay does not derail the room or shame the person arriving.

When someone walks in, the default is to put away avoidable distractions while retaining devices and supports people need. Then they receive an opening assignment: pair up or work individually, get settled, and explore the questions around the space or in the digital equivalent. Depending on group size and access needs, that runs fifteen minutes to half an hour.

Once we have a quorum, we turn up the music. That is the signal to come sit down, and we use music the same way all day, as the cue that something is changing. I welcome everyone, set up my own role briefly, and hand over to the sponsor to welcome the room in their own words. Half an hour at most.

Then I come back and set up the first round. We describe the assignment, often put it on a screen so it stays visible, show how the teams are assigned, and send people out to their breakouts.

From there the day is a rhythm. Teams work. I am walking the room, stopping at every group to make sure they understand the assignment, watching for thirty or forty minutes while the work develops. I give a warning about five minutes out so they can wrap up and finish documenting. They bring their work to the front. Most rounds, most teams report out, and I time the report-outs to keep us honest. Then straight back out to the next round. Work, share, work again.

Through all of it I check with the sponsor and design team against the stated objectives, not simply whether they are happy with the content. Participants know the scheduled breaks and may also step out as needed. The room pauses for lunch. Momentum matters, and humane participation is part of the momentum rather than its enemy.

I build the closing synthesis as the day develops, using the work people have put on the walls.

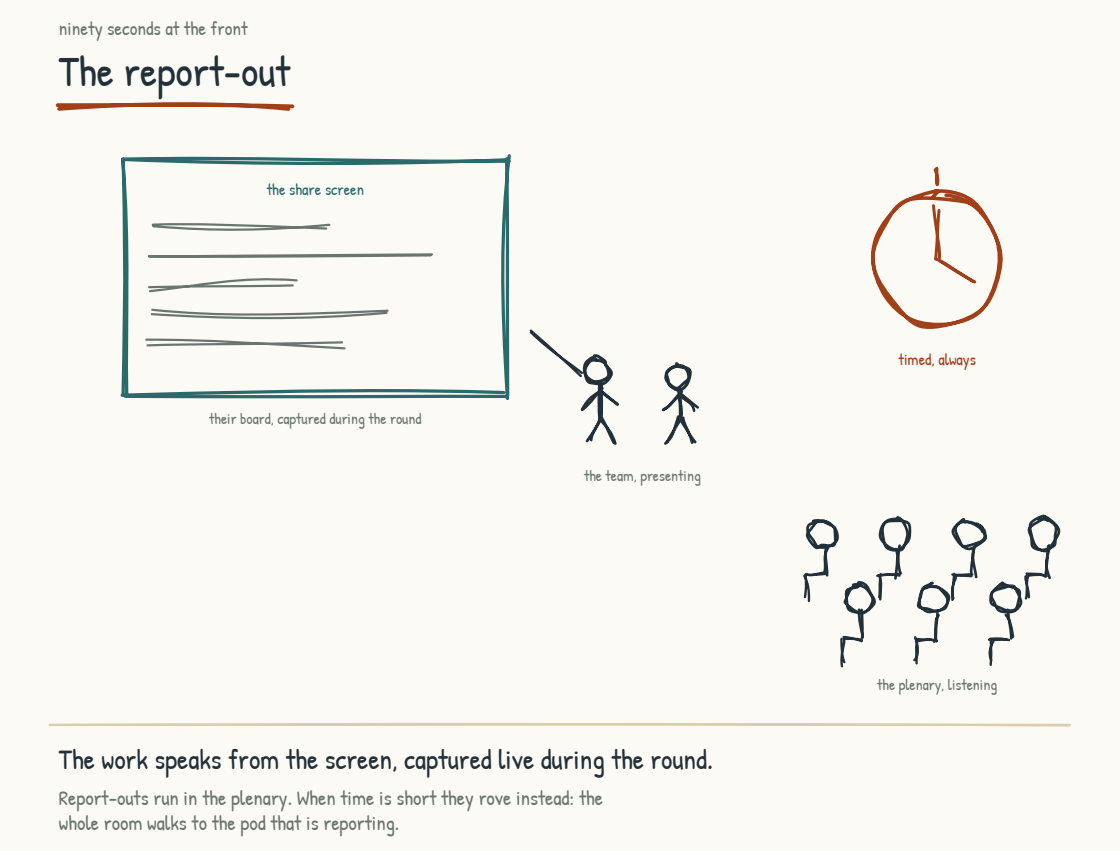

Figure 9.4j. The report-out, timed and in the plenary, with the captured board on the share screen. When time is short, the reports rove instead: the room walks to the pod.

The Scene Change

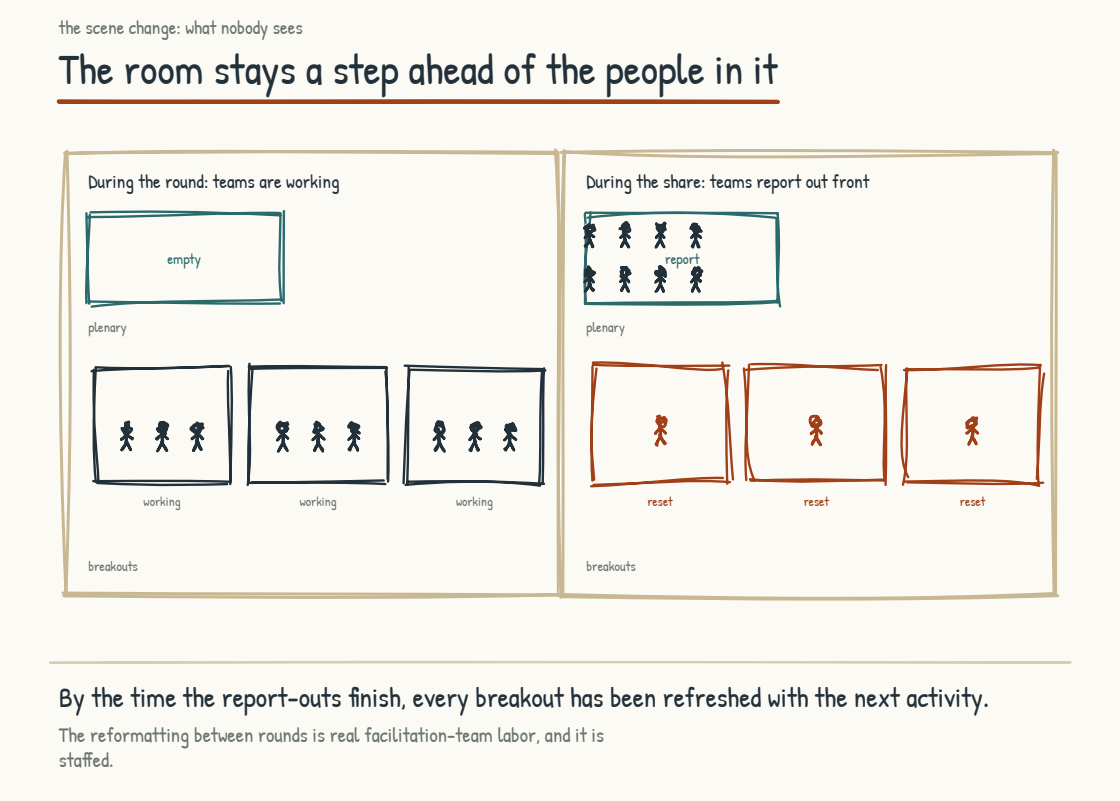

Figure 9.4k. The scene change. While teams report out at the front, the facilitation crew resets every breakout for the round that follows.

The part nobody sees is what the facilitation team does while the groups are working.

During a breakout they are preparing the next one. Organizing sticky notes. Getting flip charts ready. Then, when teams come to the front of the room to report out, the team goes into the empty breakout areas and resets them. I call it a scene change. By the time the report-outs finish, every breakout has been refreshed with the next activity, and people walk back into a space already staged for what comes next.

So while I am up front running the share-out, I am also watching my team in the background rebuilding the room. That orchestration is most of why the day feels all of one piece to the people inside it, and it is invisible by design.

> *By the time the report-outs finish, every breakout has been reset for the next activity. The room stays a step ahead of the people in it.*

Changing the Design Live

I change things when the evidence in the room warrants it. Explain changes that affect the purpose, timing, or use of the work.

If I am moving between groups and I see a team struggling, I will hand them another question or two to test their work. If I see a team lit up about one particular area, I will tell them to stay there and go deeper rather than covering my original assignment end to end. Sometimes I wait until after a report-out, change the assignment entirely, and send them back with something different. I might just write it up at the front of the room.

I have run rounds where the exact assignment is held until groups arrive, allowing the design to respond to what the room needs. Participants should still know the purpose, time, decision boundary, and how their output will be used. Openness in design is not opacity about participation.

When the Room Turns

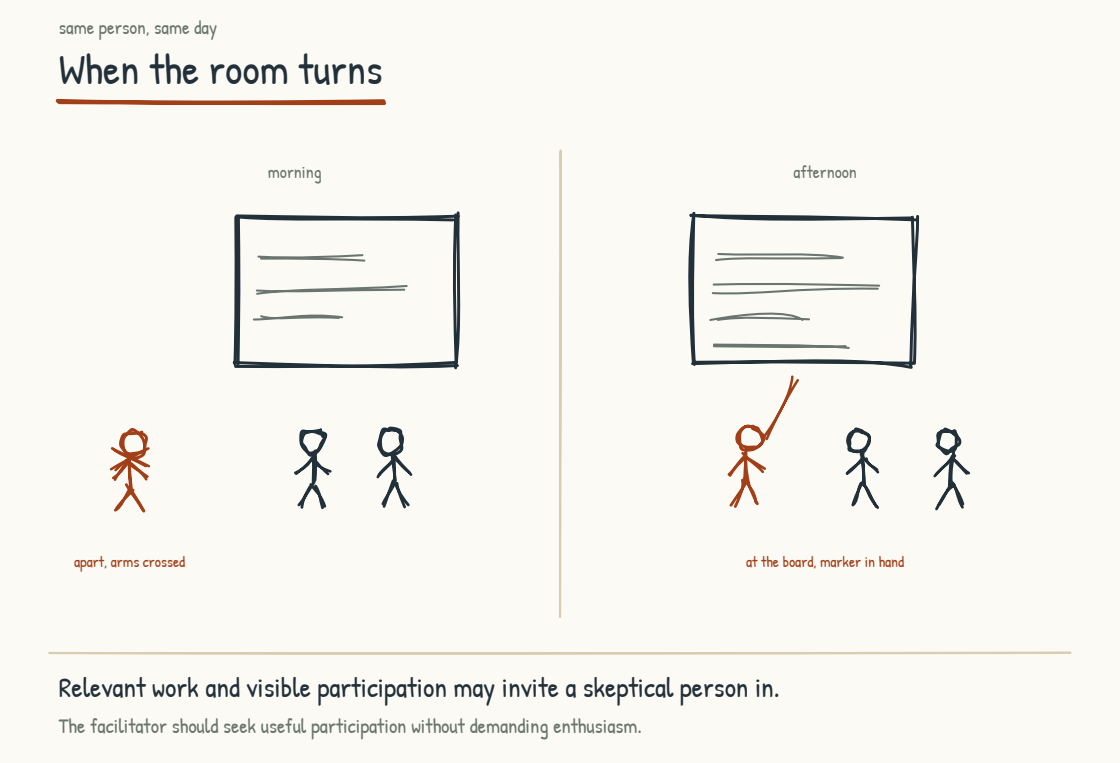

Figure 9.4l. The same person, morning and afternoon. Good design can invite engagement; continued skepticism may still carry information the room needs.

Almost every session has a few people who arrive in a bad mood, or who have never once been in a workshop worth their time and are expecting this to be the same. They have good reason to expect another wasted day.

The energy of the room can invite skeptical participants into the work. It can also pressure them to perform agreement. Some will engage later; some will remain unconvinced and may have good reason. My job is to make contribution possible and take dissent seriously, not turn every person.

By the end, I want people to understand the process and judge it worthy of trust based on what it did. The question many participants carry is: are we doing the right thing right now, and is this a good use of our time? I answer with the purpose and the connection to what comes next, not merely reassurance that they are on my path. Participants should be able to challenge the assignment and receive a substantive answer.

Figure 9.4m. The question every participant carries all day, and the facilitator answers out loud.

Afterward

In the forty-eight hours after the room closes, our team aims to inventory and digitize agreed outputs and hand the authorized set to the client team. Before photographing boards, we separate shared work from personal, confidential, privileged, or sensitive material. The delivery preserves provenance, dissent, and unresolved status rather than polishing every note into agreement. Sponsor communication thanks participants, explains next steps, and gives each audience access only to what it is entitled to see.

What the harvest looks like depends on the session. We have delivered videos, e-books, and magazines. It depends on the purpose of the work.

Three months later, my facilitation team may not be involved unless follow-through was explicitly contracted. That makes client ownership, internal capability, and named responsibility essential. The session should establish review cadence, owners, decision rights, evidence, and escalation before people leave. It should also leave a route for participants to correct the record or raise a concern after the social pressure of the room has passed.

How do I know whether it worked? Satisfaction, readiness, and sponsor judgment matter, but they are not enough. I want participant feedback that can be candid, evidence of inclusion across groups, the quality and traceability of decisions, follow-through on commitments, and later

operating outcomes appropriate to the session's purpose. Excitement is an immediate signal. Changed decisions and consequences are the test.

That is the craft, as plainly as I can set it down. It is the half of this work that does not fit in a framework, and it is the half that decides whether a room produces something real. Stage 10 closes the field guide with an account of who wrote this and why, and a statement of what we believe.

CONSIDER BEFORE YOU ACT

Look at the last working session your team ran. How many of these details were decided deliberately, and how many were inherited from how meetings usually go? The agenda, the breaks, the room, the group sizes, the way work moved between rounds. Most of the difference between a session that produces something and one that produces a nice conversation lives in those choices.

CONSIDER BEFORE YOU ACT

If you keep three things from Stage 9: 1) Strategy, operating design, workforce participation, governance, and architecture need explicit interfaces and shared evidence, whether one team provides them or several do. 2) A leadership team needs enough working alignment to decide coherently, while preserving dissent and commissioning analysis where evidence is missing. 3) The Alignment Sprint is one bounded on-ramp, built to distinguish decisions, commitments, hypotheses, and questions that require wider participation.

STAGE 10

The Voice

A personal account, and a statement of what the partnership believes.

Section 10.1

A Note on Who Is Writing This

Who has the standing to write about AI-Native?

A reader who has gotten this far is allowed to ask the obvious question. Why this person. What does he know.

I want to answer it honestly, because the honest answer is more interesting than the credentialed one and probably more relevant to whether you'd want me in a room with your team.

I think associatively. I have known this most of my life, and for much of it I did not have good language for the difference. I move between broad framings and narrow details and make connections across domains that do not usually meet. A conversation about a factory can call up something I learned in hospitality; a question about an AI architecture can expose a problem in how a leadership team makes decisions. Sometimes I see the connection before I can explain it; sometimes the room sees what I missed. I have spent years learning to slow the leap down, expose the steps, and let other people test whether the connection holds. That matters because intuition can produce insight and nonsense with the same feeling of arrival. People ask how I thought of something. The honest answer is often that the connection arrived first and the explanation came afterward. That makes the explanation, evidence, and challenge more important, not less.

I am telling you this because the work in this field guide is the product of a particular kind of mind, and you should know what kind of mind it was so you can calibrate what to do with what you have read.

The mind would not be enough on its own. What gave it traction is what I have spent my career doing. For more than two decades I have designed and led collaborative experiences for senior teams working on complex, high-stakes problems. The work has spanned industries that don't usually share rooms with each other: financial services, hospitality, defense, healthcare, retail, manufacturing, professional services, and technology. The work has spanned topics that don't usually get worked on by the same facilitator: strategy, innovation, operating model redesign, HR transformation, culture change, technology rollouts, customer experience design, and large-scale change management.

Twenty years of that variety is an epistemic position, with strengths and limits. Repetition in one domain builds depth; work across domains builds comparative pattern recognition. I have watched rooms avoid the subject they were assembled to discuss and then turn when someone finally named it. I have watched an artifact carry a decision for years, and another become wallpaper before the week ended. I have learned to ask what produces a real conversation, how power appears in collaboration, when disagreement generates insight or deadlock, and which artifacts remain useful. What might be shared across a Navy command, a hotel loyalty program, a software company's culture, and an insurance firm's strategy? Cross-domain work can reveal patterns a specialist may not encounter. Specialists can also see causal detail, history, law, and risk that a pattern-seeker misses. The work is strongest when breadth and depth correct each other, and when the facilitator knows which kind of knowledge is absent from the room.

The work has also given me a long view on what humans can do together under different conditions and what no framework guarantees. The most important principle I carry is simple: people are more likely to understand, shape, and support work they legitimately help create; work done to them without voice or recourse creates a different relationship. Participation does not confer ownership automatically, erase disagreement, or substitute for rights and authority. The principle is why I am skeptical of consulting that promises to deliver transformation as a package. Co-creation is not a ceremony added to delivery. It changes who has standing in the design.

The third thing I bring is more recent: I have been building a company of my own, from an idea to a working product, living inside the AI-Native question rather than studying it only from a

distance. Founder experience is not a credential that outranks enterprise experience. It is exposure to a different set of constraints: limited time, limited capital, incomplete roles, direct consequences, and no clean separation between the person proposing the operating model and the person who has to work inside it tomorrow.

That experience has changed how I think about everything I had written about this topic before. It has shown me, in my own work and in my own team, what it actually feels like when AI is genuinely woven through how you think, build, and decide. It has also shown me what goes wrong. The moment we noticed we were defaulting to Claude Code instead of talking to each other and had to correct it deliberately. The moments where the gain was real and the moments where the gain was illusory. The places where AI compressed timelines and the places where it produced confident output that needed to be torn up and started over. The texture of doing this kind of work rather than the theory of it.

What I have learned at this scale does not transfer cleanly to a company of a thousand people or ten thousand. Even the principles must be retested under different power, regulation, legacy systems, geography, labor arrangements, and consequences. Lived experience gives me something observation alone does not: a felt sense of the work, including its seductions and failure modes. I know the pleasure of compressing a week into an afternoon and the danger of mistaking speed for correctness. I know how quickly an AI collaborator can become the default participant in a conversation and how deliberately a human team has to reclaim the moments that should remain between people. It makes me less satisfied with maturity scores without behavior, operating-model diagrams without lived work, and answers that arrive before the room has learned to ask the consequential questions.

I am sharing all of this because the field guide you have just read is shaped by all three of these things: the way I think, the breadth of what I have facilitated, and the lived experience of building inside the question. If you have found yourself in disagreement with parts of the field guide, you should know that those disagreements are with a particular mind that has come to its positions in a particular way, and you are entirely entitled to think differently. If you have found yourself agreeing with parts of the field guide, particularly the parts that argue this conversation is more about human life than about technology, you should know that the agreement is with a position someone has come to, the long way, after a lot of years and a lot of rooms.

The offer, if you are considering bringing me into your organization, is that I would help your leadership team do this work in your context, with the people affected and on the questions that

matter. The engagement needs facilitation and operating-design capability alongside credible technical architecture. I have spent more than two decades practicing the first and recent years building inside the second; Nodalix and other qualified technical specialists bring engineering depth that I do not claim as my own. We would build and test frameworks with your team rather than deliver them as doctrine. The valuable part is what your organization accepts, rejects, corrects, and creates that would never have occurred to me. That is the part I look forward to most.

Section 10.2

What We Believe

What is worth committing to as this transition unfolds?

Everything to this point speaks in my voice because the account is mine. This last section speaks as we, and the we is specific: the human partnership behind this field guide and the people who build and facilitate alongside me. AI systems helped us draft, test, retrieve, compare, and revise, as the preceding stages describe. They are tools and collaborators in the practical sense, not authors who can hold beliefs or accountability. The people named by the we own what follows.

We have spent the length of this field guide developing the argument. What follows is what we believe, distilled to the essentials. These are commitments and working hypotheses, not findings made true by appearing in a manifesto. The full arguments and their limits are in the stages behind you.

We believe that AI-Native is finally a question about human life. Operating models, architectures, practices, and technologies express a more foundational choice: what kind of place the company will be for the people who spend a significant part of their lives there and for the people affected by its products and decisions. A leadership team that excludes that question is solving an incomplete problem, however rigorous the optimization.

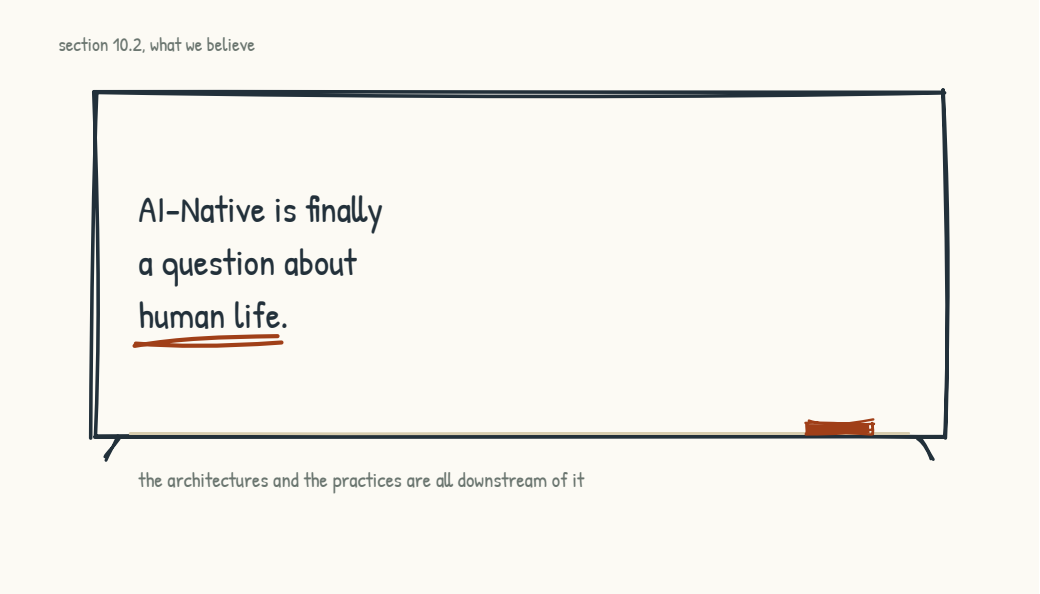

Figure 10.2a. The first belief. The rest are downstream of it.

We believe that framing humans as the residual around automation is wrong. Human value is not secured by a permanent list of tasks machines cannot perform. It rests in standing, responsibility, relationship, participation, embodiment, lived stakes, and the authority to shape consequential work. A company that protects and develops those forms of value may build stronger performance and legitimacy; it should measure both rather than treat moral commitment as a guaranteed competitive return.

We believe that institutional knowledge can become strategic advantage in the AI-Native era. It lives across people, relationships, records, documents, graphs, databases, rules, tools, and practice. No single representation captures it, and structured knowledge does not reason by itself. Fair, willing participation improves what can be preserved, but willingness also depends on workload, consent, credit, compensation, privacy, professional duty, and whether leadership's conduct makes the covenant credible. The rest must be carried through apprenticeship, redundancy, succession, and continued human judgment.

We believe that the human reckoning is foundational. Grief over expertise, threats to status, fear of obsolescence, survivor guilt, shame at starting over, anger at broken promises, and complicated relief may all be present, alongside curiosity, indifference, and hope. Leaders should not assign emotions to people or demand disclosure. They should create honest channels, change the conditions producing avoidable harm, and accept that naming emotion is neither therapy nor a guarantee of transition success.

We believe that the distinction between using AI for thinking and allowing it to substitute for thinking a person or institution still needs is consequential. The right posture depends on the task, stakes, and capability the organization must retain. Offloading can be rational; unexamined dependence can weaken learning and verification. Companies should decide deliberately, test capability over time, and protect practice where judgment remains part of the job.

We believe that up-leveling is possible and uneven. Conditions such as supported challenge, feedback, time, access, mentorship, and meaningful work can be designed, but they do not produce identical results for every person. Demanding productivity without development may meet an immediate need while spending down capability. The workforce should help define what development matters and receive fair opportunity to pursue it.

We believe that the welder's standard is a question every AI-Native leadership team should be willing to hear: if a worker met you years from now, what might they say about the work they did here? Leaders cannot answer for workers, and confidence is not evidence. Ask people directly through safe channels, examine differences across groups, and treat hesitation as a prompt for inquiry rather than a diagnosis of hidden wrongdoing.

We believe that architecture shapes whether AI capability compounds or fragments. The nine-layer stack is an analytical model, not a mandatory product sequence. Chat can be valuable and governable in a bounded use case; it becomes weak strategy when the company mistakes individual access for shared capability. Graphs fit relationship-heavy knowledge, collaboration surfaces fit shared work, and simpler stores or tools may fit other needs. The right default is deliberate, interoperable, permissioned architecture, with shared multiplayer surfaces where the work calls for them.

We believe that epistemic hygiene is a strategic discipline. Corroboration, monitoring, verification, provenance, evaluation, incident response, and the freedom to challenge an output improve the trustworthiness of AI-mediated work. They do not make it perfectly trustworthy at scale. Assurance remains bounded by the evidence, controls, people, and systems doing it.

We believe that practices help produce sustained change. Values without corresponding behavior, incentives, and accountability become slogans. Repeated practices can create conditions for learning, dignity, and human flourishing, but they can also become compulsory

ritual. Teams should examine who benefits, who is burdened, and whether the practice still serves its value.

We believe that moments that matter are a useful level of analysis for meaningful human involvement. Consequence, rights, context, responsibility, preference, and relationship help identify where human participation deserves particular investment. Some moments can be automated responsibly; others require legitimate human authority. Design, protect, and review them with the people affected rather than assuming human presence alone creates a memorable experience.

We believe that an AI-Native transition is both a portfolio of projects and a changing system. Feedback structures can reinforce or balance change, and "defeating" names our managerial category for harmful combinations. Loop maps are hypotheses to test. Mental models can be powerful places to intervene, but structures, rules, information flows, power, and immediate events matter too, and causality runs in both directions. Skilled facilitation can help a team examine its assumptions; no process guarantees change will hold for eighteen to thirty-six months.

A leadership team that has read this book has arguments, frames, cautions, and references to which it can return. It has questions to test with its own workforce, customers, systems, and evidence. It has a possible beginning. If a bounded leadership engagement fits the need, the Alignment Sprint in Section 9.2 is one concrete move. It is not the only one.

The work itself is still ahead, and this field guide is preparation, not a command to act before the evidence warrants it. A reader may move, pause, investigate, reject a recommendation, or decide that a proposed AI use should not proceed. Bringing one leadership conversation to a more honest and consequential level, or stopping one badly designed intervention, would already produce value the field guide was written to enable.

We are available to support you if you would like the support. We would be honored to do this work with you if you choose us, and we would be glad to see you do this work well if you choose someone else, or if you do it on your own. The work matters more than who supports it. The work itself is what we advocate.

This is the end of this field guide. The beginning of the work is yours.

CONSIDER BEFORE YOU ACT

If you keep three things from Stage 10: 1) Legitimate participation changes people's relationship to the work, but it does not replace rights, authority, consent, or accountability. 2) These beliefs are a distillation; the arguments, evidence, and limits live in the stages behind you. Return to the one your situation requires. 3) The field guide prepares; the decisions and consequences are yours. Improve one consequential conversation or stop one badly designed intervention and it has done useful work.

Two Ways to Read This

How to use the field guide in book mode and in graph mode, and what each does best.

A field guide of this length, on a topic moving as fast as AI, has a structural problem. Readers under the greatest pressure may have the least time to read it in full. A leadership team deep in a transition may not have a weekend for the complete argument, while a leader deciding whether to engage may not yet know which sections matter most.

The conventional response to this problem is a summary. A short version of the long version, written for people who will not read the long version. The summary helps. It also loses the work the long version is doing. The frameworks lose their grounding. The arguments lose their texture. The cross-references that hold the whole thing together as a coherent piece of work are flattened into bullet points.

We have built a different response. The full field guide is the long version, where sequence, voice, repetition, and accumulated context do work a summary cannot preserve. Alongside it, we have built a structured playbook graph and section corpus derived from the same manuscript so an agent can retrieve concepts, relationships, exercises, and source passages. The structure helps software traverse the material; the model and application perform the retrieval and inference. This gives a time-constrained reader an entry point without pretending the entry point contains the entire argument. The companion should remain synchronized with the published edition, and its answers should cite the book rather than quietly become a second authority.

The Book

Read linearly, the field guide is an argument. It builds from premise to architecture to people to practice to method, and the order matters. A reader encounters claims, counterweights, stories, and later corrections in relation to one another. The frameworks land differently because the reader has been prepared for them, and the human sections carry the structural work that came before. Linear reading also makes the author's rhetoric and assumptions easier to evaluate as a whole. It is slower and less targeted. For questions where the argument, not merely the answer, creates understanding, that cost is part of the value.

The Graph

Queried, the field guide becomes a reference. A reader can ask the companion to surface relevant passages, guide the feedback-loop diagnostic from Section 7.1, support the iceberg inquiry from Section 7.2, or retrieve the workforce-design exercise from Section 5.9. It can expose cross-section connections and help translate a framework into questions for a specific decision. That speed creates its own distortion: retrieval favors what matches the query, while the linear book can challenge the question itself or introduce a counterargument the reader did not know to request. The output remains model-generated. Open the cited passage, distinguish the book's claims from the agent's inference, inspect what may have been omitted, and do not mistake a fluent application for evidence about the reader's organization.

What the Agent Can Do Well

Locate the parts of the field guide that match your situation. You describe what you are working on, the agent surfaces the sections, callouts, and frameworks that speak to it.

Guide a diagnostic. The agent can ask the operating-system, feedback-loop, iceberg, or workforce-design questions and organize the answers. It cannot validate those answers without relevant evidence, nor should its summary erase disagreement.

Connect concepts across sections. The agent can retrieve links for a specific topic without requiring a complete linear read, while citing the passages from which the connection is drawn.

Translate frameworks into decision questions. The agent can apply the book's reasoning provisionally and ask clarifying questions. Your context, current law, evidence, rights, and accountable decision-makers still govern the choice.

What the Agent Cannot Do

The agent does not replace the work the field guide asks people and institutions to own: workforce participation, the covenant, meaningful human involvement, practices, governance, and operating design. It may help mediate or prepare a conversation. It cannot confer consent, resolve interests, hold legal or managerial accountability, or experience the consequences. The conversation and decision remain human and institutional work.

The agent also does not replace the linear read for the topics where the argument is the value. The cognitive posture argument in Section 5.4 (AI for Thinking) is an example. The argument has texture that the bulleted summary loses. The reader who comes to the agent for a quick

summary of that section will receive a competent summary and miss the texture. We recommend reading those sections linearly.

How to Access It

The graph is a companion to the field guide, available the same way the field guide is. How you use it depends on your situation.

If you have an AI assistant that supports uploaded context, the companion files may be attached or indexed subject to that product's file limits, retrieval behavior, data terms, retention, and security controls. Different assistants may truncate, chunk, or rank the material differently, so test whether answers cite the correct section and say when evidence is absent. Instructions and the graph file are at the access page linked from the back cover. Do not upload confidential company material merely because the book files themselves are intended for access, and do not assume a consumer account inherits your organization's permissions.

If you prefer the companion interface we provide, it is available at the same access page. Review its current privacy notice, retention, limitations, and terms before entering company or personal information; the interface and underlying models may change after this edition is printed.

If you are reading this in print and would rather not switch to a screen, the QR code on the back cover takes you to the access page.

The Point of Building It Both Ways

A field guide arguing that important knowledge should be usable by people and software should demonstrate the principle. The book is the authored, human-readable source. The graph and section corpus are derived, machine-usable representations. They are not equal authorities: when they diverge, the versioned manuscript governs and the derived artifacts must be rebuilt.

A leadership team has a similar design problem. Institutional knowledge needs human-readable sources of record and, where useful, structured representations that software can retrieve or traverse under inherited permissions. The derived representation should link back to its source, preserve dates and owners, expose uncertainty, and be rebuilt or corrected when the source changes. Personal notes, privileged advice, regulated data, and contested claims do not become shared knowledge merely because an agent can index them. These forms, plus owners, maintenance, provenance, and human practice, can support an operating substrate. The book and its companion are a modest worked example, not proof that every company's knowledge belongs in a graph.

Quick Reference

For return readers and working sessions. Two lists. The first helps you find a framework you remember encountering somewhere in the field guide. The second gathers all forty section questions as discussion prompts for leadership sessions. Pick one. Bring it to your team. Spend an hour on the answer.

Frameworks by Section

The major frameworks the field guide introduces, in book order. Use this list when you remember encountering a framework and want to find it again.

The Forty Inquiry Questions

The questions that head each section, gathered here as discussion prompts for leadership sessions. Pick one. Bring it to your team. Spend an hour on the answer.

§3.1 What changes when AI becomes a dimension of the operating model?

§3.2 What does the work of a company look like when AI participates in it?

§3.3 How does knowledge work change when AI is in the room?

§4.1 How does AI come to understand how a company actually works?

§4.2 How do you redesign the way work gets done when AI is now part of it?

§4.3 What happens to AI work when it stays inside private chats?

§4.4 How does an organization tell the difference between AI output that is true and AI output that sounds true?

§4.5 How do you fund a transition whose most valuable investments produce nothing to demo?

§5.1 What does the AI-Native transition look like at the leader's own desk?

§5.2 What does an AI transition feel like from inside the workforce?

§5.3 What can humans do that AI cannot?

§5.4 What is the difference between using AI to think and letting AI do the thinking?

§5.5 What becomes possible when a worker has AI alongside them?

§5.6 What makes someone care about their work, and what does their brain require of it?

§5.7 Who becomes skilled in a domain when AI lowers the barrier to entry?

§5.8 What habits of mind no longer serve the AI-Native worker?

§5.9 Who works here now, and how do we design that?

§6.1 What does a team do to keep the human work alive day to day?

§6.2 Where in the work does human presence matter most?

§6.3 What does passion have to do with AI?

§6.4 How does a new toolbox change the type of output a team generates, the way they communicate, and how they share knowledge?

§7.1 Why do AI transformations keep producing the same problems?

§7.2 What actually makes change last?

§7.3 What can the numbers tell you about whether your transition is real, and what can they not tell you?

§8.1 What does AI-Native work look like on an ordinary day?

§8.2 What does it look like when an AI-Native transition goes wrong?

End of First Edition

Becoming AI-Native | A Field Guide for Leadership Teams

www.arc5design.com | www.nodalix.ai