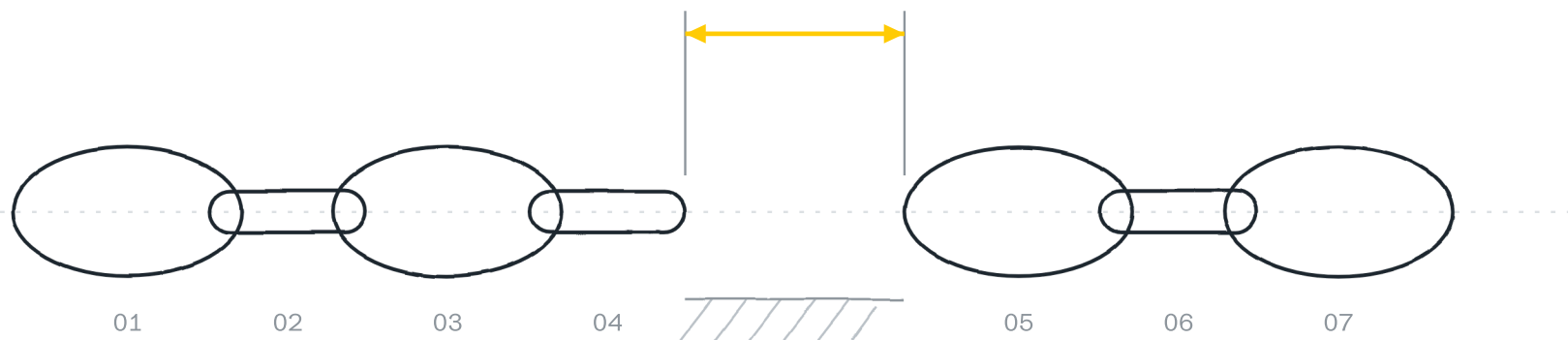


The memory gap

Modern AI is transformative and real, yet most businesses are seeing little to no real benefit. The failure begins before any tool is involved, in the condition of what the company knows.



How to read this report

Chapters one and two describe a problem and give the evidence for it. Chapters three, four and five publish the lens we use to examine it, in enough detail that you can apply it to your own company. Chapters six and seven cover what the evidence says to do about it. The last page says who we are.

01	Most companies are seeing little real benefit from AI	03
02	The tools do not learn the business	06
03	Seven links decide whether knowledge reaches the point of use	09
04	The other half of the problem is how the workflow runs	13
05	The right target is level three	15
06	AI pays back first in the same few places	18
07	Most of the first work needs no outside help	21
08	About this report	25
Exhibit 01	Three studies measured the same failure	04
Exhibit 02	The four conditions of the memory gap	07
Exhibit 03	Knowledge moves through a company in a loop	09
Exhibit 04	Some answers exist in no single document	12
Exhibit 05	The four dimensions of how the workflow runs	14
Exhibit 06	The five levels, by the sentence you would hear	15
Exhibit 07	The value is between levels two and four	17
Exhibit 08	The budgets and the returns went to different places	19
Exhibit 09	The one-workflow worksheet	22
Exhibit 10	The four fixes, in order	23

Chapter 01

Most companies are seeing little real benefit from AI

Modern AI is transformative and real, yet most businesses are seeing little to no real benefit from it. This report is about why, and about what a company can do without waiting for the technology to change.

The tools themselves are good and continue to get better. They can write, answer questions, draft documents, check work, and automate routine tasks, and each generation does these things more reliably than the last. The transformative potential is undeniable, and so is the pressure most business leaders have been under to put AI to work in their business.

The deployment record does not match the capability. Over the past two years, research groups at MIT, BCG, S&P Global, McKinsey, Gartner, and KPMG have measured what actually happened when companies tried, with different samples and different methods, and their findings agree with each other.

MIT Project NANDA
The GenAI Divide: State of AI
in Business 2025, August
2025.

MIT's Project NANDA studied enterprise AI programs and reported in August 2025 that 95 percent of enterprise GenAI pilots produced no measurable impact on profit and loss. More than 80 percent of the organizations studied had piloted the tools, and around 40 percent reported some form of deployment, so the poor results cannot be explained by companies not trying. People inside these companies were getting real personal value from the tools. That value never accumulated into anything the companies could measure.

Three more surveys

BCG, Sept 2025, 1,250+ executives. S&P Global Market Intelligence, Mar 2025, 1,000+ respondents. McKinsey, The State of AI in 2026, Aug 2026, 1,719 participants, 97 countries.

The rest of the evidence says the same thing from different angles. BCG surveyed more than 1,250 executives in 2025 and found 60 percent of companies reporting little to no value from AI, with only 5 percent achieving value at scale, and that small group getting roughly double the revenue growth of everyone else. S&P Global Market Intelligence surveyed more than a thousand companies and found that 42 percent had scrapped most of their AI initiatives, up from 17 percent the year before, and that the average organization abandoned 46 percent of its AI proofs of concept before they reached production. McKinsey’s 2026 global survey, fielded in May and June across 1,719 participants in 97 countries, found nearly nine in ten organizations using AI regularly in at least one function, while only 37 percent attributed any effect on earnings to it, the same share as a year earlier, and about 6 percent qualified as high performers.

Exhibit 01

Share of the sample studied

Three studies measured the same failure

Each bar is a different measure from a different sample. All of them point the same way.



Sources: MIT Project NANDA, The GenAI Divide, August 2025 · BCG, The Widening AI Value Gap, September 2025 · S&P Global Market Intelligence survey, March 2025. Figures as published; measures differ by study and are named per row.

Gartner
Press release, February 2025.

KPMG
Trust, attitudes and use of AI: a global study, 2025. 48,000+ people, 47 countries.

McKinsey
The State of AI in 2026, August 2026.

Two further findings point at where the failure is. Gartner predicted in early 2025 that through 2026, organizations would abandon 60 percent of AI projects that were not supported by AI-ready data, and reported that 63 percent of organizations either lack the right data management practices for AI or are unsure whether they have them. And KPMG, in a study of more than 48,000 people across 47 countries, found that 73 percent of people report personal benefit from using AI, that only 46 percent are willing to trust AI systems, and that 56 percent of users have made mistakes at work because of unverified AI output. McKinsey’s 2026 survey measures the same split: 80 percent of respondents say AI has improved their own productivity, while 37 percent of organizations can point to any effect on earnings.

80%

say AI has improved their own productivity

McKinsey, The State of AI in 2026

37%

of organizations attribute any effect on earnings to AI

McKinsey, The State of AI in 2026

56%

of users have made work mistakes on unverified AI output

KPMG global study, 2025

The KPMG and McKinsey findings show how both halves of the evidence can be true at once. The personal benefit is real, which is why adoption is close to universal. Trust in the output is missing, so people keep checking it, and work that has to be checked line by line never gets delegated. Delegation is what would move the value from the person to the business, so the gains people report never become results a company can measure.

The next chapter covers why the trust is missing.

Chapter 02

The tools do not learn the business

The stories from inside companies that piloted AI and got little from it run consistently. The tool was impressive in the demo. It wrote fluently, it answered general questions well, and for a few weeks people used it constantly.

Then someone asked it something that depended on how this company prices, or on what the policy became after last year's change. The answer came back wrong, or it came back plausible and stale, drawn from a document that had been superseded months earlier. People checked the output for a while, and the checking took longer than doing the job by hand, so they went back to doing it themselves.

MIT

The GenAI Divide names the learning gap as the core barrier, 2025.

MIT's researchers named the core barrier behind this the learning gap: the tools do not learn from the organization, do not retain context, and do not adapt to the workflow. That description treats the problem as a property of the tools, and it leaves out an earlier problem. The tools cannot learn the business because the material they would have to learn it from is in poor condition, and a better model does not repair the material.

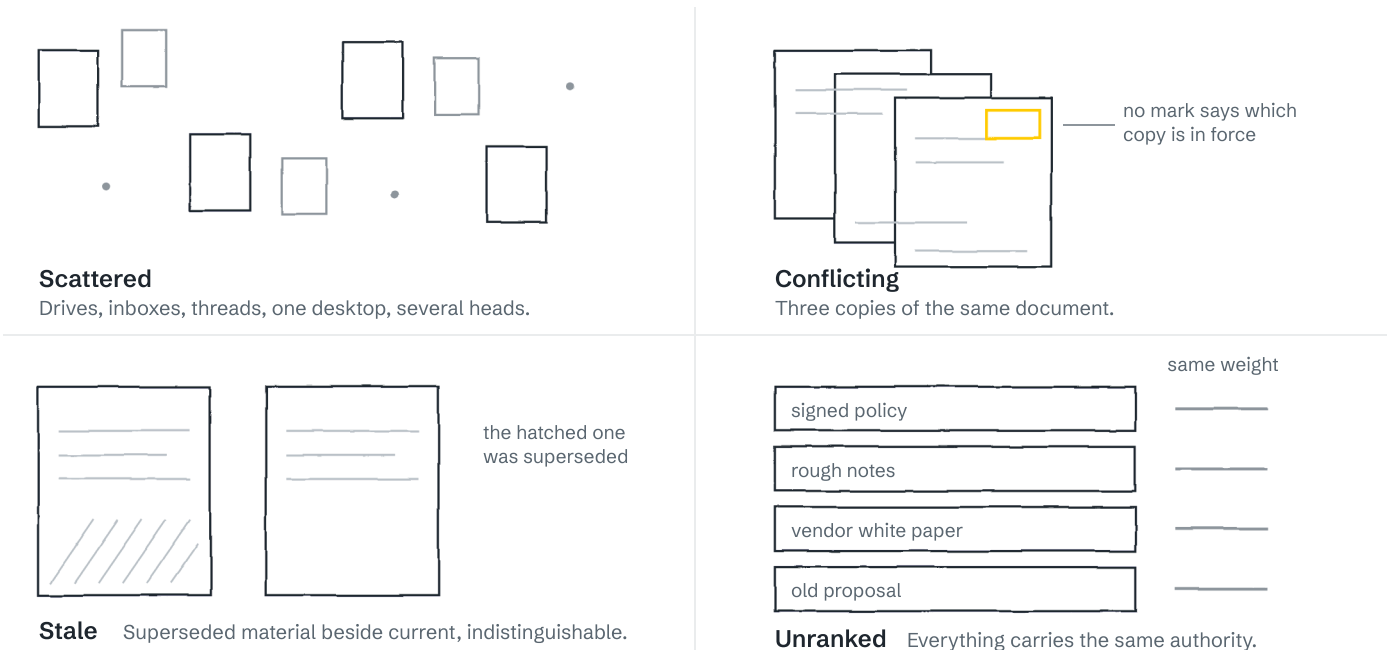
The popular version of this diagnosis, that companies never wrote their knowledge down, has the condition wrong. Almost every company has pricing sheets, templates, old proposals, procedures, and years of email. The failure is in the condition all of that material is in, and the condition has four parts.

Exhibit 02

Schematic

The four conditions of the memory gap

The knowledge exists. These are the four things wrong with the condition it exists in.



The Ermetis lens. Chapter 3 breaks the same failure into the seven links it occurs across.

When a capable AI tool is pointed at material in that condition, it behaves the way the pilot tools behaved: it returns the stale and the conflicting with total confidence, because nothing in the material tells it which version to trust. This is also why the failure survives every change of vendor. The companies in the S&P data that scrapped one set of initiatives and bought the next generation of tools were, in most cases, pointing better software at material in the same condition.

A company in this state has knowledge but no memory.

We call this the memory gap: no layer exists where what the company knows is kept current, kept where the people doing the job can reach it, marked for which version is in force, and weighted by how much each piece deserves to be trusted. The AI deployments that failed on wrong answers were failing at this layer, before the tool was ever involved. Gartner's data-readiness prediction and MIT's learning gap both point at this same missing layer.

60%

of AI projects unsupported by AI-ready data will be abandoned through 2026, by Gartner's prediction

Gartner, February 2025

OECD

AI adoption by small and medium-sized enterprises, December 2025.

The gap is not evenly distributed. Large enterprises can assign teams to it, and a growing number have. In the mid-market, closing it is nobody's day job. The OECD's 2025 study of AI adoption in small and mid-sized companies found the most frequently cited barrier to be the lack of in-house expertise, which fits what the gap requires: sustained work with judgment in it, owned by nobody inside most companies.

Chapter 03

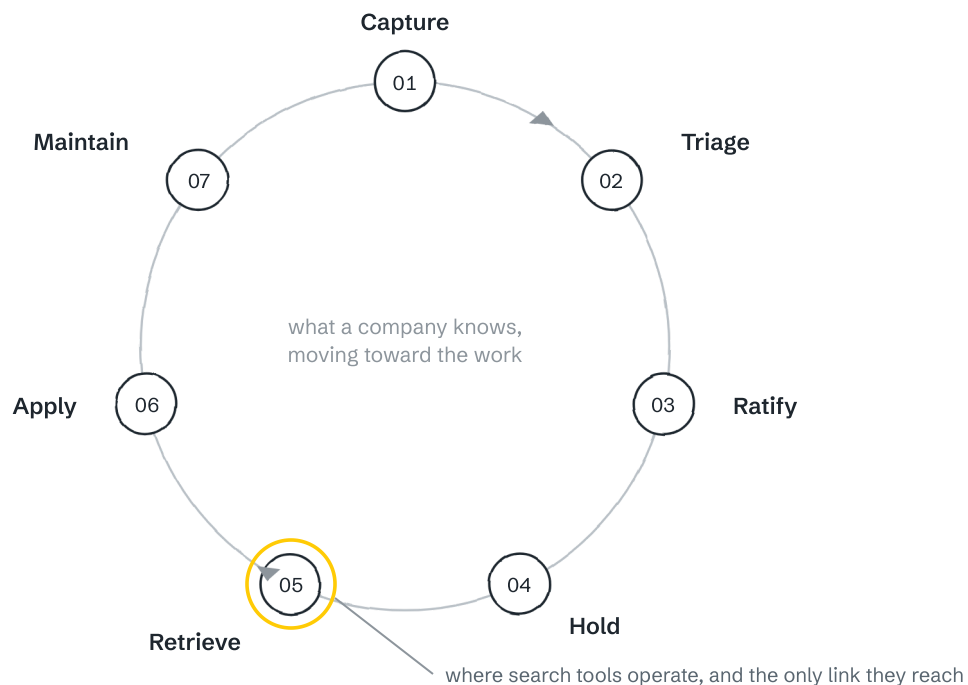
Seven links decide whether knowledge reaches the point of use

Between something being known inside a company and that knowledge changing a piece of work, it passes through seven links. Any of the seven can fail, and each failure has its own fix. This chapter goes through the links one at a time, in enough detail that you can check your own workflows against them.

Exhibit 03

The loop the links form

Knowledge moves through a company in a loop



The Ermetis lens. Each link's failure state is on the following page. The loop closes because what the workflow produces re-enters at capture.

Each link, and the way it fails. Read the right column against the workflow you know best.

01	Capture	It was never recorded anywhere The decision was made on a call or in a hallway, the job ran in a tool that keeps no record, and the lesson stayed with whoever learned it. No search recovers what was never there.
02	Triage	Everything was kept, or nothing was Both read as a policy and neither is one. The everything version is the shared drive nobody trusts. The nothing version is a company that runs on who you ask.
03	Ratify	Versions conflict, and nothing marks the current one Each holder believes theirs is current, and each of them is right about something.
04	Hold	It exists where nobody else can get to it The current one sits in a personal folder, an inbox, or a copy somebody made before the change.
05	Retrieve	The right answer cannot be told apart from the rest The right answer comes back ranked level with twenty plausible ones, and the reader cannot tell which is which.
06	Apply	The answer arrived and did not get used It arrived after the decision was made, or in a form nobody could act on inside the task they were already doing.
07	Maintain	It was right and went stale Prices moved, a job went badly, somebody left, and nothing was watching.

A search tool, including the AI assistant bundled with the licenses most companies already hold, works on one of these links: Retrieve. To a buyer it looks as though it works on all seven, because retrieval is the only part of the chain a user ever sees. The common complaint that a company bought Copilot and it gives wrong answers usually describes failures in the other six links, and a better search tool changes none of them. Retrieval over material that was never triaged, never ruled on, and not maintained returns what chapter two describes, however good the retrieval is.

ISO 30401

Knowledge management
systems: requirements. 2018.

The discipline behind these links is established, and it predates AI. ISO 30401 sets the international requirements for knowledge management systems, governance and roles included. AI raises the stakes, because for the first time software can act on what a company knows, and the condition of that knowledge becomes an operational matter.

Where knowledge enters

One reason capture fails without anyone noticing is that knowledge enters a company through more surfaces than anyone instruments. We count at least nine: conversation, written exchange, the produced work itself, system exhaust such as CRM notes and pricing overrides, rules encoded as configuration in the ERP and the quoting workbooks, field and visual material, external inbound such as trade publications and what a peer said at a conference, outcomes, and what exists only in someone's head. Most companies cover two or three, and the ones they cover are whichever surfaces a purchased tool happened to cover for them.

Two of the nine are almost universally missing, and in our judgment they are the most valuable of the nine. Outcomes go unrecorded because nothing ties a result back to the decision that produced it, so the company cannot say what it has learned across the last ten similar jobs. External inbound goes unrecorded because it does not feel like the company's own knowledge, even though the company's understanding of its market is built from it, so that understanding stays with the individual people who happened to encounter each piece.

How things connect

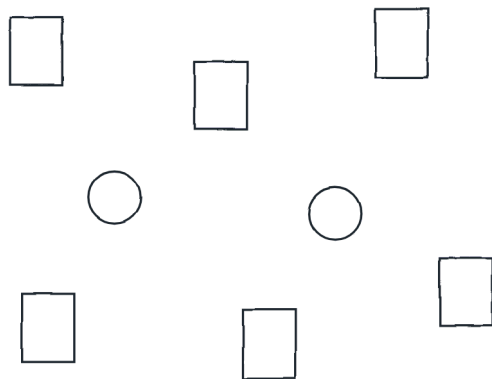
There is one more property, and it does not belong to any single link. A company talks about a stable set of things: customers, jobs, quotes, people, meetings, decisions. Those things relate to each other, and the relationships contain information that no single document states. This job was the third of its type for that customer, and that decision was taken in that meeting, about that account.

Some questions can only be answered by pulling several of those things together: what the last three jobs of this type taught us, or what we have already committed to that would conflict with this plan. No document holds answers of that kind, because they have to be assembled from material that was never filed together. Whether they can be assembled depends on whether the material was stored connected to the things it concerns, and in most companies it was not. This failure is separate from the seven, and it stays invisible until someone asks a question that needs it.

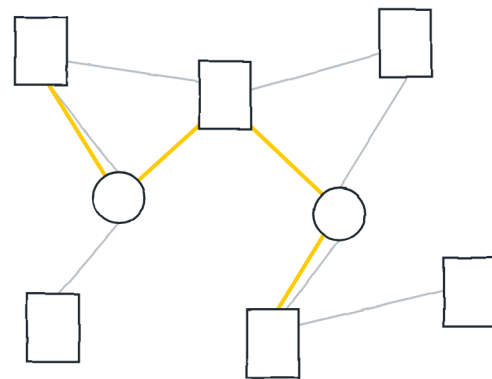
Exhibit 04

The same material, twice

Some answers exist in no single document and have to be assembled



Stored Records, each on its own.



Stored connected The answer is a path across five things.

The question class: what the last three jobs of this type taught us, and what we have already committed to elsewhere. A document search cannot reach it, because no single document holds the answer.

Chapter 04

The other half of the problem is how the workflow runs

The seven links describe the knowledge. They are half of a diagnosis, because knowledge only matters inside a piece of work, and the workflow itself can fail in ways that no amount of well-kept knowledge repairs. When we examine a workflow, four questions cover this second half.

Is there one way this runs?

In many workflows, how an instance goes depends on who picked it up, and two people running the same job would not recognize each other's version. Sometimes a described process exists in an onboarding deck, and the description is not what happens: new people are trained on the drawing and then learn the real thing from whoever sits nearest. A workflow is in good condition here when there is one described way it runs, the description is current, a named person owns it, and where instances vary, the variation is deliberate.

What moves between the steps?

Work changes hands on some trigger, carrying some content. In a weak workflow there is no agreed trigger and no agreed content, so what the receiver needs is frequently missing and the first act after most handoffs is going back for it. In a strong one, each handoff has a named trigger and a defined content, and work does not move until it is complete enough to move.

Does anyone know how it performs?

In many workflows no numbers exist at all, and how the process is doing is a matter of opinion, belonging to whoever is loudest or was most recently burned. The useful standard is modest: a small set of measures, produced on a cadence, looked at by a named person, and cut finely enough that a movement in the number produces a specific question.

Did anyone ever decide who or what does each task?

In most workflows, who does what is an accident of history. Senior people do work that is senior only in the sense that they have always done it. Where the AI question has been asked, it was usually asked at the level of a whole job, and answered by buying one tool for a whole role. The useful state is task by task: the process broken into tasks, each task with an owner by design, the reasoning for a person versus a system written down, and judgment held by people deliberately.

Exhibit 05

The second half of the lens

The four dimensions of how the workflow runs



Definition

Is there one way it runs



Handoffs

What moves between steps



Instrumentation

Does anyone know how it performs



Allocation

Who or what does each task

The Ermetis lens. Scored on the same five levels as the seven links, against the same workflow; chapter 7 turns both halves into a worksheet.

The two halves are read together, because failures in one usually explain failures in the other. The workflow shows you where things stall, and the knowledge underneath usually says why. A handoff that keeps coming back for missing input is, on its face, a definition problem, and underneath it is very often a ratification problem: nobody agrees what a complete handoff contains because nobody ever ruled on it.

The last question, allocation, decides whether AI can go into a workflow at all. A company can hold its knowledge well and still have nowhere to put an agent, because nothing has been broken down past the level of a job. Agents do tasks, and people hold the jobs those tasks roll up to; a workflow that has never been described at the task level has no place to put either the agent or the boundary around it. This is the common case, and it stays invisible to any assessment that examines only the knowledge or only the tools.

Chapter 05

The right target is level three

Any of the eleven things described so far, the seven links and the four dimensions, can be placed on the same five levels. The levels are easiest to give by the sentence you would hear inside the company.

Exhibit 06

One scale, both halves of the lens

The five levels, by the sentence you would hear

1	“You have to ask the right person.” Ad hoc	No practice exists. Whether the thing happens at all depends on who is involved and whether they remember.	
2	“It’s written down somewhere.” Aware	A practice exists in some form, and it works when a particular person pushes. This is where most companies land after buying a tool, because the tool created the practice and nothing governs it.	
3	“We know which version counts.” Standard	There is one way it runs, it is written down, a named person owns it, and it is followed most of the time. Holding it there still takes active leadership.	
4	“We can tell when it’s failing, and we act on it.” Measured	The practice produces evidence about itself, somebody looks at that evidence on a cadence, and the evidence changes what gets done. Failure becomes visible without anyone having to be burned by it first.	
5	“It gets better because people use it.” Compounding	It survives the person who built it, and use itself keeps improving it. Almost nothing in the mid-market is here, and for most companies it is not the right target.	

The ring marks the target we set on most rows; the chapter covers why the top of the scale is usually the wrong destination.

Three things about this scale are easy to miss, and they change what to do with it.

The step from level 1 to level 2 buys very little. Writing things down without governing what was written produces a store nobody trusts, and that store is what most companies own after buying a tool. Many companies did what looked like the right thing, moved from 1 to 2, and got nothing for it, because the value is earned between levels 2 and 4.

Level 3 is where additions to the company start to work. A new hire can become productive on the company's own material without a senior person narrating, because that material is written down and current, and somebody can find it. A first AI agent becomes buildable for the same reason, on the same material: there is a current version for it to cite, one place to read from, and a named person keeping it true. The new hire and the agent are reading the same thing, and the condition of that thing decides how both of them perform.

AI readiness is level 3 on a handful of links in one workflow: a smaller and more local target than most discussions of AI readiness assume.

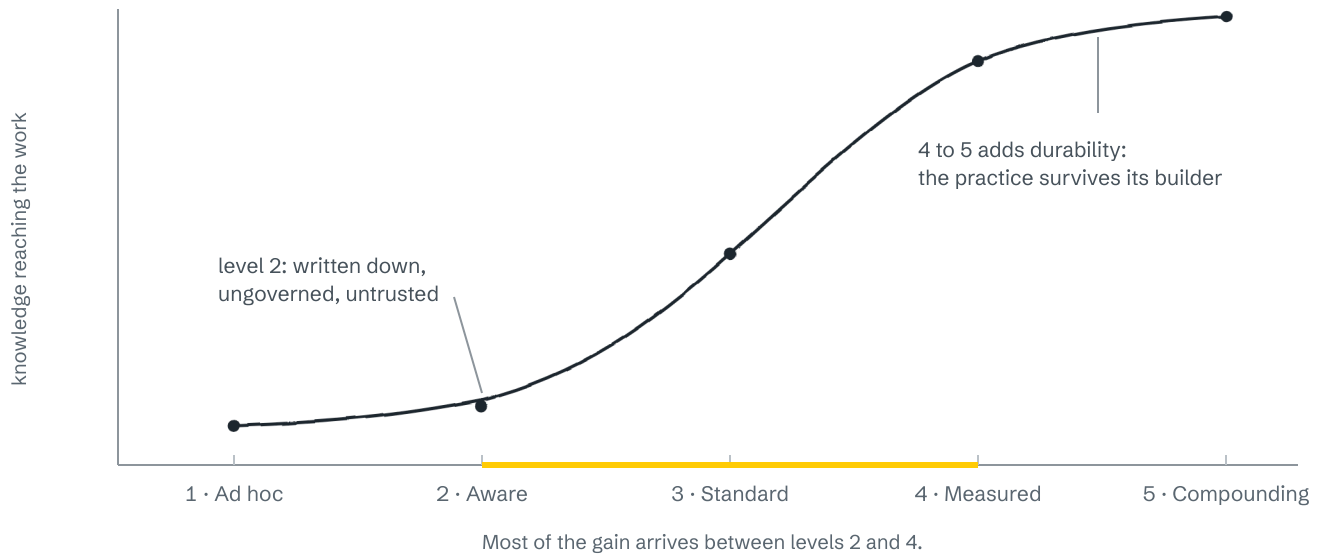
And level 5 is mostly not worth chasing. Level 5 adds durability over level 4: the practice stops depending on the person holding it up. That is worth having on the one or two links where the company's particular risk is concentrated, and a program aimed at level 5 on everything never finishes. When we place a workflow on this scale, we mark level 3 on most rows and level 4 on the row or two where the real risk is. A maturity model that recommends the top of the scale on every row is describing an engagement that never ends.

Exhibit 07

Illustrative; no numeric vertical axis

The value is between levels two and four

How much of what the company knows reaches the point of use, against the level a practice holds.



The Ermetis lens. The curve is illustrative and has no numeric vertical axis; the descriptors behind the levels are on the preceding pages.

One caveat belongs here. Most companies are fine somewhere. Some rows score well, often the storage of documents in a regulated business, and finding those rows is part of the reading. An assessment that finds everything broken should be treated with suspicion, and the rows that score well are part of what makes the rest credible.

Chapter 06

AI pays back first in the same few places

A real record also exists of where AI deployment has paid, measured properly. Four categories account for most of it, and the same few conditions repeat across all four.

Brynjolfsson, Li, Raymond
Quarterly Journal of
Economics, 2025. 5,172
customer support agents.

Customer service is the best-evidenced category in the literature. Brynjolfsson, Li, and Raymond studied 5,172 customer support agents and found AI assistance raised productivity about 14 percent on average, and about 34 percent for the least experienced agents. The assistant in the study was trained on the company’s own resolved conversations, so it was passing the practice of the company’s best agents to everyone else. The company’s own knowledge produced the gain, and the tool was the means of getting that knowledge to each person.

+14%

productivity with AI assistance, across
5,172 support agents

Brynjolfsson, Li, Raymond, QJE 2025

+34%

for the least experienced agents in the
same study

Brynjolfsson, Li, Raymond, QJE 2025

MIT
GenAI Divide deployment
analysis, 2025.

The second category is back-office document work. MIT’s deployment analysis found roughly 70 percent of AI spending going to sales and marketing, while the better returns were coming from back-office automation: the recurring, rule-bound document work in finance and operations.

Exhibit 08

Share of AI budget allocation

The budgets and the returns went to different places

Where the AI budgets went



the better returns were reported here:
back-office automation in finance and operations
MIT, The GenAI Divide, 2025

Source: MIT Project NANDA, The GenAI Divide, 2025. Allocation share as reported; the returns finding is the report's deployment analysis, relayed qualitatively because MIT publishes it that way.

Dell'Acqua et al.

Harvard Business School field experiment, 2023; published Organization Science, 2025. 758 BCG consultants.

Drafting against known rules is the third category, and the same study that supports it also measured the limits. A Harvard Business School field experiment with 758 BCG consultants found that AI assistance significantly improved results on tasks inside the technology's competence, with quality gains around 40 percent, and made results measurably worse on tasks outside it, with correctness dropping about 19 points. Both kinds of task exist within a single profession, which is why the allocation question in chapter four has to be answered task by task. Drafting against written rules, the company's own pricing, standards, and precedents, fell inside the technology's competence in this experiment, and open-ended judgment fell outside it.

The fourth category, internal question answering, has the least measurement behind it, because the cost it addresses has never been tracked: senior people in most companies spend real hours answering the same questions over and over, and no line item records that time. For this category the trust research is the operative evidence. People keep using an answering system when its answers show their sources, because a sourced answer can be verified quickly. An unsourced answer cannot be verified, and after a few early mistakes people stop trusting the system and go back to asking a colleague.

MIT
66% vs 33%: externally
partnered deployments vs
internal builds reaching
production.

Gartner
Press release, June 2025.

Two more findings apply across all four categories. MIT found that externally partnered AI deployments reached production roughly twice as often as internal builds, 66 percent against 33 percent, which suggests the job is more specialized than a demo makes it appear. And Gartner predicts that over 40 percent of agentic AI projects, the autonomous kind, will be canceled by the end of 2027 on unclear returns, escalating costs, and inadequate controls, so the fully autonomous version of this technology is not yet a dependable purchase.

The successful categories have five things in common.

-
- 01 **The task recurs**
 - 02 **The task is bounded**
 - 03 **Something written exists to check the output against**
 - 04 **The output can show its source**
 - 05 **A person owns the result**
-

None of the well-evidenced successes are autonomous or open-ended. The next chapter turns that pattern, and the lens from the three chapters before it, on your own company.

Chapter 07

Most of the first work needs no outside help

Everything in chapters three, four, and five can be applied to your own company by your own people. This chapter is the method, in enough detail to act on. It takes a few hours, and it produces a written view of where the problem is in one workflow, which is worth having before spending anything on tools.

Read one workflow

Do not try to assess the whole company. Pick one workflow, the one where the pain is: the quoting, the intake, the month-end close, the proposals. A company does not have one score. Capture can be excellent in finance and absent in the field, and an average across that describes nobody. One workflow, read honestly, will tell you more than a company-wide survey.

Sit with the people who actually run it, for an hour, and work through the eleven questions on the worksheet overleaf. Four are about the workflow itself: whether there is one way it runs, what moves at each handoff, what numbers exist about it, and whether anyone ever decided who does each task. Seven are about what the workflow runs on, one for each link in chapter three.

Score each answer against the five sentences in chapter five, roughly. A rough pattern is all the next step needs. Most workflows come back uneven, and the finding you are looking for is a pairing: where the workflow stalls, and which link underneath is causing it.

Exhibit 09

Mark one level per row

The one-workflow worksheet

Workflow: _____ · Read with the people who run it · Date: _____

The workflow itself

01	Is there one way this runs, and is the described way what actually happens? <i>Definition</i>	☐ ☐ ☒ ☐ ☐
02	At each handoff, what is supposed to move, and how often does the receiver go back for something missing? <i>Handoffs</i>	☐ ☐ ☒ ☐ ☐
03	What numbers exist about how this performs, and who looks at them? <i>Instrumentation</i>	☐ ☐ ☒ ☐ ☐
04	For each task inside it: who does this, and did anyone ever decide that? <i>Allocation</i>	☐ ☐ ☒ ☐ ☐

What the workflow runs on

05	When something is decided in this workflow, does the decision get recorded anywhere? <i>Capture</i>	☐ ☐ ☒ ☐ ☐
06	Is there a rule for what is worth keeping? <i>Triage</i>	☐ ☐ ☒ ☐ ☐
07	For the documents this depends on, is there a marked current version, with a name and a date on it? <i>Ratify</i>	☐ ☐ ☒ ☐ ☐
08	Can the people who need those documents get to them, or do working copies circulate? <i>Hold</i>	☐ ☐ ☒ ☐ ☐
09	When someone searches, can they tell the right answer from the plausible ones? <i>Retrieve</i>	☐ ☐ ☒ ☐ ☐
10	Does what exists get used inside the workflow, or do people ask a colleague because asking is faster? <i>Apply</i>	☐ ☐ ☒ ☐ ☐
11	When something changes, prices, policy, people, does anything catch the documents that just went stale? <i>Maintain</i>	☐ ☐ ☒ ☐ ☐

The five circles are the five levels of chapter five: 1 you have to ask the right person · 2 it is written down somewhere · 3 we know which version is in force · 4 we can tell when it is failing · 5 it gets better because people use it. The yellow ring marks the level worth reaching on most rows.

The Ermetis lens, in fill-in form. Copy this page freely; it works on paper.

What to fix first

Do these in order, because each step depends on the ones before it.

First, mark the current versions. For the five or ten documents the workflow actually depends on, have the person with final say declare which version is in force, put their name and a date on it, and mark or archive the rest. This takes a few hours, and of everything in this report it is the change we would make first, because people and software fail the same way against conflicting versions: nothing downstream can tell which answer to trust.

Second, put review dates on the documents that go stale: prices, rate tables, templates, standard terms. A review date commits a named person to re-confirming the document by a certain day, so staleness gets found on a schedule. Without the dates, a stale number is usually discovered after it has already gone out in a quote.

Third, start recording outcomes. When a bid resolves, when a job closes, when a client renews or does not, record the result next to the decision that produced it, in one line if one line is what will actually get written. Outcomes are the missing surface in almost every company, and recording them is what lets the company answer, a year from now, what all those similar jobs taught it.

Fourth, break the workflow into tasks and decide each one deliberately. Most tasks will stay where they are, and the decision is still worth making, because you will then know which tasks are judgment, which are routine, and which have a written rule behind them. AI can eventually take on the tasks that have a written rule and a checkable output, and it should stay out of the judgment tasks, which is the boundary the Harvard experiment in chapter six measured. Writing that boundary down is most of what a safe deployment needs.

Exhibit 10

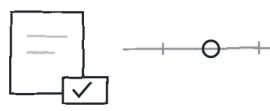
None of these needs a tool

The four fixes, in order



01 · Mark current versions

One copy in force, named and dated.



02 · Add review dates

Staleness gets a schedule.



03 · Record outcomes

Result stored next to its decision.



04 · Decide each task

Judgment stays with people.

The Ermetis lens. Each fix uses what the one before it produced; the method for each is in the text above.

What to refuse to buy

The reading also tells you what to decline.

Refuse any tool sold as fixing all of this at once. The seven links are seven different problems with seven different fixes, and the tools on the market each work on one or two of them. A tool that claims the whole chain is claiming more than any current tool delivers.

Refuse retrieval over an ungoverned store. If nothing marks current versions and nothing ranks the material, a better search tool returns the same wrong answers faster. Most companies have already made this purchase once, and many of the abandoned initiatives counted in chapter one were versions of it.

Refuse answers without sources. Output that does not show where it came from cannot be checked, and unverified output is how 56 percent of the users in the KPMG study ended up making mistakes at work. Once a system has been wrong a few times with no way to see why, people stop using it, and the company keeps paying for licenses nobody relies on.

Refuse the autonomous employee. The well-evidenced successes are all assisted and bounded, with a person owning the result, and the caution has standards behind it: NIST's AI Risk Management Framework makes human oversight core to trustworthy AI, and names fabricated output as a standing risk of generative systems.

And refuse to buy anything before the reading. A tool decision made ahead of the diagnosis is a solution chosen before the problem is known, and the scrapped initiatives in chapter one were, for the most part, bought in that order, with the tool chosen first and the diagnosis never done.

Chapter 08

About this report

Who we are

Ermetis figures out what AI should do for a business, builds it on what the company already knows, and runs it with the team as a fractional member of it. We are a small company and we take on a few clients at a time, on purpose, because what this report describes cannot be done at arm's length. We have no ties to any vendor and take no referral fees or affiliate revenue. What we set up belongs to the client, from the knowledge itself to the playbooks and configurations built on it; third-party tools stay on the client's own licenses, and leaving takes thirty days.

Who wrote it

The Memory Gap was researched and written by John Noland, the founder of Ermetis. His profile is at [linkedin.com/in/johnnoland](https://www.linkedin.com/in/johnnoland).

The free assessment

The free assessment applies the method in chapter seven to one of your own processes. Pick one process that is costing you something, fill in a short form, and spend thirty minutes on a call with us. Within two business days you get a written assessment of where that process stands and what we would do first. Choose a time at ermetis.co, or write to info@ermetis.co.

Methods and sources

Research for this report was compiled and verified in August 2026, and the second edition, in September 2026, updated chapter one with McKinsey's 2026 survey. Every figure is relayed from the named, published source cited beside it; none originates with Ermetis. Where we could check a figure against the primary document, we did: the McKinsey 2026 report, the Gartner data-readiness release, the Gartner agentic-AI release, the KPMG trust study, ISO 30401, and the NIST AI Risk Management Framework were verified against their primary publications. For the remaining figures, including the MIT, BCG, and S&P Global findings, we did not obtain the primary documents; each figure was checked across multiple independent published accounts of the named report.

The lens in chapters three through five is Ermetis's own. Its descriptors are reasoned from worked examples and from the published research cited here; they have not been calibrated against a statistical sample, and we make no claim that they have. Nothing in this report draws on client data.