WeAreBrain

THE GROUNDWORK ADVANTAGE

A C-suite handbook for building AI that creates lasting value

CONTENTS

INDEX

01 INTRODUCTION

Most of what has been published about AI in the past two years describes what the technology can do. Very little of it describes what has to be true inside an organisation before that technology produces anything of value. This handbook is about the second question, because it's the one that decides whether an AI initiative survives its first budget review.

Research from MIT and BCG, conducted separately and on different populations, points to the same uncomfortable range: somewhere between 60 and 95 per cent of enterprise AI initiatives fail to produce measurable business value, depending on how strictly "value" is defined.[1, 2] MIT's Project NANDA found that 95 per cent of generative AI pilots showed no measurable impact on profit and loss.[1] None of this is research about whether the technology works; it is research about what happens when it is asked to run on top of business processes, data and governance that were never built to support it.

That is the argument this handbook makes. Most organisations already have access to models that are more than capable of the work being asked of them. What they lack is the foundation underneath those models: explicit business objectives, trustworthy data, governance built in from the start, and infrastructure that doesn't lock the organisation into a single vendor's roadmap.

What this handbook contains

Chapter 2 sets out the principles that separate the organisations that scale AI from those that don't: data ownership, regulation as a design input, and the real cost of fragmented adoption. Chapter 3 introduces the Groundwork Model: a way of identifying and prioritising where AI creates measurable value, and of evaluating investment at the level of the groundwork rather than the individual application. Chapter 4 turns the model into a three-part sequence (deciding where to invest, building it, and scaling it) and sets out the decisions and ownership questions each phase demands. Chapter 5 addresses what AI-assisted software development does and doesn't change about those decisions. Chapter 6 closes with a self-assessment and a concrete starting sequence.

What you should expect

This handbook sets out what it takes to get lasting value from AI. It starts from the position that AI failure is a foundation problem rather than a model problem, and works through what that means for the decisions the C-suite actually controls. It is vendor-neutral by design, and the reasoning holds regardless of who executes it.

By the end, you should have a working framework for identifying where AI creates measurable value in your organisation, understand what has to be true before it can be realised reliably, and know which decisions are yours to make, and to keep, whoever you build with.

02 FUNDAMENTAL PRINCIPLES

What actually makes AI succeed

Ask ten AI vendors what determines success and most will describe their product. Ask the organisations that have scaled AI past the pilot stage, and the answer is different. PwC's 2026 research into enterprise AI programmes found that the technology accounts for only about 20 per cent of an initiative's value, with the other 80 per cent coming from redesigning the work around it, and that the companies seeing returns concentrate on a few workflows chosen by leadership and build shared, reusable components once rather than per project.[3] What that depends on is standardised processes, data good enough to trust, and governance designed in rather than added afterwards.

RAND's interviews with 65 data scientists and engineers found the inverse: the most cited causes were leadership specifying the wrong problem, and data not good enough to train on.[4] Organisations that skip that groundwork and go straight to deployment don't avoid the work; they defer it, and pay more for it later, once they're rebuilding data architecture and retrofitting governance around a system that's already in production.

This matters because it moves the real investment decision. The question stops being which AI tool to buy and becomes whether your data, your process and your governance are in a state where any AI tool could succeed. Get that right, and the specific model or vendor becomes a comparatively minor decision, replaceable without disrupting the business. Get it wrong, and no model, however capable, will produce a result the organisation can trust or scale.

That state of readiness, where data, process, and governance could support any AI tool, is what the rest of this chapter unpacks. Four things determine whether an organisation has it: whether it owns its data or has ceded that ownership to its vendors; whether regulation is designed into AI systems from the start or retrofitted around them; whether existing systems of record are treated as usable assets; and whether AI adoption is building one coherent foundation or fragmenting into dozens of disconnected tools. Each is examined in turn below, and each is a place where the failure statistics in the introduction originate.

Data ownership

Data ownership determines whether an organisation controls its own AI trajectory or remains dependent on a vendor’s pricing, roadmap, and continued existence. Owning your data means knowing where it lives, who can access it, how it flows into any AI system that touches it, and, critically, being able to move that system elsewhere without losing the value already built into it.

And this is a live commercial concern, because fragmented adoption compounds the ownership problem. Torii’s 2026 benchmark found the average enterprise now runs over 800 applications, more than 60 per cent of them outside formal IT oversight.[5] Larridin’s 2026 research found enterprises running an average of 23 AI tools, with only 38 per cent of organisations maintaining a comprehensive inventory.[6] Every one of those tools is a place data has moved to, often without anyone deciding it should. Much of that AI arrives as features embedded in software the organisation already pays for, which is why standalone tool reviews miss it.[7]

GDPR and the AI Act as design principles, not obstacles

Organisations frequently treat data protection regulation as a constraint to be managed around AI initiatives, which turns out to be an expensive habit. Deloitte's 2026 Tech Trends research found that the gap between organisations piloting agentic AI (38 per cent) and those running it in production (11 per cent) comes down to governance and integration challenges. In practice, that usually means no named owner, no audit framework, and no defined process for what happens when the system gets something wrong.[8]

Designed in from the start, the requirements of GDPR and the AI Act (explainability, data lineage, human oversight, the ability to say precisely why a system produced a given output) are also exactly what makes an AI system trustworthy enough to scale.[9, 10] Auditability, in particular, is the mechanism that lets a business rely on what the system tells it.

Existing software investments still matter

A recurring assumption in AI adoption is that legacy systems are the obstacle, and that real progress requires replacing the ERP, the CRM, the document management system, all of it, with something AI-native. In practice, the replacement is rarely necessary: what most organisations are short of is a way to connect the data inside those systems to something that can reason over it. The systems of record can, in most cases, stay exactly where they are. What changes is the layer that sits above them: the layer that understands what's inside those systems well enough to make it usable.

Why fragmentation is expensive

Worqlo's 2026 analysis puts a number on the fragmentation problem: roughly $1,800 to $2,800 per user per year in direct licensing costs, against $800 to $1,400 for a consolidated approach.[11] That is a real, budgeted cost, before accounting for the security exposure of data spread across dozens of ungoverned tools. Lenovo's 2026 survey of 6,000 enterprise employees found that between a fifth and a third of workers use AI outside formal IT oversight.[12]

Every isolated AI solution an organisation adds (another chatbot here, another document assistant there) solves a narrow problem while adding a new silo, a new integration point, and a new governance gap. Each decision is reasonable on its own; collectively, they produce exactly the pattern the statistics above describe.

What a reusable foundation delivers instead

The better approach is to build the data, governance, and infrastructure layer once, correctly. Each new use case then becomes progressively less expensive because it builds on a shared foundation instead of starting from scratch, which is the structural argument behind the model in the next chapter.

03 THE GROUNDWORK MODEL

The core idea

Most AI initiatives are evaluated one use case at a time: does this chatbot save time, does this automation reduce errors, and it's the wrong unit of analysis. The Groundwork Model evaluates AI investment at the level of the foundation an organisation is building, because the foundation determines whether the fifth use case is cheaper and faster than the first, or whether the organisation is rebuilding the same groundwork every time. That failure mode gives the model its name: the groundwork is either done once, deliberately, as an asset, or done repeatedly, invisibly, as a cost.

The model has two parts that work together: a way of identifying and prioritising where AI creates measurable value, and a foundation of governed data, security, and infrastructure that makes realising that value repeatable.

Identifying and prioritising use cases

Not every process that could use AI should be the first one built. The organisations that scale successfully prioritise three factors together, and the weighting matters as much as the factors themselves.

Volume	asks how often the process or decision happens. A use case that runs twice a week won't demonstrate value quickly enough to build organisational confidence, however well it works.
Friction	asks how much of the current time spent is genuine decision-making and how much is retrieval, formatting, chasing information or coordinating across systems; the second category is where AI adds the most measurable value fastest.
Data readiness	asks how much work is required before this specific use case can run on trustworthy, governed data. It's the step most commonly skipped, and the one the failure statistics trace back to most consistently.

Score all three together. A use case that is strong on one factor and weak on the others is a trap: high volume running on unready data is precisely how pilots fail in public.

Calculating expected business impact

Before committing resources, estimate the value a use case is likely to generate across three separate categories; blending them into a single ROI figure is how the most important one gets lost.

Direct efficiency covers hours and cost saved, and is measurable almost immediately. Quality improvement covers error reduction, consistency, and customer or employee experience; it's real, but tends to become visible with a lag. Strategic option value covers what becomes possible afterwards that wasn't possible before. It's the hardest of the three to quantify, and often where the largest long-term value sits, which is exactly why it deserves to be named explicitly before it disappears into the other two.

BCG's 2025 research across 1,803 senior executives found that 60 per cent of companies have yet to define or monitor any financial KPIs for the value AI creates.[13] Much of that is a measurement gap: the impact was never estimated with this level of structure before the project started, so there was nothing concrete to measure it against afterwards.

The foundation underneath: what Sinas provides

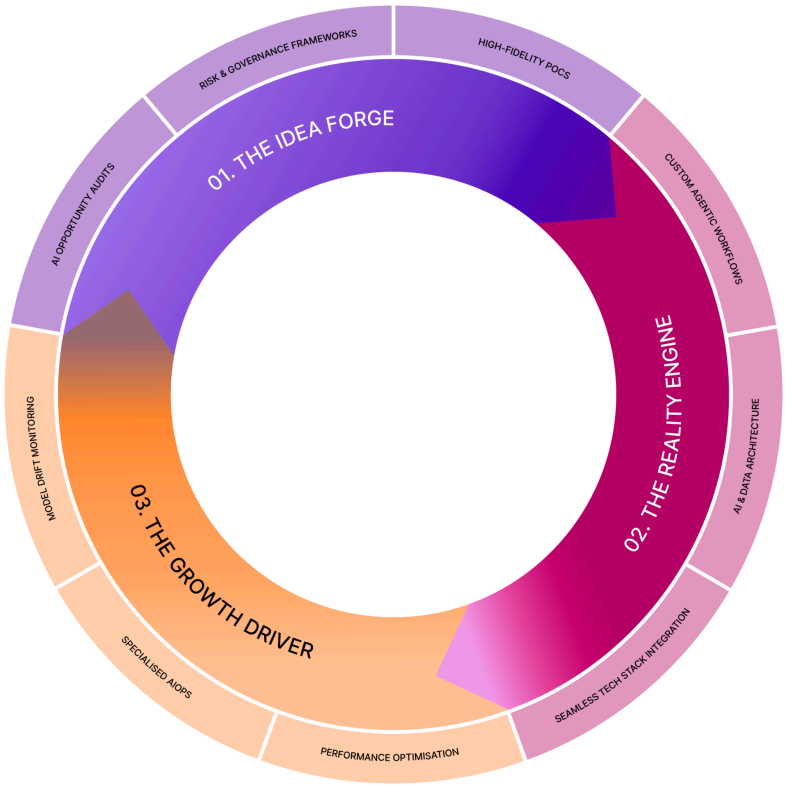
Once a use case is prioritised and its expected value estimated, it needs somewhere to run that doesn't recreate the fragmentation problem described in the previous chapter. Sinas, an open-source AI runtime platform we've developed at WeAreBrain, is our answer to that requirement. It indexes enterprise knowledge, enforces governance, and orchestrates retrieval across AI use cases, and because it's open source, it belongs to the organisation running it, not to us. We introduce it here as a concrete illustration of what the foundation layer has to provide, once, so that no subsequent use case has to rebuild it.

Every answer the platform produces is traceable to its source, which means decisions can be explained after the fact instead of reconstructed under pressure during an incident. Access controls and permissions are enforced consistently across every use case built on it, in place of the separate and usually inconsistent configuration that fragmented tooling demands. Enterprise knowledge is indexed and made available to AI systems without the data ever leaving the organisation's control. The orchestration layer is separable from the model underneath it, so a change in model pricing, availability or terms doesn't require rebuilding the application layer. And because the platform is open source, the organisation retains the ability to run it, modify it, or move away from any single implementation partner without losing what's already been built.

The foundation determines whether AI investment compounds or stalls: organisations with shared data, governance and infrastructure deliver each new use case with less effort than the last, while organisations that buy AI one tool at a time recreate the same work across a growing collection of silos.

04 HOW TO APPLY THE MODEL

Everything in the Groundwork Model is a way of reasoning about AI investment. Reasoning has to turn into a sequence of decisions and actions at some point, and that sequence looks the same regardless of who executes it, whether entirely in-house or with an implementation partner. It happens in three phases: deciding where to invest, building it, and scaling it. Each phase has its own owner, its own decisions, and its own way of failing, so they're worth taking one at a time. We've run this sequence often enough that it has become a standing methodology inside WeAreBrain. We call it Future Catalyst, and its three phases the Idea Forge, the Reality Engine, and the Growth Driver.



FUTURE CATALYST — THE THREE PHASES AND WHAT EACH DELIVERS

Phase one: deciding where to invest (the Idea Forge)

Before any use case is built, three things should exist: a named business owner for AI initiatives, with real authority and no defaulting to IT; an honest inventory of the AI activity already happening across the organisation, sanctioned or not; and a working definition of success for the first use case, agreed before the build starts.

The inventory step matters more than it sounds. With a substantial share of enterprise AI activity running outside formal IT oversight, most organisations begin this process knowing less about their current AI usage than they assume.

Using the three prioritisation factors of volume, friction, and data readiness, the first use case should score well on data readiness even if it doesn't score highest on volume or strategic value. The purpose of the first use case is to prove the foundation works end to end, from data to governance to application, with a result the organisation can trust. That's worth more, at this stage, than picking the single highest-potential use case and discovering the data behind it wasn't ready.

This phase is complete when you have a validated, quantified case for what to build first: enough evidence, short of a working system, to commit budget with confidence.

Phase two: building it (the Reality Engine)

This is where the use case, and the foundation underneath it, get built, and three questions determine how well it goes.

Who should be involved, and where responsibility sits. The business owner defines the objective and validates whether the outcome matters commercially. IT and data teams are responsible for what the foundation requires: access, governance, and integration with existing systems. An implementation partner, where one is used, brings the pattern recognition of having built this foundation before, and should be judged as much on how much of what they build is reusable for the next use case as on whether the first one works.

What should not happen. AI initiatives owned entirely by IT, with the business only consulted once something is ready to demo. That sequencing is precisely what produces the tool-sprawl and shadow-AI patterns described earlier, because if the business doesn't feel ownership of the sanctioned path, individual teams route around it.

How the organisation prepares. Preparation here is simpler than a change-management programme: the first use case has to be genuinely visible, with its objective, its expected impact and its outcome communicated honestly, including if it underperforms. Organisations that treat the first use case as a proof point people can see and evaluate build more internal trust for the second one than organisations that treat it as an IT project running in the background.

This phase is complete when the outcome has been measured directly against the three-category impact estimate made before the build; a general sense that things are going well doesn't count. If the estimate was wrong, that's useful calibration for the next one, and no reason to abandon the measurement discipline.

Phase three: scaling (the Growth Driver)

Gartner's 2026 survey of 1,303 organisations found that only 22 per cent have successfully scaled AI across multiple business units.[14] Phase three is where the foundation argument at the heart of the Groundwork Model pays off directly. If the first use case was built on governed, reusable infrastructure, the second one inherits the access controls, the auditability, and the data indexing already in place. It starts from the application layer. If it was built as a standalone project instead, the second use case starts the entire foundation-building process over again. The difference between these two outcomes is usually visible within the first year, in both cost and delivery time.

Scaling, in other words, is the payoff for having built the first use case correctly, and it begins the moment the second one does.

This phase is working when each new use case is a smaller project than the one before it: starting from the application layer, and visibly cheaper and faster to deliver.

05 AI CAN BUILD SOFTWARE. WHAT DOES THAT CHANGE?

What's genuinely changing

AI-assisted development has measurably changed how fast software gets written; code that used to take days can be scaffolded in hours, and treating that shift as a temporary novelty would be a mistake.

What doesn't change

Faster code generation does not remove the need to decide what should be built, why, for whom, and under what constraints. It doesn't resolve who owns a decision when an AI-assisted system produces the wrong output in production. It doesn't establish governance, define data access, or determine whether a given use case is worth building in the first place. Those are judgment calls, and they require understanding the business.

If anything, faster code generation raises the cost of getting the judgment wrong, because it's now possible to build the wrong thing at much greater speed and scale than before. Organisations that treat AI-assisted development purely as a productivity tool for developers, without adjusting how architecture and governance decisions get made, tend to accumulate technical and organisational debt faster than they did before, just less visibly, because the code itself looks like it was produced quickly and confidently.

What this changes about implementation partnerships

The value an implementation partner provides is shifting away from "we will write this code for you" towards "we will make sure what gets built is architected correctly, governed properly, and connected to a foundation that supports what you build next." That's a different relationship from the traditional software vendor one, and it's worth being explicit about it: the code is increasingly the easy part. The decisions around it (what to build, on what foundation, with what governance, and how it fits into everything built before and after it) are where experienced judgment still matters, and where most of the failure statistics originate.

That judgment is exactly what the Idea Forge, described in the previous chapter, exists to apply before a single line of code gets written.

06 WHERE THIS LEAVES YOU

Every idea in this handbook points to the same conclusion: the constraint on AI success, in the large majority of organisations studied by MIT, BCG, PwC, RAND, Deloitte, and Gartner, was never the model.[1, 2, 3, 4, 8, 14] It was what the model was asked to run on: the data, the governance, and the process underneath it. The reframing is the useful part, and what follows is how to act on it.

A short self-assessment

Before selecting a use case, it's worth working through a handful of questions honestly, on your own or as a short discussion among the people who'd own this work. Each one earns its place here because of what the answer reveals.

Ownership. Is there a named business owner for AI decisions, or does the mandate sit with IT by default? The answer matters because everything in chapter 4 depends on someone having the authority to prioritise and say no, and a default mandate is how initiatives end up owned by nobody.

Visibility. Could you produce a list of the AI tools already in use across the organisation, sanctioned or not, this week? If the list would take a month, that gap is the measure of how much AI activity is running outside your oversight, and of how far your data has already travelled.

Readiness. Are you able to name one process where you could estimate volume, friction and data readiness today, using what you already know, without commissioning a study? If you can, you have a candidate first use case. If you can't, the work starts further back than use-case selection.

Honesty about prior projects. If your first AI initiative underperformed, would that be visible and discussed, or quietly absorbed and dropped? The measurement discipline this handbook argues for only survives in organisations where a shortfall is treated as information, and quiet absorption is how estimates never get calibrated.

Assumptions. Is there a system of record you're assuming would need to be replaced, that could instead stay exactly where it is? Replacement is the most expensive assumption in AI adoption, and it's rarely tested before it starts driving the budget.

If most of these are easy to answer, you're closer to ready than most organisations. If several are genuinely uncertain, that uncertainty is the starting point, and the first thing to resolve.

Where to start, in order

Pulled directly from the three phases of Future Catalyst, sequenced as concrete first actions:

1. **Name an owner:** someone accountable for AI decisions, with the authority to prioritise and say no. A committee doesn't count, and neither does IT by default.
2. **Inventory what already exists:** a working list of AI tools currently in use across departments, including the ones nobody formally approved.
3. **Pick one process to measure:** apply the volume, friction and data-readiness questions to a single real process. The most ambitious one can wait; start with the one with the clearest data.
4. **Estimate impact before building anything:** a rough, written estimate of direct efficiency, quality improvement, and strategic option value, so there's something to measure against later.
5. **Build the foundation once, for that first use case:** governance and access controls that the second and third use case can reuse.
6. **Make the result visible:** report what actually happened against the estimate, including if it fell short, before deciding what to build next.

What this looks like in practice

A simplified illustration, with no specific engagement behind it: a logistics company handling shipment documentation exceptions, meaning cases where paperwork doesn't match the shipment and someone has to investigate and correct it manually. Applying the model:

Volume is high, at several hundred exceptions a month. Friction is mostly retrieval and cross-referencing rather than judgment: someone pulling records from three systems to figure out what went wrong. Data readiness is moderate: the records exist, but live in separate systems that were never designed to be queried together. Estimated impact splits cleanly into the three categories: direct efficiency from reduced handling time per exception, quality improvement from fewer downstream errors reaching customers, and a strategic option, because once the underlying data is connected and queryable, the same foundation supports exception-handling in adjacent processes without rebuilding it.

The numbers will differ from one organisation to another, while the discipline remains the same: select a use case because the underlying data is workable, and define its expected value with enough precision to measure the outcome afterwards.

None of this requires an external partner to begin; it requires treating AI adoption as organisational design work first (ownership, sequencing, governance) and a technology purchase second. Whoever leads that work succeeds by building it in this order.

07 THE GROUNDWORK CHECKLIST

A one-page summary of the handbook. Each item is expanded in the previous chapters; use this as the working reference once you've read them.

BEFORE ANYTHING IS BUILT

- [ ] A named business owner for AI decisions is in place, with the authority to prioritise and say no. A committee doesn't count, and neither does IT by default.
- [ ] A current inventory of AI tools in use across the organisation exists, including unsanctioned ones.
- [ ] Data ownership is understood: you know where your data lives, who can access it, and what it flows into.
- [ ] GDPR and AI Act requirements are treated as design inputs for any new system.[9, 10]
- [ ] Existing systems of record are treated as assets to connect to.

CHOOSING THE FIRST USE CASE

- [ ] Candidate use cases are scored on all three factors together: volume, friction, and data readiness.
- [ ] The first use case scores well on data readiness, even if others score higher on ambition.
- [ ] A definition of success is agreed before the build starts.
- [ ] Expected impact is estimated in writing across three separate categories: direct efficiency, quality improvement, and strategic option value.

BUILDING IT

- [ ] The business owner defines the objective and validates the outcome, with IT and data teams responsible for what the foundation requires.
- [ ] The foundation is built to be reused: governance, access controls, and data indexing that the next use case inherits.
- [ ] Any implementation partner is judged on how much of what they build is reusable, as well as on whether the first use case works.

- [ ] The initiative is visible across the organisation: objective, expected impact, and outcome communicated honestly, including if it underperforms.

MEASURING AND SCALING

- [ ] The outcome is measured directly against the written impact estimate.
- [ ] Results are reported before the next use case is chosen, including shortfalls.
- [ ] Each new use case starts from the application layer, inheriting the foundation already in place.
- [ ] The trend is checked over time: each use case should be cheaper and faster to deliver than the one before it. If it isn't, the groundwork is being rebuilt somewhere.

08 SOURCES

[1] Challapally, A., Pease, C., Raskar, R., & Chari, P. (2025, July). The GenAI divide: State of AI in business 2025. MIT Media Lab, Project NANDA. Reported by AI Sweden. https://my.ai.se/resources/the-genai-divide-state-of-ai-in-business-2025

[2] Boston Consulting Group. (2025). Are you generating value from AI? The widening gap. https://www.bcg.com/publications/2025/are-you-generating-value-from-ai-the-widening-gap

[3] PwC. (2026). 2026 AI business predictions. https://www.pwc.com/us/en/tech-effect/ai-analytics/ai-predictions.html

[4] Ryseff, J., De Bruhl, B. F., & Newberry, S. J. (2024). The root causes of failure for artificial intelligence projects and how they can succeed: Avoiding the anti-patterns of AI (Report No. RRA2680-1). RAND Corporation. https://www.rand.org/pubs/research_reports/RRA2680-1.html

[5] Torii. (2026). The SaaS benchmark annual report 2026. https://www.toriihq.com/saas-benchmark-annual-report-2026

[6] Larridin. (2026). 2026 state of enterprise AI report. https://larridin.com/state-of-enterprise-ai

[7] Larridin. (2026, July 2). How AI tool sprawl is draining your tech budget. https://larridin.com/blog/ai-tool-sprawl-enterprise

[8] Deloitte. (2026). Agentic AI strategy. Tech Trends 2026. https://www.deloitte.com/us/en/insights/topics/technology-management/tech-trends/2026/agentic-ai-strategy.html

[9] European Parliament and Council of the European Union. (2016). Regulation (EU) 2016/679, General Data Protection Regulation. Official Journal of the European Union, L 119, 1–88. https://eur-lex.europa.eu/eli/reg/2016/679/oj

[10] European Parliament and Council of the European Union. (2024). Regulation (EU) 2024/1689, Artificial Intelligence Act. Official Journal of the European Union. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

[11] Worqlo. (2026, May 5). Enterprise AI tool sprawl: What it's really costing you. https://worqlo.com/blog/enterprise-ai-tool-sprawl/

[12] Lenovo. (2026). Work reborn research series 2026. Reported by Help Net Security, 1 May 2026. https://www.helpnetsecurity.com/2026/05/01/shadow-ai-risks-it-oversight/

[13] Boston Consulting Group. (2025, January 15). From potential to profit: Closing the AI impact gap. BCG AI Radar. https://www.bcg.com/publications/2025/closing-the-ai-impact-gap

[14] Gartner. (2026, September 1). Gartner survey finds only 22% of organizations have successfully scaled AI across multiple business units. Reported by AIwire. https://www.hpcwire.com/aiwire/2026/09/01/gartner-finds-just-22-of-organizations-have-scaled-ai-across-business-units/