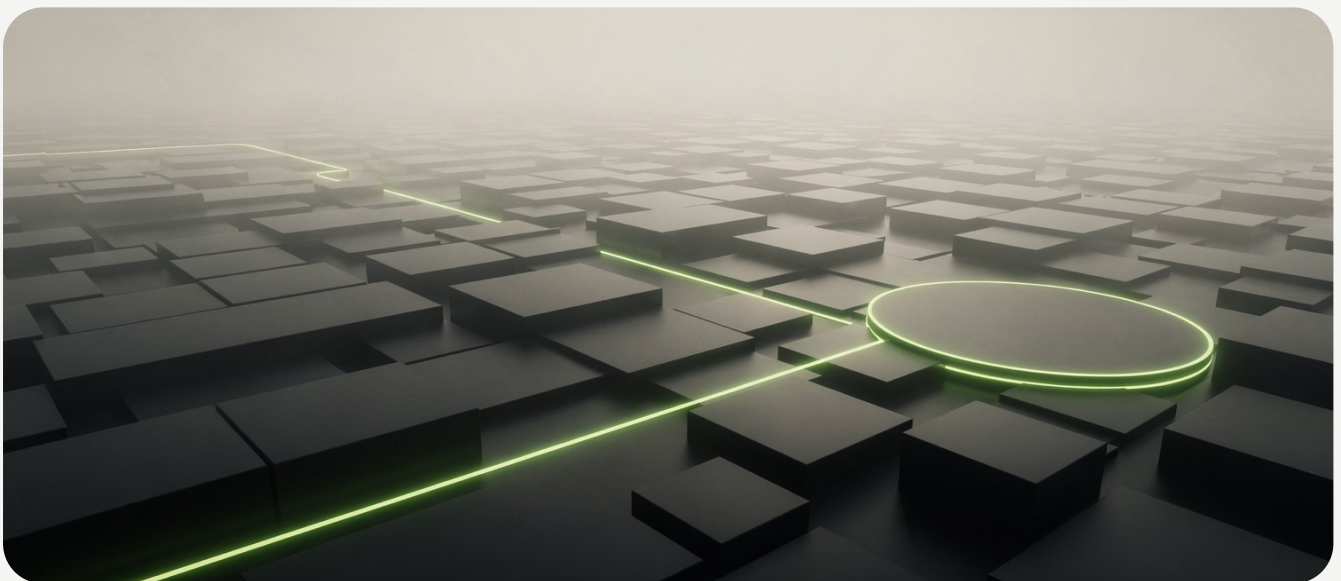


ENTERPRISE AI STRATEGY · THE OPERATIONS CHAPTER

The Last Mile of AI-Native Software

AI has collapsed the cost of building software. It has not collapsed the cost of running it. This paper supplies the chapter most AI strategies are missing: who takes AI-built software live, who runs it, and who answers the auditor. With a five-level maturity model and a drafted chapter, ready to adapt.

**42%**

of companies now abandon most of their AI initiatives, up from 17% one year earlier.

S&P GLOBAL · 2025 · N=1,006

>40%

of agentic AI projects will be cancelled by end of 2027: cost, unclear value, inadequate risk controls. A forecast, not a fate.

GARTNER · JUN 2025

~80%

of large-scale IT budgets already goes to operating and maintaining existing systems. Every new app lands on this ledger.

US GAO · JUL 2025

Contents

ES	Executive summary	03
01	The bottleneck has moved · why AI programs stall after the build	04
02	The missing chapter · run cost, shadow building, unstoppable demand	05
03	The regulatory floor is now explicit · DORA, NIS2, AI Act, BaFin	06
04	The AI-Native Operations Maturity Model · five levels, one self-assessment	07
05	Sovereignty, stated precisely · three tiers, one that gets skipped	09
06	What to require from an operations layer · eight testable requirements	10
07	The chapter, ready to adapt · a drafted operations chapter	11
08	How nyxory implements this · the vendor section, clearly labeled	12
SRC	Sources, reading notes, about the author	12

In a hurry? Read the executive summary, locate your organization in the maturity model on page 07, and lift the drafted chapter on page 11 into your strategy paper. The rest of the document is the evidence.

The failure is not in the building. It is in the last mile.

Coding agents produce working applications in hours; business units that never wrote software now commission it in an afternoon. Yet enterprise AI programs are stalling at a rising rate, and the diagnosis is consistent across every major research house. Four findings set up this paper:

- **The failure point has moved.** 42% of companies now abandon most of their AI initiatives, up from 17% one year earlier; the average organization scraps 46% of proof-of-concepts before production (S&P Global, 2025, n=1,006).
- **The cause is operational, not technical.** Gartner expects over 40% of agentic AI projects to be cancelled by end of 2027, citing escalating cost, unclear value, and inadequate risk controls, not model capability.
- **Regulation has made operations a board topic.** DORA, NIS2, the EU AI Act, and sector guidance hold the organization accountable for the operation of AI-built systems, regardless of who built them.
- **Most AI strategies have no operations chapter.** Use cases, models, data, skills, governance of the AI itself: covered. Who takes AI-built software live, who runs it at three in the morning, who answers the auditor: missing.

This paper supplies the missing chapter. Three parts of it are built to be lifted directly into your strategy document:

● THE EVIDENCE

Why operations is where AI programs stall. Assembled from Gartner, Forrester, BCG, S&P Global, MIT, the DevOps Research & Assessment program, and the European regulators themselves.

Sections 01 to 03.

● THE MATURITY MODEL

Five levels of AI-native operations, from **Backlog** to **Proven autonomy**. Locate your organization, pick a target level, and let the gap set the priority. **Section 04.**

● THE DRAFTED CHAPTER

A vendor-light operations chapter, written to be adapted into an enterprise AI strategy paper as it stands. **Section 07.**

Where this paper stands: published by nyxory, the company that operates the software your teams and their coding agents build. The evidence base is third-party throughout. Everything nyxory claims about itself is confined to the final section and labeled as ours.

The bottleneck has moved

For seventy years, the scarce resource in software was construction. That era is ending measurably: the most AI-native organizations already generate the majority of their new code with agents, and the build wave is arriving in every business unit. The delivery data, however, points the other way. The 2024 Accelerate State of DevOps report, published by the DevOps Research & Assessment program at Google, found AI tooling correlated with worsened software delivery performance for the second year running. Faster code generation without a faster path to production does not accelerate the business. It enlarges the pile of unshipped work.

The same research defines what fast looks like when the whole pipeline works: elite performers deploy on demand, with under one day from commit to production. Only about 19% of teams reach that level, and regulated enterprises skew heavily toward the clusters where lead time is measured in weeks and months. AI has industrialized the input side of a pipeline whose output side was already the constraint.

46%

of proof-of-concepts are scrapped before they reach production, at the average organization.

S&P GLOBAL · 2025

~95%

of enterprise GenAI pilots showed no measurable P&L impact. The cause: an integration gap, not model quality.

MIT · 2025 · DIRECTIONAL, 52 INTERVIEWS

19%

of teams reach elite delivery performance: on-demand deploys, under one day from commit to production.

DEVOPS RESEARCH & ASSESSMENT · 2024

The projected cancellation wave is **"a management problem wearing a technology costume."**

FORBES ON GARTNER'S AGENTIC-AI FORECAST · JUL 2026

For the strategy author, the reframe is this: the risk to your AI program is no longer that the technology cannot build what you need. It is that your organization has no industrial-grade answer for what happens after it is built.

The missing chapter

Why do strategy papers miss operations? Because the operating burden is back-loaded and the building excitement is front-loaded. Three well-established facts about the run side deserve a place in any AI strategy.

RUNNING IS MOST OF THE COST

The US GAO found roughly **80% of federal IT spend**, out of more than \$100 billion a year, goes to operating and maintaining existing systems. Industry benchmarks put maintenance at **50 to 80% of lifecycle cost**. Every AI-built application is a new entry on that ledger.

SELF-SERVICE WITHOUT GOVERNANCE FAILS

The low-code wave rehearsed this movie: **73% of organizations** planning citizen development had no governance rules for it (KPMG), and analysts attribute **30 to 50% of IT spend** to unauthorized shadow applications. AI-native building runs the same dynamic at far higher speed.

THE DEMAND IS NOT GOING AWAY

Gartner finds 17% of organizations have deployed agents, with over **60% planning to within two years**, and predicts **40% of enterprise applications** will embed task-specific agents by end of 2026. The build wave comes through the front door of every business unit.

An AI strategy that multiplies the number of applications by ten without changing the per-application cost of running them has written a cost commitment, not a productivity gain. And a strategy that answers the governance question with prohibition has picked the one policy that reliably fails against an afternoon-cheap capability: it drives the building underground, where the shadow portfolio grows without isolation, without records, and without an owner. Gartner's prescription from the low-code era, adaptive governance with explicit zones instead of blanket approval boards, applies with more force now that the builders include coding agents.

The only open question is whether the wave lands in a governed operating model or in shadow. That is a choice the strategy paper makes, explicitly or by omission.

For organizations that consider themselves late: the operations answer is useful before the AI wave and indispensable with it. An operating model that takes software live in days and keeps it accountable pays off on conventionally built software too. The operations chapter is the one part of an AI strategy that does not have to wait for the rest of the organization to become AI-native.

EXHIBIT 1 THE LIFECYCLE LEDGER OF ONE APPLICATION

BUILD

made cheap by AI

RUN • 50 TO 80% OF LIFECYCLE COST

operate, patch, secure, answer for it. The part the strategy must govern.

MAINTENANCE SHARE OF SOFTWARE LIFECYCLE COST • INDUSTRY BENCHMARKS (IEEE/ISBSG) • FEDERAL BASELINE: US GAO 25-107795

The regulatory floor is now explicit

For Europe's regulated industries, from banking and insurance to telecommunications and critical infrastructure, governed operations stopped being a preference and became contract law. Four dates define the floor:

● JAN 2025 · DORA IN FORCE

Under Regulation (EU) 2022/2554, the financial entity stays **fully responsible** for every ICT service it sources. Contracts for critical or important functions must grant **unrestricted audit rights**, control over sub-outsourcing, equivalent security standards, and a **documented, tested exit strategy** without lock-in (Art. 28, 30).

● NOV 2025 · DIRECT OVERSIGHT BEGINS

The European Supervisory Authorities designate the first **Critical ICT Third-Party Providers** under Art. 31: systemically important providers now sit under direct European supervision, with penalties up to 1% of average daily worldwide turnover per day.

● DEC 2025 · BAFIN: AI IS AN ICT ASSET

BaFin's guidance classifies AI systems as ICT assets under DORA, **across their full lifecycle including ongoing operation**. The AI-built application your business unit shipped last quarter is inside the resilience perimeter, whoever built it.

● AUG 2026 · AI ACT HIGH-RISK DUTIES

Human-oversight, data-governance, and logging obligations apply. The European Banking Authority confirmed these are **not absorbed** by existing financial regulation: they must be evidenced in parallel. The UK runs principles-based, but its Critical Third Parties regime can pull providers under direct oversight there too.

And the floor is not finance-only. NIS2 extends operational accountability, incident reporting, and management accountability to tens of thousands of essential and important entities across Europe, mid-sized manufacturers included. Whoever operates AI-built systems inherits the same question everywhere: can you show, continuously, that they run in control?

The strategic implication is sharper than "compliance matters." It is that **evidence must be a by-product of normal operation, not a project you run before each audit**. A telco-sector study makes the gap concrete:

72%

of communication service providers believe their AI is trustworthy.

TM FORUM × IBM IBV · JUN 2026

14%

can produce the evidence to prove it. That 58-point gap is what auditors, supervisors, and boards will ask about.

SAME STUDY · 130 OPERATORS

A note on names: "DORA" designates both the EU Digital Operational Resilience Act (the regulation above) and Google's DevOps Research & Assessment program (the delivery metrics in Section 01). This paper flags each use.

The AI-Native Operations Maturity Model

The model measures the last mile, from "it works" to "it runs, governed." It deliberately does not measure model quality, data readiness, or use-case selection; existing frameworks cover those. It extends the logic of the CNCF Platform Engineering Maturity Model and the DevOps Research & Assessment performance clusters into the era where software arrives faster than platform teams were ever sized for.

EXHIBIT 2 THE AI-NATIVE OPERATIONS MATURITY MODEL

LEVEL	WHO CAN TAKE SOFTWARE LIVE	LEAD TIME	WHO RUNS IT	WHAT THE AUDITOR GETS
L1 Backlog	Central IT only, via ticket queue	Months	IT, reactively	A retrospective document hunt
L2 Paved road	Professional dev teams on the internal platform	Weeks	Dev teams on call	Pipeline logs, for the platform's tenants only
L3 Governed self-service	Any team, and their coding agents, inside an IT-defined envelope	Days	Teams, with platform support	Policy checks at deploy time
L4 Operated portfolio	Any team, same envelope	Days	The operations layer itself: deploys, fixes, patches. Humans sign off.	A continuous, per-application operations record
L5 Proven autonomy	Any team, same envelope	Days	The operations layer, with portfolio economics managed	Regulator-shaped evidence on demand; exit tested

A one-question self-assessment: for the last application a business unit wanted, how many calendar days passed between "it works" and "it runs in production, with a record"? And who carried it in between? If the answer is measured in months and the carrier was a queue, the organization is at level 1 or 2, and every euro spent accelerating the build side is widening the gap.

AI-Native Operations Maturity Model · first published by nyxory, July 2026. Cite freely with attribution.

Reading the levels

1

Backlog

Building got fast; everything after "works on my machine" is a ticket. This is where prototypes die and shadow applications breed. The abandonment numbers in Section 01 and the shadow-IT spend in Section 02 are portraits of organizations held at this level while demand grows.

2

Paved road

A platform team gives professional developers golden paths. Real progress, and the ceiling of most current platform-engineering investment. But it was sized for the dev org, not for every business unit with a coding agent. Software built outside the paved road falls back to level 1.

3

Governed self-service

The envelope inverts: instead of IT executing every launch, **IT defines the boundaries**. Isolation between projects, policy checks, spend ceilings, sign-off points. Anything built inside them goes live in days. This is adaptive governance, generalized from citizen developers to coding agents.

4

Operated portfolio

The step most strategies never name. At level 3 the organization still carries the run: someone patches, someone answers the pager, someone owns the growing portfolio. At level 4, **running the software is itself delegated to the operations layer**. It ships, fixes, and patches; every action lands on a per-application record; humans move from executing operations to approving them. This is the level at which ten times the software stops implying ten times the operating effort.

5

Proven autonomy

The autonomy of level 4, now with proof: operations evidence arrives in the shape the regulator asks for. The per-application record maps to the DORA register of information, the AI Act's oversight and logging duties are evidenced by the normal operating record, and the exit path is a tested procedure, not a contractual promise. The trust gap from Section 03 closes: **"we can prove it" becomes the default state**.

Where to aim: level 3 is the ambition of most platform initiatives. The strategy question is whether to stop there. Level 3 governs the launch; level 4 takes over the run. For a portfolio that grows by tens of applications a year, the difference compounds. And the levels are not a ladder every organization must climb itself: without a platform team, buying level 4 as a service is regularly more economical than building the level-2 paved road first.

Sovereignty, stated precisely

For DACH and public-sector organizations, sovereignty belongs in the operations chapter. As a precise requirement, not a slogan. The demand is real:

88%

of German companies say an AI provider's country of origin matters.

BITKOM · 2026 · N=604

93%

would prefer a German provider, ahead of the USA at 51%.

BITKOM · 2026

2027

the year European sovereign-cloud infrastructure spend is forecast to overtake North America's.

GARTNER · FEB 2026 · FORECAST

But "sovereign" is a tiered claim, and a strategy paper should say which tier it requires:

● DATA RESIDENCY

Where the data sits. The most common tier, and the easiest to satisfy. A data-center pin on a map.

● OPERATIONAL SOVEREIGNTY

Who runs the system: under which jurisdiction the operating decisions fall, and who can be compelled to act on the infrastructure. The tier that matters most in the AI era, and the one most often skipped.

● LEGAL - CORPORATE

Under whose law the provider itself exists, and whose courts settle the dispute.

When software is operated by an automated layer, the sovereignty question follows the operator: whose systems act on your production environment, and under whose jurisdiction. An AI strategy for a regulated European organization should require the **operating layer**, not just the storage layer, to sit inside its compliance perimeter. Stated that way, sovereignty is a procurement criterion any vendor can be tested against, and cannot satisfy with a data-center pin on a map.



OPERATED IN EUROPE · THE OPERATIONAL TIER, INSIDE THE PERIMETER · NYXORY.COM

What to require from an operations layer

Whether the answer is an internal build, a service provider, or a product, the requirements are the same. A strategy paper can adopt this list as its procurement baseline. It is written to be testable.

1 Accepts software from any source

Business units, IT, external partners, and their coding agents. One operating model for all of it, or the shadow portfolio returns.

2 Governance as an envelope, not a queue

IT defines isolation, policies, spend ceilings, and sign-off points once. Everything inside goes live without a per-application project.

3 Operations taken over, not tooled up

The layer itself deploys, fixes, and patches. The measure is who carries the pager, not who owns the dashboard.

4 Every action on the record, per application

A continuous operations record that feeds the DORA register of information and supports the AI Act's logging duties, as a by-product of normal running.

5 Humans approve; the layer executes

Sign-off points are explicit and demonstrable. This is what the AI Act's human-oversight duty looks like in operations.

6 Cost governed per application

Spend ceilings and portfolio-level run-cost visibility, so the strategy's economics survive contact with year two.

7 Sovereignty at the operational tier

The operating layer, not just the data, inside the required jurisdiction. See Section 05.

8 Exit tested, not promised

Standard artifacts, documented exit procedure, no lock-in. The Art. 28(8) requirement, and the honest answer to concentration risk.

How to use this list: score every candidate, including the internal build, from 0 to 2 on each requirement, in writing. Requirements 3, 4 and 8 are the ones that separate an operations layer from deployment tooling: who carries the run, what lands on the record, and how you leave. The scoring sheet belongs in the strategy paper's appendix.

The chapter, ready to adapt

The text below is written to be lifted into an enterprise AI strategy paper and adapted. It is deliberately vendor-light; the requirements do the selecting.

DRAFT CHAPTER • ADAPT FREELY

Operations: from AI output to running systems

Our AI strategy will multiply the amount of software this organization creates. This chapter governs what happens after creation, because that is where AI programs fail: industry data shows abandonment concentrated between proof-of-concept and production, and our regulators (DORA or NIS2, the EU AI Act, and sector guidance) hold us accountable for the operation of AI-built systems regardless of who built them.

We therefore treat operations as a first-class element of the strategy, with a target maturity rather than a tool choice. Today, software built outside central IT reaches production through a ticket queue measured in months, and each new application adds permanent run cost. Our target state is an **operated portfolio** (level 4 of the AI-Native Operations Maturity Model, nyxory 2026): business units and their coding agents build; a governed operations layer takes their projects live within days, inside an envelope IT defines through isolation, policy, spend ceilings, and sign-off; the layer then runs them, with every deployment, fix, and patch on a per-application record. Humans approve; the layer executes. Evidence for auditors and supervisors is a by-product of normal operation, not a periodic project.

We will require of any operations layer, internal or external: acceptance of software from any source; governance as an envelope rather than a queue; genuine takeover of operations, not tooling; a continuous per-application record satisfying our DORA and AI Act evidence duties; explicit human sign-off points; per-application cost ceilings; sovereignty at the operational tier, not merely data residency; and a tested exit. Reference implementations of this model exist as managed services on European infrastructure, nyxory among them; the build-versus-buy decision follows our standard criteria, with time-to-value weighted for the pace of the AI program this strategy commits us to.

Adapting it is three moves:

● MOVE 1 • LOCATE

Name your organization's current level, using the one-question self-assessment in Section 04. Write the number down; it disciplines everything that follows.

● MOVE 2 • TARGET

Set the target level and the date. **Level 4 is the point where the portfolio scales without scaling the operating burden;** stopping at level 3 keeps the run on your books.

● MOVE 3 • ANCHOR

Adopt the eight requirements from Section 06 as the procurement baseline, applied to internal and external answers alike.

How nyxory implements this

Everything above stands on third-party evidence. This section is ours.

nyxory is the operated-portfolio level, delivered as a service: **you build it, nyxory runs it**. Your business units, your IT, or your partners build software with the coding agents they already use; nyxory takes those projects live and operates them: deployments, fixes, patches. Your IT keeps control of the envelope: policies, sign-off, spend ceilings, and audit answers. Projects run isolated from each other on European infrastructure, cloud-independent by design. The path from idea to production is measured in days.

Judge the model against your own application.

Bring one real project to the Idea-to-Production Workshop and leave with it live, operated, and on the record. Then decide what level 4 is worth to your strategy.

[NYXORY.COM](https://nyxory.com)

ABOUT THE AUTHOR



Tobias Nentwig

Founder & CEO, nyxory. Writes about the operations side of AI-native software: what it takes to run what agents build, governed and on the record.

TOBIAS.NENTWIG@NYXORY.COM · [NYXORY.COM](https://nyxory.com)

SOURCES AND READING NOTES

S&P Global Market Intelligence · Voice of the Enterprise: AI & ML, Oct 2025, n=1,006. Abandonment and PoC figures.

Gartner · Press releases 25 Jun 2025 and 26 Aug 2025; 2026 Hype Cycle for Agentic AI; sovereign-cloud IaaS forecast, Feb 2026.

MIT NANDA · The GenAI Divide, 2025. 52 interviews, directional; never used here without the larger-sample S&P finding beside it.

DevOps Research & Assessment (Google) · 2024 Accelerate State of DevOps. Performance clusters; AI-tooling correlation.

US GAO · GAO-25-107795, Jul 2025. Operations and maintenance share of federal IT spend.

KPMG · EMA citizen-development survey (715 firms). Governance-rule gap.

Bitkom · Künstliche Intelligenz in Deutschland, 2026, n=604. Provider-origin preferences.

TM Forum × IBM Institute for Business Value · AI trust study, Jun 2026, 130 operators.

Regulatory · Regulation (EU) 2022/2554 (DORA), Arts 28 to 31 · Directive (EU) 2022/2555 (NIS2) · ESA CTPP designations, 18 Nov 2025 · BaFin Orientierungshilfe on AI as ICT risk, 18 Dec 2025 · EBA AI Act mapping, Nov 2025 · FCA AI Update and the FSMA 2023 Critical Third Parties regime.

Forward-looking statements · Gartner's cancellation and sovereign-cloud figures and Deloitte's workload projections are forecasts by their authors, flagged as such where cited.

Suggested citation: nyxory (2026): The Last Mile of AI-Native Software. The operations chapter your AI strategy is missing. nyxory.com

nyxory is an AI-agent company: your teams and their coding agents build software, nyxory takes it live and operates it on European infrastructure, governed by your IT. Individual capabilities are still in development; we are happy to go deeper in conversation.

© 2026 nyxory GmbH. All rights reserved. This paper is provided for general information; it does not constitute legal or regulatory advice.

NYXORY

You build it. nyxory runs it.

Your teams and their coding agents build software. nyxory takes it live and operates it on European infrastructure, governed by your IT. Bring one real project to the Idea-to-Production Workshop and judge the maturity model against your own application.

NYXORY.COM · TOBIAS.NENTWIG@NYXORY.COM