

From Pilot to Production: Unifying AI and Sovereignty at Enterprise Scale

Written by:

Deepika Giri

Vice President, AI Research, IDC Asia/Pacific

Published:

September 2026

By unifying AI development, deployment, and governance on a single sovereign-ready platform, enterprises can move generative AI from stalled pilots to trusted production value at scale.

Introduction

Almost every large enterprise has now experimented with generative AI (GenAI). Very few have scaled it. That gap — between a compelling proof of concept and a system that runs reliably in production, under governance, delivering measurable value — is the defining AI challenge of the next 24 months.

The investment signal is unambiguous. IDC expects worldwide AI spending to reach approximately US\$632 billion by 2028, growing at close to 29% a year, with generative AI the fastest-moving component at a compound annual growth rate (CAGR) of nearly 60%. In Asia/Pacific (AP), AI spending is on track to reach about US\$175 billion by 2028. Budgets, executive attention, and pilots are all in place. What is scarce is production.

The reason most organizations stall is not a shortage of models or ambition. It is the absence of a unified operating model for AI: a way to develop, deploy, govern, and pay for AI on a common foundation, with the data, controls, and sovereignty guarantees that regulated enterprises require. This paper examines where enterprises are on that journey, where value is landing first, why sovereignty has moved from an IT footnote to a board mandate, and what it takes to move from pilot to production at scale. It then profiles the role a unified AI platform can play in closing that gap.

Situation overview

The adoption paradox: broad experimentation, narrow production

Adoption breadth is no longer the story. Enterprises across industries and geographies have run generative AI proofs of concept; many have committed budget. The harder truth sits underneath the headline: most of that activity is stuck in experimentation and pilot stages rather than production. IDC's Asia/Pacific Data, AI and GenAI Survey, 2025 quantifies the gap — on average, organizations report that only about 34% of their GenAI projects moved from pilot to production over the past 12 months, and of those that reached production, just around 37% achieved the key performance indicators (KPIs) their sponsors expected. Nearly half of respondents say a quarter or fewer of their production projects hit target KPIs. The maturity curve — experimentation, pilot, production — has widened at the top and stayed narrow at the point where value is actually realized.

At a glance

Key stats

According to IDC research:

- Worldwide AI spending is on track to reach around US\$632 billion by 2028 (about 29% CAGR); generative AI is the fastest-growing slice (about 59% CAGR).
- In Asia/Pacific, only about 34% of GenAI projects moved from pilot to production in the past year. Of those, just around 37% hit their target KPIs.
- Around 90% of AP enterprises are already using, planning, or actively exploring sovereign AI; about 80% run AI/GenAI workloads under data-residency requirements.
- The leading barriers to scaling AI are data readiness (57%), infrastructure and integration (55%), regulatory and compliance concerns (54%), and cost (53%).

What's important

AI value is shifting from horizontal capability to industry-specific outcomes, with banking, software, and retail leading. Sovereignty has moved from an IT footnote to a board mandate, and the ROI conversation is broadening from a single return metric to a portfolio view of AI business value.

Key takeaways

- A unified operating model, with common data, governance, cost discipline, and sovereignty built in from the start, is what separates enterprises that reach production from those stuck in pilots.
- Judge platforms on production evidence, not demos.
- Make sovereignty and responsible AI integral to their design rather than bolt them on at a later stage.

Why pilots stall: an operating-model problem, not only a technology one

The pattern behind stalled pilots is consistent. Models that perform in a controlled demonstration degrade in production, where data is fragmented across systems that were never designed to interoperate, where governance and cost controls are missing, and where the surrounding operating model — integration, security, monitoring, human oversight — was never built. This is less an AI capability problem than an AI operating-model problem: the absence of a unified data layer, governance framework, and cost discipline that make AI trustworthy and affordable at enterprise scale.

The case for a unified AI platform

Enterprises are responding by consolidating fragmented tools, environments, and point solutions onto unified platforms that standardize how AI is built and run. A unified platform is intended to shorten the path from model to production by providing common data pipelines, pre-built components, guided workflows, integrated MLOps, and governance and cost controls in one place. Rather than assembling and integrating a dozen services, teams work from a consistent foundation, which is precisely the layer where the pilot-to-production gap is won or lost.

Challenges

When IDC asked Asia/Pacific enterprises what most impedes scaling AI/GenAI, no single blocker dominates — the difficulty is that several bind at once (see Figure 1). Data readiness leads, cited by 57% among their top five barriers, followed closely by technology and infrastructure limitations, regulatory and compliance concerns, and cost.

Figure 1: Top barriers to scaling AI/GenAI in Asia/Pacific

Barrier to scaling AI/GenAI	% citing (top 5)
Data readiness (quality, availability, governance, bias)	57%
Technology and infrastructure limitations (integration, IT complexity)	55%
Regulatory, compliance, and governance concerns	54%
High cost and budget constraints	53%
Talent and skills gaps	50%
Trust and transparency concerns	49%
Change management and adoption issues	46%
ROI uncertainty	46%

Source: IDC Asia/Pacific Data, AI and GenAI Survey, 2025, n = 540 (respondents selected up to five barriers).

- Pilots stall between proof of concept and production because the surrounding operating model which includes data, integration, governance, and cost control, is missing.
- Cost unpredictability at scale, driven by graphics processing unit (GPU) consumption and inference, makes it difficult for finance leaders to underwrite AI programs with confidence.
- A widening AI skills gap limits how quickly enterprises can build, deploy, and operate AI safely; the upskilling most in demand is business-focused training on use cases and ROI (66%), followed by governance and responsible-AI training (57%).
- Fragmented data and pre-existing technology investments resist rip-and-replace, so enterprises need modular, open systems that integrate with what they already run.
- Tightening regulation and geopolitical uncertainty raise the bar on data sovereignty, compliance, and responsible AI, which are now board-level concerns.

What's important

From horizontal capability to vertical value

The conversation is shifting from horizontal capability (“We have a chatbot”) to vertical value delivered inside specific industries. In IDC’s spending data, banking sits among the top 3 AI-investing industries alongside software/information services and retail, together accounting for close to 38% of global AI spend. The sectors moving fastest from experimentation to measurable value tend to share two traits: relatively clean, well-governed data and clear, repeatable processes. Localization also matters more than it appears; in IDC’s AP survey, about 88% of enterprises call multi-language and localization support important or critical for their GenAI applications, with roughly a third rating it critical. In linguistically diverse markets, local-language models are a gateway to adoption rather than a nice-to-have. Agentic AI is the next wave of vertical value: AP enterprises are prioritizing IT (56%), security (33%), and software development (28%) for agentic automation, chiefly to gain productivity and unlock new business models.

Sovereign AI moves to the board

Sovereignty has moved from an IT footnote to a board mandate with unusual speed. IDC’s *Asia/Pacific Emerging AI Technologies Survey, 2025* shows how mainstream it has become: roughly 90% of enterprises are already using, planning, or actively exploring sovereign AI, with about 29% already using it, and around 80% run AI/GenAI workloads carrying data-residency or sovereignty requirements (28% for all data, 53% for sensitive categories only). The leading drivers are ensuring data residency and regulatory compliance (39%) and building trust through transparent AI governance (38%). Consistent with this, digital sovereignty is now the single most-cited regulatory factor shaping AP AI initiatives (59%), and 43% of enterprises name adopting sovereign AI as a responsible-AI control, emphasizing that sovereignty and responsible AI are two sides of one design choice, not separate compliance checkboxes. This also shows up in where workloads run: about 44% of AP enterprises keep AI workloads in private or dedicated environments rather than public cloud.

IDC’s forward view reinforces the direction. By 2027, roughly a third of governments are expected to mandate sovereign AI architectures for sensitive sectors; by 2029, about 70% of large enterprises are expected to move AI workloads into private, isolated cloud environments to keep data away from hyperscale models. Regionally, sovereign AI has risen from the seventh- to the second-highest government investment priority in a single year.

The ROI conversation is being redefined

As programs scale, boards are pressing for returns, and they are materializing. In IDC’s AP survey, more than nine in ten enterprises report positive ROI on AI/GenAI spend, with the most common return around 4x per dollar invested and roughly 60% reporting 4x or more. But enterprises measure that value across many dimensions — cost savings and efficiency, revenue, customer experience, risk reduction, and innovation — which is why IDC’s guidance is to shift from a narrow return-on-investment lens to a broader AI business value view. Value is nonlinear, data-dependent, and cross-functional, and total cost of ownership extends well beyond models and infrastructure to governance, retraining, and ongoing inference. Notably, when choosing how to pay for a unified AI platform, AP buyers lean toward predictable models. For instance, fixed subscription/SaaS (49%) and enterprise licensing (47%) rank ahead of pure consumption-based pricing (35%), which is a clear signal that cost predictability, not simply low cost, is what finance leaders want. The more durable approach is to manage a portfolio of use-case ROI cases across business functions, backed by predictable economics, rather than chase a single technology-ROI number.

Role of Tata Communications in enabling this transformation

Tata Communications helps enterprises operationalize AI at scale by providing a unified, sovereign AI foundation that brings together infrastructure, data, governance, model development, deployment, and lifecycle management. Rather than requiring organizations to

assemble multiple vendors and fragmented toolchains, it aims to provide a common operating model for AI, helping enterprises accelerate from experimentation to production with greater control, predictability, and compliance.

A unified platform for production-grade AI

Through Tata Communications Vayu AI Cloud, spanning GPU-as-a-Service and AI Studio, enterprises gain access to sovereign-ready GPU infrastructure, integrated AI development and MLOps, responsible-AI controls, and open integration with existing enterprise systems. This lets organisations focus on business outcomes and industry-specific use cases rather than managing the underlying complexity of AI infrastructure and operations. AI Studio, the platform layer, brings model choice, development, and lifecycle management onto that common foundation, with a Model Hub of benchmarked models, integrated MLOps, and responsible-AI tooling designed to shorten the path from model selection to a governed, production-grade endpoint.

Predictable economics and open integration

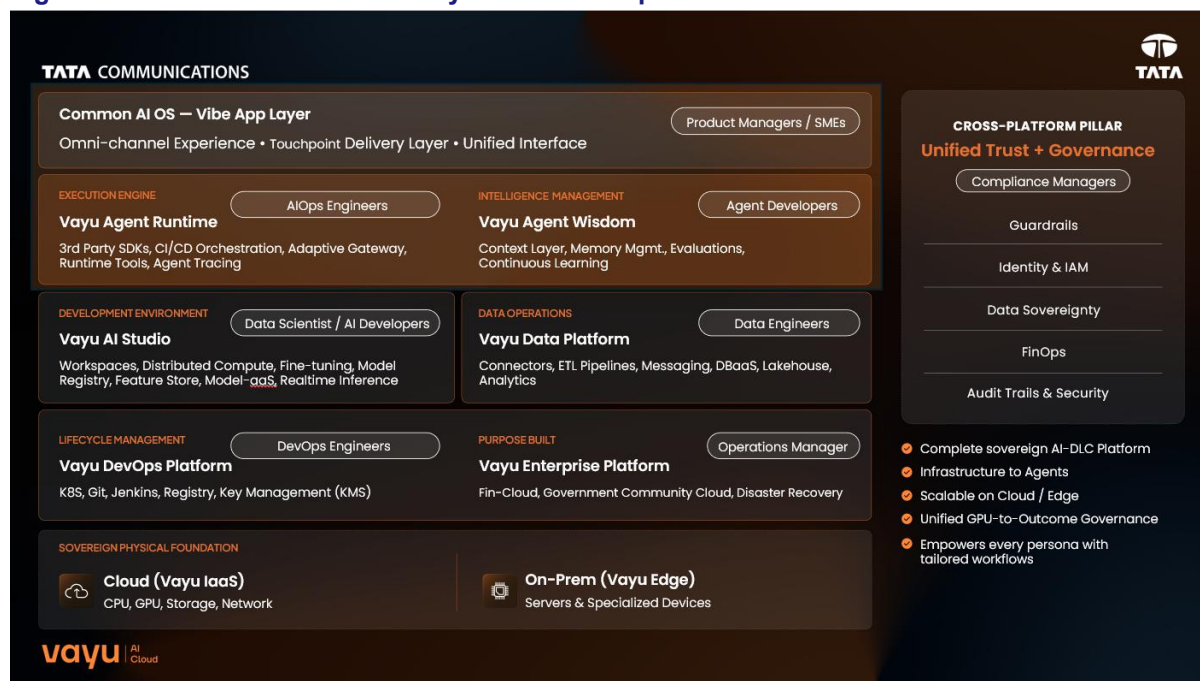
Cost predictability is a central design goal. Tata Communications positions Vayu AI Cloud around predictable pricing and improved total cost of ownership; the company cites figures such as up to 30% lower TCO and materially faster storage throughput, intended to give finance leaders more confidence when underwriting AI at scale. (These are vendor-provided figures and should be validated by prospective buyers against their own workloads.) Because most enterprises arrive with substantial existing technology investments, the platform is designed to be modular and to integrate with systems already in place rather than requiring a wholesale replacement.

Sovereign by design

Tata Communications Vayu AI Cloud is positioned as sovereign-by-design, with GPU and platform capacity hosted in-country and the data-residency, control, and governance characteristics that regulated sectors such as BFSI and government increasingly require. Combined with the platform's responsible-AI controls, this is intended to help enterprises pursue data-sovereignty and responsible-AI objectives on a single, centrally governed foundation rather than stitching together multiple external services, reducing the compliance surface area that fragmentation creates.

Figure 2 and Table 1 illustrate, at a high level, the platform's architecture and a representative adoption path.

Figure 2: Tata Communications Vayu AI Cloud — platform overview



Source: Tata Communications, 2026

Table 1: A representative phased path from foundation to autonomous value

Phase 1 (0– 6 months) Foundation	Phase 2 (6–18 months) Scale and integrate	Phase 3 (18–36 months) Autonomy and value
Consolidate data and access; stand up a governed environment; run priority use cases as production pilots.	Integrate with existing systems; industrialize MLOps; extend responsible-AI and sovereignty controls across workloads.	Introduce agentic workflows under human oversight; manage a use-case ROI portfolio; scale value across functions.

Note: Illustrative; timelines vary by enterprise maturity. Source: IDC, 2026. To be validated with Tata Communications.

IDC perspective

The enterprises that will separate from the pack over the next two years are not those with the most pilots, but those that build the operating model, which includes unified data, governance, cost discipline, and sovereignty, that lets AI run reliably in production. IDC views the consolidation onto unified AI platforms as a sound response to the pilot-to-production gap, provided the platform genuinely reduces integration effort rather than adding another layer to manage.

Tata Communications' approach is notable on two fronts. First, bringing sovereign-ready infrastructure (GPU-as-a-Service) and the AI Studio platform together under Vayu AI Cloud addresses both the compute-cost and the operating-model dimensions of the problem, which are usually solved separately. Second, anchoring the offering in sovereign-by-design, in-country infrastructure aligns with where regulated buyers in Asia/Pacific are heading: IDC's survey shows those buyers want domain data sets held within sovereign boundaries (37%), integrated responsible and explainable AI (36%), and security architectures that guard against data exfiltration (36%), capabilities a sovereign-by-design platform is well placed to address. The open question for any platform in this category is the same: can it demonstrate, with reference customers and independent evidence, that it moves organizations past the pilot wall at predictable cost, and not just that it can run a compelling proof of concept. To the extent Tata Communications can evidence that outcome, it is positioned against a large and fast-growing opportunity.

Challenges faced by customers

- **Integration with pre-existing investments.** Enterprises rarely adopt a platform on a clean slate. Integrating a unified AI platform with established data estates, security tooling, and line-of-business systems is real work, and buyers should scope this effort explicitly rather than assume it away. The value of modularity and open integration is only realized if the connectors and effort required match the enterprise's actual environment.
- **Avoiding a new form of lock-in.** Consolidating onto any single platform trades fragmentation for concentration. Buyers should confirm how portable their models, data, and pipelines remain, and how the platform interoperates with other clouds and tools, so that standardization does not harden into dependency.
- **Skills to operate the platform.** Guided workflows and pre-built components lower the barrier to entry, but operating AI responsibly at scale still requires capability in data engineering, MLOps, and governance. Enterprises should plan for the skills and change management needed to run the platform well, and treat workforce readiness as part of the program rather than an afterthought.
- **Proving ROI at scale.** Vendor-cited cost and performance figures are a starting point, not a guarantee. Buyers should validate cost predictability and business value against their own workloads and use cases, ideally through a time-boxed production pilot with agreed KPIs before enterprisewide commitment.

Guidance for technology buyers

- **Fix the foundation first.** Resolve the data foundation before scaling AI. Fragmented, poorly governed data is the most common reason pilots fail in production; assess data readiness before committing to enterprise-scale programs.
- **Buy for production.** Choose platforms on production evidence, not demo performance. Ask specifically how a partner has solved the scaling problem involving integration, governance, and cost, and not just the pilot problem.
- **Underwrite the economics.** Insist on cost transparency. Require predictable pricing models and validate GPU and inference economics against your own workloads before you commit.
- **Design in sovereignty.** Make sovereignty and responsible AI architectural, not bolt-on. Confirm data residency, control, and governance are designed into the platform and encoded in policy, not managed through external procedures.

Guidance for business buyers

- **Lead with outcomes.** Anchor programs to business outcomes, not technology milestones. Define the KPI improvements the business needs first, then make them the basis of every technology and partner decision.
- **Portfolio, not project.** Manage AI as a value portfolio. Adopt an AI business value view across benefit categories and business functions rather than chasing a single ROI number.
- **Govern from the board.** Own the sovereignty and trust agenda at the top. Treat data sovereignty, compliance, and responsible AI as board-level commitments, since they increasingly determine where and how AI can be deployed.
- **Fund the people side.** Treat the workforce transition as a deliverable. Budget for skills and change management alongside the technology so adoption keeps pace with capability.

Conclusion

The move from experimentation to production-grade, governed, sovereign AI is not a distant aspiration — the building blocks exist today. What separates leaders from followers is whether they build the operating model that makes AI reliable, affordable, and compliant at scale: unified data and development, cost discipline, responsible-AI governance, and sovereignty designed in from the start.

Unified AI platforms are a credible response to that need, and the market opportunity is large and growing quickly. IDC believes the enterprises that invest now in this foundation will build an advantage that compounds. To the extent that Tata Communications can address the challenges described in this paper, and evidence, with customers and independent proof, that its platform moves organizations past the pilot wall at predictable cost, the company has a significant opportunity for success.

About the IDC analyst



Deepika Giri

Vice President, AI Research, IDC Asia/Pacific

Deepika Giri leads IDC's Asia/Pacific research on AI platforms and strategies. Her coverage centers on the trends shaping how AI — and GenAI and agentic AI in particular — is adopted across industries, and on the interplay between the vendor ecosystem and technology buyers.

[More about Deepika](#)

Message from the sponsor

Building the Foundations for Enterprise AI at Scale

As AI adoption matures, organizations are increasingly focused on the operational challenges of scaling AI across the enterprise. Beyond model selection, success depends on factors such as data readiness, governance, infrastructure, security, compliance, sovereignty requirements, and life-cycle management.

Many organizations are evaluating how to establish a consistent operating model that enables experimentation while supporting broader deployment, oversight, and business accountability. These considerations are becoming increasingly important as AI initiatives expand across business functions and use cases, particularly in industries with heightened data residency, regulatory, and governance requirements.

Tata Communications supports enterprises in this journey through Tata Communications Vayu AI Cloud infrastructure, platform services, and governance capabilities designed to help organizations build, deploy, and manage AI workloads in a controlled, compliant, and sovereign-ready manner.

Learn more about Tata Communications' AI capabilities at [AI Cloud Solutions: Build, Train & Deploy at Scale](#)

IDC Custom Solutions

IDC Custom Solutions produced this publication. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis that IDC independently conducted and published, unless specific vendor sponsorship is noted. IDC Custom Solutions makes IDC content available in a wide range of formats for distribution by various companies.

This IDC material is licensed for [external use](#), and in no way does the use or publication of IDC research indicate IDC's endorsement of the sponsor's or licensee's products or strategies.



IDC Research, Inc.

One Beacon Street, Suite 33100, Boston, MA 02108, USA

T +1 508 872 8200

blogs.idc.com | LinkedIn [@IDC](#) | www.idc.com

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications, and consumer technology markets. With more than 1,300 analysts worldwide, IDC offers global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries. IDC's analysis and insight help IT professionals, business executives, and the investment community make fact-based technology decisions and achieve their key business objectives.

©2026 IDC. Reproduction is forbidden unless authorized. All rights reserved. [CCPA](#)