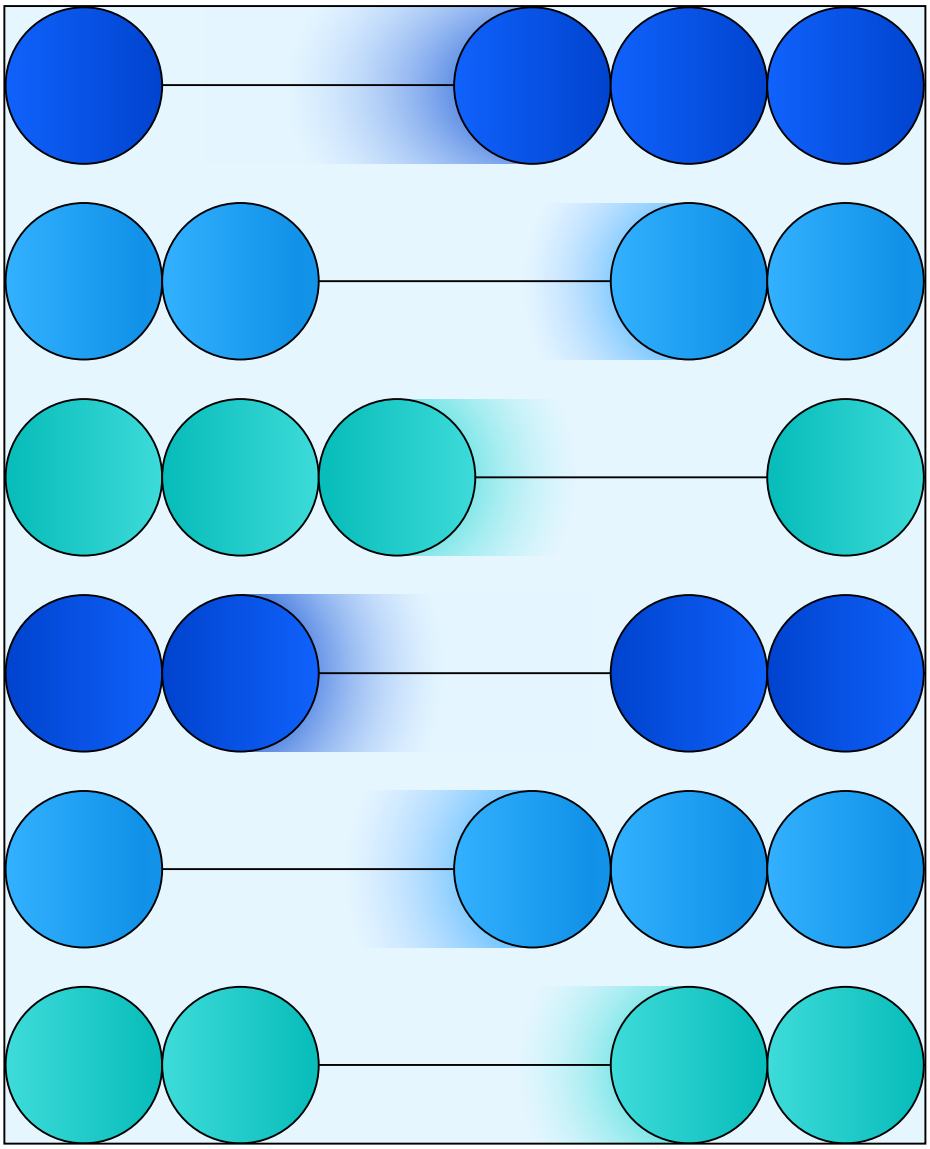


The calculus of AI sovereignty

Balancing control, flexibility, and risk

IBM

Contents

Foreword

Don’t let dependence define your AI destiny

Ana Paula de Jesus Assis

Senior Vice President and Chair, Europe, Middle East, Africa, and Asia Pacific
IBM

AI is reshaping the foundation of business—how decisions are made, how work flows, and how value is created. But as AI moves deeper into core operations, a new strategic question is emerging for leaders in every sector: Who is really in control?

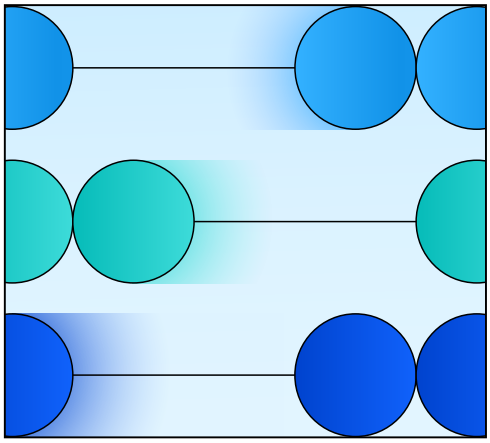
The answer isn’t straightforward. AI has introduced new forms of dependency that evolve faster than traditional governance, procurement, or technology cycles were designed to handle. That is why AI sovereignty has become one of the defining leadership issues of this moment.

The stakes are no longer technical; they are economic. As AI takes on operational decision-making, any loss of control can translate directly into margin pressure, compliance exposure, or outright business disruption. Yet this IBM Institute for Business Value (IBM IBV) report shows that only a small fraction of executives today truly understand their AI dependencies. The gap between adoption and control is widening at the very moment AI is becoming indispensable.

Sovereignty means having the ability to see your dependencies clearly, govern them consistently, and change direction when needed without breaking the business. It is the difference between relying on AI and depending on it. And what leaders increasingly recognize is that they cannot achieve this level of control without openness at the foundation.

Open source plays a pivotal role in this transition. Openness creates the optionality enterprises need to maintain leverage—access to interoperable models, transparent architectures, and ecosystems where no single vendor controls the pace or direction of innovation. Open models and open tooling make it possible to retrain, tune, or swap components without recreating entire systems. Open standards ensure that data can move across environments and that governance can follow. And open communities accelerate innovation at global scale, reducing the concentration risk that comes from depending too heavily on any one provider.

For enterprises, this combination of openness and control is what enables selective sovereignty—gaining the right level of authority where it matters most, without carrying the cost of total independence everywhere. It transforms sovereignty from a defensive posture into a strategic advantage: stronger negotiating power, more resilient operations, and the freedom to innovate without lock-in. This report provides the roadmap for doing exactly that—building AI systems that are flexible by design, resilient under pressure, and sovereign where it counts.



Key takeaways

Selective AI sovereignty lets organizations control AI where it matters most—so they can manage risk, protect operating margins, and quickly adapt when conditions change.

For most organizations, the AI estate is opaque.

Only 9% of executives say they have an excellent understanding of their dependencies on AI vendors, models, and infrastructure.

Executives want flexibility, but few have it.

71% say switching their primary AI vendor or model would be difficult if they had to do it today.

Leaders are willing to pay for flexibility—because it's worth it.

72% would accept a 20% cost increase to maintain multiple AI vendors for strategic flexibility.

Selective AI sovereignty protects profitability.

Organizations that can apply selective control across data, models, and infrastructure protect 55% more of their operating profit from AI-driven disruption.

Introduction

AI sovereignty isn't owning the stack. It's staying in control when conditions change.

Enterprises are leveling up their AI adoption. They've moved beyond isolated AI assistants to create connected networks of AI agents that reason, decide, and act on their behalf. In early 2026, CEOs said 25% of operational decisions were being made by AI. By 2030, they expect that to rise to 48%.[1]

Using agentic AI across operations could amplify productivity quickly—but wielding that power effectively requires stronger controls. That's where AI sovereignty comes in. AI sovereignty preserves the organization's ability to move data, swap models, and shift workloads across hybrid cloud environments when AI dependencies shift, whether those shifts are technical, commercial, or regulatory. It's what allows organizations to maintain AI-powered operations when conditions change, which is crucial as AI becomes more deeply embedded in how work gets done.

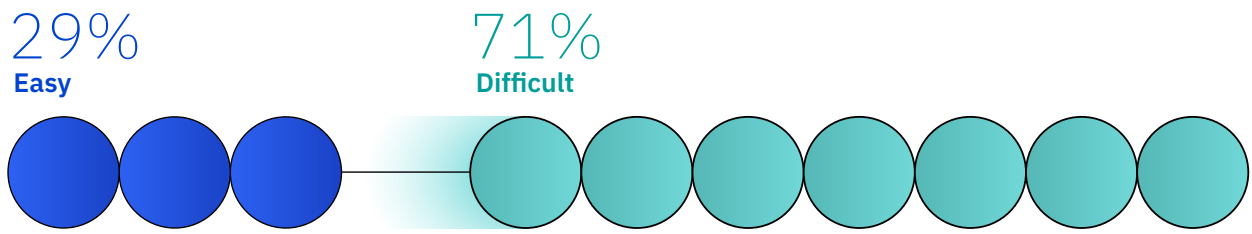
Yet most leaders don't have that control today. Only 9% of the 1,000 global executives surveyed in a new IBM IBV study, conducted with Oxford Economics, say they have an excellent understanding of their AI dependencies. 71% say switching their primary AI vendor or model would be difficult if required today (see Figure 1). "Vendor lock-in creates imbalance," says Conor Mlacak, CIO, Staples Canada. "Once you're locked in, you lose leverage."

Our research has quantified the value of that leverage. The organizations in our study that have the greatest control across their AI stack protect 55% more operating profit from AI-driven disruption than organizations with less control (see "Research methodology" on page 29).

FIGURE 1

AI lock-in is real

71% of executives say it would be difficult to switch AI vendors or models if they had to today.



Achieving this level of control increasingly depends on the ability to coordinate these elements as a system—applying policies, monitoring behavior, and orchestrating decisions consistently across a fragmented AI landscape. This is different from cloud sovereignty or multi-cloud management because AI introduces a new dependency: the AI model layer.

At the same time, evolving regulations across the US, EU, and Asia are redefining what organizations can and cannot do with AI. Many regulators are already signaling stricter expectations around data residency, model transparency, and AI accountability. These shifts will increasingly determine where data can travel, which models can be used, and how quickly organizations must respond to new compliance requirements.

But AI sovereignty isn't just a defensive play. "There is a widespread assumption that AI governance is limited to regulation, which makes it reactive. I believe governance must instead be proactive," says Raúl Cals, CIO, Unicaja. It can create upside by enabling faster decisions, better negotiating leverage, and more resilient growth in an unpredictable AI landscape.

Organizations that build AI sovereignty effectively can:

- **Increase resilience** by maintaining continuity even as dependencies change
- **Accelerate AI-driven growth and innovation** by adopting and integrating new capabilities faster
- **Enforce cost discipline** by avoiding lock-in and optimizing where and how AI runs
- **Expand market access** by meeting regulatory and jurisdictional requirements without slowing execution

Still, maintaining full control across the AI stack is neither practical nor cost-effective. That's why our research focuses on selective AI sovereignty—the ability to apply control across data, models, and infrastructure in proportion to business risk and strategic value. Not everywhere. Not equally. But where the economic and operational risks justify the investment.

For senior leaders, AI sovereignty is no longer a technology problem for IT to solve. It is central to business continuity, margin protection, and strategic flexibility. For CEOs, it's about strategic freedom. For CFOs, it's about reducing cost shocks and preserving capital flexibility. For COOs, it's about maintaining continuity under disruption. For CIOs and CTOs, it's about building architectures that can adapt while they scale. And for boards, it's about reducing concentration risk before it becomes a governance failure.

This report offers a roadmap to help these leaders enable selective AI sovereignty across the enterprise. Part one explains why AI has changed the sovereignty problem. In part two, we explore the key dimensions of AI sovereignty—data, models, and infrastructure—and show how selective control helps top performers protect profit. Part three outlines how to apply AI sovereignty across data, models, and infrastructure through a tiered operating model that aligns selective control with business priorities. We close with an action guide for CEOs, CFOs, COOs, CIOs/CTOs, and board members for the next 30 to 60 days, 60 to 120 days, and beyond.

Why selective AI sovereignty matters

CEOs see strategic flexibility at risk when AI dependencies can't shift.	CFOs absorb shock costs when data and models aren't portable.	COOs face downtime risk when enterprise AI becomes brittle.	CIOs/CTOs are limited by architectures that can't adapt under pressure.

> "Our architecture must be flexible enough to evolve, allowing us to swap components or adopt new technologies without starting from scratch."

Sergio Sánchez Gallego
CTIO, Telefónica España

Perspective

How many AI vendors does one organization need?

Most enterprises already operate in multi-vendor environments, with 28% of executives saying they have at least four AI vendors to maximize flexibility and avoid lock-in (see Figure 2). But when we examined what actually drives vendor diversity using ranking analysis, a more complex picture emerged.

We found that organizational fragmentation is the primary driver, followed by geographic and regulatory constraints. Legacy complexity and governance challenges also rank ahead of strategic intent. The deliberate use of multiple vendors to reduce risk and maintain optionality ranks last in relative importance.

In other words, most AI ecosystems are not the result of deliberate design. They are the accumulation of constraints, workarounds, and inherited decisions. That distinction matters. Without deliberate coordination, complexity fragments control—making it harder to enforce consistent policies, manage risk, or shift direction quickly.

Multi-vendor strategies only strengthen AI sovereignty when they are actively orchestrated. That means establishing common standards for data, models, and security; defining where variance is allowed and where consistency is required; and creating transparency across regions and business units. This lets leaders manage their AI vendors as a strategic portfolio, optimizing for adaptability, resilience, and long-term control.

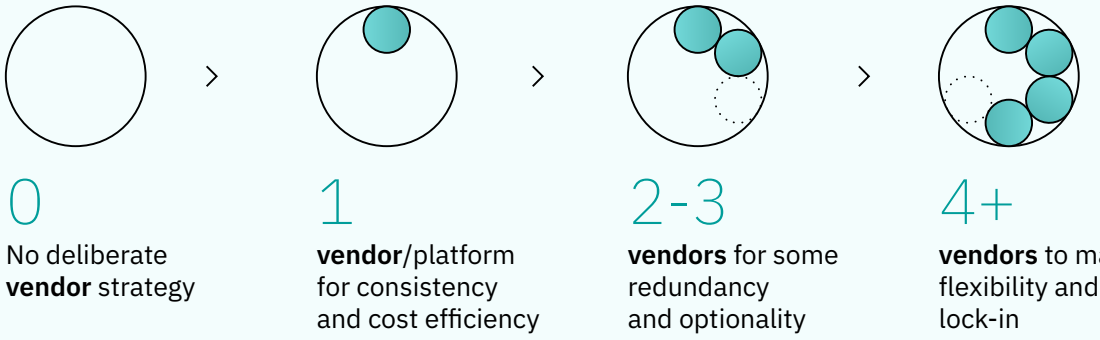
When managed deliberately, vendor diversity can provide the flexibility needed to protect profits. In fact, 72% of executives say they would accept a 20% cost increase to maintain multiple AI vendors if it improved strategic freedom. But this does not mean enterprises need multiple AI vendors in every instance. They should pay for optionality only where switching risk, criticality, and regulatory exposure justify the premium.

Perspective (continued)

FIGURE 2

Executives prefer multi-vendor AI strategies

Percentages reflect portion of executives primarily using each AI vendor strategy

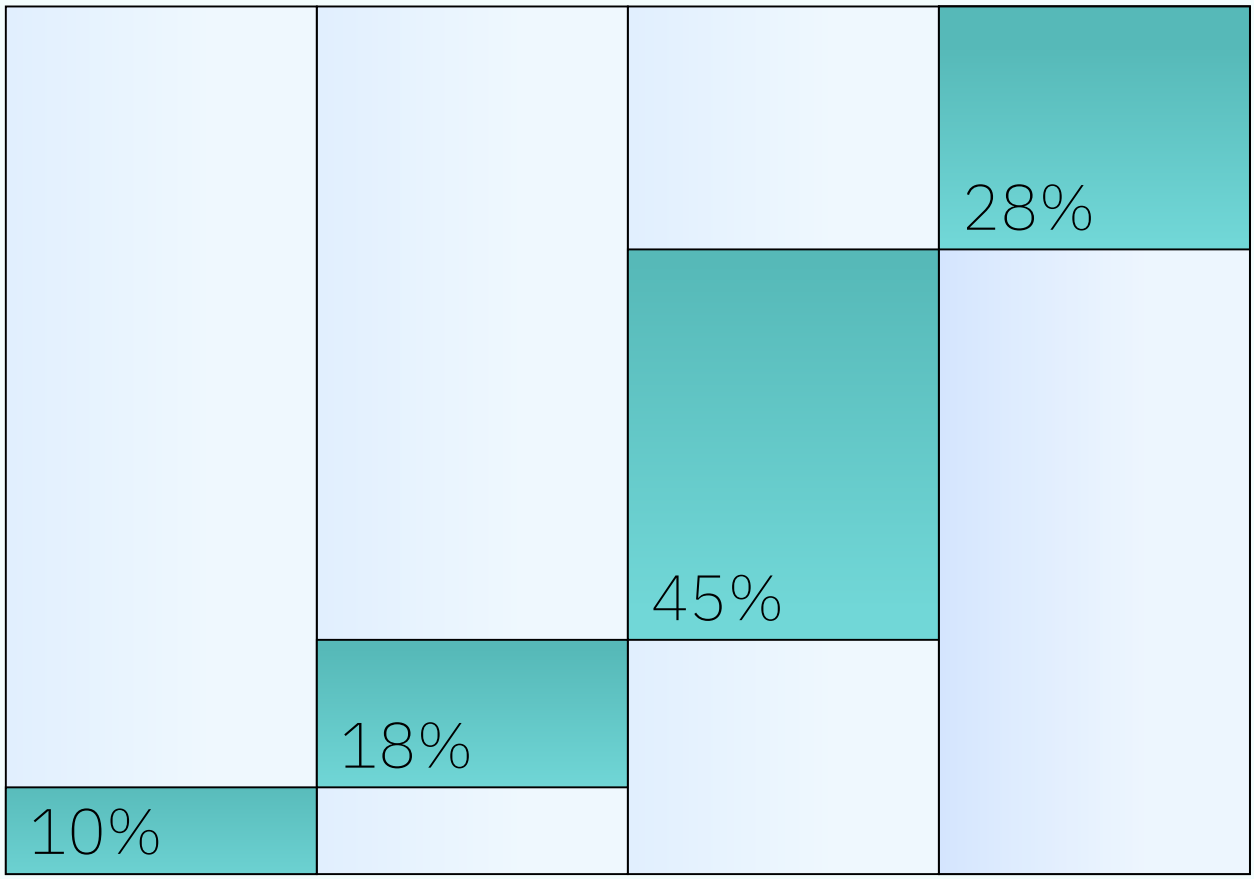


0
No deliberate **vendor** strategy

1
vendor/platform for consistency and cost efficiency

2-3
vendors for some redundancy and optionality

4+
vendors to maximize flexibility and avoid lock-in



Note: Columns do not add up to 100% due to rounding.

Part one

Why AI dependencies change the sovereignty problem

For decades, enterprises followed a consistent technology playbook: standardize, centralize, scale. It worked for ERP. It worked for infrastructure. It even worked for early cloud adoption.

It does not work for AI. What has changed is not just the pace of disruption—it is the nature of dependency itself.

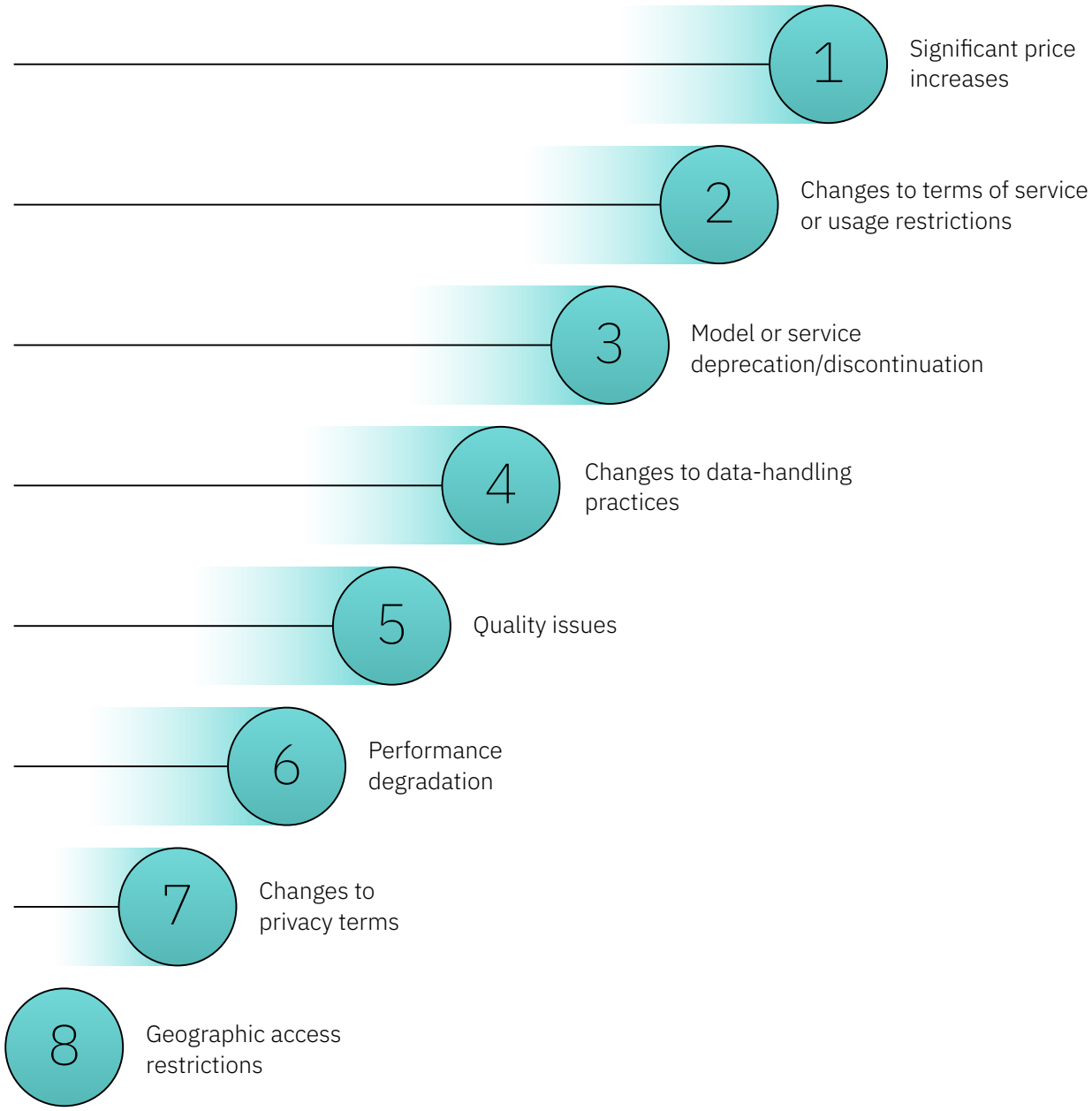
In traditional systems, dependency was concentrated in infrastructure and applications and changed relatively predictably. In AI, dependency extends into the model layer and the services around it—introducing new forms of volatility. For example, models can change behavior without formal release cycles and service terms and safety controls can be updated with little notice. As AI becomes more embedded in core processes, these dependencies put business continuity at risk.

Over the past 24 months, executives say dependencies have caused significant and unexpected shifts across the AI ecosystem: price increases, usage restrictions, model deprecations, changes to privacy and data-handling terms, performance degradation, and new geographic access limitations (see Figure 3). Because most AI architectures remain tightly coupled to particular vendors, regions, and operating assumptions, organizations are absorbing the cost of this volatility rather than adapting to it.

FIGURE 3

AI volatility threatens continuity and increases costs

Ranking reflects which unexpected changes organizations have experienced most frequently in the last 24 months



AI introduces new forms of dependency—across models, data, and services—that are more volatile, more opaque, and harder to unwind.

As a result, control must extend beyond architecture into execution—where decisions are made and acted upon in real time. AI sovereignty makes this possible. It lets organizations maintain effective control over AI data, models, and runtime environments when conditions change through:

Visibility: knowing dependencies, exposures, and points of lock-in

Governance: enforcing policy, compliance, and access boundaries

Portability: moving data and workloads across environments

Substitutability: switching models or providers when needed

Continuity: maintaining operations through disruption or degradation

But the control sovereignty requires is hard to come by. 75% of the executives who have switched or attempted to switch AI vendors in the past two years say the process was difficult due to data portability, model re-validation, compliance requirements, and technical lock-in. What should be a tactical adjustment often becomes a multi-month disruption.

To put it simply, enterprises don't really control AI. They control a tangled web of contracts, architectures, data flows, and operational choices spread across a complex estate. The implications are already visible across industries.

- In banking and insurance, model portability and auditability determine whether risk models can be validated, explained, and moved across jurisdictions.
- In healthcare and pharmaceuticals, regulated data flows and AI-assisted decisions require strict governance, traceability, and lifecycle control.
- In the public sector, defense, and critical infrastructure, mission-critical systems must operate even when dependencies fail, fragment, or fall under attack.
- In energy and utilities, operational continuity and jurisdictional control directly affect safety, reliability, and regulatory compliance.

When the underlying conditions shift, complex dependencies create costs that are often hidden until they hit the bottom line. Consider the impact of data location. When AI runs far from the data it relies on, costs climb quickly. Organizations in our study pay 2.8× more in token processing when data placement is misaligned with model execution—turning a technical choice into a recurring financial drag at scale. For a USD20 billion enterprise, it means spending roughly USD50 million more to deliver the same AI capabilities (see "Research methodology" on page 29).

To manage hidden AI costs more proactively, organizations are rethinking their platform strategies. Adding a coordinating layer—one that provides visibility across dependencies—helps ensure policies are consistently applied across systems, vendors, and environments.

Many organizations are moving mission-critical applications and data back from public cloud to on-premises or private environments—a process known as workload repatriation. Recent IBM IBV research shows that 24% of enterprises have already repatriated workloads, with another 38% planning to do so.[2] This shift, if focused on the right workloads, could help enterprises regain control, cost discipline, and resilience where it matters most.

"Rapid changes from hyperscalers—like model updates or service changes—forced us to rework our solutions constantly. To minimize disruption, we built an AI gateway layer that abstracts model dependencies, so we can switch between models and providers with minimal impact on our applications."

Anil Pradeep Kumara
Head of AI and Data, Sri Lanka Telecom

Perspective

A new blueprint for enterprise computing

Agentic AI is reshaping infrastructure requirements, as AI agents, large language models, and high-performance computing (HPC) workloads scale across increasingly distributed environments. The clear market reality is that AI and HPC now demand more than raw performance. They require platforms designed for hybrid deployment, regulatory alignment, and enterprise governance—so organizations can maintain control as conditions change.

As enterprises push AI into production and governments and regulated industries place greater emphasis on digital sovereignty, technology choices are increasingly influenced by the need for visibility, portability, and enforceable control. This is driving demand for more open, interoperable platforms—ones that can coordinate AI across environments while maintaining transparency, governance, and performance.

For example, IBM and AMD, a semiconductor company focused on high-performance and adaptive computing, are partnering to provide a sovereignty software platform that lets organizations implement continuous sovereignty controls across infrastructure, data, workloads, and operations—while maintaining the performance required for modern AI.[3]

Rather than relying on isolated tools or tightly coupled systems, organizations are beginning to treat AI as a coordinated system of operations. In this context, sovereignty is no longer just about where systems run, but how effectively they can be governed, adapted, and trusted over time. As a result, AI sovereignty increasingly depends on how well partners integrate to deliver consistent governance, shared visibility, and enforceable control across the entire AI stack.

Part two

Selective AI sovereignty: Gaining precision control

AI sovereignty is often treated as a binary choice: either the enterprise controls its AI—or it doesn't. But in practice, sovereignty is not an on/off switch. It's more like a set of dials—each tuned to a specific business impact.

Selective sovereignty reflects this reality. It gives leaders the option to reassert control when needed, without carrying the cost of full independence across the entire AI estate. Because every degree of added control carries cost, sovereignty is as much a financial problem as a technical one.

And we have measured the economic impact of getting it right. As we mentioned above, organizations with the strongest practical control across their AI stack experience less downtime and protect 55% more operating profit from AI-driven disruption. Yet these organizations are the minority. Our research shows that most organizations still have low-to-moderate control of key components across the AI stack (see Figure 4).

This lack of control translates directly into operational and financial risk. 81% of executives say that if a primary AI vendor experienced a seven-day outage, the impact would be severe or critical—effectively halting business operations. And these disruptions keep happening. Executives in our study report an average of six AI-related operational disruptions in the past two years, with vendor services being the leading cause (see Figure 5).

To help prevent this disruption, organizations need to increase control across the interdependent dimensions of AI sovereignty—data, models, and infrastructure.

FIGURE 4

Leaders lack control over their AI stack

Percentages reflect executives who have low, moderate, or high control over each AI component

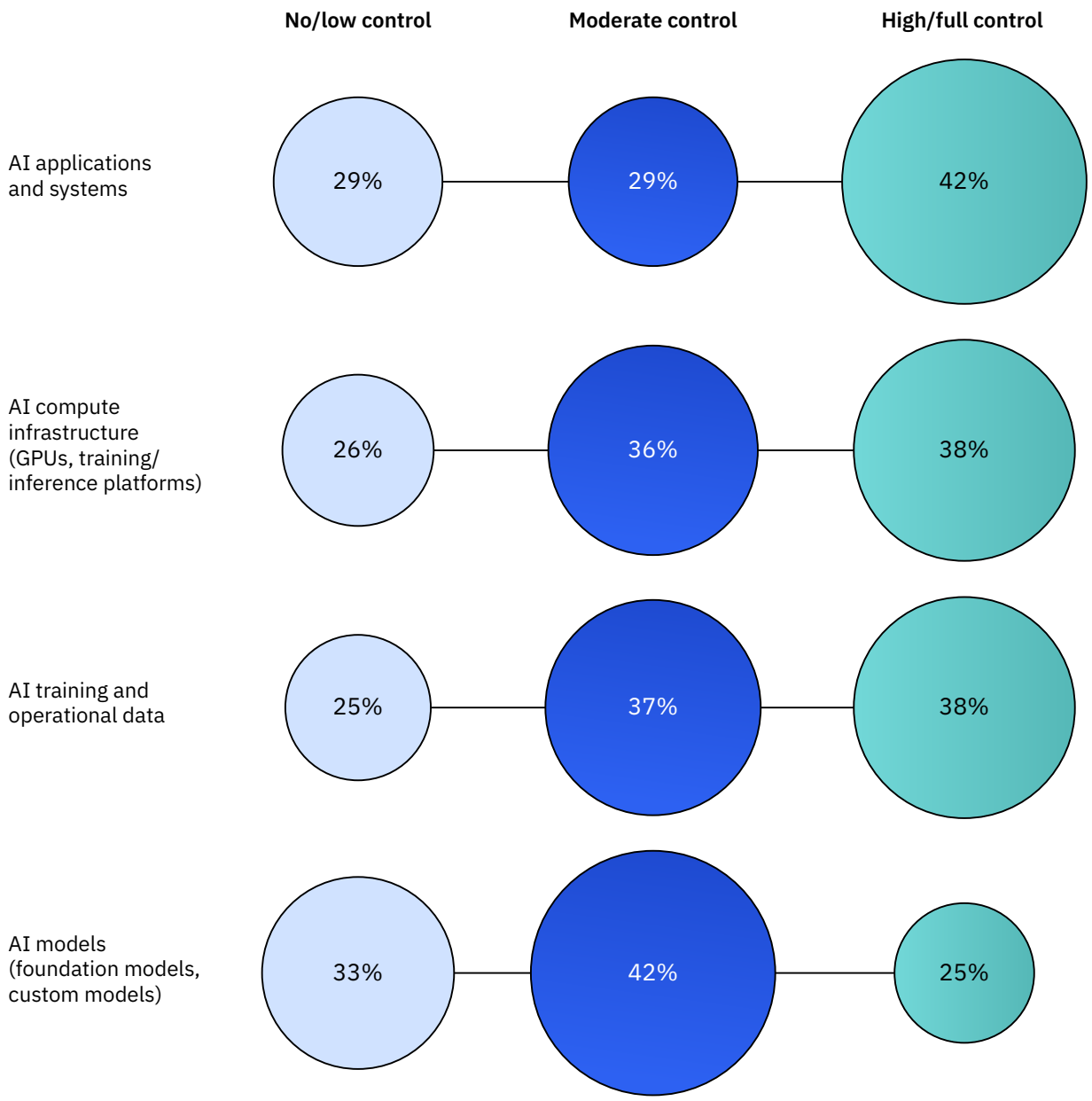
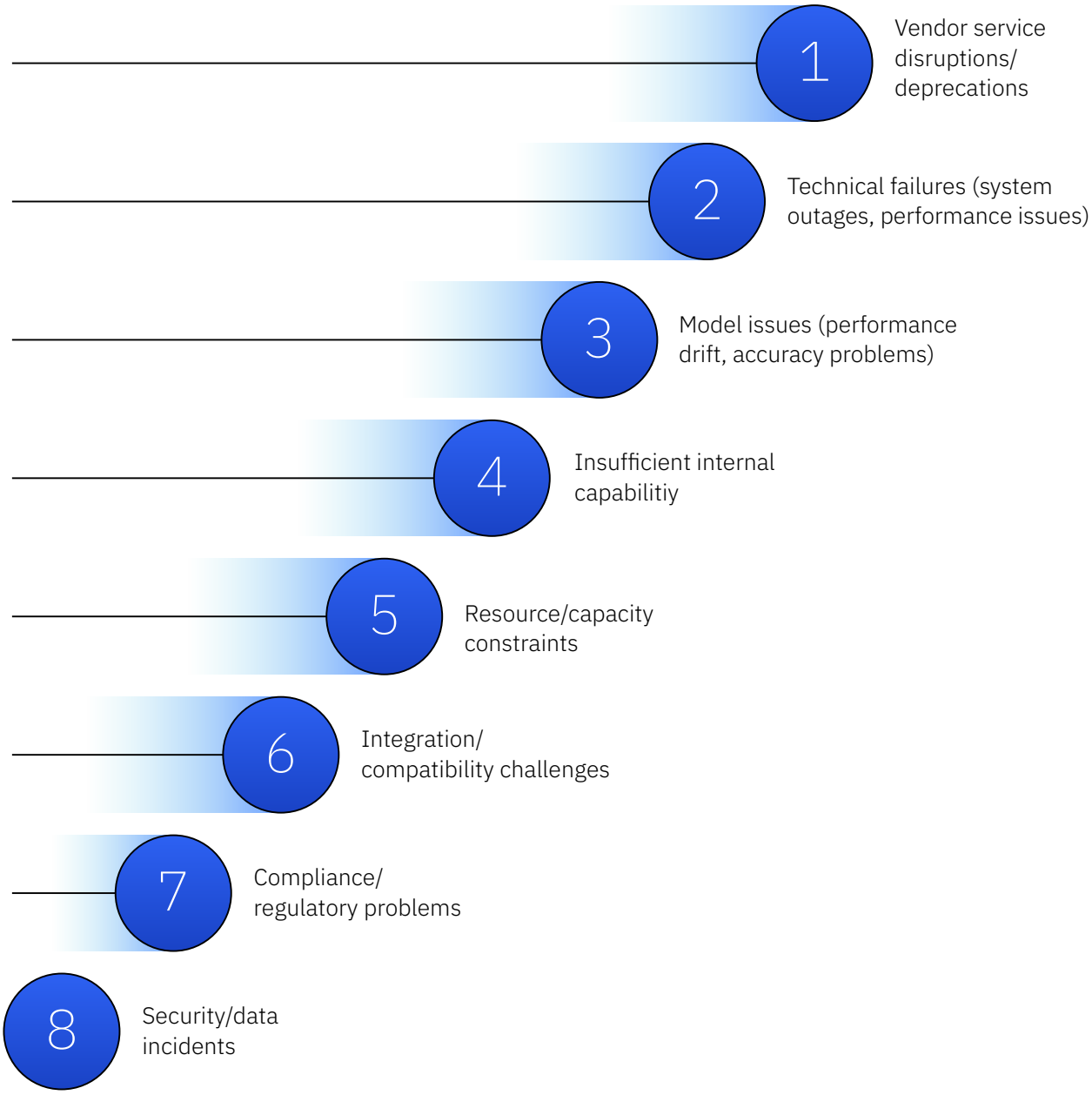


FIGURE 5

Top causes of operational disruptions to AI capabilities



Data sovereignty

CFOs and board members should ask:
Can we govern—and move—our AI fuel?

Data control is the foundation of AI sovereignty. Every model, application, and inference pipeline sits downstream from data. When data governance, access, or portability breaks, the entire stack is compromised. This is especially critical in industries such as healthcare, financial services, and the public sector, where data governance, residency, and auditability are regulatory requirements—not architectural preferences.

Across sectors, executives estimate it would take an average of 145 days to move their AI training and operational data to a different environment. This lag time creates a major risk when it comes to meeting data residency and sovereignty requirements across geographies, which 68% of executives say is a challenge. The top constraint cited: the technical complexity of data localization.

Without the ability to govern, report on, and relocate data with precision and speed, sovereignty remains theoretical. This is where platform choice becomes critical. As organizations face intensifying regulatory demands and technical complexity, some are turning to mainframe environments for their built-in governance, auditability, and resilience. Mainframes are especially valuable as AI moves into regulated, customer-facing transactions, providing the operational controls and transparency needed to maintain sovereignty and trust. Open data formats and open source data tooling can also strengthen control by reducing dependence on proprietary pipelines and making it easier to move data across environments.

Model sovereignty

CEOs and CIOs/CTOs should ask:
Can we swap models without starting over?

Model control determines how quickly an organization can respond when performance shifts, new requirements emerge, or a vendor changes terms. Yet owning a model doesn't mean you can change it at will.

In many cases, enterprises are constrained not just by the model itself, but by the surrounding tooling, tuning frameworks, and licensing terms. In fact, 57% of executives in our study say swapping or replacing a core AI model would require significant decoupling or a complete system rebuild. What should be a tactical adjustment becomes a strategic disruption—locking enterprises into decisions made months or years earlier. In regulated sectors such as banking and pharmaceuticals, model changes can also trigger revalidation, audit, and compliance cycles—further increasing the cost and complexity of switching.

Model sovereignty reduces this dependency by making it possible to swap models and associated capabilities internally or with compatible alternatives—including open source models where appropriate. Without this, enterprises cannot adjust when conditions demand it. "We accept long-term strategic relationships at the application layer," says Dalton Gouws, Group IT Director and Board Member, VWG UK Ltd. "But, at the model layer, we deliberately preserve optionality."

Infrastructure sovereignty

COOs and CIOs/CTOs should ask:
Can we move workloads without breaking the system?

Infrastructure sovereignty determines whether AI systems can scale, shift, or recover under pressure. It includes control over applications, orchestration layers, compute environments, and where workloads physically run.

In sectors such as energy, telecommunications, and defense, this level of control is directly tied to operational continuity and jurisdictional requirements, where outages or loss of control can have systemic impact. "Our architecture must be flexible enough to evolve, allowing us to swap components or adopt new technologies without starting from scratch," says Sergio Sánchez Gallego, CTIO, Telefónica España.

Most organizations struggle in this area, with 56% of executives saying it would take at least six months to move core AI systems and applications to a different vendor. To address this challenge, many organizations are adopting hybrid architectures—combining on-premises, sovereign cloud, hyperscalers, open source components, and regional providers—to optimize for flexibility and resilience.

Nearly half of executives say hybrid models deliver the best outcomes for transactional applications, and mainframe is preferred nearly five-to-one over public cloud for these use cases.[4] As AI moves into production, organizations are deliberately placing critical workloads on platforms—like mainframe—that maximize sovereignty and minimize risk.

> "Technology alone isn't resilience. In production environments, resilience means anticipation and prevention, not just recovery."
>
> **Dena Almansoori**
> Group Chief Technology and Innovation Officer, ADNOC

Part three

Turning sovereignty into an operating model

Selective AI sovereignty takes more than standards or tooling. It requires a deliberate structure for determining where the enterprise needs control, how much, and at what cost.

Our research shows that organizations can do this by classifying AI systems into tiers and applying different sovereignty expectations to each (see Figure 6). Not all AI is equal—and the operating model must reflect that.

This tiered approach establishes a shared language for decision-making, enabling CIOs, COOs, CFOs, and business leaders to align on how much control is required, where vendor dependency is acceptable, and where optionality must be engineered. It also creates a consistent way to coordinate decisions across the enterprise—managing policies, dependencies, and trade-offs with shared visibility.

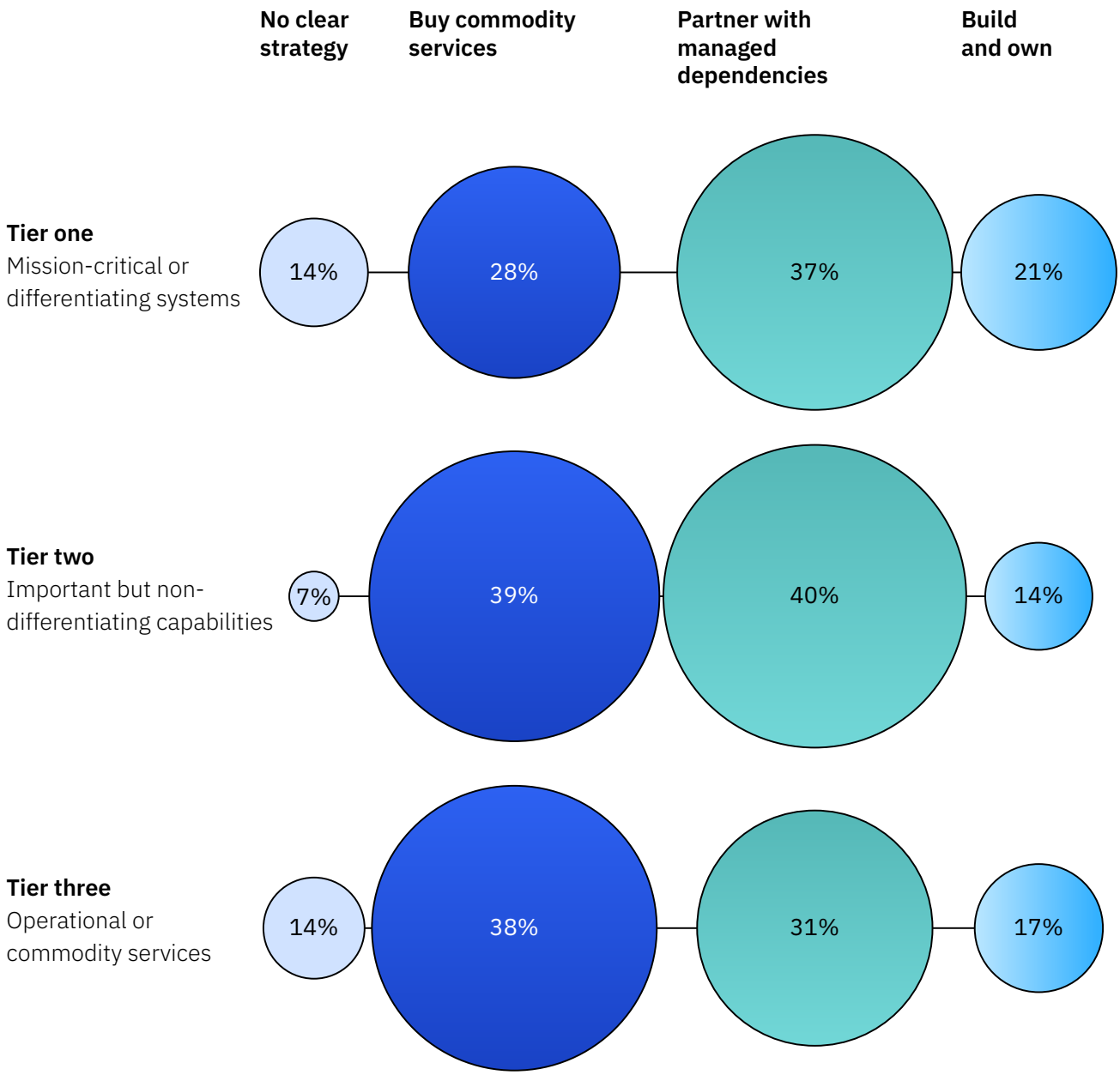
Regulatory classifications, such as the EU AI Act's high-risk categories, should inform enterprise AI tiering, but they do not replace it. Legal risk is one input, but the operating model must also reflect business criticality, dependency concentration, continuity requirements, and switching constraints.

Not all AI is equal—and the operating model must reflect that.

FIGURE 6

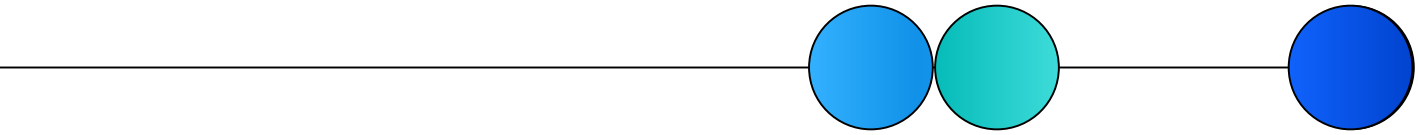
Organizations are taking a tiered approach to AI sovereignty

Percentages reflect portion of executives using each strategy, by tier of system criticality



Tier one

Mission-critical or differentiating systems



These are the systems that sit closest to the core of the enterprise: fraud engines, proprietary algorithms, core decisioning platforms, and the AI capabilities that materially influence revenue, risk, and operations. They process sensitive data and must remain reliable under stress. Here, the cost of optionality is justified—because the cost of failure, lock-in, or disruption is materially higher.

Tier one is where enterprises need the highest level of assured control. That doesn't always mean building every layer in-house—but it does mean being able to operate independently if needed. Leaders validate alternative models, maintain tested data egress paths, run outage simulations, and retain the expertise required to execute a controlled pivot. The goal is not self-sufficiency for its own sake; it is the confidence that no single dependency can dictate how critical systems operate.

In our recent research on mission-critical AI workloads, we found that performance, cost predictability, and regulatory confidence all increase when AI executes close to systems of record—particularly on platforms built for low latency and high resilience.[5]

What good looks like

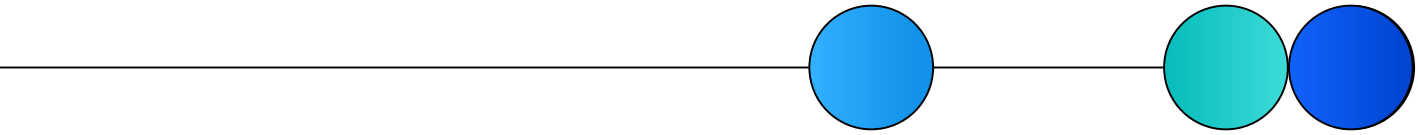
- Data can be relocated quickly and cleanly.
- Models can be tuned or swapped without system rebuilds.
- Workloads can shift to alternate environments with minimal degradation.
- Failover plans are tested, not theoretical.

"For us, control means safeguarding our strategic assets—especially core platforms and cloud infrastructure, where our dependence is absolute."

José Ovidio Silva
Group Head of Technology and Services, Campari

Tier two

Important but non-differentiating capabilities



These systems matter for performance and continuity, but they do not represent differentiated strategic advantage. Examples include customer service AI, HR analytics, and supply chain optimization—though classification depends on each enterprise's context.

For tier two, organizations can lean more on managed dependencies while retaining meaningful control over data and architecture. The objective is balance—keeping enough control to adapt to change without over-investing in flexibility that does not create meaningful economic value.

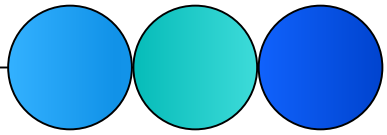
Managed dependency does not mean passive dependency. Teams must ensure modular designs, enforceable data exit rights, and clear visibility into how vendors handle model evolution and performance.

What good looks like

- Data rights and portability are contractually protected.
- Architectures remain modular, avoiding deep proprietary coupling.
- Vendor performance and model changes are monitored continuously.
- Teams can migrate or adapt systems with planned, not heroic, effort.

Tier three

Operational or commodity services



These are broadly available capabilities with low switching costs and limited strategic value, such as transcription, translation, and routine automation. Here, full sovereignty creates unnecessary cost and complexity—and erodes the economic value that selective sovereignty is designed to deliver. Vendor dependency is acceptable as long as choices are explicit and governed.

What good looks like

- Services are optimized for cost and simplicity.
- Vendor dependency is intentional, not inherited.
- Options exist but don't require upfront investment.
- Controls are proportional to business impact.

From governance to value protection

A tiered operating model converts sovereignty from an abstract ideal into a measurable source of resilience. It concentrates investment where control matters most, reduces unnecessary duplication, and prevents organizations from over-engineering low-value capabilities while under-protecting critical ones.

Organizations that implement coordinated control across data, models, applications, and compute—especially for tier one and tier two—are better positioned to absorb disruption without cascading failure. This is how they protect 55% more of their operating profit.

AI sovereignty is not about controlling everything. It is about knowing exactly what you control, where you are exposed—and ensuring you can keep critical operations running no matter what the future may bring.

Perspective

Telecommunications has its AI moment—and its message

For years, enterprises treated telecommunications providers as connectivity partners. That role is changing—quickly.

As AI sovereignty moves from policy debate to operating requirement, telcos are emerging as critical enablers. They already operate where sovereignty matters most: inside national borders, under strict regulatory regimes, and at the intersection of data, compute, and critical infrastructure.

Unlike global hyperscalers, telcos are structured for locality, jurisdictional control, and resilience. They are uniquely positioned to support enterprises that need AI systems to run close to their data, comply with domestic requirements, and remain operable when cross-border dependencies become constrained.

Across Europe, for instance, Orange Business notes that as enterprises deploy AI into core operations, the critical question is no longer just where data are stored, but who retains legal oversight and operational control.[6] In this environment, telcos are becoming more than connectivity providers. They are infrastructure partners for selective sovereignty—offering enterprises a way to localize AI workloads, reduce switching friction, and preserve optionality.

As AI architectures evolve, sovereignty will not be delivered by a single provider or platform. It will be assembled across hybrid ecosystems—linking internal infrastructure, hyperscalers, and telco-enabled sovereign environments. Enterprise leaders who recognize this early gain more than compliance. They gain flexibility.

> "We've adopted a dual vendor strategy—maintaining relationships with two hyperscalers for redundancy and cost control. This approach avoids the complexity of managing many providers while ensuring we have the flexibility and resilience to switch if needed."

Anil Pradeep Kumara
Head of AI and Data, Sri Lanka Telecom

Action guide

How to make AI sovereignty real

As AI systems take on a growing share of operational decisions, maintaining control depends on selective sovereignty—the ability to apply control across data, models, and infrastructure in proportion to business risk and strategic value. What follows is an operating playbook for executives and board members who want to make their enterprise more adaptable in the face of change.

CEO: Set the rules of control

Objective: Define where the enterprise requires sovereignty and enforce it.

What to do now (30 to 60 days)

- Approve the enterprise's three-tier classification for all AI systems.
- Require each tier one owner to explain—in plain language—their system's critical dependencies and switching constraints.

What to do soon (60 to 120 days)

- Sponsor the first AI outage simulation and require remediation plans.
- Tie executive objectives to resilience and optionality metrics.

Results to watch for

- Every leader understands tier one exposure.
- Alternatives for critical systems are validated, not assumed.

External board: Enforce discipline, not optimism

Objective: Ensure management is reducing structural exposure, not accumulating it.

What to do now (30 to 60 days)

- Request a tier one dependency map with clear continuity tiers.
- Ask for evidence of switching capability—not just contracts—including the ability to move across proprietary and open source alternatives.

What to do soon (60 to 120 days)

- Confirm an enterprise owner for AI resilience across technology, risk, procurement, and operations.
- Require trend reporting: time-to-switch, vendor concentration, and tested alternatives.

Results to watch for

- Exposure decreases quarter over quarter.
- Switching times shrink.

CFO: Fund optionality like insurance

Objective: Allocate capital to protect flexibility before disruption hits.

What to do now (30 to 60 days)

- Require tier one systems to report time-to-switch and cost-to-switch.
- Introduce a sovereignty premium based on system criticality.

What to do soon (60 to 120 days)

- Ensure contracts include enforceable data portability rights and exit SLAs.
- Track three metrics: time-to-switch, cost-to-switch, and percentage of tier one systems with validated alternatives.
- Identify tier one systems facing vendor lock-in and pricing volatility, then explore alternative strategies.

Results to watch for

- Funding flows first to data and model control—the highest-leverage investments.
- Investments are optimized for future freedom as well as short-term costs.

COO: Make operations durable

Objective: Ensure AI failure does not become business failure.

What to do now (30 to 60 days)

– Map which business processes depend on tier one AI.

– Define continuity targets: acceptable degradation, recovery time, escalation paths.

What to do soon (60 to 120 days)

– Conduct AI failover drills and operate in degraded mode.

– Implement monitoring: drift detection, audit logs, fallback workflows.

Results to watch for

– Operations continue when AI degrades.

– Teams know exactly how to respond in outage simulations.

CIO/CTO: Build the foundation

Objective: Engineer control where it matters most.

What to do now (30 to 60 days)

– Map full dependencies for all tier one systems—data, models, orchestration, compute.

– Identify where open source models or frameworks can reduce lock-in and provide viable fallback options.

– Inventory the non-portable joints creating lock-in.

What to do soon (60 to 120 days)

– First, standardize data portability formats across both proprietary and open source environments. Create internal mirrors or shadow copies of critical data. Test extraction from vendor environments.

– Next, ensure fine-tuning and retraining can occur internally or via a second provider. Validate model-swap pipelines.

– Finally, containerize where possible. Build region or provider failover. Document performance under alternative environments.

Results to watch for

– Tier one systems have tested alternatives for data, models, and runtime—including viable open source substitutes where appropriate.

– Migration takes weeks—not months.

About the authors

Hans Dekkers

General Manager, Asia Pacific
IBM
linkedin.com/in/hans-a-t-dekkers/

Hans Dekkers leads IBM's strategy, operations, and technology delivery across Asia Pacific, working with clients to accelerate hybrid cloud and AI adoption in complex and rapidly evolving markets. With leadership experience spanning Europe and Asia, he brings a strong focus on helping organizations balance innovation with regulatory compliance, data control, and local ecosystem development—key to advancing AI sovereignty.

Javier Olaizola

Global Managing Partner,
Hybrid Cloud and Data
IBM Consulting
linkedin.com/in/javier-olaizola-casin/

Javier Olaizola leads global hybrid cloud and data initiatives within IBM Consulting, helping enterprises scale AI across distributed and multicloud environments. He works closely with clients to build trusted data platforms, modernize architectures, and operationalize AI with strong governance and security. His focus includes addressing interoperability, compliance, and control—core requirements for enabling AI sovereignty at scale.

Manish Goyal

Senior Partner, Enterprise AI Strategy and Governance, Global Offering Leader
IBM Consulting
linkedin.com/in/goyalm/

Manish Goyal advises global enterprises on AI strategy, governance, and large-scale adoption, helping organizations move from experimentation to enterprise-wide deployment. He brings a combination of consulting and product experience to support trusted AI implementation, with an emphasis on data governance, risk management, and regulatory alignment. His work centers on building AI capabilities that are secure, responsible, and sovereignty-ready.

Sripriya Srinivasan

General Manager, IBM Software Products
IBM Software
linkedin.com/in/sripriya-srinivasan-385a0812/

Sripriya Srinivasan leads global software products, support, and site reliability engineering, driving the performance and resilience of IBM's enterprise software portfolio. With more than two decades of experience across data, AI, and cloud platforms, she focuses on ensuring secure, reliable operations for mission-critical systems. Her work helps organizations deploy and manage AI applications with the control, availability, and security required in sovereign and regulated environments.

Priya Kurien

Research Director
IBM Institute for Business Value
linkedin.com/in/priya-kurien/

Priya Kurien is a Research Director at the IBM Institute for Business Value, where she leads research on technology leadership and AI. With more than 25 years of experience across engineering and enterprise technology, she advises C-suite and board-level leaders on scaling AI—focusing on the structural decisions across governance, architecture, and investment that determine control, resilience, and business impact.

Robert Zabel

Infrastructure Research Lead
IBM Institute for Business Value
linkedin.com/in/rzabel/

Robert Zabel is Infrastructure Research Lead at the IBM Institute for Business Value. With more than 25 years of consulting experience, he has advised senior leaders on enterprise strategy, large-scale transformation, and technology-driven change. His work spans infrastructure modernization, hybrid cloud, and AI, focusing on how organizations evolve operating models to improve performance, flexibility, and long-term competitiveness.

Research methodology

The IBM IBV conducted a global survey between February and April 2026 to examine how organizations structure control across the AI stack and how these choices relate to resilience, performance, and operating economics. The study draws on responses from 1,000 senior executives responsible for AI, data, technology, or related enterprise capabilities.

Respondents represent organizations across 16 geographies and 17 industries, reflecting a broad range of operating environments, regulatory contexts, and stages of AI adoption. The sample is weighed toward large, complex enterprises, with organizations ranging from USD260 million to USD92 billion in annual revenue or budget.

This composition enables robust comparison across organizations with materially different AI operating models while maintaining a strong focus on enterprises actively scaling AI capabilities.

Analytical approach

A data-driven segmentation of AI control models was developed using four survey measures capturing the organization's level of control over:

- AI training and operational data
- AI models (foundation and custom models)
- AI compute infrastructure
- AI applications and systems

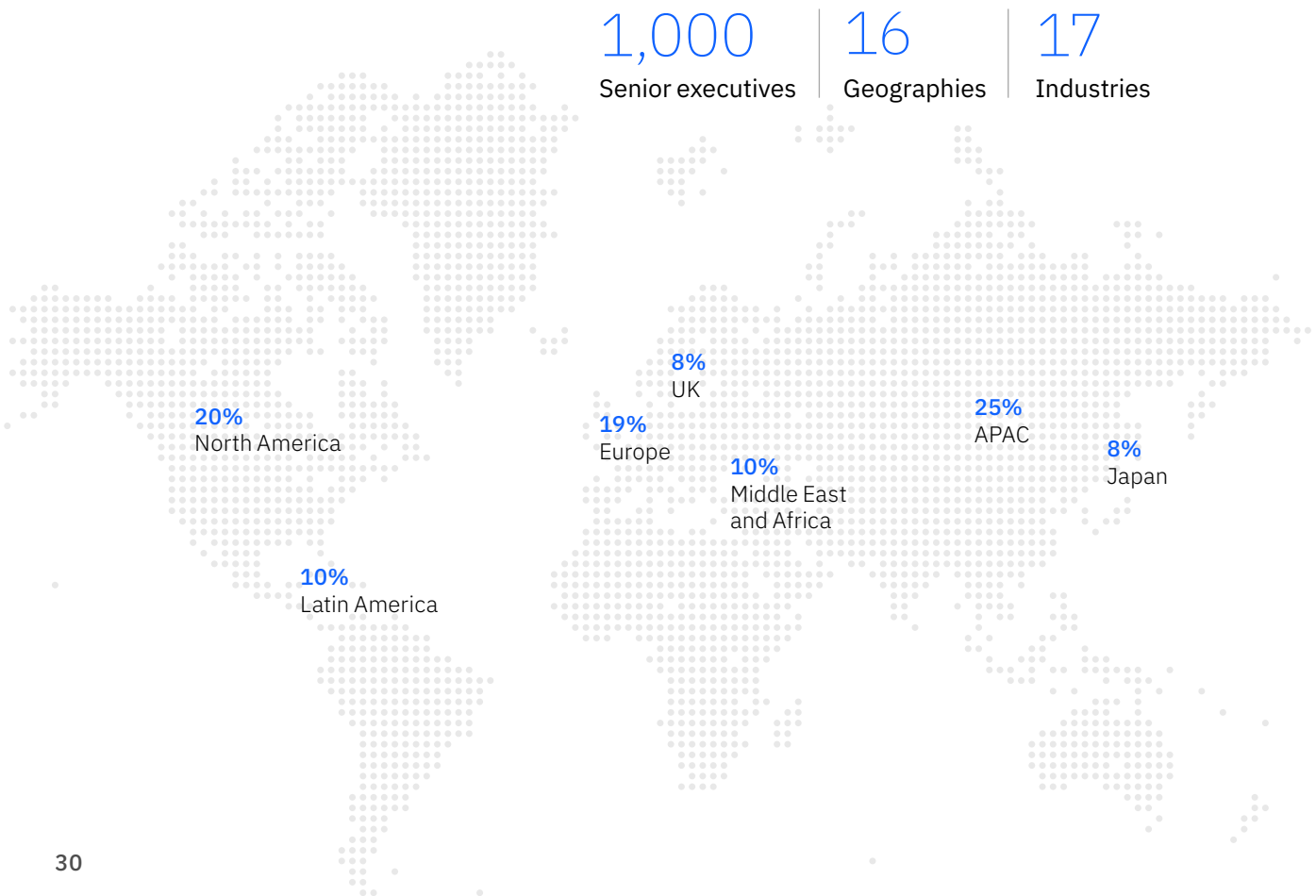
These variables were analyzed using multiple correspondence analysis (MCA) to identify the underlying multidimensional structure of AI control. MCA enables identification of structural differences in how organizations distribute control across the AI stack, rather than imposing a single linear maturity scale.

The resulting respondent coordinates were used as inputs into a hierarchical cluster analysis (Ward's method), which minimizes within-cluster variance and produces compact, interpretable segments. This approach yielded a set of distinct AI control profiles that reflect how organizations design and govern AI systems.

To validate the segmentation, analysis of variance (ANOVA) confirmed statistically significant differences across all control dimensions, indicating clear separation between groups. Post-hoc comparisons further support that segments differ meaningfully from one another.

These segments were subsequently used to estimate differences in AI-related value protection efficiency, measured as expected avoided downtime loss normalized per USD1 billion of revenue. This analysis provides the basis for the finding that end-to-end control models protect 55% more AI-related business value than other operating models.

To size the financial impact of poor data locality, we modeled a representative USD20 billion organization and compared AI delivery costs across local and remote deployment scenarios. Based on the cost patterns identified in this study, the model indicates that running AI far from the data it depends on can add roughly USD50 million in annual cost without improving capability.



How IBM can help

IBM helps organizations operationalize AI sovereignty by embedding control, compliance, and flexibility into their hybrid environments. By combining hybrid cloud platforms, AI governance capabilities, and consulting expertise, IBM enables enterprises to manage where data resides, how AI operates, and how outcomes are audited—across jurisdictions and ecosystems. Solutions such as IBM Sovereign Core provide a ready-to-run foundation with integrated control planes, continuous compliance evidence, and governed workflows, helping organizations scale AI while maintaining security, regulatory alignment, and long-term architectural flexibility. Backed by deep industry experience and an open, hybrid-by-design approach, IBM helps clients deploy AI with confidence—balancing innovation with trust. To learn more about IBM Sovereign Core, visit: ibm.com/products/sovereign-core

About Research Insights

Research Insights are fact-based strategic insights for business executives on critical public- and private-sector issues. They are based on findings from analysis of our own primary research studies. For more information, contact the IBM Institute for Business Value at iibv@us.ibm.com.

IBM Institute for Business Value

For two decades, the IBM Institute for Business Value has served as the thought leadership think tank for IBM. What inspires us is producing research-backed, technology-informed strategic insights that help leaders make smarter business decisions.

From our unique position at the intersection of business, technology, and society, we survey, interview, and engage with thousands of executives, consumers, and experts each year, synthesizing their perspectives into credible, inspiring, and actionable insights.

To stay connected and informed, sign up to receive IBV's email newsletter at ibm.com/ibv. You can also find us on LinkedIn at ibm.co/ibv-linkedin.

The right partner for a changing world

At IBM, we collaborate with our clients, bringing together business insight, advanced research, and technology to give them a distinct advantage in today's rapidly changing environment.

IBM IBV contributors

Douna Daou, Tegan Jones, Heba Nashaat, Thiago Sartori

Related reports

The mainframe advantage: Workload placement and execution decisions for mission-critical systems in the AI era.
IBM Institute for Business Value.
May 3, 2026. ibm.biz/mainframe-advantage

2026 CEO Study. *Rewiring the C-suite: The fast track to 2030.*
IBM Institute for Business Value.
May 3, 2026. ibm.biz/ceo-2026

5 trends for 2026: Capture fleeting opportunities with confidence.
IBM Institute for Business Value.
December 1, 2025. ibm.biz/tech-trends-2026

Notes and sources

1 2026 CEO Study. *Rewiring the C-suite: The fast track to 2030.* IBM Institute for Business Value. May 3, 2026. https://ibm.biz/ceo-2026

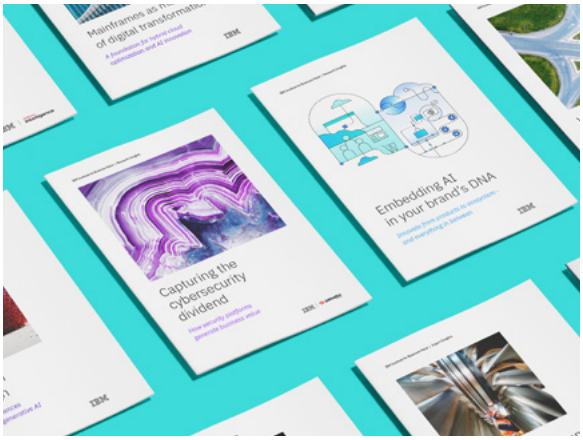
2 Pete McCaffrey, Anita Childs, Matt Lyteson, William Lobig, and Barry Baker. *The mainframe advantage: Workload placement and execution decisions for mission-critical systems in the AI era.* IBM Institute for Business Value. May 3, 2026. https://ibm.biz/mainframe-advantage

3 "AMD and IBM enhance collaboration across hybrid, sovereign and enterprise AI." IBM. May 6, 2026. https://www.ibm.com/new/announcements/amd-and-ibm-deepen-collaboration-across-hybrid-sovereign-and-enterprise-ai

4 Pete McCaffrey, Anita Childs, Matt Lyteson, William Lobig, and Barry Baker. *The mainframe advantage: Workload placement and execution decisions for mission-critical systems in the AI era.* IBM Institute for Business Value. May 3, 2026. https://ibm.biz/mainframe-advantage

5 Ibid.

6 Krásová, Tereza. "Orange Business recognizes limits of Europe's tech sovereignty." LightReading. March 26, 2026. https://www.lightreading.com/digital-transformationorange-business-recognizes-limits-of-europe-s-tech-sovereignty



Subscribe to our IdeaWatch newsletter

Just the insights. At your fingertips. Delivered monthly.

Brought to you by the IBM Institute for Business Value, ranked #1 in thought leadership quality by Source Global Research for the third consecutive year.

Research-based thought leadership insights, data, and analysis to help you make smarter business decisions and more informed technology investments.

Subscribe now: ibm.biz/ideawatch

IBM Corporation
New Orchard Road
Armonk, NY 10504

Produced in the United States of America | June 2026

16ddce7b71d48f55-USEN-01