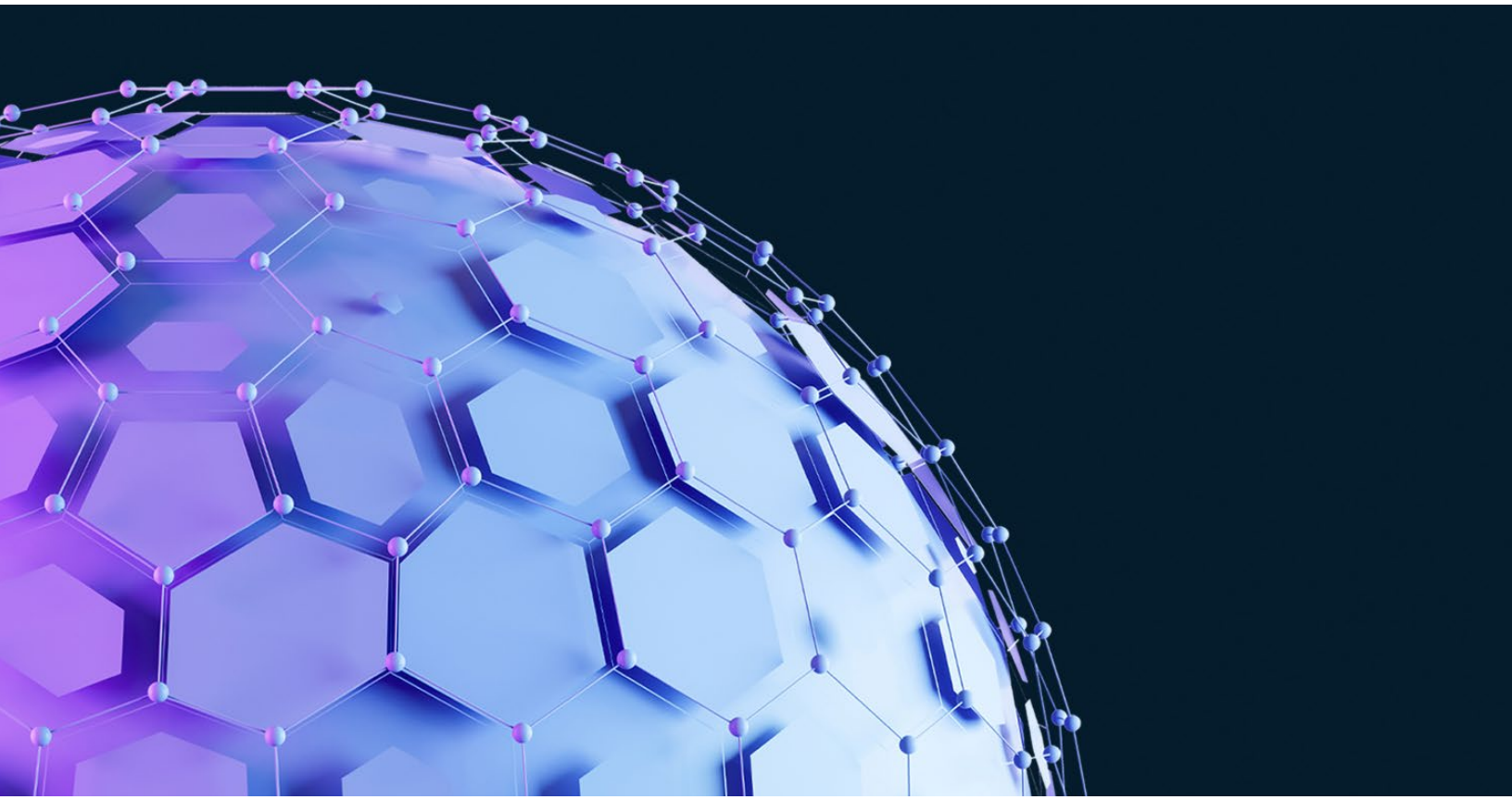
McKinsey & Company

Technology, Media & Telecommunications Practice

Sovereign AI: Building ecosystems for strategic resilience and impact

Sovereign AI is achievable only through an ecosystem effort that connects energy, compute, data, models, platforms, and applications across multiple actors.

This article is a collaborative effort by Ali Ustun, Arnaud Tournesac, Luca Bennici, and Ruben Schaubroeck, with Justin de Niese, Kaavini Takkar, Melanie Krawina, and Newfel Drahmoune, representing views from McKinsey's Technology, Media, & Telecommunications and Public Sector Practices and McKinsey Technology.



March 2026

Disclaimer:
This article is descriptive and analytical. It does not provide policy, regulatory, or national security advice, and it does not recommend specific national strategies. Examples are included to illustrate ecosystem patterns and should not be interpreted as endorsements or prescriptions.

Sovereign AI is moving from a policy debate to an economic and strategic imperative. Across governments, enterprises, and investors, leaders increasingly view the ownership of AI capabilities as central to economic competitiveness, strategic resilience, and societal trust.

Yet despite this urgency, many sovereign AI initiatives are stalling and failing to deliver their expected results. In this article, we analyze how sovereign AI efforts are being pursued and what differentiates sovereign ecosystems that successfully translate intent into scaled adoption and durable advantage. Drawing on a global survey of enterprises, providers, governments, and investors,[1] we then examine the roles different actors must play, the challenges actors face, the partnership models that consistently outperform, and a practical road map for building sovereign AI capabilities that compound over time.

Sovereign AI refers to a nation's or organization's ability to develop and control its own AI capabilities to ensure strategic independence and alignment with domestic values and laws. That said, sovereign AI does not have a single definition; rather, it is the result of the interaction between four distinct components:

— *territorial:* where data and compute physically reside

— *operational:* who manages and secures data and compute

— *technological:* who owns the underlying stack and intellectual property

— *legal:* which jurisdiction governs access and compliance

Viewed this way, sovereign AI is best thought of as a spectrum of potential solutions distributed across different tiers of sovereignty, depending on stakeholder and local circumstances (Exhibit 1).

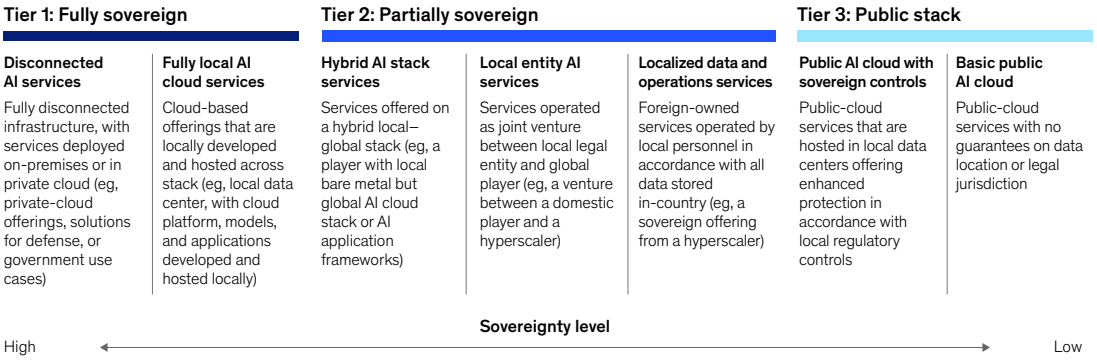
As a result, sovereign AI represents one of the largest opportunities within AI. McKinsey estimates that 30 to 40 percent of AI spending could be influenced by sovereignty requirements. This would represent a market of some $500 billion to $600 billion globally by 2030 (Exhibit 2).

But seizing that opportunity means confronting a very specific execution challenge: Success in sovereign AI is not achieved through a single policy decision, cloud contract, or "national model" announcement. Instead, sovereignty is best thought of as an ecosystem effort that connects multiple layers—energy, compute, data, models, cloud platforms, and applications—into one coherent system, managing fragmentation across ownership, operating models, and accountability, with explicit trade-offs and deliberate choices about what truly needs to be sovereign.

[1] *Tech: Forward,* "The sovereign AI agenda: Moving from ambition to reality," McKinsey, December 18, 2025.

Exhibit 1

AI workloads can be tiered along a spectrum, based on opportunities for sovereignty.



Source: *Tech: Forward*, "The sovereign AI agenda: Moving from ambition to reality," McKinsey, Dec 18, 2025

McKinsey & Company

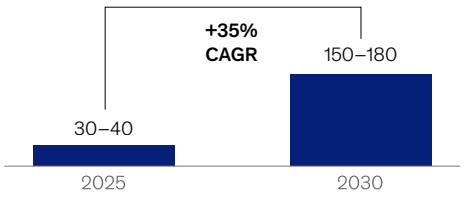
Exhibit 2

The sovereign AI opportunity, already sizable, could expand to a $600 billion market by 2030.

Global market size, total addressable market (TAM),[1] $ billion

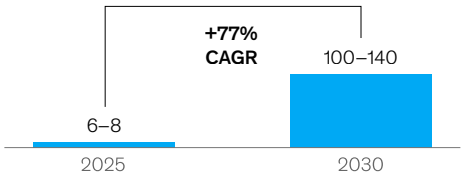
Key drivers

Applications and use cases



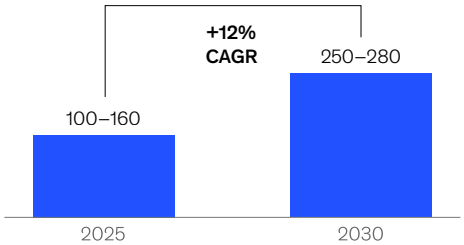
Growing local and regional regulatory and compliance requirements are leading nations to reduce dependency on global providers.

Models, data, and tooling



National and regional actors seek model independence to deliver more-relevant outcomes for local populations and to comply with data localization initiatives.

Infrastructure and compute



In response to rising geopolitical uncertainty, nations are co-investing in and building compute capacity and data centers to reduce reliance on global providers.

Total estimated demand ~150–200 ~500–600 **30–40%** of AI market

[1]TAM is equivalent to total estimated demand for sovereign AI based on baseline AI numbers and sovereignty requirements, which is different from actual spend captured by sovereign AI solutions (which evolves from low percent in 2025 to high percent in 2030, differentiated based on national archetypes).
Source: *Tech: Forward*, "The sovereign AI agenda: Moving from ambition to reality," McKinsey, Dec 18, 2025

McKinsey & Company

The sovereign AI ecosystem: Moving from 'sovereign assets' to 'sovereign assets and outcomes'

For business leaders and policymakers, a useful starting point is to change the way they think about sovereign AI. Many initiatives focus on inputs—such as GPUs, data centers, cloud regions, and national model announcements—and while those inputs matter, the prize is in the long-term outcomes, such as strategic resilience, autonomy, economic value capture, and better social outcomes.

An effective sovereign ecosystem is not necessarily one in which everything is built domestically. Instead, it is one in which key control points are sovereign by design, even if other elements of the stack may remain open to partnerships, interoperability, and competition. Our analysis shows that the most effective ecosystems operationalize "minimum sufficient sovereignty" with a repeatable decision rule: Classify workloads by the importance of regulatory issues and third-party exposure, and then assign a sovereignty tier with explicit requirements for data residency, key ownership, and access controls.

Different jurisdictions are pursuing sovereign AI through distinct ecosystem archetypes. Even jurisdictions with advanced capabilities are rarely self-sufficient across all layers and often rely on external providers in at least part of the stack (particularly in hardware and advanced compute). The following are among the different approaches we are seeing:

— *End-to-end hub and frontier AI powerhouses.* In this model, private data center operators build massive AI-ready capacity to attract hyperscalers and frontier AI labs, enabling large-scale training and advanced inference ecosystems.

— *State-led with data center or cloud execution.* This approach is led by a state to keep compute, data, and model intellectual property under national control for strategic and public sector workloads. Local cloud and data center providers execute the scale-up of sovereign compute and data platforms under state requirements.

— *Model development led by research and policy.* Led by research institutions and policymakers, regulatory and state incentives steer domestic model development and compliant data access, with cloud and data center ecosystems providing trusted environments for research compute to operationalize locally developed models.

— *Industry-led and compute-hardware-driven adoption.* In this approach, data center providers and cloud players partner with local enterprises and chip ecosystem and semiconductor leaders to create AI-capable regional platforms.

— *Policy-enabled regional hubs with strong local or regional demand.* In this model, policy enables fast permitting, power access, and investment incentives, and data center operators build and aggregate AI-ready capacity to serve public sector and regional-enterprise demand and attract anchor tenants.

Even amid these differences, effective sovereign AI ecosystems tend to share a common set of characteristics:

— *A demand-led anchor and sectoral adoption.* Leaders start by clustering demand sources (such as citizen services, health outcomes, financial integrity, critical infrastructure protection, and industrial productivity). This entails an explicit shift in public sector demand: Governments align procurement, funding, and operating mandates with the sovereign AI strategy, acting as anchor customers.

— *Sovereign AI infrastructure as a key pillar enabled by foundational inputs.* Strong ecosystems typically have a deep in-country compute backbone: data centers, high-density GPU clusters, cloud platforms, subsea cables, and low-latency networks that host and run AI workloads. The compute backbone is enabled by physical resources that make infrastructure viable: reliable and affordable power, green energy, land, and water.

— *A clear sovereignty baseline and reference architecture.* Because sovereignty is multidimensional, effective ecosystems codify what must be sovereign into a reference architecture with a set of nonnegotiable control points: data classification and permitted uses; encryption and key ownership; identity and access, logging, and monitoring model risk management and evals; and incident response and legal access pathways.

— *Trusted data governance, policies, and standards.* Governance, policies, and standards constitute an enabling force that shapes speed and scale through land and power allocation, compute import and export rules, incentives, and governance frameworks that attract demand and investment.

— *A data ecosystem and a pragmatic modular model strategy.* Localization keeps data "inside," but it does not automatically make data usable. Strong sovereign ecosystems build data products and sharing mechanisms: interoperable standards and sector consortiums that increase the quantity and quality of training and fine-tuning data. With regard to AI models, ownership builds independence, but bringing the best AI models to the country often means giving up control. As a result, countries are increasingly adopting a layered strategy: leveraging global frontier models where possible and developing or fine-tuning domain and language models where sovereignty needs and value are highest.

— *Capital that matches the stack's risk profile.* Financing fuels build-out and innovation, with mechanisms spanning public incentives, venture capital, and private equity for infrastructure, start-ups, and enterprise adoption. Energy and data centers need patient, infrastructure-style capital; models and platforms need growth capital with risk tolerance; and applications and integration need venture pathways and enterprise adoption muscle. Leading ecosystems align capital instruments to each layer.

— *Local talent that is attracted or nurtured.* The AI talent pipeline is fast emerging as a scarce resource, necessitating rapid upskilling. This includes sustained investment in education, reskilling, and lifelong-learning programs that prepare workers for new roles created by AI integration.

Taken together, these seven elements describe not a checklist, but the components of a system. Effective ecosystems are those that treat sovereignty as a coordinated design problem, aligning these components so that each reinforces the others.

The roles that matter in successful ecosystems: Governments, providers, enterprises, and investors

Building a sovereign AI ecosystem requires coordination across four distinct groups: Governments shape trust, rules, and demand; providers create the underlying technology and platforms; enterprises convert infrastructure into real economic value; and investors supply the capital and risk tolerance needed to scale.

Below, we examine each role, drawing on our survey insights to highlight the specific constraints and the actions required to move from fragmented pilots to scaled outcomes.

Governments: Act as orchestrator, investor, regulator, and anchor customer
Governments have a unique ability to turn fragmented ambition into coordinated execution and thus play a central role in leading ecosystems.

Governments set the sovereignty goalposts. They define which workloads require strong sovereignty (for example, defense, sensitive citizen data, and critical infrastructure), which can use hybrid models, and which can remain largely global. They then translate those choices into actionable controls (such as data classification, auditability, and key ownership). By creating certification regimes, governments can help standardize what "trusted" means so regulated industries can adopt quickly and repeatedly.

Governments also aggregate demand to create an adoption flywheel. Public sector demand can be bundled into multiyear frameworks anchored on a small number of interoperable, at-scale providers to justify up-front investment.

Governments also catalyze supply by enabling policy and targeted investment. Policymakers can also unlock capacity by accelerating permitting and grid readiness for at-scale infrastructure developers, enabling long-term energy planning for AI loads.

Technology providers: Build capability, localize trust, and partner for legitimacy
Providers span hyperscalers, local cloud providers, neoclouds, data center operators, telecom companies, model developers, and integrators. Leading ecosystems rarely choose either hyperscalers or local providers, instead designing architecture in which each competes and collaborates at the layer where it has advantage.

Our survey results highlight a core tension that providers must navigate. While most enterprise leaders describe sovereign AI as strategically important, sovereignty alone rarely drives the decision to switch vendors—a decision that remains dominated by price, performance, and reliability (Exhibit 3). This does not signal weak demand for sovereign options. Rather, it reflects the way enterprises operationalize risk: Sovereignty matters most for a specific subset of workloads—those involving sensitive data, regulatory exposure, or critical services.

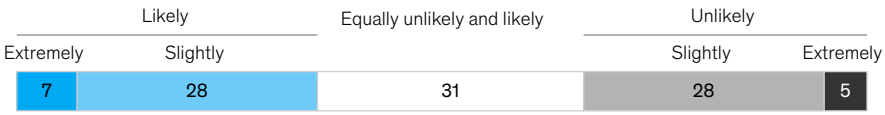
Indeed, sovereign AI offerings are perceived as being 10 to 30 percent more expensive than global alternatives (Exhibit 4). There are cases and instances in which sovereign AI players may be more advantageous than global alternatives, but those players will need to educate customers on these instances and make the case for when a premium is warranted.

For providers, the implication is straightforward. Demand for sovereign AI is real, but it is selective. Sovereignty becomes commercially relevant when it clearly reduces risk or enables deployment in regulated settings and not when it is sold as a blanket feature. In practice, this means that providers succeed when they translate ecosystem-level sovereignty requirements into offerings that are concrete and easy for enterprises to adopt at the workload level, not when asking customers to absorb complexity or uncertainty themselves.

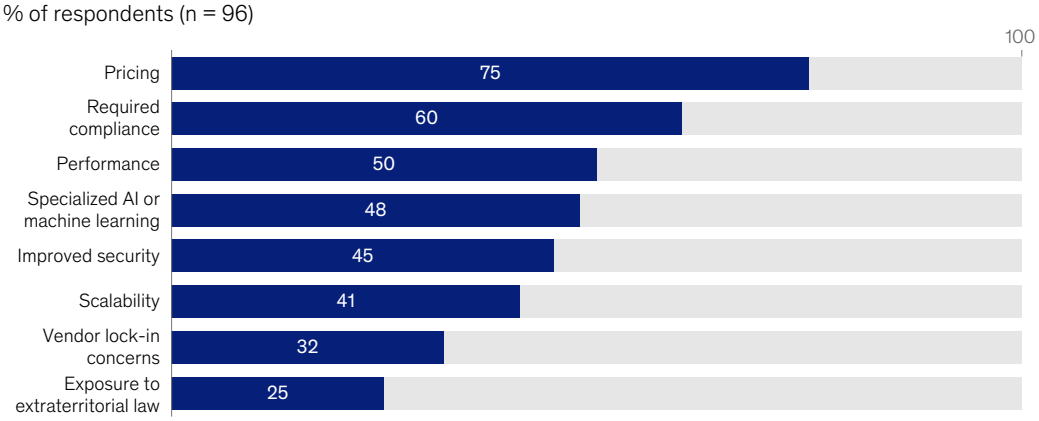
Exhibit 3

About 70 percent of enterprises are open to switching their current AI vendor for better pricing and performance.

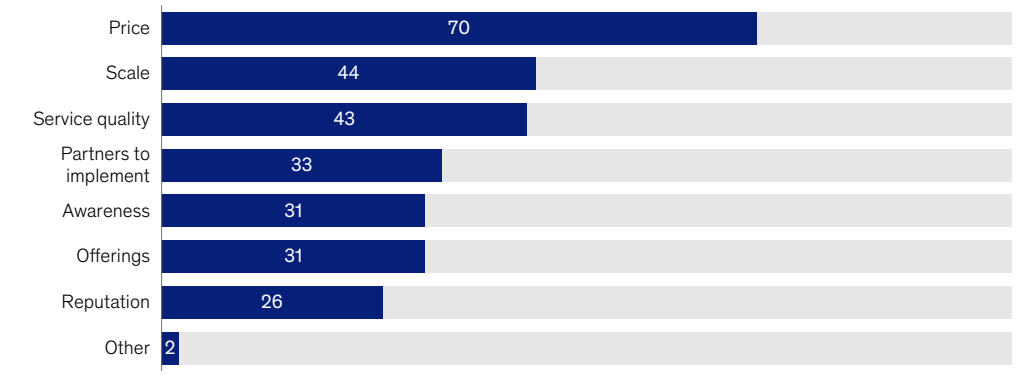
Likelihood to switch or add new vendors,[1] % of respondents (n = 96)



Factors considered when switching from current provider to a sovereign AI provider,[2]
% of respondents (n = 96)



Factors limiting use of sovereign AI provider,[3] % of respondents (n = 96)



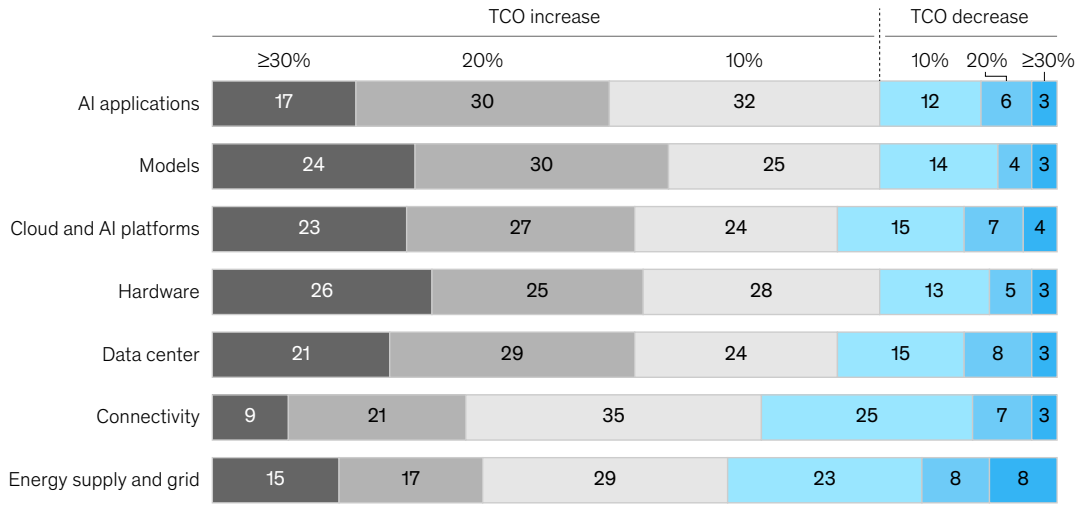
Note: Figures do not sum to 100%, because of rounding.
[1]Question: How likely are you to switch or add new vendors?
[2]Question: What are the factors you would consider when switching from your current provider to a sovereign AI provider? Respondents could select more than one answer.
[3]Question: What is limiting your use of a sovereign AI provider? Respondents could select more than one answer.
Source: McKinsey Sovereign AI CxO global survey, Dec 2025

McKinsey & Company

Exhibit 4

Sovereign AI solutions are perceived as being 10 to 30 percent more expensive across the stack.

Total cost of ownership (TCO) of sovereign solutions, compared with nonsovereign, by tech stack layer,[1] % of respondents (n = 300)



[1]Question: How does TCO compare between sovereign and nonsovereign solutions for each area?
Source: McKinsey Sovereign AI CxO global survey, Dec 2025

McKinsey & Company

Enterprises: Create demand, provide data, industrialize adoption

Enterprises are the demand engines that turn sovereignty into scaled economic value. In leading ecosystems, regulated enterprises and government-owned entities act as anchor tenants that justify ecosystem-wide investments.

Enterprise interest in sovereign AI capabilities is now widespread, but while most have it as part of their road maps for 2026, few have a detailed strategy, action plan, budgets, and workload tiering (Exhibit 5).

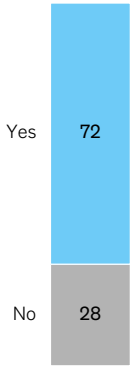
This lack of operational readiness also helps explain why sovereign cloud and AI migrations typically take three to four years (Exhibit 6). These timelines are not driven primarily by technology limitations but instead reflect the organizational work required to move regulated workloads. Sovereign AI adoption is therefore a multiyear transformation, not a simple vendor switch.

At the same time, long migration timelines should not be interpreted as a lack of technical capability. Sovereign and local providers increasingly match global alternatives on service levels, especially lower in the AI stack (Exhibit 7).

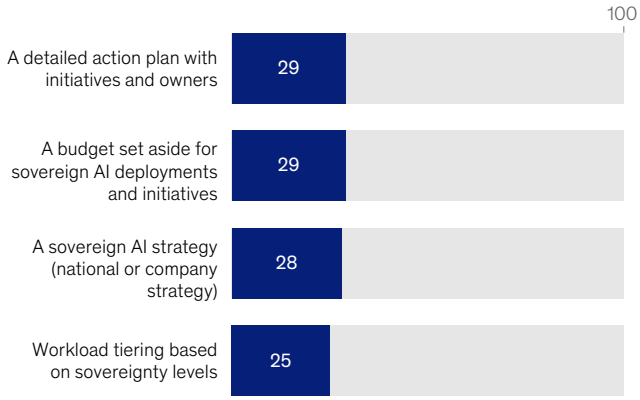
Exhibit 5

Nearly three-quarters of enterprises include sovereign AI as part of their 2026 road map—though few have concrete plans in place.

Sovereign AI is part of road map for 2026,[1] % of respondents (n = 300)



Components included in road map,[2] % of respondents (n = 300)



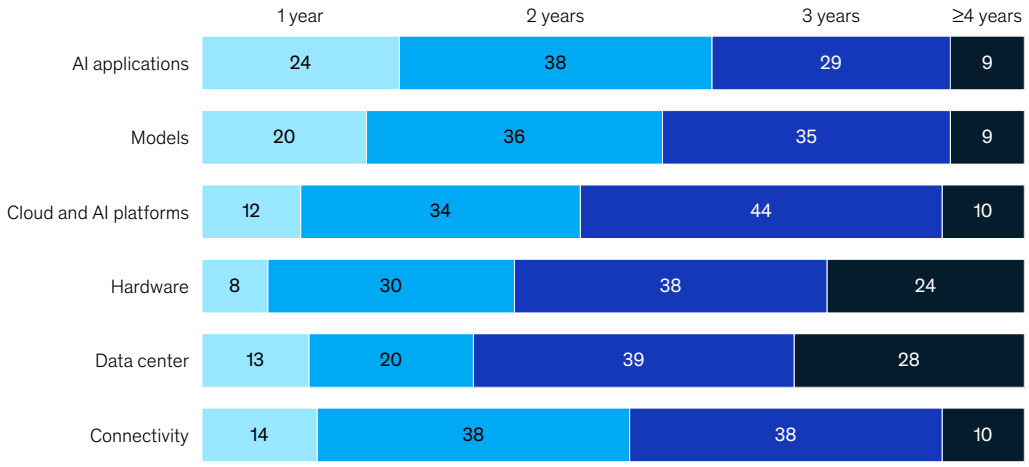
[1]Question: Is sovereign AI part of your road map for 2026?
[2]Quesiton: Which components does your company have? Respondents could select more than one answer.
Source: McKinsey Sovereign AI CxO global survey, Dec 2025

McKinsey & Company

Exhibit 6

Many enterprises expect sovereign AI migrations, especially infrastructure layers, to take three or more years.

Typical timelines for switching to sovereign AI player, by tech stack layer,[1] % of respondents (n = 96)



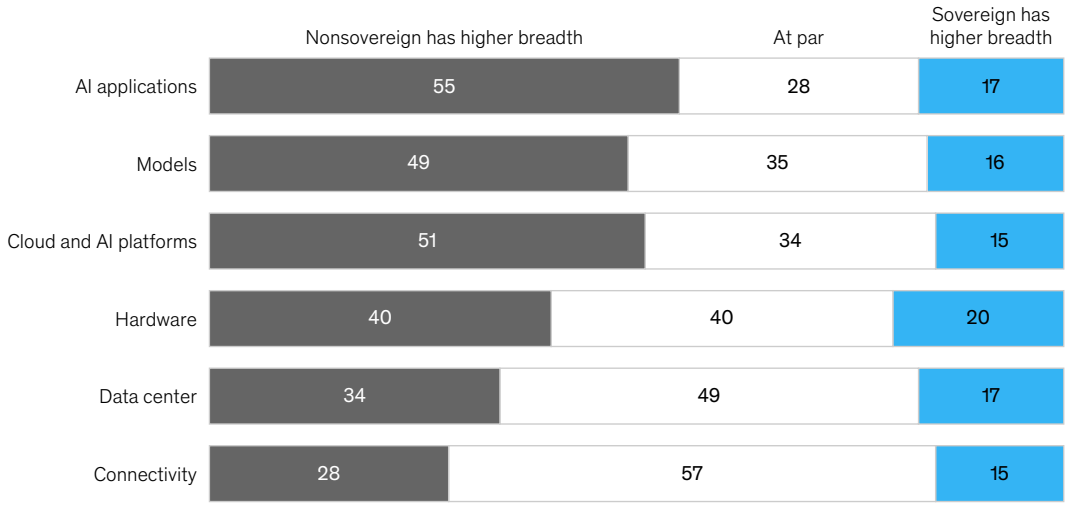
[1]Question: What are the typical timelines for switching to a sovereign AI player across the layers of the stack?
Source: McKinsey Sovereign AI CxO global survey, Dec 2025

McKinsey & Company

Exhibit 7

Enterprises generally see broader service offerings in nonsovereign AI solutions, though many infrastructure layers are viewed as comparable.

Comparison of breadth of services between sovereign and nonsovereign solutions, by tech stack layer,[1] % of respondents (n = 96)



[1]Question: How does breadth of services compare between sovereign and nonsovereign solutions for each area?
Source: McKinsey Sovereign AI CxO global survey, Dec 2025

McKinsey & Company

Sovereign AI migrations are slow not because the technology is immature but because enterprises struggle to decide where sovereignty truly matters and to adapt their operating models accordingly. As a result, providers might consider approaching sovereignty as a portfolio decision rather than as an ideology—segmenting workloads into sovereign, hybrid, and global. This avoids "all or nothing" debates and accelerates time to value while sovereignty deepens over time. Those seeking to capitalize on sovereign demand could also invest in the real bottleneck: data readiness and resolving the operating model. That means building data products with machine learning operations that can span sovereign and nonsovereign environments. Finally, they can help shape the ecosystem—joining sector consortiums, acting as early adopters and reference customers, evolving procurement, and codeveloping domain applications and models with providers and start-ups to accelerate local innovation.

Investors: Provide the capital, manage price risk, and accelerate scale

Sovereign AI is becoming a major investment theme for investors globally, especially among sovereign wealth funds, leading to an expected increase in sovereign AI investment mandates and assets under management.

Sovereign AI spans asset classes: energy, real estate, data centers, connectivity, cloud services, model development, application software, cybersecurity, and integration. Investors matter because they can help organizations bridge the "valley of uncertainty" before utilization is proven, and most expect their investments to increase across the stack, especially higher up the stack.

Across all layers, the most effective investors do two things especially well. First, they back projects with real demand and clear rules, rather than big announcements or speculative build-outs. Second, they help companies grow and exit so local innovation can scale instead of stalling.

Partnership models that outperform

Across markets, several partnership structures can move sovereign AI ecosystems from pilots to scale-up. These models succeed because they align incentives across public and private actors and reduce friction in adoption.

Sovereign AI zones with standardized controls

In this model, integrated environments combine energy, compute, secure connectivity, and compliance controls into a single operating framework. By standardizing security, data residency, and audit requirements up front, these zones reduce onboarding time and enable enterprises and providers to deploy workloads repeatedly rather than through one-off exceptions.

Examples include sovereign cloud zones offered by hyperscalers in Europe and the Middle East as well as national AI or cloud zones that bundle certified infrastructure with preapproved regulatory controls.

Demand aggregation and offtake commitments

These are multiyear frameworks that consolidate demand from the public sector and regulated industry to create predictable workloads. When paired with fast-track procurement and clear scale-up pathways, demand aggregation turns policy intent into bankable utilization.

Examples include the EuroHPC Joint Undertaking, where coordinated public demand underwrites shared supercomputing and AI capacity, with government-led frameworks that anchor early AI workloads in health, defense, or public services.

Joint operating models for sovereign environments

In this model, shared-control structures clearly define who operates infrastructure, who controls access and encryption, and how incidents and compliance are managed. These models offer a hybrid between fully state-run and fully vendor-run environments.

Examples include sovereign cloud joint ventures such as Bleu in France, which combines local operational control with global hyperscaler technology under clearly defined governance.

Model adaptation and data consortiums

Collaborative arrangements can pool data, funding, and demand to develop or fine-tune domain- and language-specific models. Governments typically set governance and evaluation standards while multiple providers compete at the application layer.

Examples include open and semi-open model initiatives such as BLOOM—a large language model created by the BigScience workshop with 176 billion parameters and trained to support 46 natural languages and 13 programming languages—as well as national language or sector-model consortiums.

Blended finance for early layers

Financing structures can combine public capital with private investment to support assets with long payback periods, such as energy, data centers, and foundational platforms. Public participation helps derisk early build-out, with a path to commercial terms as utilization scales.

Examples include public co-investment in national compute hubs and AI factories, which is often paired with long-term offtake commitments from government or regulated industries.

Across these models, the common thread is orchestration. Partnerships outperform when responsibility for aligning incentives, removing friction, and converting sovereign ambition into execution is explicit rather than assumed.

A practical road map: Three waves of ecosystem building

In practice, successful sovereign AI ecosystems tend to emerge through three overlapping waves rather than a single, linear build.

The first wave focuses on establishing the baseline and unlocking early demand. Leaders clarify which workloads truly require sovereign controls, translate those decisions into governance and procurement mechanisms, and launch a small number of lighthouse use cases large enough to justify initial investment. The goal is not completeness but credibility—creating early proof that sovereign environments can operate reliably, securely, and at scale.

The second wave concentrates on scaling shared infrastructure and data ecosystems. With demand signals in place, ecosystems expand compute and energy capacity on bankable terms, industrialize operating models, and invest in sector-specific data products and lawful data-sharing mechanisms. This is where many initiatives falter by attempting to scale infrastructure without first resolving governance, operating model, and talent constraints.

The third wave builds durable advantage and exportable capability. Ecosystems deepen specialization in selected domains, support a competitive provider landscape, and enable start-ups and integrators to scale. At this stage, trusted capabilities become not just domestic enablers but sources of regional or global differentiation.

The most common failure mode is mis-sequencing—investing heavily in shared assets before demand and governance are ready or pursuing global leadership ambitions without the data, adoption, and operating foundations required to sustain them.

Ultimately, sovereign AI is not about full-stack independence. It is an ecosystem play. Those who orchestrate coherent systems—in which sovereignty is applied deliberately at critical control points, and governments, providers, enterprises, and investors align incentives—will turn infrastructure into trusted capabilities and turn trusted capabilities into scaled outcomes.

Find more content like this on the
McKinsey Insights App
Scan • Download • Personalize

Ali Ustun is a senior partner in McKinsey's Doha office; **Arnaud Tournesac** is a partner in the Paris office; **Luca Bennici** is a partner in the Dubai office, where **Justin De Niese** is an associate partner and **Newfel Drahmoune** is a consultant; **Ruben Schaubroeck** is a senior partner in the London office; **Kaavini Takkar** is an associate partner in the Seattle office; and **Melanie Krawina** is an associate partner in the Vienna office.

This article was edited by Larry Kanter, a senior editor in the New York office.

Copyright © 2026 McKinsey & Company. All rights reserved.