Stanford University
Human-Centered
Artificial Intelligence

Issue Brief
HAI Policy & Society
July 2026

The Commercial Landscape of AI Sovereignty Offerings

Caroline Meinhardt, Juan N. Pava, Caroline Yee, James A. Landay

Introduction

"AI sovereignty" has become one of the defining buzzwords of global AI development and governance. From Europe to Southeast Asia, from the Gulf to Latin America, governments are pouring resources into strategies to reduce their dependence on a small number of foreign AI providers and increase domestic control over AI development and deployment. However, the concept remains systematically underspecified even as policy debates intensify. For some, sovereignty means achieving self-sufficiency in AI development — a practically impossible feat — while others define it as gaining more operational control or agency over various components of their AI supply chains and ensuring continued access to AI infrastructure and models. The term is invoked to describe everything from investing in national language model projects to establishing domestic chip manufacturing capabilities to implementing data localization requirements.

Key Takeaways

"AI sovereignty" is now a central concept in AI governance. Despite its ambiguity, governments worldwide are investing heavily in sovereignty strategies to guard against overdependence on foreign AI providers — and a sprawling commercial market is emerging to meet that demand.

We survey commercial AI sovereignty offerings across the AI stack to assess how different actors are designing and marketing sovereignty solutions and to what extent these meaningfully increase operational control and reduce dependencies.

Sovereignty initiatives offered by Nvidia, Microsoft, Google, AWS, and OpenAI provide solutions for increasing domestic control over computing infrastructure, data governance, and localized model deployment. However, they often reconfigure, rather than eliminate, dependence on these companies.

A growing global ecosystem of smaller companies has also oriented products around AI sovereignty, yet few interpret "sovereign" to mean fully domestic: Most solutions are built on top of foreign foundations, suggesting sovereignty is being operationalized as strategic diversification.

The core policy challenge is calibrating interdependence. Decision-makers should avoid treating sovereignty as an end goal and prioritize solutions that expand strategic choice without losing access to frontier capacity — for example, by embracing open-source AI.

Commercial actors have stepped into this conceptual morass. Alongside the policy debates, a sprawling market has taken shape in which private companies around the world are actively selling sovereign AI solutions. U.S. Big Tech is adapting its products and services for government procurement around the world, and smaller companies and startups are developing tailor-made solutions primarily for their home countries. The language of resilience and independence from foreign control now appears on company websites, investor decks, and partnership announcements from companies operating across the AI tech stack. The commercialization of sovereignty rhetoric has moved fast, and the landscape of offerings has become varied and difficult to parse. Whether these offerings deliver on their promises, or whether they simply rebrand existing dependencies to meet the moment, is a question that has increasingly been raised but has not been examined systematically.

This brief aims to clarify that landscape. We survey a range of commercial AI sovereignty offerings by both U.S. Big Tech and smaller organizations in the global AI ecosystem to identify geographic patterns and trends in the types of solutions emerging across the AI technology stack. We examine to what extent these offerings genuinely provide customers with greater agency or meaningfully reduce AI dependencies, and to what extent they may simply reconfigure or even entrench those dependencies elsewhere. In doing so, we aim to provide a more concrete understanding of the global commercial AI sovereignty landscape and add empirical grounding to what remains a largely fragmented and underspecified policy debate. We conclude with recommendations for policymakers and executives navigating this space.

Our review of commercial sovereign AI offerings demonstrates that pursuing sovereignty isn't a zero-sum game. Instead, AI sovereignty should be understood as a state's capacity to act deliberately and make independent decisions concerning the development, deployment, and governance of AI systems that fall along a spectrum of interdependent arrangements. When considering purchasing corporate AI sovereignty solutions and forming partnerships, decision-makers should prioritize expanding strategic choice without losing access to frontier capacity.

The commercialization of sovereignty rhetoric has moved fast, and the landscape of offerings has become varied and difficult to parse.

US Big Tech Offers AI Sovereignty Packages to Governments

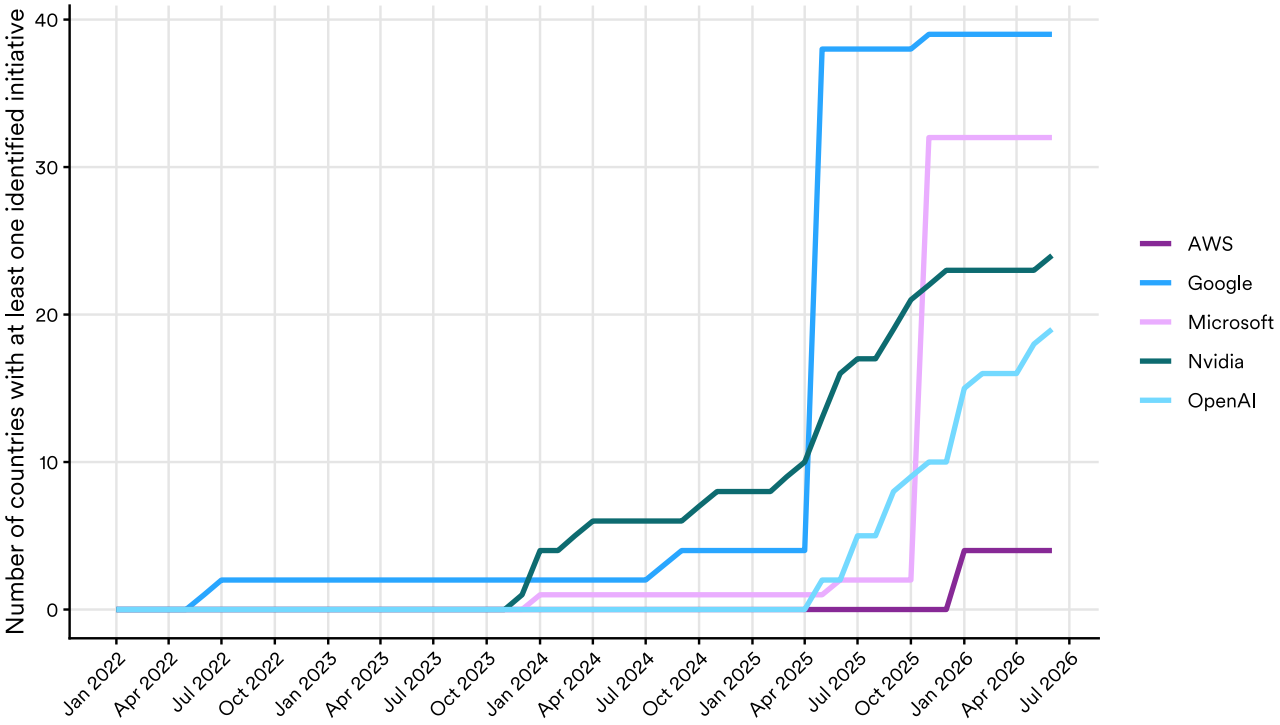
In November 2023, Nvidia's CEO, Jensen Huang, told an AI conference audience in Paris that "every region and every country needs to build their sovereign AI." In the wake of similar remarks delivered at Dubai's World Government Summit and Nvidia's official launch of its AI sovereignty initiatives in February 2024, the concept started gaining global traction. Since then, Huang and his company have become a primary driver and beneficiary of the term "AI sovereignty" and the idea that countries should prioritize national ownership of AI development — with Nvidia promising to provide the tools to make that a reality.

However, Nvidia is not alone — other major U.S. AI companies have also recently increased their commercial offerings around AI sovereignty (see Figure 1). We examined the key global partnerships and initiatives led by U.S. tech giants, reviewing the breadth of their products, services, and initiatives housed under the banner of AI sovereignty (see the Appendix for a detailed description of our methodology). Our resulting landscape analysis is not fully comprehensive. We focused on only a handful of the most dominant companies with extensive international reach (Nvidia, Microsoft, Google, Amazon Web Services, and OpenAI) and examined only their flagship sovereignty-related initiatives (rather than all international partnerships and initiatives). Even so, this selective mapping reveals a varied picture of how U.S. Big Tech is commercializing AI sovereignty concerns.

Nvidia's AI Factories and Full-Stack Sovereignty Play

The cornerstone of most AI sovereignty discussions is ensuring domestic access to the physical infrastructure underpinning AI development: the facilities where data is stored and processed, and the chips that power model training and inference. Without reliable AI infrastructure, many ambitious AI sovereignty strategies remain aspirational. This is why countries are racing to secure computing resources through massive capital investments, despite their well-documented inefficiencies and environmental costs. Yet, even as nations invest in building out this infrastructure, many remain acutely aware that relying on a single U.S. chipmaker carries its own risks — as demonstrated by U.S. export controls that have left some countries on unstable footing or altogether unable to access

Figure 1: Global Expansion of AI Sovereignty Initiatives by Provider, 2022–2026
Source: Stanford HAI, 2026



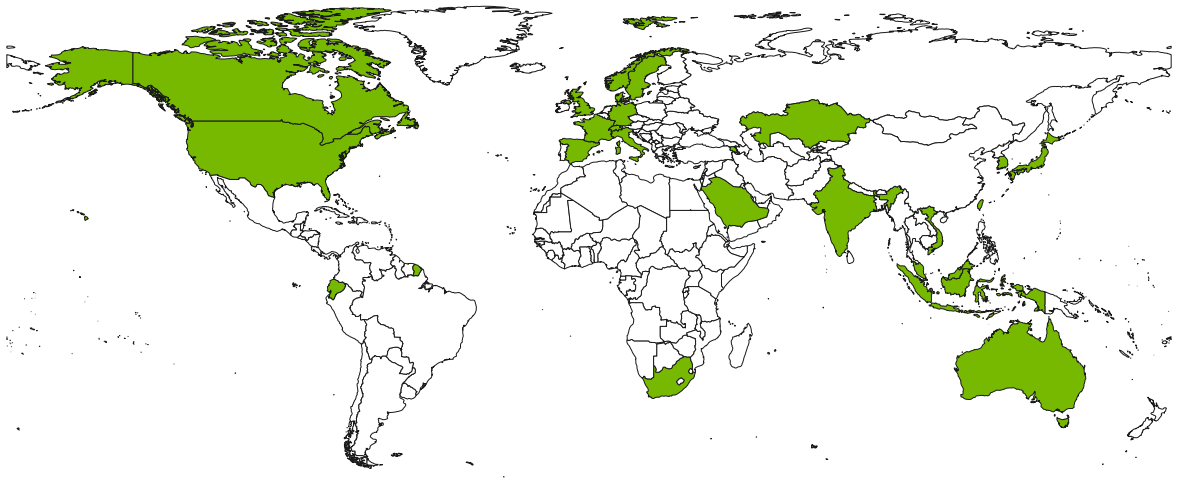
the advanced hardware their sovereign AI ambitions require. Kazakhstan, for example, faced significant procurement delays for its first supercomputer because of U.S. export controls on Nvidia chips.

Nvidia has emerged as the dominant supplier of this foundational layer, having cemented its position as the primary designer of high-performance AI chips. Through its flagship AI factory program, the company has also started positioning itself as providing the "one-stop-shop" infrastructure solution — one that promises states a path to sovereignty at this layer. The program marks a departure from earlier efforts to sell only individual components (e.g., GPUs) and goes beyond the traditional data centers designed for general-purpose data storage and retrieval. Instead, the new AI factory, whether a single data center or a cluster, is a specialized, predesigned and fully stacked turnkey computing environment optimized for AI workloads — one that, in some cases, can be deployed in as little as 90 days. The pitch is that it replaces the need for countries to independently source and integrate computing hardware, data center design, networking, software frameworks, and specialized energy solutions.

The AI factory replaces the need for countries to independently source and integrate computing hardware, data center design, networking, software frameworks, and specialized energy solutions.

Sovereign AI reportedly accounts for roughly $30 billion (or 14%) of Nvidia's revenue and is projected to reach $200 billion in the coming years. With its AI factory program, Nvidia has established a significant footprint across North America, South America, Europe, Asia, and Africa. As of June 2026, we gathered publicly available information on 25 countries — from Ecuador to Armenia to Vietnam — where Nvidia has led or supported the establishment of AI factories (see Figure 2).

Figure 2: Countries With an Nvidia AI Factory
Source: Stanford HAI, 2026



The natural partners for Nvidia in these countries are often, but not always, their domestic telecommunications operators. The companies embed Nvidia's Blackwell GPU clusters directly into their existing telecom data centers and 5G networks. Nvidia's core promise to these nations is that AI workloads will be hosted locally, meaning their data stays within their borders. As a result, AI workloads effectively bypass the legal reach of foreign authorities and are operated by the country's own secure and regulated telecommunications providers. In this way, Nvidia vows to transform national telecom providers into "sovereign infrastructure operators." For example, Qatar's Ooredoo, one of the Middle East's leading telecommunications operators, is collaborating with Nvidia on AI factory buildouts across the region. Similarly, a leading telecom operator in Ecuador, Telconet, has deployed an AI factory to support local use cases like urban safety. This partnership with telcos and governments appears to be part of a larger strategy by Nvidia to broaden its consumer base beyond U.S. hyperscalers, which are not only expected to reduce their capital expenditure on Nvidia over time, but are already actively developing in-house chips that compete directly with Nvidia hardware.

Although Nvidia has centered its sovereignty offerings around the computing layer of the AI stack, it has made meaningful inroads in other parts of the AI stack. As part of an integrated approach, the company has expanded into cloud services in direct competition with the dominant U.S. cloud service providers. And while it has scaled back some of its initial cloud efforts, Nvidia continues to pursue cloud offerings through its industrial AI cloud, launched in partnership with Deutsche Telekom in late 2025. This initiative promises nations a regionally hosted, sovereign cloud platform that pairs Nvidia's hardware with a trusted domestic telco operator, thereby reducing reliance on Google, Microsoft, and Amazon Web Services (AWS) infrastructure. Similarly, at the software and model layer, Nvidia's NeMo framework and open-weight Nemotron model family are promoted as providing governments and enterprises with a full toolkit for building sovereign AI models on local data that reflects their local languages, cultures, and regulatory requirements. The company also trains local talent on how to use its sovereign AI solutions.

Together, these offerings reflect a deliberate expansion of Nvidia's sovereign AI strategy well beyond the hardware layer. Nvidia is positioning itself not just as a hardware developer, but as the full-stack platform through which governments can advance their national AI sovereignty goals.

Microsoft, Google, and AWS Compete to Capture Sovereign Cloud Market

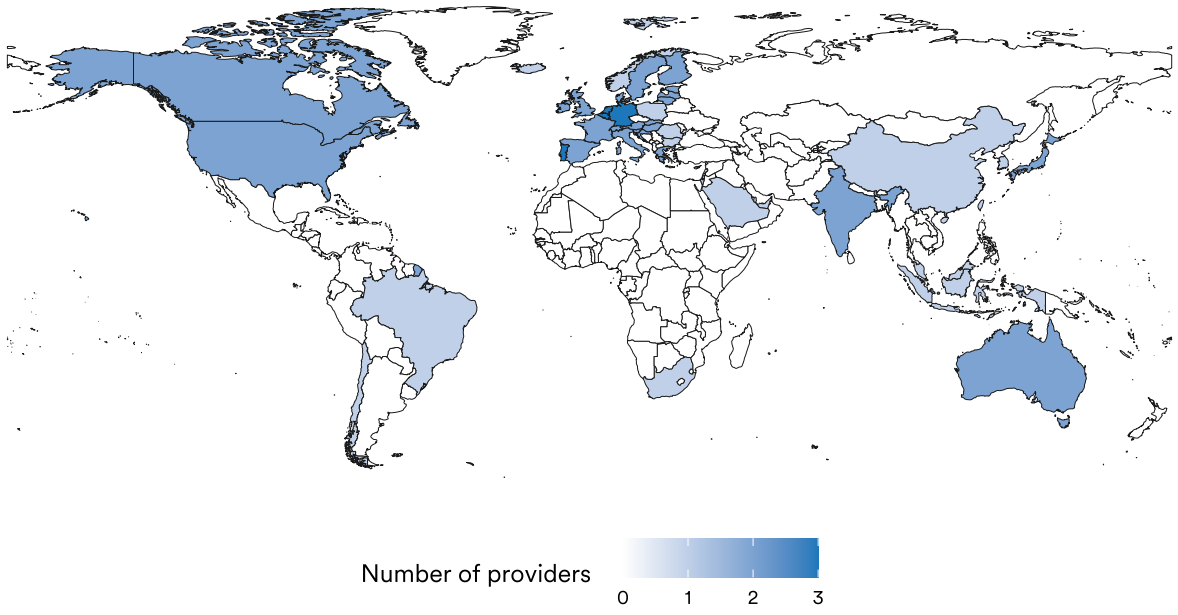
Beyond physical infrastructure, another important layer of the AI stack is cloud infrastructure, meaning the virtual compute, storage, and networking layer that sits on top of physical hardware. AI has turned cloud infrastructure into a strategic dependency, raising government concerns over data processing, operational control, and legal jurisdiction. The dominance of a few large cloud providers — especially in the United States and China — has created fears abroad over vendor lock-in and exposure to foreign surveillance or political leverage. In response, governments are moving to gain more control over their cloud resources either by building separate national or regional cloud environments, or by establishing sovereign cloud environments with the support of global hyperscale platforms. Particularly in Europe, policymakers have shown an increasing willingness to substitute U.S. hyperscaler clouds

and their services with homegrown alternatives. In response, U.S. cloud providers have ramped up their "sovereign" cloud offerings to preserve market access, even though these configurations are more costly and complicated to operate.

Microsoft, Google, and AWS have moved aggressively to capture the sovereign cloud market, and together they offer the largest and most global set of solutions (see Figure 3). While all three hyperscalers have focused on embedding sovereign controls within their public cloud architectures, they diverge in scope and structure. Whereas Microsoft and Google have pursued extensive local third-party partnerships and broadened their geographic reach beyond Europe, AWS recently moved toward developing a sovereign cloud exclusively for Europe that is isolated from its global infrastructure.

U.S. cloud providers have ramped up their "sovereign" cloud offerings to preserve market access, even though these configurations are more costly and complicated to operate.

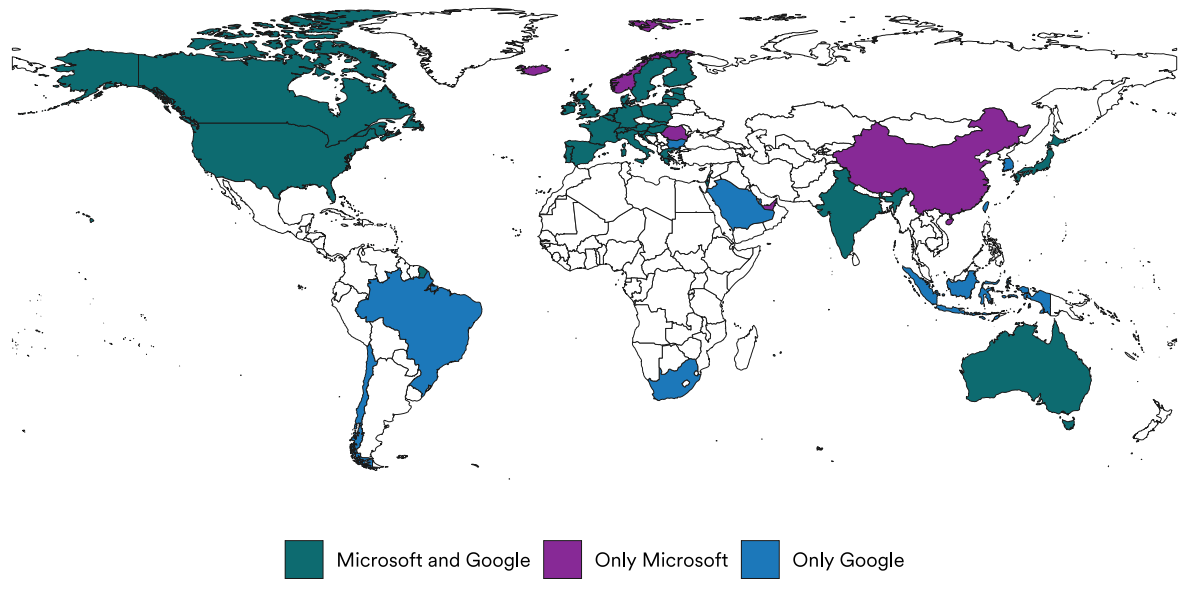
Figure 3: Countries with Sovereign Cloud Initiatives Offered by Microsoft, Google, and AWS
Source: Stanford HAI, 2026



We observe three types of offerings that these cloud hyperscalers provide:

Tier 1. Hyperscaler-operated clouds with technical "sovereignty controls": This approach allows hyperscalers to maintain full ownership and operation of the cloud infrastructure while providing a digital overlay of sophisticated software-based controls often referred to as "sovereignty controls" — such as data boundaries and confidential computing — to ensure that data is hosted locally and sensitive AI workloads are protected. These products target data concerns around the world by providing local data protections and processing requirements beyond traditional data residency commitments (see Figure 4). For example, Microsoft's Sovereign Public Cloud lets customers keep encryption keys outside of Microsoft's cloud and revoke access at any time, making cloud data unreadable to Microsoft while adding protection against extraterritorial access. Microsoft and AWS, by committing to legally contest any order to suspend their services in Europe and provide financial compensation if data is disclosed in violation of the General Data Protection Regulation (GDPR), have implicitly (and, in some cases, explicitly) admitted that these "sovereignty overlays" do not remove jurisdictional exposure to the United States.

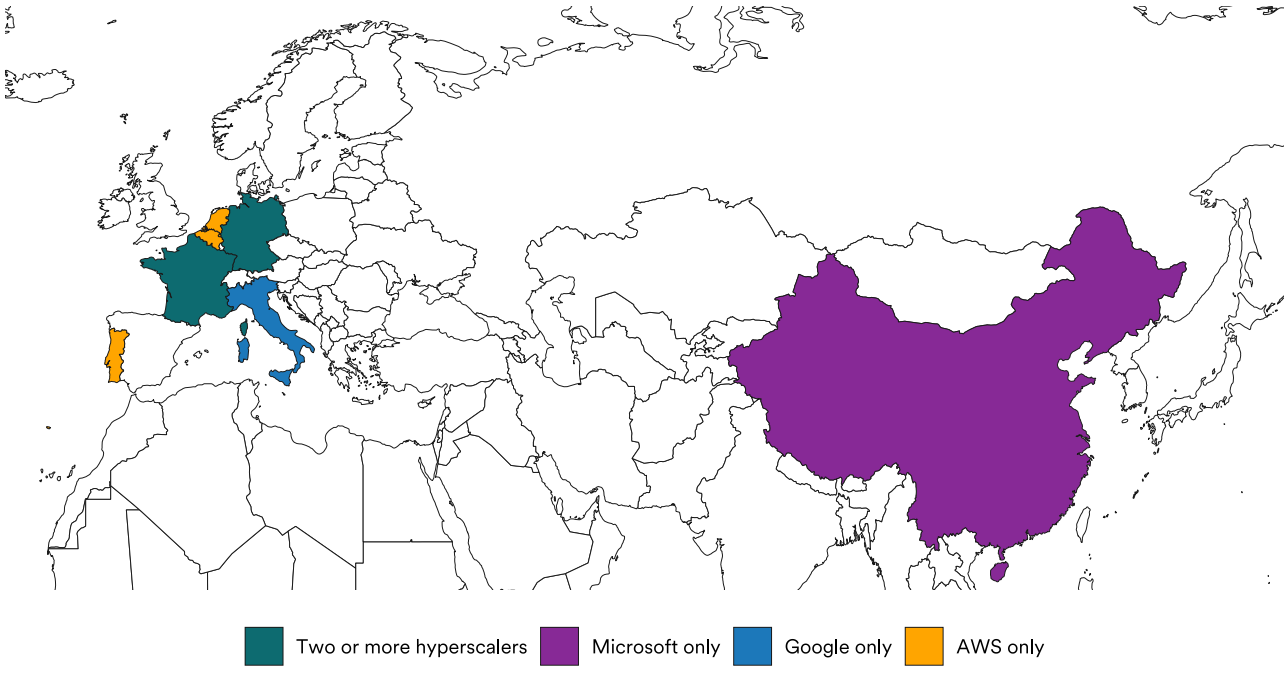
Figure 4: Countries with Added Sovereign Controls for Local Data Processing (Tier 1)*
Source: Stanford HAI, 2026



*AWS offerings exist but are not framed as sovereign cloud products and are therefore excluded

Tier 2. Partner-operated clouds: These involve a local, national entity that operates the cloud infrastructure and manages administrative access using hyperscaler technology, effectively creating a jurisdictional "shield" that blocks extraterritorial data requests from foreign authorities like those under the U.S. CLOUD Act. In France, for instance, S3NS (majority-owned by the French Thales Group, built on Google Cloud) and Bleu (joint venture between French telecom and IT services firms, built on Microsoft Azure) are both pursuing France's highest standard for cloud security (SecNumCloud qualification). Under SecNumCloud, data remains under European jurisdiction and law. Data guardian features can provide an additional layer of protection by ensuring that administrative access is restricted to personnel physically located within the EU data boundary. These partnerships aim to restructure who operates the cloud and under which legal authority, severing jurisdictional hooks that would enable foreign authorities to collect citizen data (see Figure 5).

Figure 5: Countries with Partner-Operated Clouds (Tier 2)
Source: Stanford HAI, 2026

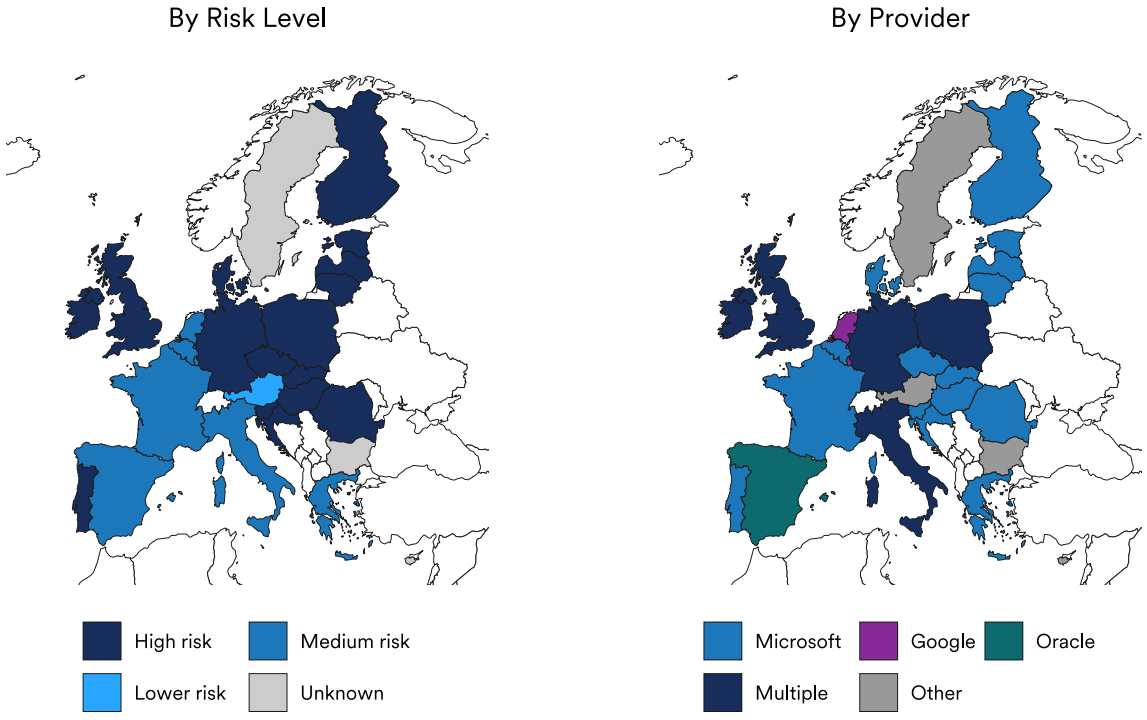


Partners: China (21Vianet), Italy (Telecom Italia / PSN), France (S3NS, Bleu), Germany (T-Systems, Delos). AWS is included via its EU Sovereign Cloud.

More recently, Amazon launched the AWS European Sovereign Cloud, which is both physically and logically separate from its global infrastructure. Unlike other partner-owned models, this cloud is wholly owned by AWS via German subsidiaries, though it is operated exclusively by EU-resident (and soon EU-citizen) staff located within the EU. While this arrangement provides operational autonomy, it offers a narrower jurisdictional shield than Google's and Microsoft's partner-owned model.

Tier 3. Disconnected and air-gapped clouds: These clouds offer a fully disconnected, standalone environment deployed on-premises or within secure facilities and managed by cleared local personnel. Adopted mostly for national security, defense, and critical infrastructure use cases, this approach enables both security and continuity under disruptions or "kill switches" (see Figure 6). These environments (e.g., Google Distributed Cloud, Microsoft Sovereign Private Cloud, AWS Top Secret Regions) are characterized by disconnection from the hyperscaler global cloud infrastructure and, in some cases, from the internet altogether. Many air-gapped deployments remain classified, with notable examples that include Malaysia's DNex-operated systems and AWS's infrastructure for UK intelligence agencies. Moreover, there are questions regarding how truly "air-gapped" these systems are, and whether strategic vulnerabilities persist — with many European defense agencies potentially still exposed to "kill switch" risks due to their deep integration with U.S. hyperscalers like Microsoft, Google, and Oracle.

Figure 6: Exposure to US Hyperscaler dependency on European Defense Contracts (Tier 3)
Source: Future of Technology Institute, 2026



OpenAI's Move from Model Provider to Sovereign AI Infrastructure Partner

While compute and cloud infrastructure determine where AI runs and who controls access to it, the models themselves determine how AI is adopted. For many governments, outsourcing model development raises concerns about relying on systems trained on foreign data, reflecting foreign values, and often underperforming when used in local languages and cultural contexts. Developing models at home is therefore seen as a path to AI outputs that are genuinely aligned with a country's specific culture and use cases, yet this approach is so resource intensive and technically challenging that it remains unfeasible for most countries that seek to reach frontier capabilities.

As one of the world's top frontier model developers, OpenAI is at the core of this debate. Yet because OpenAI models are proprietary and cannot be retrained independently by adopters, the company has said it is not aiming to help countries achieve sovereignty by providing a path so they can own their own models. Instead, the company frames model sovereignty around localized model deployment via its OpenAI for Countries initiative. The value proposition is that partner countries are given the ability to adapt OpenAI's proprietary models to local languages, cultures, and needs, and, in some cases, leverage local data to customize the models to specific use cases of national priority.

For example, several partnership announcements that are part of the OpenAI for Countries initiative stress the goal of customizing ChatGPT-based tools for public sector service integration — echoing earlier deployments in the United States. Meanwhile, a dedicated sub-pillar of the initiative, OpenAI Education for Countries, focuses on working with local ministries of education and universities to embed AI into national education systems to personalize learning, reduce administrative burdens, and build AI-ready workforces.

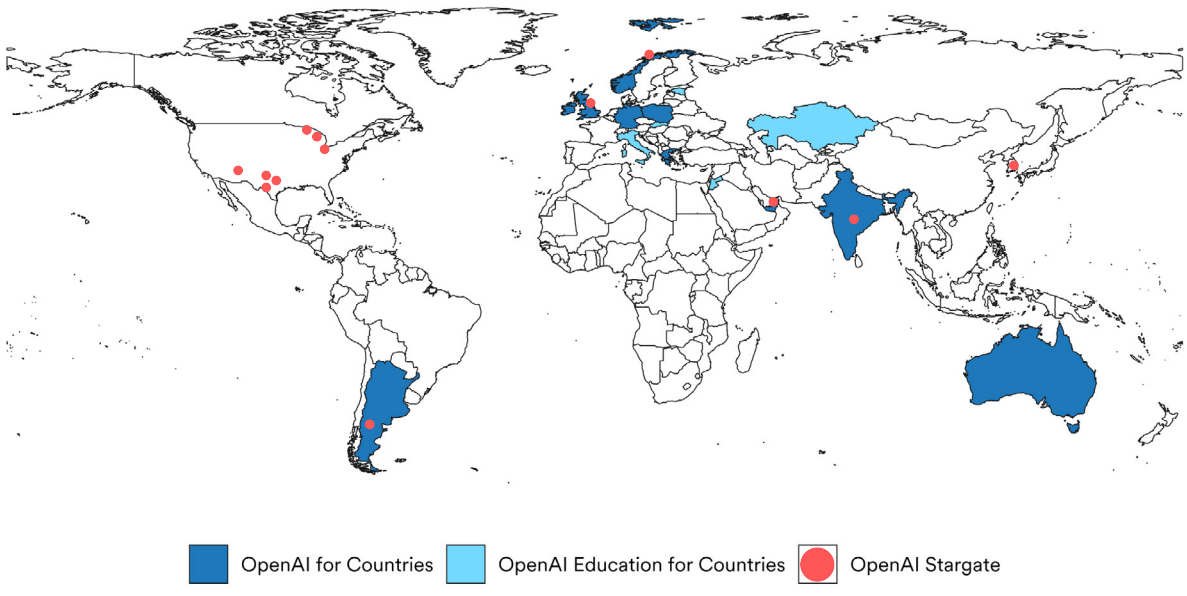
The value proposition is that partner countries are given the ability to adapt OpenAI's proprietary models to local languages, cultures, and needs.

The OpenAI for Countries initiative also represents the company's effort to evolve from model provider to national AI infrastructure partner. Originally launched in May 2025 as the international extension of OpenAI's ambitious U.S. infrastructure project Stargate, the initiative seeks to build out AI infrastructure worldwide under the banner of AI sovereignty. The most visible component involves collaborating with governments and private sector partners (e.g., telecommunication companies and cloud providers) to finance AI data center capacity capable of running OpenAI models domestically. These projects typically involve massive GPU clusters designed for both training and inference workloads, enabling countries to host AI services locally while meeting data residency and regulatory requirements. While some compute deployments occur outside the Stargate framework, most are linked to the project. When partner countries lack capital, they may trade their high-quality domestic data for access, as Iceland did in its partnership with OpenAI for language preservation.

As of June 2026, we retrieved publicly available information on 18 OpenAI for Countries initiatives currently underway (see Figure 7). About half of them involve Stargate, six are exclusively education-focused initiatives, and one (in Australia) involves a non-Stargate computing investment. While the geographic footprint of these partnerships could reflect the explicit framing of the initiative to spread "democratic AI," it is noteworthy that Stargate's first major partner was the UAE, which Freedom House classified as "not free." OpenAI has explained this contradiction by arguing that "partnering with non-democratic governments can help them become more liberal." Additionally, OpenAI recently started to distance itself from some large flagship projects in the United States, the United Kingdom, and Norway — a sign perhaps that they are dialing back on the Stargate-linked infrastructure model or trying to show financial discipline in preparation for their IPO.

In addition to these flagship sovereignty offerings, OpenAI has pursued several other international initiatives that, while not explicitly labeled as pertaining to sovereignty, address related concerns. In Europe, OpenAI launched its EU Economic Blueprint 2.0, an overview of EU AI usage data and a package of initiatives to close Europe's AI adoption gap, including programs to train 20,000 small and medium-sized enterprises with AI skills, a €500,000 youth safety research grant, and a regional adaptation of its OpenAI for Countries called "OpenAI for Europe" in 2026. In both Europe and Asia, OpenAI has introduced data residency services, which allows customers to store and process their data locally, helping them meet domestic data sovereignty requirements when using OpenAI products.

Figure 7: Countries with OpenAI Strategic Partnerships
Source: Stanford HAI, 2026



The Sovereignty Paradox: More Control, More Reliance?

Our review of prominent sovereignty-related commercial offerings by U.S. Big Tech shows that these initiatives have surged globally, with a particularly pronounced footprint in Europe. Notably, these companies are pursuing a variety of strategies and goals as they partner with customers (especially national governments) to help them operationalize the promise of "sovereignty."

Nvidia stands apart in its promise of giving countries full ownership of their AI development and deployment. By providing them with AI sovereignty tools across the entire AI tech stack — from specialized infrastructure (e.g., national telco-operated AI factories) to model training frameworks (e.g., NeMo framework) and open-weight models (e.g., Nemotron models) — Nvidia promises to enable countries to own and operate their AI tech stacks independently.

OpenAI sells access to AI capabilities rather than ownership. Under the banner of sovereignty, they offer localized access to ChatGPT, data residency within a nation's borders, and a customized user experience. The company argues that true model independence is not feasible for most countries and the goal, therefore, should be optimally adapted access. This pitch is a fundamentally different sovereignty offer — and distinctly ideological. Whereas Nvidia frames sovereignty as a neutral infrastructure and cultural capability question, the "OpenAI for Countries" initiative explicitly frames sovereignty as a geopolitical contest between democratic and authoritarian AI. The initiative is presented as a response to Chinese competitors' "authoritarian versions of AI" and is implemented in coordination with the U.S. government.

These initiatives present a fundamental paradox: They promise that countries will own their AI stack, but at the same time they deepen dependencies on U.S. Big Tech.

Meanwhile, the three cloud hyperscalers, Microsoft, Google, and AWS, offer a variety of partnership models and sovereignty controls that aim to equip countries with the tools to secure and independently operate cloud infrastructure. Partnering with these cloud hyperscalers is also an alternative to partnering directly with Nvidia, as governments may find it more efficient to rely on a hyperscaler's bargaining power to negotiate with Nvidia rather than procuring AI factories directly.

Overall, the global growth of these five AI companies' sovereignty initiatives demonstrates that they are tapping into real unmet needs and concerns. Most of these initiatives offer tangible solutions for ensuring that citizens' data, government records, and other important data remain in-country, for example. Many initiatives are also creating genuine first-time access to GPU compute for local developers who previously were "at the back of the line" for computing resources.

However, these initiatives also present a fundamental paradox: They promise that countries will own their AI stack, but at the same time they deepen dependencies on U.S. Big Tech. Adopting Nvidia's

sovereignty offerings to power your entire AI stack may ultimately reinforce vendor-lock in, as migrating to rivals like AMD or Intel would bring high switching costs, including extensive code rewrites. OpenAI's sovereignty offerings may increase localized access to cutting-edge models, but they do not allow for inspecting or modifying underlying systems. Further, given the close ties between "OpenAI for Countries" and the U.S. government, partnering countries risk becoming strategically bound to U.S. interests for access — with no viable exit if OpenAI becomes the sole gateway to frontier AI capacity, a dependency further entrenched by the Stargate project, which is poised to exclusively service its own proprietary models.

For these reasons, many sovereignty initiatives have been criticized for "sovereignty washing." While the tools these companies sell under the banner of sovereignty may genuinely improve purchaser countries' control over certain aspects of AI development and deployment (e.g., telco-operated cloud infrastructure that is compliant with regulations), they also ensure these countries will remain structurally dependent on them for the long term. Full-stack, cross-layer sovereignty offerings in particular may deepen these dependencies.

Because these companies' most viable rivals are also American entities, the underlying geopolitical anxiety — the absence of a non-U.S. alternative of comparable scale — remains fundamentally unaddressed. As such, nations remain subject to the availability of these products and partnerships and at the whim of commercial contracts, U.S. export controls, and extraterritorial data access policy. It is illustrative that, at the time of writing, OpenAI has unilaterally shelved its Stargate UK project, delivering a sharp blow to Britain's sovereign AI ambitions.

Beyond US Big Tech: The Global Sovereignty Ecosystem

There is a wide ecosystem of smaller companies around the world that have begun to develop commercial offerings grounded in AI sovereignty. Many are actively marketing themselves as practical alternatives for anyone, including governments and companies, that seek greater control over their AI systems. This section draws on a non-comprehensive review of companies using AI sovereignty language to describe their products and services, spanning hardware, cloud infrastructure, models, and full-stack platforms.

"Full-Stack" Sovereignty Platforms

We reviewed several companies that position themselves as providers of end-to-end sovereign AI stacks, offering customers a single point of control across every AI layer. In practice, however, the meaning of "full stack" varies considerably depending on the company's core competency.

For example, a UK company called SovereignAI self-describes as "the global champion of sovereign AI systems," providing its customers with ownership "from the soil to the token" — from the physical land and power infrastructure that support data centers to AI model outputs. The company builds and scales data centers that are designed to run intensive AI workloads while keeping them secure. The key selling point is that customers retain full legal and operational control over their data and AI systems, which all stay in the UK on physical infrastructure they control. Another UK company, Nscale, advertises its full-stack AI infrastructure that powers AI systems "from ground to cloud." Nscale rents out GPU compute and offers

cloud services that runs on data centers they own and operate that are located primarily in Europe. They also provide the storage, networking, and management tools needed to train and customize AI models.

Other companies that sell "full-stack" sovereignty platforms in reality focus their offerings on the model and application layers of the tech stack. For instance, Sarvam.ai, advertised as India's "full-stack sovereign AI platform," builds foundation models trained on primarily Indian datasets and sells a suite of enterprise products including consumer-facing voice AI and conversational agents. The promise is that when an Indian citizen uses a voice assistant powered by Sarvam, their voice data is processed on Indian servers under Indian law, in their own language, by a model that understands their dialect. The U.S. company MeetKai promotes itself as "the only end-to-end provider ensuring true sovereignty" through an AI platform that provides an entire AI software suite (reasoning models, agents, APIs, a chat and voice user interface, safety tools, and more) that can be trained on a country's own data and run entirely inside local data centers.

These offerings reflect an attempt to increase cross-layer operational control over AI development and deployment.

Stack Layer-Specific Sovereignty Offerings

Other companies do not claim to address sovereignty issues across the entire stack, choosing instead to target their sovereignty marketing at one specific layer of the tech stack or one component of AI development.

Some of these firms target the sovereign production of AI chips. France's SiPearl designs chips to supply Europe with their own sovereign hardware; Japan's Rapidus is manufacturing cutting-edge chips domestically; and Mastiska is positioning UAE-developed silicon as a sovereign hardware option for the Global South. These initiatives show a growing recognition that even the most locally operated clouds or models ultimately depend on hardware that, for most countries, is sourced abroad. Even countries without a history of semiconductor manufacturing (e.g., India and the UAE) increasingly feel pushed to build chip design capability, even if the road is long and uncertain.

Several other companies are positioning themselves as sovereign alternatives to the main U.S. cloud providers. In Europe, these include regionally certified providers like the French OVHcloud, which was selected as the sovereign cloud provider for the European Central Bank's digital euro project. OVHcloud and other companies like Germany's STACKIT emphasize GDPR compliance and European data residency among their core sovereignty promises. STACKIT goes even further in its sovereignty framing by calling for an independent Europe and positioning itself as the European hyperscaler that can make that happen. In Asia, providers like SIAM.AI (Thailand's first sovereign AI cloud) are similarly framing their offerings around keeping data within national or regional borders. Meanwhile, Core42, the infrastructure arm of Abu Dhabi-based G42, positions itself as the Gulf region's sovereign alternative to U.S. hyperscalers, promising to keep government and enterprise data under national jurisdiction.

At the model level, a growing number of companies are building AI models and tools explicitly trained on local languages and cultural contexts. For example, Widelabs, Awarri, Kakao, QazCode, and Sahabat-AI target Brazilian, Nigerian, South Korean, Kazakh, and Indonesian languages and contexts, respectively. These companies' sovereignty argument rests on

linguistic and cultural specificity; they insist that a model trained elsewhere on different data cannot faithfully preserve a given culture's linguistic nuance, regardless of where it is hosted. The argument extends beyond linguistics: In predominantly English-speaking Australia, the startup Sovereign Australia AI hopes to build a foundation model trained on Australian data and designed for Australian contexts. While many of these offerings (and others by model developers such as France's Mistral and Germany's Aleph Alpha) are released openly with widely available model weights, the dominant approach to localization continues to rely heavily on adapting American-made open-weight foundation models.

Most of these companies appeal to their domestic or regional markets by explicitly positioning themselves as local alternatives or even substitutions for products and services offered by leading U.S. tech companies. However, many also operate as multinationals that market their offerings around the world, which suggests that sovereign AI may be less about technological self-sufficiency than about diversifying foreign dependencies. For example, the UK's Nscale has a presence in the United States, Portugal, and Iceland, with further expansions planned. Cross-border coalitions are also being formed in the name of sovereignty, with the Canadian enterprise-grade AI startup Cohere, for example, acquiring German LLM developer Aleph Alpha in the name of sovereign AI and offering businesses and governments an alternative to U.S. Big Tech. Meanwhile, smaller U.S. tech companies are also selling sovereignty solutions abroad. For instance, MeetKai, which brands itself as "the Sovereign AI Company," has an established presence in Brazil, Pakistan, and Kazakhstan.

A Fragmented Sovereignty Marketplace that Retains Dependency Problems

Our review reveals a colorful landscape of commercial sovereignty offerings across many regions around the world. While our selective review of companies' sovereignty-related products and services does not allow for sweeping conclusions, it offers interesting signals.

Companies differ in what layer(s) of the stack they address with their offerings and how they define their target market. Some, like the UK's Sovereign AI, explicitly orient their offerings to high-sensitivity sectors like government, defense, healthcare, and finance; where data control and jurisdictional security are critical. Others, like Germany's STACKIT, take a more horizontal approach, marketing GDPR-compliant cloud solutions "for companies of all sizes and industries."

Many of the most mature and advanced companies are actively backed by their governments. In these cases, sovereignty is not just a marketing claim but a stated policy objective, with governments directing funding, structuring procurement, and, in some cases, selecting specific companies to build out domestic AI capacity on their behalf.

Sarvam AI, for example, was chosen by India's IT Ministry to develop the country's first indigenous foundational model under the "IndiaAI mission." In Korea, the government selected five elite research teams consisting of public and private partnerships to build foundation models and provided funding, compute, and data center support as part of a national "Sovereign AI Foundation Model" project. In Nigeria, Awarri partnered with the government to develop

N-ATLAS, the country's first open-source multilingual LLM designed to support African languages. These examples demonstrate the symbiotic relationship between governments pursuing AI sovereignty and companies that need funding, resources, and support to develop products that advance their country's sovereignty ambitions.

A through line across these examples is an interpretation of "sovereign" that does not mean fully domestic. All of these companies maintain notable foreign dependencies — whether on Nvidia chip designs, TSMC chip manufacturing, Microsoft cloud partnerships, or LLMs built by foreign labs. The sovereignty claim, despite the scope of what the company is marketing, appears to focus less on the origin of the components and more on where the stack is deployed and under whose legal jurisdiction it operates. For example, Sovereign Australia AI offers compute, data centers, LLMs, and AI applications specifically for Australian markets, yet they still rely on Nvidia's chips.

Finally, it is significant that several companies have emerged in the United States that market themselves as viable alternatives to other, globally dominant U.S. companies. For example, CoreWeave positions itself as a purpose-built alternative to AWS, Microsoft, and Google, arguing that organizations running serious AI workloads should control their specialized infrastructure instead of depending on the Big Three. According to the company website, "The CoreWeave AI Cloud delivers the speed, scale, and reliability that your AI needs to thrive — and that hyperscalers simply can't match." The company recently went public in one of the largest U.S. tech IPOs since 2021, a clear sign that theirs is a model gaining substantial traction. WebAI uses framing that is even more blunt, declaring, "AI belongs to you, not big tech," and arguing that the future of AI will be a constellation of sovereign, specialized models rather than a single, dominant provider.

These examples show that the sovereignty conversation is playing out at the level of the individual enterprise and at the national or jurisdictional concern. Full control of your AI stack means you are not beholden to any one company that could withhold its capabilities at any moment. Anthropic's dispute with the U.S. Department of War over the company's refusal to allow the U.S. military to adopt its models for certain uses strengthened the sovereignty argument these companies had already embraced.

The sovereignty claim, despite the scope of what the company is marketing, appears to focus less on the origin of the components and more on where the stack is deployed and under whose legal jurisdiction it operates.

Conclusion

The AI sovereignty landscape is far more nuanced than dominant narratives suggest. The range of commercial sovereign AI offerings we survey around the world demonstrate that, in practice, pursuing AI sovereignty does not mean making a binary choice between dependence and independence — it involves a series of decisions that fall along a spectrum of interdependent arrangements.

The sovereign AI products and initiatives marketed by U.S. Big Tech around the world address real and urgent government concerns about control, resilience, and autonomy of their AI supply chains. These offerings promise increased domestic operational control over computing infrastructure, access to air-gapped cloud computing environments, and the deployment of localized AI models. However, they more often reconfigure, rather than fully eliminate, AI supply chain dependencies. In some cases, they may even entrench them more deeply. Cross-stack sovereignty solutions (e.g., Nvidia's AI factories), in particular, demonstrate this dilemma: They may promise greater control and integration across different layers of the AI tech stack, but they can also tighten long-term vendor lock-in, reducing interoperability and hardening reliance.

Local and regional alternatives to U.S. AI solutions and components do not fully resolve these dependencies either. They provide countries with more optionality and often more operational control, better regulatory and cultural alignment, and long-term resilience in certain aspects. Yet, these providers usually rely on the same foundational inputs (e.g., AI chips and cloud infrastructure) that are predominantly offered by U.S. Big Tech, reinforcing dependence on a narrow set of upstream providers.

The core policy challenge of AI sovereignty is not eliminating dependence but calibrating interdependence.

Companies appear to be betting that countries will not choose between "buying" sovereignty solutions from global incumbents or "building" it domestically, but that they will do both. While AI sovereignty narratives among government leaders center on reducing dependence on U.S. AI developers and increasing domestic control over AI development, in practice, their partnerships with different companies offering AI sovereignty solutions indicate that many increasingly accept the reality that certain dependencies will endure.

The core policy challenge of AI sovereignty is not eliminating dependence but calibrating interdependence. Sovereignty should be viewed as the capacity to shape and negotiate dependencies, not avoid them entirely. Policymakers and industry leaders should avoid treating sovereignty as an end goal and prioritize solutions that expand strategic choice. When purchasing AI hardware and software, or when forming partnerships with full-stack AI providers, governments and industry will need to accept certain dependencies on U.S. providers and developers. They should also consider alternate pathways for reducing dependency. Open-source AI solutions represent one common and meaningful option for lowering dependency specifically at the model layer of the stack. Open-

weight models (models released with access to the weights that can be run, fine-tuned, and inspected independently from the developer) and even more so fully open-source AI (that provides additional access to training data, code, architecture, and licensing rights), can offer countries greater flexibility to adapt models to local languages and regulatory requirements, while also enabling independent model hosting and adaptation, thereby avoiding complete vendor lock-in. Importantly, sovereignty strategies that do not consider the role of open-source AI risk normalizing fragmentation and political overreach. Yet, once again, adapting and deploying open-weight models or fully open-source systems at scale typically also requires proprietary hardware and cloud infrastructure, meaning they reduce rather than eliminate upstream dependencies.

Policymakers must be acutely aware of persistent, underlying dependencies and how commercial sovereignty offerings, especially those marketed as delivering "full-stack sovereignty," merely recalibrate rather than eliminate dependencies. The policy focus should be on building targeted domestic capacity where critical while leveraging external capabilities where efficient.

Appendix

Data collection methodology

This analysis draws on a structured review of publicly available information to identify country-level AI sovereignty initiatives across major technology providers. The objective was to capture only those initiatives that are explicitly framed in terms of sovereignty, rather than general-purpose infrastructure or commercial deployments.

For Nvidia, initiatives were included only when the company or its official partners explicitly referred to deployments as "AI factories." References to generic GPU clusters, data centers, or AI infrastructure were not considered sufficient for inclusion. Identification relied primarily on official announcements, company blogs, and partner communications, with additional corroboration from credible reporting where necessary.

For OpenAI, inclusion required explicit participation in one of the following initiatives: OpenAI for Countries, OpenAI for Europe, or OpenAI Education for Countries. Coding relied on official OpenAI announcements where available, supplemented by credible secondary reporting.

Cloud-related sovereignty initiatives were identified by two criteria that reflect different levels of control.

- "Tier 1" initiatives include cases where providers offer *enforceable constraints on data processing within a defined jurisdiction*, such as region-specific data boundaries or in-country processing guarantees. Inclusion required evidence that these controls extend beyond storage to encompass data processing and, where relevant, administrative access, and that they are presented as addressing sovereignty concerns. Google's Data Boundary configurations and Microsoft's newer localized inference commitments were included in this category because they explicitly frame data-processing location as a sovereignty-relevant feature. By contrast, AWS's regional infrastructure and data residency capabilities were not included in this tier, as they are not marketed as sovereignty-oriented offerings within the period analyzed.
- "Tier 2" initiatives capture more advanced forms of operational sovereignty, defined by *the presence of locally anchored entities responsible for operating cloud infrastructure and/or managing administrative control*. Inclusion required evidence that a domestic legal entity (e.g., joint venture, partner, or nationally governed operator) plays a substantive role in governance and access control, and that the arrangement is framed in terms of sovereignty or jurisdictional insulation. This includes partner-operated cloud models developed by Google and Microsoft, as well as AWS's European Sovereign Cloud, which, while not partner-operated, reflects a comparable effort to localize operational control within a specific jurisdiction.
- Finally, the analysis recognizes a third category of sovereignty-oriented systems, which are fully disconnected or air-gapped environments, but does not systematically map these geographically due to limited public disclosure.

Across providers, the identification process prioritized official documentation such as press releases, technical blogs, and partner announcements as primary sources, supplemented by credible secondary reporting when needed.

Stanford University's Institute for Human-Centered Artificial Intelligence (HAI) applies rigorous analysis and research to pressing policy questions on artificial intelligence. A pillar of HAI is to inform policymakers, industry leaders, and civil society by disseminating scholarship to a wide audience. HAI is a nonpartisan research institute, representing a range of voices.

The views expressed in this policy brief reflect the views of the authors. For further information, please contact **HAI-Policy@stanford.edu.**

Caroline Meinhardt is the policy research manager at the Stanford Institute for Human-Centered Artificial Intelligence (HAI).

Juan N. Pava is a research fellow in the Tech Ethics & Policy Rising Scholars Program at Stanford University's McCoy Family Center for Ethics in Society.

Caroline Yee is a research fellow in the Tech Ethics & Policy Rising Scholars Program at Stanford University's McCoy Family Center for Ethics in Society.

James A. Landay is the Denning Director of Stanford HAI, Professor of Computer Science, and the Anand Rajaraman and Venky Harinarayan Professor in the School of Engineering at Stanford University.

Disclaimer

Stanford HAI has received financial support from several companies referenced in this brief, including Nvidia, Google, Microsoft, and AWS. James Landay serves as a consultant to Google.

Consistent with Stanford HAI's fundraising and research policies, no outside entity sets the Institute's research agenda or determines the outcome or dissemination of its research. The authors conducted this research independently and retained full authority over the analysis, findings, recommendations, and editorial decisions presented in this brief.

Stanford University
Human-Centered
Artificial Intelligence

Stanford HAI: 353 Jane Stanford Way, Stanford CA 94305-5008

T 650.725.4537 **F** 650.123.4567 **E** HAI-Policy@stanford.edu **hai.stanford.edu**