

Considerations regarding Sovereign AI and National AI Policy

By David L. Shrier^{1,3}, Ayisha Piotti², Alex Pentland^{3,4} & A. Aldo Faisal^{1,5,6}

¹ Imperial College London, ² ETH Zurich, ³ Massachusetts Institute of Technology, ⁴ Stanford University, ⁵ Universität Bayreuth, ⁶ Alan Turing Institute

This paper was made possible through a collaboration with Payments Network Malaysia Sdn Bhd (Paynet). The authors would like to gratefully acknowledge the assistance of Farhan Ahmad and Endry Lim in crafting and revising this document, as well as research assistance from Adele Jashari and Gemma Bagchi and technical contributions from Yves-Alexandre de Montjoye.

March 2025

Contact:
david.shrier@imperial.ac.uk
aldo.faisal@imperial.ac.uk

Executive Summary

The rapid evolution of artificial intelligence (AI) is reshaping global power dynamics, prompting countries to reassess their technological strategies and reduce dependency on foreign systems. As AI becomes increasingly integral to national security, economic prosperity, and cultural identity, nations are prioritizing **sovereign AI**, the development and control of AI capabilities on their own terms. This growing trend stems from the desire to maintain autonomy over critical technologies, reduce vulnerabilities to foreign influence, and ensure that AI systems align with national values and interests.

Nations are driven by several key motivations in pursuing sovereign AI. Politically, there is a clear need to ensure national security by controlling critical infrastructure and avoiding risks associated with reliance on foreign technologies, including potential data breaches and cyber threats. Economically, AI offers profound opportunities for innovation and growth, allowing countries to tailor technological advancements to their unique economic needs, protect intellectual property, and boost global competitiveness. At the same time, culturally, sovereign AI offers the opportunity to ensure that AI systems reflect a nation's social values, ethics, and societal norms, particularly in areas where cultural context plays a crucial role, such as language processing.

However, the pursuit of sovereign AI is not without significant challenges. The resources required for developing and sustaining independent AI capabilities go far beyond mere technological investment. Nations must evaluate their access to critical resources like energy, water, and hardware, as well as their ability to develop or acquire the necessary infrastructure. Additionally, there is the challenge of cultivating a skilled workforce capable of advancing AI research, developing applications, and managing ethical concerns, all of which are essential to the long-term sustainability of sovereign AI.

Governance also plays a critical role in ensuring that AI technologies serve the public good while adhering to national and international standards. As nations push for greater control over AI systems, there is a need for regulatory frameworks that promote transparency, accountability, and fairness, while also allowing for innovation. Balancing the pursuit of AI autonomy with global cooperation will be essential for navigating the complex legal and ethical landscape that surrounds AI technologies.

Executive Summary

To navigate these challenges, countries can pursue a range of strategic approaches that fit their national priorities. Some may choose to prioritize sovereignty in specific sectors, such as defence, healthcare, or finance, where national security and economic stability are paramount. Others may adopt hybrid models that integrate sovereign AI capabilities with international collaboration in less sensitive areas. The development of open-source AI systems and decentralised models also offers potential pathways for enhancing control and security while mitigating risks associated with centralization.

Sovereign AI is not a one-size-fits-all approach, and nations must navigate a spectrum of strategic choices that balance independence with global integration. While full AI self-sufficiency may be an ideal for some, in reality, most governments will need to adopt hybrid models that leverage regional alliances, federated AI governance, and cost-efficient AI infrastructure sharing. Emerging approaches, such as the use of pre-trained models like DeepSeek, demonstrate that AI sovereignty can be achieved without prohibitive infrastructure costs. Similarly, international Sovereign AI partnerships can provide access to compute resources, research collaborations, and shared regulatory frameworks without compromising national control. We are proposing a new framework for combining forces, exploiting cooperation opportunities while maintaining strategic independence by forging a federation of sovereign AI initiatives. To this end we have established Sovereign-AI.org that acts both as a register and forum of sovereign initiatives and their resources, enabling concerted actions in this space.

In summary, this paper examines how nations can structure their AI sovereignty strategies in ways that align with their economic capacity, technical expertise, and security priorities, while ensuring interoperability in an increasingly interconnected AI landscape.

The Case for Sovereign AI – Definition and Strategic Motivations

Sovereign AI refers to a nation's ability to develop, control, and regulate artificial intelligence systems independently, ensuring that these technologies align with national security, economic interests, and ethical values. The push for AI sovereignty is driven by the increasing reliance on AI in critical infrastructure, cybersecurity, economic growth, and geopolitical competition. Governments recognize that foreign dependence on AI technology providers—many of whom are concentrated in a few dominant economies—poses risks to data security, economic competitiveness, and national strategic autonomy. However, sovereign AI is not a binary goal; it exists on a spectrum of strategic choices, where nations must balance self-reliance with global AI interoperability.

Executive Summary

The motivation for sovereign AI stems from several key factors. Economic independence and industrial competitiveness play a central role, as nations with strong AI capabilities gain a strategic advantage in automation, financial systems, and innovation-driven industries. Countries investing in sovereign AI seek to retain the economic benefits of AI domestically rather than relying on foreign AI solutions that extract value without reinvesting in local ecosystems. For example, South Korea's National AI Strategy and Canada's Pan-Canadian AI Strategy both emphasize building AI-driven economies while ensuring national control over research and data.

Security is another major driver. AI is deeply embedded in national defense, cybersecurity, and intelligence operations, and relying on external providers increases vulnerabilities to cyber threats, algorithmic bias, and surveillance risks. The U.S. Department of Defense's Joint Artificial Intelligence Center (JAIC) exemplifies a national effort to develop AI capabilities internally to minimize foreign dependencies. Similarly, China's "Intelligentized Warfare" doctrine integrates AI across military and cyber operations, ensuring that AI remains a strategic asset under national control.

Regulatory and ethical concerns further reinforce the need for sovereign AI. Nations want to ensure that AI development aligns with their regulatory frameworks, particularly regarding data privacy, transparency, and human rights. The European Union's AI Act is a prime example of a regional approach to AI sovereignty, establishing clear rules that reflect European values while reducing reliance on external AI governance models. In contrast, China's AI regulatory system prioritizes state oversight and algorithmic control, reflecting its governance priorities. Countries that lack sovereign AI risk being subject to the regulatory decisions of AI-dominant nations, which may not align with their own legal or ethical standards.

However, full AI sovereignty is not realistic or necessary for all nations. The high costs of compute infrastructure, semiconductor access, and AI expertise make full self-sufficiency impractical for most governments. Instead, many nations pursue a hybrid model, where sovereign AI is developed for critical sectors like defense, finance, and healthcare, while global AI partnerships remain essential for commercial and research applications. Emerging models like DeepSeek demonstrate how nations can leverage cost-efficient AI technologies without excessive infrastructure investment, allowing for partial sovereignty rather than complete independence.

Sovereign AI is ultimately a strategic decision rather than a universal necessity. Some nations will pursue full independence, others will adopt a hybrid AI sovereignty model, and many will participate in federated AI collaborations to pool resources while maintaining autonomy. The future of sovereign AI will depend on each nation's ability to balance self-reliance with global AI integration, ensuring that sovereignty does not come at the cost of innovation, efficiency, or international cooperation.




Altogether, the motivation for sovereign AI reflects a desire for greater self-reliance and the ability to shape technology in alignment with national interests and values. As AI continues to reshape the world, this pursuit represents a strategic step toward ensuring that nations retain control over their technological futures. These include:



AI is increasingly deployed in cyber defence, military intelligence, and national security operations. Sovereign AI ensures that critical defence systems remain independent of foreign control, mitigating risks related to cyber espionage, algorithmic bias, and AI weaponization.

For instance, the U.S. Department of Defense's Joint Artificial Intelligence Centre (JAIC) has focused on integrating AI across military applications. DOD strategies state that AI will transform the character of warfare, and failure to adopt AI technology could hinder the capability of war fighters to defend the nation. DOD is making organizational changes and investing billions of dollars to incorporate AI technology, such as establishing the Joint AI Centre to accelerate the delivery of AI thus minimising reliance on external vendors. Similarly, China's "Intelligentised Warfare" doctrine seeks to embed AI across its defence strategy, emphasizing the use of autonomous systems, cyber AI, and AI-enhanced decision-making.





Nations that fail to invest in sovereign AI for defence purposes risk being dependent on foreign AI suppliers, potentially exposing military and intelligence operations to vulnerabilities.

According to an IDC report, Artificial Intelligence Will Contribute \$19.9 Trillion to the Global Economy through 2030 and Drive 3.5% of Global GDP in 2030. Countries that develop sovereign AI capabilities will be positioned to lead in AI-driven sectors such as finance, healthcare, autonomous systems, and manufacturing.

Key examples of national AI economic strategies include:

- **National Strategy for Artificial Intelligence of Korea (2019)**, outlines a comprehensive plan to position South Korea as a leading nation in the AI era. The strategy emphasises regulatory flexibility, comprehensive support for AI startups, and the creation of a skilled workforce, aiming to position Korea at the forefront of the global AI landscape.
- Through the **Pan-Canadian Artificial Intelligence Strategy**, the Government of Canada is investing in efforts to drive the adoption of artificial intelligence across Canada's economy and society. The second phase of the strategy bridges Canada's world-class talent and research capacity with programs to enable commercialization and adoption to help ensure that Canadian ideas and knowledge are mobilized and commercialized here at home
- The European Union has outlined plans to invest €1 billion annually in artificial intelligence (AI) through its **Horizon Europe** and **Digital Europe** programmes. This initiative aims to mobilize additional investments from the private sector and Member States to reach a total annual investment of €20 billion over the course of the digital decade. Economic AI sovereignty also reduces dependency on foreign AI infrastructure, cloud services, and computing resources, which are currently dominated by U.S.-based firms like Amazon Web Services (AWS), Microsoft Azure, and Google Cloud.

Data sovereignty is an essential component of AI governance, ensuring that AI models comply with national privacy regulations and protect sensitive information from foreign surveillance.

For instance:

- The European Union's General Data Protection Regulation (GDPR) mandates strict data localization and user privacy protections.
- China's Data Security Law (DSL) requires that sensitive national data be stored domestically.
- India's Personal Data Protection Bill seeks to establish national data governance frameworks for AI applications.

The push for sovereign AI is, therefore, closely linked to the ability to enforce domestic AI and data protection laws without reliance on foreign entities.



Overview of Technology:

What is a GPT?

A GPT, or Generative Pretrained Transformer, is an advanced artificial intelligence model designed to understand and generate human-like text. It excels in various tasks, including writing, answering questions, summarizing content, translating languages, and even creating code. GPT models achieve this by analyzing vast amounts of text data during training and learning complex patterns in language, context, and structure.

The development of GPTs is part of a broader evolution in AI that has unfolded over more than 80 years. Key milestones in this journey include the advent of machine learning, where computers began to "learn" from data without explicit programming, and deep learning, which uses neural networks to model intricate patterns.

Two revolutionary breakthroughs that accelerated progress in natural language processing (NLP)

The Open-Sourcing of TensorFlow in 2015, a widely used deep learning framework made advanced AI research more accessible, enabling rapid innovation and collaboration across industries. Then the Introduction of Transformer Architecture in 2017 introduced a new way for AI models to process information, allowing them to focus on different parts of input data, like words in a sentence, more effectively. Transformers form the foundation of GPT and similar models, significantly improving their ability to understand and generate coherent, context-aware text.

Together, these advancements have propelled GPTs to the forefront of AI, making them versatile tools with applications in education, research, creative industries, and beyond.





Key Use Cases of AI Across Industries

Artificial intelligence is transforming various industries by introducing innovative solutions and enhancing efficiency. Here are some key use cases across major sectors:

Financial Services

AI is revolutionizing the financial industry through advanced capabilities such as fraud detection, where machine learning algorithms analyze patterns to identify suspicious activities. It also enables personalized banking experiences, offering tailored financial advice based on individual customer behavior and preferences. Algorithmic trading leverages AI to execute high-speed, data-driven market transactions, while robo-advisors provide accessible and automated investment guidance.

Healthcare

In the healthcare sector, AI is making significant strides in improving patient outcomes and operational efficiency. It assists in diagnosing diseases with precision, supports preventative care by predicting health risks, and accelerates drug discovery by analyzing complex biological data. Additionally, AI-powered tools streamline medical transcription, reducing administrative burdens for healthcare professionals and improving documentation accuracy.

Education

AI is reshaping education by creating personalized learning experiences that adapt to individual student needs and abilities. Intelligent tutoring systems provide tailored instruction and feedback, while AI-driven analytics help educators identify learning gaps and optimize teaching strategies. Virtual assistants and chatbots support students with instant answers to queries, and AI tools facilitate content creation and management, making education more accessible and engaging.

Strategic Building Blocks:

As sovereign AI strategies evolve, it is crucial to examine the foundational pillars that enable their success. From infrastructure and regulatory mechanisms to workforce development and cybersecurity resilience, nations must develop robust frameworks that ensure both innovation and security. The following sections dissect these building blocks, offering insights into how different models of AI sovereignty can be pursued.

To successfully implement sovereign AI capabilities, countries must address a range of strategic elements.

Governance and
Regulatory
Framework

Pro-Innovation
AI Policy

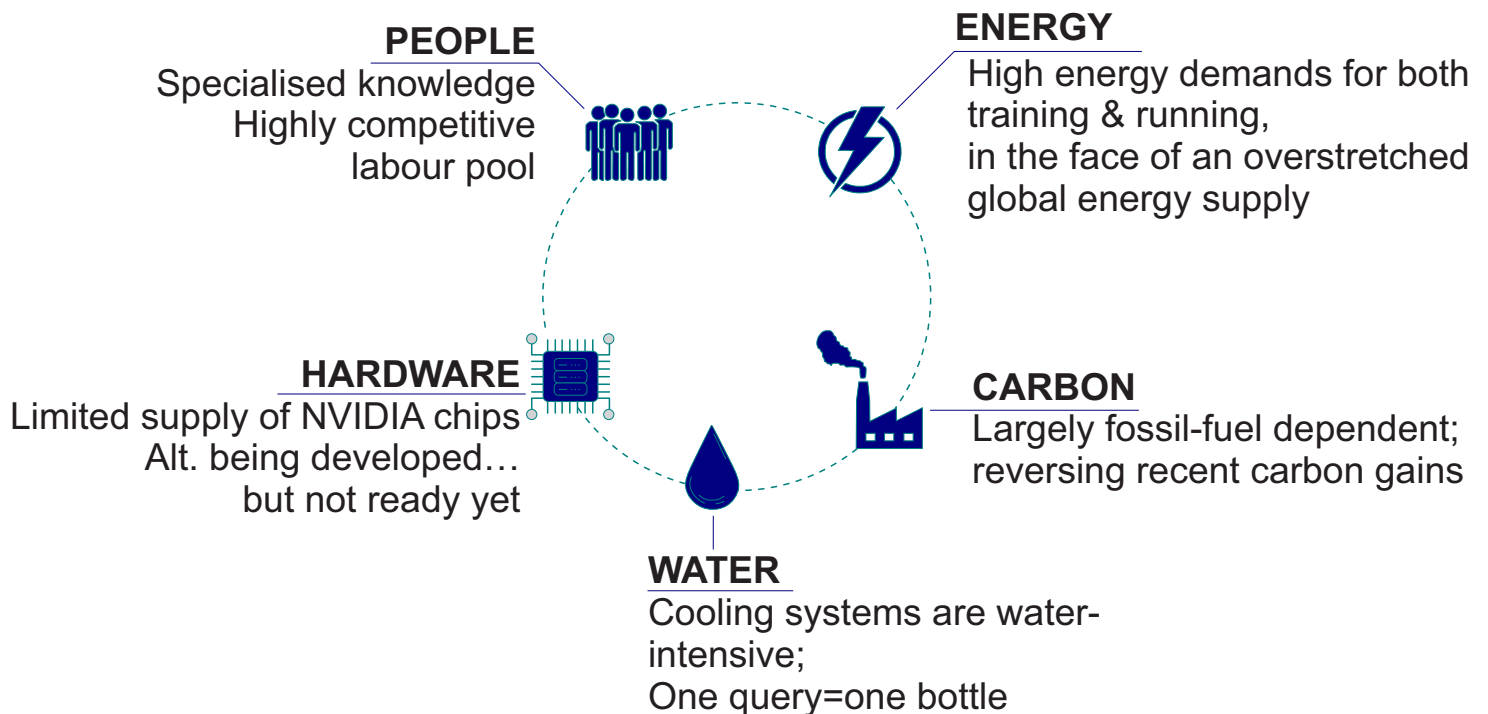
National AI Security
& Safety

Data Sovereignty,
Authenticity, Consent,
and Provenance

Digital Infrastructure
Requirements

Socio-technic Systems Considerations

There are 5 key building blocks in order to implement Sovereign AI capabilities, which have a number of embedded challenges



Determining Sovereign Value Proposition

Policymakers play a critical role in evaluating their nation's infrastructure and capabilities to determine the feasibility and benefits of developing sovereign AI. This assessment involves analyzing key factors necessary for the creation and sustainability of such systems. By systematically evaluating these factors, nations can identify their strengths, address gaps, and make informed decisions about pursuing sovereign AI initiatives. This approach ensures that investments are aligned with national priorities and capabilities, paving the way for effective and sustainable AI development.

These elements are critical for ensuring the sustainability and effectiveness of sovereign AI systems.

- **Skilled Talent** - A robust talent pool is essential for the development and deployment of sovereign AI. This includes AI researchers, engineers, data scientists, and policymakers who understand AI ethics and governance.
- **Energy Access** - High-performance AI systems demand substantial energy resources. Governments must ensure reliable and sustainable energy supplies while minimizing environmental impacts.
- **Hardware Resources** - Sovereign AI relies on advanced computing hardware, including chips, servers, and storage systems. Policymakers need to evaluate their ability to source, produce, or acquire these critical components domestically, minimizing dependency on foreign suppliers.
- **Carbon Footprint Management** - Large-scale AI systems can have a significant environmental impact. Policymakers must assess the carbon footprint of AI systems and implement measures to minimize emissions, aligning AI development with sustainability goals.
- **Water Availability** - Data centres, which are integral to AI operations, often rely on water for cooling. Nations must ensure that water resources are available without compromising ecological sustainability.

Policy Frameworks

Principles and Regulations:

The governance of AI must be rooted in core global principles that ensure transparency, safety, inclusivity, accountability, and fairness. These principles serve as the foundation for national and international AI regulations. A critical example of AI regulation is the European Union's AI Act, which adopts a risk-based approach, classifying AI applications into unacceptable, high-risk, and low-risk categories. The Act establishes strict compliance requirements for AI systems used in sensitive areas such as biometrics, healthcare, and law enforcement while fostering innovation in non-risk applications.

The Council of Europe Binding Convention on AI further solidifies international commitments to AI governance by ensuring that AI technologies adhere to human rights, democratic values, and the rule of law. These principles and regulatory frameworks must be adapted to local contexts, ensuring that AI regulations align with national security policies, economic structures, and ethical considerations.

For AI regulations to be effective, they must transition from principles to enforceable laws through structured legislative efforts. Governments must establish regulatory bodies that oversee AI compliance, create standards for AI auditing and risk assessment, and develop legal mechanisms for AI accountability in high-impact sectors. The role of public-private partnerships is also crucial in defining sector-specific guidelines that can balance regulation with innovation.

The policy landscape for AI governance spans three levels: foundational principles, regulatory frameworks, and enforcement mechanisms.

- **Principles** - Broad ethical standards guiding AI governance, such as transparency, accountability, and fairness.
- **Regulatory Frameworks** - Legal instruments defining compliance obligations, including AI-specific legislation (e.g., the EU AI Act) and sector-based risk classifications.
- **Implementation Mechanisms** - Institutional structures enforcing AI policies, including auditing bodies, certification protocols, and liability frameworks.

This layered approach ensures that AI governance is both adaptive and enforceable, allowing nations to regulate high-risk AI applications while fostering innovation in lower-risk sectors. Governments must avoid regulatory fragmentation, ensuring alignment between national policies and international AI governance standards.

Creating an Enabling Environment

A jurisdiction can foster an AI-friendly ecosystem by establishing coherent and flexible policies that promote AI innovation while ensuring responsible development. The following strategies contribute to an enabling AI environment:

1. Governments must conduct policy audits to identify legal inconsistencies or gaps that hinder AI adoption.
2. AI policies should implement adaptive regulatory models, such as regulatory sandboxes, which allow AI developers to test emerging technologies under controlled conditions.
3. Strengthening government AI expertise through training programs, research collaborations, and AI-focused task forces ensures that policymakers can effectively regulate AI.
4. Providing tax incentives, grants, and funding programs for startups and research institutions fosters a competitive AI ecosystem.
5. AI governance should avoid fragmented regulations by ensuring that AI policies are harmonized across sectors (e.g., finance, healthcare, defence).
6. Governments must raise awareness about AI risks and benefits, fostering public trust in AI applications.

Sector-specific regulatory approaches are critical, as AI's impact varies across industries. AI in finance requires strict risk assessment protocols, while AI in healthcare necessitates medical AI validation. Additionally, international collaboration is vital, as AI regulations must align with global standards while respecting national sovereignty.

Horizon Scanning

A dynamic AI regulatory environment requires continuous monitoring of global AI policies. Horizon scanning involves analysing AI governance models in key economies, including:

- The European Union AI Act and GDPR set strict AI governance standards.
- In the United States, the NIST AI Risk Management Framework guides voluntary AI governance.
- The Japan Society 5.0 Initiative promotes AI-driven digital transformation.
- The OECD & WEF organizations provide global AI policy frameworks for responsible development.
- The IEEE & UAI ethics and standardization efforts influence AI best practices.

By leveraging insights from international AI strategies, policymakers can develop robust national AI strategies that align with global best practices.

The development of sovereign AI involves a range of strategic considerations, each of which presents distinct opportunities and challenges. Governments must evaluate their national priorities, technological capabilities, and geopolitical realities before committing to specific AI development paths. This section explores key pathways for sovereign AI, ranging from highly autonomous national strategies to more collaborative and decentralized approaches.

1. Industry-Specific Sovereign AI

Many governments recognize that a fully sovereign AI ecosystem may not be feasible in all sectors. Instead, they may choose to prioritize AI sovereignty in strategic industries such as financial services, defence, healthcare, and cybersecurity. By focusing on industry-specific AI, nations can ensure that critical areas remain protected from external control while still benefiting from international AI collaborations. For instance, in financial services, sovereign AI can help enhance fraud detection, automate compliance, and secure transactions while remaining compliant with domestic financial regulations.

2. Creating National AI Capabilities

A comprehensive approach to sovereign AI involves significant investments in national AI capabilities, including the development of publicly funded AI research centres, high-performance computing infrastructure, and national AI datasets. Countries such as China and the United States have already established government-backed AI initiatives that foster innovation, talent development, and industry collaboration. Investing in domestic AI startups and enterprises can further strengthen national AI ecosystems, ensuring that key innovations remain under domestic control rather than being dominated by foreign technology providers.

3. Partnering with Big Tech

While full AI sovereignty is a long-term goal, many nations recognize the short-term benefits of partnering with multinational technology firms to build AI capacity. Strategic collaborations with major AI firms such as Google, Microsoft, Amazon, and NVIDIA can provide access to cutting-edge AI models, infrastructure, and expertise. However, such partnerships must be structured with robust legal and regulatory safeguards to protect national interests, ensuring that domestic data, intellectual property, and security concerns are not compromised.

4. Adapting Open-Source AI

An alternative approach to AI sovereignty is leveraging open-source AI models and frameworks. Open-source AI development allows countries to customize and deploy AI solutions without being locked into proprietary technologies. The emergence of open-source AI models, such as those developed by OpenAI, Meta, and Stability AI, provides nations with an opportunity to develop independent AI capabilities while benefiting from global advancements. However, maintaining national expertise in AI adaptation and security is essential to mitigate risks associated with external dependencies and vulnerabilities.

5. Integrating Digital Initiatives

Governments that are developing national AI strategies can benefit from aligning AI investments with existing digital initiatives. National projects related to digital identity, cloud computing, e-government services, and cybersecurity can be integrated into AI strategies to enhance efficiency, security, and scalability. For example, countries like Estonia and Singapore have successfully incorporated AI-driven digital identity systems into government services, improving public sector efficiency and security. Ensuring that AI aligns with broader national digital transformation initiatives can maximize the benefits of sovereign AI investments.

6. Exploring Decentralized AI

A growing trend in AI development is the exploration of decentralized AI architectures. Instead of relying on centralized national AI infrastructure, countries can develop federated learning models and decentralized inference systems that distribute AI training and decision-making across multiple nodes. Decentralized AI systems provide greater resilience, security, and adaptability while reducing the risks associated with centralized data storage and processing. Governments considering decentralized AI must invest in secure network architectures, privacy-preserving AI techniques, and interoperable AI standards to ensure the effectiveness of such systems.

7. Adopting a “Wait and See” Approach

For some nations, the most strategic approach to sovereign AI may be to observe technological advancements and global policy trends before committing to large-scale investments. Given the rapid evolution of AI technologies, governments must continuously assess emerging risks, opportunities, and regulatory frameworks. A “wait and see” strategy does not imply inaction; rather, it involves strategic monitoring, pilot programs, and risk assessments to ensure that AI policies remain flexible and responsive to future developments.

Balancing Sovereignty with Innovation

The challenge for policymakers is finding a balance between AI sovereignty and global collaboration. While full independence in AI development may be desirable in strategic sectors, many nations will benefit from hybrid approaches that integrate local AI capabilities with international partnerships. Ensuring that sovereign AI policies remain adaptable, forward-thinking, and aligned with global AI standards will be crucial for long-term success.

Moving Forward:

International cooperation plays a vital role in mitigating the risks of AI monopolization and ensuring equitable technological progress. **The Global Partnership on AI (GPAI)**, an OECD-affiliated initiative, fosters AI collaboration across diverse economies, including India, Canada, and the European Union. Additionally, regional groupings like **ASEAN and the African Union** offer frameworks for shared AI infrastructure, capacity-building programs, and ethical governance models. These cooperative mechanisms provide alternative pathways for nations that lack the resources to develop sovereign AI independently but seek to retain regional control over AI governance.

Governments must explore adaptive cooperation strategies beyond traditional diplomatic alliances. **Ad hoc coalitions** between nations with aligned interests, such as cross-border data-sharing agreements and open-source AI research initiatives, can enhance technological resilience without necessitating full AI sovereignty. Nations should strategically assess when collaboration enhances sovereignty rather than undermining it.

The successful implementation of sovereign AI requires structured policy planning, strategic partnerships, and continuous assessment of technological and regulatory developments. To ensure a balanced approach to AI sovereignty, the following key steps should be considered:

- Governments should create a dedicated task force comprising experts from policy, technology, industry, and academia to conduct in-depth feasibility studies on sovereign AI strategies. This group should assess existing national AI capabilities, infrastructure needs, and potential regulatory frameworks to determine the most effective approach for AI sovereignty.
- Given the cross-border nature of AI development, participation in international AI policy discussions is critical. Nations should actively engage with organizations such as the OECD, UNESCO, the Global Partnership on AI (GPAI), and the European Commission to exchange best practices, align AI governance strategies, and explore areas of cooperation. Such dialogues will help ensure that AI sovereignty policies remain adaptable and harmonized with international AI regulations.
- Policymakers should craft a comprehensive sovereign AI strategy that balances national security interests with economic and technological advancement. This policy should include investment in AI infrastructure, legal and ethical guidelines for AI development, and strategic collaboration models that foster innovation while ensuring data security and regulatory compliance.