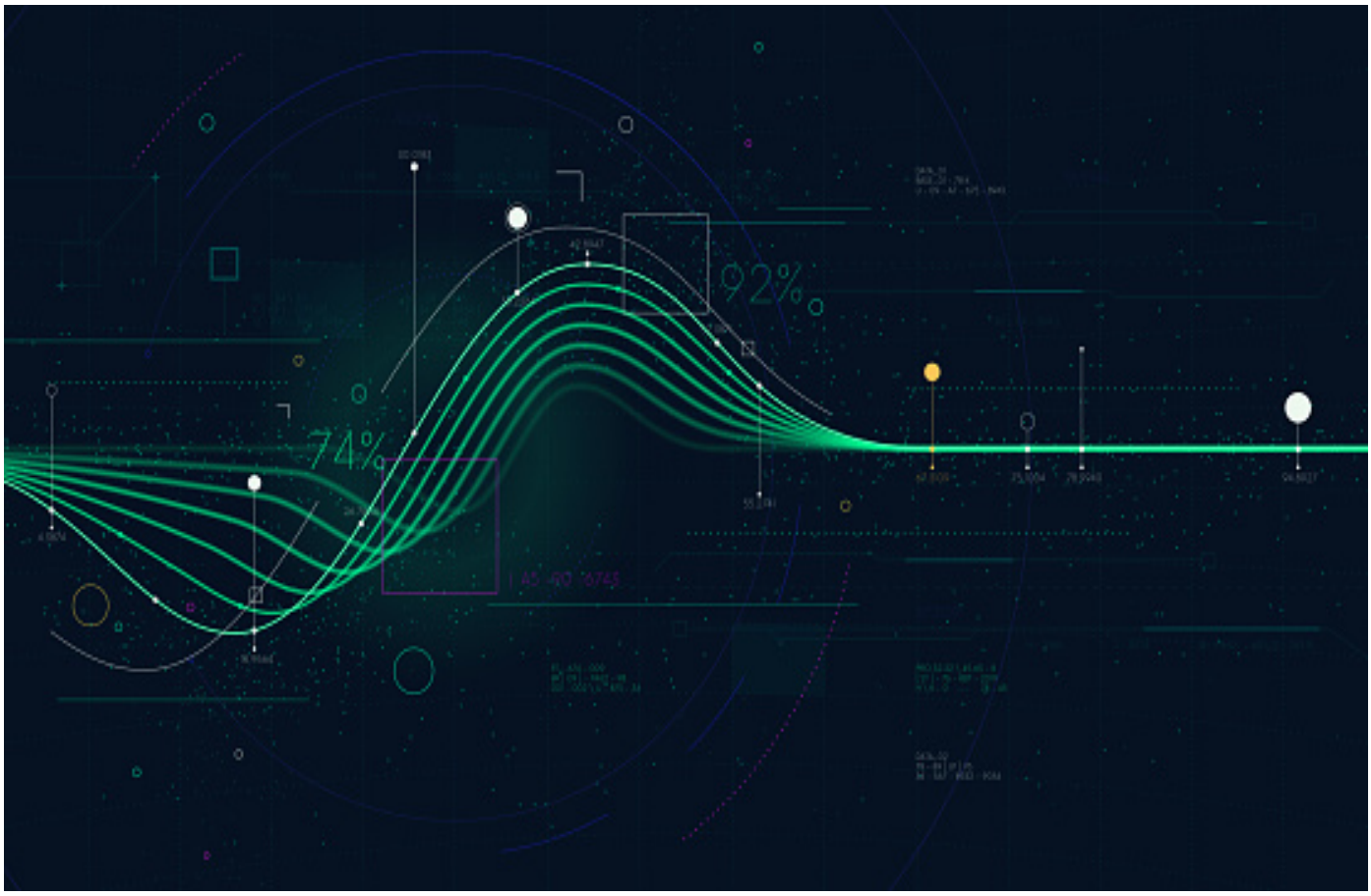



# Considerations Regarding Sovereign AI and National AI Policy

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.

**Prepared by:**

David L. Shrier[1,3],
Ayisha Piotti[2],
Alex Pentland[3,4],
Aldo Faisal[1]

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

Contact:

**david.shrier@imperial.ac.uk**

**aldo.faisal@imperial.ac.uk**

November 2024

## Executive Summary

The concept of sovereign AI is gaining traction globally as nations seek to develop AI capabilities independent of a few US-based or PRC-based Big Tech companies (FANGMs plus BATs). Sovereign AI initiatives are under way in domiciles (or regional authorities within countries) including Singapore, Japan, Germany, France, the UK, KSA, UAE, India and others. Countries must balance the significant financial and technological investments that may be required with other national priorities. Options for nation-states, supranational federals (e.g., EU) and international alliances (e.g., ASEAN) include building sovereign AI, collaborating regionally, and leveraging open-source initiatives.

Sovereign AI will participate in the broader structural shifts that AI is introducing globally, not only to economic and business structures, but in fundamental ways that people interact with each other, with companies, and with governments.

**This is an excerpt of a detailed report on Sovereign AI. The full report can be downloaded at [http://sovereign-ai.org/](http://sovereign-ai.org/)**

# Part I – Context

## Motivation for Sovereign AI:

Sovereign nations are initiating AI projects over which they have greater control and influence due to a number of factors, spanning political, economic and cultural spheres. These include


Source: Trusted AI Alliance https://aicollab.org

## Overview of Technology:

**What is a GPT?**:

- GPT (Generative Pretrained Transformer) is a powerful AI model that can perform various tasks like writing, answering questions, and translating languages by understanding and generating human-like text.

- **History and Evolution**: AI has evolved over 80 years, with significant milestones like machine learning and deep learning. The open-sourcing of TensorFlow and transformer technology in 2015 and 2018, respectively, revolutionized AI, particularly in natural language processing.

**Use Cases:**

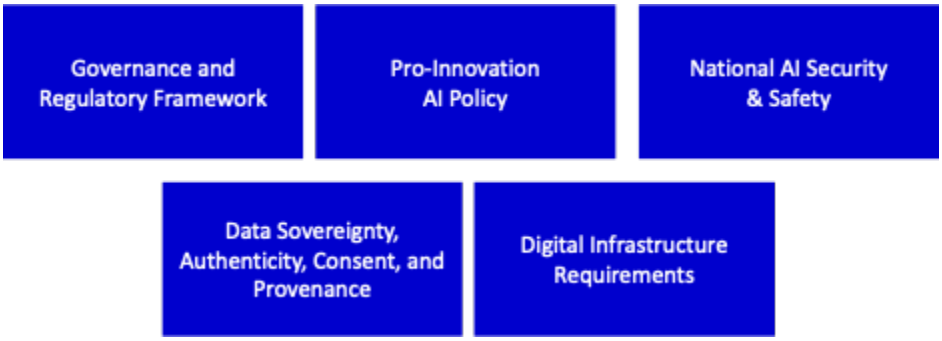
- **Financial Services**: AI enhances fraud detection, personalized banking, algorithmic trading, and financial advice.

- **Healthcare**: AI aids in diagnosis, preventative care, drug discovery, and medical transcription.

- **Education**: AI supports personalized learning, administrative efficiency, language learning, and writing assistance.

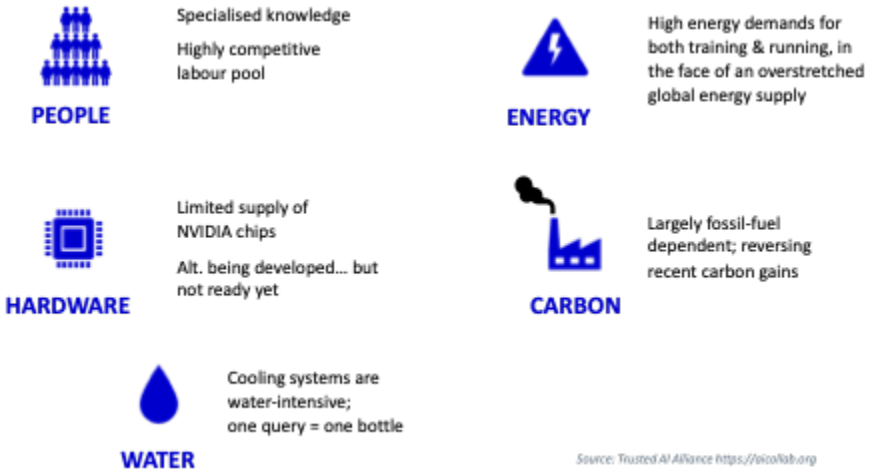
## Strategic Building Blocks:

Several critical elements form the strategic building blocks of Sovereign AI.



## Part II – 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 need to assess their domicile's infrastructure and capabilities against critical factors needed for creation and maintenance of sovereign AI, namely access to power, water, hardware, talent, and carbon.

## Part III — Policy Frameworks

- **Principles and Regulations**:
  Core global AI principles such as transparency, safety, inclusivity, accountability, and fairness are emphasized, with examples like the EU AI Act and the Council of Europe Binding Convention demonstrating how these principles evolve into binding regulations. Adaptation to local contexts and the transition from principles to enforceable laws are highlighted.

- **Creating an Enabling Environment**: A domicile can foster an AI-friendly ecosystem through identifying policy gaps, adopting flexible regulations, building public sector capacity, supporting innovators, enhancing regulatory coherence, and raising public awareness. Sector-specific regulatory approaches and international collaboration are also crucial.

- **Horizon Scanning**: Brief overview of AI policies in various countries (e.g., EU, Japan, USA) and multilateral organizations (e.g., OECD, WEF, IEEE, UN), highlighting their focus areas and frameworks for responsible AI development. These insights aim to inform national AI strategy by drawing on global best practices and standards.

## Part IV –Strategic Considerations

Policymakers have several options to take into account with respect to Sovereign AI:

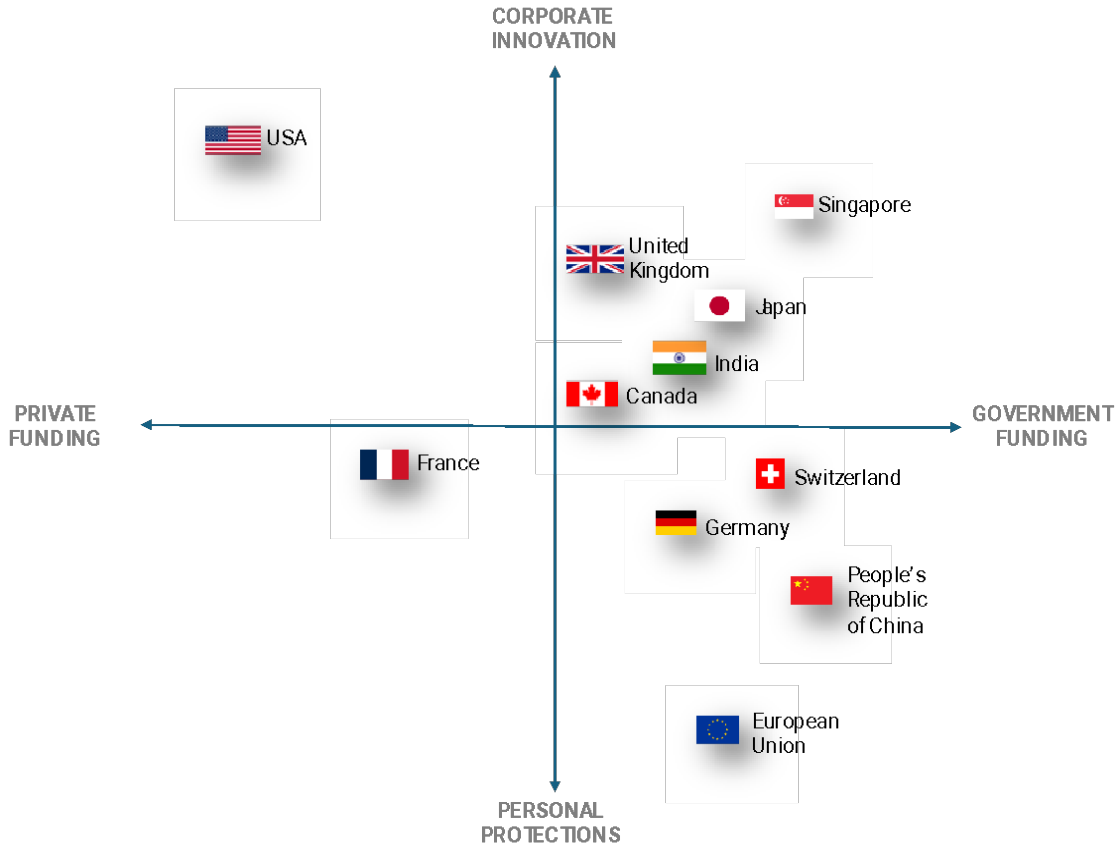
- **Industry-Specific Sovereign AI**: Focusing on specific sectors like financial services.
- **Creating Sovereign AI**: Developing national AI capabilities with significant investments.
- **Partnering with Big Tech**: Collaborating with companies that support local capabilities.
- **Adapting Open Source Code**: Leveraging open-source AI projects for non-aligned development.
- **Harvesting Benefits of Other Initiatives**: Integrating national projects like digital identity into the AI strategy.
- **Exploring Alternatives to Sovereign AI:** investigate other choices versus a full-blown sovereign, AI, such as decentralized inference systems.
- **Wait and See**: Monitoring the technology landscape while assessing risks and opportunities.

## Moving Forward:

- Form a high-level working group to assess options.
- Participate in multilateral dialogues for insights.
- Develop refined sovereign AI policy.

# SELECT EXHIBITS

When we compare across countries and regions, we can provide a simplified analysis on the dimensions of innovation vs. personal protection, and reliance on private funding versus government support:



## Comparison of AI Policy Initiatives by Domicile

| | EU | Japan | Singapore | France | USA | Germany | Switzerland | UK | Canada | China | India |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Values/ Ethics** | Human-centric and trustworthy AI | Human-centric AI and fairness | Fairness, ethics, accountability and Transparency (FEAT) | Transparency and Fairness | Democratic values and human rights | Responsible and public welfare-oriented AI | Human-centric approach | Transparency and explainability | Alignment with international human rights and responsible governance | Alignment with socialist values and national security, ensuring ethical AI development. | Ethical AI development, emphasizing fairness, transparency, and accountability |
| **Economic** | Support for AI start-ups and SMEs | Support for Startups and R&D, Innovation | Economic and Business support (SMEs), National AI R&D plans | Innovation and Economic support of SMEs, R&D | Innovation and Competition, investments in R&D | Research and Innovation, Technology Leadership, support for startups | Encouraging innovation | Investment in R&D and economic growth | Economic growth and commercialisation-support for business | Aiming to lead the global AI market by 2030 with significant investments in AI infrastructure. | AI for national development, enhancing sectors like agriculture, healthcare, and education |
| **Safety** | Ethical guidelines for AI development and regulation | Privacy Protection, ensuring security | Trusted development and deployment, protection of data | Data protection and Privacy | AI Bill of rights, Safe, secure and trustworthy development of AI, privacy and civil liberties | Data security, explainability | Transparency, Traceability and Explainability. Safe, robust systems | Public Trust and Responsible Innovation | Transparency, accountability and trust | Introduced measures to regulate synthetic content and ensure AI safety | Compliance with stringent data privacy requirements |
| **Risk management** | Risk-based approach for legislation | Guidelines to address AI risks, transparency and accountability | Accountability, Responsibility-testing and assurance | Transparency of algorithms, ethical code for programmers | Evaluation and policies to test and mitigate AI risks, consumer protection | Setting national AI standards, Protecting personal rights | Accountability and liability. Putting people first | Clear, adaptable and trustworthy regulatory regime. Responsible innovation | Creation and adoption of standards related to AI | Focused on national security and strict AI regulations | Ethical guidelines ensuring that AI is deployed transparently and safely |
| **International Collaboration / Standards** | Ethical Framework and Standards across the EU | International Collaboration and Standards | Global Outlook and international collaboration | European and International Cooperation | International Cooperation | International cooperation and social dialogue | Engagement in Global AI governance | Global Leadership and cooperation, promote cross-border interoperability | International collaboration and standards | Setting national AI standards that align with international norms | Active in global AI partnership |

| | EU | Japan | Singapore | France | USA | Germany | Switzerland | UK | Canada | China | India |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Infrastructure** | Infrastructure Investment - AI factories, data infrastructure | Data management and infrastructure | Infrastructure Investment and Data management, sectoral focus | Infrastructure and Data management | Infrastructure investment and data management | Infrastructure investment - computational resources, AI competence centers | Digital Infrastructure | Investment in digital Infrastructure- smart homes, electro mobility | Research institutes, AI compute infrastructure | Investing heavily in AI infrastructure and core technologies | AI across sectors to drive inclusive growth and modernization |
| **Workforce** | Education, skills and Talent Development | Education and skills development | National Competency, Education and Talent Development | Education and Talent Development | Investment in Education and AI training, focus on workers rights | National training strategy | Investment in education and building new skills | Create new AI-related jobs | Development research Talent, education and training | Promoting AI-driven productivity in key industries | Upskilling and integrating AI in public services |
| **Legal Framework** | Liability and Legal Framework | Fair competition and market regulation, Soft law | Soft Law and flexible governance | Legal Framework via EU AI Act | (Proposed) Federal and state regulations | Legal Framework via EU AI Act | Soft law guidelines,AI regulation, accountability and proposed legislation | Soft law approach | (Proposed)federal regulation and ethical framework | Active approach – Implemented laws like the draft AI Law and generative AI rules | Existing laws govern AI, with new legislation in progress |
| **Commentary** | Lack of emphasis on changes to workforce/ work environment, job insecurity | Ethics Ineffectiveness, Corporate Agenda | Good Adaptability | Implementation Challenges, Investment needed | Favoring Big Tech, regulatory Gaps | Data Protection, Talent shortage | Skills training, public skepticism | Economic focus, alignment versus enforcement on ethical standards | Challenges concerning regulatory clarity and public trust | Stringent security standards posing challenged for AI service providers | AI generated content under the government strict supervision |

*Source: Trusted AI Alliance analysis, Imperial College London 2024*

**When we look at research output, the top countries publishing AI research (aside from the US) also see government supporting significant private sector activity in AI:**

| Country | AI Publications (in 000s) 2016-2020 | AI Index Rank |
|---|---|---|
| China | 76.3 | 2 |
| United States | 44.4 | 1 |
| India | 27.0 | 14 |
| United Kingdom | 16.0 | 4 |
| Japan | 13.0 | 12 |
| Germany | 12.9 | 8 |
| France | 10.9 | 13 |
| Spain | 9.7 | 21 |
| Italy | 9.1 | 23 |
| Australia | 8.5 | 15 |
| Canada | 8.2 | 5 |
| South Korea | 6.5 | 6 |
| Brazil | 6.0 | 35 |
| Poland | 5.8 | 27 |
| Iran | 5.7 | *Not ranked* |
| Turkey | 5.6 | 39 |
| Russia | 4.8 | 30 |
| Taiwan | 4.8 | 26 |
| Singapore | 4.1 | 3 |
| Hong Kong | 4.1 | 32 |
| Malaysia | 3.8 | *Not ranked* |
| Netherlands | 3.6 | 11 |
| Switzerland | 3.3 | 9 |
| Portugal | 3.2 | 29 |
| Saudi Arabia | 2.9 | 31 |

*Sources: Statistica 2023; Tortoise Media 2023*

Volume of publication does not automatically confer overall AI competitiveness, when taking into account other factors like talent availability, infrastructure, ease of doing business and other dimensions of the operating environment, and development of research into commercial activity.

# About the Authors

## David Shrier

David Shrier is a Professor of Practice, AI and Innovation, with Imperial College London, where he leads the Trusted AI Alliance (a multi-university collaborative focused on building responsible & trustworthy AI). He has helped over 100 governments to develop technology policy including advising the European Parliament on the EU AI Act, and leading the team that created the Commonwealth Fintech Toolkit. Other relevant work includes creating the Leadership and Diversity for Regulators programme when he was with University of Oxford, helping more than 50 nations shape financial inclusion policy. His latest book, *Basic AI: A Human Guide to Artificial Intelligence* was published in 2024.

**Name**
david.shrier@imperial.ac.uk
**Title**
Professor of Practice

## Ayisha Piotti

Ayisha Piotti is the head of the AI Policy Summit, an annual event jointly organized by the ETH Zurich Center for Law & Economics and RegHorizon. She has over 20 years of experience in the private & public sector, including with the United Nations.

**Name**
ayisha.piotti@imperial.ac.uk
**Title**
Head of the AI Policy Summit

## Alex Pentland

Alex Pentland has helped create and direct the MIT Media Lab and the Media Lab Asia in India, and is a HAI Fellow at Stanford. He is one of the most-cited computational scientists in the world, and *Forbes* declared him one of the "7 most powerful data scientists in the world" along with Google founders and the Chief Technical Officer of the United States. He co-led the World Economic Forum discussion in Davos that led to the EU privacy regulation GDPR, and was one of the UN Secretary General's "Data Revolutionaries" helping to forge the transparency and accountability mechanisms in the UN's Sustainable Development Goals.

**Name**
sandy@media.mit.edu
**Title**
HAI Fellow at Stanford

## Aldo Faisal

Aldo Faisal holds a prestigious UKRI Turing AI Fellowship, and is a Professor of AI & Neuroscience at Imperial College London and holds the Chair in Digital Health at the Universtät Bayreuth (Germany). He is director of the £50 million UKRI Centers in AI for Healthcare and AI for Digital Health in London, and a Director of the Alan Turing Institute. He is actively engaged in developing open source Generative AI systems.

**Name**
aldo.faisal@imperial.ac.uk
**Title**
Professor of AI & Neuroscience at Imperial College London

## Origins of this Report

In the spring of 2023, Farhan Ahmad, the CEO of Payment Systems Malaysia (PayNet), participated in the World Innovation Network (TWIN) roundtable organised by Prof. Rob Wolcott of University of Chicago and Northwestern. At this roundtable, Mr. Ahmad met David Shrier, Professor of Practice in AI & Innovation with Imperial College London. Ahmad subsequently visited the Imperial campus and met with Shrier and his colleagues to discuss the importance of AI for Malaysia's financial system, and opportunities to leverage academic research and insights to address fraud and cyber crime.

This led to a multi-year collaboration with Imperial's Trusted AI Alliance, a network of AI researchers spanning multiple institutions, seeking to deliver responsible and trustworthy AI for the world. As the conversations evolved, it emerged that Malaysia was revisiting its national AI strategy and was seeking to evaluate the broader context of AI policy and global interdependence, as well as address the needs for AI systems that are tailored to suit Malaysia's citizens.

In the course of this exploration the project team realised that the same questions that Malaysia faces, and the principles to develop answers to those questions, are also relevant to a broader set of policymakers and innovators in numerous other domiciles. PayNet authorised the creation of a report that not only could be useful in Malaysia, but that could help other ASEAN countries (as Malaysia assumes the Chair of the ASEAN group) and other political bodies globally, in navigating national AI policy and sovereign AI.