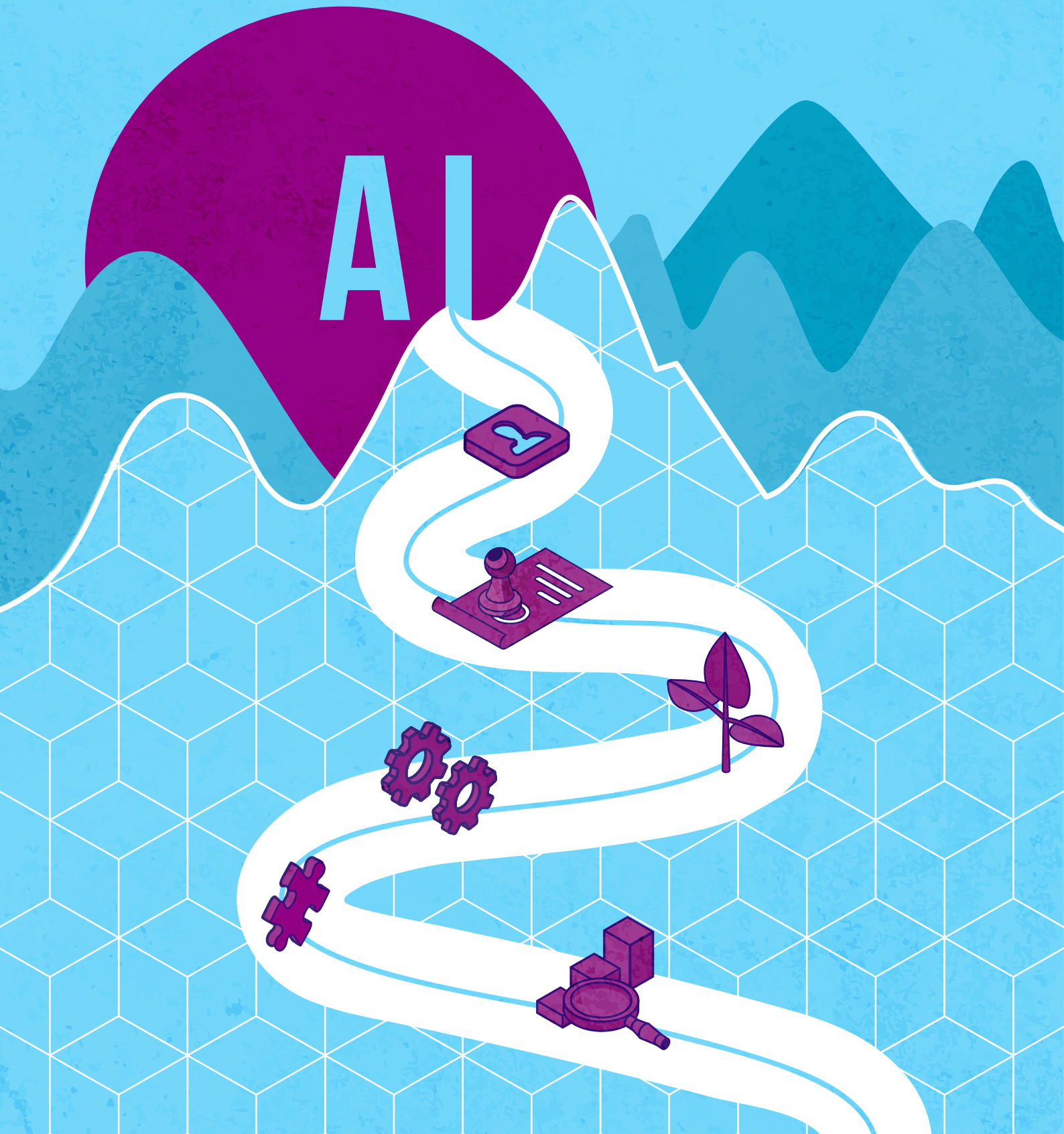
DATASPHERE
INITIATIVE

SANDBOXES FOR AI

TOOLS FOR A NEW FRONTIER



About the Datasphere Initiative

The Datasphere Initiative is a think and do tank that catalyzes meaningful dialogues and co-creates actionable and innovative approaches to respond to data challenges and harness opportunities across borders. Our mission is to equip organizations to responsibly unlock the value of data for all.

For more information, visit www.thedatasphere.org or contact info@thedatasphere.org.

About this report

This report was first developed as a working paper presented at the Global Sandbox Forum Inaugural Meeting in July 2024 and then further developed in an online roundtable in December 2024 and through individual consultations with experts. This report summarizes the findings of a study on sandboxes in the field of Artificial Intelligence — those that have been announced, are in development, or have been completed. The study aimed to identify patterns across the examples analyzed, focusing on the reasons behind their creation, as well as the timing and processes involved in their development. The report presents why regulators and companies should consider sandboxes for AI and provides recommendations on when and how to implement them. In addition to documentary research, we organized Chatham House-style events with experts to gather their insights as the research progressed. Quotes from the consulted experts are included throughout the report, without disclosing their names or positions.

TABLE OF CONTENTS

Acknowledgements

When preparing this report, the Datasphere Initiative team held consultations with various stakeholders in the form of in-person and online meetings. The report has however been drafted under the sole responsibility of the Datasphere Initiative and does not engage the actors listed. The Datasphere Initiative is thankful to all those who shared their expertise and experience which helped the production of the report.

Datasphere Initiative team: **Lorrayne Porciuncula**, Executive Director, **Bertrand de La Chapelle**, Chief Vision Officer, **Maria Marinho**, Director of Research, **Sophie Tomlinson**, Director of Programs, **Mariana Rozo-Paz**, Policy, Research and Project Management Lead, **Morine Amutorine**, Lead Africa Sandboxes Forum, **Gabriel Souto**, Research Assistant, **Risper Onyango**, Research Assistant.

Contributors in their personal capacity: **Lee Chein Inn**, Deputy Director, Development of Data Driven Tech of IMDA Singapore, **Carolyn Nguyen**, Board Member, Datasphere Initiative, **JiaYen Ong**, Independent Lawyer, **Adam Zable**, Datasphere Fellow 2023-2024, **Sam Jungyun Choi**, Lawyer, Covington & Burling LLP, **Alex Moltzau**, Policy Officer, European AI Office, **Thiago Moraes**, PhD Scholar, Vrije Universiteit Brussels, **Romanas Zontovičius**, ICT Industry Manager at Innovation, Agency Lithuania, Government of Lithuania, **Livio Rubino**, Project Officer, AI Office, European Commission, **Mei Lin Fung**, Cofounder, PCI, **Norberto Andrade**, Consultant, **Becky King**, Project Manager, OECD, **Andrea Motta**, UK-Brazil Tech Hub Director, British Government.

The report counted with editorial and design support from **Barbara Miranda**, Design Thinking Lead of the Datasphere Initiative.

EXECUTIVE SUMMARY

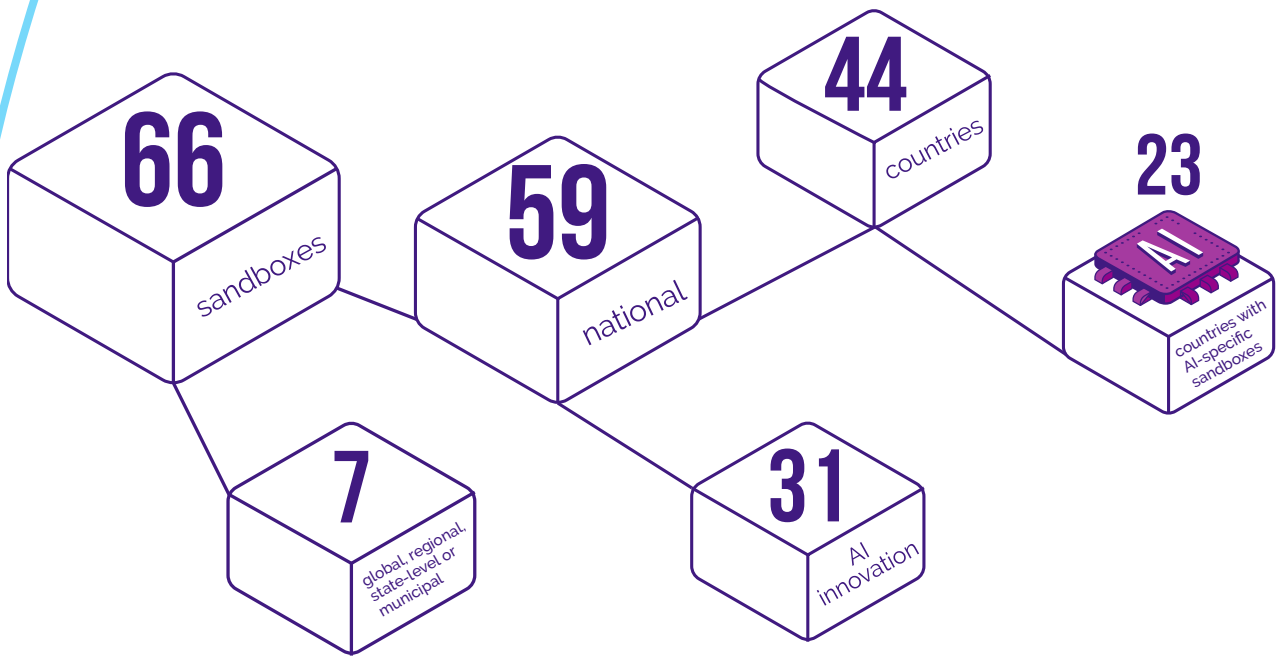
Artificial Intelligence (AI) offers **unprecedented opportunities** to enhance economic development and human welfare. However, realizing these benefits will depend on how effectively we understand how AI technologies work, are built, and manage the substantial risks and potential harms associated with them. Data governance is the cornerstone of responsible and ethical AI development, but holistic, inclusive and robust AI and data governance models require iterative, agile and safe spaces to be developed and refined.

Traditional regulatory processes are often linear, with a single point of consultation followed by a fixed outcome. This rigidity can often stifle innovation and lead to outdated policies that fail to address emerging technologies' complexities and risks. **Sandboxes, defined as safe spaces to test new technologies and practices against regulatory frameworks or experiment with innovative uses and means of governing data**[1], have emerged as a tool to help regulators better adapt and address factors of innovation throughout a regulatory cycle. Sandboxes can help policymakers develop flexible, innovative and participative approaches to testing new regulatory and technical models.

This report explores **sandbox initiatives and outlines the potential of sandboxes as tools for the regulation and development of AI and data-driven technologies**. The report aims to identify patterns across the examples analyzed, focusing on the reasons behind their creation, as well as the timing and processes involved in their development. It maps sandboxes for data, AI and technology from around the world, highlighting sandboxes for AI innovation which are categorized by country, the coordinating authority responsible, duration, type, and topics tested, providing a comprehensive overview of their scope and impact.

[1] For a short introduction to sandboxes view *here*.

As of January 2025, research by the Datasphere Initiative reveals that there are **66 sandboxes related to data, AI or technology worldwide**, 59 of which are national sandboxes and the remaining 7 are global, regional, state-level or municipal sandboxes. Among the 59 national sandboxes, **31 are specifically designed to foster AI innovation**, focusing on areas such as machine learning, AI development, and data-driven solutions. At least 44 countries have implemented or are in the process of developing national sandboxes, with **23 countries actively planning or operating AI-specific sandboxes** to address emerging technological challenges and opportunities. These sandboxes are not only proving grounds for technological innovation, but also critical spaces to test and refine governance frameworks, ensuring that they remain responsive and effective in addressing emerging challenges.



The report presents a number of **concrete cases of sandboxes for AI** and analyzes why regulators, companies and civil society should consider sandboxes for AI. It provides an assessment of the incentives and reasons for when to do it. Despite the broad diversity of experiences around the world, it is found that a structured approach is needed for sandbox implementation. The report lays down a blueprint for how to implement sandboxes for AI, outlining a five-phase approach: design, planning, execution, communication and engagement, and closure and evaluation.

Key insights

Sandboxes could be essential tools for AI governance – As AI technologies evolve at an unprecedented pace, traditional regulatory frameworks struggle to keep up. Sandboxes provide a controlled, adaptive environment for regulators, industry, and other stakeholders to experiment with new AI governance models, reducing regulatory uncertainty and fostering responsible innovation.

The use of sandboxes for AI is rapidly expanding – The report identifies a growing number of AI-specific sandboxes. These sandboxes serve multiple purposes, from clarifying regulatory parameters to enabling AI experimentation under real-world conditions.

Cross-border and cross-sectoral collaboration is still underdeveloped – Despite the global nature of AI, sandboxes for AI are mostly developed at the national level with limited collaboration between countries or sectors. Greater coordination across jurisdictions and regulatory bodies could help harmonize AI governance and avoid regulatory fragmentation.

Different types of AI sandboxes serve distinct purposes – AI sandboxes are categorized into regulatory, operational, and hybrid models. Regulatory sandboxes help test AI compliance with existing laws, operational sandboxes focus on data-sharing and experimentation, and hybrid sandboxes combine both approaches. National sandboxes have a different function from sub-national and cross-border ones. Each has specific advantages depending on the regulatory and innovation needs.

Transparency, trust, and inclusivity are critical for success – Sandboxes must balance flexibility and accountability. A lack of transparency and clear public reporting on sandbox outcomes can undermine public trust and limit their effectiveness. Meaningful participation from civil society and academia is often missing but necessary to ensure AI sandboxes serve the broader public interest.

As countries grapple with the fast-evolving nature of AI and increasingly acknowledge a need to innovate with new regulatory tools and frameworks. This report builds on the work of the Datasphere Initiative[2] and its Global Sandboxes Forum[3] to share analysis and recommendations on sandboxing in the age of AI and how to foster these new regulatory tools in the interest of people and the planet.

[2] Datasphere Initiative (2022), *Sandboxes for data: creating spaces for agile solutions across borders*, Datasphere Initiative.

[3] Datasphere Initiative (2024), *Global Sandboxes Forum*, Datasphere Initiative.

INTRODUCTION

Artificial Intelligence (AI) relies on vast stores of data to function effectively, identifying patterns, making predictions, and performing some tasks which in the past required human intelligence. Generative AI (GenAI) creates new content, such as text, images, audio, or code, based on patterns learned from large datasets driven by advances in deep learning and neural networks, alongside increased computational power. AI offers unprecedented opportunities to enhance human welfare by transforming productivity, human capability, and decision-making, impacting sectors like healthcare,[4] supply chain management,[5] and education.[6] However, realizing these benefits will depend on how effectively we understand how AI technologies work, are built, and manage the substantial risks and potential harms associated with them.

As AI's potential to influence political narratives and electoral outcomes grows, global unease increases.[7] AI's integration into healthcare, finance, education, and entertainment raises ethical, legal, and social issues, with concerns about systemic biases, job displacement, and privacy. The interconnectedness of AI's impact across multiple sectors and its ability to affect a wide range of stakeholders make it clear that a new model of governance and participation is essential to ensure its responsible development and deployment.

Despite the current public attention towards AI, it has become clear that its rapid evolution has outpaced existing regulatory frameworks, which remain anchored in traditional, inflexible command-and-control models.[8] This mismatch has limited the ability of regulators to experiment, adapt, and respond quickly to new developments. Public sector institutions, constrained by cultural and resource limitations, have often struggled to keep up with the dynamic nature of AI innovation.

Regulatory experimentation models have been adopted to address these challenges. These models are grounded in an agile, data-driven approach to regulation, focusing on adaptive learning through collaborative relationships with stakeholders. The objectives are to reduce informational asymmetry, create a reliable testing environment for new products and processes, enhance the resilience of emerging regulations, and foster innovation. One of these tools is sandboxes.[9]

---

[4] Saraswat et al (2022), *Explainable AI for Healthcare 5.0: Opportunities and Challenges*, IEEE Access.

[5] Dash et al (2019), *Application of Artificial Intelligence in Automation of Supply Chain Management*, Journal of Strategic Innovation and Sustainability.

[6] Bhutoria (2022), *Personalized education and Artificial Intelligence in the United States, China, and India: A systematic review using a Human-In-The-Loop model*, Computers and Education: Artificial Intell.

[7] Bell and Korinek (2024), *AI's economic peril to democracy*, Brookings / Nick Robins-Early (2023), Disinformation reimagined: how AI could erode democracy in the 2024 US elections, The Guardian.

[8] Sinclair, D. (1997). *Self-regulation versus command and control? Beyond false dichotomies*. Law & Policy.

[9] OECD (2024), *Regulatory experimentation: Moving ahead on the agile regulatory governance agenda*, OECD Public Governance Policy Papers.

Sandboxes were originally developed by financial technology (FinTech) regulators from the enclosed software testing environments used by programmers. Sandboxes have now evolved into crucial tools for the exploration and regulation of data-driven technologies like AI.[10] Since 2015, countries such as the United Kingdom,[11] Norway[12] and Singapore[13] have designed and implemented sandboxes to test emerging technologies including AI. Countries in the Global South such as South Africa[14] and Brazil[15] have also begun to use sandboxes to test and understand regulatory impacts of digital technologies. Additionally, enacted in 2024, the European Union[16] AI Act (EU AI Act) requires EU member states to establish AI sandboxes as part of its comprehensive strategy to regulate AI across the Union. Article 57 of the EU AI Act details how these sandboxes are intended to provide a controlled testing environment in which innovators and regulators will work together to identify risks and ensure compliance with the EU AI Act and potentially other EU regulations.[17] The diversity of sandbox implementations and references in AI legislation reflects a spectrum of objectives, from strict compliance to fostering innovation, and varies widely globally.

As countries grapple with the fast-evolving nature of AI and increasingly acknowledge a need to innovate with new regulatory tools and frameworks, clarity on the lessons from these experiences becomes increasingly important.

The report begins by mapping the landscape of AI sandboxes worldwide, offering a **global overview** of how different jurisdictions are experimenting with these regulatory tools. It categorizes sandboxes into regulatory, operational, and hybrid models, analyzing their objectives, governance structures, and geographic distribution. By examining case studies and identifying patterns in AI sandbox initiatives, this section provides a foundational understanding of the role sandboxes play in AI governance.

Following the global mapping, the report delves into the **rationale for sandboxing AI** — exploring why regulatory sandboxes are increasingly seen as a necessary mechanism for managing AI's risks and uncertainties. Traditional regulatory approaches often fail to keep pace with AI advancements, leading to either regulatory gaps or overly restrictive policies that hinder innovation. Sandboxes offer an alternative, providing **controlled environments for iterative learning, stakeholder engagement, and real-time regulatory experimentation.** The section outlines the key benefits of sandboxes, including their ability to bridge regulatory asymmetries, foster trust among public and private actors, and create a structured process for assessing AI risks before full-scale deployment.

---

[10] Datasphere Initiative (2022), *Sandboxes for data: creating spaces for agile solutions across borders*, Datasphere Initiative.

[11] Medicines and Healthcare products Regulatory Agency (2024). *AI Airlock: the regulatory sandbox for AIaMD*. GOV.UK.

[12] Datatilsynet (2021). *Regulatory privacy sandbox*. Datatilsynet.

[13] Infocomm Media Development Authority (2023). *Privacy Enhancing Technology Sandboxes*. IMDA.

[14] Intergovernmental Fintech Working Group (2021). Regulatory Sandbox. IFWG.

[15] Brazilian Data Protection Authority (2024). *ANPD's Call for Contributions to the regulatory sandbox for artificial intelligence and data protection in Brazil is now open*. ANPD.

[16] The EU is currently made up of 27 countries in Europe: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden.

[17] EU Artificial Intelligence Act (2024), *Article 57: AI Regulatory Sandboxes*, Future of Life Institute.

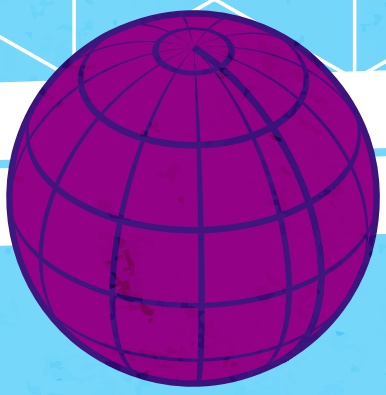
The third section of the report examines the **timing of AI sandbox implementation**, answering the critical question of **when** sandboxes should be used in the regulatory process. Sandboxes can be established at different moments — **before regulation is developed** to explore risks and governance needs, **during regulation drafting** to refine policy approaches, or **after regulation is enacted** to test compliance and ensure effective implementation. The section also highlights how different stakeholders — governments, start-ups, large corporations, and civil society—engage with sandboxes at varying stages, depending on their incentives and regulatory certainty. Understanding when to deploy a sandbox is essential to maximizing its impact and aligning it with broader policy objectives.

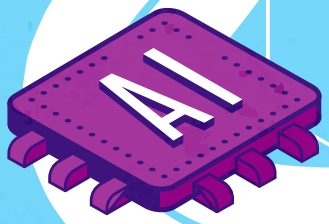
Finally, the report provides a **practical guide on how to implement sandboxes for AI**, detailing a structured, five-phase approach: initiation, planning, execution, communication and engagement, and evaluation. It emphasizes the need for **clear governance frameworks, stakeholder coordination, risk mitigation strategies, and iterative learning mechanisms** to ensure that sandboxes effectively contribute to AI governance. The section also highlights the importance of **transparency, cross-sectoral collaboration, and knowledge-sharing** to avoid duplication of efforts and enhance the global regulatory landscape. As AI continues to evolve, the ability of policymakers to design, test, and refine regulatory approaches through sandboxes will be a key determinant of their success in fostering innovation while safeguarding public interest.

This report builds on the work of the Datasphere Initiative[18] and its Global Sandboxes Forum[19] – a collaborative space dedicated to exchange sandbox ideas, best practices, and real-world applications, to share analysis and recommendations on sandboxing in the age of AI and how to foster these new regulator tools in the interest of people and the planet.

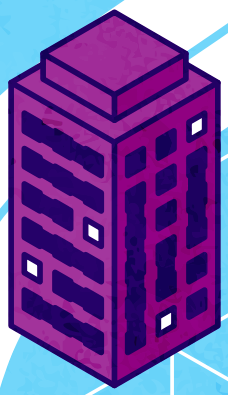
[18] Datasphere Initiative (2022), *Sandboxes for data: creating spaces for agile solutions across borders*, Datasphere Initiative.

[19] Datasphere Initiative (2024), *Global Sandboxes Forum*, Datasphere Initiative.





GLOBAL OVERVIEW OF SANDBOXES FOR AI AND DATA



This report presents the results of mapping sandboxes related to data, AI and technology since 2017. The information is organized by country, the coordinating authority responsible, duration, type, and topics tested, providing a comprehensive overview of their scope and impact. By examining these sandboxes together, the report underscores how technology and data-related sandboxes can provide the foundational resources and infrastructure necessary for AI experimentation, while AI sandboxes enable the development of intelligent applications that leverage and enhance data utility. In fact, the interdependent relationship between AI and data highlights that effective and responsible AI development relies on robust data ecosystems,[20] while advancements in AI further unlock the value and potential of data.

The formal introduction of regulatory sandboxes in 2015 by the UK Financial Conduct Authority marked a milestone in the development of experimental regulatory approaches.[21] Initially designed to accommodate innovation in financial services, such as FinTech, digital payments, and cryptocurrencies, sandboxes were inspired by controlled environments used by software developers to test code without affecting the broader system. Since then, the sandbox model has been adapted for use in a wide range of sectors, including telecommunications, transportation, utilities, aviation, and data governance. Sandboxes are now deployed at various levels—national, subnational (municipal and regional), and even cross-border—and can take the form of regulatory, operational, or hybrid models.[22] Moreover, the establishment of a legal and institutional framework for personal data protection over the past decade has encouraged national data protection authorities to experiment with regulatory and operational solutions for the emerging applications of AI systems. For this reason, some sandboxes in the area of personal data protection have been integrated into the list of AI-related sandboxes.

There is still no consensus in the literature on how sandbox experiments can be organized and categorized. The World Bank classifies sandboxes as: policy-focused, product or innovation-focused, thematic, and cross-border.[23] The Datasphere Initiative developed a simpler classification based on the primary objective of the experiment, organizing the sandbox experiences into two main types: regulatory and operational.[24] Broadly speaking, regulatory sandboxes can help clarify regulatory parameters and improve regulators' abilities to respond to sectoral needs, while operational sandboxes enable stakeholders to access pooled data resources to explore new uses of data. There are also the so-called hybrid sandboxes that provide a mix of these two objectives.

---

[20] World Economic Forum. Global Future Council on the Future of Data Equity. (2024). *Advancing Data Equity: An Action-Oriented Framework*. White paper.

[21] United Kingdom Financial Conduct Authority (2017), *Regulatory sandbox lessons learned report*.

[22] Datasphere Initiative (2022), *Sandboxes for data: creating spaces for agile solutions across borders*, Datasphere Initiative.

[23] Appaya, M. S., Gradstein, H. L., & Haji Kanz, M. (2020). *Global Experiences from Regulatory Sandboxes*, World Bank.

[24] Datasphere Initiative (2022), *Sandboxes for data: creating spaces for agile solutions across borders*, Datasphere Initiative.

Regulatory sandboxes offer a collaborative, time-bound setting where innovators can test novel technologies and practices within existing regulatory frameworks under the supervision of regulators, before the new product or service is offered to the wider market. When developed with appropriate guardrails, this model is beneficial for assessing new technologies while minimizing risks to the broader economy.[25] Regulatory sandboxes are not merely theoretical spaces but dynamic environments where real-time interaction occurs between innovators and regulators, helping to refine both technology and business models. Outputs of a regulatory sandbox can include products or services market launches, streamlining of compliance procedures, among others. For regulatory sandboxes, some are designed to align innovators with existing regulations, enhancing the oversight capabilities of regulators, sometimes intending to effect regulatory modifications, changes or revisions to underlining legislation. Others, especially those aimed at fostering scientific and industrial innovation, may temporarily relax certain regulations to encourage experimentation.

The case of South Korea provides an example of a regulatory sandbox for AI. This initiative initiated under the "5 Regulatory Innovation Acts" of 2019 and further developed in subsequent years, emphasizes economic growth and technological innovation. Since 2019, these sandboxes have seen significant evolution, with new laws, regulations, and guidance enhancing their effectiveness. They provide tailored pathways for companies to test new technologies or business models within defined limits, often bypassing traditional regulatory barriers. South Korea's case highlights the potential of flexible regulation to serve as a tool for strategic technological leadership. The country's sandbox system offers a flexible regulatory environment that encourages testing and quick adaptation to technological needs. This approach explicitly shifts the regulatory paradigm from a traditionally restrictive framework to one that prioritizes innovation by permitting activities unless they are explicitly prohibited.[26] Today, five different government agencies in South Korea each operate their own sandbox, in industrial convergence, regulation free zones (regional innovation), information and communication technology (ICT), FinTech, and smart city.[27]

Operational sandboxes, on the other hand, are secure platforms where datasets and other resources are pooled and accessed by various stakeholders to explore new data applications. They enable collaborative analysis and testing of datasets in a controlled environment.[28] Outputs from an operational sandbox can include open platforms or services or open datasets for public use. Operational sandboxes, too, vary widely in their objectives, scope, and impact.

[25] United Nations Department of Economic and Social Affairs (2021), *Sandboxing and experimenting digital technologies for sustainable development*, Future of the World Policy Brief.

[26] Heonyoung Kwon (2019),. "한국형 규제샌드박스의 현황과 향후 과제 *[Current status and future challenges of Korean regulatory sandbox]*, KISO Journal.

[27] European Chamber of Commerce in Korea (2020), *Guide to the Regulatory Sandbox*, European Chamber of Commerce in Korea.

[28] Datasphere Initiative (2022), *Sandboxes for data: creating spaces for agile solutions across borders*, Datasphere Initiative.

In **Australia**, the Australian Competition and Consumer Commission launched the Consumer Data Right (CDR) Sandbox to help participants test and improve their data sharing solutions under the Consumer Data Right legislation. Its primary objective is to enhance the quality of data sharing offerings, reduce barriers for businesses, and foster innovation in the early stages of software development. The scope of this sandbox is relatively narrow at present, focusing on the financial and energy sectors the CDR covers, but the government intends to continuously expand the CDR to other sectors. Its impact is primarily on helping companies comply with data sharing regulations, improving their data sharing services, and ultimately empowering consumers with greater control over their personal data.[29]

In contrast, in **Colombia**, the Ministry of Information Technology and Communications launched the Data Sandbox Collaborative Space in 2021 to enable public entities to carry out pilot projects in analytics and Big Data. It aims to leverage Big Data technologies to address public sector challenges and foster innovation within government entities by providing a secure environment for public entities to develop and refine solutions to public and citizen problems. This operational sandbox allows public entities to pool data resources, experiment with data-driven solutions, and enhance their capabilities in handling large-scale data through the use of advanced Big Data technologies.[30]

The scope and intended impact of the Colombian sandbox are wider than those of Australia's. While the Australian sandbox focuses only on CDR data sharing, Colombia's sandbox addresses critical societal issues such as identifying flood-prone areas, monitoring rural land markets, and estimating poverty through a multidimensional index. The Australian sandbox aims to improve compliance and service quality in data sharing under the CDR regulatory framework, whereas the Colombian sandbox aims to enhance the data handling capabilities of public entities and promote collaborative experimentation to improve public services and inform policy decisions.

Finally, **hybrid sandboxes** combine elements of regulatory and operational sandboxes, offering a controlled and collaborative environment where innovators can test new technologies, datasets, and practices while navigating existing regulations under regulatory supervision. Hybrid sandboxes, such as the Data Protection Panel and Sandbox in **Estonia**, provide a blend of innovation support and data governance. Led by the Ministry of Economic Affairs and Communications, an advisory board of experts assists public sector institutions in projects that involve complex issues related to personal data processing and reuse. In the **European Union**, the Metis Sandbox allows data holders to test their datasets for interoperability within existing data portal infrastructure, ensuring high quality and compliance with strict metadata standards. By reducing back-and-forth iterations with data aggregators like the Europeana team, the sandbox streamlines the preparation of datasets for publication. Additionally, it enables data holders to preview their datasets on the Europeana.eu platform, facilitating the exploration of potential use cases and synergies with existing records.

Sandboxes can operate at various levels of governance—local, national, and international. Local sandboxes address community-specific needs and contexts, like Zurich's AI sandbox, which assesses and implements AI projects while granting participants access to new data sources.[31]

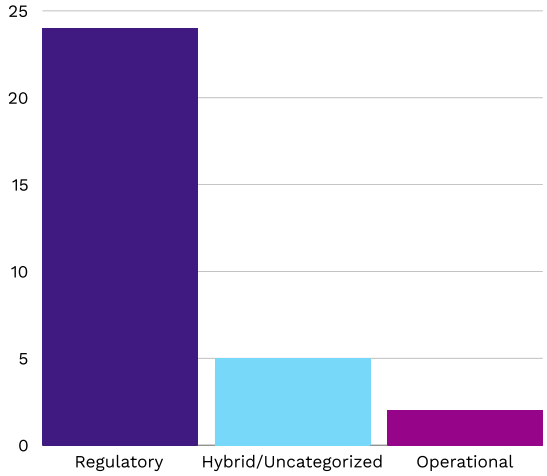
[29] Consumer Data Right (2024), *Consumer Data Right Sandbox*, Commonwealth of Australia.

[30] Ministry of Information and Communications Technologies, *Data Sandbox Collaborative Space*, Government of Colombia.

[31] Zurich (2024), *Innovation Sandbox for Artificial Intelligence (AI)*.

National sandboxes are the most common type seen today. In 2025, research by the Datasphere Initiative identified 23 countries that have implemented or are developing one or more national sandboxes for AI. Overall, there are 31 sandboxes for AI that have been developed or are underway, including 24 regulatory sandboxes, two operational, and five hybrid or uncategorized.[32] The graph below (Figure 1) illustrates this distribution of national AI sandboxes by type. Notably, regulatory sandboxes constitute the majority, highlighting a global emphasis on testing AI technologies within controlled policy environments.

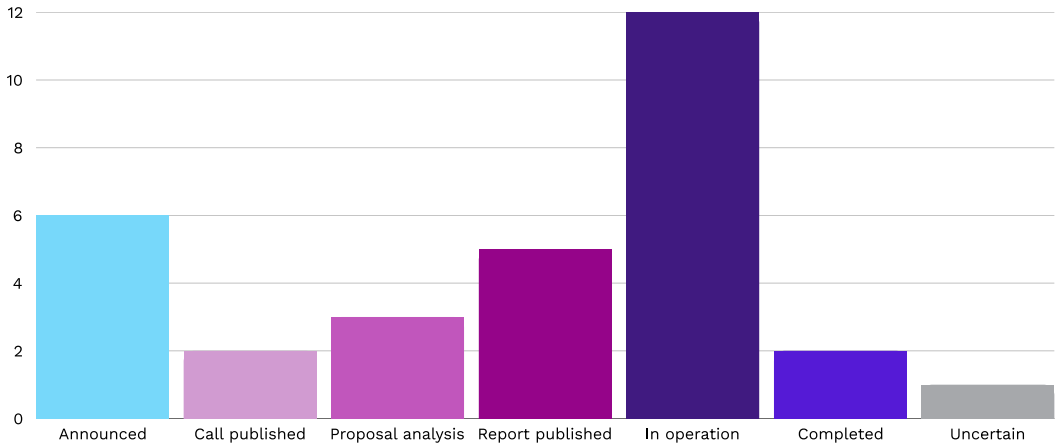
Figure 1. National AI Sandboxes by Type



Note: The total number of sandboxes represented in this graphic is 31, comprising 24 regulatory sandboxes, 5 hybrid or uncategorized sandboxes, and 2 operational sandboxes.

The national AI sandboxes mapped in this research are at various stages of development, as illustrated in the graph below (Figure 2). The stages reflect the lifecycle of each sandbox, from initial announcements to post-completion analysis. Specifically, six sandboxes are in the "announced" phase, indicating that these countries have announced their intent to launch a sandbox but have not yet moved forward with the implementation. There are two sandboxes in the "call published" phase, where countries have begun calling for applications to join the sandbox. Another three sandboxes are in the "proposal analysis" phase, where submitted proposals are being reviewed by a designated committee. The "in operation" phase is the most populated, with 12 sandboxes actively running. This is followed by two sandboxes in the "completed" phase, having concluded their testing or implementation period. Five sandboxes have reached the "report published" stage, where results and findings have been analyzed and published in formal reports. Lastly, there is one sandbox categorized as "uncertain," where the status and outcomes remain unclear.

Figure 2. Stages of Development of National AI Sandboxes



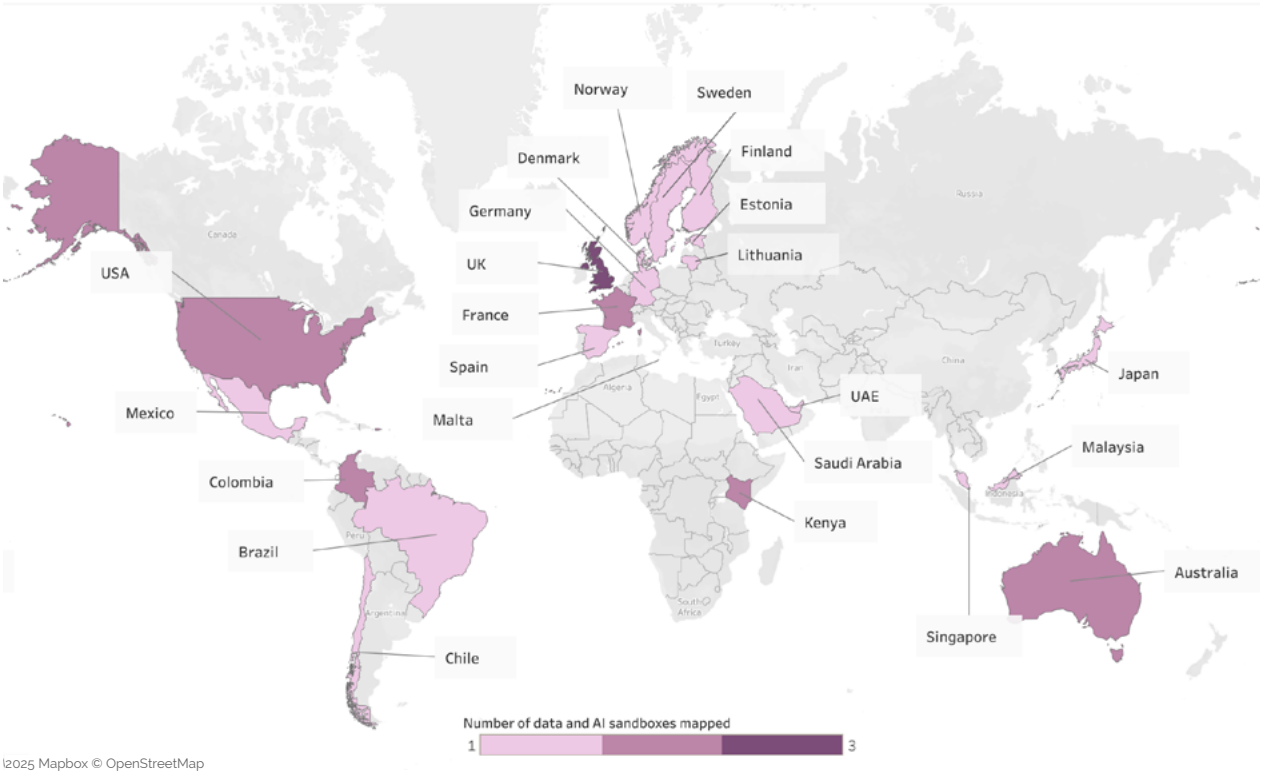
Note: The total number of sandboxes represented in this graphic is 31, distributed across various stages of development: 12 in operation, 6 announced, 5 with reports published, 3 in the proposal analysis phase, 2 with calls published, 2 completed, and 1 with an uncertain status.

[32] This does not include the numerous fintech sandboxes that exist around the world. Moreover, some countries have more than one sandbox.

While few **cross-border sandboxes** exist today, they hold immense potential to facilitate international collaboration and harmonize regulations across jurisdictions. The EU's Europeana Metis Sandbox exemplifies this potential. This cross-border operational sandbox allows data holders to test datasets for interoperability within the EU's existing data portal infrastructure. By ensuring datasets meet strict metadata standards and maintaining high quality, the Metis Sandbox reduces the need for multiple iterations with data aggregators like the Europeana team. It also provides a platform for data holders to preview their datasets, exploring potential use cases and synergies with existing records. In this way, cross-border sandboxes like the Europeana Metis Sandbox help improve data aggregation, quality, and workflow processes on an international scale, enabling smoother collaboration across borders in areas like cultural heritage and data management.[33] It is also expected that there will be some level of cross-border collaboration stemming from the implementation of the sandbox provision within the EU AI Act in the near future.

Below is a map with a global overview of the known sandboxes for AI innovation and a table with the cataloged information on sandboxes for AI Innovation around the world, as of January 2025 (Figure 3).

Figure 3. Map of Sandboxes for AI innovation around the world, as of January 2025



Note: The categorization of a sandbox as a sandbox for AI innovation is based on the explicit topic of focus or description of their objectives.

[33] Europeana (2022), *Metis Sandbox Training*, Europeana Knowledge Base.

Table 1. National Sandboxes for AI Innovation Around the World

Name of Sandbox	Country	Competent Authority	Years of Operation	Type	Topic/Focus
Consumer Data Right Sandbox	Australia	Australian Competition and Consumer Commission	2022-	Operational	Technical solutions to data sharing under the Consumer Data Right
Enhanced Regulatory Sandbox	Australia	Australian Securities and Investments Commission	2020-	Regulatory	Testing innovative financial services or credit activities, including AI-driven technologies
AI Regulatory Sandbox	Brazil	Brazilian National Data Protection Authority	In development	Regulatory	Machine learning (ML)-driven technologies, including generative AI
AI Regulatory Sandbox	Chile	Ministry of Economy, Development and Tourism	In development	Regulatory	Experimenting with new uses of AI within the framework of the National AI Policy
Sandbox on privacy by design and by default in Artificial Intelligence projects	Colombia	Superintendence of Industry and Commerce	Uncertain	Regulatory	AI projects in the design stage which involve the processing of personal data
Data Sandbox Collaborative Space	Colombia	Ministry of Information Technology and Communications	2021-	Regulatory	Pilot projects in Analytics and Big Data, empowering public entities to leverage Big Data technologies in a collaborative environment
Regulatory Sandbox on AI	Denmark	Danish Data Protection Authority	In development	Regulatory	Data protection
Data Protection [Panel and] Sandbox	Estonia	Ministry of Economic Affairs and Communications	Uncertain	Other	Projects that are a priority in the development of the digital state; data processing that impacts fundamental rights
Aurora AI Programme	Finland	Ministry of Economic Affairs and Employment	2017-	Regulatory	Testing AI solutions for the public interest
CNIL Personal Data Sandbox	France	National Commission on Informatics and Liberty (CNIL)	2021-	Regulatory	Projects developing innovative goods or services linked to the processing of health data (first edition); education technology (EdTech) (second edition); Projects involving the use of AI in public services (third edition)

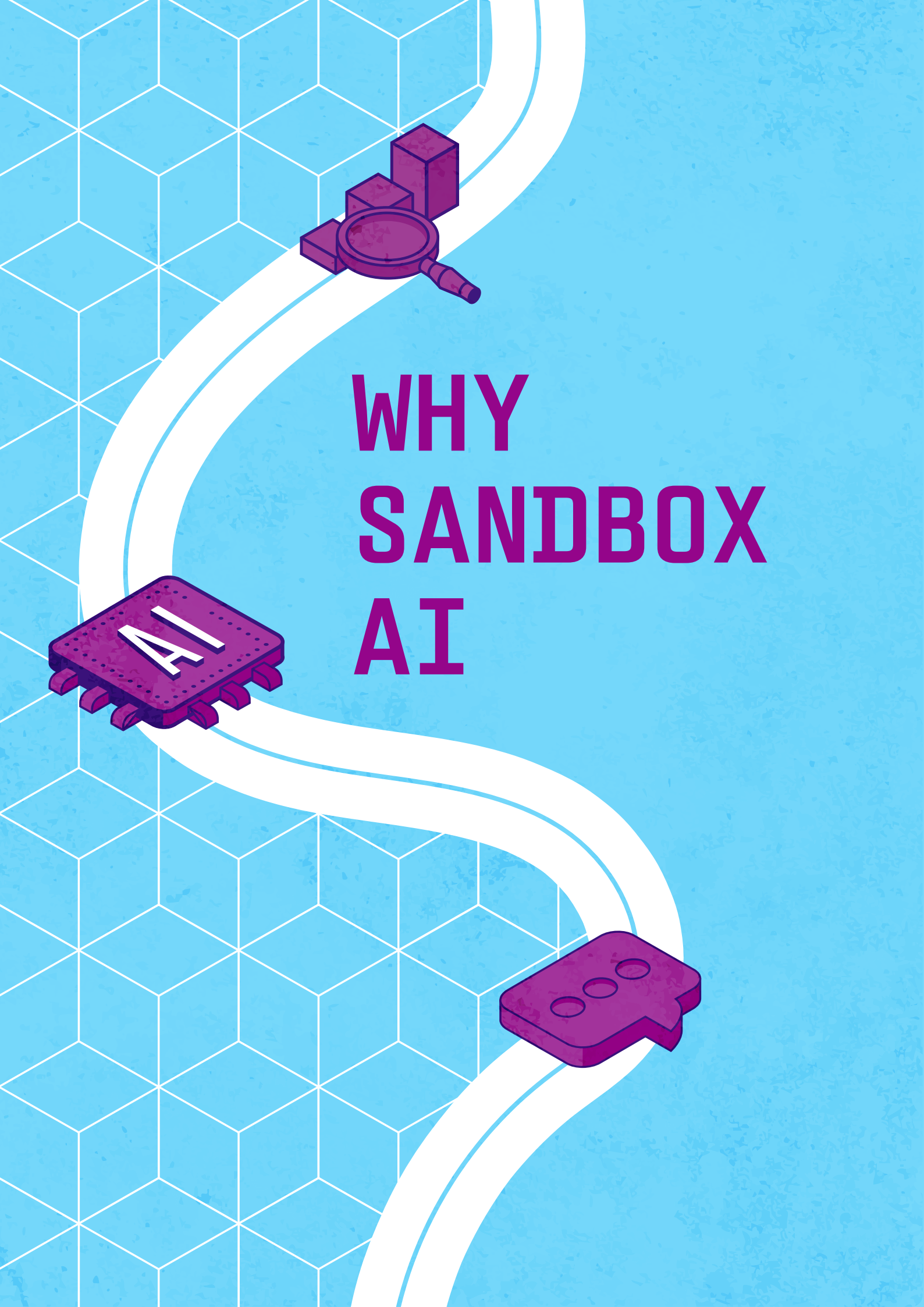
Name of Sandbox	Country	Competent Authority	Years of Operation	Type	Topic/Focus
AI Sandbox	France	CNIL	2023-	Regulatory	There are 8 projects that have been selected to participate in the sandbox, which focuses on AI in public services, that will receive personalized support with respect to emerging legal and technical issues relating to data protection and AI regulation
Sandbox - Delivery Robot Hamburg for CEP delivery	Germany	Office for Regulatory Sandboxes at the Federal Ministry for Economic Affairs and Climate Action	2022-	Unknown	Delivery robots, telemedicine, autonomous bus, eGovernment
Regulatory Sandbox	Japan	Cabinet Secretariat, Secretariat of New Form of Capitalism Realization Headquarters	2018-	Regulatory	Cutting-edge technologies and business models in any sector including AI
Regulatory Sandbox	Kenya	Communications Authority of Kenya	2024-	Regulatory	ICT products and services such as AI-driven services, e-learning platforms that utilize AI, e-health solutions, internet of things etc.
Regulatory Sandbox	Kenya	Capital Markets Authority	2024-	Regulatory	Innovative technologies in the finance and capital markets sector
AI Sandbox	Lithuania	Ministry of the Economy and Innovation of the Republic of Lithuania	2024	Regulatory	AI for the public sector
AI Sandbox	Malaysia	Ministry of Science, Technology and Innovation in collaboration with Nvidia and The Malaysian Research Accelerator for Technology & Innovation	In development	Hybrid	Artificial intelligence in general
Technology Assurance Sandbox	Malta	Digital Innovation Authority	2020-	Unknown	Emerging technologies

Name of Sandbox	Country	Competent Authority	Years of Operation	Type	Topic/Focus
Regulatory Sandboxes for AI	Mexico	Mexican Academy of Cybersecurity (AMCID)	In preparation	Regulatory	Exploring data, antitrust, and telecommunications regulations to coordinate a sandboxes approach to emerging technologies like AI
Regulatory Sandbox for Artificial Intelligence	Norway	Norwegian Data Protection Authority	2020-	Regulatory	Developing AI solutions that comply with data protection regulations (1st, 2nd, and 3rd editions); projects addressing regulatory uncertainties in complex data sharing, the EU GDPR's provisions on automated decision-making, secondary data uses, etc (4th, 5th editions)
Data and Privacy Regulatory Sandbox	Saudi Arabia	Saudi Authority for Data and Artificial Intelligence	2023-	Regulatory	Solution/service/business model that falls under Data and Privacy laws and regulations or is a Privacy Enhancing Technology (PETs) solution
Privacy Enhancing Technology Sandbox	Singapore	Infocomm Media Development Authority (IMDA)	2022-	Regulatory	Projects using PETs to collaborate on data
Generative AI Evaluation Sandbox for Trusted AI	Singapore	Infocomm Media Development Authority (IMDA)	2023-	Regulatory	Evaluation and Testing of trustworthy GenAI
Regulatory Sandbox on Artificial Intelligence	Spain	Ministry of Economic Affairs and Digital Transformation	In development	Regulatory	Artificial Intelligence
Regulatory Sandbox on Data Protection	Sweden	Swedish Authority for Data Protection	2023-	Regulatory	Topics related to "gray area issues" in data protection law
RegLab	United Arab Emirates	Ministry of Cabinet Affairs of the Future and the Dubai Future Foundation	2019-	Regulatory	New technologies such as AI – use cases involve self-driving vehicles, health data repository and disease registry, electric vertical takeoff and landing (eVTOL) aircraft

Name of Sandbox	Country	Competent Authority	Years of Operation	Type	Topic/Focus
Information Commissioner's Office (ICO) Regulatory Sandbox	United Kingdom	Information Commissioner's Office	2020-	Regulatory	Various including consumer health tech, immersive technology and virtual worlds, next generation search, personalized AI
AI Airlock	United Kingdom	Medicines & Healthcare Products Regulatory Agency	In development	Regulatory	AI as a Medical Device
NayaOne's AI Sandbox	United Kingdom	Conducted by NayaOne but endorsed by the Centre for Data Ethics and Innovation and Department for Science, Innovation and Technology	In development	Operational	NayaOne is a Sandbox-as-a-Service provider to tier 1 financial services institutions, world-leading regulators, and governments to allow them to test data and models in order to address key concerns in AI deployment
AI Testbeds	United States of America	Department of Energy	In development	Operational	The testbeds will be used to explore how the future of AI hardware may evolve to be more efficient and how the risks associated with the use of AI can be effectively managed and secured, as well as research on, development and testing of PETs
Advanced Analytics Platform for Machine Learning (CAP-M) AI Sandbox	United States of America	Science and Technology Directorate (S&T) - housed within the Department of Homeland Security (DHS) - and the Cybersecurity and Infrastructure Security Agency (CISA)	In development	Operational	The sandbox will allow researchers to collaborate and test analytical approaches and techniques in combating cyber threats

Note 1: This table includes only national sandboxes for AI, defined as those explicitly focused on machine learning, AI development, and data-driven solutions, as stated in their objectives or descriptions. Of the 66 sandboxes related to data, AI, or technology identified by the Datasphere Initiative, 59 are national, and of those, 31 specifically focus on AI—all 31 of which are captured here. By focusing on national AI sandboxes, this table highlights the countries running these initiatives and provides additional details on the sandboxes they operate, offering insights into how governments foster AI innovation through structured experimentation.

Note 2: The table includes key details on each sandbox's name, country of focus, competent authority (institution responsible for the sandbox), years of operation (active period), type (e.g., regulatory, operational, or hybrid), and topic/focus (in areas such as data protection, health, smart cities, ICT, and finance).


WHY
SANDBOX
AI
AI

The rapid expansion of AI technologies has heightened public awareness of a host of regulatory challenges that policymakers are struggling to address. Concerns around AI's rapid integration into various sectors are becoming well documented and stem from the systemic biases and potential discrimination, impacts on workforce[34] and the lack of transparency in how data is sourced and used, complicating accountability and fairness.[35]

The drive for more innovation in public sector regulation is not new. For many decades, public institutions have catalyzed pilot programs, experimental regulatory approaches and safe harbor provisions in sector-specific cases and regulatory frameworks to encourage some level of controlled spaces for innovation.

Sandboxes have emerged as a valuable governance tool for both public and private sectors. They provide a controlled space where new technologies, data-driven systems, and business models can be tested and refined under real-world conditions, enabling iterative learning and policy adaptation. By bridging the gap between the rapid pace of technological advancements and the slower evolution of regulatory frameworks, sandboxes can facilitate innovation while promoting safety and responsibility. Moreover, they have the potential to build trust among regulators, both within a single jurisdiction and, in a few cases across borders, and assure the public that new practices have been thoroughly scrutinized.

For AI, the sandbox model is especially promising. Given the speed of technological change, the uncertainty surrounding potential risks, and the wide array of actors involved in the development of sandboxes for AI, experimentation is essential to understanding how best to balance innovation with the need for safeguards. It is through continuous experimentation within these controlled environments that the challenges of AI governance can be identified, addressed, and refined, ensuring that the benefits of AI are realized in a responsible, ethical, and inclusive manner.

When considering why to use a sandbox model for AI it's important to consider the different incentives and challenges regulators, companies and civil society alike may have when deciding to design or participate in an AI sandbox.

Rationale for regulators

This section answers the question of why to adopt sandboxes by analyzing three key factors: the speed of technological change, the challenges of cross-sectoral and cross-border regulation, and the limitations of command-and-control models.

SPEED OF CHANGE

The rapid pace of AI technology development, exemplified by the frenetic evolution of GenAI, poses significant challenges to existing regulatory frameworks, which often lag behind technological advancements. The faster a technology spreads; the less control regulators have over the market to mitigate risks and protect consumers from harm.

[34] Korinek and Juelfs (2022), *Preparing for the (non-existent?) future of work*, Brookings Center on Regulation and Markets.

[35] Nikitina, S. (2024). *Multilingual, Not Monolithic: Why Diverse Language Data Matters for Generative AI*.

One example of emerging risks for data protection regulators relates to webscrapping. GenAI developers typically train their models on data gathered through extensive scraping of the open internet, presenting complex legal questions involving copyright and personal data protection laws.[36] The practice of web scraping can result in the infringement of intellectual property rights, both by training models on copyrighted material without permission and by generating content that may replicate or closely resemble copyrighted works.[37] Web scraping also holds implications for data protection laws; for example, if a model is trained on sensitive commercial or personal data, there is currently no way for the model to subsequently 'unlearn' that data, as required by some data protection laws.[38] EU data protection authorities have made significant strides in issuing guidelines on data scraping. For instance, the Dutch authority released a series of guidelines for data scraping performed by organizations and private individuals. The data scraped usually involves personal data, which is why the guidelines underscore that their processing activities must comply with the EU GDPR.[39] Similarly, the Italian data protection authority published recommendations for controllers to protect personal data from web scraping. The recommendations include: (i) creating restricted access areas; (ii) including dedicated provisions, such as prohibiting web scraping, in the terms of service; (iii) implementing traffic data monitoring; and (iv) implementing bot access restrictions.[40]

Moreover, AI's ability to generate content at scale, particularly through GenAI, complicates the information landscape by making it difficult to distinguish between real and fabricated content.[41] This can undermine trust and accountability in economic and social institutions, especially during elections, when actors can misuse unregulated systems to manipulate and misinform, threatening the integrity of democratic processes.

The pace of change in AI technology and its impacts on society calls for a solution that is resilient enough to adapt regulatory measures based on the type of innovation, associated risks, and the scope of new AI-based products. Indeed, the speed of change in the AI field requires regulators to innovate more in the process of drafting regulations, rather than focusing solely on the regulatory text itself. The information asymmetry between regulators and innovative companies requires the adoption of collaborative strategies by regulators. These strategies should aim to better understand the nature of new AI systems, how they function, the purposes for which they are developed, the appropriate metrics for measuring risks, and the types of data required to ensure the effective deployment of governance tools. Based on this information, regulators can adapt their tools to manage risks and develop more effective control and monitoring measures.

---

[36] Tiedrich (2024), *The AI data scraping challenge: How can we proceed responsibly?*, OECD

[37] Holloway, Cheng, and Dickenson (2024), *Will copyright law enable or inhibit generative AI?*, World Economic Forum.

[38] Falconer (2023), *Privacy in the age of generative AI*, Stack Overflow.

[39] Autoriteit Persoonsgegevens (2024), *Richtlijnen scraping door private organisaties en particulieren*, Dutch Data Protection Authority.

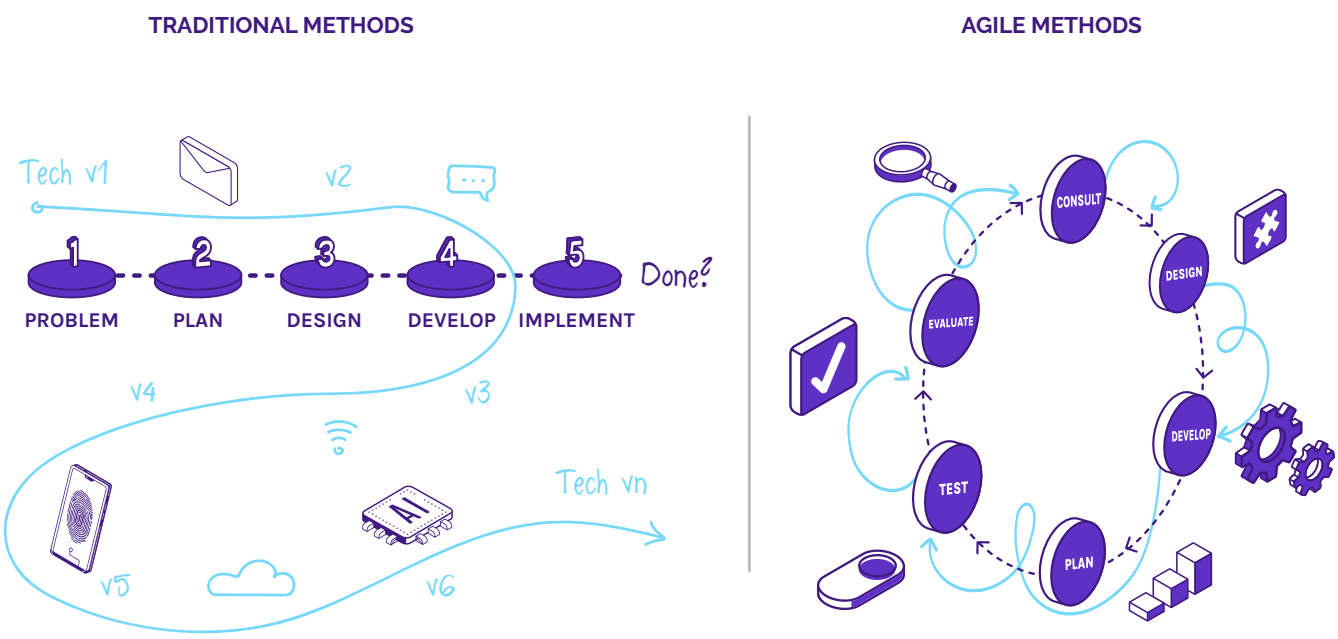
[40] Garante per la Protezione dei Dati Personali (2024), *Web scraping ed intelligenza artificiale generativa: nota informativa e possibili azioni di contrasto*, Italian Data Protection Authority.

[41] Ofcom (2024), *Future Technology and Media Literacy: Understanding Generative AI*.

A lack of transparency around how leading-edge models are trained and algorithms are developed exacerbates these challenges, making it difficult to ensure accountability and trust in AI systems. This opacity makes it difficult to address issues such as "hallucinations," where GenAI produces fabricated content. Such unreliable outputs pose significant challenges for determining liability and ensuring the reproducibility of scientific results.[42] In fact, there might be a need to augment the notion of "peer review" in the era of AI, and rethink about the new institutions that might be needed to ensure the reproducibility of science in the digital era.

Sandboxes aim to bridge the gap between rapid technological innovation and the slower pace of regulatory evolution, facilitating innovation while ensuring safety and compliance (Figure 4). They can build trust between regulators within a country or between countries, and can also build public trust not only by giving consumers greater assurance that novel practices that emerge from the sandbox have been subjected to regulatory scrutiny, but also by identifying when these practices are not compliant with regulations and intervening to prevent their proliferation in the marketplace.[43] As AI and other data-driven technologies continue to integrate into various global sectors and applications, the strategic use of sandboxes could become a cornerstone of the effective governance of data and emerging technologies in general.

Figure 4. Traditional methods vs agile methods of regulation



[42] Aaronson and Thakur (2024), *We Need to Talk about AI Reproducibility*, Centre for International Governance Innovation.

[43] Organisation for Economic Cooperation and Development (2022), *Harnessing the power of AI and emerging technologies*, OECD Digital Economy Papers.

CROSS-SECTORAL AND CROSS-BORDER CHALLENGES

As AI technologies become increasingly integrated into the critical infrastructure of various sectors — from public service delivery and healthcare to military and commercial decision-making — the need for comprehensive and adaptive governance frameworks becomes urgent.

The global nature of AI development further complicates regulation, as the essential computational infrastructure, expertise, and data often flow across national borders. This necessitates an unprecedented level of international cooperation to avoid a fragmented global regulatory landscape. Such coordination is currently lacking, with political polarization and fundamental issues like the definition of AI itself still lacking a commonly agreed standard with clearly defined terms.[44] Most legislators, especially in the European Union, seem to be converging on the definition of AI published by the Organisation for Economic Co-operation and Development (OECD), but a remaining challenge is agreeing on a definition of concepts like *autonomy* and *adaptiveness*[45] and fostering global convergence or interoperability on AI standards and governance approaches.

Regulatory strategies and data governance frameworks differ significantly between regions.[46] Inconsistent regulatory environments can stifle innovation by small and medium-sized enterprises (SMEs) and create competing standards, erecting barriers to new entrants and limiting data sharing crucial for advanced AI development. They can also place companies in regions with stringent AI regulations at a competitive disadvantage, increase compliance costs significantly, and enable AI developers to "forum shop" or exploit regulatory differences for the most beneficial regulations. While there are promising initial efforts, such as the Council of Europe Convention on AI – the first international treaty of its kind –,[47] much more international coordination is needed to harmonize regulations, balance protection from harm with scientific innovation, and ensure the global adoption of AI technologies.[48]

By enabling developers and regulators to collaboratively test new systems and datasets across borders, sandboxes can help ensure that AI technologies are safe, compliant, and aligned with societal values before their widespread deployment. Given AI's cross-border development and deployment – with data coming from various countries, systems developed in other places and deployed across regions and sectors –, traditional national and sectoral governance approaches can quickly become obsolete. A collaborative cross-border and cross-sectoral approach is essential, especially since a single regulatory body, perspective, or set of cultural values cannot effectively govern AI. For example, a competition regulator addressing monopolistic practices cannot alone tackle the ethical, social, and economic implications of AI. Additionally, concepts like transparency vary significantly across cultures, making it impractical for one country to adopt another's AI governance framework wholesale. Sandboxes – and

---

[44] Marsden (2017), *Artificial Intelligence Defined: Useful list of popular definitions from business and science*, digitalwellbeing.org.

[45] OECD AI Principles. *https://oecd.ai/en/ai-principles*.

[46] Benizri et al (2023), *A Comparative Perspective on AI Regulation*, Lawfare.

[47] Council of Europe (2024). *Council of Europe adopts first international treaty on artificial intelligence - Portal*.

[48] Council of Europe (2024), *The Framework Convention on Artificial Intelligence*, Council of Europe.

cross-border sandboxes like Ecobank's sandbox in Africa[49] or the Europeana Metis Sandbox[50]– could further foster collaboration across jurisdictions, stakeholders and sectors. These cross-border sandboxes not only encourage alignment and harmonization of regulatory approaches, but also allow space for each country to embed its unique cultural values and regulatory priorities, ensuring that diverse cultural and value systems are respected while still fostering international cooperation.

Additionally, AI cannot be regulated from a single sectoral perspective, given the multiplicity of involved regulators and its broad pool of training data and interconnected applications. The level of sensitivity of the data used (e.g., risks of negative impacts, privacy considerations, and varying cultural norms) as well as its accessibility (e.g., how transparent, interoperable, accurate or complete is the data) are all factors[51] that vary across sectors and hence points to the need of a multisectoral approach. For instance, AI health applications require coordination between health and ICT ministries and sandboxes can play a role in connecting multiple regulators around a regulatory challenge.

Both operational and regulatory sandboxes allow stakeholders to explore innovative practices that do not fit neatly within traditional regulatory frameworks. This flexibility is crucial in a landscape where AI and data-driven technologies evolve rapidly, helping to ensure that data governance keeps pace with technological advancements and societal needs. Sandboxes can also support AI development by facilitating rights-respecting data sharing and access by incubating technologies such as Privacy-Enhancing Technologies (PETs).[52]

THE NEED FOR AGILITY: RESPONSIVENESS AND ADAPTABILITY

Traditional regulatory methods and approaches are being challenged in the context of AI. Governments are slowly understanding a need to test and run AI systems to identify the potential challenges and their interaction with a regulatory environment, often shifting the regulatory paradigm in the process.

For example, as AI technology becomes more advanced, the potential for monopolistic control over information and technology increases.[53] Large corporations with access to vast amounts of data and computational power, as well as with greater economic resilience and ability to comply with various national regulatory regimes, can dominate the market, stifling innovation and creating barriers for smaller entities or competitors.[54] Furthermore, the increasing prevalence of partnerships among key players could reinforce or extend existing market power positions across the value chain.[55] This concentration of power across the global supply chain raises questions about fairness, accessibility, and the overall health of the competitive landscape.

---

[49] *Ecobank's Fintech Sandbox*.

[50] Europeana (2022), *Metis Sandbox Training*, Europeana Knowledge Base.

[51] World Economic Forum. Global Future Council on the Future of Data Equity. (2024). *Advancing Data Equity: An Action-Oriented Framework*. White paper.

[52] Datasphere Initiative (2022), *Sandboxes for data: creating spaces for agile solutions across borders*, Datasphere Initiative.

[53] Price (2023), *Allowing big tech to monopolize AI is risky business*, Digital Content Next.

[54] Staff in the Bureau of Competition & Office of Technology (2023), "*Generative AI Raises Competition Concerns*, Federal Trade Commission.

[55] Competition & Markets Authority (2024), *AI Foundation Models Update paper*, Government of the United Kingdom.

In this context, sandboxes could play a crucial role in supporting competition regulations by providing a controlled environment where emerging technologies can be tested without allowing dominant players to distort the market.[56] While sandboxes traditionally focus on technical issues like safety and privacy, they can also be designed to foster fair competition. By enabling smaller firms to test their innovations alongside larger corporations, sandboxes can create opportunities for new entrants to demonstrate their technologies and ideas, leveling the playing field. However, it's important to ensure that participation in sandboxes does not unfairly advantage certain companies, especially those with more resources. Regulatory frameworks for sandboxes should be structured in a way that promotes innovation from a diverse range of players, helping to prevent market dominance by a few large entities and encouraging a more competitive and dynamic ecosystem.

Overall, existing regulatory frameworks – including privacy, intellectual property, data protection, content moderation – may prove outdated and often unfit to address the challenges highlighted before. Either because they were developed without consideration of new AI and technological advancements or simply because they cannot approach the systemic societal and economic issues that digital technologies like AI are giving rise to, traditional regulations are falling short and need to be revisited.

Navigating these challenges requires developing agile regulatory approaches that are as dynamic and adaptable as the technologies they aim to govern. This includes creating frameworks that respond to current AI capabilities and are flexible enough to evolve with future technological developments.

The experimental nature of sandboxes can be invaluable for testing innovative approaches to the governance of emerging technologies. They provide a controlled setting to experiment with and understand the implications of new technologies and data practices. This hands-on approach allows for real-time problem-solving and adaptation, which is crucial in a rapidly evolving digital landscape.

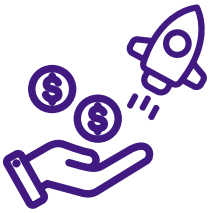
Regulatory sandboxes have demonstrated significant benefits in sectors like fintech over the past decade and these advantages are increasingly being applied to other areas of technology and policy.[57] They provide a structured environment where businesses can engage with regulators and receive tailored feedback on compliance requirements. They reduce uncertainty and facilitate market entry, as long as risk mitigation strategies like knowledge-sharing mechanisms, sound eligibility criteria, awareness campaigns and transparent results-sharing are effectively put in place to prevent unfair advantages. This enables companies to refine their technologies and business models with greater confidence. Additionally, sandboxes offer a platform for iterative testing and adaptation, allowing innovators to adjust their approaches based on real-time insights and regulatory advice, which accelerates the development process and helps ensure market readiness. For regulators, they offer direct exposure to advanced technologies and business models, preparing them for future challenges. Additionally, they can help regulators develop more informed and adaptive regulatory frameworks, foster industry collaboration, and build capacity within regulatory bodies.[58]

---

[56] Crampes, C., & Estache, A. (2023), *Efficiency vs. equity concerns in regulatory sandboxes*, Toulouse School of Economics.

[57] Appaya and Haji (2020), *Four years and counting: What we've learned from regulatory sandboxes*, World Bank Blogs.

[58] Datasphere Initiative (2022), *Sandboxes for data: creating spaces for agile solutions across borders*, Datasphere Initiative.



Rationale for companies

While the benefits of sandboxing from a regulators perspective are slowly emerging and becoming more broadly recognized, private sector actors may need to have incentives and trust building mechanisms to participate in a sandbox. SMEs may not have the time and resources to participate in a sandbox which can require time and personnel to contribute effectively. At the same time, larger companies may fear attracting regulatory scrutiny and feel uncomfortable sharing information in the public domain in the form of the reports or press releases that may report on the sandbox.

Regulators must clearly articulate the benefits for participating in a sandbox to the private sector including large companies who may have strong legal and policy teams to navigate regulatory uncertainty instead of participating in a direct analysis within a sandbox. Alternatively, for start-ups and SMEs, with stretched time and resources, if regulators are able to offer funding or the limited sharing of data-sets, as incentives for sandbox participation, that might also make them more appealing.

One of the main advantages of joining an AI regulatory sandbox is the opportunity to engage directly with regulators. Even if formal AI regulations are not yet in place, sandboxes offer a unique chance to gain insight into potential regulatory expectations and align products with future legal frameworks. This proactive approach helps reduce compliance risks once regulations are in place. Furthermore, operational sandboxes may provide companies with access to a pool of data or a safe space to test and refine their technologies in real-world conditions without the fear of penalties, accelerating product development. Additionally, they foster collaboration between innovators and industry stakeholders, enabling companies to identify and address potential operational challenges before full-scale deployment. Companies may also gain access to formal verification or quality certifications that enhance their credibility and legitimacy, demonstrating to stakeholders—such as customers, investors, and partners—that their solutions have been rigorously tested and validated within the framework of a regulatory sandbox.

Participating in either an operational or regulatory sandbox requires companies to evaluate several key factors:

1. **Alignment with sandbox objectives:** Companies must assess how well their product or process aligns with the sandbox's testing goals. The alignment between the sandbox's objectives and the company's development stage, as well as the type of testing required, enhances the company's chances of benefiting from the experience.

2. **Qualified personnel and ongoing engagement:** Companies should ensure they have qualified personnel who can actively participate in the sandbox testing process. Effective involvement demands continuous commitment, expertise, and the flexibility to adjust the product based on feedback gathered throughout the testing phase. Knowledge and assessment of the potential financial or associated costs of participating in a sandbox can also be a significant factor to assess.

3 **Data confidentiality and competitive advantage:** It is crucial for companies to assess the level of access granted to the data shared with the regulatory authority. Maintaining confidentiality and safeguarding competitive advantages is vital for protecting the company's market position and avoiding potential misuse of proprietary information.

4 **Long-term impact on regulatory relationships:** Companies should consider how insights gained from the sandbox will affect future acceptance by regulators and other stakeholders. Understanding how the sandbox experience will influence relationships with regulatory bodies and bolster market confidence is essential for shaping the company's strategic position.

Rationale for civil society

Civil society plays a crucial role in ensuring that AI governance is inclusive, transparent, and accountable to the public interest. **While sandboxes have primarily been framed as tools for policymakers and industry, their potential to serve as a mechanism for broader societal engagement is often overlooked.** Sandboxes can provide civil society organizations (CSOs) with a unique opportunity to contribute to AI governance by scrutinizing risks, advocating for human rights protections, ensuring that emerging technologies align with public values and guaranteeing that underserved and underrepresented communities' perspectives are heard.

By fostering collaboration between regulators, companies, and civil society groups, sandboxes can serve as a space for surfacing public concerns and integrating them into AI governance frameworks. In fact, **one of the key advantages of engaging civil society in sandboxes is their ability to bring transparency into AI development processes**. Given the opacity of many AI systems — particularly in relation to data collection, algorithmic decision-making, and model training — civil society participation in sandboxes can help identify potential biases, ethical concerns, and unintended social consequences before these technologies are widely deployed.

Meaningful civil society participation in sandboxes requires intentional design choices. CSOs often lack the resources, technical capacity, and access to proprietary data that companies and regulators possess. In fact, none of the sandboxes researched to date have actively engaged civil society as key stakeholders in their processes. **Equipping civil society with the necessary skills and tools to engage effectively in sandbox processes is thus essential** to ensuring AI sandboxes generate meaningful benefits aligned with the public interest.

Sandboxes must include mechanisms for public interest representation, such as structured participation from human rights organizations, consumer protection groups, and digital rights advocates. Transparency in sandbox processes—including public reporting on outcomes, clear documentation of lessons learned, and open engagement with affected communities—is essential to ensuring that AI sandboxes contribute to broader democratic oversight rather than reinforcing industry-driven governance.

Sandboxes can help civil society organizations develop stronger evidence-based advocacy. By engaging in sandbox processes, CSOs can test policy recommendations, monitor AI's real-world impacts, and gather empirical data to support their impact. In contexts where AI governance remains fragmented or underdeveloped, sandboxes could serve as critical mechanisms for bridging knowledge gaps and fostering more inclusive policymaking.

While AI sandboxes offer both opportunities and challenges for CSOs—who often lack the resources to participate fully—their inclusion is crucial to ensuring that AI innovations prioritize public value, equity, and social good. Ultimately, **integrating civil society perspectives into AI sandboxes enhances their legitimacy and effectiveness.** By ensuring that AI governance is not solely shaped by industry and regulatory interests, but also by those advocating for fundamental rights and public accountability, sandboxes can serve as a powerful tool for advancing both innovation and societal well-being.

In 2024, the Institute of the Future of Work (IFOW) announced the creation of the Responsible AI Sandbox.[59] This civil society-led sandbox, which will be the first of its kind, could set a precedent for how civil society can contribute to shaping responsible AI adoption through sandboxes. Its collaborative model—engaging regulators, industry partners, and experts—offers active testing and evaluation of AI applications in the workplace. This approach not only strengthens public interest-driven AI use and regulation but also ensures that AI's impact on work is evaluated through a lens of fairness, accountability, and worker well-being.

"SANDBOXING": A NEW MINDSET

It is important to recognize that sandboxes pose a new mode of interaction for public and private actors. For this reason, the when, why, and how to sandbox can be unclear to regulators and companies, who may have understandable hesitancy to engage in new modes of interaction. Therefore, it is essential for regulators and companies to form a new mindset for the application of sandboxes that considers the key elements of a sandbox. Firstly, the question of when to sandbox is an important consideration. For example, regulators may consider sandboxing when there is a new regulation as this creates an alignment and incentive by the public authority to improve its current consultation mechanisms, which may be unsuitable for the rapid pace of technology development. On the other hand, a regulator may also consider implementing the use of sandboxes prior to or during the development of a regulation as having a more exploratory approach early on may be critical in shaping a regulation that adapts to the fast-paced nature of technological developments.There is no one answer to when to sandbox, but it is clear that flexibility is needed to adjust to the objectives set out by the regulators and companies.

To understand why and how to use AI sandboxes, it's crucial to first shift the regulatory culture. Indeed, regulatory experiments require a new mindset from regulators—one focused on continuous learning and adaptive responses based on data-driven solutions. They also introduce a new approach to risk and failure, where the focus is not on punishment, but rather on viewing failures as learning opportunities. This shift is only possible with improved monitoring tools and agile responses, enabling regulators to quickly correct mistakes and minimize damage. This approach requires a change in the relationship between regulators, businesses, and civil society, prioritizing cooperation over control. Transparency must guide the rules of engagement between regulators and other stakeholders so that trust is built.

[59] Institute for Future of Work (2024), *Responsible AI Sandbox*.

Sandboxes contribute to this by setting clear goals for participants, regulators, and civil society to collaborate, establishing processes to collect data and foster trust. Regulators must be prepared to use the data gathered throughout the process to refine their activities. As such, they need to incorporate regular decision reviews into their routines, ensuring these decisions are informed by data. This transformation requires training and new models to help regulators improve their processes and build the capabilities necessary for success. Cultural change is essential for sandboxes to deliver the expected results. This is one of the key challenges for the success of sandboxes in general, and the same holds true for AI.

Quotes from consulted stakeholders

We usually start a sandbox for a very specific purpose. And usually it is in the situation where we see that there is interest in moving into some new technology or adoption of certain frameworks. And there's also uncertainty from the regulator side as well."

"The sandbox is an important tool for us to understand how the regulations and the policies will intersect with the actual implementation."

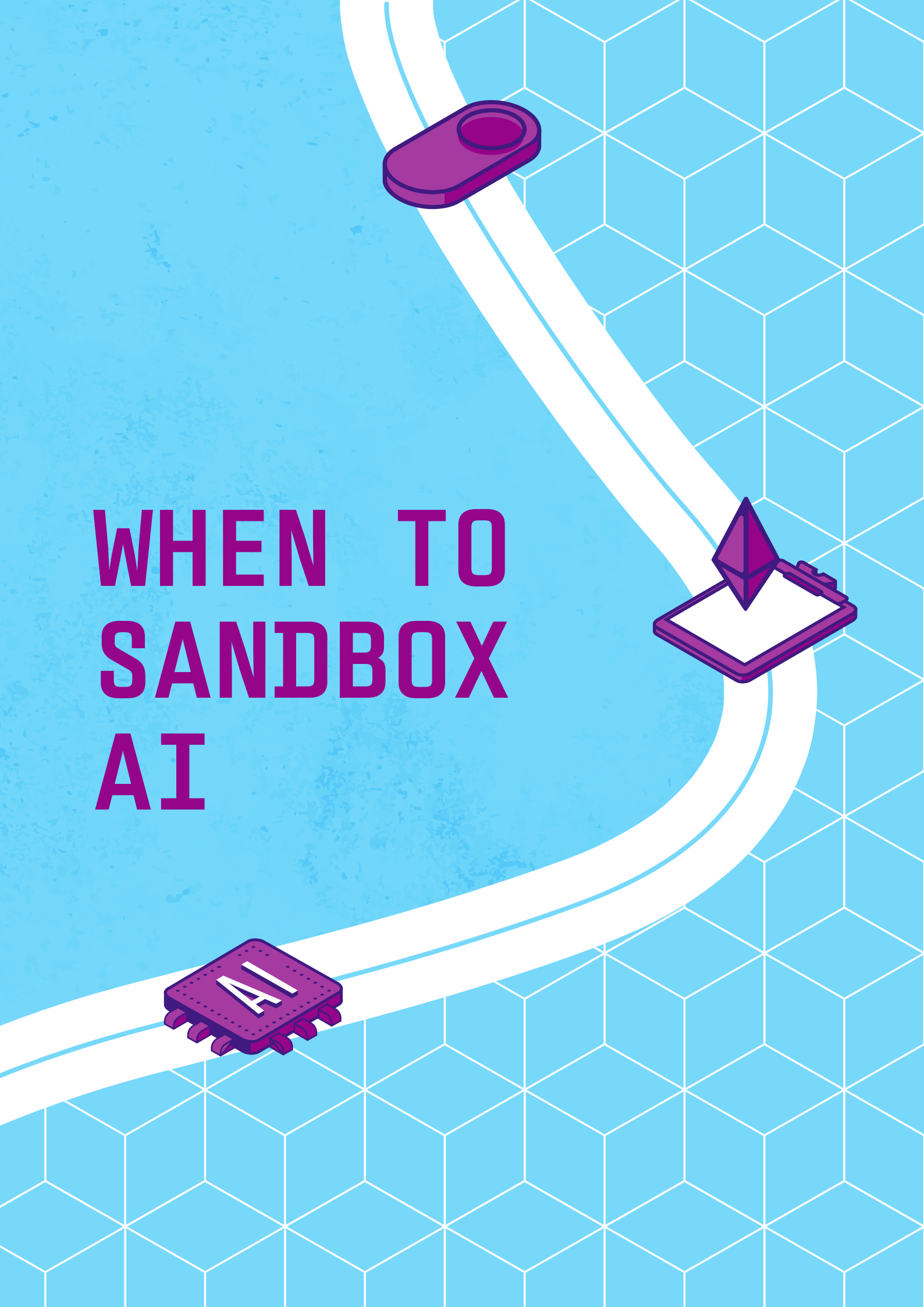
"I think companies have preferred to put their house in order more privately, and then try not to attract regulatory scrutiny."

"In the EU's AI Act the text has certain provisions that seem to provide a kind of a partial safe harbor for enforcement, which I thought was interesting and could make regulatory sandboxes more attractive."

"We see that as really a trust building mechanism to bring the regulators and the actors together. To create this safe environment where they will know specifically what will be shared. What mechanism, what risks are being taken into consideration."

"At least it's like a testing ground. You know that you can test your solution before you put more money into certification.

WHEN TO SANDBOX AI



Answering the question of when to implement a sandbox should not be framed as an opposition between regulation and innovation, but rather as a preparedness test. In the context of experimental regulation, the timing of launching a sandbox depends on the specific goals to be achieved. However, the objectives of fostering innovation, creating a trusted testing space, and encouraging adaptive learning are not necessarily distinct, but could all be addressed by the same sandbox over time, and even simultaneously, depending on the regulations. These goals will always remain central, regardless of when the initiative is launched, and have far-reaching implications for the timing and practice of sandbox development and implementation. Broadly speaking, there are three different moments when sandboxes could be implemented: before the development of regulation, during the development of regulation, or to test compliance of an existing regulation (Figure 5).

Figure 5. Types of sandboxes based on moment of development regarding regulation

1 Sandboxing before the development of a regulation

Many regulators wonder whether they can design a sandbox before a specific regulation related to that innovation has been put in place, whereas others see this as the start of understanding better the scope and aspects of a regulatory effort. In the case of AI, existing laws on product safety and data protection already address many challenges, providing a legal foundation for sandboxes. These sandboxes, in turn, help identify gaps in existing legislation and highlight areas where updates or new laws may be needed. Another reason for a sandbox may be to explore and research key issues before developing or even embarking on a regulation, recognizing that every regulator must work within the framework of existing laws. This allows them to identify gaps that need to be addressed in future regulations. A regulator also may set up a sandbox as a means to encourage innovation or co-develop policies between regulators or between regulators and other stakeholders to further the responsible development of science and technology. This moment can be beneficial for situations where the technology has not yet reached a high level of diffusion in society, and when the regulator still lacks knowledge about the new technology or its applications.

The experiences of Brazil and Norway illustrate the adoption of sandboxes before the creation of dedicated AI regulations. Additionally, they highlight a common scenario resulting from the cross-sectoral nature of AI: the intersection between data protection frameworks and future AI regulations.

The Brazilian case is particularly interesting because its objective is to understand how algorithmic transparency can be demonstrated.[60] This experiment is based on two provisions of Brazilian data protection legislation (Articles 6 and 20), but the results of the sandbox are expected to go beyond mere compliance objectives. It could provide lawmakers with valuable evidence on how this issue may be addressed in future AI legislation.

The Norwegian Data Protection Authority launched its sandbox initiative as part of the National Strategy for Artificial Intelligence. Since 2023, the sandbox has evolved into a permanent initiative.[61] In December 2023, the fifth edition of the program was launched, highlighting the growing interest from companies in participating and the valuable contributions they make to enhancing the authority's understanding of the challenges in implementing responsible AI systems. The establishment of a continuous learning environment focused on AI systems and their connection with personal data protection laws allows the Data Protection Authority to stay updated on technological advancements. This environment also provides an opportunity to test solutions from previous program editions in the context of new projects. As a result, it generates valuable evidence that will be crucial in shaping adaptive and effective regulatory responses.

2 Sandboxing during the development of a regulation

In developing AI regulation, insights drawn from various sandbox models can offer valuable guidance, providing a framework for balancing innovation with ethical and legal considerations. One such example is the UK government's AI Opportunities Action Plan,[62] released on January 13, 2025, which emphasizes its commitment to supporting AI innovation while ensuring robust regulatory oversight. As part of this plan, the government pledges to fund regulators to enhance their capabilities, allowing them to better manage and support AI growth in priority sectors such as autonomous vehicles, drones, and robotics. This funding will enable regulators to establish pro-innovation initiatives, including regulatory sandboxes, that help companies safely test AI-driven products in real-world environments. Additionally, the government intends to require all regulators to report annually on how they have facilitated AI innovation and growth within their respective sectors.

In the Global South, Singapore's regulatory model for AI is built on adaptability and collaboration, with a strong focus on fostering innovation while ensuring ethical and legal compliance. This approach requires significant commitment from government agencies, which must not only understand the specific challenges faced by companies but also proactively identify and address potential ethical and

[60] ANPD's Call for Contributions to the regulatory sandbox for artificial intelligence and data protection in Brazil is now open: *https://www.gov.br/anpd/pt-br/assuntos/noticias/anpds-call-for-contributions-to-the-regulatory-sandbox-for-artificial-intelligence-and-data-protection-in-brazil-is-now-open*

[61] Norwegian Data Protection Authority (2023). *Evaluation of the Norwegian Data Protection Authority's Regulatory Sandbox for Artificial Intelligence*.

[62] Privacy Laws & Businesses (2025). *Government commits to fund regulators to enable AI growth*.

legal concerns that could arise within AI sandboxes. These controlled environments allow companies to test and refine their AI technologies, with regulators working closely to provide tailored solutions for each case. When necessary, the regulatory framework itself is adjusted to better align with emerging challenges and technological advancements.

Understanding Singapore's approach to sandboxes requires an appreciation of the broader national AI ecosystem. In 2019, Singapore launched its National AI Strategy (NAIS), which outlined key goals and principles for the development of AI in the country. A central objective of this strategy was to create a progressive and trusted environment for AI innovations, balancing technological advancement with ethical standards and public confidence. To support this vision, Singapore introduced its Model AI Governance Framework,[63] which set a global benchmark for companies and captured the attention of policymakers worldwide. Integral to Singapore's regulatory approach is its collaboration with the private sector, which plays a critical role in shaping AI policy and ensuring that emerging technologies align with the public interest.

Rather than adopting a single overarching law, Singapore's regulatory strategy for AI is sector-specific, allowing for tailored, industry-driven regulations that address the unique challenges posed by AI in areas such as finance, healthcare, and transportation. The Infocomm Media Development Authority (IMDA), which leads AI governance initiatives, has embraced an experimental approach, working closely with companies and experts to develop effective guidance. This collaborative effort resulted in the creation of various initiatives, including the launch of a sandbox for Privacy-Enhancing Technologies (PETs) in 2022.[64]

In 2023, IMDA expanded its efforts further with the introduction of AI Verify, a governance testing framework paired with a software toolkit.[65] This initiative allows businesses to assess their AI systems for compliance with ethical and regulatory standards, ensuring that technologies are transparent, robust, and aligned with Singapore's governance principles. In 2024, IMDA launched the Generative AI Evaluation Sandbox, developed in partnership with another government agency tasked with fostering innovation. This sandbox focuses on testing generative AI models, providing a safe space for businesses to explore cutting-edge technologies while ensuring adherence to regulatory frameworks.

3 Sandboxing to test compliance with an existing regulation

The EU AI Act requires member states to establish AI sandboxes as part of its comprehensive strategy to regulate AI across the EU. Article 57 of the Act details how these sandboxes are intended to provide a controlled testing environment in which innovators and regulators will work together to identify risks and ensure compliance with the EU AI Act and potentially other EU regulations.[66] Sandboxes aim to foster

[63] Infocomm Media Development Authority (2024), *Singapore proposes framework to foster trusted Generative AI development*, Infocomm Media Development Authority.

[64] Infocomm Media Development Authority (2022), *Privacy Enhancing Technology Sandboxes*, Infocomm Media Development Authority.

[65] Infocomm Media Development Authority (2023), *First of its kind Generative AI Evaluation Sandbox for Trusted AI by AI Verify Foundation and IMDA*, Infocomm Media Development Authority.

[66] EU Artificial Intelligence Act (2024), *Article 57: AI Regulatory Sandboxes*, Future of Life Institute.

innovation and competitiveness, as well as support legal certainty, facilitating regulatory compliance, and promoting the sharing of best practices. The insights gained from these sandboxes are intended to better enable regulators to effectively apply the EU Act to emerging sectors and use cases, and where relevant, may be aimed at influencing policy amendments.[67]

Spain has been a pioneer of AI sandboxes in the EU, which was inaugurated by the enactment of Royal Decree 817/2023, which established the "RD Sandbox."[68] This controlled environment enables the testing and evaluation of high-risk AI systems, offering participants guidance on compliance and fostering innovation. The insights from the RD Sandbox are expected to culminate in a report of good practices that will inform future national regulations. Additionally, Royal Decree 729/2023 has created the Spanish Agency for the Supervision of Artificial Intelligence (AESIA), a dedicated body tasked with ensuring compliance with AI regulations, conducting inspections, raising awareness, and providing expert advice.[69] While specific AI laws in Spain are still under development, the National AI Strategy (2020–2025) provides a cohesive framework aligning state, regional, and sectoral strategies with EU policies. By embracing these measures and adhering to international principles such as those of the OECD, Spain has demonstrated its commitment to advancing ethical, innovative, and well-regulated AI development through its sandbox.

When to Implement a Sandbox from a Private Sector Perspective

AI sandboxes are usually initiated by government agencies or authorities but are also sometimes developed by private companies to experiment with regulatory frameworks and emerging technologies. While traditionally used as a regulatory tool for governments to test commercialized AI applications, sandboxes are increasingly being implemented in diverse contexts to address challenges and harness the benefits of AI development.

For instance, Harvard University's operational AI sandbox provides students with a controlled environment to explore generative AI tools in alignment with university guidelines, featuring access to leading LLMs and capabilities like data visualization and image generation.[70] Similarly, Meta's AI sandbox allows advertisers to experiment with generative AI features such as text variation and background generation to improve ad performance and usability.[71] In New Zealand, the Nelson AI Sandbox offers a secure space for individuals to test innovative AI applications,[72] while MITRE's Federal AI Sandbox supports the U.S. government by providing AI supercomputers to train specialized models for weather prediction, cybersecurity, and benefits processing.[73] These examples demonstrate how AI sandboxes can address diverse challenges while fostering innovation. To ensure responsible AI development, open communication between governments and organizations managing AI sandboxes is critical for enhancing transparency, collaboration, and trust.

---

[67] EU Artificial Intelligence Act (2024), *Article 57: AI Regulatory Sandboxes*, Future of Life Institute

[68] White & Case (2024). *AI Watch: Global regulatory tracker - Spain*.

[69] Ibidem.

[70] Harvard University (2024). *AI Sandbox | Harvard University Information Technology*.

[71] Meta (2024). *AI Sandbox for Meta Advantage suite*.

[72] Nelson CBD (2024). *Nelson AI Sandbox*.

[73] MITRE (2024). *Federal AI Sandbox*.

Answering the question of *when* to do a sandbox is only the starting point. Success depends heavily on the design and execution of the sandbox. The next section dives into *how* to do a sandbox for AI.

Quotes from consulted stakeholders

Sandboxes are not an ordinary affair. They take financial and human resources."

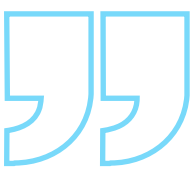
"Regulators and policymakers may want both pre and post, whereas you know industry in particular, small medium enterprises and start-ups may want something prior to regulation. However larger companies may wait until there is certainty in what the regulation is all about. So I think that trying to tease out the nuance in this, is very, very important."

"The value of the sandbox really comes from the part where we will shape the policy together with the industry. And hence I personally feel that it's better, and easier to do it before the regulation".

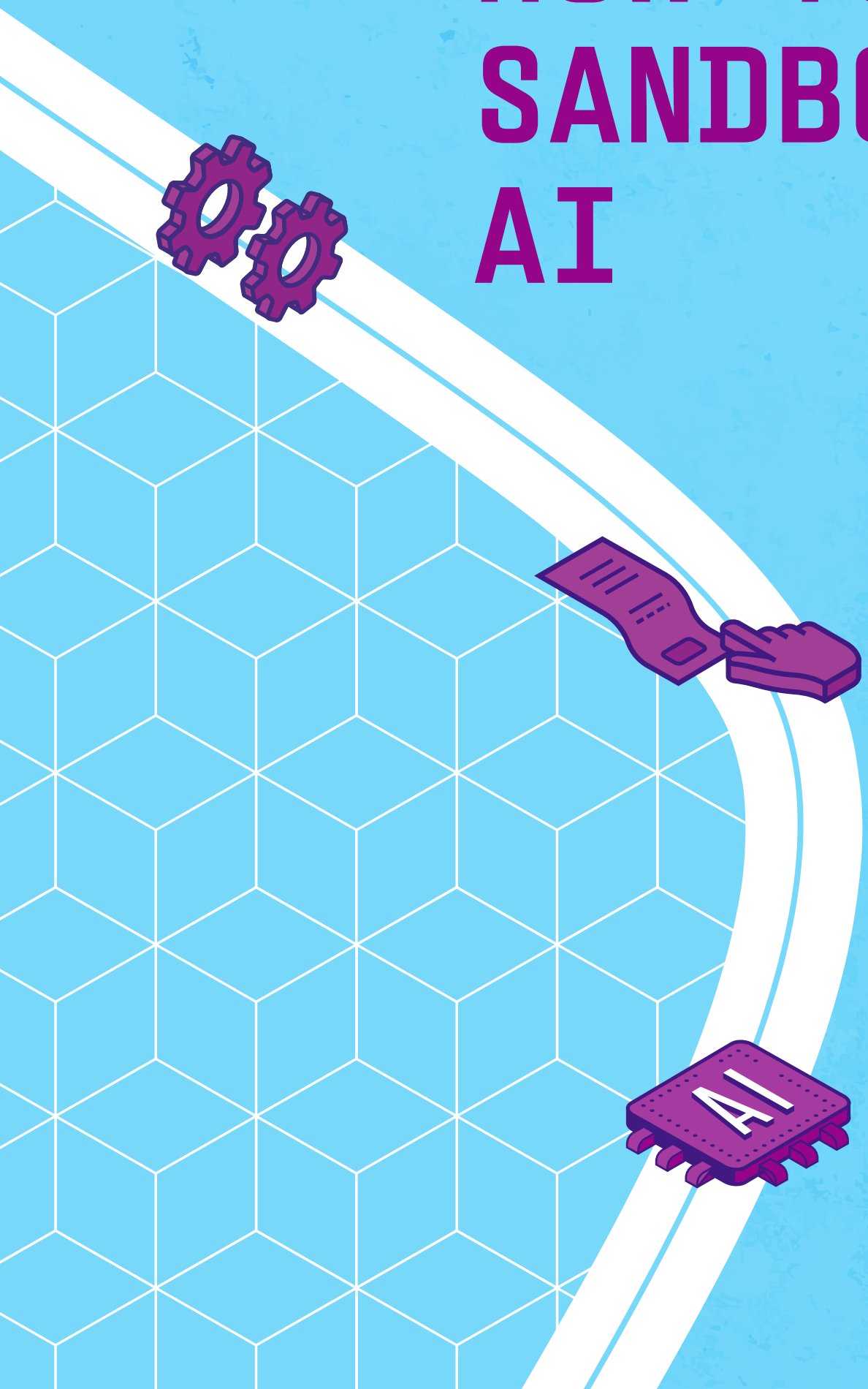
"Sandboxes can be used to review the relevance of existing sectoral regulations to AI applications as well."

"The when actually depends also on the goal, and how you see the sandbox's connection to your regulatory strategy."

"I believe that every single government will need to establish something to facilitate the setting up of sandboxes in a replicable manner, according to their own national arrangement.



HOW TO SANDBOX AI



Sandboxes for AI are a relatively new but rapidly evolving tool for addressing the unique challenges posed by AI as a data-intensive technology. While the concept is gaining traction globally, determining how to structure and implement these sandboxes remains a complex task, as AI regulation must balance innovation with risk mitigation while navigating uncertainties around governance, ethical concerns, and technological advancements. Understanding how to establish an effective AI sandbox is critical, as it shapes the way policymakers, businesses, and researchers experiment with AI solutions in a controlled environment.

A well-designed sandbox can accelerate responsible AI development, promote regulatory clarity, and foster international collaboration. However, the process is not straightforward, requiring each jurisdiction to tailor its approach to local legal, economic, and technological contexts. While regulatory sandboxes for AI are a recent practice, a number of initiatives have already demonstrated their effectiveness in addressing the challenges of AI as a data-intensive technology. This section highlights three examples from Norway, France, and Singapore.

These selected cases provide a very diverse panorama on how to implement sandboxes for AI, but a structure could be extrapolated from these experiences towards a common structure on how to do sandboxes in this space.

This section concludes by offering a blueprint for designing a sandbox for AI.

SANDBOXES IN PRACTICE

The cases of Norway, France, and Singapore illustrate the diverse ways in which AI sandboxes can be designed and implemented, each tailored to distinct regulatory needs and technological priorities. While all three initiatives aim to foster responsible AI development and provide regulatory clarity, they differ in scope, focus, and stakeholder engagement.

Norway's sandbox, embedded within the national data protection authority, takes a broad and iterative approach, expanding over time to address evolving challenges in AI and data governance (Box 1). France's sandbox operates with a more sector-specific focus, selecting a different priority area each year to align with national digital policy goals (Box 2). Meanwhile, Singapore's sandbox, with its emphasis on generative AI and industry-wide collaboration, showcases a forward-looking model that seeks to establish standardized benchmarks for AI evaluation (Box 3). Despite these differences, all three examples underscore the importance of regulatory guidance, multi-stakeholder collaboration, and the development of best practices to navigate AI's complexities.

The following case boxes provide a closer look at each initiative, outlining their objectives, implementation strategies, and key takeaways for designing AI sandboxes in different contexts.

Box 1. Norway's Regulatory Sandbox for Artificial Intelligence

Context

The Norwegian Data Protection Authority, Datatilsynet, established its Regulatory Sandbox in 2020 under the National Strategy for Artificial Intelligence. Initially funded through supplemental appropriations in the National Budget, the sandbox became a permanent fixture with the 2023 National Budget.[74] Its primary goal at first was to stimulate "privacy-friendly" innovation in AI.[75] The rationale for this initial focus on AI stemmed from its vast potential to transform public and commercial sectors and improve the general population's quality of life. The Norwegian government established the sandbox as a proactive measure to address the significant challenges regarding AI's personal data usage, providing a controlled environment to develop compliant and ethical AI solutions.

While data was always a primary focus of the sandbox, its first three iterations were dedicated entirely to developing responsible and ethical AI solutions that complied with data protection regulations. However, recognizing that data protection challenges and the need for privacy-friendly solutions extend across various technological advancements, the government subsequently expanded the sandbox's scope.[76] Now in its fifth iteration, the sandbox supports projects addressing not only AI but also regulatory uncertainties in complex data sharing, the EU GDPR's provisions on automated decision-making, and secondary data uses, among other topics.[77]

Objectives

Norway's Regulatory Privacy Sandbox has three main objectives: stimulating privacy-friendly innovation and digitalization, enhancing the Norwegian Data Protection Authority's expertise in new technologies, and ensuring that technological advancements benefit society while safeguarding individual rights.[78] Datatilsynet aims to build public trust by developing and implementing technological solutions within a governance framework that emphasizes accountability, transparency, explainability, and the protection of fundamental rights. Datatilsynet provides tailored guidance to selected projects, develops best practices based on insights gained, and facilitates collaboration and networking among participants and external experts.

[74] Lystad (2022), *Datatilsynet jubler: sandkassa får permanent støtte [The Norwegian Data Protection Authority rejoices: the sandbox receives permanent support]*, Computerworld.

[75] Datatilsynet (2024), *Doorkeeper, exit report: Intelligent video monitoring with data protection as a primary focus*, Norwegian Data Protection Authority.

[76] Markussen (2023), *Evaluation of the Norwegian Data Protection Authority's Regulatory Sandbox for Artificial Intelligence*, Norwegian Data Protection Authority.

[77] Datatilsynets (2021), *How to apply to join the sandbox?*, Norwegian Data Protection Authority.

[78] Datatilsynets (2021), *How to apply to join the sandbox?*, Norwegian Data Protection Authority.

Stakeholder engagement

The sandbox emphasizes broad stakeholder engagement, involving diverse participants from both private and public sectors. The selection process involves a rigorous evaluation of applications by an internal interdisciplinary group within the Norwegian Data Protection Authority, which interviews all applicants.[79] An external reference group, comprising members from Innovation Norway, the Norwegian Computing Centre, the Equality and Anti-Discrimination Ombud, and Tekna (a union for graduate technical and scientific professionals), assists in assessing the public benefit of potential projects.[80] The final selection is made by a steering committee composed of the Authority's management. Supporting stakeholders play a crucial role, with Datatilsynet coordinating with other sandbox-managing entities within the Norwegian government and being part of a European sandbox network that includes the British ICO and French CNIL. Datatilsynet has also engaged with authorities in several other countries interested in the Norwegian sandbox and has shared experiences with master's students and doctoral fellows who have written assignments about the sandbox.[81]

Datatilsynet actively engages the public and stakeholders through its website, where it publishes project plans, final reports, and other relevant sandbox news. The agency also maintains a sandbox newsletter and a podcast series called SandKasten, in which it shares insights from ongoing sandbox projects. Furthermore, Datatilsynet organizes workshops and seminars to disseminate learnings from the sandbox and has leveraged both national and international media, participating in over 30 conferences and events.[82]

Reporting

Datatilsynet publishes experiences from both ongoing and completed projects, sharing insights that can benefit other organizations and contribute to broader understanding and improvement in privacy practices.[83] An example from their website can illustrate how projects that successfully navigate the sandbox conclude with the development and refinement of a product or service that complies with privacy regulations and incorporates feedback and insights gained during participation. Take Doorkeeper, a Norwegian start-up that developed an intelligent video monitoring system. Through the sandbox, Doorkeeper addressed regulatory challenges, explored alternative designs, and implemented data minimization and robust security measures. This collaboration ensured their solution not only complied with data protection laws but also set a precedent for privacy-friendly innovations in the security industry, exemplifying how sandbox participation can lead to the successful integration of privacy principles in technological advancements.[84]

[79] Ibidem.

[80] Markussen (2023), *Evaluation of the Norwegian Data Protection Authority's Regulatory Sandbox for Artificial Intelligence*, Norwegian Data Protection Authority.

[81] Datatilsynet (2024), *Doorkeeper, exit report: Intelligent video monitoring with data protection as a primary focus*, Norwegian Data Protection Authority.

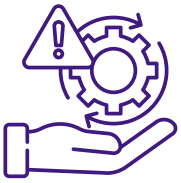
[82] Datatilsynet (2024), *Doorkeeper, exit report: Intelligent video monitoring with data protection as a primary focus*, Norwegian Data Protection Authority.

[83] Datatilsynet (2024), *Reports*, Norwegian Data Protection Authority.

[84] Datatilsynet (2024), *Reports*, Norwegian Data Protection Authority.

Impact assessment

In 2023, Datatilsynet partnered with an external consulting company to assess the sandbox's impact and effectiveness. The evaluation revealed that the sandbox successfully met the needs of participating entities and generated valuable insights into AI, data protection, and ethical technology use. Key recommendations from the evaluation included enhancing the sandbox's technical expertise, improving communication strategies to reach a broader audience, and fostering systematic collaboration with other public and private actors. While the sandbox was generally well-received, the evaluation highlighted areas for improvement, such as providing more detailed final project reports, streamlining internal processes for faster project approval, and recruiting a more diverse range of projects.[85]



Risk management

Notably, this sandbox ensured protections for previously-existing intellectual property (IP), allowing participants to retain ownership of any IP they bring into the sandbox collaboration. As a risk-assessment practice, this is an interesting example of an active responsible practice.

Box 2. France's Personal Data Sandbox and its AI iteration

Context

The French Data Protection Authority (CNIL) launched its personal data sandbox in 2021 to provide focused support to innovative projects that prioritize data privacy from their inception. CNIL offers direct engagement through its legal and technical teams to clarify regulatory requirements, provide practical advice, and audit developed solutions to ensure GDPR compliance.[86]

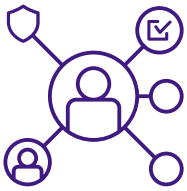
Objectives ans scope

Each year, CNIL invites applications from organizations developing products and services within a selected sector, contrasting with Norway's broader approach that accepts any project dealing with data or AI as outlined in its guidelines. The French sandbox's first iteration focused on innovations in the field of digital health, and its second on educational digital tools (EdTech). The latest edition targets the integration of AI in public services. As with Norway's sandbox, this choice was driven by a recognition of AI's significant potential to transform public sector operations. According to the CNIL, AI can significantly improve service delivery by improving efficiency, reducing administrative burdens, and enhancing user accessibility. It can assist public officials in optimizing time management, prioritizing tasks, and maximizing the utility of data. The sandbox addresses the challenges related to AI systems relying on substantial volumes of personal data, ensuring innovations comply with strict data protection standards while improving public service delivery.[87]

[85] Markussen (2023), *Evaluation of the Norwegian Data Protection Authority's Regulatory Sandbox for Artificial Intelligence*, Norwegian Data Protection Authority.

[86] CNIL (2021), *Bac à sable » données personnelles de la CNIL : appel à projets 2021 [CNIL personal data "sandbox": 2021 call for projects]*, CNIL.

[87] CNIL (2023), *Digital health and EdTech: the CNIL publishes the results of its first "sandboxes"*, CNIL.



Outcomes and engagement

Throughout their participation, sandbox projects receive ongoing feedback and support from CNIL, with findings and recommendations published on its website afterwards to foster broader compliance and innovation. Previous recommendations from the digital health and EdTech sandboxes emphasized data minimization, clear data retention policies, robust anonymization to protect privacy, and transparency in data processing.[88] However, the analysis of the sandbox program and communication of its broader impacts remain somewhat constrained, with limited engagement beyond the participating organizations. This may restrict the application of insights gained to broader sectoral technology development or regulatory adjustments.

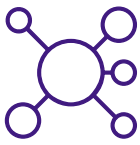
Box 3. Singapore's Generative AI Evaluation Sandbox for Trusted AI

Context

Singapore's Generative AI Evaluation Sandbox, spearheaded by the Infocomm Media Development Authority (IMDA) and established in 2023 in partnership with the AI Verify Foundation, illustrates a unique approach to sandbox implementation that diverges from traditional frameworks. This initiative brings together major multinational companies to evaluate trusted AI products using a newly crafted Evaluation Catalogue developed by the IMDA. The Catalogue delineates standardized methods and benchmarks specifically for Large Language Models (LLMs), setting foundational criteria for GenAI evaluation.[89]

Objectives

The collaboration between IMDA and the AI Verify Foundation seeks to forge a common standard for GenAI evaluations that not only mitigates risks but also fosters safe adoption, thereby enhancing assessment capabilities across the AI ecosystem. The sandbox is tailored to pinpoint and address the particular gaps in GenAI assessments, and develop benchmarks for model performance, focusing especially on domains relevant to the unique needs of Singapore.



Engagement

The sandbox involves a diverse array of participants, including model developers, application developers, and third-party testers, as well as the Singapore Personal Data Protection Commission.

Reporting

Given how recent the GenAI Evaluation Sandbox is, detailed public information about the sandbox remains scarce.

As the examples illustrate, sandboxes can be designed and implemented in very different ways for unique purposes and environments. While the purpose and scope of an AI sandbox can vary greatly depending on the legal context, culture, and regulator, there are specific phases in developing a sandbox to consider.

[88] CNIL (2023), "*Sandbox": CNIL launches call for projects on artificial intelligence in public services*, CNIL.

[89] Infocomm Media Development Authority (2023), *First of its kind Generative AI Evaluation Sandbox for Trusted AI by AI Verify Foundation and IMDA*, Infocomm Media Development Authority.

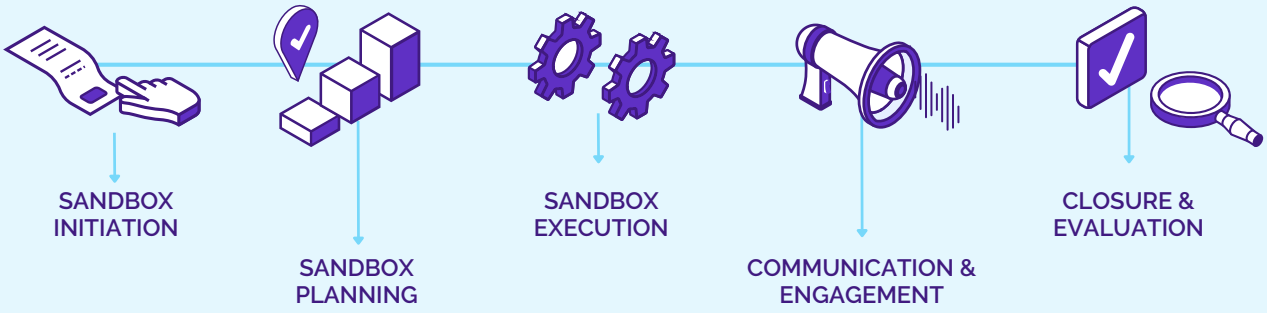
TOWARDS A BLUEPRINT FOR SANDBOXES

A well-structured AI sandbox typically follows a systematic process that includes several key phases: **initiation, planning, execution, closure, and evaluation**. These phases ensure that the sandbox is effectively designed, implemented, and assessed. Additionally, **communication and engagement** is an ongoing aspect that permeates all stages, ensuring that stakeholders are informed, involved, and aligned throughout the development of the sandbox - here included as a fourth step of a five-step process.

Below is the **five-phase approach** of sandbox development (Figure 6):

1. **Initiation** – Defining objectives, scope, stakeholders, and regulatory context.
2. **Planning** – Establishing governance frameworks, risk mitigation strategies, and participant selection criteria.
3. **Execution** – Launching the sandbox, monitoring AI testing, and facilitating real-time regulatory feedback.
4. **Communication and Engagement** – Ensuring transparency, sharing lessons learned, and engaging external stakeholders.
5. **Closure and Evaluation** – Assessing outcomes, documenting insights, and refining future sandbox iterations.

Figure 6. Sandbox development phases



1 INITIATION

The **initiation** phase marks the beginning of the sandbox process, where the objectives and scope are defined. This phase involves identifying the key stakeholders, understanding the specific regulatory or technical challenges to be addressed, and establishing the sandbox's goals. A crucial activity during this phase is to **evaluate the organization's preparedness to implement a sandbox**, which includes identifying the strengths and areas for development. This evaluation helps determine whether the organization has the necessary financial, human, and technical resources to establish and operate the sandbox effectively. It also involves assessing if existing resources or initiatives can be leveraged or if additional funding and support are necessary to ensure the successful implementation of the sandbox. To ensure alignment with existing regulatory frameworks, both an **external policy framework assessment**—evaluating alignment with regulatory authority policies, innovation policies, and legal frameworks—and an **internal organizational assessment**—establishing criteria to assess the sandbox's maturity—must be conducted. Additionally, implementing a **sandbox maturity assessment** would provide a benchmark for readiness against established models.

Clarifying regulatory grey areas is already complex and challenging, let alone in the light of more innovation and uncertainty. It is essential to choose an issue area of importance to all parties. One way to identify key issues and encourage buy-in early on is to actively survey innovators and other stakeholders such as civil society, and apply their insights. As part of this phase, a **stakeholder and partner map** needs to be developed to clarify roles and potential collaborations. Direct engagement with key stakeholders before kickstarting the sandbox will provide valuable feedback to shape the sandbox's future direction. Identifying issues of equal importance, uncertainty and public interest will help to prioritise the sandbox activities and resource allocation, while maintaining focus on the problems it is aiming to solve.

The choice of AI issue will attract support and resources, maintain focus on outputs, and motivate progress towards defined outcomes. A critical ingredient is to carefully identify and delineate the **core regulatory or operational question(s)** that the sandbox needs to answer for each innovation. After the issue area has been identified and innovators involved, the first step of the sandbox will be to tightly define the critical question(s). This preparatory work will provide the clarity and focus needed for success.

A structured approach needs to be taken to refine the sandbox concept, which could include training to build foundational understanding, as well as workshops to consolidate outcomes from the initiation phase. Moreover, to ensure sustainability, a detailed **cost analysis** needs to be carried out to ensure that each phase of the sandbox is financially feasible and well-resourced.

2 PLANNING

During the **planning** phase, a detailed strategy for the sandbox is developed. This includes defining the regulatory framework, technical infrastructure, and support mechanisms needed for the participants. A key component of this phase is the creation of a **Sandbox Governance Framework**, which outlines the rules, responsibilities, and oversight mechanisms necessary to maintain order and ensure the ethical operation of the sandbox. This framework helps regulate how participants interact with the sandbox, ensuring that all parties adhere to established standards. It also defines processes for mediating conflicts between the competent lead(s) (particularly if different organizations are co-leading the sandbox) or between leads and participants and observers.

Additionally, a **Controls and Data Flows Framework** should be put in place to address critical concerns such as confidentiality, data protection, and intellectual property protection. A detailed plan for securing data, protecting sensitive information, and ensuring that intellectual property rights are respected is also developed to mitigate risks and protect all stakeholders. To complement these safeguards, an **Information Collection Framework** is established to develop a robust system for collecting and analyzing data throughout the sandbox's lifecycle. This includes defining key metrics, establishing data collection methods, and outlining procedures for interpreting findings to inform future policy decisions.

To ensure the effective management and execution of the sandbox, a **Sandbox Project Management Framework** is defined, detailing processes, deadlines, roles, and responsibilities across the sandbox lifecycle. To ensure partners work well together towards a shared strategic goal, it is essential to articulate and agree on the key goals. Setting a clear overarching and underlying set of goals will also surface key points of difference in a way that allows them to be resolved, at least at the level of principle, before work begins.

Since participation in the sandbox requires a structured and transparent application process, templates and forms for applying to the sandbox are developed. A draft **Sandbox Call** is also prepared, along with materials explaining the call content, how to apply, and frequently asked questions to guide potential participants.

Not every participant in a sandbox will already have run or participated in a sandbox. In order to bridge the potential differences between different ways of operating or participating in a sandbox, a dialogue between stakeholders should be structured.

3 EXECUTION

The **execution** phase is when the sandbox is put into action. Key activities during this phase include the **proposal selection process**, where submitted applications are reviewed, and a **draft of the call results** is prepared to ensure transparency in participant selection. Once participants are confirmed, an onboarding process is conducted, providing sandbox guidelines that clearly outline expectations, rights, duties, schedules, and communication channels to ensure all participants understand the framework in which they will operate.

Launching the sandbox

The launch moment involves key steps for rolling out the sandbox, including undertaking the selection of participants, setting up the necessary technical and regulatory infrastructure, and managing resources to ensure smooth operations. This step also requires defining clear milestones, coordinating logistics, and ensuring that participants have access to the necessary tools, data, and regulatory support. One key component when launching the sandbox is ensuring that the public in general knows the status of the sandbox and is informed about the envisioned objectives and expected outcomes of the intervention.

Testing and monitoring

Testing and monitoring are crucial components of the execution phase, where participants experiment with their innovations in a controlled environment while adhering to the sandbox's guidelines. Regulators continuously monitor participants' activities, collect relevant data, and offer ongoing guidance to ensure the testing process aligns with the sandbox's objectives. Throughout this period, operational management plays a critical role in addressing challenges, ensuring compliance, and troubleshooting any issues that arise. A comprehensive overview of each project's progress is developed over a determined period (preferably a three-month period) to assess advancements, identify challenges, and provide necessary support.

Workshops play a key role in this phase, facilitating knowledge exchange and engagement between participants, regulators, and other stakeholders. A series of workshops need to be conducted to provide structured moments for participants to share insights, refine their approaches, and learn from one another. Additionally, these workshops help refine the sandbox's technical and governance approaches by incorporating real-time feedback and addressing unforeseen challenges.

Iterative adaptation and evaluation

The execution phase also includes **iterative adaptation**, where feedback and insights gathered during testing lead to adjustments and refinements. A structured report is compiled based on participant feedback, capturing lessons learned and potential improvements. To enhance regulatory alignment, a dedicated workshop with regulators is conducted to refine and adapt the sandbox's framework based on real-world findings, ensuring that policies evolve in tandem with innovation.

Further adjustments may be necessary based on ongoing monitoring, ensuring that both technical and regulatory aspects remain responsive to emerging needs. This continuous refinement process allows for the early identification of policy gaps and the development of responsible sandbox practices.

This iterative process allows technologies to be improved continuously, ensuring they are aligned with both regulatory standards and real-world conditions. The flexible, responsive nature of this phase helps businesses innovate while maintaining compliance with legal and ethical frameworks. By integrating structured monitoring, participant engagement, and regulatory dialogue, this phase ensures that the sandbox remains a dynamic and effective tool for responsible innovation. A final review at the end of this phase will consolidate findings, document key takeaways, and prepare the groundwork for scaling successful innovations beyond the sandbox environment.

4 COMMUNICATION AND ENGAGEMENT

Throughout all the stages, **communication and engagement** are crucial. Regular updates, feedback sessions, and open dialogue between regulators, innovators, and other stakeholders ensure alignment and trust. This ongoing interaction helps to identify emerging issues, address concerns promptly, and ensure that the sandbox remains relevant and effective throughout its lifecycle. Monthly updates on project progress and next steps should be provided regularly to keep stakeholders informed and aligned with the sandbox's goals. Effective communication and engagement ensure the sandbox serves as a dynamic tool for innovation and regulatory development.

To ensure consistent and coordinated communication, a detailed **communication and engagement strategy plan** should be developed at the outset of the sandbox process. This will guide messaging, communication channels, and stakeholder engagement activities throughout the lifecycle of the sandbox. Defining the sandbox's core, active and supporting stakeholders is key to an effective communication and engagement strategy. For regulatory sandboxes, *core stakeholders* are the most senior representatives from the involved regulators and may also include other regulatory authorities, ministries and judicial authorities as appropriate, depending on the issue areas. For operational sandboxes, *core stakeholders* may also include key entities responsible for coordinating and contributing with data and datasets. These are 'core' stakeholders because their active involvement and support is essential to the establishment, funding and operation of the sandbox.

Active stakeholders are the innovators, i.e. the firms and organisations responsible for an innovative technology, process or model. The industries or sectors they emerge from may be part of pre-sandbox consulting and engagement carried out in Step 1 above. While the innovators will be participating in the sandbox, rather than managing its operations, their input on issue identification and definition, goals and timelines is essential.

Supporting stakeholders will be all other stakeholders relevant for consultations regarding positive and negative externalities of sandboxes. If the sandbox issues deal with, for example, personal data, or particularly sensitive categories such as health data, the stakeholders need to include people who can speak for the interests of data subjects. This may include civil society organizations and academics or other experts, or even information campaigners to ensure the public understands what is being done. While these stakeholders will not make regulatory or operational decisions, they should be part of how the sandbox is set up, and consulted appropriately as it develops. **Mechanisms for open and regular dialogue** are essential to earn and build trust, and to maintain strong working relationships between all stakeholders.

Defining the roles of each stakeholder will be important to keep the necessary people involved while also ensuring they take an active role in the sandbox as needed. Political buy-in of core stakeholders is key, especially from the relevant agencies, departments and regulators. There may often be a 'watching brief' involvement by various entities within a given government. While transparency and capacity-building are vital, and will help to implement the outcomes of the sandbox and funnel its insights into policymaking, decision-making within reasonable timelines relies on each core stakeholder having sufficient authority and being actively engaged.

There is a trade-off between having sufficient involvement amongst other regulators, agencies and ministries to ensure buy-in and knowledge dissemination, and a dedicated decision-making core to ensure efficiency. Agreement and alignment among state actors in a clear direction is crucial. The core stakeholders need to coordinate and assign responsibility and decision-making for design, implementation, supervision and steering of the sandbox, and also to agree on its day to day working practices.

To enhance public awareness it is crucial to engage in external communication to ensure the broader public is kept informed about the sandbox's progress and key milestones.

As part of **engagement and transparency efforts**, it is also crucial to establish mechanisms for **receiving and incorporating stakeholder feedback**. This may include the creation of an Ombudsman role or other structured feedback channels to address concerns and improve sandbox operations. Additionally, events such as workshops and roundtables should be organized to facilitate stakeholder engagement, share progress updates, and discuss sandbox insights. Presentation materials for regulators should be developed to showcase sandbox results, ensuring that findings are communicated effectively to diverse audiences, including policymakers, industry leaders, and the public.

5 CLOSURE AND EVALUATION

Once the testing period concludes, the **closure** phase begins. This phase involves finalizing the results of the sandbox, compiling insights, and determining whether the tested technologies are ready for broader deployment. Exit strategies will be developed for participants, ensuring that they can transition smoothly out of the sandbox and into the broader market or regulatory environment. The closure phase also includes comprehensive documentation that captures key aspects of the sandbox process, including methodologies, findings, challenges, and good practices. The sandbox's outcomes will be analyzed through a formal review, and a *post-mortem* analysis report will be created to reflect on successes, challenges, lessons learned and areas for improvement. Impact assessments will be carried out to evaluate the broader effects of the sandbox, considering its influence on regulatory landscapes, innovation, and societal benefits. Recommendations for future improvements or additional sandbox initiatives may also arise during this phase.

Finally, **evaluation** is an essential phase that takes place after closure. This step involves assessing the overall effectiveness of the sandbox, reviewing its impact on the regulatory landscape, and determining its value to the participants and society at large. This phase also includes a workshop to collect feedback from all stakeholders and participants, sharing insights that will improve future sandbox projects. Recommendations for future improvements or the need for additional sandbox initiatives may also arise during this phase. The **final public report** will be shared to communicate key findings, lessons learned, and recommendations to a wider audience. A **digital archive** of all relevant materials will also be created, including methodologies, findings, challenges, and best practices, ensuring that future sandboxes can learn from these experiences.

As part of incorporating learnings into future projects, an **action plan** will be developed to help participants adopt the advice and insights gained from the sandbox experience into their internal procedures and strategies. Additionally, a **final report with internal recommendations** will be created to provide actionable insights for improving future sandboxes. A prototype or suggestions for future regulation drafts may also be developed based on the findings of the sandbox.

Indeed, any sandbox needs to be evaluated and improved iteratively. Therefore, it will be essential to measure – or even just assess qualitatively – the success or failure of each sandbox. This is essential both for public accountability, and also to find out what works and what doesn't, to inform future sandboxes. As so few regulatory sandboxes for AI have existed to date, and have dealt intensively with a relatively small number of innovations, benchmarking through quantitative measures to evaluate success would be limited. While elaboration of quantitative indicators is encouraged, qualitative indicators are more likely to be useful at this earlier stage (Box 4).

Box 4. Key qualitative indicators to be considered when evaluating sandboxes

- Were the regulatory decisions implementable? What, if any, impacts have they had on the success or otherwise of the innovator? On competition and choice in home markets?
- What use has been made of the sandboxes' findings or decisions? Can they be scaled up or applied elsewhere? If it is legally possible, do they form precedents, either formally or informally?
- Has the sandbox produced learnings that can be implemented in regulation and laws, or international agreements, treaties, codes of conduct, etc? Or to guide the development of similar innovations?
- For operational sandboxes, what new uses, business and organization practices and models have they enabled? What research has been made possible, and has it been beneficial?
- Which public services or policies have been informed or improved by the outcomes of the operational sandbox?
- What additional or different metrics/reporting requirements would be useful for future iterations of this same sandbox?[90]

[90] The Global Sandboxes Forum provides services and expertise to support countries in the methodologies needed to develop a sandbox in the context of AI. For more details on the methodology and steps involved, see the services here: *https://sandboxes.thedatasphere.org/*

Quotes from consulted stakeholders

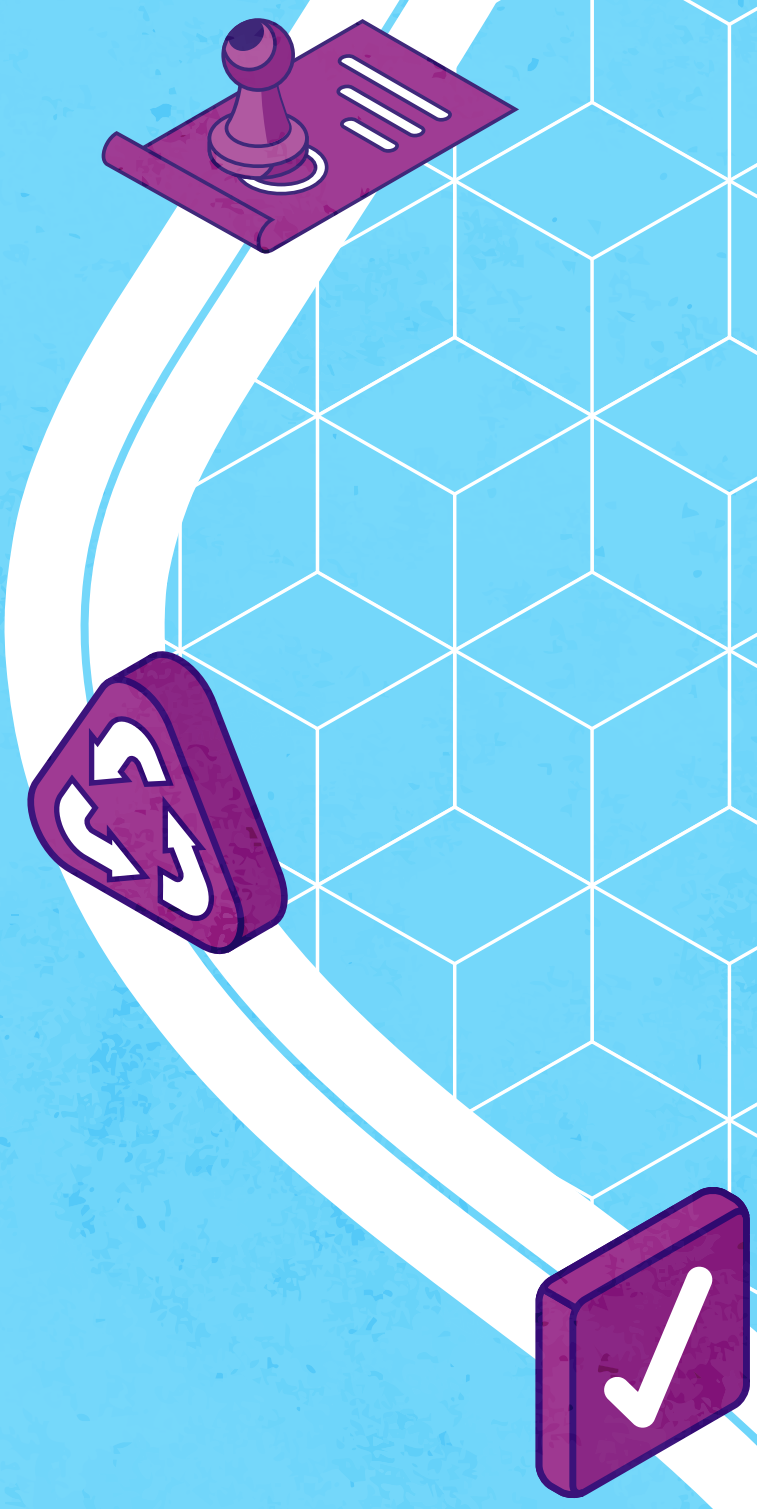
Sandboxes cannot exist in a vacuum. They need to exist in an ecosystem of efforts that need to be connected. I think that's why needs, resources and the different incentives and initiatives need to be mapped out".

"We have to really look at the challenges and the objectives. And what is the context like? To create the building blocks of that sandbox. So it's not 100% replicable. But I think in more general terms, there could be some science."

"Who is going to measure the appetite for risks, for example, what kind of risks are we willing to take with that exercise in particular?"

"We've been talking with the various members who might be a primary user of our sandbox, and what seems to happen is a sort of accelerator approach."

"We are basically sandboxing sandboxes. We are trying to see if it's possible to develop a gold standard and best practice. And I do believe in some years we will come to a conclusion that the answer is yes, but flexibility will always be part of it, even in the methodology itself.



CONCLUSION

Sandboxes hold significant potential in regards to AI governance. Nevertheless, their development and implementation present significant challenges that must be addressed to fully leverage their benefits. The inherently global nature of AI, characterized by technology inputs and outputs frequently crossing national borders, complicates regulatory oversight by any single national sandbox operator. Additionally, AI cannot be regulated from a single perspective, given the multiplicity of involved regulators, its broad pool of training data and interconnected applications.

Sandboxes for AI are more often being developed within AI communities, without necessarily leveraging learnings from sandboxes in other sectors. While efforts exist to facilitate dialogue among countries on establishing their own sandboxes, there remains a notable lack of cross-country collaboration within the sandbox initiatives themselves. For instance, the EU AI Act promotes the use of sandboxes as key to its compliance mechanism, yet their implementation would benefit from cross-country coordination and learning from other sandbox experiences beyond the EU and across sectors like health, finance, and ICT. Research from the Datasphere Initiative has pointed to the lack of information about sandbox experiences worldwide as well as the lack of transparency regarding sandboxes' results, impact, challenges, and best practices. This underscores the urgent need to promote cross-regional, cross-sectoral and cross-jurisdictional knowledge sharing to ensure that lessons learned truly enrich sandbox evolution.

A hermetic approach can lead to inefficiencies and unnecessary duplication, compounding the already substantial resources and time required to set up and run a sandbox. For example, while a sandbox might be effective within its own country, the lack of knowledge sharing and collaboration with other countries means that best practices and successful strategies are not disseminated, limiting the overall effectiveness and innovation potential globally. Additionally, even within a single jurisdiction, the effectiveness of sandboxes can be undermined by a lack of transparency, as citizens may not have visibility into how resources are being allocated or the impact these initiatives are generating. Building public trust and ensuring that sandboxes deliver clear public value are critical to their success in a dynamic regulatory environment. To maximize their effectiveness, sandboxes must foster both public trust through transparency and global innovation through collaborative information sharing.

As noted above, sandboxes often require significant investment in terms of time and expertise to manage complex technologies and ensure compliance under regulatory regimes which often lack clarity in their application (hence the technology's inclusion in the sandbox in the first place). They are usually designed to address specific, narrowly defined problems, limiting their ability to handle cross-sectoral issues or adapt to rapidly evolving AI technologies that span multiple regulatory domains. The experimental nature of sandboxes also means that they are typically not suited for large-scale deployment without significant modifications. These challenges underscore the need for careful planning, robust design, and clear objectives when designing and implementing sandboxes.

Recommendations

To fully leverage the potential of sandboxes for AI policymakers, governments, the private sector, international organizations and other stakeholders should focus on the following areas.

1. **Ensure that AI sandboxes are designed with clear objectives and measurable outcomes** – Each AI sandbox should have well-defined goals, evaluation metrics, and a structured process for assessing effectiveness. That is why it is essential to invest significant effort and resources into the initiation and planning phases of sandboxes. This will allow for continuous improvement and inform future regulatory developments.

2. **Develop structured methodologies for sandbox implementation** – While AI sandboxes are still evolving, common frameworks and methodologies should be established to guide their design, execution, and evaluation. Developing a "science" of sandboxes will help ensure consistency, scalability, and replicability across different jurisdictions.

3. **Enhance transparency and public accountability** – Sandbox operators should ensure clear public reporting on objectives, processes, and outcomes. Open access to findings will build trust, facilitate learning, and increase the adoption of best practices. Cooperation should be prioritized over control and transparency a guiding principle.

4. **Expand civil society engagement in AI sandboxes** – AI governance should not be left solely to governments and industry. Civil society organizations, academia, and consumer groups should be actively included in sandbox discussions to provide oversight, identify ethical concerns, and advocate for public interest protections.

5. **Foster interoperability between AI and data governance sandboxes** – AI development is inherently data-driven. Policymakers should better integrate AI sandboxes with data governance sandboxes to ensure responsible data use while enabling innovation.

6. **Foster capacity-building related to AI, data and sandboxes** – Training programs and educational initiatives should be developed to build capacity among regulators, industry participants, and civil society to equip them with the tools and knowledge to run and participate in sandboxes. These programs should focus on the unique challenges posed by AI technologies, including ethical considerations, data governance, and compliance with evolving regulations.

7. **Support regulatory agility and adaptability through iterative learning** – Sandboxes should not be one-time experiments but ongoing processes for regulatory learning and adaptation. Sandboxes require a change in mindset. Regulators should continuously refine their policies based on sandbox findings and engage in an iterative process to keep pace with AI advancements.

8. **Leverage sandboxes for AI to test AI risk mitigation measures** – Regulatory sandboxes should incorporate mechanisms to test AI explainability, bias mitigation, privacy-enhancing technologies, and other safeguards before AI applications are widely deployed.

9 **Promote sectoral and cross-sectoral AI sandbox initiatives** – Given AI's widespread impact, more sandboxes should address AI use cases across different sectors such as health, finance, and smart cities. Additionally, multi-sector sandboxes should be encouraged to tackle cross-sectoral and transversal policies related to AI governance challenges that span multiple regulatory domains, ensuring collaboration among different regulatory agencies in the same jurisdiction.

10 **Strengthen cross-border collaboration on sandboxes for AI** – Regulators and policymakers should coordinate efforts to foster interoperability of regulations and promote international cooperation across borders, as well as to reduce forum-shopping and enable businesses to draw more value out of experimentation spaces, ensuring that AI innovations can thrive in a consistent and supportive global environment.

This report provides a starting point to further exploration and research on sandboxes and their potential to support AI governance and drive responsible innovation. More research into sandbox responsible practices and sharing of lessons, and stories behind sandbox development and implementation is necessary to uncover the real opportunities and risks of leveraging new regulatory tools such as sandboxes to handle the speed and scale of AI advancement and deployment. Additionally, there is a need for communities from all sectors –working in sandboxes or not– to come together and share knowledge, experiences and learnings on safe spaces for collaboration and enhanced cross-pollination. The Datasphere Initiative plans to work with partners to build a sandbox science and community of practice for AI and other emerging technological and policy disruptions, guided by goals of responsibly unlocking the value of data for people and the planet.


thedatasphere.org