DATASPHERE INITIATIVE

Advancing global AI governance: Exploring adaptive frameworks and the role of sandboxes

AI Action Summit | Paris, France
February 11, 2025

INTRODUCTION

The official side event titled "**Advancing Global AI Governance**: Exploring Adaptive Frameworks and the Role of Sandboxes" took place during France's AI Action Summit. It was organized by the Datasphere Initiative and the ICC, involving roundtable and panel discussions with numerous international experts.

Andrew Wilson, Deputy Secretary General - Policy for the ICC, welcomed participants and highlighted the urgency of exploring new approaches to AI governance in a time of geopolitical tension. He referred to the International Chamber of Commerce's four-pillar model for AI governance that includes principles and codes of conduct, regulation, technical standards, and industry self-regulation. He encouraged attendees to reflect on how best to combine risk mitigation with innovation incentives.

Lorrayne Porciuncula, Executive Director of the Datasphere Initiative, compared sandbox experimentation to learning to swim, emphasizing that practical, hands-on methods are necessary to fully understand and govern AI. She thanked the UK Government and the Hewlett Foundation for their support of the Global Sandboxes Forum, launched previously as an official side event of the G20, and stressed the importance of strengthening a "science of sandboxes" and a global community of practice on adaptive governance.

Session One: The Need for Adaptive Frameworks for AI Governance

Moderated by **Sophie Tomlinson**, Director of Programs at the Datasphere Initiative, this session brought together five lead discussants: **Amanda Craig**, Senior Director of Responsible AI Public Policy from Microsoft, **Lucia Russo**, Economist and Policy Analyst from the OECD, **Yoichi Iida,** Assistant Vice Minister for International Affairs from Japan's Ministry of Internal Affairs and Communications, **Rachel Adams**, CEO and Founder from the Global Center on AI Governance, **David Leslie,** Director of Ethics and Responsible Innovation Research from the Alan Turing Institute, **Benjamin Chua,** Senior Manager from Singapore AISI, and **David Buckley**, Policy Consultant from OpenMined.

Amanda Craig pointed out the vast proliferation of regulatory frameworks around AI. She underlined the challenge of ensuring consistent governance practices across multiple sectors and legal jurisdictions and encouraged participants to focus on identifying common governance practices that can operate effectively under diverse regulatory conditions.

Lucia Russo highlighted the OECD's leadership in promoting interoperability and introduced the Hiroshima AI Process and its associated Code of Conduct, which emphasize voluntary commitments and transparency.

Yoichi Iida explained how Japan led discussions under the G7 Presidency in Hiroshima [Summit], focusing on a supply-chain approach that assigns roles to every stakeholder along the AI life cycle, from developers to end-users, and [encourages] companies to publicly disclose risk mitigation efforts [for effective co-governance].

Rachel Adams underscored the viewpoint of African countries that sometimes approach AI regulation from a sovereignty angle, seeking tangible benefits from AI while dealing with power imbalances and market concentration by global actors. She also cautioned that not all states can meet strict global standards, potentially leaving them excluded from vital data flows.

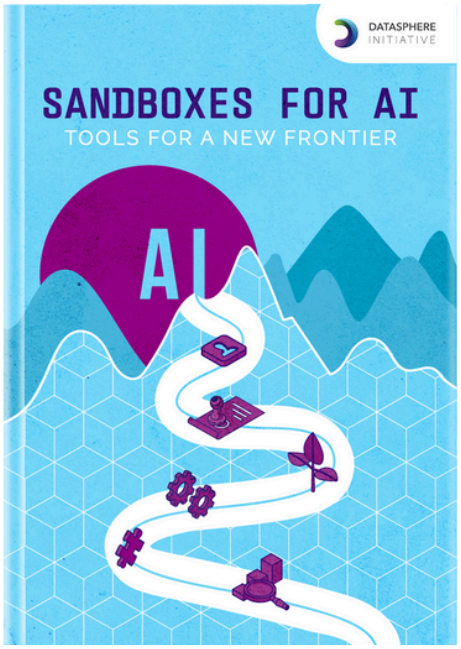
David Leslie further elaborated on the risks of standardization that does not account for cultural and power asymmetries. He argued that meaningful interoperability must consider global equity and that AI governance often overlooks real-world infrastructural disparities. In addition to these core speakers, other participants intervened with examples from India, Kenya, and the Council of Europe, raising concerns about the challenges of combining national needs with cross-border AI solutions.

Benjamin Chua highlighted how Singapore has developed open-source toolkits for both traditional AI and generative AI, ensuring interoperability with major existing principles. He described Singapore's work in launching one of the first third-party assurance ecosystems for AI, bringing together independent auditors to evaluate real-world AI applications.

David Buckley explained the need for secure technical infrastructures and "privacy-enhancing technologies" that allow regulated, privacy-respecting audits of AI systems without exposing developers' sensitive data. He added that, for OpenMined, a global public good approach and open-source philosophy are instrumental in fostering trustworthy AI.

Launch of the Datasphere Initiative Report on AI Sandboxes

After this first session, **Lorrayne Porciuncula** presented the Datasphere Initiative's new report devoted to AI sandboxes. She explained that the researchers had mapped 66 sandbox initiatives operating at global, regional, or national levels, grouped as regulatory sandboxes, operational sandboxes, or hybrid models. She noted that cross-border collaboration is still at an early stage, with many sandboxes focusing on national or subnational experimentation.



The report highlights why some governments and private actors decide to pursue sandboxes before or during regulation, while others opt to set them up after new rules come into effect as a means to verify compliance or gather feedback. The presentation stressed the need for transparency, civil-society engagement, and capacity-building so that public authorities, businesses, and other actors can realize the full potential of sandbox experimentation. It also included a roadmap for designing and executing a sandbox, from laying initial groundwork to setting the goals, running the tests, publishing outcomes, and concluding with evaluation.

Finally, Lorrayne announced that the Datasphere Initiative would continue hosting thematic webinars, publish a new regional Africa-focused report as part of the Africa Sandboxes Forum, and organize a Sandbox Summer School to further global dialogue and capacity.

Session Two: Policy Tools Fit for AI: Why and When to Sandbox

Moderated by **Lorrayne Porciuncula**, the second session focused on when sandboxes should be implemented and under what conditions they are most effective. Speakers included **Félicien Vallet,** Head of the AI Department from CNIL in France, **Alex Moltzau**, Policy Officer from the European AI Office at the European Commission, **Cyril Seck**, Digital Strategy Advisor from Africa Centres for Disease Control and Prevention, **Sam Jungyun Choi,** Lawyer from Covington & Burling LLP, **Matt Prewitt** from RadicalxChange, and **Anne Josephine Flanagan**, Sandbox expert from GSF.

Félicien Vallet explained that CNIL's sandboxes do not provide special exemptions from GDPR obligations but instead offer a structured environment for sector-specific initiatives, such as in healthcare, to ensure privacy by design. These sandboxes allow CNIL to deepen its guidance and produce insights that can be generalized to a broader ecosystem.

Alex Moltzau noted that the forthcoming EU AI Act mandates each Member State to establish a regulatory sandbox by August 2026. These sandboxes, which must be free for startups, aim to help innovative actors navigate compliance in a rapidly changing AI space. He recognized the practical challenges of budgeting, staffing, and aligning with parallel EU measures.

Cyril Seck discussed a prospective regional sandbox at Africa CDC focusing on health-related AI applications. He cited low digital infrastructure in parts of Africa, limited connectivity, and a small specialized workforce but said that well-designed sandboxes could help localize AI solutions and build trust among diverse stakeholders across more than 50 countries.

Sam Jungyun Choi described how clients in various industries, ranging from large technology developers to smaller enterprises, often hesitate to join sandboxes for fear of revealing IP or facing inconsistent regulation across different fields. She noted that stronger guarantees on confidentiality, combined with a multi-regulator approach, could incentivize participation.

Matt Prewitt drew attention to how major technological innovations can reshape the balance of power. He proposed that sandboxes might serve as venues for testing more equitable governance strategies or data-sharing models that offset existing power asymmetries.

Anne Josephine Flanagan reminded the audience that regulation is only one tool for steering AI. She argued that codes of conduct, industry standards, and dedicated spaces for experimentation can foster innovation while reducing risk. However, she also noted the need to avoid over-reliance on any single approach.

Session Three: Opening the Black Box: How to Sandbox for AI

Moderated by **Bertrand de La Chapelle**, Chief Vision Officer at the Datasphere Initiative, this session delved into the operational dimensions of setting up and running sandboxes. The panelists were **HaeOk Choi,** Research Fellow, Directorate for System Innovation from the Science and Technology Policy Institute of Korea, **Sarah Heck,** Head of Policy Planning and Programs from Anthropic, **Caroline Louveaux**, Chief Privacy Officer from Mastercard, **Thiago Moraes**, PhD Scholar from Vrije Universiteit Brussels, **Keith Sabilika,** Senior Fintech Specialist, Financial Technology from the Financial Sector Conduct Authority of South Africa, **Raphael von Thiessen**, Programmleiter KI-Sandbox from the Office of Economy in the Canton of Zurich, and **Romanas Zontovičius**, ICT Industry Manager at Innovation from Innovation Agency Lithuania. Several participants emphasized the need to set clear and achievable objectives before launching a sandbox.

Caroline Louveaux recounted Mastercard's experience with a privacy-enhancing technology sandbox in Singapore, noting that pre-defined outcomes, timelines, and data-protection safeguards were key to success.

Keith Sabilika described how multiple financial regulators in South Africa collaborate to guide fintech innovators in a more coherent manner and prevent overlapping compliance burdens.

HaeOk Choi showcased Korea's highly active regulatory sandbox program, which includes approximately 1403 approved projects, with 46 involving AI. She mentioned that KCCI support centers, one-stop-application processes, rapid approval process, enhanced pre-consultation and advisory support, and fast-track approvals can accelerate innovation, but also that ensuring interoperability across ministries, duplication in demonstration project, and ensure public trust remains a challenge. She also emphasized the need for a transition to a risk-based regulatory framework to effectively respond to emerging industries such as AI.

Raphael von Thiessen explained the Zurich AI sandbox, which does not provide any relaxation of legal requirements but instead offers regulatory clarity and partnerships to help AI companies conduct real-world tests. This regional approach can be more flexible and responsive to local startup ecosystems. He also cautioned that expectations for sandboxes should remain realistic, as they are a useful but not all-encompassing mechanism.

Romanas Zontovičius pointed out that earning industry trust is critical. He discussed how Lithuania intends to encourage companies by providing expert consultations and legal advice, making it clear that the sandbox environment is intended to support experimentation in good faith, rather than trap participants in punitive oversight.

Sarah Heck, from Anthropic, added that frontier AI companies are looking for spaces where they can candidly present evidence of potential risks without fearing immediate sanctions. Such trust is essential for robust oversight that benefits both developers and society.

Thiago Moraes stressed the importance of civil society involvement to enhance transparency and legitimacy, reminding participants that sandboxes can become even stronger when community organizations and academic researchers are able to monitor processes and share expertise from the outset.

Conclusion

As the event concluded, **Bertrand de La Chapelle** observed that trust was a common denominator throughout the discussions. Companies must feel that sandbox involvement will not expose them to undue risk, while regulators and civil society want to ensure that real accountability mechanisms are in place.

He pointed out that AI governance is not a static endeavor and that sandboxes should be seen as iterative spaces that adapt to the rapid evolution of AI technologies. He encouraged everyone to keep sharing experiences and lessons learned through initiatives such as the Global Sandboxes Forum and further collaborative research and capacity-building.

Participants were reminded that the complexities of AI and its worldwide reach require inclusive problem-solving. True collaboration among government agencies, private companies, civil society, and academia is pivotal for success. Sandboxes provide a structured means of testing, discovering, and sharing insights, but they must operate alongside well-crafted regulations, technical standards, industry codes, and broader accountability frameworks.

The newly released report by the Datasphere Initiative offers policymakers, startups, and researchers an essential reference for designing and implementing effective AI sandboxes. By building global links, encouraging cross-border cooperation, and involving diverse communities, sandboxes can help develop the flexible yet robust governance structures needed for AI's future.

> *"If you're not going to get convergence, how will you manage coherence?"*
>
> Amanda Craig, Microsoft

> *"Transparency and information sharing are what ultimately lead to accountability and capacity building."*
>
> Lucia Russo, OECD

> *"[All AI stakeholders including] AI end-consumers have to recognise their own responsibilities. The Hiroshima [AI] Process is a [voluntary] co-governance mechanism."*
>
> Yoichi Iida, Ministry of Japan

> *"Without principles to guide regulation, regulation can be misplaced."*
>
> Rachel Adams

> *"Asymmetries of power need to be addressed before we talk about interoperability on a global level."*
>
> David Leslie

> *"OpenMined must be global and a public good, and the best way to achieve this is to make the technology open source."*
>
> David Buckley, OpenMined

> *"The specificity of the AI standards we have is that it is not a regulatory sandbox, meaning that you cannot derogate from these obligations."*
>
> Félicien Vallet, CNIL

**Quotations may be paraphrased or shortened for brevity or clarity.*

> *"Singapore has developed open-source testing toolkits for both traditional AI and generative AI, underpinned by frameworks which are interoperable with the key principles that are out there. We are also launching the first third-party assurance pilot."*
>
> Benjamin Chua, Singapore AISI

> *"We need to learn as we go, especially since many decision-makers may not understand the transformative power of digital."*
>
> Cyril Seck, Africa Centres for Disease Control and Prevention

> *"I think it's really important for society, and for this approach in Europe, to consider the impact of these technologies, what they actually do, and whether they will actually work."*
>
> Alex Moltzau, European AI Office

> *"Whenever you have a really big technological revolution on the scale of AI, somewhere in society someone loses a lot of bargaining power, as we see in almost every major technological revolution."*
>
> Matt Prewitt, RadicalxChange

> *"If a company decides to join a sandbox, they need to communicate smoothly with the regulator."*
>
> Sam Jungyun Choi, Covington & Burling

> *"The early stages of sandboxes are important to understand exactly what you want to achieve and what outcomes you want to reach, and to identify all stakeholders who can bring the right skill sets and knowledge to make it a success."*
>
> Caroline Louveaux, Mastercard

**Quotations may be paraphrased or shortened for brevity or clarity.*

> *"For companies to be comfortable participating in sandbox efforts, it is essential to create an environment that ensures the safety of proprietary information. This is especially true in situations where the information involved could be market sensitive in nature."*
>
> Anne Josephine Flanagan, Sandbox Expert

> *"There is a need for a traditional risk-based regulatory framework that can effectively respond to emerging industries such as AI."*
>
> HaeOk Choi, Science and Technology Policy Institute of Korea

> *"Involving multiple regulators in the sandbox is important especially when dealing with emerging and complex innovations such as crypto assets and AI."*
>
> Keith Sabilika, Financial Sector Conduct Authority

> *"Very good communication with the regulator is essential for the success of a sandbox."*
>
> Romanas Zontovičius, Government of Lithuania

> *"When you are looking to engage stakeholders, you need incentives, and you have to determine what those incentives will be."*
>
> Thiago Moraes, PhD Scholar, Vrije Universiteit Brussel

> *"We need spaces for regulators to feel that they can really understand and experience the technology in order to regulate effectively."*
>
> Sarah Heck, Anthropic

> *"The regional layer benefits from close contact with companies and is better able to understand their challenges."*
>
> Raphael von Thiessen, Department of Economic Affairs of the Canton of Zurich

**Quotations may be paraphrased or shortened for brevity or clarity.*