Sandboxes for data:
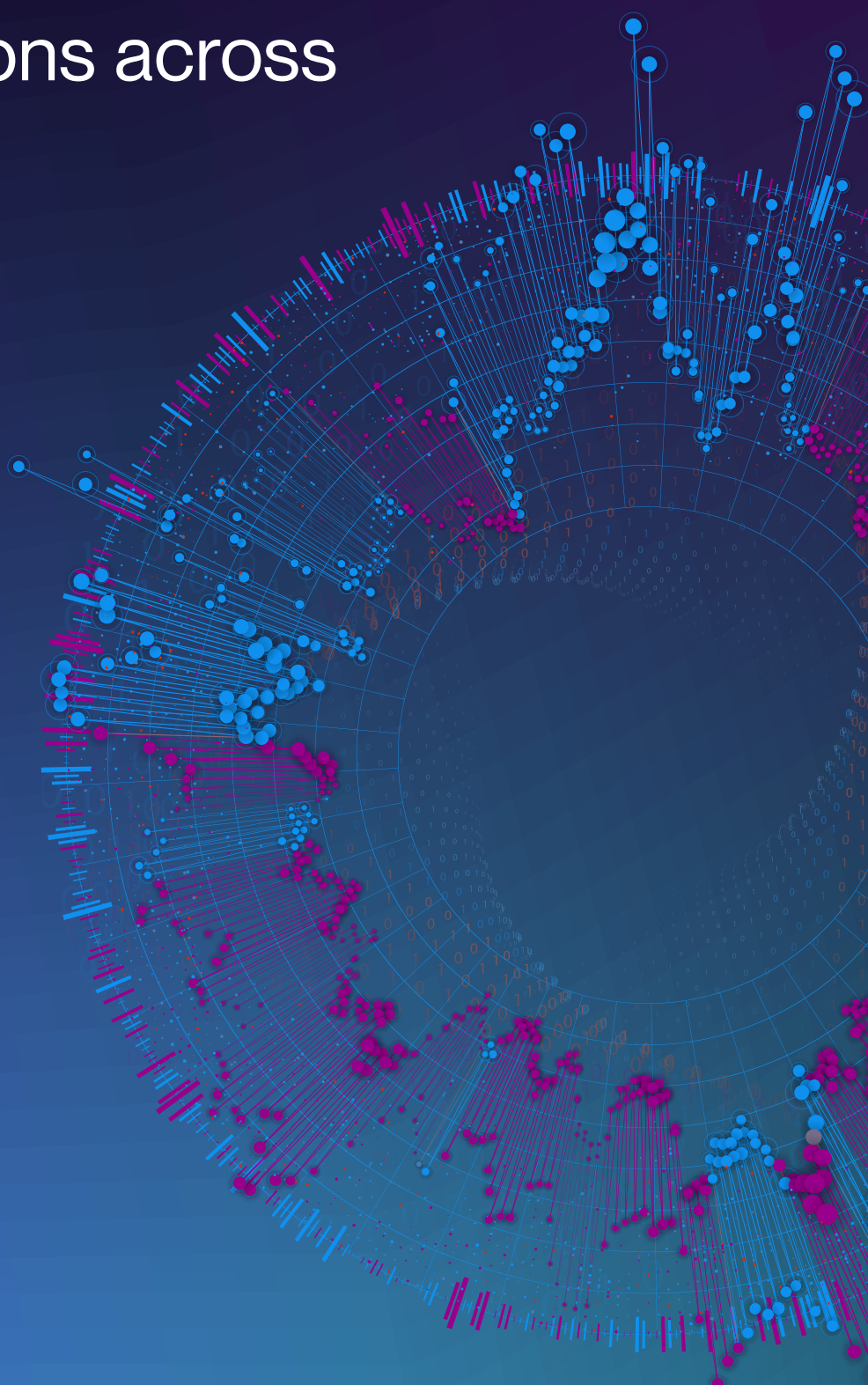creating spaces for
agile solutions across
borders

UK Government

The Datasphere Initiative is a global network of stakeholders fostering a holistic and innovative approach to data governance. By cultivating dialogue and connecting communities, the Datasphere Initiative connects sectoral silos and people to build a collaboratively governed Datasphere and responsibly unlock the value of data for all.

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

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

EXECUTIVE SUMMARY

Broadly speaking, operational sandboxes are testing environments where hosted data can be accessed and used, while regulatory sandboxes are collaborative processes where regulators and firms evaluate new technologies within a regulatory framework. More succinctly, operational sandboxes actually handle data, and regulatory sandboxes provide dialogue and guidance on how data is handled.[1]

This report considers existing kinds of sandboxes, and seeks to apply lessons learnt towards the designing and implementation of cross-border sandboxes for data.

This report's typology of regulatory sandboxes shows that they vary according to the flexibility or leeway permitted by regulators, how directly the findings are used to shape or inform future laws, and if any incentives are offered to innovators to participate. These differences are shaped by the legal and cultural environment of each national regulator. Sandboxes also vary in goals, scope and scale.

Sandboxes are an agile response to the 'Datasphere', the complex system encompassing all types of data and their dynamic interactions with human groups and norms. This report argues that the new and complex challenges brought by data require innovative governance mechanisms to address them. Sandboxes are one such mechanism. They can reduce regulatory uncertainty, help emerging innovators, build capacity within and cooperation between regulators, and increase regulatory clarity and compliance. However, they are also resource-intensive, can increase risk and face difficulty in scaling up.

Cross-border sandboxes for data can address these challenges by using careful design and issue identification and definition, choice of regulatory partners, stakeholder engagement and active management of information asymmetry and risk. Ongoing assessment and iterative improvement in the design and operation of the sandbox will also be essential.

This report analyses a range of cross-sectoral innovations – emerging intermediaries, privacy enhancing technologies (PETs) and browser-based consent management to identify areas of regulatory opportunity that sandboxes could address. It also summarises challenges and opportunities in wider applications of data-flows like innovation and trade, health data, mobility and biodiversity genomics where sandboxes could be used to tackle friction and bottlenecks.

Currently, cross-border sandboxes for data are just emerging. An initiative by the Association of Southeast Asian Nations ASEAN and the GSM Association (GSMA) is just beginning. Provision for sandboxes has been made in agreements such as the Canada-United States-Mexico (CUSMA) trade agreement, which refers to cross-border sandboxes for financial services, and the Digital Economy Partnership Agreement (DEPA) between Singapore, Chile and New Zealand. These sandboxes are still in the planning phase. This report and its roadmap are a contribution to the evolving discussions on how to implement cross-border sandboxes for data.

---

1. IMDA and PDPC (2019), *Public-Private Data Collaboration Case Study*, Infocomm Media Development Authority.

Cross-border sandboxes have the potential to improve regulatory capacity-building and cooperation across borders, increase innovation, competition and choice in many markets, enhance compliance and reduce regulatory arbitrage, and make data more available and accessible across borders and sectors.

The report's roadmap sets out a preliminary guide for how to design and initiate cross-border sandboxes for data. It also sets out key questions that must still be addressed, including how to ensure compliance across borders and what kind of internationally binding commitments innovators and regulators can make.

The report is structured in six sections. **Section 1** defines the different types of sandbox, describes their uses, sets out a typology of key differences between regulatory sandboxes, and summarises the emerging examples of cross-border sandboxes. **Section 2** sets out the lessons learnt from sandboxes so far, and how they have been applied to data. **Section 3** looks at innovations in technology and new approaches to data - emerging intermediaries, PETs and browser-based consent management - to identify issues sandboxes could be used to address. It also examines policy or sector challenges where sandboxes could help to bridge divides for innovation and trade, health, mobility and biodiversity genomics. **Section 4** sets out the main challenges of implementing cross-border sandboxes, and how these challenges could be addressed. **Section 5** outlines the process or model for setting up cross-border sandboxes for data. **Section 6,** the conclusion, summarises the main findings and identifies future questions to be answered for the successful design and implementation of a cross-border regulatory or operational sandbox.

INTRODUCTION

The fast pace of technological change and increasing reliance on new technologies have transformed societies and the way we create, use, consume and share data. At the same time, this unprecedented leap and the disruptive nature of technological developments have challenged existing institutional and regulatory arrangements, and how social and economic value is distributed in society.

Data is created, collected, stored, processed, accessed, used, shared and destroyed in numerous geographic locations. Yet, discourse about data tends to focus on only a few of these components: the location of its storage, processing and sharing. In a context where cloud services have proliferated worldwide, the actual location of data can be hard to determine, sometimes by design. Indeed, not only is it moved almost at the speed of light along its processing workflow, but it can be replicated in diverse locations for reasons of security (e.g. geo-redundancy and infrastructure resilience) or efficiency (e.g. speed of delivery and specialisation of networks, such as content delivery networks). It can also be split into many different and distributed pieces (e.g. sharding) to facilitate processing.[2] The rapid and complex movement of data between countries, and the growth of innovative business models that take advantage of this, can bring regulatory complexity. Innovators may face many and overlapping applicable laws, creating uncertainty about compliance. Governments may respond to concerns about the rapid movement and international processing of data with measures, such as data localisation, that can unnecessarily curtail the broad social and economic benefits of data.

The importance of data, both personal and non-personal, in our society calls for a more holistic conceptual framing of data governance models and ongoing innovation in how they are applied. While data protection law continues to apply to personal data, there are often inconsistencies or gaps in understanding and agreement about how and when it should be enforced. Considering data governance through the lens of what we may call the Datasphere, that is, "*the complex system encompassing all types of data and their dynamic interactions with human groups and norms*", rather than using incomplete analogies (e.g. data as oil), is a first step.[3] A second is recognizing that innovative governance, collaboration and enforcement mechanisms are needed. A systems approach to data governance should not only rely on expecting changes in attitude from the various actors of the global Datasphere, but also on the institutional frameworks that can enable these changes.

There is currently a lack of mechanisms to bridge the existing policy silos in health, mobility and genomic biodiversity data described in this report. Iterative collaboration involving all categories of stakeholders is a priority prerequisite to the transdisciplinary collaboration data governance requires. On that basis, regulators, firms and other stakeholders could experiment with dynamic and modular arrangements, such as through cross-border regulatory sandboxes.

---

2. De La Chapelle, B. and L. Porciuncula (2021), *We Need to Talk About Data: Framing the Debate Around Free Flow of Data and Data Sovereignty*, Internet & Jurisdiction Policy Network.

3. De La Chapelle, B. and L. Porciuncula (2022), *Hello Datasphere — Towards a Systems Approach to Data Governance*, Datasphere Medium.

Regulatory sandboxes are collaborations that bring together regulators and organisations with new technologies and processes to test the innovations against the regulatory framework. Regulatory sandboxes have a potential use anywhere innovation is hampered by uncertainty about how regulation applies, or where regulation is emerging and risks are not fully understood. Sandboxes may even help to level the playing field between established and emerging players, by providing a structured and low-risk way to test out challenging new ideas that would otherwise need significant institutional backing before they could even contemplate market entry.

Regulatory sandboxes for data have been used by regulators in the United Kingdom, Norway, Canada, Singapore, Colombia and several other countries. So far, they have all focused on the application of national regulations within each regulator's jurisdiction. While several issues need to be resolved to make them work, cross-border regulatory sandboxes have significant potential to address barriers to international data-flows, particularly regulatory uncertainty between jurisdictions and challenges that require multi-disciplinary collaboration.[4]

Cross-border regulatory sandboxes could also respond more agilely to the rapidly growing field of technologically-enabled and decentralised bottom-up innovations; initiatives such as data intermediaries (e.g. data trusts, fiduciaries, collaboratives), PETs, and browser-based solutions for consent management. The potential for cross-border sandboxes to clarify key issues around each of these innovations and widen their use is explored in the report. Taking inspiration and learning from agile methods, emerging business models and technological solutions could inform the development of new solutions to data-related challenges, and also drive innovation in governance mechanisms themselves.

As this report shows, the stakes are high on tackling the emerging global challenges such as climate change and public health by removing barriers in cross-border data flows, such as flexibility and clarity around regulation. Regulatory frameworks and their traditional modes of implementation need to evolve. They need to be able to accommodate new approaches to address present challenges and to future-proof them to accommodate ongoing, rapid technological and business model change. There is a clear need to understand how this agile regulatory model can strengthen cross-border, sectoral and cross-sectoral data collaboration.

This research report describes how regulatory sandboxes for data work today, analysing their opportunities, benefits and challenges, to set out how they can begin to be applied to cross-border data flows and governance. It presents case studies in diverse policy areas - international data flows, health, mobility and biodiversity - to show where specific challenges might be addressed by cross-border sandboxes. To date, a fully operational cross-border regulatory sandbox for data has yet to be launched. This report sets out an initial model for how the first cross-border regulatory sandbox for data could work, and identifies questions that need to be answered for the process to begin.

---

4 BMWi (2019), *Making Space for Innovation The handbook for regulatory sandboxes*, Federal Ministry for Economic Affairs and Energy, Germany.

01

Understanding Regulatory Sandboxes

Definitions and background

Sandboxes were originally modelled on the enclosed software testing environments that programmers create to test and run code and applications without risking the overall system they work in. This practice was adapted in at least two different ways, evolving into regulatory sandboxes and operational sandboxes (also known as data spaces).

Regulatory sandboxes are time-limited collaborative endeavours involving regulators, service-providers and other relevant stakeholders to test innovative technology and data practices against regulatory frameworks. They were first used by financial regulators that adapted the software testing environment of sandboxes to create closed environments where new 'fintech' is tested - both to check compliance with regulations and, in some jurisdictions, to investigate if the regulations themselves needed to be updated. The sandbox setting ensures that the risks either to consumers or the financial system are contained.

Regulatory sandboxes are not physical spaces. They are a collaborative process through which technological innovations and new business processes and models are explored with regulators. The Consultative Group to Assist the Poor (CGAP), a partnership on financial innovation hosted by The World Bank, defines a sandbox as "*a framework set up by a regulator that allows fintech start-ups and other innovators to conduct live experiments in a controlled environment under a regulator's supervision.*"[5] The majority of experiences of regulatory sandboxes has been limited to domestic use, within the domestic boundaries of national regulators.

In the past five years, regulatory sandboxes have begun to be adopted by different regulators. Countries including the UK, Germany,[6] Norway, Canada, Colombia[7] and Singapore have set up sandboxes that companies can enter to test new ways of collecting, using or moving data around.

Operational sandboxes are secure, collaborative data-spaces that pool data-sets and resources together. They are hosted by one entity and accessed by several others, to serve new combinations and uses of data. They can be created by regulators or government agencies to test capabilities on actual data-sets, or by a coalition of actors to pool resources together, sometimes through technologically enabled decentralised approaches (e.g. data collaboratives, fiduciaries, commons), in order to explore or encourage their use. Examples of operational sandboxes of this type include:

- A sandbox created by Health Data Research UK (HDRUK), the United Kingdom's national institute for health data services. The HDRUK Sandbox creates; "*a virtual testing environment for products, services, innovations or knowledge that require access to large scale data to improve the health of patients and the population*".[8] It permits access to the data to approved researchers.

Broadly speaking, operational sandboxes are testing environments where hosted data can be accessed and used, while regulatory sandboxes are collaborative processes where regulators and firms evaluate new

---

5. CGAP (2018), *Regulatory Sandboxes: What Have We Learned So Far?,* Consultative Group to Assist the Poor Blog.

6. BMWi (2019), *Making Space for Innovation The handbook for regulatory sandboxes*, Federal Ministry for Economic Affairs and Energy, Germany.

7. RedIPD (2021), *Colombia Data Protection Authority launches innovative regulatory Sandbox on privacy by design and by default in artificial intelligence projects*, Red Iberoamericana de Protección de Datos.

8.HRDUK (2019), *Call for applications to test new £37.5M UK health data research services,* Health Research Data UK Blog.

technologies within a regulatory framework. More succinctly, operational sandboxes actually handle data, and regulatory sandboxes provide dialogue and guidance on how data is handled.[9]

Drilling down further into regulatory sandboxes, there is significant variation between them. The next section presents a typology and description of how regulatory sandboxes vary, and why.

Typology of regulatory sandboxes

Regulatory sandboxes are an "incubation model"[10] for innovative products, services or business models. They are generally time limited, i.e. a sandbox is tasked to explore a specific technology or practice for a limited amount of time, and when that exploration is complete, the firm leaves the sandbox. However, regulatory sandboxes may also have renewed calls for participants. Sandboxes usually have safeguards against risks. The risks include data breaches, re-identification of personal data, or other misuses of data. Broadly speaking, sandboxes vary in the following ways:

- **Goals:** Given sandboxes' origins in fintech, some sandboxes are part of an existing ecosystem of business development and start-up support.[11] Their goals may be part of a country's broader strategy to champion itself as a base for new technologies, services and business models. More broadly, regulators' goals can be strategically aligned more towards economic innovation and growth than regulatory compliance and minimising risk. Sandbox goals are influenced by the same factors that determine regulatory leeway and discovery; sandboxing may be placed within existing legal frameworks or as a means to explore and test new ones.

- **Scope:** Issues dealt by sandboxes can vary from being narrowly focused (e.g. concerning a particular technology or regulatory gap) or having a broad scope (e.g. tackling innovation in a sector or policy area as a whole).

- **Scale:** Sandboxes vary in the scale of the initiatives and stakeholders engaged in the sandboxes (either on the type or number of actors) and also in the geographic reach of their mandate (local, national or transnational, that is, cross-border). Regulators vary regarding the time and resources they can allocate, and this may limit the number and scope of activities taken into a sandbox.

> **Sandboxes vary in the scale of the initiatives and stakeholders engaged in the sandboxes (either on the type or number of actors) and also in the geographic reach of their mandate (local, national or transnational, that is, cross-border).**

9. IMDA and PDPC (2019), *Public-Private Data Collaboration Case Study*, Infocomm Media Development Authority

10. Alaassar, A., AL. Mention and TA. Helge (2021), *Exploring a new incubation model for FinTechs: Regulatory Sandboxes* Technovation.

11. Alaassar. A., A. Mention and T. Aas (2021), *Exploring a new incubation model for FinTechs: Regulatory sandboxes*, Volume 103, Technovation.

Drilling down on goals, sandboxes are shaped by differing national contexts and strategies.[12] Sandboxes' goals can also vary in the following ways:

- **Regulatory leeway -** Depending on the regulator's legal environment, sandboxes can offer varying degrees of flexibility on compliance for the duration of the sandbox. While the OECD defines a regulatory sandbox as "*a limited form of regulatory waiver or flexibility for firms, enabling them to test new business models with reduced regulatory requirements*",[13] not all regulators are legally able to offer this. Sandboxes differ on whether they fully waive or partially relax regulatory rules for the period of testing and development, and what incentives, if any, they can offer firms.

- **Direct impact of regulatory discovery -** Sandboxes are intended for truly novel technology or services whose success would broadly serve the public interest. They are also motivated by what policymakers and regulators can learn, potentially guiding future enforcement and legislation.[14] This use of a structured and interactive process of information-sharing drives transparency. Sandboxes differ on how directly their learnings are applied to policy; from the regulator directly triggering a pre-legislative consultation on a specific regulatory change, to providing contextual information to the ongoing evolution of policy discussions.

- **Availability of incentives for firms to participate -** Sandboxes offer varying incentives to innovators to participate. These can involve direct financial support for legal, technical and other costs involved in participation (see Box 1. Singapore's 'FinTech Regulatory Sandbox', below). More commonly, the main incentives are regulatory clarity and certainty, and the knowledge and insights gained and shared from focused interaction with regulators.

Each of these elements strongly reflects the legal and cultural environment of national regulators, and is a way in which national regulatory sandboxes may differ. The table below, sets out, from low to high, the degree to which each element can be present in regulators and their contexts (Table 1).

> **Sandboxes are intended for truly novel technology or services whose success would broadly serve the public interest. They are also motivated by what policymakers and regulators can learn, potentially guiding future enforcement and legislation. This use of a structured and interactive process of information-sharing drives transparency.**

12. Parenti (2020), *Regulatory Sandboxes and Innovation Hubs for FinTech,* Policy Department for Economic, Scientific and Quality of Life Policies.

13. Attrey, A., M. Lesher and C. Lomax (2020), *The role of sandboxes in promoting flexibility and innovation in the digital age,* Going Digital Toolkit Note, No. 2.

14. Fintech Notes No. 8 (2020), *Global Experiences from Regulatory Sandboxes, Finance, Competitiveness and Innovation Global Practice,* World Bank Group.

⊞ **Table 1.** **Typology of regulatory sandboxes**

Element/ Flexibility	Low	Medium	High	Opportunities and risks
Regulatory leeway	No leeway at all. Sandbox exists purely for dialogue and clarification on regulatory compliance.	Regulator cannot suspend regulations, but can offer other measures which provide for a relaxation of enforcement by the regulator, e.g. 'letter of regulatory comfort'. *E.g. UK Information Commissioner's Office, which has issued such letters to participants in its regulatory sandbox.*	Regulator can temporarily suspend regulations for the duration of the sandbox. *E.g. Monetary Authority of Singapore, 'Sandbox Plus' suspend regulations for the duration of the sandbox. Moreover, German regulators in some sectors - including aviation and media - have temporarily exempted innovators from compliance, using 'experimentation clauses', but not (yet) within a sandbox.*[15]	A high degree of regulatory leeway means the regulator can offer greater incentives to innovators to join the sandbox and use insights to directly encourage specific innovations. It also increases risks to consumers, and potentially the whole sector or economy, as regulations are suspended.
Direct impact of regulatory discovery	Regulator uses insights and examples from the sandbox broadly to provide input to the evolution of regulation but does not use sandbox process to launch or trigger direct changes to legislation.	Regulator collects insights and examples from sandbox to guide future targeted legislation. *E.g. Economic Affairs Ministry of Germany designed a regulatory sandbox for autonomous electrical watercraft that is also directed to pro-actively generate insights into the future development of the relevant rules.*	Regulator can initiate regulatory / legislative change based on sandbox findings. *E.g. Personal Data Protection Commission of Singapore has used sandbox finding to directly trigger regulatory change. (For details, See Box 2. below)*	A high degree of regulatory discovery means the regulator can use sandbox findings to directly trigger proposed legislative changes. This increases regulatory agility and responsiveness, but risks favouring innovators' interests at the agenda-setting stage of legislation.
Availability of incentives for firms to participate	Regulator offers non-financial incentives including regulatory clarity (especially to new entrant firms without access to costly legal expertise), dialogue and learning.		Regulator can offer financial incentives - e.g. cover legal and operational costs - for firms entering the sandbox. *E.g. Monetary Authority of Singapore, 'Sandbox Plus' and Singapore's 'FinTech Regulatory Sandbox'. (Box 1)*	A high level of incentives to firms can encourage new market entrants to enter the sandbox and help improve competition and innovation. The risk is inequitable or inefficient spending of resources and potential breaches of 'state aid' rules to industry, where those rules exist.

Source: Datasphere Initiative (2022)
Note: 'Low', 'medium' and 'high' indicated the comparable degree of each element, i.e. low to high degrees of regulatory leeway, direct impact of regulatory discovery and availability of incentives for firms to participate.

---

15. BMWi (2019), *Making Space for Innovation The handbook for regulatory sandboxes*, Federal Ministry for Economic Affairs and Energy, Germany.

As the typology above shows, there is significant variety between different regulatory sandboxes which reflects the statutory role of the regulator, transnational, national and sectoral regulations, as well as the broader role that risk plays in the regulatory culture. Singapore, for example, was an early and highly motivated adopter of regulatory sandboxes in fintech, not least as this is an area of strategic importance[16] for the country's industrial strategy (Box 1). The resource-intensity and degree of synergy required between the public sector and the private sector in this approach should not be underestimated.

⬡ Box 1. Singapore's 'FinTech Regulatory Sandbox'

Singapore's 'FinTech Regulatory Sandbox' is run by the Monetary Authority. It "enables financial institutions and fintech players to experiment with innovative financial products or services in a live environment but within a well-defined space and duration." It is only open to banks or firms bringing a new technology not yet used in Singapore, i.e. applicants must be 'first movers'.

The sandbox relaxes the specific legal and regulatory requirements of the Monetary Authority of Singapore for the duration of the sandbox, but also includes safeguards to contain the consequences of any failures. When firms leave the sandbox they must comply with regulations.

The sandbox is part of a broader system to encourage financial innovation. Its 'Sandbox Express' started in 2019 and gives faster market testing. Companies are funded half the cost of entering it, up to SGP 500,000, and the sandbox is connected to a 'first Fridays' deal-making network.

One graduate of the sandbox is Investax.io, a "leading investment and trading platform for Digital Securities and Security Tokens of global private markets deals including venture capital, private equity, real estate , digital SPACs and more."[17]

Source: MAS (2021), Sandbox Plus, Monetary Authority of Singapore Blog.

---

16. FF (2020), *Fintech powerhouse: Understanding the rise of Singapore,* Fintech Futures Blog.

17. InvestaX (n.d), *Homepage Website,* (accessed 29 March 2022).

⬡ **Box 2. Regulatory Discovery: use of a regulatory
sandbox for data to update the law**

The Personal Data Protection Commission (PDPC) of Singapore started a regulatory sandbox for data in 2017. It operates in consultation with the communications and media regulator, the Infocomm Media Development Authority (IMDA). By the end of 2019, it had already worked with almost thirty companies. The sandbox has three stages:

1. Engagement – in-depth discussions with organisations using novel technologies, to see how regulations apply
2. Guidance – when novel technologies expose ambiguity or silence in existing regulations, the regulator can issue specific guidance. The guidance is published for everyone to read, minus any commercially sensitive information.
3. Policy prototyping – the regulator collaborates with industry "to co-create new policies to address gaps identified".

A gap was identified by the sandbox in the use of personal data for business innovation. Industry wanted more clarity on how data could be used for "*product development, operational improvements, and understanding customers better. So, together with the PDPC, a private sector-led committee spent about four months formulating a new policy proposal*."

The proposal that was identified by the sandbox and prototyped with industry then went through the normal consultation and legislative process.

Benefits: Companies get the clarity needed to deploy new technologies and methods, and the PDPC gets "*greater confidence to adjust our data protection policies so that they remain relevant to developing industry practices while ensuring a high level of data protection*".

Risks*:* 'Policy prototyping' is a targeted and intentional process to apply what is learnt in the sandbox to develop new legislation. Regulatory sandboxes in Singapore expect the regulator to stand in place of citizens and users, meaning input from other stakeholders comes much later, when a policy proposal has already been made. This may put other interests - for example those who want to limit the use of data to protect privacy - at a disadvantage.

Source: Zee Kin, Y. (2019), _Keynote Speech by Deputy Commissioner at AI and Commercial Law: Re-imagining Trust Governance and Private Law Rules_, Singapore Management University Blog.

Due to the varying goals, scope, scale (as well as degrees of regulatory leeway, direct impact of regulatory discovery and availability of incentives for firms to participate), it follows then that costs, benefits and risks vary along with different types of regulatory sandboxes. Experience in Singapore, for example, shows the benefits and risks related to one of the variables - regulatory discovery (Box 2). It shows that a simplistic reading of the typology - that higher levels of regulatory agency make sandboxes better - would not be correct. The different approaches of national regulators each have their merits.

Despite the limitation of metrics or case studies on Singapore's relaunched fintech sandboxes, likely due to how recent they are, it could be concluded that regulatory discovery - as a variable to distinguish between different national approaches to sandboxes - does not produce a simple rubric of 'higher is better'. While the total number of regulatory sandboxes for data globally is still far too small to draw firm conclusions, variations in 'high' to 'low' regulatory discovery align somewhat with whether the goals of the sandbox seem to prioritise innovation over privacy.

Aside from the availability or targeting of resources, the typology highlights some key overall differences between national regulators which may come into play when cross-border data regulatory sandboxes are designed. The most obvious difference between regulators is of course the legal and regulatory framework they use. However, the typology shows that what regulators fundamentally differ on is their scope of regulatory action. Cross-border sandboxes will need to work in a focused and transparent way to surface and address these differences so that regulators can work together within their differing levels of agency. Although this is a significant challenge, the typology provides a scale of the impacts sandboxes can have, given the right conditions and sufficient motivation to unlock benefits.[18]

To summarise, this typology is an initial exercise, based on a very small number of known regulatory sandboxes for data. It shows that there are significant differences in how those sandboxes operate, and suggests that these differences are based in different legal regimes but also in national strategies. While these differences are not insurmountable, they do surface issues that would need to be addressed in designing and running a cross-border regulatory sandbox for data; particularly its goals.

Emerging cross-border experiences

Cross-border sandboxes involve regulators from two or more jurisdictions that aim to achieve some regulatory alignment or certainty for new technologies or activities. While most experiences of regulatory sandboxes are national, there is a small number of emerging cross-border and transnational regulatory sandboxes.

In the financial sector, cross-border regulatory sandboxes can provide "*streamlined licensing and reciprocal license arrangements, reducing the regulatory burden on firms looking to scale*".[19] Their goals typically include improved regulatory alignment and enhanced ability of firms to scale up, regionally or globally.[20] The Global Financial Innovation Network (GFIN) (Box 3) and the Pacific Islands Regional Initiative (RIRI) (Box 4) provide illustrative examples of an ongoing cross-border regulatory sandbox for financial services .

---

18. Given how recent many sandboxes experiences are and that reporting is not available for all of the few experiences, evaluations of their effectiveness or inferences on their ability to scale should be limited.

19. WBG (2020), *Global Experiences from Regulatory Sandboxes, Finance, Competitiveness and Innovation Global Practice, Fintech Note No. 8* , The World Bank Group.

20. WBG (2020), *Global Experiences from Regulatory Sandboxes, Finance, Competitiveness and Innovation Global Practice, Fintech Note No. 8* , The World Bank Group.

⬡ **Box 3. Global Financial Innovation Network (GFIN)**

GFIN is a platform for regulatory cooperation and collaboration by over 50 national and state-level financial regulators. GFIN's cross-border testing capability is "*an environment that allows firms to consecutively or concurrently trial and scale new technologies, products or business models in multiple jurisdictions, gaining real-time insight into how a product or service might operate in multiple markets.*" Its first intake of testers was scheduled to begin in late 2021. One of the two firms involved is ATOME:Matter, a metadata management firm that organises "data reporting requirements into easily understandable data concepts".

Led by the Lithuanian financial services regulator, a group of regulators is overseeing the assessment of ATOME:Matter's activities by a Polish assessor that specialises in data reporting standards. The goal is to come up with standardised ways of reporting sustainability metrics like carbon capture that work both for financial firms and the regulators who need to be able to interpret them. The sandbox is expected to conclude and report back in 2022.

Sources: GFIN (2022), *The GFIN Cross-border Testing*, The Global Financial Innovation Network Blog.

⬡ **Box 4. The Pacific Islands Regional Initiative (RIRI)**

The Pacific Islands Regional Initiative (RIRI) plans to launch a world first, a regional regulatory sandbox for financial services, in 2022. The sandbox will be run by central banks in Fiji, Papua New Guinea, Samoa, Seychelles, Solomon Islands, Timor-Leste, Tonga and Vanuatu, and will create a fintech testing environment shared by all eight countries.

The sandbox will have a single application process and will allow innovators to test their services in multiple jurisdictions. It aims to "present the Pacific as a single market to build scale and attract solutions that would previously have been inaccessible within each individual country." While it is not yet clear how the sandbox will manage differing legal regimes, it is expected to drive both collaboration and capacity-building between central banks.

The overall goal is to generate foreign investment by helping firms tap into a larger, more diverse customer base, to reduce regulatory and legal choke-points, and ensure that new fintech-based products and business models are sustainable.

As the sandbox has not yet launched, it is too soon to identify company case studies or draw any broader conclusions on how this cross-border sandbox overcomes the challenges of regulatory harmonisation or equivalence.

Sources: Shust, P. (2022), *Pacific sandbox unites regulators in supporting fintech innovation*, Alliance for Financial Inclusion Blog.

Despite the challenges to their implementation (see Section 4), the opportunities cross-border sandboxes could create are significant. Sandboxes that deal with data could improve regulatory certainty, improve data-flows, and allow new and innovative firms better access to customers and services abroad. From the regulator's perspective, they could reduce regulatory arbitrage and forum-shopping and build more consistent compliance, based on mutual support, collaboration and alignment among the involved regulators.

To date, the number of cross-border sandboxes is too small, and it is too soon for them to create and report on results, for researchers to reach concrete conclusions about their use, efficacy and scale. It seems likely, however, that as well as agreeing on overall regulatory goals, successful cross-border sandboxes will need to inhabit a 'sweet spot' where their national frameworks are not harmonised - and so require little or no extra intervention - but not so different that collaboration and alignment are impossible.

02

Drawing lessons from Regulatory Sandboxes

A few lessons have been drawn from the use of regulatory sandboxes in the past decade, albeit largely within the fintech world.[21] Broadly speaking, the following lessons could be applicable to other policy areas:

- Sandboxes can be an effective way to **reduce legal and regulatory uncertainty**, giving innovators access to a controlled testing environment, and this subsequently facilitates market entry, enhancing overall competition and access to innovation.[22]

- Sandboxes can provide clear **benefits for regulators**, giving them first-hand contact with the latest technological developments and solutions. This builds capacity and helps regulators to anticipate developments driven by technological change and future business models.[23]

- When developed in a multistakeholder and collaborative way to include the expertise both of innovators and impacted actors, sandboxes can help to **build consensus and trust**, and **spur innovative collaborations** among different stakeholders, including across borders.

- Where regulatory requirements are unclear or missing, or where they create barriers to entry that may be disproportionate to risks, sandboxes can help to **develop clarity and certainty**, and to communicate them to market actors. Where regulators are driven by transparency requirements and ethos, communicating sandbox outcomes can help drive greater public and political understanding of evolving technologies, practices, rules and norms.

- However, depending on how they are constituted and run, sandboxes can also **fail to capture and mitigate risk**, both to citizens and the overall economy. As elsewhere, regulators often work with **information asymmetry**, whether in the financial services or technology sectors. Regulators may not have the knowledge or skill to design sandboxes that anticipate and control for risk, as initiatives in the sandbox become more innovative.

How sandboxes could be useful for data

Data now underpins and reflects practically all economic sectors and social activities. Digital data, personal and non-personal, private and public or government held, is organised in datasets of diverse sizes and types. Data can largely stem from the most *public spheres*, such as data concerning the planet, its resources and infrastructures (e.g. climate, oceans, wildlife, national parks, submarine cables, pipelines, ports, railroads), to more *private spheres*, such as data on *private activities* of public companies, private companies (e.g. trade secrets), and of people (e.g. doctor appointments). However, this is not a clear cut distinction and classifications have blurred, overlapping and moving boundaries.

---

21. Attrey, A., M. Leshe and C. Lomax (2020), *The role of sandboxes in promoting flexibility and innovation in the digital age, Going Digital Toolkit Note,* No. 2.

22. Appaya, S., Amah and M. Haji (2020), *Four years and counting: What we've learned from regulatory sandboxes* World Bank Blogs.

23. SP (2021), *Sandbox benefits for supervisors and regulators,* Sandbox in Spain Blog.

The Datasphere can be categorised as a complex adaptive system;[24] a multitude of individuals and organisations can, individually or in aggregate have a great impact on the entire system. Moreover, chains of actions and reactions produce emergent trajectories for the whole of society, not determined by any single authority.[25] An added level of complexity in the Datasphere, however, is that differently from natural systems, public and private actors have a capacity to intervene at a meta level, to modify the norms applicable to others regarding the collection, processing, access or use of data.

The new and complex challenges brought by data and the Datasphere require innovative governance mechanisms to address them. In the absence of robust international cooperation between different stakeholders, numerous uncoordinated unilateral measures have been adopted, including those enforcing data localisation.[26] Moreover, many of these measures have been taken as a result of siloed approaches to data governance - where topics including privacy, cybersecurity, national security, content moderation, digital trade, or taxation are addressed separately, rather than interrelatedly.

Given the major interdependencies between sectors and policy issues, a systems approach to data that fosters innovative and holistic approaches is urgently needed. Sandboxes could be used as a testing space for more dynamic arrangements to address challenges and unlock opportunities in the Datasphere, offering the kind of flexibility required by a system's approach that needs iterative adjustments to elaborate webs of positive and negative feedback loops governing emergent behaviours.

Regulatory sandboxes have already been tested to address challenges related to data, particularly personal data, in domestic contexts. For other types of data, however, such as those concerning public spheres or non-personally identified data, they have not yet been tested, despite the potential positive spillovers of unlocking the responsible sharing of those types of datasets, either through operational sandboxes or by clarifying and flexibilizing regulations through regulatory sandboxes.

Sandboxes for data could be used to work through the potential regulatory issues of a new data-related product, service or business model that may not fit within existing regulations but that could nonetheless offer broader public benefits such as promoting innovation, competition or access. As test-beds for innovation, sandboxes can help navigate through rapid and consequential technological and economic changes, including the impacts of technologies and practices such as artificial intelligence, ubiquitous computing, blockchain adoption, digital currencies and of the multiple uses of different types of data and their non-linear value-chains.

Sandboxes for data can be leveraged to address particular data issues or to test new products and business models, including:

- New technologies, services and business models that do not easily fit within current regulatory frameworks, or whose novel mix of positive and negative externalities require a regulatory assessment to understand their impact on risk.

- Cross-sectoral innovation, involving different and potentially conflicting national regulatory frameworks.

24. Carmichael, T. and M. Hadzikadic (2019), *The Fundamentals of Complex Adaptive Systems, Complex Adaptive Systems*, Research Gate.

25. De La Chapelle, B. and L. Porciuncula (2022), *Hello Datasphere — Towards a Systems Approach to Data Governance*, Datasphere Medium.

26. Cory and Dascolly (2021), *How Barriers to Cross-Border Data Flows Are Spreading Globally, What They Cost, and How to Address Them*, Information Technology and Innovation Foundation Blog.

- Products or policies that affect currently under-served communities,[27] for example new products and services to improve financial inclusion, or policy initiatives for under-served groups across different policy silos.

Sandboxes are spaces to experiment, hold dialogue, learn, imagine and co-create new ways to ensure technological development benefits everyone, including not only private sector and data regulators, but also related actors, such as civil society, academics, data subjects and even other regulators, both nationally and internationally. In this context, sandboxes may provide an additional, pro-active way for regulators to test and ultimately develop *ex ante* regulation that enables the trusted use of data.

They can also be leveraged for actors to pool resources together. Cases where that would be particularly useful could be:

- Actors from the same sector.
- Actors across sectors to explore interdependencies and positive and negative externalities.
- Research entities getting access to portions of public or private data for explicitly mentioned use cases.
- Private data shared with public authorities (e.g. mobility data for smart cities).
- Public data made available for private entities to develop services.
- Specific arrangements to maintain certain data exchanges even in situations of political tension (e.g. environmental data, epidemic-related data sharing, etc.).

Benefits of regulatory sandboxes

Sandboxes have the potential to offer important benefits in the field of data and regulation; agility and responsiveness, rich sources of qualitative and quantitative evidence for current and future regulation, facilitated product and service development and market entry for firms and public sector organisations, along with reduced legal uncertainty, capacity-building for regulators, fostering partnerships and building consensus and trust, and filling the gap where regulation is absent or unclear.

Sandboxes may also facilitate better evidence-based policy-making. They can be a key part of a data-grounded regulatory policy and legislative development, in addition to enriching and extending the knowledge, capacity, ongoing learning and development needs of regulators.

When they work well, sandboxes operate bi-directionally; they help organisations to understand and apply regulation to novel technologies, services and business models and they help regulators understand emerging practices and models, and feed into regulatory development. For emerging initiatives and technologies (Section 3), sandboxes can be a crucial space to test conceptually challenging new ideas that do not obviously or easily fit into current frameworks.

It is important to note that regulatory sandboxes are not the same as deregulation, nor are they used with simply this end goal in sight. On the contrary, where the legal situation is unclear or meaningful legislation has yet to be created, they can provide a stopgap solution to determine and impose the relevant rules in real time.

---

27. Jenik, I. (2017), *Regulatory Sandboxes and Financial Inclusion,* CGAP (Consultative Group to Assist the Poor) Blog.

However, while sandboxes clearly have huge potential to answer many questions and solve problems, it is important to look into what kinds of issues in the data regulatory sphere existing sandboxes have actually helped resolve. The experience from one of the recent sandboxes on monetary analysis from the United Kingdom Information Commissioner's Office sets out both the challenges and lessons learnt (Box 5). It shows that sandboxes can provide enormous value, but also that they take focused, collaborative work grounded in careful preparation to ensure concentration on bounded questions, and a close fit between the activity to be sandboxed and the public interest served by dedicating resources to doing so.

Box 5. United Kingdom's Regulatory Sandbox on Monetary Analytics and FutureFlow

FutureFlow is a United Kingdom start-up in the use of monetary analytics to detect electronic financial crime. It runs a platform where financial institutions contribute pseudonymised transactional data in bulk to be analysed by regulators, agencies and the financial institutions themselves. The firm was started "*with a vision that money has identity*".

FutureFlow entered the United Kingdom's Information Commission Office (ICO) sandbox in mid-2019 and both parties worked together through the first year of the pandemic. The joint objectives of FutureFlow and ICO were to:

- Map the firm's data flow and clarify responsibilities at each point
- Develop a system to measure and manage the risk of re-identification of anonymised data
- Develop compliance documentation
- Set out measures for ongoing security and integrity of the data

The key regulatory questions were;

- Given the complexity of FutureFlow's processes, who was the data controller and the data processor, and when? (These legal roles determine responsibilities.)
  - FutureFlow achieved clarity on the question of which role it had.
- Could the pseudonymised customer financial data be considered anonymous?
  - The data was found to be possible to re-identify, and suitable security responses were developed to mitigate the risk of it being leaked.

FutureFlow left the sandbox in August 2020 when this work was complete.

Benefits:

FutureFlow, an innovative start-up responding to the need to reduce money-laundering, got regulatory clarity on key questions and significant overall insight into data protection.

The ICO got "*valuable insight into the financial sector and how banks and other financial institutions might choose to leverage and share data .. to detect ..financial crime. The ICO's work with FutureFlow will help to influence [their] views and any future work on how large organisations can anonymise, pseudonymise and share data for the purposes of tackling financial crime in a compliant and secure manner while maintaining individuals' rights to privacy.*"

Sources: ICO (2020), *Regulatory Sandbox Final Report: Future Flow,* Information Commissioner's Office.

One of the few, if not the only example, of a cross-border regulatory sandbox for data that has been started is that of the ASEAN-GSMA Regulatory Pilot Space for Cross-Border Data Flows (Box 6). Despite the lack of information concerning results and implementation status of the ASEAN-GSMA pilot, observations on its design could help inform a model for future cross-border regulatory sandboxes for data:

- All or most countries involved have privacy laws that align with the OECD and ASEAN data principles. However, the national laws vary considerably. Possibly, this example is of a region with sufficient overlap in the broad objectives of regulations but significant differences in their enactment, i.e. the 'sweet spot' necessary for a cross-border regulatory sandbox for data to work.

- The sandbox includes a mandatory accountability mechanism, i.e. a requirement for firms to show they have implemented binding safeguards to ensure their compliance. This suggests pre-binding mechanisms may be a useful or necessary component of such sandboxes to reduce risk and secure trust.

- While outcomes are not yet available, the ASEAN-GSMA Regulatory Pilot Space shows there is strong will amongst some regulators to actively explore the interoperability potential of national data governance regimes.

⬡ Box 6. ASEAN-GSMA Regulatory Pilot Space for Cross-Border Data Flows

ASEAN and the GSMA[28] came together in 2019 to put forward a proposal for the creation of a regulatory sandbox focused on the flow of cross border data amongst participating countries in the Asia-Pacific region. ASEAN and the GSMA believe that "the regulatory sandbox can be a stepping stone towards a formal mechanism for cross border data flows".

Two of ASEAN's strategic priorities are:

- ASEAN Data Classification Framework (initiative led by Indonesia)
- ASEAN Cross-Border Data Flows Mechanism (Singapore)

The pilot space (originally called a sandbox) supports the implementation of these priorities "*by creating a safe space in which policymakers can understand and test possible policy solutions to facilitate cross-border data flows.*"

The premise of the sandbox is to give the opportunity to different countries to address the various cybersecurity concerns associated with cross border data flows and, in doing so, innovation and the evolution of the digital society will not stall.

The sandbox is intended as a regulatory space shared by national regulators of the participating governments. It only involves firms who request to be allowed to send personal data across borders as applicants. Its goal is to allow countries with differing personal data laws to begin to exchange data, while ensuring their rules are followed.

The sandbox is intended as a stepping stone to broader data-flows between the participating countries; it is "*a bridging solution while ASEAN Member States develop their data privacy frameworks and interoperable mechanisms for cross-border data flows.*"

No implementation information is yet available on the pilot space.

Sources: GSMA (2019), *ASEAN Regulatory Pilot Space for Cross-Border Data Flows,* GSM Association and GSMA (2019), *Operationalizing the ASEAN Framework on Digital Data Governance A Regulatory Pilot Space for Cross-Border Data Flows,* GSM Association.

---

28. GSMA (2019), *Operationalizing the ASEAN Framework on Digital Data Governance A Regulatory Pilot Space for Cross-Border Data Flows,* GSM Association.

While cross-border regulatory sandboxes are still just in their infancy, they already have proven potential to bring regulators in different countries together to analyse specific technologies and initiatives, and to increase the flow of cross-border data to support innovation.

So far, most domestic regulatory sandboxes have not intended to include stakeholders beyond regulators and firms. While regulators stand statutorily in the place of representing the broad public interest, their varying goals may favour some stakeholders and interests over others. On the other hand, cross-border sandboxes could be proved to be most useful in assessing new technologies, services and business models whose testing requires the participation of different stakeholders - i.e. not just the regulator and sponsoring firm or organisation - or whose opportunities or impacts are cross-border and need to be assessed by two or more national regulators or accreditation bodies.

03

Identifying opportunity areas for cross-border regulatory sandboxes for data

Cross-border regulatory sandboxes for data have good potential, but the coordination challenges mean that few have been attempted. This section examines substantive areas of innovation with a view to explore what might be achieved by cross-border sandboxes in priority areas, first focusing on specific innovations whose implications they can help to tease out, and secondly, on sector-specific challenges they could help to address.

Innovations that sandboxes can support

This section considers four emerging approaches and technologies that rely on the free flow of data, often across borders. These cross-cutting applications have the potential to change how we use and even think about data. Sandboxes are a key way to enable practical dialogue and build consensus-based approaches to sustaining and improving the data-flows on which these innovations depend.

Emerging intermediaries

Different types of data intermediaries are gradually emerging, including data fiduciaries, trusts, cooperatives, commons and stewardship.[29] Data cooperatives are "*organisations that collect data from .. members, process and monetise the pooled data, and compensate the members for their individual contributions.*"[30] They establish an "*ecosystem of trust*"[31] among members, based on control of one's data, collective bargaining power and compensation for use of the data.

Related structures are 'data commons', exchange networks and cooperative initiatives that pool data held by multiple organisations to create a shared resource that serves, for instance, open data projects. A core element to many of these initiatives is the adoption of strong common-law based fiduciary duties of care and loyalty towards patrons.[32] Many aim to increase control over data and distribute the benefits from its management. Examples include Our Brain Bank,[33] a patient-led data collaboration for brain cancers, and DECODE,[34] a distributed European data commons that uses distributed ledgers. Others such as Deeper Edge LLC in the United States incorporate fiduciary duties into web-based service offerings they provide in the marketplace.[35] Some online companies have even promulgated terms of service that expressly adopt a digital fiduciary role with patrons.[36]

The World Economic Forum (WEF) has been exploring concrete ways that data intermediaries can advance digital agency for ordinary human beings.[37] A working group of leading experts recently released a report iden-

29. Hardings, J. (2018), *A data trust provides independent, fiduciary stewardship of data*, Open Data Institute Blog.

30. Mehta, S., M. Dawande and V. Mookerjee (2021), *Can data cooperatives sustain themselves?* The London School of Economics and Political Science Blog.

31. Mehta, S., M. Dawande and V. Mookerjee (2021), *Can data cooperatives sustain themselves?* The London School of Economics and Political Science Blog.

32. Whitt. R, (2020), *Old School Goes Online: Exploring Fiduciary Duties of Loyalty and Care in the Digital Platforms Era,* 36 Santa Clara High Tech. L.J. 75.

33. Our Brain Bank (n.d), *Homepage Website,* (acessed 28 March 2022).

34. DECODE (n.d), *Homepage Website,* (acessed 28 March 2022).

35. Deeper Edge (n.d), *Homepage Website,* (accessed 28 March 2022).

36. DataLucent (n.d), *Terms & Conditions Website,* (accessed 28 March 2022).

37. WEF (n.d), *Data Intermediaries: Trusted Digital Agency,* World Economic Forum Blog.

tifying how such digital agents can help people navigate complex data ecosystems.[38] An earlier WEF initiative in October 2021 led to the adoption of ethical data practices by the City of Helsinki on behalf of its citizens, aiming at mitigating some of the trade-offs between innovation and privacy.[39] Importantly, these new intermediary models are at work in both personal and non-personal data. In non-personal data, exchange networks and cooperatives pool data held by multiple organisations to create a shared resource, for example Driver's Seat Cooperative rideshare in the United States,[40] that collects mobility data from many drivers and sells it to cities, and Amazonia Socioambiental,[41] a Latin American initiative mapping threats to the Amazon.

There are many variations, but data cooperatives broadly aim to provide alternatives to maximal extraction and concentration of data and its exploitation by technology platforms and other organisations. This is done by creating intermediary organisations that sit between data-subjects and the firms and institutions that wish to use their data. These intermediary organisations typically let people set the terms of how their data is used, and share the financial benefits of its use with them.

This is a key opportunity area for sandboxes. Few alternative data governance initiatives have robust and publicly available decision-making processes and lines of accountability.[42] There is also a current gap between legal and theoretical frameworks and actual practice,[43] which opens up an opportunity for sandboxes to engender supportive and creative regulatory spaces in which new data models and organisations can be tested.

There is a strong public interest benefit in developing cross-border sandboxes to sustain a practical dialogue between regulators and new data organisations. While many data-subjects whose data is handled by traditional, data-extractive platforms may wish to restrict the flow and use of their data by third parties, people who have signed up for nontraditional data intermediaries may wish the opposite, because they have been able to set controls for how it is used. Arguably, the current personal data protection regulatory model is conceived of as a way to monitor and minimise data extraction and exploitation. New organisations and emerging models seeking to promote end-user empowerment via data do not fit this frame, however, and would need to be part of an ongoing dialogue and iterative process to mould themselves to the regulatory model and vice versa, as both evolve.

Consent may be a distinctive aspect of how emerging intermediaries interact with regulation on cross-border transfers. Regulators are increasingly challenging blanket consent opt-ins[44] and 'dark patterns'[45] used by many firms and platforms to say their data-subjects have agreed to the transfer and use of their data. However, emerging intermediaries can have positive *ex ante* consent of data subjects (or originators/owners, for non-personal data) for transfers. They risk being caught up in restrictions on abusive consent practices and may need models like sandboxes to resolve them. Additionally, their positive consent models may also have something constructive to offer to how regulators frame these issues going forward.

38. WEF (2022), *Advancing Digital Agency: The Power of Data Intermediaries*, World Economic Forum Blog.

39. Tedeneke, A. (2021), *First-of-Its-Kind Blueprint for Data Policy Adopted by City of Helsinki*, World Economic Forum Blog.

40. Driver's Seat (n.d), *Homepage Website*, (accessed 28 March 2022).

41. RAISG (n.d), *Homepage Website*, (accessed 29 March 2022).

42. Insights, M., S. Baack and M. Maxwell (2020), *Who is Innovating? | Global Landscape Scan and Analysis of Initiatives*, Mozilla Blog.

43. Insights, M., S. Baack and M. Maxwell (2020), *Who is Innovating? | Global Landscape Scan and Analysis of Initiatives*, Mozilla Blog.

44. Utz, C., M. Degeling, S. Fahl, F. Schaub and T. Holz (2019) *(Un)informed Consent: Studying GDPR Consent Notices in the Field*, ACM Digital Library.

45. Zhu, C. (2021), *Dark patterns — a new frontier in privacy regulation,* Reuters Blog.

The conceptual and institutional innovation of emerging intermediaries presents at least as many questions as the new technologies and services. These entities' goals and points of contact with regulatory regimes differ because of the agency and choice exercised *ex ante* by data-subjects, and the questions raised may be best answered in the focused environment of a sandbox.

Privacy Enhancing Technologies (PETs)

Privacy Enhancing Technologies (PETs) have been used or proposed since the late 1990s as ways for individuals and organisations to minimise the personal data collected and used.[46] While they have rarely fulfilled their promise - some have suggested this is largely because they tend to push the burden of responsibility back onto individuals[47] - recent developments in PETs have the potential to change this. These include homomorphic encryption and a related set of technologies, trusted execution environments, multiparty computation, differential privacy methods which effectively anonymize data, and federated analytics which allow the data to remain where it is, and bring the processor's capabilities to it.

These technologies have different practical applications and varying levels of privacy protection, but they can be used to change the traditional model of cross-border transfers which assumes data leaving one jurisdiction where it has certain protections and requirements, and entering another where these are lessened. The new generation of PETs can bring the processing to where the data currently is, insert 'noise' into data-sets that de-identify data while maintaining much of its utility, or even encrypt the data so the processor never actually handles the data.

Homomorphic encryption, for example, is an application of public key cryptography that allows work to be done on encrypted data without it being decrypted first. It can support cross-border, cloud processing of personal data that ensures subjects' privacy and the integrity of the data. It may support the use of personal data for predictive analytics in ways that do not disclose that data to third parties,[48] meaning transfers or access to data simply do not occur. Current regulatory frameworks have uncertain application in this field - how does processing of data occur when access or transfers itself have not? Homomorphic encryption enables services that do not easily fit current conceptions and definitions, but considerable work needs to be done to understand how it works in practice, particularly when the processing is done on data from another jurisdiction. This requires careful and focused examination and dialogue with innovators, something sandboxes can excel at providing.

A cross-border sandbox for PETs with one or more potential service-providers per different technology types and a range of affected participants could:

- Tease out applicability of definitions and terms, implementation questions and appropriate regulatory touch-points.

---

46. Goldberg. I., D. Wagner and E. Brewer (1997), *Privacy-enhancing technologies for the Internet,* University of California, Berkeley,

47. Stalder, F. (2002), *The Failure of Privacy Enhancing Technologies (PETs) and the Voiding of Privacy,* Vol 2/7, Journal of Sociological Research Online.

48. Bradley-Silverio Donato, J. and L. Porciuncula (2021), *Data governance is in desperate need of innovation, Zama and the Datasphere; a discussion about data governance,* Zama Medium.

- Dynamically facilitate capacity-building for regulators, service-providers and civil society to better understand this emerging technology and its broader implications.

- Establish practical guidelines, risk frameworks and parameters for broader future use of the technology, while giving regulatory safeguards and certainty to all.

Moreover, taking on inspiration from certifications processes that emerged from the open source[49] or B-corp[50] movements, sandboxes could provide a step towards interoperability and scalability of PETs. Sandboxes could be used to understand their commonalities and differences, exchange and advance on the good practices, including defining protective guardrails upon which greater innovation could be built, and on setting the foundation for standardisation and models for mutual recognition or cross-border certification, providing more certainty to innovative solutions.

Browser-based solutions for consent management

It is a widely acknowledged problem that pop-ups as a means of consent management are not fit for purpose.[51] A cross-border regulatory sandbox could be used to allow companies to work with regulators on potential international standards on the general approach and/or technological solutions to obtain meaningful consent from the user.

Cookie banners as a consent mechanism in web applications are poorly implemented and arguably fail to obtain meaningful, fully informed consent from users, something that is required by law in the United Kingdom and other international law.[52] From a user perspective, they can be a constant annoyance through repeated pop-ups and confusing permission tick-boxes. Many users report consent fatigue from repeatedly clicking on seemingly meaningless cookie pop-ups. As a result, they may click 'consent' simply to quickly access a website, rather than make an informed and willing decision.

There is a broad international consensus that consent pop-ups do not work as intended and are a problem, but no agreed solution or standard to fix the problem. The United Kingdom Information Commissioner's Office (ICO) has encouraged fellow G7 data protection and privacy authorities to work together to overhaul cookie consent pop-ups,[53] so as "to ensure people's privacy is more meaningfully protected and businesses can provide a better web browsing experience".[54]

The solution is not clear cut, and different actors may have preferences, for example, web browsers, software applications and device settings which allow people to set lasting privacy preferences of their choosing, rather than having to click pop-ups on each website.[55] While this approach is already technologically possible and compliant with data protection law, a cross-border sandbox could encourage technology firms and standards

49. Lerner, J. and J. Tirole (2005), "The Economics of Technology Sharing: Open Source and Beyond", *Journal of Economic Perspectives*, 19 (2): 99-120.

50. Poponi, S., A. Colantoni, S.R.S. Cividino, and E.M. Mosconi (2019), "The Stakeholders' Perspective within the B Corp Certification for a Circular Approach", *Sustainability* 11, no. 6: 1584.

51. Yahoo (n.d), Consent Website, (accessed 29 March 2022).

52. Noyb (2021), noyb aims to end "cookie banner terror" and issues more than 500 GDPR complaints, NOYB Blog.

53. ICO (2021), ICO to call on G7 countries to tackle cookie pop-ups challenge, Information Commissioner's Office Blog.

54. Santos, C., N. Bielova and C. Matte (2020), Are cookie banners indeed compliant with the law?, Technology and Regulation Blog.

55. Delli Santi, M. (2021), Eprivacy Regulation and Privacy Automation, Open Rights Group Blog.

organisations to work with regulators to further develop and roll out privacy-oriented solutions to this issue, based on common international standards.

Sectors and policy areas that sandboxes can support

This section of the report looks at a selection of case studies on challenges in cross-cutting policy areas and sectors where sandboxes might be applied to find collaborative solutions, and in ways that actively serve the public interest and address global challenges:

- Innovation, trade, competition and digital policies
- Health and its impact on inclusion
- Mobility and its effects on social and environmental sustainability
- Biodiversity and its value for research, science and development.

The next subsection, on data flows and innovation, trade, competition and digital policies, explores how some of the concerns driving policies like data localisation might be addressed within sandboxes, and in ways that particularly support small, new and innovative firms. It highlights, in particular, how sandboxes are already beginning to be built into multilateral digital economy agreements. The three additional case studies in this section analyse friction in data-flows in three different but globally critical areas; health, mobility and biodiversity genomics. The particular challenges in these areas are explored, and opportunities are identified where improved data-flows would support global goals such as inclusion, sustainability and development.

Innovation, trade, competition and digital policies

Innovation, trade, competition and broader digital policies are affected by data through numerous complex value chains and positive and negative externalities. First, cross border data flows comprise companies' communications with customers, suppliers, and their various affiliates. Second, data flows are both an integral part of global supply chains and logistic channels and an enabler of globally supplied communications services such as data transmission networks and online distribution platforms. These two features are what is commonly referred to as digitally enabled trade and are vital resources for both goods and services producing entities. Data transmission or "flow" enables other products or services, such as social media, e-books, computer software, audio-visual products, or accountancy, architectural or medical services that are delivered directly or through intermediary platforms, to business clients or end consumers. Emerging technologies will foster an ever-increasing role for data flows. ICT technologies such as cloud-enabled networks and servers, Internet of things (IoT) communication links, high-bandwidth fifth generation (5G) mobile networks, high-capacity data storage and processing (big data), distributed networks and the algorithms that make AI possible will soon be widely adopted and stimulate a further surge in global data flow.

Knowledge has always been valuable. In the global Datasphere,[56] information, encoded as data, has become a resource that facilitates trade in knowledge more effectively than ever before. The challenge for the global economy is to harness this value of data responsibly, avoiding abuses and managing risks. Challenges related

---

56. Chapelle, B. and L. Porciuncula (2022), *Hello Datasphere — Towards a Systems Approach to Data Governance,* The Datasphere Initiative.

to social well-being, competition concerns and curbing illicit online activities (e.g. fraud, money laundering, data hacking, ransomware and critical infrastructure attacks) are daunting. Negotiators and experts take it for granted that these are extremely important problems for governments to resolve, whose implications go well beyond trade interests. As such, much of the discussion on data flows has centred on how to balance commercial interests and privacy objectives and, more recently have resurfaced historic concerns around economic imbalances and of claims for a more inclusive digital economy.

As governments grapple with these challenges, the past decade has witnessed increasing resort to various forms of data localisation, whereby data must be stored or processed locally, and cross-border transmission can be subject to conditions or even blocked entirely.[57] The appeal of such measures is their simplicity, but a problem with such measures is that they may be far too blunt an instrument that overshoots the mark, causing consequences beyond the intended objectives. Businesses large and small have pointed out that data flow restrictions hinder the flow of data that not only has become integrated into business models employed around the world, but also stymies the data flows that enable trade in goods and services and that convey digital services. For digital products, data flow restrictions would hinder trade all-

> **For digital products, data flow restrictions would hinder trade all-together, as much as a border measure would impede the import of traditional goods. These measures have emerged not perhaps as the best solution, but more as a step that is easier to take and more rapidly implemented.**

together, as much as a border measure would impede the import of traditional goods. These measures have emerged not perhaps as the best solution, but more as a step that is easier to take and more rapidly implemented. The concern is that such a blunt instrument may reduce the benefits of globally accessible information for business, trade and societies, yet without effectively safeguarding citizens, consumers or individual privacy.

Unjustifiable data flow restrictions limit access to lower-cost services and more innovative markets, thus inhibiting business model innovation and adversely impact other economic, social, scientific opportunities that the free flow of data can generate. While large companies would rather not assume the added costs of data restrictions, particularly when they are more burdensome than necessary to meet their goals, the costs are ultimately passed on to their business clients and consumers. Yet, start-ups and SMEs are much less able to withstand added costs, which range from higher prices for data storage and processing services for the business software they need, to the added regulatory burden of implementing or verifying their compliance with measures, often across multiple jurisdictions. Other cost-saving, efficient and more secure business options for SMEs, such as cloud-enabled services and distributed networks, may not be feasible. In fact, if restrictions on data flows become near-universal, the technologies that underpin these services means that they may, at worst, be virtually impossible to trade globally and, at best, offer reduced benefits to the consumers.

---

57. Cory, N. and L. Dascoli (2021), *How Barriers to Cross-Border Data Flows are Spreading Globally, What they Cost, and How to Address Them,* Information Technology and Innovation Foundation Blog.

Cross-border regulatory sandboxes can support innovative projects involving cross-border data flows and could be used to cover a range of difficult and complex issues in a mutually supportive environment.

Provisions calling on the use of sandboxes are included in a few existing trade agreements, such as the Canada-United States-Mexico (CUSMA) trade agreement,[58] which refers to cross-border sandboxes for financial services, and the Digital Economy Partnership Agreement (DEPA), which includes a provision for sandboxes for data (Box 7). As with other examples of trade agreements,[59] both CUSMA and DEPA discourage data localisation while recognising that there should be leeway for rights-protecting national regimes whose mechanisms are proportionate to their objectives. Most importantly, the fact that regulatory sandboxes are already recognised within some recent trade agreements is encouraging.

⬡ **Box 7.** Digital Economy Partnership Agreement (DEPA)

The Digital Economy Partnership Agreement (DEPA) signed in June 2020 between Singapore, Chile and New Zealand is an example that explicitly proposes regulatory sandboxes as a pathway to address cross-border data issues. DEPA's article 9.4 on data innovation expressly recognises regulatory sandboxes as potentially beneficial in producing innovative solutions for seeking a better-calibrated policy balance for measures governing cross-border data flows:

1. *"The Parties recognise that cross-border data flows and data sharing enable data-driven innovation. The Parties further recognise that innovation may be enhanced within the context of regulatory data sandboxes where data, including personal information, is shared amongst businesses in accordance with the Parties' respective laws and regulations.*

2. *The Parties also recognise that data sharing mechanisms, such as trusted data sharing frameworks and open licensing agreements, facilitate data sharing and promote its use in the digital environment to: (a) promote innovation and creativity; (b) facilitate the diffusion of information, knowledge, technology, culture and the arts; and (c) foster competition and open and efficient markets.*

3. *The Parties shall endeavour to collaborate on data-sharing projects and mechanisms, and proof of concepts for new uses of data, including data sandboxes, to promote data-driven innovation."*

 DEPA is the first of its kind to include a provision on a dedicated sandbox for data. The objective of this provision was to allow a space for addressing emerging issues in a collaborative manner, rather than committing to binding rules. While such a sandbox has not yet been realised in the case of DEPA, other ongoing projects related to connectivity and e-invoicing are being implemented between the parties. Finally, the inclusion from such a provision in the trade agreement expresses a positive mandate for an agile approach for international regulatory collaboration which could be crucial to encouraging regulatory experimentation and innovation.

Source: MTI (n.d), *The Digital Economy Partnership Agreement (DEPA),* Ministry of Trade and Industry Singapore Blog.

---

58. GOC (n.d), *Canada-United States-Mexico Agreement (CUSMA) - Chapter 17 - Financial Services,* Government of Canada Blog.

59. DFAT (n.d), *CPTPP outcomes: Trade in the digital age,* Australian Government Blog.

One greatest challenge facing negotiators seeking to include data flows policies and practices in their agreements is that examples of what constitutes regulatory best practices in emerging data flows issues are not yet evident.[60] Discovering and gathering evidence on what works best (or may not work as intended), allowing for adjustments to policy, is one of the most compelling reasons why regulatory sandboxes for exploring data policies will be useful for all stakeholders.

Box 8. Cross-border Regulatory sandboxes within trade agreements

In the trade agreements that deal with cross-border data flows, permission for experimentation is crucial. Data localisation requirements have hindered the free flow of data - frequently while failing to create meaningful citizen protections - and they are spreading.[61] Both the CUSMA and DEPA pave the way for sandboxes to be one way to rethink legal and technology approaches as alternatives to data localisation measures, while giving space and support to proportionate rights protections and risk management.

The flexibility and experimental nature of regulatory sandboxes, which have proven successful in the context of fintech, could be transferable qualities for data flows discussions. Including regulatory sandboxes in trade agreements could be a gateway to ensure further regulatory collaboration as part of ongoing bilateral or regional trade agreements. Of particular concern related to trade agreements at the national level are the competing interpretations of international obligations as captured in free trade agreements and national data protection law. This in itself could be a topic for dedicated sandboxes themselves. Moreover, sandboxes may also unlock investment and level the playing field between established and emerging players, such as small and medium enterprises.

A potential drawback of using cross-border regulatory sandboxes to clarify or implement solutions to trade-related issues could be that this results in its own patchwork of data provisions agreed bilaterally, hampering future multilateral agreements. Moreover, even where not explicitly mentioned, the use of sandboxes for regulatory experimentation are certainly not precluded by other trade agreements or the World Trade Organization framework of rules, so long as the measures comply with principles such as transparency and non-discrimination and are calibrated so their effect is proportional to the policy objective, rather than being unduly burdensome to trade.

Health

Health data is all data "related to health conditions, reproductive outcomes, causes of death, and quality of life".[62] More health-related data is generated today than ever before. From wearables and apps[63] to the explosion of research and public health interventions since the beginning of the pandemic, an unprecedented amount of health data - both personal and non-personal - is now generated, collected and transmitted.

60. OECD (2022), *Recommendations and Guidelines on Regulatory Policy, OECD website* (accessed on 29 March 2022).

61. Cory, N. and L. Dascoli (2021). *How Barriers to Cross-Border Data Flows are Spreading Globally, What they Cost, and How to Address Them*, Information Technology and Innovation Foundation Blog.

62. McGraw-Hill Concise Dictionary of Modern Medicine (2002), *Definition of Health,* The McGraw-Hill Companies, Inc. Blog.

63. NRC (2020), *The Internet of Bodies Will Change Everything, for Better or Worse,* Rand Corporation Blog.

This brings great public benefits. The first COVID DNA sequences were globally shared, using standard data access and sharing agreements, through the GISAID data network developed after the 2005 H5N1 avian influenza. This greatly accelerated work on COVID vaccine development.[64] Pooling and sharing health data can accelerate research and innovation and support better decision-making in public health. Conversely, growing health inequality can be driven by the lack of data or access to it, or its inappropriate use.

Pooling and sharing personal health data can create collective benefits, but impose costs and risk on individuals if they are identifiable from the data.[65] Privacy is essential - the release of personal data regarding reproductive health has raised concerns that it could even lead to human rights abuses.[66] These are not abstract concerns to the people whose data are pooled. The same surveys that show support for health information sharing often simultaneously express serious privacy concerns,[67] creating challenges for public trust in data sharing initiatives. The drive by public health systems and large insurers to use health data to train and develop AI can lead to conflict with data protection laws and confuse the public.[68]

Nonetheless, broad support exists in governments and international organisations[69] for the sharing of knowledge, intellectual property and data to "advance scientific discovery, technology development and broad sharing of the benefits of scientific advancement".[70] Generating data of sufficient scale and quality is so expensive that it is often the province of governments. From the United Kingdom's Biobank[71] to the United States' AllofUs Research Program[72] to China's 100,000 person genome project[73] to the African Genome Variation Project,[74] national and regional efforts hold promise to create breakthrough scientific knowledge.

Previous initiatives identified the need for scientific and medical data to be findable, accessible, interoperable and reusable,[75] as has been the case in other fields mentioned in the case studies in this report. But the potential benefits of health-related research and development as a global public good[76] must be articulated in frameworks that both facilitate effective sharing and mitigate the known and tangible risks of that sharing.[77]

Today, policy frameworks or detailed operational guidelines have emerged for sharing some types of health data across borders; genomes, Alzheimer's brain images, and more.[78] While there have been recent efforts to foster

64. GISAID (n.d), *About Us Website,* (accessed 30 March 2022).

65. Hansson, M., H. Lochmüller and O. Riess *et al.* (2016), *The risk of re-identification versus the need to identify individuals in rare disease research,* Eur J Hum Genet 24, 1553–1558 , Nature Blog.

66. MGuardian (2022), *El Salvador woman punished under strict abortion law freed after 10 years,* The Guardian Blog. TE (2021), *Texas's bounty-hunting abortion law could remain on the books for a long time,* The Economist Blog.

67. MPEW (2021), *Most Americans Want to Share and Access More Digital Health Data,* The Pew Charitable Trust Blog.

68. BBC (2021), *DeepMind faces legal action over NHS data use,* BBC News Service Blog.

69. Gall, C. and E. Suzuki (n.d), *Big Data: A new Dawn for Public Health?,* OECD iLibrary Blog.

70. Gall, C. and E. Suzuki (n.d), *Big Data: A new Dawn for Public Health?,* OECD iLibrary Blog.

71. Biobank UK (n.d), *Homepage Website,* (accessed 30 March 2022).

72. National Institute of Health (n.d), *Homepage Website,* (accessed 30 March 2022).

73. Su, Y. (2017), *China Genome Project: China launches 100,000 Genomes Project*, CGTN Beijing Blog.

74. Sanger (n.d), *Collaboration Website,* (accessed 30 March 2022).

75. Wilkinson, MD., M. Dumontier and IJ. Aalbersberg et.al, (2016), *The FAIR Guiding Principles for scientific data management and stewardship*, National Library of Medicine Blog.

76. Smith, R.D. and L. MacKellar, (2007), *Global public goods and the global health agenda: problems, priorities and potential*, Global Health 3, 9, Globalisation and Health Blog.

77. Malin, B. and K. Goodman & Section Editors for the IMIA Yearbook Special Section (2018), *Between Access and Privacy: Challenges in Sharing Health Data*, 27(1), 55–59, *Yearbook of medical informatics*, National Library of Medicine Blog.

78. Global Alliance for Genomics and Health (n.d), *Homepage Website,* (accessed 30 March 2022). The Alzheimer's Disease Neuroimaging Initiative (n.d), *Homepage Website,* (accessed 30 March 2022).

more systematic sharing of health data, through processes launched by the WHO[79] and the G7[80] to respond to the pandemic, as well as initiatives such as European Health Data Space[81], broadly speaking, health data sharing is still made more complex by a patchwork of international agreements - multilateral[82] and institutional - and national laws, both related to general privacy, confidentiality, specific health privacy laws, ethics and consent. There is also an even greater diversity of legal instruments and norms, including contracts and consent processes that implement those laws.

Health data is also not uniform. Viral genome data is different to personal medical records or data about the social determinants of health such as income, race, gender, or locality. Thus, successful data-sharing regimes most easily form within scientific disciplines, with each kind of information requiring different systems, metadata and analytics, and often carrying different legal obligations.

There is a significant opportunity for integration across these data-sharing regimes, especially, for example, in constructing long-term, observational studies connecting regular clinical care data to detailed research data.[83] Ability to track real world evidence of the performance of medicines (e.g. from larger and more diverse data-sets)[84] could let data science stratify the sub-populations for whom medicines do and don't work. Data-sharing between nations could also address representational inequities in reference to data and promote inclusion and diversity in data-sets.[85]

> **Health data is also not uniform. Viral genome data is different to personal medical records or data about the social determinants of health such as income, race, gender, or locality. Thus, successful data-sharing regimes most easily form within scientific disciplines, with each kind of information requiring different systems, metadata and analytics, and often carrying different legal obligations.**

But that integration will often come into contact with multiple levels of law, from local confidentiality requirements at the state level, to national laws about research ethics, to international regimes about data protection. Integration also faces divergent interests and concerns about competition, compliance or unfair exploitation. This can

---

79. WHO (2021), *World Health Assembly agrees to launch process to develop historic global accord on pandemic prevention, preparedness and response*, World Health Organisation Blog.

80. The Global Pandemic Data Alliance (n.d), *Homepage Website*, (accessed 30 March 2022).

81. EC (n.d), *European Health Data Space*, European Commission Blog.

82. WHO (n.d), *International Health Regulations*, World Health Organisation Blog. Centre for Disease Control and Prevention (n.d), *International Health Regulations Website,* (accessed 30 March 2022). Venkatasubramanian A (2020), *The Case for Health Data Regulation*, The Graduate Institute Geneva Blog.

83. Mowry Ellen M., A. Bermel Robert and W. James R et al (2020), *Harnessing Real-World Data to Inform Decision-Making: Multiple Sclerosis Partners Advancing Technology and Health Solutions (MS PATHS)*, Frontiers Blog.

84. FDA (2021), *Considerations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products*, US. Food and Drug Administration Blog.

85. Gurven, M. D. and E. Lieberman (2020), *WEIRD bodies: mismatch, medicine and missing diversity. Evolution and human behavior, Official journal of the Human Behavior and Evolution Society*, *41*(5), 330–340.

lead to the withholding of even basic information about publicly funded data such as clinical trials, requiring enforcement of open data mandates. Cross-border regulatory sandboxes may be helpful at addressing some of the issues related to the sharing of data for research and science across borders (Box 9).

⬡ Box 9. Cross-border sandboxes for health data

Cross-border regulatory sandboxes could be implemented when clarification is needed on how international medical research interacts with data protection laws and other regulations across jurisdictions. For example, data protection laws may entitle a data subject to request all copies of data about that individual be deleted. However, if their data has been packaged and widely shared for scientific use, it can be complex to fulfil that request (or impossible, if those data packages have been broadly downloaded). Other ambiguities exist around data protection laws, general duty of confidentiality, laws around genetic data, and national research law. The classification of genetic data as sensitive, for example, can interact in unexpected ways with specific national laws around research protections and new uses of data for medical products. Existing data protection rules and different requirements around sharing may make this difficult, particularly for multi-country research projects. A cross-border regulatory sandbox may be useful in navigating some of these challenges.

Such a regulatory sandbox might also benefit from a tight coupling to an operational sandbox that would leverage models of trusted research environment (TRE) or other privacy enhancing technologies (PETs). A TRE would allow for trusted users to "visit" the data in an enclave where analysis is facilitated. Findings are exportable rather than the data itself. Deploying a core TRE could reduce transaction costs and allow broad, diverse user-base access to shared health data. Similarly, systems which use PETs, such as OpenSafely[86], may support the analysis of datasets without disclosing personal information.

The combination of sandboxes with either TREs or PETs could be used to enable rapid, targeted and evidence-based progress in health related data-sharing, helping to inform the development of any cross-jurisdictional policy framework or detailed operational guidelines for this activity. TREs inside a regulatory sandbox could potentially help to address power imbalances in the ability to compute data, subsidising data processing and storage for data scientists in data-sharing countries. TREs could limit the risks of allowing operation in a permissive regulatory sandbox with visible and versioned tracking of how data scientists use the data, and enable compliance with regulatory terms, citation of key peers, and return of value to patients whose data drives discovery.

Mobility

Mobility data is data related to the movement of people and things. It has expanded enormously as technology makes it easier to generate and use. Much human mobility data is generated by GPS receivers embedded in devices such as mobile phones, which use GPS and cellular networks to track users and provide location-based services. Companies such as mobile phone networks, social media platforms, application providers and

---

86. Open Safely (n.d), *Homepage Website,* (accessed 30 March 2022).

operating system owners collect location data and use it to plan and provide services, target search results and advertising, deliver real-time mapping and traffic information, and monetise it, including through licensing or transferring it to third parties. These mobility data-sets and their derivatives range from household level consumer habit analysis, origin-destination trip databases that estimate commuter habits and commercial good flows, or dashboards that track transponders of logistics fleets in near-real-time.

Mobility data of individuals is treated as personal data by many ethical and legal frameworks; at a granular level, it can be used to identify people's habits, workplaces and homes.[87] The limits on uses required for human mobility data can encourage caution in organisations considering sharing or licensing this data. Principles of data federation, affirmative consent, anonymisation and aggregation, and computational trust assurance can help to safeguard the data. While formal governance systems often exist for bilateral data-sharing between organisations in different countries – for example, within the European Union – consistent international data-sharing frameworks for human mobility data have yet to emerge.

Mobility data is not only about people, but about the movement of things. This data supports planning and operations in logistics and the delivery of goods, and underpins the entire supply chain. Mobility data is crucial across the private and public sectors. It is used to support sophisticated investments in commercial and domestic real estate, through the analysis of patterns of activity and land use. Policymakers and planners use it to do short and long range infrastructure planning, gauge the need for public services and optimise where they are delivered to improve transport safety, and much more. Logistics firms use mobility data for a wide range of essential activities, from route-planning to compliance with regulations about working hours for hauliers, to extending the lifespan of vehicles and other assets through proactive maintenance.

Mobility data is the basis for transport policy and planning, both within and between countries. For planners, mobility data is central to understanding and alleviating bottlenecks in complex adaptive traffic systems, be they a seaport, airport, railyard, motorways or surface streets. Traffic data also underlies decisions on improvements to hubs such as border crossings, ports and airports. Transnational highway and rail networks are built based on analysis of transportation patterns and how they may be developed in the future.

As climate change deepens, there is an acute need to integrate and expand the use of mass transportation that bridges different providers and countries and also enhancing the transparency and consistency of emissions data. This requires sharing reports and aggregated data collected by national-level actors, including communication providers who collect location data. It also requires better integration of mobility data collected by multiple actors in the transit and transport space. But while significant improvements are being made, year on year, much wider adoption of mobility data frameworks and networks is needed.

Infrastructure planning is increasingly driven by commitments to a sustainable climate future. Road-building and use are carbon-intensive activities. In addition to emissions of heavy machinery and temporarily increased congestion during the construction process, the whole supply chain of components for civil engineering generates significant carbon and other greenhouse gases emissions. Better infrastructure planning, based on good mobility data, is essential for building and maintaining alternative, less carbon-intensive forms of

---

87. De Montjoye, Y.A., C. Hidalgo and M. Verleysen *et al* (2013), *Unique in the Crowd: The privacy bounds of human mobility,* Sci Rep 3, 1376, Nature Blog.

transport. Data is also central to policy making that incentivises or subsidises less carbon-intensive travel choices by individuals and groups. It supports better understanding of and support for telework and remote work that complements local economic development while improving environmental factors. The vast and interlocking systems of transit, transport, planning and environmental change require sophisticated analyses and complex optimisations that depend on the rapid, real-time flow of data across integrated systems.

However, while the uses and benefits of mobility data are many, the ability to transfer it across borders to support integrated transport systems and decision-making is constrained by both technical and legal barriers. Technical challenges in the cross-border integration and comparative analysis of mobility data are partly due to differences in local or national mobility habits and infrastructure. For example, a mobility mode choice algorithm attempts to estimate what conveyance a traveller is using, such as car, bicycle, or bus, from the rate and speed at which they are travelling across a road network. Someone whose data showed them travelling at a high rate of speed on a surface street might be assumed to be in a car in the United States, but may be more likely to be using a bicycle in Denmark and a motorised scooter in south east Asia. Algorithms and analytics developed in one region need to be recalibrated to be put into a local context.

Also, significant technical challenges to data, meta-data and systems interoperability include errors and inconsistencies in the identification and labelling of items like lanes, road segments or intersections. Standardisation work in labelling and geocoding – building a common, authoritative way to communicate locations – can run into resistance from proprietary systems used by navigation and location-based service providers.

On mobility data, one response to these technical (and to some extent competition), is the development of systems like that proposed by SharedStreets, a "global, non-proprietary system for describing streets."[88] It uses open source software to reconcile road network data between different databases. Another approach to the issue of discrepancies between different systems uses open source remote sensing data such as open access satellite data[89] to develop authoritative digital identifiers to enable data exchange.

Another way to smooth the flow of mobility data is the establishment of regional mobility data exchanges that allow transport operators, planners, policymakers, service providers and others to share and access mobility data. In Europe, the Mobility Data Space was kickstarted by the German government in 2019 and is a platform for sharing and accessing mobility data "*from major vehicle manufacturers to ride-share services, from public transport operators to navigation software companies, from research institutes to bike-sharing companies*".[90] This initiative is still in its early days, so analysis of its ability to overcome the identified technical issues is premature.

Similarly, for greenhouse gas emissions data, data sharing across stakeholders is hindered by the complexity of collecting data across value chains and a lack of interoperability between technology solutions. Moreover, existing emissions standards and methods leave significant room for interpretation and result in data inconsistency. A solution being designed by a partnership of stakeholders is developing a data exchange network enabling

88. SharedSheets (n.d), *Homepage Website*, (accessed 29 March 2022).

89. Zeeshan, A. (2020), *Building Footprints and AI Do you know your buildings well enough to win?* Attentive AI Tech Blog.

90. Mobility Data Space (n.d), *Homepage Website*, (accessed 29 March 2022).

businesses to share their GHG emissions data in a verified and confidential manner across industries and value chains.[91] Cross-border sandboxes could test some of the emerging solutions in both mobility and related data (such as greenhouse gas emissions) and advance their use if their benefits are proved scalable (Box 10).

⬡ Box 10. Cross-border sandboxes for mobility data

For most countries, emerging ethical, legal, regulatory, and technical challenges create barriers to the cross-border sharing of mobility data. Organisations which make presumptions about pooling and sharing of data can create governance risk unless internal audit systems or other mechanisms are in place to ensure compliance. Mobility data as a whole includes both personal and significant amounts of non-personal data, meaning it is not always clear which type of legal regime may apply. Another issue which may create uncertainty is the origin, provenance and accuracy of mobility data; it ranges from personally identifiable location data all the way to machine-generated and targeted data about vehicles, locations, aggregated traffic, and so on.

Regulatory sandboxes could reduce the friction caused by regulatory uncertainty about transfers between countries that do not have bilateral frameworks or equivalent regimes. Sandboxes could point to the most useful emerging technical standardisation initiatives; unpack when and if personal and non-personal data within mobility data-sets need to be treated in different ways; and – potentially - fast-track the cross-border transfer and integration of data whose use serves advancement in critical areas.

There is also significant opportunity for operational sandboxes to be employed to work on labelling, identification and treatment of personal and non-personal data, and the combination of multiple data-sets for testing and use by multiple actors.

Biodiversity Genomics

Biodiversity genomics is the study of biological variation in all its dimensions, from the foundational DNA layer to organisms and ecosystems, phylogeny and function.[92] Biodiversity work has always relied heavily on genetic resources to generate scientific findings that contribute to conservation and the bio-economy. Today, the huge reduction in the cost of genome sequencing, along with the use of algorithms, have put genomics at the heart of biodiversity research. Once an impossibly ambitious goal, providing a reference genome for all of earth's species is now an achievable mission.[93]

---

91. Wbcsd (2021), *Carbon Transparency Partnership publishes new guidance to enhance consistency of emissions data*, World Business Council for Sustainable Development Blog.

92. Droge, G., K. Barker and J. Deck (n.d), *Genomic Biodiversity*, TDWG Blog.

93. Lewin, H. A., E. Robinson and J. Kress (2018), *Earth BioGenome Project: Sequencing life for the future of life*, PNAS Blog.

The public and private sectors have invested substantially in local, regional and international biodiversity genomics initiatives, including Genomics Aotearoa,[94] GROW Columbia,[95] Bat1K,[96] Vertebrate Genome Project,[97] European Reference Genome Atlas (ERGA),[98] Africa BioGenome Project[99] and the Earth BioGenome Project.[100] The information generated has the potential to be a global public good, an encyclopaedia of knowledge that forms the basis for scientific discoveries that can be applied to disease outbreak preparedness, bio-economy stimulation and the protection and conservation of biodiversity. To effectively preserve and manage biodiversity sequencing information, the data needs to flow in a streamlined and standardised way.

Biodiversity genomics research typically follows the openness norms and agreements in the field of human genomics. The Human Genome Project drove the rapid release of prepublication data as the 'common heritage of mankind', through the Bermuda principles[101] which support efficient and collaborative sequencing. The genomics community applies the FAIR data principles[102] that data be findable, accessible, interoperable and reusable - through public repositories including the International Nucleotide Sequence Database Collaboration[103] (INSDC) repositories which is composed of the National Centre for Biotechnology Information,[104] European Nucleotide Archive[105] and DNA Database Of Japan.[106]

Frictionless access is a cornerstone of genomic data use. INSDC policy[107] calls for unrestricted, licence-free use of sequencing information. In 2019, its repositories contained the sequencing information of over 300,000 species, consisting of quadrillions of sequences which were partly or completely downloaded 34 million times per year by over ten million unique users. As sequencing costs continue to fall, the INSDC continues to grow exponentially with the number of sequences doubling every eighteen months. The value of biodiversity genomics data depends heavily on being able to make large-scale comparisons between sequences, so unrestricted use and access by scientists and researchers is fundamental to maximising the scientific value of data. INSDC data is transformed, parsed, and exchanged with over 1,600 downstream databases, each creating additional value that depends on the flow and availability of that data.

However, some countries prefer to establish national data repositories for all genomic data collected within their jurisdiction, for example, the China National Genebank.[108] Indigenous Peoples and Local Communities

94. Genomics Aotearoa (n.d), *Homepage Website* (accessed 29 March 2022).

95. Grow Colombia (n.d), *Homepage Website* (accessed 29 March 2022).

96. Bat1K (n.d), *Homepage Website* (accessed 29 March 2022).

97. The Vertebrate Genomes Project (n.d), *Homepage Website* (accessed 29 March 2022).

98. The European Reference Genome Atlas (n.d), *Homepage Website* (accessed 29 March 2022).

99. African BioGenome Project (n.d), *Homepage Website* (accessed 29 March 2022)

100. Earth Biogenome Project (n.d), *Homepage Website* (accessed 29 March 2022).

101. Jones K, M., R. Ankeny and R. Cook-Deegan (2018), *The Bermuda Triangle: The Pragmatics, Policies, and Principles for Data Sharing in the History of the Human Genome Project*, J Hist Biol 51, 693–805, Springer.

102. GoFair (n.d), *Fair Principles*, GoFair Blog.

103. International Nucleotide Sequence Database Collaboration (n.d), *About INSDC Website* (accessed 29 March 2022).

104. National Centre for Biotechnology Information (n.d), *Homepage Website* (accessed 29 March 2022).

105. EMBL-EBI (n.d), *Homepage Website* (accessed 29 March 2022).

106. DDBJ (n.d), *Homepage Website* (accessed 29 March 2022).

107. International Nucleotide Sequence Database Collaboration Policy (n.d), *Policy Website* (accessed 29 March 2022).

108. China National GeneBank (n.d), *Homepage Website* (accessed 29 March 2022).

(IPLCs) are also creating alternative data-sharing solutions based on the CARE Principles for data sharing.[109] These principles prioritise the biocultural value and intergenerational connection to biodiversity and its associated data, over what is viewed as the westernised technical and procedural value of biodiversity data, promoted by the INSDC. For example, the Genomics Aotearoa data repository[110] in New Zealand was designed to respect indigenous data sovereignty.[111] It stores the data of taonga - culturally salient species - with access permissions governed by Māori partners.

Another example of how the treatment of genomic data is being adapted is the European Reference Genome Atlas (ERGA). The ERGA is a consortium funded through the EU programme, Horizon Europe 2020, to sequence all life across Europe. While most of its data will be available through the INSDC, to respect the choices of both Sami and Inuit Peoples, ERGA has partnered with the Local Context Hub.[112] This will digitally associate all samples and associated sequencing data obtained from Indigenous or local community lands, or that contains traditional knowledge, with a Biocultural[113] or Traditional Knowledge[114] Label. These labels empower IPLCs to assign permissions, protocols and provenance to their species samples and data. Adaptations like this one show how genomic data can be governed collaboratively and respectfully to the people and places it comes from, and ensure its benefits revert to them, while also enriching a broader public good.

The international regulatory landscape that biodiversity genomic research fits inside is the 2003 Convention on Biological Diversity (CBD).[115] The CBD has three major objectives; conservation of biological diversity, the sustainable use of its components, and fair and equitable sharing of the benefits from the use of genetic resources. It requires data-users to have prior informed consent and mutually agreed terms that ensure benefit-sharing. The CBD is now working on how to ensure the benefits from sequencing can be distributed through a multilateral system. Its technical discussions are focused on how to ensure the free flow of data while enabling the traceability needed to support benefit-sharing.

Within the existing infrastructure, some of the main obstacles to data use in biodiversity genomics have been identified by the INSDC as stemming from:

- **Inadequate metadata** – just 16% of biodiversity genome sequences held by INSDC have country of origin information, and only 6% link directly to their associated genetic resources.[116] INSDC is strengthening its metadata requirements. This will increase traceability for benefit-sharing.

- **Permanent identifiers** – the permanent identifiers associated with data-sets that are generated on publication in scientific journals do not link back to access and benefit-sharing numbers. This impedes tracing and monitoring.

---

109. CARE Principles for Indigenous Data Governance (n.d), *Homepage Website* (accessed 29 March 2022).

110. The Aotearoa Genomic Data Repository (n.d), *Homepage Website* (accessed 29 March 2022).

111. United States Indigenous Data Sovereignty Network (n.d), *Homepage Website* (accessed 29 March 2022).

112. Local Contexts (n.d), *Homepage Website* (accessed 29 March 2022).

113. Local Contexts (n.d), *Labels Website* (accessed 29 March 2022).

114. Idem.

115. UN (1997), *Convention on Biological Diversity,* United Nations.

116. CBD (2020), *Combined Study on Digital Sequence Information in Public and Private Databases and Traceability,* WHO and UNEP.

While work is ongoing, tracking and monitoring technologies across the whole life cycle of genetic research and genome sequences are not yet fully developed and implemented. This impedes access and benefit-sharing, constraining the data's use as a public good. This may be due to low demand or the complexity of navigating the regime of bilateral or multilateral benefit-sharing mechanisms. Other factors include:

- **Lack of implementation resources** – inadequate or absent procedural implementation impedes ethical and legal sample collection within countries, with knock-on effects on international collaboration and scientific discovery.

- **High transaction costs** – the cost of permits constrain both national research and international collaborations.

- **Inconsistent implementation** – the domestic implementation of access and benefit-sharing regimes varies (from restrictive to enabling), can be influenced by national agendas that do not prioritise science, and are challenging for researchers with few legal resources. The amount of time it takes to navigate these systems can also be incompatible with the short-term manner in which research grants are funded.

- **Concerns of Indigenous Peoples and Local Communities (ILPCs)** – In countries where the rights of ILPCs have not been recognised, these people may not receive benefits from their genetic resources and traditional knowledge. Further, some ILPCs had profound concerns about how the CBD and its protocols were negotiated.

- **Commercial versus non-commercial use** – Implementing protocols do not codify distinct processes for commercial and non-commercial access and use. This has hindered the flow of biodiversity data solely for non-commercial conservation.

These issues will influence if and how the CBD is expanded to include genomic sequencing data. A multilateral solution to address these challenges has been favoured amongst the non-commercial biodiversity research community, despite the lack of clarity on how benefits would be fairly and equitably shared. Options discussed include monetary mechanisms such as micro-levies on sequencing reagents and disposables, user fees for bio-based products, or for high income countries to pay into an international biodiversity fund that would be distributed to fulfil the objectives of the CBD.[117] Regulatory sandboxes could help test those solutions and clarify a roadmap for future multilateral solutions (Box 11).

> **While work is ongoing, tracking and monitoring technologies across the whole life cycle of genetic research and genome sequences are not yet fully developed and implemented. This impedes access and benefit-sharing, constraining the data's use as a public good. This may be due to low demand or the complexity of navigating the regime of bilateral or multilateral benefit-sharing mechanisms.**

---

117. Scholz, A.H., J. Freitag and C.H.C Lyal *et al.* (2022), *Multilateral benefit-sharing from digital sequence information will support both science and biodiversity conservation, Nat Commun* 13, 1086, Nature Blog.

⬡ **Box 11. Regulatory sandboxes for biodiversity genomics data**

Cross-border regulatory sandboxes could be useful in a context, such as that of biodiversity genomics, where many parties and interests need to be reconciled in an iterative process to equitably and efficiently ensure the free flow of genomic data for the scientific progress and benefit of future generations. Regulatory sandboxes for data are a possible way to work collaboratively across jurisdictions on resolving specified and discrete issues of the broader biodiversity genomics regime to advance the following objectives:

- **Mainstreaming FAIR data-sharing** – with richer metadata for better interoperability and benefit-sharing.
- **Scalable comparative genomics** – to encourage the large-scale, computationally intensive comparisons of sequencing information that produce scientific knowledge.
- **Minimise administration and legal complexity** – avoiding bilateral solutions which mean researchers have to negotiate access with multiple regimes.
- **Mitigate jurisdiction-shopping** – levelling the playing field by facilitating research in countries with better access and benefit-sharing, rather than in more restrictive ones.
- **Capacity-building** – codifying the non-monetary benefits, including technology transfer, training, scholarships and collaborations, which play a key role in bridging divides.

The sandboxes would need to be carefully structured and actively managed to ensure they foster the active participation of indigenous peoples and local communities as well as low to middle income countries. Sandboxes for biodiversity genomics could be a testing space to embed the lived experiences and perspectives of a diverse set of stakeholders in a robust and inclusive data-sharing and management ecosystem that ensures the free flow of data to secure a globally critical public good.

Benefits unlocked by cross-border sandboxes

Additional benefits that can potentially be unlocked, if cross-border sandboxes can address the identified challenges, include:

• Regulatory capacity-building and collaboration

- Sandbox discovery and deliberation provide valuable insights and actionable precedents (where legally appropriate) that inform broader compliance and regulatory decisions.
- Regulators learn from other legal traditions and regulatory cultures, applying new insights to domestic decision-making and future legislative input.
- Sandbox collaboration between regulators can increase formal and informal cooperation that improves compliance and enforcement in other areas.

• Competition and choice

- Smaller and innovative companies whose services may need regulatory certainty to attract investment, but may not have the resources to pursue it, get clarity and the opportunity to scale up across borders.
- Consumers, companies and the economy as a whole benefit from greater choice as cross-border data innovation is supported and makes more innovative products and services available.

- Innovative firms and organisations throughout the economy benefit from transparent cross-border decision-making that clarifies regulation and supports greater regulatory alignment between jurisdictions.

• Reduced arbitrage

- Sandboxes might help reduce risks of regulatory arbitrage and forum shopping, allowing for more consistency.
- Increased mutual support and collaboration among regulators towards achieving greater interoperability on key issues, such as compliance.

• Increased data availability

- Fewer or narrower data localisation measures in specific jurisdictions where risk and perceived risk has been reduced, and suitable mechanisms for transfers developed.
- Data available for use that previously would not have been transferred.

• Increased data accessibility

- Data is easier to use, more 'interoperable', thanks to good data use principles.
- Where appropriate, use of data unlocks value for more people, including benefit-sharing with indigenous groups and stakeholders of emerging data intermediaries (e.g. data cooperatives).

• Better compliance

- Less burden on regulators and their resources if applicability and content of rules is clarified.
- Streamlined licensing and lower licensing requirements on firms scaling up across borders.

04

Mapping challenges of cross-border sandboxes for data

Challenges of cross-border sandboxes for data

Along with the opportunities and benefits that cross-border sandboxes would bring, there may also be drawbacks, risks and challenges. For them to work well, they need to 'square the circle' of applying a systems approach to data governance, providing certainty about compliance with differing legal regimes, while operating in a focused, fast and agile way.

In addition to different regulatory and legal frameworks, regulators from different countries operate in different cultural understandings of risk, for example, how much of the risk or potential downside of data transfers and management should be borne by firms, data-subjects or other parties. Additionally, a key issue with all regulatory sandboxes - the fact that they cannot easily be used at scale - is magnified at the regional or international level.

Moreover, to the extent that sandboxes may test alternatives to data localisation measures, they may not respond to all concerns that catalyse such measures. A significant driver of data localisation is also national competitiveness and the desire for quick and easy law enforcement access to data, in addition to concerns about national security data. Such concerns may not be dealt with as easily by sandboxes.

A key concern with sandboxes is that they are often costly, resource-heavy and can only deal with one, narrowly focused problem at a time. They place a significant time and resource burden on regulators, but there are concerns that they do not scale well, i.e. deal effectively with many issues or cross-sectoral ones.[118] However, cross-border sandboxes have the potential to provide a targeted capability to answer key questions to the satisfaction of multiple regulators, while potentially mutualising costs.

There are also challenges of funding, varying resources and capability among regulators. Resource allocation needs to be measured against the regulator's wider statutory goals, and also to reflect both private (i.e. firm-directed) and public benefit. This may limit both the scope and scale of sandboxes. Regulators - particularly those outside financial services regulation - cannot always devote significant resources to 'blue sky' technologies, emerging business models and broadly scoped initiatives that tackle promising but challenging and complex issues. Particularly for recently established regulatory authorities, which is the case of many that hold the mandate over personal data protection, daily demands and resource constraints may be a concrete impediment in designing, engaging and evaluating sandboxes - national or cross-border.

These challenges can be addressed and reduced in the design and implementation of cross-border sandboxes through:

- Mitigating the complexity of different legal regimes by tackling **tightly focused and boundaried regulatory questions** based on clear preparatory work to identify the issues, as a requirement before initiating the sandbox.[119] Issue selection prioritises key regulatory questions, and targeted issue definition can mean the applicability of decisions to a range of innovations across markets.

- Sandboxes may have difficulty dealing with questions that lie across different sectors or the domains of different regulators or agencies. However, this can be mitigated by **engaging adjacent regulators and others**

---

118. Lesher, M. (2020), *Bringing new digitally enabled products and services to market: Sandboxes and the role of policy experimentation* , VoxEU Blog.

119. ICO's Regulatory Sandbox on Monetary Analytics and FutureFlow is a striking example of the effectiveness of setting clear and limited regulatory questions for the sandbox process to solve.

in the sandbox, and also by ensuring sandboxes are developed along consistent lines, such as the German government's sandbox strategy[120] or the roadmap for cross-border sandboxes for data set out in this report.

- Understanding that risk in some areas - personal health data, for example, or the potential economic impacts of financial innovations - may have an outsized, non-linear impact, should things go wrong. Information asymmetries for regulators may be multiplied in the international context; there is consequently a need to proceed cautiously while **planning to mitigate risk accordingly**.

- **Developing metrics and methods for assessing likely public benefits of sandbox** applicant innovations - especially in relation to the resources of the regulator and the limited ability to scale up the relative time and resource intensity of sandboxes. Areas where cross-border sandboxes may simply be inappropriate must be identified, e.g. when data may prompt national security or law enforcement considerations.

- **Sharing the costs of identification** of stakeholders, coordination, implementation and evaluation of results of cross-border sandboxes with other partners.

Challenges specific to operational sandboxes

Operational sandboxes, i.e. 'data spaces' or sandboxes that actually manage data, have additional, specific challenges. Centralising data can create a target for malign actors to use data against a donor's wishes or to expropriate data to an unprotected digital locale. Thus, there is significant current research exploring how decentralised technologies or approaches, such as federated computing, differential privacy, homomorphic encryption, or decentralised autonomous organisations (DAOs) supported by the blockchain can support larger data pools.

However, at present these technologies are computationally intensive. Large data-sets stored close to large processors with fast internet connections remain the standard. Such data-sets may emerge out of private collaboration inside the sandbox, cross-border public or pooled investments, or crowd-sourced data-gathering efforts. However, for the time being, they will tend to converge on centralised sites because of a legally compliant, locally available data-set which is, for now, arguably the easiest to work with. Accordingly, any operational sandbox will require a long-term commitment to independent systems to hold players accountable to the terms of joining, as well as to ensuring appropriate penalties for defecting actors.

These challenges can be addressed and reduced in the design and implementation of operational sandboxes by:

- Securing external validation to ensure security and integrity of data.
- Setting broader goal and performance metrics, and using evaluations of the sandboxes to iteratively improve their design.
- Fostering accountability and transparency.

In summary, given the variety of stakeholder's interests and regulators' goals, differences between their legislative and regulatory frameworks, and cultural understanding of and capacity for risk, it is clearly a significant challenge to create functioning cross-border operational sandboxes. However, the possibility to access and make available technologies, initiatives, products and services that have the potential to solve long-standing data dilemmas is a considerable incentive.

---

120. FMEACA (2021), *Making Space for Innovation; The Handbook for Regulatory Sandboxes,* Federal Ministry for Economic Affairs and Climate Action, Blog.

05

A preliminary roadmap for cross-border sandboxes for data

This section outlines the process for setting up a minimum viable product for a cross-border sandbox for data. Such sandboxes can be either purely regulatory or purely operational, or a combination of both. They can operate across boundaries - including jurisdictional, regulatory framework and sector boundaries (public, private, disciplines, etc.) - seeking to test and implement a holistic approach to international data governance. Some aspects of operational sandboxes require specialist approaches, particularly to the identification and mitigation of risk related to directly hosting data. However, all cross-border sandboxes for data share the need for collaboration between stakeholders to responsibly unlock the value of data by promoting the sharing of both data and advice.

What could the minimum viable product version of a desired sandbox look like? One approach is to begin with at least two sufficiently aligned (albeit not fully harmonised) national regulators who share a similar interest in an innovation or issue area. The key task then is to work through an iteration of the sandbox with just a few stakeholders, and only after that to increase challenges like lower regulatory alignment and more partners. However, it may also be possible to identify multiple stakeholders in several jurisdictions who are all keenly interested in resolving a key issue area or enabling a specific innovation. This approach may rely less on like-minded regulators forging a path, and more on an emerging coalition of interest that appoints a secretariat or sub-group to work on the sandbox. The steps set out below will be of use to both approaches.

The steps draw upon and apply the lessons learnt in the report for setting up and operating a cross-border sandbox for data. They include choice of partner, jurisdiction and issue; setting strategy and goals; agreeing format, duration and funding; defining coordination processes and mechanisms, mitigating information asymmetry and risk; and preparing for assessment. This guide is primarily addressed to regulators as the stakeholders most likely to initiate any type of cross-border sandbox for data, and policymakers who may seek to support the process.

Step 1. Choosing partner(s) and issue(s)

Clarifying regulatory grey areas is already complex and challenging, let alone in the light of more innovation and uncertainty. Starting with choosing a regulator in a country (or countries) with reasonably aligned regulatory regimes, for example, another common law or civil law jurisdiction, will be helpful. It may be necessary to explore options and possibilities with several potential partners. Choosing regulators with which one already has relationships and even some working ties may help. It may also be beneficial to start with regulators who have experience of developing sandboxes in a domestic context. It is also important that they have similar motivation and commitment to establishing the sandbox, as the inevitable challenges may overcome a weak commitment, and an uneven distribution of work and risk could hamper the chances of success.

Similar legal and cultural understandings of concepts such as risk, privacy and the role of the regulator in relation to innovation are key. Even where similar concepts, definitions and values are shared, it will be important to be explicit in collaboratively articulating them and agreeing on the ones the sandbox will be driven by.

Also critical is to identify the issue areas that the sandbox will prioritise. 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. 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 issue will attract support and resources, maintain focus on outputs, and motivate progress towards defined outcomes.

- Once the foundational elements of the sandbox are set up, the issue area prioritisation, identified by a pro-active outreach process, will help to identify and attract innovators to apply. National regulators and other government stakeholders will also need to publicise the sandbox so that all relevant stakeholders have a chance to participate. Depending on the range of stakeholders to be involved, outreach on issue identification and definition may also need to be directed to civil society, academia and the technical community.

- As with national sandboxes, it is essential to identify applicants whose innovations are truly novel and which raise questions that there is a reasonable public interest in answering, for example, by providing decision precedents with broader application and that can drive learning, both for innovators and regulators.

- The final, 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.

Step 2. Setting strategy and goals

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.

The typology of national regulatory sandboxes (Section 1) showed some key ways in which different jurisdictions can differ. While not every participant in a cross-border sandbox will already have run a national sandbox, they may still differ from others on legal and cultural questions such as regulatory leeway and flexibility, or even expectations about the role of a regulator in balancing data rights with innovation and growth. In order to bridge the potential differences between different ways of operating a sandbox, partners should structure a dialogue on the following questions:

- How much regulatory leeway or flexibility does the regulator have in relation to this sandbox, i.e. can they provide flexibility on compliance for the duration of the sandbox? Related, is their goal to work within their own legal and statutory framework, or to work on testing and developing it? Answering these questions sets expectations of the degree of flexibility each partner can exercise, and whether they see the sandbox as a way to test out different frameworks. If one partner's goal is simply clarifying the application of regulations, and another is interested in using the sandbox to adapt its regulatory framework, they need to know at the beginning, and articulate expectations so that the sandbox creates equal value for both.

- What are the legal and policy expectations of each regulator in relation to concepts such as privacy, innovation or the sharing of any benefits from the use of personal or non-personal data (if relevant)? Are these seen as a trade-off, and if so, what is the implicit or explicit expectation of the regulator in their national context? How will competing values or priorities be reflected in the shared goal of the sandbox?

- What are the goals in relation to transparency, learning and public interest? How much of what is learnt will be made available to other firms, regulators, and civil society/academia at large? If the establishment of precedent is important, how will this be communicated and what status will it have?

These questions to trigger goal discussion and setting are just an opening. They try to surface the differences between existing national sandboxes identified in this report, and recognise that these differences come from varying national strategies and approaches, as much as from statutory regulatory constraints. As more national and cross-border data sandboxes emerge, more insights from them may drive a wider dialogue about goal-setting, aimed at articulating what might otherwise remain implicit differences that may emerge or present issues late in a process.

In addition to setting a broad strategic goal that articulates the interests and values in play, the sandbox needs to have working level goals that communicate its ambition, set targets to evaluate progress against, and fit within funding and reporting cycles of the national sponsors. For instance, clear and measurable operational targets can be set through goals such as: "Set up a cross-border regulatory sandbox for data comprised of Regulators A and B, define oversight, funding, internal decision-making and working practices, and fully clarify and graduate three innovations within twenty-four months", or "Set up a cross-border operational sandbox bringing together at least five collaborators to support innovation on carbon emission accounting, with defined group boundaries, data security provisions, monitoring systems, and dispute resolution processes, in order to produce a scalable solution within twelve months".

Step 3. Identifying and engaging stakeholders

Defining the sandbox's core, active and supporting stakeholders is key.

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 organisations 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 on 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. Critical initial issues this cross-jurisdictional group will need to resolve include:

- Oversight and ultimate decision-making, and measures to foster accountability and trust

- Responsibility and processes for coordination, communication and escalation

- Ways of working - virtual, hybrid, in-person (and if so, where)

- Ongoing or periodic evaluation, including criteria and reporting

Step 4. Agreeing format, duration and funding

Early consensus on the type, working practices and duration of the sandbox is crucial.

- What kind of sandbox is envisioned? Will it be a regulatory sandbox, focused on clarifying and resolving how to apply - and potentially to evolve - regulation? Or will it be an operational sandbox, bringing together data-sets and collaborators to support new innovation? Or will it combine both, trialling new uses of data alongside governance testing?

- If it is a regulatory sandbox, how many issues or cases will it work on? What will be the engagement, application and selection processes for innovators?

- What resources need to be allocated? For an operational sandbox, what additional technical resources, processes, risk identification and mitigation will be needed to protect data and assure the operation of the sandbox?

- For how long will it operate?

Core stakeholders need to agree on a realistic timetable and allocate sufficient resources. Cost specifically is a real concern and challenge in the context of regulatory and public sector initiatives. This type of sandbox may be more demanding of resources and funding than national sandboxes, due to the increased complexity, coordination costs, more complex decision-making, and the likelihood of greater information asymmetries and higher risks (see step 5). The higher commitment and costs needed for success are a key reason for focusing the sandbox on issues of shared importance. Key questions to be resolved include:

- What is required in terms of person-hours from each regulator, and additional resources to support them?

- What are the additional costs and organisational resources required to support cross-border working methods?

- Do these costs make the ultimate goals worthwhile achieving? Both in terms of the immediate sandbox goals and the broader goals of trialling and developing a new regulatory capability?

- Should innovators pay? If so, all or part of the additional costs? This can be prohibitive and may defeat the purpose of using sandboxes to encourage innovation amongst new market entrants.

Step 5. Addressing key issues, information asymmetry and risks

Risk in some areas - personal health data, for example, or the potential economic impacts of financial innovations - may have an outsized, non-linear impact if things go wrong. Information asymmetries for regulators may be multiplied in the international context. In addition to the organisational and procedural issues of the sandbox, core stakeholders need to explore and mitigate the key issues of information asymmetry and risk.

Information asymmetry between the regulator and the innovators and/or market under consideration means decisions under uncertainty and with incomplete knowledge. Regulators do this every day. However, at the cross-border level, information asymmetry is greater, drawing in different markets, practices and even languages. More time and resources may need to be built into the sandbox processes so that regulators can become sufficiently knowledgeable about emerging innovations in relevant markets. The sandbox will need to:

- Build in extra time for discovery, both for acquiring and processing essential information

- Consider allocating extra funding or resources to acquire expert input on specific topics

Most importantly, the sandbox will need to be designed to recognize and accommodate a higher level of risk. Risk encompasses the risks to people and the economy of innovations that, for example, leak data or compromise it in some way. Mitigating this may mean enhanced security measures for operational sandboxes, such as external validation to ensure security and integrity of the data. For regulatory sandboxes, it may require more active monitoring and discovery of business practices and the innovative technology itself.

Working cross-border, different regulators may also have different approaches to risk, determined legally and in the statutory responsibilities of the regulator and also by the content of regulation. They may also have different cultural acceptance and understanding of risk, and varying wider economic structures that allocate risk and negative externalities. The core stakeholders need to:

- Identify all forms of risk and their safeguards and mitigations:
    - What risks are there for users including citizens, and third parties? To the economy or sector?
    - Where does liability lie? Is this adequate?
    - Have these risks been identified and mitigated, and insured if necessary?
    - What processes does the sandbox have for dealing with problems?

- Agree on appropriate binding commitments and mandatory enforcement mechanisms for the firms or organisations entering the sandbox

- Agree on and get appropriate institutional buy-in for any necessary national enforcement of commitments and allocation of risk

Additionally, the dialogue and discovery on risk should be done early on so that any legal steps to indemnify actors can be taken, or to create an appropriate legal structure for the sandbox, if that is required.

Step 6. Preparing the criteria for evaluation

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 data 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:

- 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 organisation 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?

- For both, with the benefit of experience, what additional or different metrics/reporting requirements would be useful for future iterations of this same sandbox?

Evaluations of the sandboxes should be targeted at measures that will help to iteratively improve their design and develop best practices that can be used by other sandboxes. Moreover, the model of cross-border sandboxes for data itself should be evaluated in order to be tailored and improved after different iterations.

This preliminary roadmap for cross-border sandboxes for data starts from the assessment that sandboxes are challenging and resource-intensive but can produce targeted, high-quality decision-making with a broad impact. The roadmap recommends starting with somewhat aligned regulators working on identified issues of equal importance and urgency to both.

The insights, working methods and experiences gained in this setting can then be built on. They can be used in sandboxes with more regulators, and stakeholders from more parts of government or from civil society and academia, move into different issue areas, increase the scope of issues dealt with, and of course increase the number and through-put of issues. The model and experiences can be built upon and extended, and also taken as a blue-print by regulators in different regions and issue areas to adapt and implement as they see fit.

06

Conclusion

This report lays out key information and insights on regulatory sandboxes and the policy silos and cross-border data issues they may be used to address. It highlights aspects of emerging technologies, approaches and data models – including data intermediaries, PETs and browser solutions – that may be compatible with using cross-border sandboxes to facilitate their adoption internationally.

To the extent that data measures, frameworks and governance models may be driven by concerns about the lawful treatment of data – both personal and non-personal – cross-border regulatory sandboxes for data can be used to establish and bolster compliant but agile data governance. The ASEAN example of a national regulator-led sandbox to enable data-flows between countries shows that this is a promising area of development. At the same time, cross-border sandboxes will not necessarily address all the concerns that underline data governance. But for regulators and policy makers with the will and resources to use cross-border sandboxes in a targeted way, they can establish and sustain data-flows, while also building relationships and trust.

Cross-border regulatory and operational sandboxes for data are still in their infancy, therefore we do not yet know the scope and scale of issues they can resolve. As with any new decision-making model, significant time and resources may be needed to pilot and establish their institutional frameworks as well as formal and informal working mechanisms. This report presents a range of issues which sandboxes could be used to resolve. Where innovation is impeded by uncertainty, there is an opportunity to apply the sandbox model to unlock opportunities, particularly for emerging firms and data intermediaries.

However, just because an idea is a good one does not mean it will succeed in implementation or be widely used. The success of cross-border sandboxes in dealing with many issues which may otherwise remain unresolved will depend on early use cases that build knowledge, confidence and momentum. Success also depends on clearly delineating the questions to be resolved, and maximising transparency of process and outcomes. Cross-border sandboxes can solve problems other methods struggle with, but they need to be started off in the right way. From a social science perspective, path dependence builds network effects; how one launches and initially implements an idea has more impact on whether it becomes a dominant force. Getting the early decisions right, to show and build on how the model works best, can put cross-border sandboxing at the forefront of solving more, tougher and bigger issues.

> **The success of cross-border sandboxes in dealing with many issues which may otherwise remain unresolved will depend on early use cases that build knowledge, confidence and momentum. Success also depends on clearly delineating the questions to be resolved, and maximising transparency of process and outcomes. Cross-border sandboxes can solve problems other methods struggle with, but they need to be started off in the right way**

A key challenge identified in this report is to involve regulators from jurisdictions that have some regulatory similarities but sufficient differences to make sandboxing attractive and worthwhile. Multilateral treaties, free trade agreements or legislation-derived bilateral contracts and model clauses can be either too time-consuming or too rigid. Reliance on consent as a basis for transfers has been undermined by consumer fatigue and distrust,

and the use of 'dark patterns' by websites.[121] Early sandboxes that involve regulators from similar but not identical data regulatory traditions could build the early success and institutional know-how to tackle more complex and larger scale issues, in relationships of higher regulatory divergence and with more regulators.

Cross-border sandboxing also has a path to overcoming issues of scale. A concern frequently raised is that sandboxes are resource-intensive and only tackle one discrete issue at a time. However, cross-border sandboxes have the potential to provide a targeted capability to answer key questions to the satisfaction of multiple regulators. In settling well-chosen questions of high importance and public benefit, regulators can multiply the benefits and provide more than enough precedent and experience to justify the investment.

Finally, this report has addressed many key issues, but **some questions remain**. Questions which are important to answer before the design and implementation of a cross-border regulatory or operational sandbox include:

- Can legal restrictions regarding mandates to store data locally, local supply requirements, and the effects of conflicts between laws in different jurisdictions be made more flexible by regulatory sandboxes?

- How do cross-border sandboxes operate in the context of mixed data-sets of personal and non-personal data? Particularly if data is from a policy area or sector with different regulatory compliance structures, e.g. if it touches on competition?

- What should be the fiduciary responsibilities of cross-border data intermediaries or operational sandboxes?

- What does enforcement and compliance look like in cross-border sandboxes? What kind of internationally binding commitments can parties to the sandbox make, and how accessible is enforcement to those potentially affected?

- How do regulators dealing internationally handle information asymmetries of the firms, innovations and industries they may be dealing with, which can be multiplied at this level?

These are key questions for further research. Sandboxes can further many policy objectives, while balancing sometimes competing priorities like innovation and risk. They are one method, but a potentially significant one, for responsibly unlocking the value of data for all.

---

121. Christine Utz, Martin Degeling, Sascha Fahl, Florian Schaub, Thorsten Holz (2019) *(Un)informed Consent: Studying GDPR Consent Notices in the Field,* ACM Digital Library.

thedatasphere.org