Europe's AI Sandboxes:

Navigating Regulatory Evolution

FORHUMANITY

Content

1 Introduction 3

2 Sandboxes in general 4

3 Sandboxes and the EU AI Act 13

4 Sandbox innovation 24

5 The possible impact of the sandbox 33

6 Illustrative Case Study - Financial Services 35

7 Conclusions 37

8 APPENDIX: Comparison of Commission and Parliament texts 41

9 About ForHumanity 55

References 56

1 INTRODUCTION

Regulatory experimentation through combinations of soft law, hard law, and novel collaboration approaches is gaining ground in the increasingly busy field of technology regulation [1].

The EU's proposed AI Act[1] [2] includes several measures to protect innovation, of which sandboxes are one. Sandboxes are a kind of regulatory experimentation, however, the term sandbox is also used in computer science to describe an isolated environment used for testing or analysis [3]. Sandboxes conjure up an image of a physical environment, with the walls of the sandpit preventing uncontrolled effects. However, in reality they are mostly a legal construct, rather than a physical space. Exactly where sandboxes are physically hosted and operated is not clear at present[2].

The regulatory sandboxes envisaged by the AI Act are the **largest such regulatory experiment ever attempted**. While there is no requirement for AI providers or deployers to utilise sandboxes, there are benefits for both regulators and innovators.

This report defines and describes sandboxes from a regulator and an innovators perspective, analyses their history and modalities, explains how the EU's AI Act envisages sandboxes, identifies potential areas of innovation that could assist with the delivery of sandboxes, and analyses the potential impact of the sandboxes. It also points to differences between the Commission proposal and the Parliament proposal, which have significant differences. The differences are also listed in an appendix, section 7.3.

It ends with an illustrative case study of how an AI sandbox could work in the context of a common financial services use case.

AI SANDBOX DEFINITION [4]

A regulatory sandbox is a controlled environment established by a public authority that facilitates the safe development, testing and validation of innovative AI systems for a limited time before their placement on the market or putting into service pursuant to a specific plan under regulatory supervision

[1] Although the AI Act is properly described using the phrases 'proposed' or 'draft', hereafter in this document it will just be referred to as the AI Act for brevity. Where a specific version is referred to, it is explicitly referenced.

[2] Practically speaking, it may be economically and logistically difficult for regulators to provide a hosted environment given the infrastructure demands of some AI systems. On the other hand, where data is being shared between organisations solely within the context of the sandbox, it may be very appropriate.

2 SANDBOXES IN GENERAL

2.1 HISTORY OF SANDBOXES

Although legal experimentation began in financial services as early as 1999 [5], the operation of regulatory sandboxes can be traced back only to the early 2010s when the concept emerged as a response to the challenges faced by regulators in keeping pace with the rapid advancements in technology, particularly in the financial services sector.

Faced with large-scale technological innovation in regulated markets, regulatory actors have four main approaches: do nothing, allow flexibility on a case-by-case basis, provide a structured context for experimentation or plan to reform/adapt better regulation. Recently, structured contexts for experimentation have become popular - this is a sandbox.

A regulatory sandbox is a framework set up by one or more regulators to collaboratively test innovations by one or more third parties in a controlled environment, operating some kind of special exemption, allowance, or other limited time-bound exception, and under the regulator's supervision.

In 2015, the Financial Conduct Authority (FCA) in the United Kingdom launched the world's first regulatory sandbox. The FCA's sandbox aimed to foster innovation in the FinTech industry while maintaining consumer protection and financial stability [6]. It provided a controlled environment for FinTech startups and companies to test their innovative products and services under relaxed regulatory requirements.

In 2016, the Monetary Authority of Singapore (MAS) introduced the FinTech Regulatory Sandbox [7]. MAS aimed to position Singapore as a leading FinTech hub by providing a conducive environment for experimentation. The sandbox allowed companies to test their FinTech solutions, receive guidance from regulators, and validate their business models before obtaining full regulatory approval.

Following the success of the UK and Singapore sandboxes, several other countries and regions adopted the concept. Regulators launched sandboxes in Australia, Canada, Hong Kong, Malaysia, Thailand, and the United Arab Emirates. These sandboxes catered to their respective jurisdictions' specific needs and regulatory landscapes, facilitating innovation in FinTech and other sectors.

While regulatory sandboxes initially focused on FinTech, the concept expanded to other industries. Sandboxes were established in sectors like health technology [8] and energy technology [9]. This diversification allowed for innovation in various domains while ensuring compliance with sector-specific regulations.

Over time, regulatory sandboxes have evolved rapidly. Regulatory authorities incorporated lessons learned from the early sandboxes, refined their approaches, and introduced guidelines and frameworks [10] to enhance the effectiveness of sandboxes. Authorities engaged in knowledge sharing and international collaboration to exchange best practices and shape regulatory approaches.

In 2020, the European Council adopted a set of conclusions on the role of regulatory sandboxes and experimentation clauses in an innovation-friendly, future-proof, sustainable and resilient EU regulatory framework [11]. The European Council defines regulatory sandboxes *'as concrete frameworks which, by providing a structured context for experimentation, enable where appropriate in a real-world environment the testing of innovative technologies, products, services or approaches (...) for a limited time and in a limited part of a sector or area under regulatory supervision ensuring that appropriate safeguards are in place'*. The inclusion of experimental instruments in the regulation of AI can be partially explained by the need to accommodate rapid development and complexity [12].

The history of regulatory sandboxes demonstrates the recognition by regulators and policymakers of the need to balance innovation and regulation in rapidly evolving industries. Sandboxes have provided a platform for collaboration, experimentation, and learning, enabling regulators to adapt and develop effective frameworks that support responsible innovation.

Regulatory sandboxes offer an environment where innovators can conduct limited tests of their innovations with fewer regulatory constraints, real customers, less legal risk, and enhanced dialogue with regulators. As discussed later, these environments can be used as a pro-innovation measure to support the AI Act.

2.2 GENERAL OBJECTIVES OF SANDBOXES

From a legal perspective, a regulatory sandbox refers to a controlled and temporary framework established by regulatory authorities to facilitate innovation in regulated industries. The sandbox operates under specific guidelines and regulatory frameworks tailored to the needs of the participating industry. The sandbox framework enables regulators to closely monitor participants' activities, assess risks, and gather valuable insights to provide guidance and inform future regulatory approaches.

The specific objectives of sandboxes can vary based on the regulatory context. Some typical goals include:

- **Promoting innovation:** Sandboxes aim to foster innovation by providing a controlled space where startups and companies can test and develop new products, services, or business models. Sandboxes encourage experimentation and creativity by removing spe-

cific regulatory barriers or offering flexible regulatory frameworks.

- **Regulatory learning:** Sandboxes allow regulatory authorities to gain first-hand experience and insights into the regulatory implications of new technologies and innovations. This learning process helps regulators stay informed about emerging trends, risks, and opportunities in the FinTech sector. By actively participating in the sandbox environment, regulators can refine their regulatory approaches and make informed decisions regarding policy changes or updates.

- **Regulatory compliance:** While promoting innovation, sandboxes ensure participating companies adhere to relevant regulatory requirements. Regulatory authorities establish specific guidelines for operating within the sandbox, outlining the compliance measures that participants must follow. This helps regulators understand and assess potential risks associated with innovative FinTech solutions.

- **Consumer protection:** Protecting consumers is a crucial objective in some sectors. By monitoring and overseeing the activities of companies in the sandbox, regulatory authorities can assess potential risks and ensure that consumer interests are safeguarded. This objective involves establishing mechanisms to handle customer complaints, ensuring fair treatment, and addressing privacy and data security concerns. Sandboxes can limit consumer impact through sandboxes in terms of numbers, but also in other ways such as obtaining enhanced consent from consumers.

- **Collaboration and knowledge sharing:** Sandboxes enable regulators, innovators and customers to collaborate and share insights. By fostering dialogue and collaboration, sandboxes can enhance the understanding of emerging technologies, business models, and regulatory challenges. This objective facilitates the development of informed and practical regulatory frameworks that can adapt to the rapidly evolving technology landscape.

- **Market monitoring:** Sandboxes provide regulatory authorities with an opportunity to monitor developments in the market closely. By observing the behaviour and impact of new products, services, or business models, regulators can assess potential risks to society, personal data, market stability, competition, and financial integrity. This objective helps regulators pro-actively address any emerging challenges.

2.3 KEY CHARACTERISTICS OF SANDBOXES

Sandboxes are an experimental regulatory approach that also provide the opportunity to pro-actively refine regulatory guidance. This approach differs from evidence-based methods, which rely on existing data, expertise and scientific information, as sandboxes actively seek new data points and stakeholder learning [5].

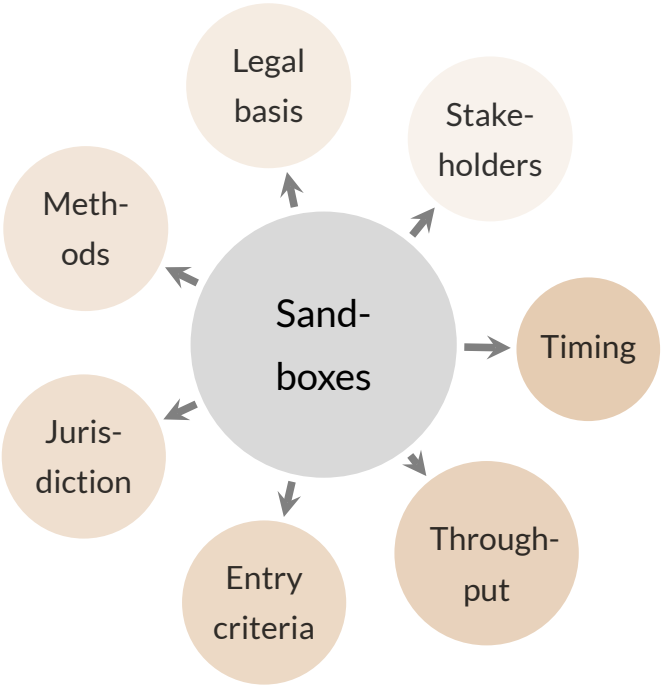


FIGURE 1: KEY CHARACTERISTICS OF SANDBOXES

2.3.1 ENTRY CRITERIA

In some countries, anyone can apply for a sandbox. In other countries, they may need to be already licensed or approved for operation. For example, in the UK, a firm may apply for limited regulatory authorisation before full authorisation to offer financial services.

Other criteria that can be used include:

- The innovation may need to be sufficiently novel in some sandboxes to deserve special focus [13].
- Regulators may also determine that there is adequate regulatory clarity already available about the proposed innovation.
- The innovation may need to be of benefit to consumers or society.
- The innovation may also need to be at a certain level of development. For example, testing an innovation that does not yet function may be impossible.
- Regulators may also consider particular risk factors like market stability.
- The applying organisation may need to have sufficient management controls on entry to protect against risks.
- The regulators may consider whether the company trying to innovate is too big. Some have proposed that Google, Amazon, Meta and Microsoft should not be allowed to participate as they are already 'too big to fail' [14]. Indeed, these companies may prefer to

litigate their use of innovations post-hoc.

2.3.2 STAKEHOLDERS

Regulatory sandboxes typically involve stakeholders who play different roles in the sandbox ecosystem. Regulatory sandboxes provide a unique platform for regulators and innovators to collaborate, experiment, and learn in a controlled environment. This engagement can also significantly benefit the development of international or harmonized standards by fostering collaboration, identifying regulatory gaps, providing feedback, and facilitating the alignment of regulatory approaches.

Research focusing on qualitative analysis of the interactions between regulators and regulatees in sandboxes claims that [15]:

- Regulatees benefit from access to informal and formal networks, either in the provision of advice or cross-border introductions.
- Knowledge exchanged during the sandbox increases regulator understanding of constraints and risks arising from new technologies, improving monitoring practices.
- The frequency of interaction increases the understanding of each party.
- The interaction improves the regulatees risk management practices.
- The creation of a common language positively affects knowledge exchange.

The same research also warns that:

- Asymmetrical information exchange restricts regulatees' willingness to share best practices.
- Regulators' unwillingness to make regulatory changes negatively impacts regulatees' testing manoeuvrability.

The specific stakeholders involved in a sandbox can vary depending on the jurisdiction and objectives of the sandbox. Here are some examples of stakeholders that could be involved in regulatory sandboxes:

- **Regulatory Authorities:** Regulatory authorities, such as financial regulators, healthcare regulators, energy regulators, or other relevant regulatory bodies, play a central role in establishing and overseeing regulatory sandboxes. They define the framework, guidelines, and regulatory exemptions for participants. Regulatory authorities are responsible for monitoring and assessing the risks associated with sandbox activities and ensuring compliance with applicable regulations. Data protection and privacy authorities play

a crucial role when sandbox activities involve collecting, processing, or sharing personal data.

- **Industry Participants:** startups, technology companies, established financial institutions, healthcare technology providers, energy technology companies, or other innovative firms are the primary participants in regulatory sandboxes. These companies bring their innovative products, services, or business models to the sandbox for testing, validation, and development. They collaborate with regulators, adhere to sandbox guidelines, and provide feedback to shape future regulations.
- **Consumers/Users:** Consumers or end-users of the products or services being tested in the sandbox are important stakeholders. Their feedback, experiences, and protection are considered during sandbox testing. The involvement of consumers helps regulators understand the potential impact of innovative solutions on consumer rights, privacy, and overall user experience.
- **Civil Society:** Civil society refers to the diverse and interconnected network of voluntary associations, organizations, and individuals that exist beyond the realms of government and the market. It encompasses a wide range of non-state actors, such as non-governmental organizations, community groups, advocacy organizations, and philanthropic institutions, which collectively engage in social, cultural, and political activities aimed at promoting civic engagement, social cohesion, and the protection of individual and collective rights within a society. Such groups can bring unique perspectives to regulatory considerations in sandboxes.
- **Regulatory Support and Advisory Bodies:** Some regulatory sandboxes involve support or advisory bodies that provide participants guidance, expertise, and assistance. These bodies can include innovation hubs and regulatory innovation teams.
- **Industry Associations and Standards Organizations:** Industry associations and standards organisations relevant to the sector covered by the sandbox may be involved as stakeholders. These organisations provide industry-specific expertise, promote best practices, and contribute to developing sector-specific regulations. They collaborate with regulatory authorities and participants to ensure that sandbox activities align with industry standards and norms.
- **Academia:** The involvement of Academic stakeholders can be particularly use in the context of nascent and emerging technologies that are under active research.

The involvement of stakeholders can vary depending on the nature of the sandbox and the specific sector it covers. The collaboration and interaction among these stakeholders are essential for the success and effectiveness of regulatory sandboxes.

2.3.3 LEGAL BASIS

By legal basis, we mean the means by which the the sandbox has authority has to waive, relax or not enforce certain obligations. A sandbox can be created by a legislative body in new regulation itself or by an existing regulator with existing regulation.

The extent to which regulators can exercise discretion depends on their legal basis for decision-making. This may be difficult in a cross-regulatory context. Some argue that single regulator sandboxes entrench existing regulatory borders, reduce economies of scale and create superfluous restrictions [16].

2.3.4 METHODS

Methods can firstly vary by legal approach, in the context of the legal basis. For example, sandboxes can experiment by derogation or by devolution [17]. Derogation means that rules or guidance are put aside for participants in the sandbox in exchange for alternative rules. Devolution implies geographic (e.g. state) or vertical domain waivers (e.g. national security). One US example of this is for the testing of UAVs[3], FAA[4] regulations were waived for specific participants, in a specific geographic area, for 30 months [18].

Methods can also vary based on the way that the regulator interacts with the regulatee. Regulators in a sandbox can provide bespoke individual guidance, that is, customised guidance provided to the innovator. Regulators can 'provide comfort' about what they consider compliant behaviour and their approach to enforcement [19]. They can also provide 'no action letters', stating that they won't enforce something.

Other legal methods may include commitments about protection of intellectual property to the innovator, and protection from civil liability.

In many sandboxes, the action taken, or participation in the sandbox itself, is time-limited. Theoretically, the time limit depends on the context and is not set [19]. In practice, it usually varies between six and twelve months [17].

Methods are likely to vary across sectors and risk profiles. For example, the methods used by a medical regulator dealing with a large-scale test of an AI system that poses risk to life, may be quite different to those used for a creditworthiness assessment.

[3] Unmanned Aerial Vehicle
[4] Federal Aviation Authority, USA

2.3.5 JURISDICTION

By jurisdiction we mean the geographical authority, or the authority within a specific sector or the authority to make legal decision over a particular matter.

Some jurisdictions opt for international collaboration by establishing agreements or partnerships with other countries or regulatory bodies. This approach promotes knowledge exchange, shared experiences, and harmonization of regulatory approaches across borders. It enables companies to operate in multiple jurisdictions and facilitates cross-border innovation.

However, it has not so far been very successful. In May 2022, the Global Financial Innovation Network (a network of financial services regulators) launched a cross-border sandbox spanning 23 regulators [20]. Thirty-eight firms applied, and it was noted that the entry criteria for each regulator were different, creating issues in the application process. Following assessment, only nine firms made it through to 'testing development', and two firms made it through to the live testing phase. A critical conceptual challenge pointed to in the report on the sandbox is that regulators were unwilling to engage in the harmonization of requirements, which led to difficulties in selecting innovators to progress.

Some jurisdictions, such as the EU, support cross-border recognition of licenses awarded by one state and are pursuing internal cross-border sandbox initiatives [21].

2.3.6 TIMING

Of particular importance is whether the sandbox is operating in the 'real-world' during the sandbox.

Sandboxes can operate ex-ante, that is, before placing a product or service on the market. They can also operate ex-post, in a market surveillance or with limited approval for the operator. Operating ex-ante is the easiest option, as it means that consumers and markets cannot be harmed during the sandbox operation. Typically, operating ex-ante means considering design, implementation and test results from a product perspective and management systems from an organisational perspective.

Operating ex-post or in a market surveillance model may involve allowing specific customers to use the application or allowing it to be used in a limited way. This can allow greater learning, as regulators can see the results of their guidance in the real world, and innovators can see the impact of the guidance on their innovation or business model.

Operating ex-post has implications that regulators need to consider carefully. What harm is possible to consumers? What risks may manifest affecting fundamental rights, health or safety? Does particular consent or information need to be given by affected citizens?

2.3.7 THROUGHPUT

Regulatory sandboxes have yet to be shown to be scalable [22] for innovators. FCA sandboxes in the UK typically have 18-24 participants in each cohort in comparison to the 50,000 companies the FCA regulates [23], and each is over-subscribed several times [24]. This is presumably due to the cost and risk to the regulators, explored in the next section. Existing sandboxes dealing with AI applications specifically typically have a cohort in single digits number of participants.

2.4 BENEFITS AND RISKS OF SANDBOXES

Case-by-case experimentation and provision of regulatory clarity through no-action letters and restricted licenses has upsides and downsides for regulators and innovators [22].

For regulators, one downside is the risk of liability for decisions and the effort associated with bespoke choices. The degree to which they may have liability depends on the specific legal context, their competency, and the information symmetry with the innovator [17]. The assessment and analysis criteria that they use may not capture the effect of a product on the market or consumer risk [25].

When regulators make decisions, they have high levels of information exchange with the participants and can adjust their approach more easily. They can also benefit from first-hand experience with innovation, thus building capacity [17].

For innovators, the process of obtaining such regulatory comfort is often costly outside of a sandbox (or innovation hub). Firms need to procure legal advice and develop applications and reports. Each application will require in-depth development. On the other hand, this can be more cost-effective than taking legal risks or acquiring legal advice in emerging regulatory areas. Additionally, the public association with the regulator may encourage investment [15]. Regulators can also offer financial incentives [7].

Ensuring equitable involvement in sandboxes is crucial for the market. The overall ecosystem gains advantages from the improved regulatory transparency they offer. As described, those engaging in the sandbox experience a degree of regulatory comfort for specific compliance aspects. As adhering to regulations can incur substantial expenses, this presents a compelling motive to take part in a sandbox, thus bolstering competition. However, the regulator must strike a balance in this aspect. They must guarantee that non-participants in the sandbox don't perceive any bias or injustice.

Sandbox participants receive insights and counsel from other stakeholders that propel their innovation forward. This has the potential to create a perception of unfair competition, where the regulator elevates the market potential of a single player at the expense of those who opt

out of the sandbox program.

In conclusion, the key risks of sandboxes relate to regulator capacity, equitable involvement of the market and the entry criteria for selecting participants. The key benefits are the regulatory learning for regulators, the cost benefits for innovators, as well as the attractiveness of regulatory association to investors.

3 SANDBOXES AND THE EU AI ACT

3.1 INTRODUCTION

In the previous section we describe the general characteristics of sandboxes. In this section we look specifically at the AI Act in Europe, and its provisions relating to sandboxes.

Many stakeholders are concerned about the impact of the AI Act on innovation. The proposers of the Act have primarily addressed this through a focus on technical standardisation and the presumption of conformity [26] and regulatory sandboxes.

In the Commission draft [2], member states are not required to set up an AI sandbox. However, in the latest Parliament draft [4] the are required to set up at least one. However, they may do it jointly with another member state. Sandboxes can also be set up at a more local level or by the EC. Parliament's version also introduces a firm definition of sandboxes, instead of leaving it to the recitals.

The EU AI Act attempts to meet many of the objectives listed previously: fostering innovation, ensuring compliance, and regulatory learning. These objectives are described in more detail in the Parliament version. Notably absent is any focus on the consumer.

A new Article 53a added in the most recent Parliament [4] amendments states that the sandboxes shall facilitate the development of tools and infrastructure for testing, benchmarking, assessing and explaining dimensions of AI systems.

The following section analyses the proposed AI sandboxes in the AI Act in the context of the characteristics previously described, and sandboxes that have mobilised at the time of writing.

Although the first sandbox proposed under the EU's AI Act is in Spain, other countries have worked with AI systems in other sandbox contexts, for example data privacy. In the UK, Norway and France, Data Privacy Authorities have launched regulatory sandboxes that are working with single digit numbers of AI companies, in specific targeted areas of innovation [12]. In Germany, a regulatory sandbox offered a testbed for 7 months for autonomous delivery robots. Russia also introduced a regulatory sandbox for AI technologies in 2021.

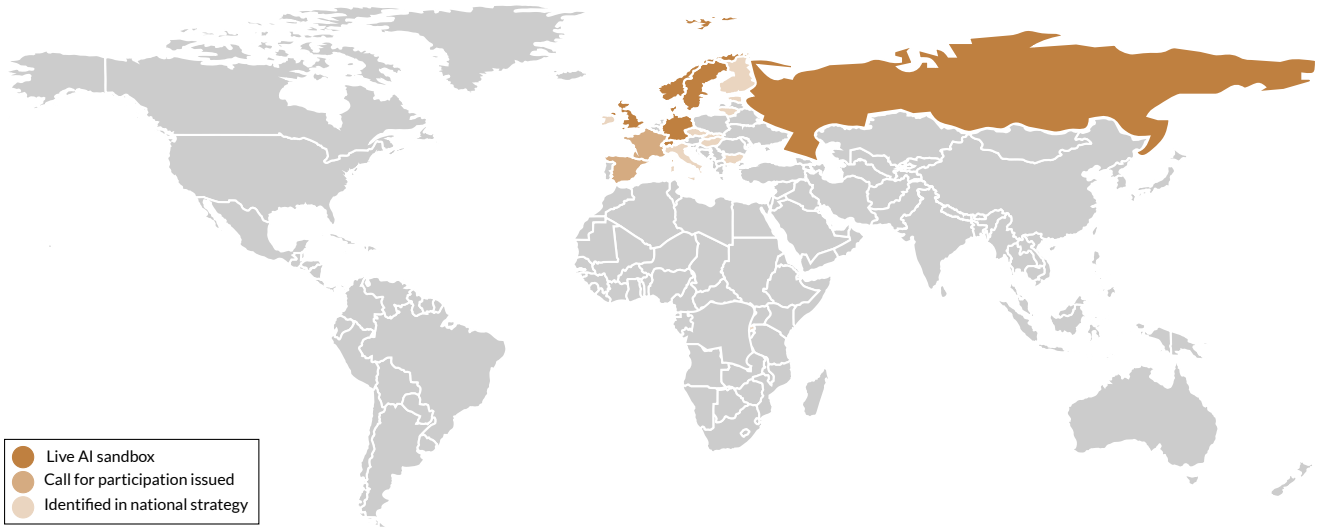


FIGURE 2: MOBILISATION OF AI SANDBOXES

Figure 2 shows the current state of AI sandboxes across the world. In Europe, most member states have published a national AI strategy [27] that references AI regulatory sandboxes. Two member states are at the call for participation stage (Spain and France), and there are sandboxes that mention AI in their objectives are live in four member states: Germany, Norway, Denmark and Sweden.

3.2 ENTRY CRITERIA

The Commission text is clear that the process must be transparent, fair, and open to any AI provider who meets specified criteria. Implementing acts will provide more details.

The Parliament version of the AI Act seems to be limiting participation to AI providers[5], which this report recommends extending to deployers (see section 3.1).

It is clear that specified criteria for participation will be provided. This is important, as regulatory experimentation tends to be viewed as unfair by those in industry that do not participate and could even be subject to world trade disputes - noting that China has already raised five trade disputes in relation to the AI Act [28].

If the sandbox results in some providers getting a lot of free advice at the expense of the provider's competition, then the market may end up worse off [29]. Full transparency of regulatory findings and guidance while protecting personal and private information will maximise the positive impact of the sandbox.

Some legal researchers [30] also point to Article 20 of the Charter of Fundamental Rights. There is existing case law relating to discrimination in relation to the equitable application of regulatory experimentation.

[5] 'regulatory sandboxes are open to any applying prospective provider of an AI system'

JUDICIAL OPINION OF ADVOCATE GENERAL POIARES MADURO, CASE C127/07 SOCIÉTÉ ARCELOR ATLANTIQUE [31]

[It is] in the very nature of legislative experimentation that tension with the principle of equal treatment should arise. The very idea of 'learning by doing' requires that the new policy be applied to only a limited number of its potential subjects to begin with.

As a result, the scope of the policy is artificially circumscribed so that its consequences can be tested before its rules are extended, if appropriate, to all operators who might, in the light of its objectives, be subject to it. That said, recognition of the legitimacy of legislative experimentation cannot invalidate any criticism that might be levelled against it from the point of view of the principle of equal treatment. The discrimination which experimental legislation inevitably entails is compatible with the principle of equal treatment only if certain conditions are satisfied.

The experimental measures must first of all be transitory. That is indeed the case with the Directive. Article 30 provides for a review of the Directive on the basis of experience and progress achieved in the monitoring of emissions of greenhouse gases with a view to including other industrial sectors and emissions of other greenhouse gases in the greenhouse gas emission allowance trading scheme. In application of that provision, the Commission proposed the inclusion of aviation activities [...]

Second, the scope of the trial measure must be defined in accordance with certain objective criteria.

In terms of maximising regulatory learning, diversity of participation from a technology and use case perspective should be maximised. In terms of maximising national innovation, the perceived national benefit should be maximised. These goals may conflict and there is likely to be latitude for regulators to affect throughput by limiting the novelty of innovation. If a sandbox applicant represents a similar use case and technology to one already examined, should the sandbox admit the new organisation? The AI sandboxes that are mobilised now generally require novel innovation and range between the general [32], selecting specific sectors [33], companies and use cases [34] [35].

Currently, the Parliament text gives priority access to SMEs established within the EU, and projects specifically intended to produce socially and environmentally beneficial outcomes.

3.3 STAKEHOLDERS

The Commission's text proposes that sandboxes must facilitate the involvement of other relevant actors within the AI ecosystem, including the public and private sectors. The Parliament version of the Act [4] gives examples[6] such as notified bodies and standardisation organisations, SMEs, start-ups, enterprises, innovators, testing and experimentation facilities, research and experimentation labs and digital innovation hubs, centres of excellence and individual researchers. This is a much wider and more complex set of stakeholders than participated in the FinTech-generation regulatory sandboxes.

The AI Act was originally intended to primarily put obligations on the AI technology providers and still does in large parts of the text. However, various factors have influenced the interpretation of this during the development of the regulatory framework.

Firstly, it has become clear during the last year that the market has shifted towards foundation models [36] [37]. These models are redistributed business to business and then retrained by AI deployers, thus inheriting most legal obligations directly. It is likely that an AI deployer making a minor modification to a purchased foundation model would then be considered an AI technology provider, incurring many additional obligations.

Secondly, international technical standards have progressed. They now describe [38] the difference between an AI Technology Provider, AI Data Provider, AI Platform Provider, AI Integrator, AI User/Operator and AI Subject. The obligations in the AI Act, and likely supporting standards body of work, are starting to look a lot more distributed.

The AI standardisation community has developed an AI management system [39] and impact assessment for deployers to apply. Deployers bear the brunt of the obligations in standards as they can pass obligations down the supply chain, and only they can control important risk management techniques (guardrails) like human-in-the-loop strategies, operator training, testing of inputs, data pipelines and ongoing monitoring. In progress standards on topics like AI bias place explicit technical requirements on different AI stakeholders, including AI deployers.

While none of these standards are yet harmonized by the European Commission, if emerging consensus on technical best practice in standards is following a more nuanced and defined approach than the legal text, the standards will be difficult to ignore on the ground in an enforcement context.

Given this complexity, it needs to be clarified that the AI sandboxes should not limit themselves to AI providers and SMEs. There is plenty of need for AI data providers, deployers[7] and other

[6] Article 53a

[7] AI deployers are the organisations under whose authority the system would be operated. They are also sometimes referred to as AI users

stakeholders to participate. For example, one study in medical AI applications states that *'From a regulatory perspective, the performance of AI-based systems should be tested under real-world conditions in the hands of the intended users and not as stand-alone devices. Only then can we expect to rationally adopt and improve AI-based decision support and to accelerate its evolution.'* [40].

Civil society stakeholders have also called for different objectives from the proposed AI sandboxes:

- The European Digital SME Alliance has called for sandboxes to provide individualised guidance to SMEs on whether particular systems should be classified as high-risk [41].
- The European Trade Union Confederation has called for the use of AI in the workplace to be excluded from regulatory sandboxes [42].
- BCS, The Chartered Institute of IT in the UK, recommended to the UK government that the use of AI sandboxes should be encouraged beyond a purely regulatory need - for example, to test the correct skills and registration requirements for AI assurance professionals and how best to engage with civil society and other stakeholders. [43].

The Parliament amendments [4] introduce a new Article 29a requiring an impact assessment for high-risk AI systems. Within that text, it also requires deployers to make efforts to involve representatives of the persons or groups of persons that are likely to be affected by the high-risk AI system. These include equality bodies, consumer protection agencies, social partners and data protection agencies.

3.4 LEGAL BASIS

It is not currently clear what the legal basis for the sandbox will be. This is because the Commission version of the AI Act [2] differs significantly from the Parliament's recent amendments [4]. The Parliament version mandates at least one sandbox per member state and provides a legal basis for the AI sandbox, whereas the Commission version does not require this, and may require additional national legislation [30].

Parliament's version [4] sets out that participants remain vulnerable to liability legislation with no exemptions. However, provided they follow guidance, no administrative fines shall be imposed by the authorities.

However, it is notable that the AI Act still needs to be agreed upon, and yet at least one jurisdiction has launched a sandbox in advance of the final text [44]. Even if the Act were finalised, it is not clear that it would be at all easy to determine compliance. For example, Article 10 of the AI Act [2] contains the following requirement:

> *Training, validation and testing data sets shall be relevant, representative, and to the best extent possible, free of errors and complete. They shall have the appropriate statistical properties, including, where applicable, as regards the persons or groups of persons on which the high-risk AI system is intended to be used. These characteristics of the data sets may be met at the level of individual data sets or a combination thereof.*

Many will want to know more precisely what this means given a particular context. There are three routes to obtain clarity, for a standard to be granted harmonized status, for a regulator to provide guidance, or for a court to make a decision based on the intent of the law.

Standards provide best practices, and there are standards that may in future develop into harmonized standards on AI bias by implementing a management system [39], specific bias treatment measures [45] and standardised data quality measures [46] that can be applied to training data. This may be sufficient if the EC agrees this provides a presumption of conformity, but it only covers due process rather than outcomes.

The important matter is really whether the trained, resulting system demonstrates sufficiently equivalent accuracy across, for example, demographic groups. We want to know that the AI system does not show unwanted bias; the data used for training isn't necessarily important. But how sufficiently accurate is sufficient? There is no way to answer this without understanding the system's context or intended purpose. An organisation will need to conduct a risk and impact assessment to support a view that fundamental rights and values would not be unduly affected. However, it is difficult to see how a regulator, court or sandbox could accept or reject that view without an adequate and comparable alternative.

A separate but related concern is that meeting the concerns of all stakeholders may not actually be possible [47], and tradeoffs may ultimately have to be made between accuracy for the average person and accuracy for a minority group.

Looking at AI sandboxes mobilising around the world, they generally work within existing data privacy regulation and enforcement ecosystem as legal basis. With some exceptions:

- The AI sandbox in Germany was conducted in 2019 and represents a very specific use case to test autonomous delivery robots with an exemption from two road and vehicle related regulations [35]. This specific sandbox lasted 7 months and offered a test bed for a specific innovation.
- In Russia, a federal law in 2021 [48] enabled experimental legal regimes to be established. The conditions for these regimes includes the necessity that the current regulation contains restrictions that impede innovation. It is envisaged to be effective in the context of different regulators. It establishes a time limit that cannot be more than three years, and

applies within a specified territory. This follows a Moscow sandbox [49] started in 2020 and open to those who *engage in the development, creation, introduction, implementation or sale of artificial intelligence technologies or individual goods, works or services based on them.*

- Norway [50] and Denmark [51] have also stated an intent to also use the EU's Ethics Guidelines for Trustworthy AI to help with decision making in their early data privacy sandbox focused on AI [52].

3.5 METHODS

According to the Commission's text, the detail of the sandbox modalities will be deferred to implementing acts. Still, it is notable that the latest Parliament text requires competent authorities to provide guidance to achieve compliance with the regulations. It goes further in requiring them to provide guidance on identifying risks, and to test and demonstrate the effectiveness of those mitigation measures for those risks. On the other hand, if risks to fundamental rights, democracy, rule of law, health and safety or the environment cannot be mitigated immediately - the sandbox operators can suspend the testing process and the participation in the sandbox.

Sandbox authorities are also required to cooperate within the framework of the AI Office[8]. The Parliament text provides more detail about the reporting processes than the Commission version, and importantly requires detailed implementation reports to be published online.

Parliament's version sets out that participation in the AI regulatory sandbox is limited to a period appropriate to the project's complexity and scale.

In order to create a structured regulatory learning feedback loop, the Parliament's version also states that if a participant complies with the guidance given to them during the sandbox, then on exit they receive presumption of conformity. This is documented in an exit report, which market surveillance authorities and notified bodies are required to take into account in future conformity assessments.

One of the derogations that the AI Act allows for explicitly is concerning the privacy rights of individuals. It allows personal data to be processed solely for the legal purpose of the sandbox. In effect, no consent is required. This derogation is only allowed in specific situations with specific risk treatments. Still, it is also unclear how it could apply before the AI Act takes force.

[8] This is referred to as the AI Board in the original Commission proposal

UK ICO SANDBOX: ONFIDO

Onfido provides remote biometrics identification software, using AI, to end clients such as financial services. From July 2019 to August 2020 it worked inside the ICO sandbox to successfully mitigate bias risks in its solution [53].

The regulator did not provided any advice for the technical mitigation of the risk, but provided significant regulatory comfort that enabled it to mitigate risks. Specifically, the ICO gave Onfido confidence that although race labels necessary for bias testing were special category data, they could rely upon the public interest purpose for processing.

This is an example of a legal grey area, that when presented in a sandbox enabled a company to improve the trustworthiness of its AI solution.

Additionally, the Act calls for national data protection authorities to be associated with the operation of the AI sandbox, where the systems are processing personal data. Both versions of the text also add a public interest derogation, data that was collected for another purpose can sometimes be processed in the sandbox without an additional legal basis (e.g. consent). The situations where this can be done vary between the versions, however, public health and safety, environmental protection and critical infrastructure resilience are included in both. If this derogation is relied upon then there are a number of other requirements relating to security, auditability,

AI sandboxes that are operated now by data privacy regulators are inconsistent in methods either exempting participants from enforcement measures using no-action letters [32], exempting participants during the development phase [50] only, and providing no exemptions and instead working through co-creation [33] [50] [54].

Consistently all regulatory sandboxes aim to give individualised guidance in some form.

3.6 JURISDICTION

Jurisdictional issues are likely to be frequent following implementation of the AI Act.

Firstly, there may be variations in the interpretation of the AI Act between member states competent authorities. Secondly, and more importantly, there are many other pieces of legislation that may be relevant. In Austria, for example, there are medical laws unrelated to AI that may affect the use of AI services [30]. AI innovation may not simply need to comply with the AI Act, but indeed consider changes to other norms, conventions and laws. Finally, in many cases data privacy authorities will be just as a relevant as competent authorities under the AI Act

The AI Act encourages cross-border work within the European Union but makes little refer-

ence to any international or sector-specific collaborations. Key to such international collaboration will be the speed at which harmonized European standards are available and the degree to which they diverge from international standards[9]. This mechanism of globalised multi-stakeholder standardisation is a good way to collaborate internationally.

HORIZONTAL VERSUS VERTICAL STANDARDS

Harmonized standards, and sometimes regulation, are often described as horizontal and vertical. Verticals in this context are specific industries or applications.

For example, examining the harmonized standards published in relation to the EU's Machinery Directive [55], a general horizontal standard is *ISO 12100 - Safety of machinery - General principles for design - Risk assessment and risk reduction*. More specific requirements are sometimes included in vertical standards such as *ISO 10218-1 - Robots for industrial environments - Safety requirements - Part 1: Robots*.

It is difficult, if not impossible, to specify detailed requirements for topics like accuracy and oversight in horizontal standards. However, terminology, processes and metrics can be defined horizontally.

Sector-specific collaborations, for example, for connected and autonomous vehicles [35], would seem like clear opportunities that may benefit from special focus later.

HEALTHCARE SANDBOX IN SWEDEN

In Sweden [34], two healthcare providers wanted to evaluate jointly training and exchanging machine learning (ML) models for predicting readmission of heart failure patients. It was unclear whether there was a legal basis for exchanging data. The regulator determined that the data probably could not be shared if it was secret.

Further research [56] on this use-case has been conducted focussing on how privacy enhancing technologies such as fully homomorphic encryption could assist with this problem.

3.7 THROUGHPUT

The Commission proposal contains no specific guidance about the number of sandboxes or their throughput. The Parliament version [4] requires that Sandboxes must keep up with the

[9] I.e. from ISO/IEC

demand for participation, and that they must also be free of charge to SME[10] participants. At least one sandbox is required operate to similar principles in each member state, and the number of startups producing high-risk AI is expected to increase significantly. Therefore, the proposed AI sandbox is likely to be the largest-scale regulatory experiment ever conceived.

While budget funding, expert resourcing and sufficient training will likely be the key factors that affect the scale of sandboxes, driving those factors is likely to be the appetite of a member state or regulator to drive industry innovation, in contrast to producing regulatory clarity. Clearly, that same budget and resources will be required to support national supervisory authorities, notified and conformity assessment bodies - potentially creating a resource bottleneck across Europe.

All the sandboxes in operation run between three and twelve months and have four to eight participants in each cohort.

In terms of estimating the amount of work required in sandboxes, it is worth considering that existing AI sandboxes focusing on data privacy regulations go into far less detail than can be expected of AI Act related sandboxes. For example, in the UK [53] a regulator appears to have focussed more on helping answer a particularly regulatory question, rather than conducting a full technical assessment of the AI system.

3.8 TIMING

The AI Act is clear that sandboxes apply before a system is placed on the market, however that is a legal concept. Placing on the market could include, as previous drafts did, an extended 'real-world testing' mechanism. The Coreper draft [57] of the AI Act included a whole Article on real-world testing and laid out a series of requirements including informed consent. Importantly, this was considered to be separate to a regulatory sandbox. The Parliament amendments [4] have not referenced this Article and it's current state is unclear. It could be that real-world testing is not permitted, or it could be that real-world testing can be conducted in sandboxes without officially placing the product on the market.

In many cases, it is questionable how much a sandbox could really achieve without real-world end-users. If the non-real end-users are not representative of the final population, there is a possibility that the sandbox may not identify risks to certain types of consumers, who were not early adopters. Even if real end-users are participating, it will be necessary to ensure that they are representative of the actual target population.

Notably, the AI sandbox in Norway [50] says that they specifically include the ongoing imple-

[10] Small and medium size enterprises, defined as employing less than 250 persons. They should also have an annual turnover of up to EUR 50 million, or a balance sheet total of no more than EUR 43 million.

mentation of AI systems. While it is not ruled out by other sandboxes, most focus on product development (ex-ante).

3.9 SUMMARY COMPARISON

The following Table 3.9 provides a high-level comparison of historical sandboxes with both texts of the AI Act. Further information on the AI Act differences is available in the Appendix.

	HISTORICAL SANDBOXES	COMMISSION	PARLIAMENT
Entry criteria	Can require novelty, societal benefit, existing licensing, particular size/type of company	Must be transparent, fair and open. Implementing acts to provide more detail	Limits participation to AI providers. Requires easy access at Union level. Specifies more about the process. Free entry for SMEs.
Stakeholders	Typically limited to one participant and one or more regulators	Not specified	Notified bodies, standardisation bodies, TEFs, researchers etc.
Legal basis	A sandbox can be created by a legislative body in new regulation itself or by an existing regulator with existing regulation.	May require additional national legislation	Mandates at least one sandbox per member state and provides a legal basis for the AI sandbox.
Methods	No-action letters Individual Guidance	Methods largely to be set out in implementing acts.	Provide guidance on identifying risks, demonstrate the effectiveness of mitigation measures. Provide presumption of conformity on exit. Public transparency of exit reports.
Jurisdiction	Usually national or regional	Encourages cross-border work	Greater role for co-operation through the AI Office
Timing	Ex-ante and ex-post	Ex-ante	Ex-ante
Throughput	Typically small-scale	Optional	Mandatory per state and must keep up with demand

TABLE 1: COMPARISON OF HISTORICAL SANDBOXES WITH THE AI ACT

4 SANDBOX INNOVATION

Although sandboxes themselves are a legal innovation, by involving more stakeholders and technology to implement the regulation, further innovation opportunities can be identified. This section addresses such innovations.

4.1 INNOVATION INSIDE THE SANDBOX

We looked before at the different types of characteristics of sandboxes in a taxonomy. Now, using Figure 3 we can explore how different potential innovations can affect those characteristics.

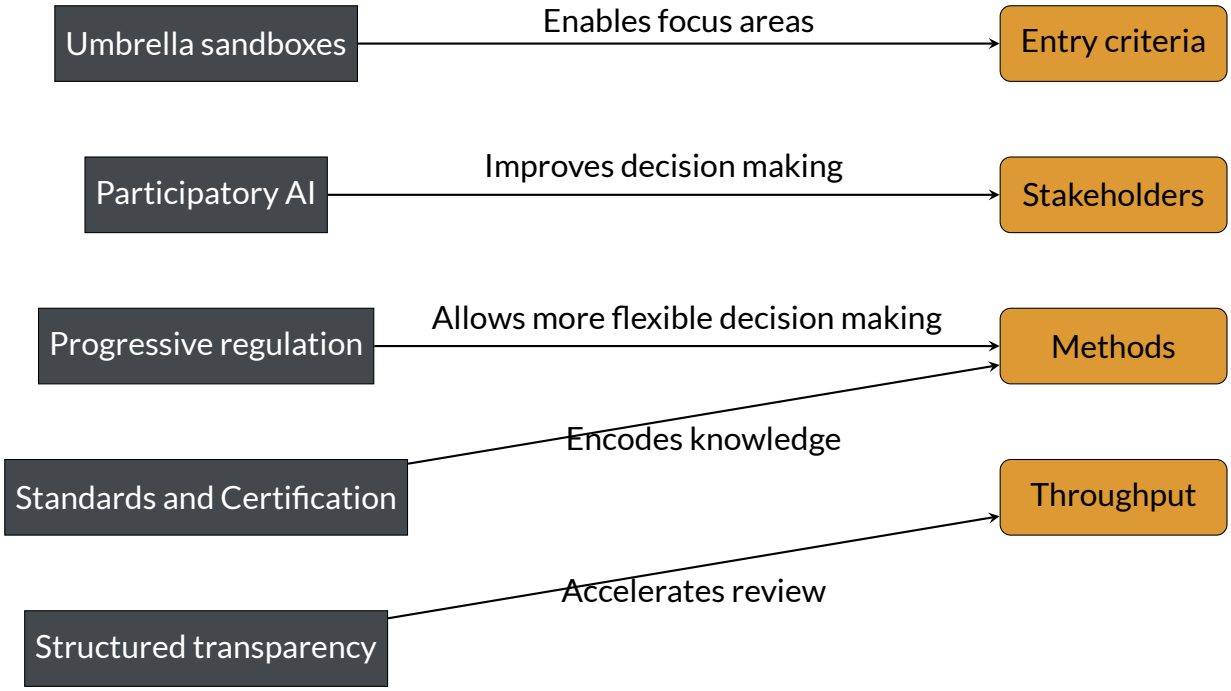


FIGURE 3: MAPPING SANDBOX CHARACTERISTICS TO INNOVATION MEASURES

4.1.1 STRUCTURED TRANSPARENCY

A recent Ada Lovelace report about the AI supply chain identifies [58] that there are not only many types of accountable stakeholders in the supply chain of AI systems but also many different potential accountability configurations. It concludes that transparency mechanisms between the stakeholders are essential for AI regulation.

A recently published ISO/IEC Technical Report [59] covering functional safety engineering and AI (both of which are horizontal topics) places a lot of importance on explainability, in that it is

key to using existing safety standards with AI. This is because explainability allows for scrutiny of how the system is working. However, many systems will not be able to rely on explainability. Instead, many will turn to governance processes, verification and validation procedures, and ongoing system monitoring.

Governance procedures must include understanding the provenance of the training data and the processes used to train ML models. This is a particularly relevant piece of information that is likely to be needed to be exchanged. Various initiatives, such as Google's Model Cards [60] and the UK's algorithmic transparency standard [61], attempt to address this, but no standard has yet emerged. Hugging Face, the open source (but for-profit) AI provider has developed additional guidance on model cards, including machine-readable versions [62].

Various groups have called for more advanced, technically-led regulatory innovation:

- A report by the Future Society in 2020 called on the European Commission to integrate both ex-ante and ex-post compliance mechanisms in the same governance system, design shared testing facilities, and use agile approaches [63].
- Work completed by the Global Digital Foundation describes [64] an information flow in the chain of assurance, passing an 'assurance file' between actors to provide information from the development and use of the AI system.
- Researchers in Finland [65] also propose that sandboxes utilise ML monitoring throughout the sandbox AI system lifecycle to facilitate continuous experimentation and learning.
- Some early research [66] investigates how privacy enhancing technologies can be combined with structured transparency to enable AI governance.

The AI Act contains an article about record keeping and logging, and the European Commission, in the AI standardisation request [67], has asked CEN/CENELEC, as the lead European Standardisation Organisation, to design technical standards supporting logging. It is possible to envisage that the amount of data available about the behaviour of AI systems will grow large and be monitored in an increasingly real-time manner. In fact, in Article 54, which covers the use of data in the public interest, mentions effective monitoring to mitigate risks.

A large-scale transparency system for high-risk AI systems could involve the continuous observation, assessment, and oversight of AI deployments at a significant scale. Such a system would aim to detect potential issues, ensure compliance with regulations and ethical guidelines, and maintain accountability. Key components and considerations for a large-scale monitoring system for AI systems could include:

- **Data Collection:** The monitoring system would gather data from deployed AI systems,

which may include input data, output predictions, system logs, performance metrics, user feedback, and any relevant contextual information. This data serves as the basis for monitoring and analysis. These systems may also collect ground truth, where possible.

- **Model Performance Evaluation:** Regular evaluation of AI model performance is crucial. The monitoring system can employ techniques like A/B testing, model comparison, and statistical analysis to assess model accuracy, robustness, and generalisation across different data sets and scenarios. Continuous monitoring ensures that models maintain high-performance levels and can identify degradation or concept drift over time.
- **User Feedback and Complaints Handling:** Incorporating user feedback and complaints handling mechanisms allows for a user-centric monitoring approach. The system can collect and analyse user feedback, complaints, or reported incidents to understand user experiences and address any issues promptly. This feedback loop helps improve the AI system's usability, effectiveness, risk management process and adherence to user expectations.

Considerations around data privacy, security, scalability, and resource allocation would be crucial for the successful implementation of such a system.

Key innovation questions that need to be addressed in this area are:

- Should mechanisms for transparency be standardised? Or be industry or open-source led?
- Should standardised transparency mechanisms reach into the development process, into the governance processes, or even into real-time operation?

4.1.2 PROGRESSIVE REGULATION

Looking at FinTech again, some research [22] identifies an option for better regulation. It identifies some market trends that may affect future regulatory sandboxes:

- Increasingly diverse geographical distribution of startups and increasing scale.
- Increasing opportunity for automated regulation driven by technical innovation.
- Increasing amounts of data available to inform regulation.

The same research encourages innovation in how regulatory sandboxes are implemented. Innovation could lower the barriers to entry, as more innovators could participate. For example, innovation could reduce information asymmetry and increase transparency. It also suggests a staged(progressive) implementation of regulation:

- Providing non-binding advice through some form of innovation hub
- Providing a testing and test and experimentation facility
- Providing a regulatory sandbox (or umbrella sandbox) which widens the scope of testing
- Providing a restricted licensing scheme while innovators grow their revenue and then progression to a full license.

Using this approach, the regulatory complexity and cost should therefore grow in proportion to the risk and revenue, embedding proportionality.

FIGURE 4: PROGRESSIVE REGULATORY FLOW

Some AI-specific policy research [68] suggests that sandboxes can be viewed as ex-ante and ex-post placing on the market. It envisages a virtual sandbox that supports the evaluation of conformity of the AI-based system with regard to technical specifications, horizontal and vertical regulation, and ethical principles in a controlled and limited testing environment. Once conformity has been verified, sandboxes can be used to interface with the deployed AI-based asset via the established monitoring plan, so that information about its post-market functioning can be collected and processed. This information is used by the national supervisory authority to evaluate compliance.

One example of why this may be necessary is that AI systems can continue to learn after being placed on the market[11]. In the USA, the FDA[12] has issued draft guidance[69] for pre-determined change control plans for ML enabled medical devices. These anticipate modifications in advance and determine them to be safe, removing the need for new submissions to the regulator when the model is updated.

Other research has highlighted the importance of auditability of AI systems given the envisaged focus on certification in the implementation of the AI Act [68]. Yet further research suggests that a focus on governance and management practices in sandboxes will help small businesses move from development to concrete implementation [70].

[11] This can be the result of scheduled retraining based on new data observed, reinforcement learning, or other approaches.

[12] Food and Drug Administration

4.1.3 STAKEHOLDER INTEGRATION

A Nesta report on Anticipatory Regulation [71] differentiates adaptive and anticipatory regulation. Anticipatory regulation brings more inclusion and engagement than a classical regulatory sandbox.

Anticipatory Regulation includes a wide variety of stakeholders, many of which are directly involved in the research and evidence building activities. Autonomous vehicle (AV) testbeds, for example, involve the coordinated actions of regulators, local authorities (often cities and regional governments), research institutions and technology companies.

This section outlines some ways that different stakeholders can be integrated into the sandbox process.

UMBRELLA SANDBOXES

The FCA called for the concept of an umbrella sandbox to be set up by the private sector as a non-profit [13] authorised by the regulator. The umbrella could act as a regulated entity, removing a barrier to entry into financial services. Rather than just using real data, data-sharing agreements and the use of privacy-enhancing technologies could deliver significant additional innovation testing benefits.

Disposing of the need for licensing may less relevant to the AI Act in some sectors. Data-sharing agreements between innovation partners could be helpful, especially if it was driven by technology and shared amongst market participants. These might be especially relevant at the sectoral level.

SINGAPORE'S AUTONOMOUS VEHICLE INITIATIVE

Singapore created an industry committee with public and private sector members to oversee integration of autonomous vehicles after the Land Transport Authority gave greater flexibility around transport laws to test AVs on public roads.

A test and experimentation facility was also created to improve AV technology in both a live and laboratory environment.

One benefit of umbrella sandbox is that they can potentially target particular use cases. For example, an industry actor could wish to enable innovation in a specific domain with related datasets and regulatory challenges. This domain might benefit from many different applications of AI, and different companies could collaborate, combining aspects of innovation hubs, testing and experimentation facilities and a regulatory sandbox.

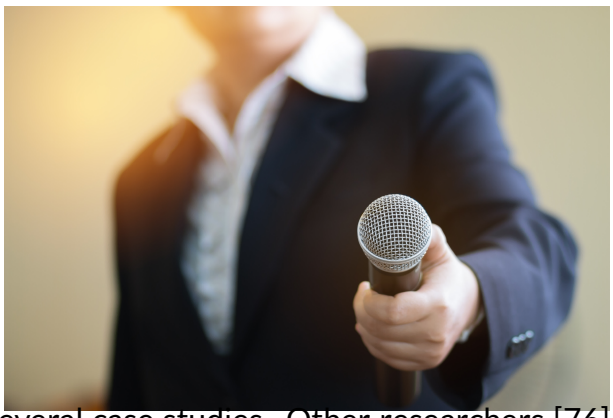
For example, medical diagnostic datasets might be identified that multiple companies would

benefit from access to. An umbrella sandbox could establish a data trust as a service, [72] discuss with regulatory actors appropriate legal methods for data sharing, or appropriate privacy enhancing technologies that could be applied. This would then enable datasets to be provided, similar to a testing and experimentation facilities. Analysis with suitable stakeholders about the risks associated with the datasets could be conducted in both a centrally-enabled and use-case specific manner. Where innovative solutions are further developed, engagement between the innovator and the regulator is facilitated by the umbrella sandbox.

This approach supports innovation in both AI and regulation, in a way where efficiency is maximised and thus cost is minimised for all parties.

PARTICIPATORY AI

Participatory AI is the involvement of members of the public in an AI project or intervention, incorporating perspectives and experience. Public participation is an approach used in other fields [73] and is particularly relevant given the concern about the societal impact of AI.



A wide range of participatory approaches are possible. The Norwegian AI sandbox [74] focussed on openness and published plans, insights and examples to the public. The sandbox should benefit the public and the market, not only AI technology providers. textbf A Nesta report [75] from 2021 proposed a more integrated operational framework for public stakeholder engagement in developing AI, including several case studies. Other researchers [76] conclude that participatory AI is hindered by corporate profit motives and concerns over corporate exploitation, suggesting it may be more effective outside of a corporate environment.

CARE QUALITY COMMISSION - AI REGULATORY SANDBOX

The CQC ran three sandboxes in 2019/2020 [77] focused on AI digital triage, screening and diagnostics, and personal assistants. They found it valuable to be able to draw upon the contributions of people with lived experience of care. This helped keep a focus on enabling services to provide the right care for the people who need it.

Another approach used in European standardisation is that trade unions, consumers and SMEs are given a funded voice at the table of relevant technical standards projects. This ensures some level of public participation when reaching a consensus on soft law in the form of standards. The same approach could be used in sandboxes, drawing both public views and alternative stakeholders into the process.

At the moment civil society and social stakeholders are barely mentioned in the AI Act in the

context of sandboxes. Even the Parliament version only mentions their involvement in the context of AI solutions that are specifically intended to provide socially and environmentally beneficial outcomes.

STANDARDS AND CERTIFICATION DEVELOPMENT

Regulatory sandboxes can provide valuable benefits to the development of international and/or harmonized standards in several ways:

- Regulatory sandboxes serve as testing grounds which enable the identification of regulatory gaps, ambiguities, or inefficiencies that may hinder the effective deployment of their technologies. Close observation and interaction with sandbox participants allow regulators to gain insights into these regulatory issues. This first-hand understanding of the challenges participants face can inform discussions on the need for international or harmonized standards to address these gaps.
- The experiences and insights gained within regulatory sandboxes can be shared with relevant standardisation bodies. Sandbox participants, including technology developers, industry representatives, and regulatory authorities, can contribute to standardization discussions by providing feedback, case studies, and lessons learned.
- Regulatory sandboxes can also serve as testing grounds for international standards in development. By piloting or implementing international standards within the sandbox environment, regulators can assess their applicability, effectiveness, and practicality. This testing phase allows for iterative improvements and refinements to the standards, ensuring they are fit for purpose and adaptable across different jurisdictions.

In addition to standards development organisations, other organisations can develop certification mechanisms that support the AI Act, such as ForHumanity. The sandbox provides a platform by which these certification mechanisms can be trialled in the context of the regulation to see if they can provide alternative mechanisms of compliance.

4.2 RELATED INNOVATION ACTIVITIES

In this section, we discuss innovation activities that are not directly in scope of regulatory sandboxes, but are related. By using innovation hubs to disseminate non-binding advice, demand for regulatory sandbox support can be reduced. By using test and experimentation facilities to establish industry benchmarks, decision making in sandboxes can be improved.

4.2.1 REGTECH

The growing pace of technical innovation is driving more granular and bespoke regulation that is being supported by a new area of innovation, RegTech.

Generally speaking, RegTech solutions are software or technology-driven tools that help organisations automate and streamline their regulatory compliance processes. These solutions leverage technologies such as AI, ML and big data analytics to address regulatory challenges efficiently.

REGTECH EXAMPLES INCLUDE:

- Solutions for transaction monitoring leverage advanced analytics and anomaly detection techniques to identify suspicious activities and potential financial crimes. They analyse transactional data in real-time, flagging suspicious patterns or unusual behaviours for further investigation and helping organisations comply with AML regulations.

- Solutions for regulatory reporting automate the collection, validation, and submission of required regulatory reports to relevant authorities. These solutions ensure accurate and timely reporting while reducing manual efforts and improving data quality.

- Solutions for risk management to assist organisations in identifying, assessing, and managing regulatory risks. These solutions use data analytics and predictive modelling to analyse risks, monitor compliance gaps, and provide real-time risk insights. They help organisations pro-actively mitigate risks and maintain compliance.

RegTech has historically been driven by FinTech regulatory innovation, but many solutions are coming to market to help with these goals in the AI domain. As AI systems are continually being updated, it makes sense that technical solutions to monitor them should be used in order to achieve the throughput of regulatory events required. In order to drive RegTech innovation for AI compliance, it could be useful to have a text and experimentation facility dedicated to AI compliance tools themselves.

4.2.2 INNOVATION HUBS

Other types of innovation facilities exist in addition to regulatory sandboxes. However, the literature is not consistent in its terminology describing them as innovation hubs, testing and experimentation facilities and data spaces. Innovation hubs do not provide the same level of individualised or binding guidance compared to sandboxes, therefore significantly reducing the

cost and risk for the regulator.

Some researchers [78] define an innovation hub as simply a portal for a regulatee to obtain a non-binding response from a regulator. The same researchers recommend an innovation hub be run in parallel with the regulatory sandbox. Concluding that it supports innovation and regulatory learning better than a regulatory sandbox and at a greater scale.

While the EU has established a similar concept of digital innovation hubs, they do not include regulatory guidance within their scope. Nor is an innovation hub likely to produce regulatory change in the same way a regulatory sandbox might. Nevertheless, properly funded and managed to provide advice on compliance to innovators may help reduce the financial burden of complying.

Both versions of the text suggest that at least SMEs should be given access to guidance on the implementation of the regulation. The wording in the Parliament version is stronger and more specific about the way this should be achieved.

4.2.3 TESTING AND EXPERIMENTATION FACILITIES

The FCA in the UK offers interesting innovation facilities focussing on providing synthetic, public or anonymised high-quality financial data sets and over 1000 APIs[13] [6]. This can also be thought of as a 'data space' or a 'testing and experimentation facility' - a secure environment that pools resources together. SMEs, technology providers, regulators or governments can also create these to test capabilities on datasets [17] and dedicate or create physical environments [35].

This concept can be extended beyond regulatory innovation to truly supporting innovators with a platform to experiment, validate, and refine their technical solutions while adhering to specific regulatory requirements and guidelines.

A testing and experimentation facility can also involve the creation of simulated environments or testbeds where participants can deploy and test their software, systems, or prototypes. The sandbox environment often includes specific datasets, simulated user interactions, or simulated production environments to mimic real-world scenarios. This can be achieved through data, for example, the FCA sandbox might offer synthetic financial transactions that AI anti-money laundering innovations can be tested with. In turn, these activities can lead to bench-

[13] Application Programming Interfaces are procedures that allow the creation of applications that access the features or data of another application

marking, that is, establishing standards of performance for a use case. In this example, it is clear that regulatory experimentation to support innovation drives secondary innovation that helps enable regulation.

The EU launched four projects to build similar facilities in 2023 [79]. These four projects will build environments supporting healthcare, agriculture, manufacturing and smart cities.

ML benchmarks are pairings of performance metrics and datasets applied to a specific algorithmic objective. Given any ML model, it is probably possible to select a dataset and a performance metric that shows it performs well. Standardised benchmarks, for example for the detection of malaria species from blood smears can be built upon specific open datasets [80].

Benchmarks can enable regulators to evaluate whether a solution is providing state-of-the-art compliance with the objectives of the AI Act, particularly around accuracy. An illustrative example of this is included later in this report in section 5.2.

5 THE POSSIBLE IMPACT OF THE SANDBOX

5.1 MARKET SIZING

In this section we look at the likely number of AI applications that will be subject to the AI Act, and the likelihood they will seek advice from a Sandbox. A number of factors affect this:

The number of AI systems in scope of regulation
It is hard to determine the full impact of the AI Act, in part as we do not know in full how AI will be used. For example, all toys aimed at children under 14 in Europe require conformity assessment under sectoral regulations. To what extent will toy manufacturers look to use generative AI to enhance toys?

The EU's own impact assessment [81] estimated that 5-15% of AI applications would be high-risk and calculated that the cost of compliance (over time) would be 4-5% of the spend on those applications.

The impact assessment assumes that the average spend on an AI application development is €170k. That would mean that their low estimate for AI spending in Europe in 2025 (€30 billion) represents hundreds of millions of AI systems. Even assuming only 5% of those are regulated, that is still more than eight million AI systems made in Europe that will be regulated.

The stakeholders that are regulated

The impact assessment is also assuming that the creation of an ML model in Europe creates an instance of compliance. In fact, the creation of an ML model anywhere that is intended to be used in Europe (or with European citizens as subjects) falls in scope.

It is also more likely the deployed AI system that is more likely to be regulated, not the ML model. Many requirements of the Act and its supporting standards will put significant obligations on deployers, as they are the ones who control the context of use and the overall risk management system.

As a result, the EU may have underestimated the likely demand from innovators for guidance in the sandbox framework, and other types of measures to support innovation.

The frequency of regulatory events
Under the AI Act, a system must re-undergo conformity assessment if it is in scope of Act, and is substantially modified. It is not yet fully understood how substantial that modification should be. The prevailing assumption is that the system must be changed in a way that puts it outside of the parameters or results from the previous conformity assessment. This would potentially include retraining an existing model, unless the previous conformity assessment had included assertions about the potential range of change resulting from retraining, and verified it.

The timing of the regulatory events
Speaking in abstract terms, least regulatory clarity will exist at the point until harmonised standards are published, this will increase when notified bodies start taking decisions or sandboxes start giving guidance. Then, when enforcement starts we can expect further clarity.

The further we are down the path from the current point, the lower the demand for sandboxes will be. As innovation hubs start to provide general and non-binding guidance based on clarity gained from external events, it will be less necessary to seek the individualised guidance a regulatory sandbox offers.

5.2 CHALLENGES FOR THE EU'S AI SANDBOXES

As currently planned, the EU's AI sandboxes will be the largest scale regulatory experimentation activity ever attempted. Scaling to meet the market demand will be a huge challenge. The difference between the number of AI systems being placed on the market and the number that existing sandboxes can handle is stark.

It also seems inevitable that one of the key challenges for all sandboxes will be getting the entry criteria and application process right. This will be important to ensure the goals of protecting competition and innovation are achieved, but also to ensure that novel innovations are prioritised. Only by consuming novel innovations or applications is regulatory learning achieved. Another key challenge will be efficiently codifying that learning in future regulatory guidance

or standards, so that the next organisation can get general guidance from an innovation hub rather than individualised guidance from a regulatory sandbox.

Any delay to harmonized standards could result in a very high demand for individualised guidance, coupled with a more expensive decision-making process for a regulator.

6 ILLUSTRATIVE CASE STUDY - FINANCIAL SERVICES

The Bank of England compiled ML case studies from the firms it regulates [82], and one of the most common areas of use was credit underwriting. ML is used to support lending decisions, typically as part of a wider scoring process and sometimes as a direct input to an automated underwriting process. In this illustrative case study, a European non-bank lender plans to use an internally developed ML model - without human supervision - to complete credit card underwriting in full. The ML model has access to the applicant's historical financial data through interoperability with multiple banks and is producing a prediction of whether the applicant will default on a payment in the first year.

The firm is aware that creditworthiness assessment is a high-risk use case in the context of the proposed EU AI Act[14] and that it needs to plan for compliance. The firm decides to pursue harmonized standards compliance and engages an external consultant. The consultant implements an AI quality management system based on a harmonized standard but also tells the firm that they also need to consider use-case specific risks in relation to accuracy and human oversight, and compliance is not gained automatically from implementing the quality management system. Harmonized standards for conformity assessment of creditworthiness models do not exist yet. However, the consultant is able to provide general requirements and guidance from cross-sector standards on testing, bias, oversight and quality. The firm engages with a national **innovation hub** for guidance - how can they show that they have sufficient accuracy and oversight? The innovation hub is easy to reach but is not able to provide answers. As there are no harmonized standards covering the use case, and there is no regulatory guidance to refer to yet, they recommend applying to the local **financial services AI sandbox**.

While waiting for a response from the sandbox, the firm reviews its internally identified risks. The accuracy of the model is sufficient for business purposes, however, it is 25% less accurate for women, which will result in some women being incorrectly denied credit. It is possible for an applicant to request a human review, but the firm believes it is unlikely many people will request this.

The firm is accepted into the sandbox and has its first meeting with regulators to scope the engagement. It is decided that the financial services regulator will take the lead. The key sand-

[14] Annex III 5.b

box objective for the firm is to identify how to determine whether 25% lower accuracy for one gender is acceptable to the regulator, the regulator also wishes to review the risk assessment to ensure it is complete. The regulator provides a letter confirming they will not take any enforcement action in relation to the sandbox discussions, and the firm provides the full risk assessment documentation, training data and test results from the ML model to the regulator to review.

Some time passes, and the regulator calls a meeting. This time they have added a **civil society stakeholder** who represents marginalised groups in financial services. The regulator explains that in their inspection of the test results, they were unable to understand how lending decisions were made in many cases. The civil society stakeholder believes that there is an unjustified correlation between seasonal income and a negative lending decision. The regulator also cautions that seasonal income could correlate with membership of a protected group and that this should be investigated.

The regulator then provides a short overview of a FinTech **test and experimentation facility** that includes synthetically generated historical financial data from across the industry. While the firm's intended way of using the ML model is novel, the application of predicting creditworthiness is not. The regulator is able to point to a benchmark across several other ML models that shows less than a 5% accuracy difference between genders. The firm is asked to investigate both of these issues.

A few weeks pass and the firm requests a follow-up meeting, including a senior data scientist. The data scientist explains changes that have been made to the model that have increased the likelihood of seasonal workers obtaining loans. They have also retrained their model using data from the test and experimentation facility and have been able to reduce the accuracy difference between genders to 5%. However, in doing so, the overall accuracy of the model has degraded by 10%, which compromises the business case.

The regulator indicates that the firm can still use the more generally accurate model, but it must explain the limitations clearly on the application website and encourage applicants to request a human review if they believe the decision was incorrect, and they must submit to ongoing monitoring of the review rate.

Considering this from a business perspective, the firm decides to proceed with the discriminatory but more accurate model. The regulator follows up in writing, providing approval for the firm to begin transacting, subject to the stated conditions, and **reporting back monthly on**

review statistics.

After six more months and five more meetings between the regulator and the firm, sufficient reviews are being requested by applicants that the regulator is convinced the appeal mechanism is providing a sufficient risk treatment. The firm **exits the sandbox with full approval** from financial services and AI regulators.

The financial services and AI regulators meet as a follow-up, and they **update regulatory guidance** for firms who are producing ML for certain purposes to compare performance to certain benchmarks. They also publish a report showing how oversight and redress mechanisms can assist with mitigating the impact of technical issues.

7 CONCLUSIONS

The EU's current proposed texts for the AI Act both lay out a vision for a network of sandboxes across the EU, each with different stakeholder groups. At the same time, this is a much greater ecosystem of stakeholders than in previously studied FinTech sandboxes. These will need to 'keep up with demand' and be free to SMEs, which could consume significant resources to deliver in practice. It is expected that regulators will struggle with expert capacity in the sandboxes.

Given the focus in the literature put on transparent entry criteria [29], it is notable that the current text confirms it will publish criteria but delegates the actual criteria to future implementing acts. The current text also speaks little to the methods of the AI sandboxes, only compelling the sandboxes to provide guidance.

Figure 5 below illustrates the key factors that this report highlights that can help ensure regulatory sandboxes have a positive impact on innovations, as highlighted throughout this report.

7.1 DEMAND

The proposal for sandboxes in the AI Act needs to also be viewed in the context of other mechanisms to place products on the market:

- An organisation can, in many circumstances, simply self-assess itself against harmonized standards. However, these standards do not yet exist, and the timeline is in question.
- An organisation can also submit to a conformity assessment review by a notified body, a process that is required for some high-risk use cases.
- An organisation could submit to an independent third party certification scheme that is itself in line with harmonized standards.

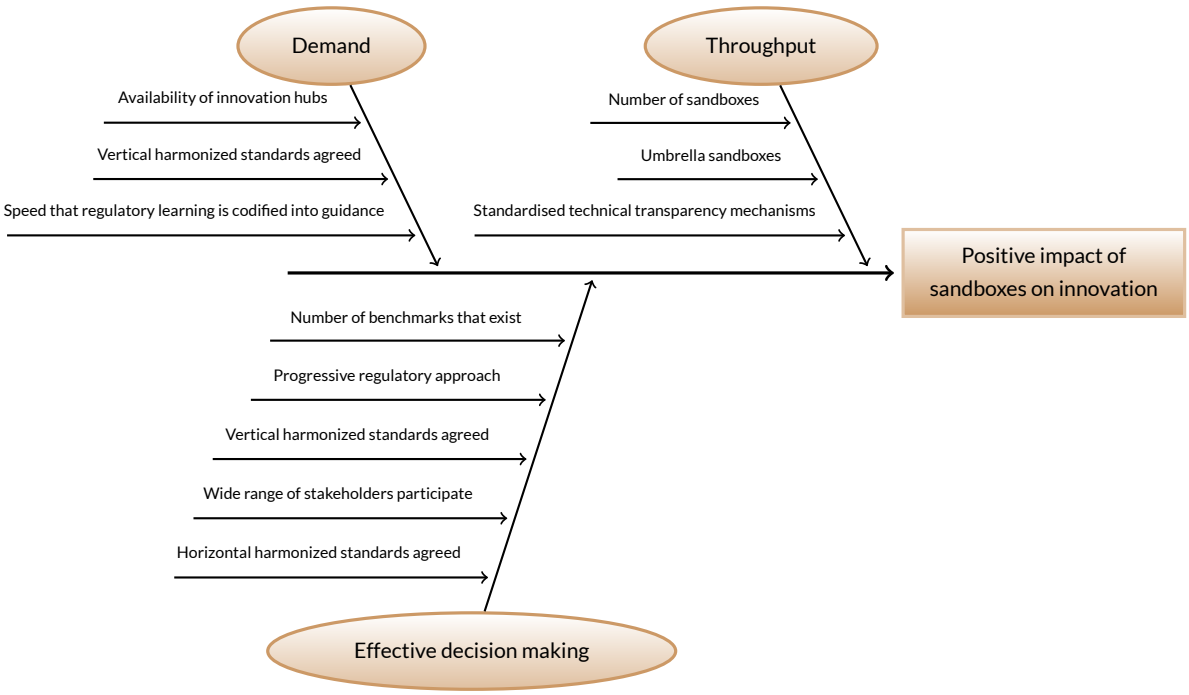


FIGURE 5: FACTORS DRIVING THE POSITIVE IMPACT OF AI SANDBOXES ON INNOVATION

- An organisation can simply put a product on the market, claiming compliance with the law.

Put in this context, it would seem unlikely that a sandbox would necessarily be a preferred route for innovators unless they are breaking new ground. However, if standards are delayed [83], it may be the cheapest option to get regulatory clarity. If this becomes the case, and sanboxes do manage to meet demand, then the viability and profitability of notified bodies could also be threatened [84].

At a national level, it is clear that competent authorities need to invest in national innovation hubs to accompany sandboxes - or other ways for companies to obtain non-binding or general regulatory comfort.
The availability and effectiveness such mechanisms, and the availability of vertical harmonized standards, are key to ensuring that AI providers and deployers can get advice quickly and cheaply. However, innovations hubs require regulatory guidance, which in turn requires sandboxes.
Without the implementation of vertical harmonized standards well in advance of the AI Act taking effect, demand for sandboxes will likely outstrip supply.

7.2 EFFECTIVE DECISION MAKING

The sandboxes are also likely to suffer from significant difficulty in making decisions, and the early lack of a complete set of standards will increase the pressure on this process. Regulators will likely be forced to seek guidance from competent authorities, which in turn may need to seek guidance from the envisaged AI Board, or its independent pool of experts.

To accelerate decision making decisions, and thereby overall efficiency, Europe should enable the provision of data into testing and experimentation facilities, leading to benchmarks of different solutions[a].

[a] See section 4.2.2

The provisions for sandboxes to test with real data and users should be clarified in the AI Act, and a progressive regulatory approach should be used[a].

[a] See section 4.1.1

Finally, a wide range of stakeholders should be involved to input into sandboxes. The Parliament text is preferred in this regard, however invitations should be extended to civil society groups in a wider range of scenarios[a] to make inclusive decisions. They also need to find ways to involve AI deployers into regulatory sandboxes.
Facilitating the involvement of standardisation experts directly in the sandboxes will also create positive feedback loops and reduce the cost of enforcement post-standardisation. This will enable regulators to understand the state of the art of best practices, and standardisation experts to find gaps in detail and coverage.

[a] See section 4.1.2

7.3 THROUGHPUT

Obviously, the number of sandboxes that are mobilised is key - and the Parliament text proposes to increase this by requiring each member state to initiate a sandbox. While establishing an effective enforcement regime at scale is an inevitable challenge, there are opportunities that regulators can seize. Innovative approaches can help with throughput and accelerating the objectives of sandboxes.

It could be possible to create a Testing and Experimentation Facility in relation to transparency of general AI system meta-data[a]. This facility could leverage the AI Act's logging requirements to drive RegTech innovations that support governance and regulation throughout the AI system lifecycle. This facility could also test assurance tools, methods and certification mechanisms from third parties. Ultimately, such innovations will reduce the cost of enforcement, and sandboxes, and enable a more progressive approach to regulation during product development.

[a] See section 4.2

Standardising the exchange of technical information between stakeholders in the complex AI ecosystem can significantly reduce effort and uncertainty in the assessment process. Solutions can range between standardised exchange of 'model cards', through to standardised monitoring procedures leveraging the logging requirements of the AI Act[a].

[a] See section 4.1

Creation of sandboxes that leverage third-party assessment schemes can alleviate resource demands[a]. Focus on particular sectors with common needs, potentially integrated with Test and Experimentation Facilities, can also produce innovative results.

[a] See section 4.1.3

Umbrella sandboxes led by industry stakeholders focussed on specific data or policy related areas may maximise efficiency and throughput for all parties[a].

[a] See section 4.1.3

8 APPENDIX: COMPARISON OF COMMISSION AND PARLIAMENT TEXTS

At the time of writing, there is a Commission [2] and a Parliament [4] version of the AI Act.

These differ in relation to sandboxes in the following normative ways. Except where converted to normative text, the recitals are not compared.

TOPIC	COMMISSION	PARLIAMENT
Sandbox definition Article 3(44g)	*Only defined in recitals*	A controlled environment established by a public authority that facilitates the safe development, testing and validation of innovative AI systems for a limited time before their placement on the market or putting into service pursuant to a specific plan under regulatory supervision
Mandate Article 53, para 1	AI regulatory sandboxes established by one or more Member States competent authorities or the European Data Protection Supervisor shall provide a controlled environment that facilitates the development, testing and validation of innovative AI systems for a limited time before their placement on the market or putting into service pursuant to a specific plan. This shall take place under the direct supervision and guidance by the competent authorities with a view to ensuring compliance with the requirements of this Regulation and, where relevant, other Union and Member States legislation supervised within the sandbox.	Member States shall establish at least one AI regulatory sandbox at national level, which shall be operational at the latest on the day of the entry into application of this Regulation This sandbox can also be established jointly with one or several other Member States;

Mandate Article 53 para 1(a) (new)	Not covered	Additional AI regulatory sandboxes at regional or local levels or jointly with other Member States may also be established;
Mandate Article 53 para 1(b) (new)	Not covered	The Commission and the European Data Protection Supervisor, on their own, jointly or in collaboration with one or more Member States may also establish AI regulatory sandboxes at Union level;
Throughput Article 53 para 1(c) (new)	Not covered	Establishing authorities shall allocate sufficient resources to comply with this Article effectively and in a timely manner;
Scope Article 53 para 1d (new)	Not covered	AI regulatory sandboxes shall, in accordance with criteria set out in Article 53a, provide for a controlled environment that fosters innovation and facilitates the development, testing and validation of innovative AI systems for a limited time before their placement on the market or putting into service pursuant to a specific plan agreed between the prospective providers and the establishing authority;

Objectives Article 53 para 1(e) (new)	Not covered	The establishment of AI regulatory sandboxes shall aim to contribute to the following objectives: (a) for the competent authorities to provide guidance to AI systems prospective providers providers to achieve regulatory compliance with this Regulation or where relevant other applicable Union and Member States legislation; (b) for the prospective providers to allow and facilitate the testing and development of innovative solutions related to AI systems; (c) regulatory learning in a controlled environment.
Methods Article 53, para 1(f) (new)	Not covered	Establishing authorities shall provide guidance and supervision within the sandbox with a view to identify risks, in particular to fundamental rights, democracy and rule of law, health and safety and the environment, test and demonstrate mitigation measures for identified risks, and their effectiveness and ensure compliance with the requirements of this Regulation and, where relevant, other Union and Member States legislation;

Methods Article 53, para 1g (new)	Not covered	Establishing authorities shall provide sandbox prospective providers who develop high-risk AI systems with guidance and supervision on how to fulfil the requirements set out in this Regulation, so that the AI systems may exit the sandbox being in presumption of conformity with the specific requirements of this Regulation that were assessed within the sandbox. Insofar as the AI system complies with the requirements when exiting the sandbox, it shall be presumed to be in conformity with this regulation. In this regard, the exit reports created by the establishing authority shall be taken into account by market surveillance authorities or notified bodies, as applicable, in the context of conformity assessment procedures or market surveillance checks;
Jurisdiction Article 53, para 2	Member States shall ensure that to the extent the innovative AI systems involve the processing of personal data or otherwise fall under the supervisory remit of other national authorities or competent authorities providing or supporting access to data, the national data protection authorities and those other national authorities are associated to the operation of the AI regulatory sandbox.	Establishing authorities shall ensure that, to the extent the innovative AI systems involve the processing of personal data or otherwise fall under the supervisory remit of other national authorities or competent authorities providing or supporting access to personal data, the national data protection authorities, or in cases referred to in paragraph 1b the EDPS, and those other national authorities are associated to the operation of the AI regulatory sandbox and involved in the supervision of those aspects to the full extent of their respective tasks and powers;

Methods Article 53 para 3	The AI regulatory sandboxes shall not affect the supervisory and corrective powers of the competent authorities. Any significant risks to health and safety and fundamental rights identified during the development and testing of such systems shall result in immediate mitigation and, failing that, in the suspension of the development and testing process until such mitigation takes place.	The AI regulatory sandboxes shall not affect the supervisory and corrective powers of the competent authorities, including at regional or local level. Any significant risks to fundamental rights, democracy and rule of law, health and safety or the environment identified during the development and testing of such AI systems shall result in immediate and adequate mitigation. Competent authorities shall have the power to temporarily or permanently suspend the testing process, or participation in the sandbox if no effective mitigation is possible and inform the AI office of such decision;
Methods Article 53 para 4	Participants in the AI regulatory sandbox shall remain liable under applicable Union and Member States liability legislation for any harm inflicted on third parties as a result from the experimentation taking place in the sandbox.	Prospective providers in the AI regulatory sandbox shall remain liable under applicable Union and Member States liability legislation for any harm inflicted on third parties as a result of the experimentation taking place in the sandbox. However, provided that the prospective provider(s) respect the specific plan referred to in paragraph 1c and the terms and conditions for their participation and follow in good faith the guidance given by the establishing authorities, no administrative fines shall be imposed by the authorities for infringements of this Regulation;

Jurisdiction Article 53 para 5 (new)	Member States' competent authorities that have established AI regulatory sandboxes shall coordinate their activities and cooperate within the framework of the European Artificial Intelligence Board. They shall submit annual reports to the Board and the Commission on the results from the implementation of those scheme, including good practices, lessons learnt and recommendations on their setup and, where relevant, on the application of this Regulation and other Union legislation supervised within the sandbox.	Establishing authorities shall coordinate their activities and cooperate within the framework of the AI office; Establishing authorities shall inform the AI Office of the establishment of a sandbox and may ask for support and guidance. A list of planned and existing sandboxes shall be made publicly available by the AI office and kept up to date in order to encourage more interaction in the regulatory sandboxes and transnational cooperation; Establishing authorities shall submit to the AI office and, unless the Commission is the sole establishing authority, to the Commission, annual reports, starting one year after the establishment of the sandbox and then every year until its termination and a final report. Those reports shall provide information on the progress and results of the implementation of those sandboxes, including best practices, incidents, lessons learnt and recommendations on their setup and, where relevant, on the application and possible revision of this Regulation and other Union law supervised within the sandbox. Those annual reports or abstracts thereof shall be made available to the public, online;

Entry criteria and methods Article 53 para 6	The modalities and the conditions of the operation of the AI regulatory sandboxes, including the eligibility criteria and the procedure for the application, selection, participation and exiting from the sandbox, and the rights and obligations of the participants shall be set out in implementing acts. Those implementing acts shall be adopted in accordance with the examination procedure referred to in Article 74(2).	The Commission shall develop a single and dedicated interface containing all relevant information related to sandboxes, together with a single contact point at Union level to interact with the regulatory sandboxes and to allow stakeholders to raise enquiries with competent authorities, and to seek non-binding guidance on the conformity of innovative products, services, business models embedding AI technologies; The Commission shall proactively coordinate with national, regional and also local authorities, where relevant;
Stakeholders Article 53 para 6 (new)	Not covered	For the purpose of paragraph 1 and 1a, the Commission shall play a complementary role, enabling Member States to build on their expertise and, on the other hand, assisting and providing technical understanding and resources to those Member States that seek guidance on the set-up and running of these regulatory sandboxes;

Entry criteria Article 53a (new) 1, 2a-c	Not covered	1. In order to avoid fragmentation across the Union, the Commission, in consultation with the AI office, shall adopt a delegated act detailing the modalities for the establishment, development, implementation, functioning and supervision of the AI regulatory sandboxes, including the eligibility criteria and the procedure for the application, selection, participation and exiting from the sandbox, and the rights and obligations of the participants based on the provisions set out in this Article; 2. The Commission is empowered to adopt delegated acts in accordance with the procedure referred to in Article 73, no later than 12 months following the entry into force of this Regulation and shall ensure that: a) regulatory sandboxes are open to any applying prospective provider of an AI system who fulfils eligibility and selection criteria. The criteria for accessing to the regulatory sandbox are transparent and fair and establishing authorities inform applicants of their decision within 3 months of the application; b) regulatory sandboxes allow broad and equal access and keep up with demand for participation; c) access to the AI regulatory sandboxes is free of charge for SMEs and start-ups without prejudice to exceptional costs that establishing authorities may recover in a fair and proportionate manner;

Stakeholders Article 53a (new)(d)	Not covered	regulatory sandboxes facilitate the involvement of other relevant actors within the AI ecosystem, such as notified bodies and standardisation organisations (SMEs, start-ups, enterprises, innovators, testing and experimentation facilities, research and experimentation labs and digital innovation hubs, centers of excellence, individual researchers), in order to allow and facilitate cooperation with the public and private sector;
Methods Article 53a (new) (e-g)	Not covered	e) they allow prospective providers to to fulfil, in a controlled environment, the conformity assessment obligations of this Regulation or the voluntary application of the codes of conduct referred to in Article 69; (f) procedures, processes and administrative requirements for application, selection, participation and exiting the sandbox are simple, easily intelligible, clearly communicated in order to facilitate the participation of SMEs and start-ups with limited legal and administrative capacities and are streamlined across the Union, in order to avoid fragmentation and that participation in a regulatory sandbox established by a Member State, by the Commission, or by the EDPS is mutually and uniformly recognised and carries the same legal effects across the Union; g) participation in the AI regulatory sandbox is limited to a period that is appropriate to the complexity and scale of the project.

Scope Article 53a (new) (h)	Not covered	the sandboxes shall facilitate the development of tools and infrastructure for testing, benchmarking, assessing and explaining dimensions of AI systems relevant to sandboxes, such as accuracy, robustness and cybersecurity as well as minimisation of risks to fundamental rights, environment and the society at large
Innovations hubs and TEFs	Not covered	Prospective providers in the sandboxes, in particular SMEs and start-ups, shall be facilitated access to pre-deployment services such as guidance on the implementation of this Regulation, to other value-adding services such as help with standardisation documents and certification and consultation, and to other Digital Single Market initiatives such as Testing & Experimentation Facilities, Digital Hubs, Centres of Excellence, and EU benchmarking capabilities;
Data privacy Article 54 para 1	In the AI regulatory sandbox personal data lawfully collected for other purposes shall be processed for the purposes of developing and testing certain innovative AI systems in the sandbox under the following conditions:	In the AI regulatory sandbox personal data lawfully collected for other purposes may be processed solely for the purposes of developing and testing certain AI systems in the sandbox when all of the following conditions are met:

Public interest Article 54 para 1(a)	the innovative AI systems shall be developed for safeguarding substantial public interest in one or more of the following areas:	AI systems shall be developed for safeguarding substantial public interest in one or more of the following areas: (ii) public safety and public health, including disease detection, diagnosis prevention, control and treatment; (iii) a high level of protection and improvement of the quality of the environment, protection of biodiversity, pollution as well as climate change mitigation and adaptation; (iii a) safety and resilience of transport systems, critical infrastructure and networks.
Public interest Article 54 para 1(a)	the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security, under the control and responsibility of the competent authorities. The processing shall be based on Member State or Union law;	*deleted*
Public interest Article 54 para 1(c)	there are effective monitoring mechanisms to identify if any high risks to the fundamental rights of the data subjects may arise during the sandbox experimentation as well as response mechanism to promptly mitigate those risks and, where necessary, stop the processing;	there are effective monitoring mechanisms to identify if any high risks to the rights and freedoms of the data subjects, as referred to in Article 35 of Regulation (EU) 2016/679 and in Article 35 of Regulation (EU) 2018/1725 may arise during the sandbox experimentation as well as response mechanism to promptly mitigate those risks and, where necessary, stop the processing;

Public interest Article 54 para 1(d)	any personal data to be processed in the context of the sandbox are in a functionally separate, isolated and protected data processing environment under the control of the participants and only authorised persons have access to that data;	any personal data to be processed in the context of the sandbox are in a functionally separate, isolated and protected data processing environment under the control of the prospective provider and only authorised persons have access to that those data;
Public interest Article 54 para 1(f)	any processing of personal data in the context of the sandbox do not lead to measures or decisions affecting the data subjects;	any processing of personal data in the context of the sandbox do not lead to measures or decisions affecting the data subjects nor affect the application of their rights laid down in Union law on the protection of personal data;
Public interest Article 54 para 1(g)	any personal data processed in the context of the sandbox are deleted once the participation in the sandbox has terminated or the personal data has reached the end of its retention period;	any personal data processed in the context of the sandbox are protected by means of appropriate technical and organisational measures and deleted once the participation in the sandbox has terminated or the personal data has reached the end of its retention period;
Public interest Article 54 para 1(h)	the logs of the processing of personal data in the context of the sandbox are kept for the duration of the participation in the sandbox and 1 year after its termination, solely for the purpose of and only as long as necessary for fulfilling accountability and documentation obligations under this Article or other application Union or Member States legislation;	the logs of the processing of personal data in the context of the sandbox are kept for the duration of the participation in the sandbox;
Public interest Article 54 para 1(j)	a short summary of the AI project developed in the sandbox, its objectives and expected results published on the website of the competent authorities.	a short summary of the AI system developed in the sandbox, its objectives, hypotheses, and expected results, published on the website of the competent authorities;

Entry criteria Article 54a (new)(1)	Not covered	Member States shall promote research and development of AI solutions which support socially and environmentally beneficial outcomes, including but not limited to development of AI-based solutions to increase accessibility for persons with disabilities, tackle socio-economic inequalities, and meet sustainability and environmental targets, by: (a) providing relevant projects with priority access to the AI regulatory sandboxes to the extent that they fulfil the eligibility conditions; (b) earmarking public funding, including from relevant EU funds, for AI research and development in support of socially and environmentally beneficial outcomes; (c) organising specific awareness raising activities about the application of this Regulation, the availability of and application procedures for dedicated funding, tailored to the needs of those projects; (d) where appropriate, establishing accessible dedicated channels, including within the sandboxes, for communication with projects to provide guidance and respond to queries about the implementation of this Regulation. Member States shall support civil society and social stakeholders to lead or participate in such projects;

Entry criteria Article 55 para 1(a)	provide small-scale providers and start-ups with priority access to the AI regulatory sandboxes to the extent that they fulfil the eligibility conditions;	provide SMEs and start-ups, established in the Union, with priority access to the AI regulatory sandboxes, to the extent that they fulfil the eligibility conditions;
Entry criteria Article 55 para 1(b)	organise specific awareness raising activities about the application of this Regulation tailored to the needs of the small-scale providers and users;	organise specific awareness raising and enhanced digital skills development activities on the application of this Regulation tailored to the needs of SMEs, start-ups and users;
Innovation hubs Article 55(c)	where appropriate, establish a dedicated channel for communication with small-scale providers and user and other innovators to provide guidance and respond to queries about the implementation of this Regulation.	utilise existing dedicated channels and where appropriate, establish new dedicated channels for communication with SMEs, start-ups, users and other innovators to provide guidance and respond to queries about the implementation of this Regulation;
Jurisdiction Article 56b (new)	Not covered	The AI Office shall carry out the following tasks: .... j) assist authorities in the establishment and development of regulatory sandboxes and to facilitate cooperation among regulatory sandboxes;

9 ABOUT FORHUMANITY

ForHumanity is a 501(c)(3) non profit organisation and ForHumanity Europe is a French 1901 Association, dedicated to addressing risk associated with Ethics, Bias, Privacy, Trust, and Cybersecurity in artificial intelligence and autonomous systems.

ForHumanity uses an open and transparent process that draws from a pool of over 1600+ international contributors from 89 countries to construct audit criteria, certification schemes, and educational programs for legal and compliance professionals, educators, auditors, developers, and legislators to mitigate bias, enhance ethics, protect privacy, build trust, improve cybersecurity, and drive accountability and transparency in AI, algorithmic and autonomous (AAA) systems. ForHumanity works to make AAA Systems safe for all people and makes itself available to support government agencies and instrumentalities to manage risk associated with AI and autonomous systems.

Our mission is to *examine and analyse downside risk associated with the ubiquitous advance of AI, algorithmic and autonomous systems and where possible to engage in risk mitigation to maximise the benefits of these systems… ForHumanity*.

ForHumanity Europe was supported by Huawei UK in the production of this report.

References

[1] Ryan Hagemann and Jennifer Huddleston Skees. "SOFT LAW FOR HARD PROBLEMS: THE GOVERNANCE OF EMERGING TECHNOLOGIES IN AN UNCERTAIN FUTURE". en. In: *Colorado Technology Law Journal* (Feb. 2018).

[2] *Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS*. en. 2021. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206 (visited on 08/28/2023).

[3] Katerina Yordanova. *The Shifting Sands of Regulatory Sandboxes for AI*. July 2019. URL: https://www.law.kuleuven.be/citip/blog/the-shifting-sands-of-regulatory-sandboxes-for-ai/ (visited on 08/07/2023).

[4] *Texts adopted - Artificial Intelligence Act - Wednesday, 14 June 2023*. en. URL: https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.html (visited on 07/09/2023).

[5] Gestel„ and Dijck. "Better Regulation through Experimental Legislation." In: *European Public Law 17.3 (2011): 539-553*. (2011).

[6] Financial Conduct Authority. *Digital Sandbox*. en. May 2020. URL: https://www.fca.org.uk/firms/innovation/digital-sandbox (visited on 08/04/2023).

[7] *MAS enhances FinTech Regulatory Sandbox and announces technology and data sharing platform*. en. URL: https://www.cliffordchance.com/content/cliffordchance/insights/resources/blogs/talking-tech/en/articles/2021/11/mas-enhances-fintech-regulatory-sandbox-and-announces-technology.html (visited on 07/08/2023).

[8] *MOH | Licensing Experimentation and Adaptation Programme (LEAP) - A MOH Regulatory Sandbox*. URL: https://www.moh.gov.sg/home/our-healthcare-system/licensing-experimentation-and-adaptation-programme-(leap)---a-moh-regulatory-sandbox (visited on 07/08/2023).

[9] Fresh Media. *Major Sandbox for Energy Infrastructure and Data Launched in Lithuania*. en. URL: https://ignitisgrupe.lt/en/news/major-sandbox-energy-infrastructure-and-data-launched-lithuania (visited on 07/08/2023).

[10] World Bank. *How to Build a Regulatory Sandbox - A Practical Guide for Policy Makers*. Sept. 2020. URL: https://documents1.worldbank.org/curated/en/126281625136122935/pdf/How-to-Build-a-Regulatory-Sandbox-A-Practical-Guide-for-Policy-Makers.pdf (visited on 07/08/2023).

[11] European Council. *Council Conclusions on Regulatory sandboxes and experimentation clauses as tools for an innovation-friendly, future-proof and resilient regulatory framework that masters disruptive challenges in the digital age*. Nov. 2020. URL: https://data.consilium.europa.eu/doc/document/ST-13026-2020-INIT/en/pdf (visited on 07/08/2023).

[12] Sofia Ranchordas. *Experimental Regulations for AI: Sandboxes for Morals and Mores*. May 2021. URL: https://ssrn.com/abstract=3839744.

[13] Financial Conduct Authority. *Regulatory sandbox*. en. Tech. rep. Nov. 2015.

[14] Hilary Allen. "Regulatory Sandboxes". en. In: (2019).

[15] Ahmad Alaassar, Anne-Laure Mention, and Tor Helge Aas. "Exploring how social interactions influence regulators and innovators: The case of regulatory sandboxes". en. In: *Technological Forecasting and Social Change* 160 (Nov. 2020), p. 120257. ISSN: 00401625. DOI: 10.1016/j.techfore.2020.120257. URL: https://linkinghub.elsevier.com/retrieve/pii/S0040162520310830 (visited on 07/08/2023).

[16] Dirk A. Zetzsche et al. "Regulating a Revolution: From Regulatory Sandboxes to Smart Regulation". en. In: *SSRN Electronic Journal* (2017). ISSN: 1556-5068. DOI: 10.2139/ssrn.3018534. URL: https://www.ssrn.com/abstract=3018534 (visited on 07/08/2023).

[17] *Sandboxes for Responsible Artificial Intelligence*. en-US. URL: https://www.eipa.eu/publications/briefing/sandboxes-for-responsible-artificial-intelligence/ (visited on 07/08/2023).

[18] *The role of sandboxes in promoting flexibility and innovation in the digital age*. en. Going Digital Toolkit Notes 2. June 2020. DOI: 10.1787/cdf5ed45-en. URL: https://www.oecd-ilibrary.org/science-and-technology/the-role-of-sandboxes-in-promoting-flexibility-and-innovation-in-the-digital-age_cdf5ed45-en (visited on 08/07/2023).

[19] Sofia Ranchordás. "Experimental Regulations and Regulatory Sandboxes – Law Without Order?: Special Issue Experimental Legislation in Times of Crisis, Sofia Ranchordás & Bart van Klink (eds.)" en. In: *Law and Method* (Nov. 2021). ISSN: 2352-7927. DOI: 10.5553/REM/.000064. URL: https://www.bjutijdschriften.nl/doi/10.5553/REM/.000064 (visited on 08/07/2023).

[20] Global Financial Innovation Network. *GFIN Cross Border Testing Initiative Cohort 1 Report*. May 2022. URL: https://static1.squarespace.com/static/5db7cdf53d173c0e010e8f68/t/62baeaac3ec4851f313afe78/1656416941725/GFIN+Cross-Border+Testing+Initiative+Cohort+1_0+external+2_FINALFINAL.pdf (visited on 08/08/2023).

[21] *Sandbox Project - EBSI -*. URL: https://ec.europa.eu/digital-building-blocks/wikis/display/EBSI/Sandbox+Project (visited on 08/08/2023).

[22] Dirk A. Zetzsche et al. "Regulating a Revolution: From Regulatory Sandboxes to Smart Regulation". en. In: *SSRN Electronic Journal* (2017). ISSN: 1556-5068. DOI: 10.2139/ssrn.3018534. URL: https://www.ssrn.com/abstract=3018534 (visited on 07/09/2023).

[23] *About the FCA | FCA*. URL: https://www.fca.org.uk/about/what-we-do/the-fca (visited on 07/09/2023).

[24] *Regulatory Sandbox accepted firms*. en. Mar. 2022. URL: https://www.fca.org.uk/firms/innovation/regulatory-sandbox/accepted-firms (visited on 08/07/2023).

[25] Saule T Omarova. "Technology v Technocracy: Fintech as a Regulatory Challenge". en. In: *Journal of Financial Regulation* 6.1 (June 2020), pp. 75–124. ISSN: 2053-4841. DOI: 10.1093/jfr/fjaa004. URL: https://academic.oup.com/jfr/article/6/1/75/5858319 (visited on 07/09/2023).

[26] Alessio Tartaro, Adam Leon Smith, and Patricia Shaw. *Assessing the impact of regulations and standards on innovation in the field of AI*. Feb. 2023. URL: http://arxiv.org/abs/2302.04110 (visited on 08/07/2023).

[27] *Coordinated Plan on Artificial Intelligence | Shaping Europe's digital future*. en. June 2023. URL: https://digital-strategy.ec.europa.eu/en/policies/plan-ai (visited on 08/13/2023).

[28] *WTO Trade Concerns*. URL: https://tradeconcerns.wto.org/en/stcs/details?imsId=736&domainId=TBT&searchTerm=AI (visited on 08/08/2023).

[29] *How to Build a Good Regulatory Sandbox | Mercatus Center*. en. Apr. 2019. URL: https://www.mercatus.org/economic-insights/expert-commentary/how-build-good-regulatory-sandbox (visited on 08/07/2023).

[30] Thomas Buocz, Sebastian Pfotenhauer, and Iris Eisenberger. "Regulatory sandboxes in the AI Act: reconciling innovation and safety?" en. In: *Law, Innovation and Technology* (Aug. 2023), pp. 1–33. ISSN: 1757-9961, 1757-997X. DOI: 10.1080/17579961.2023.2245678. URL: https://www.tandfonline.com/doi/full/10.1080/17579961.2023.2245678 (visited on 08/21/2023).

[31] OPINION OF ADVOCATE GENERAL and POIARES MADURO. *Case C-127/07 Société Arcelor Atlantique*. 2008. URL: https://curia.europa.eu/juris/document/document.jsf;jsessionid=F0540C161AF2597E8EAEACC34F33426F?text=&docid=67733&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=1958679 (visited on 08/21/2023).

[32] *Regulatory Sandbox*. en. July 2023. URL: https://ico.org.uk/for-organisations/advice-and-services/regulatory-sandbox/ (visited on 08/15/2023).

[33] *"Sandbox": CNIL launches call for projects on artificial intelligence in public services*. en. URL: https://www.cnil.fr/en/sandbox-cnil-launches-call-projects-artificial-intelligence-public-services (visited on 08/14/2023).

[34] IMY. *Swedish Authority for Privacy Protection, IMY, finishes first Sandbox Pilot*. Tech. rep. Mar. 2023. URL: https://www.imy.se/globalassets/dokument/ovrigt/first-regulatory-sandbox-pilot---english-summary.pdf (visited on 08/14/2023).

[35] Federal Ministry for Economic Affairs and Energy (BMWi), Germany. *Making space for innovation - The handbook for regulatory sandboxes*. en. Tech. rep. 2019.

[36] Rishi Bommasani et al. *On the Opportunities and Risks of Foundation Models*. July 2022. URL: http://arxiv.org/abs/2108.07258 (visited on 08/07/2023).

[37] VentureBeat. *Foundation models: 2022's AI paradigm shift*. en-US. Sept. 2022. URL: https://venturebeat.com/ai/foundation-models-2022s-ai-paradigm-shift/ (visited on 08/07/2023).

[38] ISO/IEC. *ISO/IEC 22989:2022, Information technology — Artificial intelligence — Artificial intelligence concepts and terminology*. en. 2022. URL: https://www.iso.org/standard/74296.html (visited on 08/08/2023).

[39] ISO/IEC. *ISO/IEC FDIS 42001:2023, Information technology — Artificial intelligence — Management system*. en. URL: https://www.iso.org/standard/81230.html (visited on 08/08/2023).

[40] Philipp Tschandl et al. "Human–computer collaboration for skin cancer recognition". en. In: *Nature Medicine* 26.8 (Aug. 2020), pp. 1229–1234. ISSN: 1078-8956, 1546-170X. DOI: 10.1038/s41591-020-0942-0. URL: https://www.nature.com/articles/s41591-020-0942-0 (visited on 08/21/2023).

[41] Ian European Digital SME Alliance. *DIGITAL SME Position Paper on the EU AI Act*. en. Tech. rep. Sept. 2023. URL: https://www.digitalsme.eu/digital/uploads/DIGITAL-SME-Position-Paper-AI-Act-FINAL-DRAFT-1.pdf (visited on 08/07/2023).

[42] *Commission's proposal for a regulation on Artificial Intelligence fails to address the workplace dimension*. en. URL: https://www.etuc.org/en/document/commissions-proposal-regulation-artificial-intelligence-fails-address-workplace-dimension (visited on 08/07/2023).

[43] BCS, The Chartered Institute for IT. *Helping AI grow up without pressing pause*. May 2023. URL: https://www.bcs.org/media/10542/helping-ai-grohttps://www.bcs.org/media/10542/helping-ai-grow-up.pdfw-up.pdf.

[44] *Launch event for the Spanish Regulatory Sandbox on Artificial Intelligence | Shaping Europe's digital future*. en. June 2022. URL: https://digital-strategy.ec.europa.eu/en/events/launch-event-spanish-regulatory-sandbox-artificial-intelligence (visited on 08/02/2023).

[45] ISO/IEC. *ISO/IEC DTS 12791, Information technology — Artificial intelligence — Treatment of unwanted bias in classification and regression machine learning tasks*. en. 2023. URL: https://www.iso.org/standard/84110.html (visited on 08/08/2023).

[46] ISO/IEC. *ISO/IEC DIS 5259-2, Artificial intelligence — Data quality for analytics and machine learning (ML) — Part 2: Data quality measures*. en. 2023. URL: https://www.iso.org/standard/81860.html (visited on 08/07/2023).

[47] ISO/IEC. *ISO/IEC TR 24027:2021*. en. Nov. 2021. URL: https://www.iso.org/standard/77607.html (visited on 08/08/2023).

[48] *Russia introduces regulatory sandboxes for digital innovation*. en. URL: https://cms-lawnow.com/en/ealerts/2020/10/russia-introduces-regulatory-sandboxes-for-digital-innovation (visited on 08/14/2023).

[49] CMS. *Experimental AI regime to be introduced in Moscow*. en. URL: https://cms-lawnow.com/en/ealerts/2020/06/experimental-ai-regime-to-be-introduced-in-moscow (visited on 08/14/2023).

[50] Datailsynet. *Framework for the Regulatory Sandbox*. en. URL: https://www.datatilsynet.no/en/regulations-and-tools/sandbox-for-artificial-intelligence/framework-for-the-regulatory-sandbox/ (visited on 08/14/2023).

[51] The Danish Government. *National Strategy for Artificial Intelligence*. en. Tech. rep. Mar. 2019.

[52] EU High-Level Experts Group. *Ethics guidelines for trustworthy AI | Shaping Europe's digital future*. en. Apr. 2019. URL: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (visited on 08/14/2023).

[53] ICO. *Regulatory Sandbox Final Report: Onfido*. Tech. rep. 2020. URL: https://ico.org.uk/media/for-organisations/documents/2618551/onfido-sandbox-report.pdf (visited on 08/14/2023).

[54] *Innovation Sandbox for AI*. en. URL: https://www.zh.ch/en/wirtschaft-arbeit/wirtschaftsstandort/innovation-sandbox.html (visited on 08/14/2023).

[55] European Commission. *Summary of references of harmonised standards published in the Official Journal – Directive 2006/42/EC 1 of the European Parliament and of the Council of 17 May 2006 on machinery, and amending Directive 95/16/EC*. URL: https://ec.europa.eu/docsroom/documents/55576 (visited on 08/28/2023).

[56] Rickard Brännvall, Helena Linge, and Johan Östman. "Can the use of privacy enhancing technologies enable federated learning for health data applications in a Swedish regulatory context?" en. In: June 2023, pp. 58–67. DOI: 10.3384/ecp199006. URL: https://ecp.ep.liu.se/index.php/sais/article/view/718 (visited on 08/30/2023).

[57] European Commission and laying down harmonised rules on artificial intelligence (Artificial Intelligence. *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - draft for coreper*. Nov. 2022. URL: https://artificialintelligence eu/wp-content/uploads/2022/11/AIA-CZ-Draft-for-Coreper-3-Nov-22.pdf (visited on 07/09/2023).

[58] Ada Lovelace Institute. *Expert explainer: Allocating accountability in AI supply chains*. en-GB. URL: https://www.adalovelaceinstitute.org/resource/ai-supply-chains/ (visited on 07/31/2023).

[59] ISO/IEC. *ISO/IEC DTR 5469, Artificial intelligence — Functional safety and AI systems*.

[60] Margaret Mitchell et al. "Model Cards for Model Reporting". en. In: *Proceedings of the Conference on Fairness, Accountability, and Transparency - FAT* '19* (2019), pp. 220–229. DOI: 10.1145/3287560.3287596. URL: http://arxiv.org/abs/1810.03993 (visited on 04/26/2020).

[61] *Algorithmic Transparency Recording Standard Hub*. en. URL: https://www.gov.uk/government/collections/algorithmic-transparency-recording-standard-hub (visited on 08/02/2023).

[62] *Model Card Guidebook*. URL: https://huggingface.co/docs/hub/model-card-guidebook (visited on 08/21/2023).

[63] Inês Fernandes Godinho, Cláudio R. Flores, and Nuno Castro Marques. "CONSULTATION ON THE WHITE PAPER ON ARTIFICIAL INTELLIGENCE - A EUROPEAN APPROACH". en. In: *ULP Law Review* 14.1 (Feb. 2021), pp. 157–167. ISSN: 21846219. DOI: 10.46294/ulplr-rdulp.v14i1.7475. URL: https://revistas.ulusofona.pt/index.php/rfdulp/article/view/7475 (visited on 07/09/2023).

[64] Global Digital Foundation. *Trust and assurance in the AI Assurance Eco-system A Multi Actor Governance Framework (MAGF)*.

[65] Nitin Sawhney and Ana Paula Gonzalez Torres. "Devising Regulatory Sandboxes and Responsible Practices for Designing AI-based Services in the Finnish Public Sector". en. In: *International Workshop on AI Compliance Mechanism (WAICOM 2022), 35th International Conference on Legal Knowledge and Information Systems (JURIX 2022)* ().

[66] Emma Bluemke et al. *Exploring the Relevance of Data Privacy-Enhancing Technologies for AI Governance Use Cases*. Mar. 2023. URL: http://arxiv.org/abs/2303.08956 (visited on 09/17/2023).

[67] European Commission. *DocsRoom - European Commission*. URL: https://ec.europa.eu/docsroom/documents/52376 (visited on 08/02/2023).

[68] Natalia Díaz-Rodríguez et al. "Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation". In: *Information Fusion* 99 (2023), p. 101896. ISSN: 1566-2535. DOI: https://doi.org/10.1016/j.inffus.2023.101896. URL: https://www.sciencedirect.com/science/article/pii/S1566253523002129.

[69] *FDA Predetermined Change Control Plan: AI/ML-Enabled Device Software Functions*. URL: https://www.mwe.com/insights/fda-issues-draft-predetermined-change-control-plan-for-machine-learning-enabled-device-software-functions/ (visited on 08/21/2023).

[70] Letizia Tomada. "Start-ups and the proposed EU AI Act: Bridges or Barriers in the path from Invention to Innovation?" In: *JIPITEC* 13.1 (Apr. 2022). ISSN: 2190-3387. URL: https://www.jipitec.eu/issues/jipitec-13-1-2022/5511.

[71] Harry Armstrong and Jen Rae. "A working model for anticipatory regulation". en. In: (Nov. 2017).

[72] Richard K. Lomotey, Sandra Kumi, and Ralph Deters. "Data Trusts as a Service: Providing a platform for multi□party data sharing". en. In: *International Journal of Information Management Data Insights* 2.1 (Apr. 2022), p. 100075. ISSN: 26670968. DOI: 10.1016/j.jjimei.2022.100075. URL: https://linkinghub.elsevier.com/retrieve/pii/S2667096822000180 (visited on 09/17/2023).

[73] The Regulatory Review. *The Value of Public Participation in Rulemaking | The Regulatory Review*. en-US. Sept. 2017. URL: https://www.theregreview.org/2017/09/25/scalia-public-participation-rulemaking/ (visited on 08/04/2023).

[74] Kristin Undheim, Truls Erikson, and Bram Timmermans. "True uncertainty and ethical AI: regulatory sandboxes as a policy tool for moral imagination". en. In: *AI and Ethics* (Nov. 2022). ISSN: 2730-5953, 2730-5961. DOI: 10.1007/s43681-022-00240-x. URL: https://link.springer.com/10.1007/s43681-022-00240-x (visited on 07/08/2023).

[75] Aleks Berditchevskaia. "Participatory AI for humanitarian innovation". en. In: ().

[76] Lara Groves et al. *Going public: the role of public participation approaches in commercial AI labs*. June 2023. URL: http://arxiv.org/abs/2306.09871 (visited on 07/31/2023).

[77] Care Quality Commission. *Evaluation of CQC's regulatory sandboxing pilot*. URL: https://www.cqc.org.uk/what-we-do/how-we-work-people/evaluation-cqcs-regulatory-sandboxing-pilot (visited on 08/14/2023).

[78] Ross P. Buckley et al. "Building FinTech Ecosystems: Regulatory Sandboxes, Innovation Hubs and Beyond". en. In: *SSRN Electronic Journal* (2019). ISSN: 1556-5068. DOI: 10.2139/ssrn.3455872. URL: https://www.ssrn.com/abstract=3455872 (visited on 08/07/2023).

[79] *Sectorial AI Testing and Experimentation Facilities under the Digital Europe Programme | Shaping Europe's digital future*. en. May 2023. URL: https://digital-strategy.ec.europa.eu/en/activities/testing-and-experimentation-facilities (visited on 08/14/2023).

[80] Natasha Lepore et al., eds. *Processing and Analysis of Biomedical Information: First International SIPAIM Workshop, SaMBa 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Revised Selected Papers*. en. Vol. 11379. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2019. ISBN: 978-3-030-13834-9 978-3-030-13835-6. DOI: 10.1007/978-3-030-13835-6. URL: http://link.springer.com/10.1007/978-3-030-13835-6 (visited on 08/28/2023).

[81] European Commission. *Impact Assessment Accompanying the Proposal for a Regulation of the European Parliament and of the Council LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS*. en. European Commission, 2021. URL: https://digital-strategy.ec.europa.eu/en/library/impact-assessment-regulation-artificial-intelligence (visited on 08/07/2023).

[82] *Machine learning in UK financial services*. en. Aug. 2023. URL: https://www.bankofengland.co.uk/report/2022/machine-learning-in-uk-financial-services (visited on 08/15/2023).

[83] *The EU's AI Act Is Barreling Toward AI Standards That Do Not Exist*. en. URL: https://www.lawfaremedia.org/article/eus-ai-act-barreling-toward-ai-standards-do-not-exist (visited on 08/08/2023).

[84] *Key enforcement issues of the AI Act should lead EU trilogue debate*. en-US. URL: https://www.brookings.edu/articles/key-enforcement-issues-of-the-ai-act-should-lead-eu-trilogue-debate/ (visited on 07/10/2023).