Getting Regulatory Sandboxes Right: Design and Governance Under the AI Act

Claudio Novelli[1], Philipp Hacker[2], Simon McDougall[3], Jessica Morley[1], Antonino Rotolo[4][1], Luciano Floridi[1,4]

[1] *Digital Ethics Center, Yale University, 85 Trumbull Street, New Haven, United States*
[2] *European New School of Digital studies, European University Viadrina, Gr. Scharrnstr. 59, Frankfurt, (Oder), Germany*
[3] *Digital Ethics Center, Yale University, 85 Trumbull Street, New Haven, United States*
[4] *University of Bologna, Alma AI and Department of Legal Studies, Via Zamboni 27/29, Bologna, Italy*

Abstract. Regulating emerging technologies involves balancing the mitigation of risks with the promotion of innovation; a balance frequently seen as a zero-sum "dilemma of control". Regulatory sandboxes offer a practical way to address this dilemma by enabling controlled, evidence-based testing of new technologies. In this article, we examine the regulatory sandbox framework introduced by the EU Artificial Intelligence Act (AIA). We argue that the AIA's multi-level governance structure represents a shift from traditional sandbox models by prioritizing regulatory learning over technological disruption and expanding public interest considerations to include strategically aligned commercial innovations. Afterwards, we identify governance challenges across three sandbox phases—pre-testing, testing, and post-testing—and propose structured solutions. Our analysis suggests that effective sandbox governance requires specific mechanisms: tailored entry criteria, precise pipeline placement guidance, and multi-agency coordination in pre-testing; experimental realism and continuous risk classification updates during testing; and clear graduation criteria with robust transition support in post-testing.

[1] Supported by *EU Regulatory Sandboxes for AI* (EUSAiR) - DIGITAL-2024-AI-ACT-06-SANDBOX (101195535).

1. Introduction

Regulating emerging technologies frequently involves striking a delicate balance between imposing necessary obligations and maintaining innovation momentum. Popularised by (Collingridge 1980), the "dilemma of control" clearly describes this inherent challenge: regulating too early risks stifling innovation, while regulating too late allows potential harms to proliferate unchecked. This dilemma is commonly framed either as a negative-sum game—where neither regulation nor innovation wins—or as a zero-sum game, where one benefits only at the expense of the other. However, this representation oversimplifies reality, as other factors beyond regulation often play more significant roles in hindering innovation (Bradford 2024), including limited access to funding, inadequate infrastructure, skill shortages, market monopolies, and unfavorable economic conditions. Moreover, it is ethically problematic because it implies that societies must accept either the risks of under-regulated, potentially harmful technologies or innovation stagnation with significant opportunity costs.

Traditional regulatory approaches struggle with this dilemma due to their static, inflexible frameworks, unable to quickly adapt to technological uncertainties. Fortunately, practical ways exist to overcome this dilemma. Regulatory sandboxes—controlled testing environments providing dedicated regulatory support and often temporary relaxation of specific legal requirements—represent one practical mechanism for achieving proportionate regulation of emerging technologies, including AI. Acting as policymakers' laboratories, sandboxes enable iterative refinements and evidence-based adjustments to regulatory frameworks, offering innovators and regulators insights into how emerging technologies function under realistic but controlled conditions. However, sandboxes also present challenges, particularly regarding their integration into existing complex regulatory systems. Currently, a lack of standardized guidelines for sandbox design and operation can sometimes create, rather than resolve, governance risks.

To address this gap, we comprehensively analyze the regulatory sandbox framework emerging from the European Union's Artificial Intelligence Act (AIA), making a twofold contribution. First, we compare the AIA's sandbox model with previous sandbox experiences to clearly map its distinctive multi-level governance structure (Sections 3 and 4). We argue that the AIA departs from traditional models by adopting a broader understanding of innovation—prioritizing regulatory learning over technological disruption, aligning effectively with AI's incremental development patterns. Additionally, the AIA implicitly expands the notion of public interest beyond traditional social-good considerations to include commercially driven innovations that align with the EU's strategic objectives, such as developing efficient alternative LLMs or reducing reliance on external cloud services.

Second, we identify critical governance challenges associated with sandboxes across three essential phases—pre-testing, testing, and post-testing—and offer structured, actionable strategies to address them (Section 5). Our analysis underscores the importance of specific governance levers at each

stage: during pre-testing, tailored entry criteria, precise pipeline placement guidance, structured multi-agency coordination, transparency measures, and sunset clauses; during testing, maintaining experimental realism, continuously updating risk classifications, and realistic oversight; and during post-testing, clear graduation criteria, algorithmic audits, and robust support mechanisms to facilitate smooth market transitions. The paper concludes by summarizing core insights and highlighting further considerations for policymakers, regulators, and AI providers (Section 6).

2. Regulatory Sandboxes under the AIA: Preconditions for Effective Functioning

Regulatory sandboxes are defined in various ways, often differing significantly depending on whether the definition is offered by a (legal) academic or not (Makarov and Davydova 2021). Trying to maintain a legal stance on this concept, we can define regulatory sandboxes as provisional legal regimes that enable structured and tightly scoped experimental governance: regulators provide support to participants, but also often temporarily relax, adapt or choose not to enforce rules to facilitate controlled testing of innovative technologies or business models – that are not yet available on the market – while maintaining oversight to assess risks, evaluate regulatory gaps, and iteratively refine rules in response to empirical evidence. The partial suspension of sanctions within the sandbox serves as a balance mechanism between innovation and risk: it benefits innovators by reducing the likelihood and cost of non-compliance, and it assists regulators by enabling the early implementation of safeguards and proactive risk mitigation measures (Allen 2019a; 2019b; Zetzsche et al. 2017).

Accordingly, regulatory sandboxes should not be conflated with broader regulatory engagement mechanisms, which typically involve engaging and advising innovators within the bounds of existing legal frameworks and do not entail any derogation from current regulations.

The European AI Act (AIA) puts regulatory sandboxes front and centre to help drive AI innovation. EU countries need to set up at least one national sandbox–or join forces with others–to make sure there's solid coverage across the board. These sandboxes can be physical, digital, or a mix of both, but they need the right resources to actually work. The idea is simple: give AI developers a safe space to test their tech, stay on the right side of the law, and get clearer guidance from regulators. It's also about making it easier for startups and small businesses to get to market, while encouraging collaboration and sharing know-how.

A core aspect of the AIA's conceptualization of regulatory sandboxes – balanced alongside the promotion of innovation – is the resolution or mitigation of legal uncertainty. Sandboxes are, in fact, designed to help both (prospective) providers and regulators. For regulators in particular, they serve as a means of evidence-based regulatory learning (Recital 139, AIA). As we argue in Section 3.1, this uncertainty-reducing potential may, in fact, constitute the primary criterion for determining the eligibility of projects and participants for admission to a sandbox.

The AIA sets out rules on regulatory sandboxes in Chapter VI (Articles 57–63), which centers on measures supporting innovation. Accordingly, regulatory sandboxes are intended to balance the mitigation of risks – particularly regarding fundamental rights, health, and safety – with the need to provide a supportive context for developers, researchers, and prospective providers of AI systems (as indicated in Recital 138).

Some key needs to bear in mind when setting up a regulatory sandbox include:

(1) Pinpointing which non-essential legal rules can be safely relaxed–rules big enough to let firms experiment, small enough to keep core protections (e.g., civil liability) intact, and always ring-fenced by strict oversight. Crucially, not all regulations are equally flexible–some directly safeguard consumers and markets, while others are more administrative/procedural. For instance, waiving rules that mitigate fundamental risks, such as transparency, product safety, or redress mechanisms, demands compensating safeguards. These might include requiring explicit customer consent, limiting the scope or duration of tests, or imposing enhanced monitoring. By contrast, administrative hurdles–such as lengthy licensing processes or rigid capital requirements in financial services–may be more readily eased within a sandbox, as they do not inherently weaken consumer protections.

(2) Plugging the sandbox into existing European AI resources–digital-innovation hubs, data spaces, test beds, open-source tools–so teams can reuse infrastructure instead of starting from scratch.

(3) Securing a sustainable setup: pooled or multi-year funding, staff that scale with demand, and continuous training for supervisors (EUSAiR 2025 Roadmap, 5-8). We shall deepen these aspects in Section 6.

3. Inside the AIA Regulatory Sandbox: An Applicant Perspective

Any effort to understand and describe the functioning of a regulatory sandbox must begin by recognising that there is no single, standardised model. Of course, each model is inherently shaped by the legal framework within which it operates. However, for the early stage, broad and cross-cutting legislation like the AIA, marked by its overlap with numerous other regulatory frameworks, the overall legal landscape remains complex and, in some areas, unresolved. In these situations, we can still attempt to address uncertainties by examining how regulatory sandboxes are currently being implemented and operated.

To achieve this, we adopt a comparative approach which draws on existing literature, such as (Seferi 2025). Another useful source are the studies commissioned by the EU Parliament (Parenti 2020) or by EU competent authorities (ESMA, EBA, EIOPA 2018).[2] These studies identify fundamental patterns and best practices based on 87 sandbox use cases across financial services, digital technologies (including blockchain), healthcare,

[2] More specifically, the Policy Department for Economic, Scientific and Quality of Life Policies Directorate-General for Internal Policies.

telecommunications, energy, and other industries, covering a range of legal systems.

By building on how regulatory sandboxes operate in other fields and countries, we can develop more informed expectations about how they will function under the AIA.

3.1. Eligibility criteria

Chapter VI of the AIA outlines the operational rules for regulatory sandboxes but does not explicitly define eligibility criteria, that is, the parameters used to assess whether applicants qualify for sandbox entry. Article 58(1)(a) delegates this task to the EU Commission, which is expected to specify the criteria through implementing acts. However, independently of the implementing act(s), it is still possible to extract the principles and constraints that will guide the development of the eligibility criteria from the AIA itself.

Generally, the AIA intends for regulatory sandboxes to be inclusive rather than restrictive. Rather than limiting entry to a niche group of "cutting-edge" companies, the AIA opens sandbox participation broadly to any provider or prospective provider of AI systems. This is explicitly referenced in Recitals 139 and 141, and Article 3. In keeping with this aim, the AIA intends for eligibility to depend on objective criteria set by the Commission's implementing act and potentially further specifications by the relevant national sandbox authority. Of course, these criteria may differ to some extent between jurisdictions or industries (within the limits of the Commission's implementing act), but all must be transparent and publicly available to ensure that applicants can assess their likelihood of qualifying before applying.

That the AIA intends for AI regulatory sandboxes to be inclusive in design does not, however, mean that there will be no prioritization of sandbox applicants. Instead, the AIA stresses the importance of supporting Small and Medium-sized Enterprises (SMEs) (Recital 143 and Article 62(1)). This priority is operationalized in two key ways: (a) if two applicants meet the baseline eligibility criteria and score similarly in other respects, SMEs may be given precedence; (b) Article 58(2)(d) stipulates that participation in the sandbox should be free of charge for SMEs, except in exceptional circumstances where authorities may recover costs.

To anticipate how the eligibility criteria may look under the AIA, we can gain lessons from what happens in other sandboxes. In particular, Seferi highlights a set of recurring eligibility criteria (Seferi 2025), with three primary criteria emerging as the most significant:

1) The *degree of innovativeness* is the most frequently used eligibility criterion (with 74 occurrences) and assesses how novel and groundbreaking a proposed project or solution is. Specifically, it evaluates whether a project addresses clear market gaps or introduces entirely new functionalities, often through emerging technologies or novel business models. Both the Italian Financial Services Regulatory Sandbox and the UK Financial Conduct Authority's (FCA) Regulatory Sandbox use this standard to assess applicants.

The FCA's Regulatory Sandbox specifically requires "genuine innovation" (Financial Conduct Authority 2015, 7) as an eligibility criterion. This explicitly requires cutting-edge financial innovations rather than incremental product enhancements. A genuinely innovative solution under these criteria might be a novel approach to delivering financial services to underserved groups, whereas minor improvements to existing services would not qualify.[3]

This binary distinction poses challenges when applied to AI, as AI's transformative impact often lies in incremental optimizations rather than radical reinvention. Many AI applications, especially those based on machine learning, significantly refine existing processes such as automating customer service, enhancing predictive analytics, or streamlining supply chains. However, these incremental improvements do not necessarily align with traditional definitions of "genuine innovation," which focus on radically disruptive breakthroughs. For example, an AI system improving retail demand forecasting by 15% offers meaningful efficiency gains but may not be viewed as redefining the sector. To qualify as genuinely innovative, an AI-driven project may likely need to go beyond automation and create new market opportunities or solve market gaps in new ways. For example, an AI-powered credit scoring system that leverages alternative data sources to expand financial access for the unbanked would likely meet this threshold–particularly if it provides a significantly better alternative to traditional credit evaluation methods in ways that have not yet entered the market.

The AI Act (AIA) implicitly addresses this challenge by adopting a broader definition of innovativeness. Although placed within Chapter VI ("Measures in Support of Innovation"), suggesting innovativeness will remain essential, the AIA is unlikely to restrict eligibility solely to radical disruption. Two main rationales support this broader interpretation:

The technological rationale, as anticipated, is that AI systems often optimize existing processes rather than create entirely new ones. Machine learning is frequently deployed to enhance efficiency, as, for instance, in supply chain management or customer service. Improvements can be significant, but they typically do not represent a complete departure from existing solutions, either technologically or procedurally.

The hermeneutic rationale is that the AIA does not explicitly require applicants to demonstrate "genuine novelty" or disruptive potential, as is common in other sandbox frameworks (or IP rights). Instead, the Act repeatedly prioritizes legal uncertainty as a key trigger for sandbox participation (Recital 139). This suggests that the primary rationale for inclusion is the need for regulatory clarity or guidance, regardless of whether the system is "disruptively" novel. The AIA's acceptance of non-disruptive innovations finds further justification in the pragmatic realities of non-EU companies assessing market entry strategies. Many commercially operational AI systems – particularly

[3] Link: https://www.fca.org.uk/firms/innovation/regulatory-sandbox/eligibility-criteria.

those developed in jurisdictions with less stringent regulations – may not qualify as innovative by conventional sandbox standards, yet their operators (and EU regulators) face genuine uncertainty about EU compliance pathways and the feasibility of EU market entry.

In other words, innovativeness matters, but primarily to the extent that the project raises unusual or emergent technical or legal questions, or provides an opportunity for providers and regulators to "learn by doing" in areas of uncertain compliance. Consequently, unlike some sandbox regimes (e.g., fintech), the AIA would not explicitly exclude incremental innovations from eligibility. The focus would likely be on whether the project presents non-trivial questions or challenges related to compliance with the Act's obligations.

2) *Public interest or societal benefit* is a less common but still significant eligibility criterion, appearing in 67 cases. It assesses the broader societal impact of a project, including its contributions to public welfare and market benefits. More specifically, this means that a project or solution "advances the public interest or [...] generates positive social impacts, such as improving accessibility, promoting sustainability, or enhancing public and private services" (Seferi 2025, 150).

A positive indicator of eligibility may be, for instance, the project's ability to expand access to essential services for underserved and marginalized communities or to support environmental conservation. For instance, the UK's 'AI Airlock' for AI as a medical device requires that applicants demonstrate that their product or prototype will provide a benefit to patients and public health. Conversely, a negative indicator may be that the project exclusively benefits privileged groups or contributes to environmental strain.

It is, however, important to recognise that criteria related to societal benefit are, by their nature, subjective and often qualitative. The EU's understanding of the value of innovation has evolved subsequent to the passing of the AIA, with the publication of the Draghi Report in September 2024, and the subsequent acceptance of its overarching points by European Commission President Ursula von der Leyen. The Report argues that increased productivity through innovation is essential to preserving the EU values of equity and social inclusion. As such, the scope of societal benefits expands beyond its traditional boundaries to cover innovation that supports the EU's global competitiveness and economic growth. In this context, commercial innovations that support the EU's competitiveness with the US and China, such as developing alternative LLMs or reducing the EU's reliance on US cloud infrastructure, would meet the public interest criterion. In any case, such value-subjectivity may make it more challenging to reach agreement as to whether a particular technology meets the public benefit criterion.

The AIA does not explicitly list public reason or societal benefit as a condition for accessing regulatory sandboxes. This omission may appear unexpected, given that its risk-based categorisation and

obligations for AI deployers are firmly grounded in the EU's cultural and legal values, which prioritise societal welfare. The AIA's prohibition of harmful practices, such as biometric categorisation, and its focus on safety and trustworthiness, reflect a clear principle: AI development should not be driven solely by profit or private interests but should aim to serve the public good.

That said, claiming that public interest plays no role in AIA regulatory sandboxes would be inaccurate. The AIA does address it, notably in Article 59(1), which allows for "further processing of personal data for developing certain AI systems in the public interest". This exception grants data-sharing flexibilities within the sandbox for AI systems addressing substantial public-interest domains, such as public health (e.g., disease detection), environmental protection, and energy sustainability. Projects targeting recognised public interest areas may benefit from enhanced data-use permissions or enhanced engagement with regulators, resulting in expedited regulatory guidance. Conversely, projects without a clear societal benefit are not excluded but may not access the same advantages.

Looking ahead, public interest in the regulatory sandbox mechanism could still be integrated into the AIA's implementing acts more explicitly. To align with the Act's overarching framework, such eligibility criteria would need to reflect its specific conception of public interest, which spans from safeguarding fundamental rights, ensuring safety, promoting transparency, and fostering innovation. Challenges remain in balancing these values, as the AIA lacks detailed guidance on prioritising or reconciling them in practice, nor, one could argue, is it necessarily intended to do so. Such judgments may be more appropriately addressed on a case-by-case basis, for instance, through judicial interpretation or administrative decision-making. In this respect, a degree of legal ambiguity surrounding eligibility may be both inevitable and, to some extent, desirable.

3) *Level of maturity* is a similarly significant criterion, appearing 63 times. It assesses whether a project is sufficiently advanced for meaningful experimentation within a sandbox and considers s several factors, including technical capability, financial sustainability, and scalability(Seferi 2025, 151). However, technological readiness for sandbox candidates should not be equated with that of fully developed and tested technologies; it should be evaluated based on their potential for structured experimentation. The level of maturity is closely tied to timing. Applicants must demonstrate that their innovations are at a stage where they are ready to be tested so that the sandbox serves as a platform for actionable experimentation rather than premature trials (Jeník and Lauer 2017). Required levels of maturity will differ between types of sandboxes (e.g., early-stage versus late-stage).

A positive indicator of maturity – especially for a medium- or late-stage sandbox – may be having a functional and testable working

prototype, secured funding, a clear business model, and a demonstrated understanding of relevant regulatory requirements. For example, in the Hong Kong Monetary Authority's (HKMA) Fintech sandbox, participants must demonstrate measures to protect customers' interests, such as by providing timely and fair compensation for any financial losses caused by trial failures (Everhart 2020). In jurisdictions that apply this criterion–such as Slovakia's NBS Financial Services Regulatory Sandbox–project maturity is also assessed based on the presence of a comprehensive risk management strategy, including contingency planning for premature termination and threat identification. Similarly, the Zurich AI sandbox requires candidates to demonstrate the necessary technological expertise as a mandatory eligibility criterion. Conversely, a project at the concept stage or one that completely disregards legal and compliance requirements is unlikely to meet the eligibility threshold for such a sandbox type.

When it comes to AI, we can predict that, for instance, an AI company proposing a self-learning regulatory compliance system without any developed code or proof of feasibility would likely fail to qualify. Similarly, an AI solution that fails to acknowledge GDPR requirements for data quality and privacy (even when submitted for the AIA sandbox) would struggle to meet the necessary maturity criteria.

Among the three criteria analyzed, technological maturity or readiness appears to align most closely with what the AIA already mandates. From the outset, the AIA indicates that a certain degree of maturity is expected. Article 3(54-55) defines a sandbox plan as an agreement between the applicant and the competent authority, requiring clear objectives, testing methods, and a defined timeframe (further detailed in Article 58(2)(b)). These requirements inherently assume that the applicant has a workable model or at least a feasible approach–otherwise, the authority would have nothing to oversee or test.

Further support for this eligibility criterion can be found in Article 57(5), which states that the sandbox applies to AI systems "for a limited time before their being placed on the market or put into service." This suggests that applicants must already be developing a real solution–not yet commercially deployed, but advanced enough for meaningful testing.

The importance of the level of maturity (or technological readiness) can also be inferred from how risks arising from the sandbox experimentation are handled. Article 57(6)–(7) and Recital 139 require sandbox participants to cooperate with authorities to identify and mitigate significant risks to fundamental rights or health and safety. If any "significant risks to health and safety and fundamental rights" arise and cannot be mitigated, the authority may suspend or terminate sandbox participation (Article 57(11)). By implication, the system must have some basic risk controls or mitigation strategies in place.

For all these reasons, while we cannot yet read any reference to "maturity", the AIA implies that completely unformed or purely

> conceptual ideas are unlikely to gain acceptance since you need (1) a credible plan, (2) the ability to run meaningful tests and (3) minimal risk-mitigation controls.

Other, somewhat secondary, eligibility criteria were also identified (Seferi 2025, 151–52). One relates to the applicant's ability to identify and mitigate potential risks (4), which is closely linked to the maturity level of both the project and the organization behind it. The ability to identify and mitigate potential risks aligns well with the AIA's risk-based nature, which requires deployers, including prospective ones, to establish internal risk management systems.

Eligibility also depends on whether the project falls within the supervisory jurisdiction of the relevant authority, meaning that sector-specific legislation overseen by that authority governs it (5), and whether the project demonstrates a clear need for testing or experimentation (6).

(5) ensures that the project is subject to specific legislation that is explicitly within the authority's regulatory mandate. While some sandboxes—such as the Zurich AI Sandbox and to some extent the AIA itself—are sector-agnostic, covering domains like autonomous systems, sustainability, and healthcare (von Thiessen 2025), their flexibility remains bounded by applicable regulations. Accordingly, a national competent authority overseeing AIA-aligned sandboxes would reject projects involving technologies excluded under the AIA, such as autonomous weapons.

(6) focuses on whether the project can clearly justify the added value of testing in a controlled environment, showing that sandbox participation would deliver meaningful benefits, particularly through direct regulatory guidance provided by the authority. The AIA strongly implies the relevance of this criterion. In fact, applicants will likely need to show why regulatory supervision is necessary, as per the sandbox plan (Article 3(54) and Article 58(2)(b)), which requires detailing objectives, testing methods, and expected outcomes. Thus, an entity applying for sandbox participation should demonstrate a genuine need to clarify how to comply with the AIA (and possibly other Union or national rules) or to test risk-mitigation measures in a controlled environment.

Finally, some additional and less frequently cited eligibility criteria include the existence of clear exit strategies, the potential of the applicant's project to increase legal certainty by clarifying ambiguous or grey areas, and the submission of a transparent and complete application that ensures all necessary details are provided for a streamlined selection process. These criteria also seem to be implied by the AIA, as already pointed out. So, for instance, Article 57(7) mandates an exit report, signaling that projects should have a defined scope and testing lifecycle, and Article 57(9) highlights legal certainty as a core sandbox objective, favoring projects that help clarify ambiguous regulatory areas.

3.2. The operational procedure

Typically, the operational procedures of regulatory sandboxes progress through the following phases: application, preparation, testing, validation, and exit (Everhart 2020; OECD 2023). Below, each phase is described comprehensively as it pertains specifically to the regulatory sandboxes under the AIA.

3.2.1 Application Phase

Under the AIA, regulatory sandboxes operate primarily according to Articles 57 and 58. As mentioned, SMEs, especially start-ups, enjoy priority access to these sandboxes (Article 62(1)(a)). Their participation must be free of charge or fees proportionately reduced (Article 58(2)(d)), and microenterprises may fulfill certain quality-management requirements more simply (Article 63(1)).

Generally, the sandbox application process involves periodic (cohort-based), continuous (on-tap), or hybrid admission methods. Yet, unlike other sandboxes, such as the European Blockchain Regulatory Sandbox, which operate on periodic (cohort-based) or continuous (on-tap) admission models, the AIA does not explicitly define admission intervals. Instead, it allows national authorities the flexibility to establish their own admission frameworks, adapting to the needs of different AI ecosystems.

Applicants typically submit comprehensive documentation, including eligibility proof, team background and funding details, partnerships with financial institutions if required, a project intent letter, and a detailed testing and exit plan tailored to the sandbox environment (Seferi 2025, 155). Under the AIA, once an applicant–whether a provider or prospective provider–is admitted to the sandbox, the principal document governing their participation is the "sandbox plan" (Article 3(54)).

3.2.2 Preparation Phase

Upon acceptance, the authority typically defines case-specific testing parameters (e.g., error rates, data accuracy, system uptime) and establishes close collaboration with participants during testing (ESMA, EBA, EIOPA 2018, 18). It specifies test scope, conditions for modifying or terminating tests, success evaluation criteria (such as cost savings or improved security), sandbox duration, and post-sandbox transition plans.

The same happens under the AIA, where the sandbox plan is drafted jointly by the participant and the relevant competent authority. It must describe the participant's objectives (for instance, verifying compliance with certain AIA requirements or clarifying legal uncertainties such as how to conduct the fundamental rights impact assessment), as well as the planned duration of sandbox participation (which may vary according to the AI project's complexity and scale) and the testing and validation methodology.

Ultimately, the sandbox arrangement continues only as long as needed for the participant to meet the objectives stated in the sandbox plan. Competent authorities have the discretion to extend this period based on the project's complexity (Article 58(2)(h)). However, if a participant repeatedly fails to comply with the sandbox plan or if unmitigated high risks arise, the competent authority may suspend or permanently terminate that participant's sandbox activities (Article 57(11)).

3.2.3 Testing Phase

During the sandbox phase, participants receive tailored guidance from the competent authority regarding compliance with the AIA and relevant EU or national rules (Article 57(6)–(7)). In practice, this means the competent authority helps participants interpret the AIA's obligations, identify the relevant articles and requirements (e.g., risk management under Article 9 or data governance under Article 10), and verify risk-mitigation measures (Article 57(6)). Depending on the project's scope and the type of data involved, multiple competent authorities may jointly supervise the sandbox–for example, where the AI system processes personal data (including sensitive data), data protection authorities might also participate. The next section on governance addresses this multi-authority setup in more detail.

Inside the sandbox, providers can train, test, and validate their AI systems without immediately triggering all the AIA obligations that apply once the system is placed on the market. Training refers to developing or refining an AI model by exposing it to relevant datasets. Testing involves validating the AI system's performance, reliability, and safety before a commercial rollout. This may involve simulating and testing real-world scenarios, checking for accuracy, bias, or vulnerabilities, and refining algorithms as needed. It may also involve pilot deployments under controlled conditions or smaller trials with real or synthetic data. Monitoring means ongoing observation and assessment of the system's behavior: the provider and the regulator track whether the AI continues to meet relevant standards–especially for fundamental rights, health, and safety. Ultimately, by providing developers with access to a controlled, pre-market environment with more flexible rules, the aim of AIA sandboxes is to enable experimentation – for example with different data sources and algorithmic approaches – without exposing the public to undue risk (Recital 139, Article 57(5)).

Of course, there is a limit to how much can be achieved through simulated studies alone. Thus, if specified in the sandbox plan (Articles 57(5), 59, 76(2)), real-world testing may also take place inside the sandbox (an option that differs from the real-world testing regime outside sandboxes, Article 60). If severe risks arise during testing (including real-world testing), the authority can suspend or terminate the activity (Article 57(11)). However, according to Article 76(2)(1) AIA, real-world testing inside the sandbox needs to comply with the same requirements as real-world testing outside the sandbox, specified in Article 60. However, such real-world testing does not count as placing the system on the market or putting it into service. This distinction means that full AIA obligations—especially those for high-risk AI—do not yet apply. Only the terms agreed upon in the sandbox plan are binding. Participants may test high-risk AI with real users (who must give explicit consent, per Article 61), without triggering the full compliance regime. This ability to test under real-world conditions while temporarily exempt from high-risk requirements constitutes the first important legal relaxation constitutive of the sandbox under our definition.

A further significant legal relaxation in the AIA concerns personal data processing. Notably, while the GDPR imposes strict requirements on personal data handling, Article 59 of the AIA allows limited exceptions within sandboxes, particularly for data reuse, monitoring, and retention (Recital 140; Göbel and von Kruedener 2024, 756-757). This provision enables certain data uses that would not

be permitted under the GDPR alone, especially for training AI systems, aligning with GDPR Article 6(4) on secondary data use for purposes not initially contemplated during collection (Piltz and Weiss 2025, 93).

Article 59(1) also permits processing of sensitive data, which is significant given the broad interpretation of such data by the CJEU (e.g., Meta v. Bundeskartellamt).[4] Importantly, while personal data created in the sandbox may not be shared outside of it (Art. 59(1)(e)), the model itself (very likely irrespective of whether the model itself is understood as personal data, see Novelli, Casolari, et al. 2024) can be used outside of the sandbox even if trained making use of Art. 59 AIA; the legislator clearly did not have the model in mind when talking about personal data in this context, as it always refers to models as "model" in the AIA. Furthermore, if the trained system couldn't be used outside the sandbox, Article 59 would serve little purpose, especially given the AIA's emphasis on exit strategies.

To better illustrate the testing phase, consider a startup developing an AI-based clinical decision-support tool for healthcare professionals (classified as high-risk under the AIA). Once admitted to the sandbox, the startup can collect and use real-world data under the supervision of the competent authority and, if needed, a data protection authority. Under Article 59, it can even use data, including health data, collected in other contexts for the training of the model, if all requirements of said article are fulfilled. Such an oversight regime likely improves patient safety and proper data usage and helps the startup refine its algorithms (e.g., handling rare diseases as edge cases). Regulators can identify potential biases or high rates of false positives/negatives, and the startup can conduct pilot tests in controlled clinical settings with extra safeguards, before the system is "placed on the market" or introduced in hospitals (Leckenby et al. 2021). Indeed, the first cohort of the UK's AI Airlock included both OncoFlow - an AI system intended to help clinicians provide their patients with personalized management plans – and SmartGuideline – an LLM intended to help clinicians interact with clinical guidelines.[5]

Crucially, during this sandbox phase, the startup does not need to comply with every operational or documentation requirement that applies to a fully commercialized AI system (such as extensive post-market monitoring or standard labeling obligations). Instead, it is exempted from certain obligations as long as it adheres to the sandbox plan and follows the regulator's guidance. This allows the startup to gain real-world feedback and close regulatory support early on, without the full administrative burden of a commercial-ready AI system. Once the product is truly market-ready, all standard obligations take effect.

Interestingly, Recital 138 of the AIA specifies that AI regulatory sandboxes may be established in physical, digital, or hybrid formats. This opens the possibility for AI models to be tested in simulated environments, such as digital twins, which virtually replicate real-world scenarios. Leveraging such virtual environments may be advantageous as it allows comprehensive testing of AI models while reducing potential negative impacts, operational constraints, and

[4] Judgment of the Court (Grand Chamber) of 4 July 2023, Case C-252/21, ECLI:EU:C:2023:537, para. 73.

[5] https://www.gov.uk/government/publications/ai-airlock-pilot-cohort/ai-airlock-pilot-cohort.

resource demands associated with traditional, physical testing setups (Novelli et al. 2025).

3.2.4 Validation Phase

As the sandbox procedure nears its conclusion, reporting and documentation become particularly important. In general, participants must submit regular test results in line with the previously agreed parameters. Based on ongoing observations and feedback, necessary adjustments may be made to the innovation itself or its implementation strategy. In fact, it is during this phase that participants may also identify the need for additional actions, such as engaging external professionals for audits or cybersecurity assessments (Everhart 2020). The same empirical data may be used by regulators to adjust existing regulations. The duration of the experimentation phase is typically pre-defined, usually ranging from 3 to 36 months (Seferi 2025, 157).

In comparison to other regulatory sandboxes, the AIA not only mandates comprehensive reporting at exit but also places particular emphasis on detailed documentation as proof of regulatory compliance and a tool for regulatory learning. This approach is expected to facilitate a smoother market entry post-sandbox (which is quite complex given the AIA rules on licensing and certification) (Novelli, Hacker, et al. 2024). Additionally, the AIA requires annual public reporting and the exchange of best practices, a feature less commonly found in traditional sandboxes.

3.2.5 Exit Phase

The exit phase concludes the sandbox lifecycle with transition-focused actions. The sandbox duration remains flexible: it may end early if the participant achieves its objectives, opts to discontinue the AI system, or if the competent authority terminates testing due to disproportionately high risks that are not adequately addressed.

Upon the participant's request, the authority provides a written statement of the activities successfully completed in the sandbox (Article 57(7)). At the close of the sandbox period, the competent authority issues a detailed exit report on the testing and development, which may serve as evidence in future conformity assessments or market surveillance checks, potentially streamlining those processes (Article 57(7)–(8)).

However, as learned from other sandbox experiences, exit scenarios vary in success, ranging from most to least favorable: transition to full deployment (sometimes in the form of an authorisation[6]), limited approval to operate, extension of the testing period, and project termination. The applicant may apply for a full or limited license to operate in the broader market in the first two scenarios, if such a license is needed. In the third scenario, where additional data or adjustments are needed, participants might request an extension of the testing

[6] Sometimes it doesn't. In the ICO Sandbox, for instance, departing sandbox participants received a letter confirming that–based on the information they had supplied–the authority regarded their current personal-data processing as GDPR-compliant. The letter also made clear, however, that the ongoing responsibility for compliance always rests with the data controller.

phase. In the fourth scenario, unfavorable experimental results cannot justify any license, if required, and guarantees must be provided that obligations to existing customers are fulfilled and market stability is maintained.[7]

The competent authorities must respect confidentiality throughout this process, especially concerning trade secrets and personal data. Nevertheless, with the participant's consent, lessons learned or results may be shared to promote best practices or to inform the Commission and Board (Article 57(8)).

As long as the participant adheres to the sandbox plan and follows the authority's guidance in good faith, no administrative fines under the AIA will be imposed for any infringements that occur within the sandbox (Article 57(12)). Yet, the sandbox framework does not exempt the participant from liability if it causes harm to third parties, and existing civil or criminal liability under EU or national law still applies (Article 57(12)).

4. Outside the sandbox: the governance under AIA

A distinctive advantage of the AIA regulatory sandbox framework lies in its explicit multi-level governance structure, combining local, national, and EU-level oversight and coordination mechanisms. Unlike many traditional sandbox frameworks, which typically rely solely on single-level governance structures with limited coordination, the AIA sandboxes benefit from harmonized guidelines and oversight from the European Commission and the European AI Board. This multi-layered approach enhances regulatory consistency, reduces the risks of regulatory arbitrage, and facilitates broader dissemination of regulatory learnings and best practices across jurisdictions, providing unique systemic advantages compared to more fragmented or isolated sandbox approaches.

Under the AIA, this modular, multi-layered governance model allows Member States (or groups thereof) to meet the overarching goal of fostering AI innovation in flexible ways. Specifically, Article 57(1) of the AIA requires each Member State–acting through its national competent authorities–to ensure the establishment of at least one AI regulatory sandbox at the national level. To fulfill this obligation, Member States may either:

a) Create a new, dedicated sandbox, or
b) Participate in an existing sandbox–domestically or cross-border with other Member States–as long as the national territory is sufficiently covered.

Additionally, Article 57(2) permits Member States to set up further local or regional sandboxes alongside the mandatory national one.

In practice, the public bodies designated as "competent authorities" are usually those charged with enforcing the AIA in each Member State: e.g., a single authority (e.g., a digital or market-surveillance authority) or multiple agencies with shared responsibilities. Importantly, such authorities coordinate with specialized regulators, like data protection or consumer protection authorities, whenever sandboxed AI systems intersect with those regulators' mandates.

[7] Some of these exit strategies are report in this policy report by the World Bank, here: https://digitalregulation.org/a-case-for-ict-regulatory-sandbox/.

The multi-layered governance architecture becomes more complex, considering that different entities establish their own sandboxes. For instance, Article 57(3) allows the European Data Protection Supervisor to create a dedicated regulatory sandbox for AI systems used by EU institutions, bodies, offices, and agencies, in which case the EDPS assumes the role of (national) competent authority for that sandbox so that the EU institutions can test AI solutions they intend to adopt within a controlled environment. For example, if the European Anti-Fraud Office develops a machine-learning system to detect anomalies in financial transactions, testing such AI solution would fall within the scope of the EDPS regulatory sandbox, as EU institutions are not subject to national oversight in areas where the EU has its own supervisory framework (e.g., data protection).

Anyhow, whether it is a national or supranational competent authority, Article 57 and Recital 139 of the AIA specify that their core functions include defining rules, such as eligibility criteria or the format of exit reports, and establishing procedures. This entails clearly outlining objectives, ensuring ongoing supervision, regularly verifying compliance with legal requirements, and intervening if unmitigated risks arise (e.g., suspending the sandbox, as discussed earlier). They also need to lay down protective measures for participants, including rules on using personal data (as per Article 58), and set the timeline for the sandbox. All of this has to align with the Commission's rules on how to run these sandboxes (aka the implementing acts).

The twist is that the AIA grants national competent authorities a degree of discretion in exercising their supervisory powers within the bounds of relevant laws. Discretion manifests in several specific ways. First, authorities can adjust the sandbox framework to fit local realities. For example, while the AIA prioritises SMEs, authorities might further prioritise certain SMEs based on sector-specific goals, such as environmental sustainability, also depending on what the local economy needs. They may also simplify administrative requirements to encourage participation by local SMEs or startups with limited administrative capacity.

Second, authorities can adjust their approach to supervision, adopting a more proactive or reactive stance based on the perceived risk level of projects: higher-risk initiatives may require closer monitoring, while lower-risk projects could benefit from lighter oversight. Similarly, the duration of the sandbox may vary: e.g., predictive systems for public health crises might warrant longer timeframes, whereas simpler applications like customer-service automation could operate within shorter, focused periods.

Third, determine intervention, suspension, or termination thresholds (as per Article 57(11)), including acceptable risk levels during testing and triggers for mandatory corrective actions. And it can be context-specific as well. So, for instance, some authorities might intervene early if strong algorithmic biases are detected in AI systems for public administration, while others may allow extended experimentation if biases remain within acceptable limits, especially in a field or applications perceived as less sensitive.

Finally, depending on the specifics of each sandbox, authorities may exercise discretion in determining the extent and methods of collaboration with other regulators, such as data protection or consumer protection agencies. This coordination could range from informal consultations to formalized partnerships,

with the depth of cooperation tailored to the risk level of the AI model in question.

After participation in the regulatory sandbox, the authority issues an exit report, which is accessible to market surveillance authorities, notified bodies, the Commission, the AI Office, and the Board (for different tasks that we shall outline in the next sub-section). The exit report must include details on regulatory lessons, AI system performance, identified risks, and data quality to support certification processes. The report may also suggest modifications or further development steps. Additionally, the sandbox provides an opportunity to validate or reassess the initial risk classification of AI systems (e.g., low, medium, high, or unacceptable risk).

A distinctive feature of the AIA regulatory sandboxes is the requirement for authorities to submit annual reports to the AI Office and the European AI Board, as well as a final report upon the sandbox's conclusion (Article 57(16)), for transparency reasons and to assess the utility of the sandbox experience.

4.1. The Role of the Commission and EU-Level Coordination

To ensure a clear line of accountability against this backdrop of complexity, the national competent authority holds primary responsibility for the day-to-day oversight of regulatory sandboxes, including maintaining direct contact with sandbox participants and exercising immediate enforcement powers. Member States are tasked with ensuring that their competent authorities are adequately resourced financially and in terms of human resources.

Different national regulators may be brought in by the lead sandbox authority to co-supervise specific aspects within their jurisdiction. This occurs particularly when personal data protection or other specialized regulatory areas are involved. In such cases, data protection authorities or other relevant regulators must be formally associated with sandbox operations (Article 57(10)), ensuring compliance with GDPR and sectoral laws within their respective mandates.

However, national authorities are not the only institutional bodies involved. As outlined in Recital 143, the European Commission plays a critical role in providing the implementing acts to establish common principles for sandboxes across all Member States (e.g., eligibility criteria, terms and conditions, and standardised templates to ensure broad alignment). Additionally, the Commission may analyse annual reports submitted by national sandbox authorities. Although these reports are not directly addressed to the Commission, it can access them indirectly through the AI Board and the AI Office. The Commission uses these reports to inform its broader responsibilities under the AIA, especially for one of the most salient tasks, that is, enabling it to propose regulatory adjustments based on insights gained from sandbox testing (Novelli, Hacker, et al. 2024). For instance, if sandbox testing reveals an unanticipated high risk related to the accuracy of facial recognition algorithms, the Commission may propose amendments to delegated or implementing acts, such as introducing explicit accuracy thresholds.

Apart from national authorities and the Commission, various AI ecosystem actors also have advisory and technical roles. These include testing facilities,

research labs, innovation hubs, and civil society organisations (Recital 139; Articles 57–58). Their contributions focus on testing, data provision, risk mitigation best practices, and standardization input.

Sandbox coordination also involves the European AI Board and the AI Office. The AI Board can offer recommendations for regulatory updates based on sandbox reports. The AI Office manages a registry of sandboxes (Article 57(11)) and supports cross-border collaboration by sharing best practices, incidents, and recommendations. It may also assist Member States by providing templates for exit reports, test methodologies, or cooperation models for supervisory authorities.

5. Governance Challenges (and How to Handle Them)

The descriptive overview done so far can inadvertently create the impression that regulatory sandboxes are static activities–brief checkpoints innovators engage with merely as tick-box exercises–rather than dynamic processes designed to actively facilitate regulatory compliance for innovators and regulatory learning for authorities. Hence, conceptualizing regulatory sandboxes explicitly as processes rather than isolated activities is critical. Viewed this way, regulatory sandboxes encompass three distinct and interconnected phases: pre-testing, testing, and post-testing. Each of these phases has unique operational and governance requirements, and consequently, each presents its own distinct design and implementation challenges. Recognizing this is essential. As emphasized in the introduction, without structured transitions and effective learning mechanisms connecting each phase, regulatory sandboxes risk becoming isolated experiments rather than pathways to responsible AI deployment.

Below, we analyze the governance challenges at each stage, explaining their origins, implications, and offering actionable strategies to address them.

6.1. Pre-testing phase challenges and strategies

The pre-testing phase involves preparatory challenges before sandbox entry, including defining clear eligibility criteria, determining optimal pipeline placement, managing multi-agency coordination, preventing regulatory arbitrage, attracting suitable companies, integrating sandboxes with existing EU ecosystems, and avoiding unintended expansions of scope (scope creep). These early choices have a lasting impact and require structured, proactive management to ensure clarity, fairness, and coherence.

a) *Pipeline Placement.* Pipeline placement within regulatory sandboxes refers to determining the optimal stage – early conceptual design, mid-stage development, or late-stage pre-market deployment – at which an AI system should enter the sandbox. It also refers to the possibility for AI providers to revisit earlier development stages for updates or rechecks as needed. Under the AIA, the timing of sandbox entry significantly influences the supervisory approach and governance mechanisms applied to the sandboxed AI systems:

- Early-stage sandbox entrants in the initial stages of development require a guidance-oriented and collaborative approach to supervision. Authorities at this stage should proactively support compliance by assisting applicants in interpreting and integrating regulatory requirements into their system designs, providing tailored advice (Article 57(6)).
- Mid-stage sandbox entrants, for example those currently engaged in model refinement, require more structured, more continuous supervision, from multiple relevant authorities. These authorities should coordinate and actively monitor adherence to predefined sandbox plans, ensuring rigorous compliance with risk mitigation measures and data protection standards (Article 59). Non-compliance or failures at this stage should trigger targeted interventions or even suspension of the AI system's progress within the sandbox.
- Late-stage sandbox entrants, for example those in the throes of pre-market validation, require stringent supervision characterised by rigorous compliance reviews and conformity assessments. At this stage, the sandbox functions as a critical compliance checkpoint, filtering out systems that fail to meet essential regulatory requirements before they reach the market (Article 57(7–8)).

Strategy. Sandbox eligibility criteria and guidelines should explicitly delineate the different priorities, expectations, and goals of different pipeline entry points[8]. Guidelines for early- and mid-stage entrants should emphasise the importance of proactive compliance support, make clear that entry at this point is (a) designed to help both regulators and innovators address regulatory uncertainties early in the AI system's lifecycle, and (b) most suited to AI systems with uncertain risks. Conversely, guidelines for late-stage entrants should stress that entry at this point in the pipeline is designed to (a) help innovators and regulators manage critical risks and ensure robust market-readiness, and (b) is most suited to systems with clearly defined high risks (Recital 139 & Article 57).

b) *Multi-agency collaboration and overlapping oversight***.** We have seen that the regulatory sandbox framework under the AIA involves multiple authorities. National competent authorities must establish at least one regulatory sandbox, potentially involving regional or joint arrangements (Art. 57(1)–(2)). The European Data Protection Supervisor (EDPS) may set up a separate sandbox for EU institutions (Art. 57(3)). National data protection authorities (DPAs) must oversee data compliance, especially given the frequent use of personal data (Arts. 57(10), 59(1)). Other regulators—covering fundamental rights, safety, cybersecurity, or sectors like transport and energy—may also participate where relevant (Arts. 57(4), 57(10)).

[8] It is important to note that very early or very late stages are often excluded by eligibility criteria in existing regulatory sandboxes across various fields, a practice likely to be mirrored under the AIA.

The involvement of multiple authorities necessitates careful coordination and governance models, as it can lead to overlapping jurisdictions and potential conflicts. For example, AI systems placed in the sandbox might simultaneously fall under the oversight of general AI regulatory authorities and specialized sectoral regulators (healthcare, transport, energy, environment). As a result, regulatory sandboxes become convergence points where oversight from various legal frameworks intersects, complicating the clear delineation of regulatory responsibilities and accountability.

Regarding the EDPS sandbox and national sandboxes, direct regulatory overlap should generally remain limited due to their distinct primary focuses–the EDPS sandbox predominantly targets EU institutions and agencies. However, indirect overlaps could still occur, especially when AI systems involve cross-border data transfers or testing in multiple jurisdictions. For instance, if an EU agency develops an AI system that requires practical testing involving personal data collected nationally or implementation within Member States, coordination between the EDPS and relevant national authorities would become essential.

Strategy. A practical control strategy for mitigating these risks revolves around structured information and resource sharing. For instance, in the UK, the Information Commissioner's Office (ICO) and the FCA successfully implemented collaborative mechanisms by embedding staff within each other's sandbox initiatives. Such an arrangement facilitated alignment on overlapping issues like anti-money laundering (AML) and anonymisation and enabled joint participation in innovation-focused activities, such as jointly-run hackathons.

At the same time, we cannot assume that this approach will automatically succeed in the EU context. Concrete efforts are needed to mitigate the heightened risks of confusion and tension arising from the EU's complex regulatory and political landscape. Although the AI Board and the Commission's implementing acts strive for uniformity, the existence of separate sandboxes may nonetheless yield divergent interpretations regarding, for instance, compliance requirements and risk mitigation strategies. Projects involving both EU institutions and national entities may raise questions of jurisdiction and primary responsibility, potentially causing confusion among participants regarding which sandbox or supervisory framework to engage with. Finally, differences in funding, expertise, and infrastructure between national authorities and the EDPS may further compound these challenges. In essence, unlike the ICO and FCA coordinated approach, EU regulators have to face the challenge of systemic information sharing, underscoring the need for clear guidance and strong collaboration to manage multi-agency oversight. Cross-border cooperation and extensive sharing and exchange of use cases and testbeds will be crucial to facilitate the harmonization in the interpretation of AIA concerning regulatory questions on innovative AI systems.

c) *Regulatory arbitrage and race to the bottom.* Regulatory arbitrage refers to strategies firms employ to exploit differences or gaps between regulatory regimes to achieve similar economic benefits while avoiding stricter regulations (Pollman 2019). Exploiting the weaknesses of regulatory sandboxes, arbitrage can be a cascade effect of scope creep. Such arbitrage is particularly amplified within regulatory sandboxes, which inherently adopt a lighter-touch regulatory environment. Historical precedents such as Arizona's fintech sandbox, which weakened consumer protections (Allen 2019b, 312) , illustrate how regulatory arbitrage can encourage jurisdictions to progressively lower regulatory standards, known as a 'race to the bottom. Larger or more permissive sandboxes may exacerbate this trend because they attract more companies, increasing competitive pressures among jurisdictions. To remain competitive, jurisdictions may feel compelled to further relax oversight, potentially sacrificing consumer protection and regulatory rigor in exchange for economic benefits such as increased investments or job creation (Allen 2019b).

Race to the bottom can also be a direct side effect of close collaboration – aka pressure or lobbying – between regulators and firms, favoring undue private and very partial pressure (a phenomenon also known in the literature as regulatory capture). Increased engagement—and thus an elevated risk of regulatory capture—is common in sandbox environments due to frequent regulator-firm interactions. Therefore, transparency is essential to ensure waivers or tailored rules remain justified and not preferentia. Cases like Italy's sandbox, where no public information has been provided about the measures authorizing approved experiments, demonstrate how a lack of transparency can lead to opaque decision-making and potential unfair advantages for participants (Ranchordas and Vinci 2024). Interestingly, broader eligibility scopes can heighten the risk of regulatory capture by incentivizing firms to lobby aggressively for favorable regulatory adjustments.

Under the AIA, Article 58(1)(b) delegates the definition of eligibility criteria to the European Commission's implementing acts, but Member States retain considerable discretion regarding their application (Article 57). Differences in interpretation and implementation–such as varying entry criteria (defining innovation levels), operational rules, supervisory intensities, and intervention thresholds–may encourage providers to engage in "sandbox shopping," selecting jurisdictions that offer less stringent compliance requirements or more relaxed oversight. Additionally, a particular concern arises from SME prioritisation under Article 62(1), which waives participation fees and simplifies quality-management obligations. While beneficial for innovation, this provision could incentivise Member States to relax oversight to attract startups competitively. Jurisdictions with stricter standards may then face pressure to lower requirements to retain competitiveness.

Strategy. Although this is a significant challenge, the AIA already incorporates specific safeguards against excessive regulatory arbitrage (e.g., harmonization and standardization duties). Indeed, some degree of

jurisdictional competition, or sandbox shopping, might be both inevitable and even beneficial, provided it occurs transparently and maintains regulatory balance. Essential safeguards outlined by the AIA include harmonization efforts through the AI Board and AI Office to ensure minimal EU-wide standards, particularly in safeguarding fundamental rights, alongside periodic reporting, sharing of best practices, and comprehensive annual reports. Yet, as harmonisation consistency mechanisms to address such issues, something more could be done, like limiting regulatory relaxations through ex-post evaluation sunset clauses which concretely means that any waivers, exemptions, or tailored regulatory conditions (e.g., on data governance) granted within sandboxes would automatically expire after a defined period unless their renewal is explicitly justified through a rigorous post-trial evaluation. This might help keep regulatory sandboxes experimental and evidence-based rather than becoming default practices that incentivize regulatory arbitrage or pressure jurisdictions into competitive deregulation.

d) *Attract "good" companies*. A key challenge in regulatory sandbox implementation is ensuring appeal to high-potential companies–those with strong innovation and scalability prospects. Certain sandbox features, such as compliance burdens and time limitations (e.g., participation caps tied to project complexity under Art. 58(2)(h) of the AIA), may deter ambitious firms seeking rapid growth or large-scale testing. Additionally, the close regulatory oversight within sandboxes can be perceived as overly restrictive, discouraging participation. This issue has been observed in FinTech sandboxes, where criticism centered on excessive bureaucracy and limited pathways to commercialisation. Many successful startups bypassed formal sandboxes altogether, opting instead for direct market entry or informal testing frameworks. This risk is especially pertinent in the context of the EU's focus on competition with other economic blocs, as in many cases, organisations will be considering whether to base product development and related trials in the EU or elsewhere.

Strategy. While the AIA recognises this risk–encouraging simplified procedures for SMEs and startups (Art. 58(2)(g))–further measures are needed to attract high-growth firms. These could include: (1) demonstrating viable, commercial-scale pathways post-sandbox (e.g., fast-track licensing, market access support) and (2) aligning sandbox incentives with those of successful accelerator programs (e.g., funding opportunities, investor exposure).

Further, a proven way to attract strong companies is through structured stakeholder engagement. Although the AI Act references strategic plans (e.g., exit strategies), it does not yet explicitly require a stakeholder engagement planì. Early and systematic dialogue with industry players can clarify the commercial advantages (e.g., investor networks), mitigate concerns about restrictive oversight or intellectual property risks, and help sandbox designers align with market demands, track innovation trends, and streamline procedures. Crucially, such

engagement also equips local regulators with the necessary domain-specific insights to exercise informed discretion in tailoring sandbox operations to their local environments and economies. A virtuous example is, as pointed out by a report of the World Bank Organization, the Bank of Sierra Leone, which revised its sandbox framework after industry feedback to simplify procedures (Jeník and Duff 2021).

e) *Integration with the existing EU ecosystem*. To support the growth of innovative SMEs and startups, it's essential to connect them with existing EU frameworks. In fact, by providing structured environments that accompany them in the innovation cycle that leads to regulatory sandboxes, different infrastructures and initiatives in the EU ecosystem can offer support that simplifies compliance requirements and allows these innovators to focus on developing their technologies.

Strategy. AI Factories offer necessary resources to develop advanced AI models and provide access to cutting-edge technology that can significantly reduce the time and cost associated with bringing a product to market. Concurrently, European Digital Innovation Hubs (EDIHs) are the major point of entry and information for users and providers of AI systems, facilitating digital transformation across various sectors by offering training, networking opportunities, and access to funding. Finally, Testing Experimentation Facilities (TEFs), leveraging their AI expertise, can provide tailored support for establishing sandbox environments (EUSAiR Project 2025; cf. EU Commission 2025[9]). By collaborating closely, TEFs can concentrating on technical development or sectorial issues (being devoted to specific market domains) while sandboxes address legal considerations as they are directly managed by national competent authorities. By collaborating with these existing structures, regulatory sandboxes can enhance their value proposition, also providing SMEs the technical and operational support needed to thrive.

f) *Scope creep.* In regulatory sandboxes, the gradual and unplanned expansion of scope is a debated, but mostly theoretical issue (Knight and Mitchell 2020). Borrowed from project management (Lewis 2002), the concept of scope creep here describes the uncontrolled broadening of eligibility criteria or deviation from a sandbox's original objectives. Additionally, regulators themselves might inadvertently or deliberately misuse sandboxes for purposes beyond safe experimentation, undermining compliance standards.

Under the AIA, scope creep is particularly pertinent due to the Act's explicit intent to foster inclusivity by welcoming all providers or prospective providers. Although promoting broad participation is beneficial, this openness might inadvertently allow numerous incremental, minimally innovative projects into the sandbox.

[9] https://euagenda.eu/publications/a-brief-guide-about-the-european-ai-innovation-ecosystem

Moreover, scope creep concerns not only entry to the sandbox but also access to specific privileges within it. Particularly problematic is the flexible "public interest" criterion under Article 59(1), which participants may opportunistically leverage to access special sandbox benefits, such as enhanced personal data-use permissions, by framing their projects as fulfilling public interest objectives.

Strategy. More granular eligibility requirements may be of some help in this context. Just think about the criteria for the level of maturity (or technological readiness). Rather than relying on a single, undifferentiated "maturity" metric, regulators can specify which dimension of readiness they are measuring–manufacturing, commercial, or both–tailor the required threshold to the domain so that a life-critical area like healthcare faces stricter expectations than, say, logistics, and evaluate separate layers such as data readiness, operational fit, and governance within a transparent multi-dimensional framework (Lavin et al. 2022). A clear, disclosed methodology of this kind yields more balanced decisions and markedly reduces opportunities for regulatory arbitrage.

Having said that, a balanced strategy should be taken as an excessive focus on preventing scope creep should not lead to the opposite problem: overly restrictive eligibility criteria could result in cohorts of sandbox participants that are disappointly few in number, and/or non-representative, undermining the generalizability of evidence generated in the sandbox – e.g., about the cost or duration of regulatory adaptation for sandbox's participants – with other cohorts (Ranchordás 2021).

6.2. Testing phase challenges and strategies

The testing phase covers governance challenges encountered during active experimentation within the sandbox, specifically regarding integrating risk assessment frameworks, maintaining a realistic testing environment, and facilitating effective regulatory learning from sandbox outcomes.

a) *Integrating the AIA risk assessment model.* One key issue involves integrating the AIA's approach to risk assessment into regulatory sandbox procedures. This integration itself poses challenges due to a fundamental tension: sandboxes are intended precisely for testing and refining AI systems whose risks are uncertain or unknown, yet formal compliance under the AIA demands that risks be systematically identified, classified, and mitigated from the outset. As previously noted (section 4), this implies that providers and sandbox authorities will establish a provisional risk categorization of the AI system in the mandatory sandbox plan prior to sandbox entry and subsequently adjust this classification based on outcomes observed during sandbox testing, documented explicitly in the exit report.

Strategy. For this approach to be effective, sandbox authorities must be trained in the AIA's risk assessment methodologies and capable of applying its classification criteria consistently. Effective governance thus demands coordination among national authorities and with the AI Office, particularly regarding complex or cross-cutting risk factors (such as fundamental rights or cybersecurity). This may be especially challenging in the early stages of the AIA going into force, as local regulators seek to apply the AIA with no previous experience; this challenge should then diminish over time as those regulators gain experience and further guidance and precedents are shared. In this context, the need for standardization and speed is even greater than in regular (i.e., post-development) risk assessments, as sandbox procedures must accommodate legal oversight within a compressed timeframe.

It must also be noted that sandbox-specific risk assessments may differ significantly from real-world conditions, requiring potentially extensive adjustments post-sandbox. As a result, after a product exits the sandbox and is placed on the market, a comprehensive reassessment of risk may be necessary, taking into account real-world use and conditions not fully captured within the controlled sandbox environment.

b) *Realistic testing environment*. Poorly designed experiments have low internal validity (uncertainty about whether outcomes result from the intervention or external factors) and external validity (limited generalizability), undermining their policy relevance (Ranchordás 2021). To maximise benefits for participants–such as reduced regulatory uncertainty, guidance, market credibility, and investor confidence–the internal functioning of a regulatory sandbox must strike a careful balance: it should provide meaningful concessions to foster innovation while still maintaining a realistic regulatory environment where possible. Participants must be able to understand the difference between the sandbox environment and the real world, so they can make realistic plans for full deployment. Otherwise, participants could face severe compliance challenges upon exit, such as GDPR obligations, data localization requirements, or strict consent management rules. Without proper preparation, these regulatory shifts could render their innovations commercially impractical.

Strategy. To create a realistic regulatory environment, sandboxes could start with minimal, well-calibrated regulatory concessions–enough to encourage innovation–and then, in some cases, gradually reintroduce standard requirements over the testing period. In many cases, the sandbox participant is looking to experiment with one particular proposition or product feature. For example, if a participant seeks only to gauge customer reactions to a specific product feature–such as the rate of adoption of a biometric recognition technology perceived as either convenient or intrusive–the reintroduction of additional regulatory requirements (e.g., compliance with the AIA or GDPR around biometrics) might offer little to

no value. In such cases, a brief and focused sandbox trial, tailored precisely to the specific experimental goals, could be more beneficial.

Equally important is replicating the multi-agency oversight participants will encounter after exiting the sandbox. Access to realistic data is another critical factor. High-quality, real-world datasets–or synthetic alternatives that closely mirror actual conditions–enable accurate customer feedback, user testing, and compliance stress-testing within the sandbox. For example, incorporating simulated audits by data protection regulators during sandbox testing can help participants prepare for real compliance assessments.

To make regulatory sandbox outcomes more realistic and reliable, some scholars have proposed using control groups, as in scientific experiments. This approach compares sandbox participants against firms operating under standard regulatory conditions, providing a real-time benchmark (Ranchordás 2021, 27). Hence, control groups improve internal validity by clarifying whether observed outcomes result from regulatory intervention or external factors. They also strengthen external validity by helping regulators assess whether sandbox-tested innovations can scale under real-world conditions. Yet, implementing control groups remains challenging, especially in emerging fields like AI, where few comparable firms operate under full regulatory compliance.

Currently, the AIA does not prioritise "experimental realism" in regulatory sandboxes, with a focus on short-term innovation support, particularly for SMEs and startups. This is understandable as an initial goal for a new sandbox regime. A medium-term goal could be to develop more complex and rigorous testbeds that simulate the full regulatory environment. This could improve the long-term scalability and reliability of lessons drawn from sandbox experiments.

c) *Regulatory learning***.** Reviewing and integrating insights from regulatory sandboxes constitutes a critical challenge. Regulatory sandboxes frequently fail to realise their full learning potential due to inherent design flaws, such as their casuistic approach and non-representative participant cohorts (Ranchordás 2021). Consequently, policymakers–including regulators and other key stakeholders–should prioritise establishing a robust, standardised methodology applicable across different national competent authorities (Dimitropoulos and Hacker 2016).

Strategy. Effective regulatory learning presupposes the existence of reliable, standardised success and failure metrics, along with clear procedures for translating sandbox outcomes into regulatory action–such as rule modifications or updated guidance. Without such mechanisms, sandboxes risk becoming isolated experiments with no systemic impact.

In this regard, the AIA provides only limited operational clarity. While Articles 57(7), (8), and (16) require exit reports and annual reports documenting sandbox results and lessons learned, the Act does not establish a harmonised mechanism for integrating these insights into the regulatory framework. Furthermore, the role of institutional bodies is

constrained: the AI Office and the Board may access exit reports only with explicit consent (Article 57(8)), and there is no provision for a centralized, automatic review or synthesis of sandbox findings across Member States. This lack of coordination inhibits horizontal learning and cross-jurisdictional regulatory coherence.

Additionally, although Article 57(16) states that sandbox insights should inform potential revisions to the AIA text (and its implementing acts), the Act lacks a structured mandate for how findings are to be channeled into formal legislative updates, standardization processes, or the development of codes of conduct (Article 58(2)(e)). Likewise, the AIA does not establish a formal mechanism for addressing conflicts or tensions revealed in sandboxes between AI-specific rules and sectoral or fundamental rights regimes (e.g., data protection). No cross-sectoral trigger exists to prompt regulatory coordination or legal reform in response to such conflicts, limiting the sandbox's capacity to drive systemic regulatory improvement.

Therefore, the legislator must establish a binding, harmonised mechanism, with specific metrics, to translate sandbox insights into regulatory updates and coordinated cross-sectoral reforms. In the startup phase of sandboxes, sharing use cases and testbeds among national competent authorities will be crucial to support the development of these metrics, especially since technical thresholds are still undefined.

6.3. Post-testing phase challenges and strategies

The post-testing phase involves three distinct categories of challenges: **post-sandbox compliance**, which covers regulatory decisions about market readiness and the enforcement of rules after exiting the sandbox; **transition challenges**, which concern the practical hurdles companies face in moving from sandbox participation to full market deployment; and **success evaluation**, which assesses whether the sandbox itself has effectively balanced innovation, market competition, and regulatory learning objectives.

a) *Post-sandbox compliance*. After testing, regulators must decide if the firm can operate under standard or modified rules. This depends on whether the firm met its objectives—e.g., technical feasibility, user uptake, or functional performance—and whether the product shows real value at scale. Regulators also assess whether the sandbox trial ran safely, with effective safeguards and no major breaches (e.g., fraud, consumer harm, or compliance failures) (Jeník and Duff 2021). Another key consideration is whether the firm is prepared for full regulatory compliance post-sandbox (Rathnam 2024), a challenging transition that some navigate by using test results to meet licensing requirements, as seen in the FCA Regulatory Sandbox. Operational viability is also scrutinised, including whether the company has the necessary resources, technology, and business model to succeed commercially. Additionally, regulators review the firm's exit strategy, since sandbox programs, like Singapore's MAS guidelines, often

require participants to outline a clear transition plan upfront. Finally, the decision is informed by detailed reporting from the firm throughout the testing period, covering performance metrics, risk management, and compliance. These reports help regulators gauge factors like consumer adoption rates, adherence to sandbox rules, and the innovation's broader market potential before granting approval for wider deployment.

Strategy. For AI-specific sandboxes, these foundational criteria should be supplemented with additional requirements. Participants may need to conduct algorithmic audits to detect biased outcomes and demonstrate concrete measures to prevent discrimination alongside standard consumer protection assessments.

The AIA implicitly accommodates these criteria, though with varying degrees of explicitness. Article 57(7) directly emphasises regulatory compliance readiness, while Article 57(5) implies requirements for operational viability through sandbox plan evaluations. However, effective implementation demands that Member States and EU institutions provide sandbox authorities with sufficient resources–including financial support, technical expertise, and specialized training in AI governance. Moreover, to ensure consistency, the Commission should establish clearly measurable graduation criteria and standardised guidance – e.g., via implementing acts – for post-sandbox compliance.

b) *Success evaluation***.** Evaluating whether a regulatory sandbox has been successful is crucial for all stakeholders (also for the abovementioned regulatory learning reasons). Regulators must determine if the sandbox effectively balances innovation and control or if alternative approaches (e.g., innovation hubs or forbearance) would be more effective. Sandbox designers and regulators must evaluate if adjustments are required, such as participant recruitment or resource allocation. Participants also need clarity on whether sandbox participation aligns with their goals.. However, measuring success presents challenges, particularly in selecting appropriate metrics.

Strategy. Success criteria should include impact on innovation, market competition, consumer protection, and regulatory learning.

The impact on innovation can be evaluated by examining whether the sandbox facilitated the development of new ideas into viable products or services. Relevant metrics include the number and diversity of innovations, the rate at which sandbox-tested solutions progressed to full market deployment (Zetzsche et al. 2017; Herrera and Vadillo 2018), and the level of venture investment attracted (Goo and Heo 2020). For example, research has found increased venture funding, enhanced patenting activity, and improved survival rates for sandbox participating firms (Cornelli et al. 2024) in the case of the FCA Regulatory Sandbox.

The sandbox's ability to promote market entry and competition can be assessed through the speed and cost of obtaining licenses or regulatory authorization, subsequent access to funding, the conversion rate (i.e., how

many participants successfully transition from testing to full market operation), and the emergence of new business models, especially from startups or non-traditional actors. These new entrants often introduce more efficient or productive approaches, contributing to overall sector dynamism.

Additional dimensions include consumer protection and risk mitigation, which can be assessed through metrics such as the number of safety incidents, consumer complaints, or direct feedback on user satisfaction.

Equally important is regulatory learning. Success in this area can be evaluated through post-sandbox assessments that explore whether regulatory staff gained insights from working with innovative firms (maybe through exit reports)[10], whether the sandbox led to the identification and resolution of regulatory gaps (Appaya 2020)[11], and whether it fostered sustained engagement between regulators and innovators. The frequency and quality of feedback loops, such as workshops or scheduled check-ins, can serve as proxies for this interaction.

Under the AIA, many of these metrics are explicitly or implicitly justified through its key sandbox objectives, such as improving legal certainty for innovators, facilitating knowledge-sharing across authorities, and enabling evidence-based regulatory learning. Notably, Article 57(9)(e) emphasises the support of AI innovation–particularly from SMEs and startups–by encouraging safe experimentation and accelerating market entry.

Thus, in the context of the AIA, a meaningful success evaluation framework would align with how well each of these objectives is met. When EU Member States implement these AI sandboxes, they should wonder: Did the sandbox lead to clearer regulatory guidance (regulatory learning)? Did it help an AI startup bring a product to market (innovation and market access)? And was this achieved without compromising users' fundamental rights (consumer protection)? Such questions should also inspire their design.

c) *Transition challenges*. Entering a regulatory sandbox is not an end in itself; the ultimate goal for participants is a successful and sustainable market launch at scale. Consequently, some of the most significant supervisory and governance challenges arise from sandbox experimentation to full market deployment during the transition phase. To address these challenges, sandboxes, including those established under the AIA, typically require clear, structured exit plans. Nevertheless, certain transition

[10] In this regards, the Joint ESAs report in Europe found that running sandboxes helps supervisors "increase their knowledge about financial innovations, the risks and opportunities they entail" (p.3). Link: https://www.esma.europa.eu/sites/default/files/2023-12/ESA_2023_27_Joint_ESAs_Report_on_Innovation_Facilitators_2023.pdf.

[11] For instance, Malaysia's central bank (BNM) tweaked its sandbox approach after early cohorts – the evaluation prompted creation of a "Sandbox Express" to streamline testing for lower-risk innovations, showing the regulator learned and adapted the process to better meet industry needs (Appaya 2020, 41).

challenges remain inevitable, and regulators should proactively guide and prepare participants for these realities.

Strategy. Firstly, completing sandbox testing successfully does not guarantee immediate market approval. Companies often encounter a compliance "cliff edge" post-testing, meaning they must quickly establish comprehensive compliance measures or fulfill capital requirements previously relaxed within the sandbox. To mitigate this risk, firms should develop detailed plans for the post-sandbox phase well in advance, anticipating compliance demands and potential costs. Secondly, sandbox testing may reveal the need for legal or policy changes, aligning with regulatory learning objectives. However, if these legislative changes are delayed, companies exiting sandboxes might struggle to fully scale operations due to persistent regulatory constraints (Appaya 2020). We should not forget that delays in transition can result in losing first-mover advantages, potentially benefiting competitors who are not bound by sandbox limitations or larger firms capable of faster adaptation. Regulators must, therefore, remain agile, committing to prompt policy adaptations and, where necessary, granting temporary regulatory waivers to support innovators until permanent legislative adjustments are in place. Thirdly, companies leaving sandboxes frequently face resource gaps. Sandboxes usually provide structured support such as regulatory guidance, mentorship, and industry connections. However, this support typically diminishes or ends after exit. Consequently, firms must independently secure funding and expertise to effectively scale their operations. Proactively addressing these resource needs–including fundraising, hiring compliance personnel, and scaling technical systems–is essential for a smooth transition.

Additionally, supervisory challenges previously highlighted, such as interagency coordination, substantially affect transition ease. Post-testing supervision and governance thus deserve equal attention alongside initial sandbox design.

Under the AIA specifically, transition challenges may be particularly acute. While benefits for sandbox participants–especially SMEs and start-ups–are substantial during the entry phase (e.g., simplified eligibility criteria), fewer supports are available post-sandbox. Although Article 57(7) of the AIA mentions that conformity assessments might be accelerated based on sandbox exit reports, the vague formulation "accelerating [...] to a reasonable extent" may invalidate the effort. Moreover, the AIA sandbox environment primarily evaluates regulatory compliance rather than operational scalability or large-scale performance. While the AIA allows sandbox testing under real-world conditions (Article 57(5)), inherent constraints limit full market replication (and scaling beyond EU borders introduces additional complexity).

To strengthen the transition process under future implementations of the AIA, several solutions are recommended: (1) operationalise genuinely simplified and accelerated conformity assessments explicitly linked to sandbox exit documentation; (2) establish explicit regulatory feedback

loops; (3) broaden conditions for real-world testing and encouraging its use–for example, by introducing provisions for "conditional market access" if real-world testing meets predefined benchmarks, similar to the supportive role played by sandbox exit reports; (4) Encourage proactive consumer trust-building initiatives within the sandbox plans; (5) develop funding pathways, including a European AI Innovation Fund, for successful sandbox alumni.

6. Conclusions

The point of departure for this article was a familiar–but often overstated–dilemma: how regulators can impose necessary obligations on AI developers without stifling the innovation they aim to foster. Rather than forcing a zero-sum choice between regulation and innovation, regulatory sandboxes enable regulators and innovators to collaboratively test, refine, and proportionately adjust regulatory frameworks in response to real-world evidence. However, not all sandboxes achieve this balance equally effectively. As our analysis of the proposed regulatory sandboxes under the AIA demonstrates, their effectiveness ultimately depends on their design, specifically, on the robustness and coherence of the governance mechanisms established throughout their lifecycle.

Specifically, our analysis has highlighted essential governance levers across all three sandbox phases. During the pre-testing phase, clearly defined eligibility criteria, precise pipeline placement guidance, structured multi-agency coordination, and integration with the broader EU innovation ecosystem are essential. In the testing phase, maintaining experimental realism, continuously updating risk classifications, and replicating realistic regulatory oversight are critical for ensuring that sandbox experimentation closely mirrors real-world conditions. In the post-testing phase, structured transitions, sustained support mechanisms, clearly defined exit pathways, and rigorous yet streamlined conformity assessments are necessary to ensure that successful sandbox innovations transition smoothly to market deployment, thereby effectively supporting regulatory learning.

If these governance levers are pulled coherently across all three phases, regulatory sandboxes can fulfil the AIA's twin goals: (a) offering SMEs and start-ups an intelligible route to market, and (b) providing regulators with an evidence base for iterative rule-making firmly rooted in fundamental-rights protection. Conversely, sandboxes may become mere token experiments, failing to protect fundamental rights and consumer interests, and ultimately hindering rather than helping innovators.

Future research should explore how governance practices might be tailored to the risk level of AI systems to enhance the proportionality of sandbox oversight. Further investigation into cross-jurisdictional regulatory harmonization could help reduce arbitrage risks and support the emergence of coherent global standards. Longitudinal studies on the downstream impacts of sandboxed innovations would provide valuable insights into their sustainability and market effects. Finally, examining how sandboxes interact with broader

innovation support mechanisms—such as hubs and accelerators—could enhance their integration and overall contribution to the AI innovation ecosystem.

Bibliography

Allen, Hilary J. 2019a. 'Regulatory Sandboxes'. *George Washington Law Review* 87 (3): 579–645.

———. 2019b. 'Sandbox Boundaries'. *Vanderbilt Journal of Entertainment & Technology Law* 22 (2): 299–322.

Appaya, Mandepanda Sharmista. 2020. 'Global Experiences from Regulatory Sandboxes', November. https://policycommons.net/artifacts/1246170/global-experiences-from-regulatory-sandboxes/1801669/.

Bradford, Anu. 2024. 'The False Choice Between Digital Regulation and Innovation'. SSRN Scholarly Paper. Rochester, NY. https://doi.org/10.2139/ssrn.4753107.

Collingridge, David. 1980. *The Social Control of Technology*. St. Martin's Press.

Cornelli, Giulio, Sebastian Doerr, Leonardo Gambacorta, and Ouarda Merrouche. 2024. 'Regulatory Sandboxes and Fintech Funding: Evidence from the UK*'. *Review of Finance* 28 (1): 203–33. https://doi.org/10.1093/rof/rfad017.

Dimitropoulos, Georgios, and Philipp Hacker. 2016. 'Learning and the Law: Improving Behavioral Regulation from an International and Comparative Perspective'. *Journal of Law and Policy* 25 (2): 473–548.

ESMA, EBA, EIOPA. 2018. 'FinTech: Regulatory Sandboxes and Innovation Hubs'. ESMA, EBA, EIOPA. https://www.esma.europa.eu/sites/default/files/library/jc_2018_74_joint_report_on_regulatory_sandboxes_and_innovation_hubs.pdf.

EUSAiR Project. 2025. 'EUSAiR Project'. EUSAiR. 2025. https://eusair-project.eu/.

Everhart, Jonathan R. 2020. 'The Fintech Sandbox: An Overview of Regulatory Sandbox Regimes'. *Southern Journal of Business and Ethics* 12:64–73.

Financial Conduct Authority. 2015. 'Regulatory Sandbox'.

Göbel, Maximilian, and Alexis von Kruedener. 2024. 'KI-Reallabore Und Innovationsförderung in Der KI-VO'. *GRUR-Prax*, 2024.

Goo, Jayoung James, and Joo-Yeun Heo. 2020. 'The Impact of the Regulatory Sandbox on the Fintech Industry, with a Discussion on the Relation between Regulatory Sandboxes and Open Innovation'. *Journal of Open Innovation: Technology, Market, and Complexity* 6 (2): 43. https://doi.org/10.3390/joitmc6020043.

Herrera, Diego, and Sonia Vadillo. 2018. 'Regulatory Sandboxes in Latin America and the Caribbean for the FinTech Ecosystem and the Financial System'. *IDB Publications*, March. https://doi.org/10.18235/0007982.

Jeník, Ivo, and Schan Duff. 2021. 'How to Build a Regulatory Sandbox : A Practical Guide for Policy Makers'. Text/HTML 161314. World Bank Group. https://documents.worldbank.org/en/publication/documents-reports/documentdetail/en/126281625136122935.

Jeník, Ivo, and Kate Lauer. 2017. 'Regulatory Sandboxes and Financial Inclusion | CGAP Research & Publications'.

https://www.cgap.org/research/publication/regulatory-sandboxes-and-financial-inclusion,
https://www.cgap.org/research/publication/regulatory-sandboxes-and-financial-inclusion.
Knight, Brian, and Trace Mitchell. 2020. 'The Sandbox Paradox: Balancing the Need to Facilitate Innovation with the Risk of Regulatory Privilege'. *South Carolina Law Review* 72 (2). https://scholarcommons.sc.edu/sclr/vol72/iss2/7.
Lavin, Alexander, Ciarán M. Gilligan-Lee, Alessya Visnjic, Siddha Ganju, Dava Newman, Sujoy Ganguly, Danny Lange, et al. 2022. 'Technology Readiness Levels for Machine Learning Systems'. *Nature Communications* 13 (1): 6039. https://doi.org/10.1038/s41467-022-33128-9.
Leckenby, Emily, Dalia Dawoud, Jacoline Bouvy, and Páll Jónsson. 2021. 'The Sandbox Approach and Its Potential for Use in Health Technology Assessment: A Literature Review'. *Applied Health Economics and Health Policy* 19 (6): 857–69. https://doi.org/10.1007/s40258-021-00665-1.
Lewis, James P. 2002. *Fundamentals of Project Management*. AMACOM. http://archive.org/details/fundamentalsofproolewi_1.
Makarov, Vladislav O., and Marina L. Davydova. 2021. 'On the Concept of Regulatory Sandboxes'. In *'Smart Technologies' for Society, State and Economy*, edited by Elena G. Popkova and Bruno S. Sergi, 1014–20. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-59126-7_112.
Novelli, Claudio, Javier Argota Sánchez-Vaquerizo, Dirk Helbing, Antonino Rotolo, and Luciano Floridi. 2025. 'A Replica for Our Democracies? On Using Digital Twins to Enhance Deliberative Democracy'. SSRN Scholarly Paper. Rochester, NY. https://papers.ssrn.com/abstract=5190735.
Novelli, Claudio, Federico Casolari, Philipp Hacker, Giorgio Spedicato, and Luciano Floridi. 2024. 'Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity'. SSRN Scholarly Paper. Rochester, NY. https://doi.org/10.2139/ssrn.4694565.
Novelli, Claudio, Philipp Hacker, Jessica Morley, Jarle Trondal, and Luciano Floridi. 2024. 'A Robust Governance for the AI Act: AI Office, AI Board, Scientific Panel, and National Authorities'. *European Journal of Risk Regulation*, September, 1–25. https://doi.org/10.1017/err.2024.57.
Parenti, Radostina. 2020. 'Regulatory Sandboxes and Innovation Hubs for FinTech | Think Tank | European Parliament'. Policy Department for Economic, Scientific and Quality of Life Policies Directorate-General for Internal Policies. https://www.europarl.europa.eu/thinktank/en/document/IPOL_STU(2020)652752.
Piltz, Carlo, and Alexander Weiss. 2025. 'Datenschutzrechtliche Rechtsgrundlagen Für Das Training von KI-Modellen'. *EuDIR*, 2025.
Pollman, Elizabeth. 2019. 'Tech, Regulatory Arbitrage, and Limits'. *European Business Organization Law Review* 20 (3): 567–90. https://doi.org/10.1007/s40804-019-00155-x.
Ranchordás, Sofia. 2021. 'Experimental Regulations and Regulatory Sandboxes: Law without Order?' *Law and Method* 2021. https://doi.org/10.5553/REM/.000064.
Ranchordas, Sofia, and Valeria Vinci. 2024. 'Regulatory Sandboxes and Innovation-Friendly Regulation: Between Collaboration and Capture'. *Italian Journal of Public Law* 16 (1): 107–39.

Rathnam, Lavanya. 2024. 'Regulatory Sandboxes: How They Help Fintech Innovate Safely'. *Planet Compliance* (blog). 19 June 2024. https://www.planetcompliance.com/compliance-software/regulatory-sandboxes-how-they-are-helping-fintech-innovate-safely/.
Seferi, Fabio. 2025. 'A Comparative Analysis of Regulatory Sandboxes from Selecred Use Cases: Insights from Recurring Operational Practices'. In *Regulatory Sandboxes for AI and Cybersecurity - Questions and Answers for Stakeholders*. Cybersecurity National Lab. https://cybersecnatlab.it/wp-content/uploads/2025/02/CybersecNatLab-White-Paper-Regulary-Sandboxes.pdf.
Thiessen, Raphael von. 2025. 'Learning From The Ai Sandbox In Zurich: A Practical Perspective'. In *Regulatory Sandboxes for AI and Cybersecurity - Questions and Answers for Stakeholders*. Cybersecurity National Lab. https://cybersecnatlab.it/wp-content/uploads/2025/02/CybersecNatLab-White-Paper-Regulary-Sandboxes.pdf.
Zetzsche, Dirk, Ross Buckley, Janos Barberis, and Douglas Arner. 2017. 'Regulating a Revolution: From Regulatory Sandboxes to Smart Regulation'. In *Fordham Journal of Corporate & Financial Law*, 23:31. https://ir.lawnet.fordham.edu/jcfl/vol23/iss1/2.