Thinking Outside the Box?

Regulatory Sandboxes as a Tool for AI Regulation

Hannah Ruschemeier[1]

[1] Fernuniversität Hagen, Universitätsstraße 27, 58084 Hagen
hannah.ruschemeier@fernuni-hagen.de

Abstract. This legal article deals with the concept of regulatory sandboxes as part of experimental regulation to meet the challenges of digitalisation for legal regimes. It focuses on a critical analysis of the new provisions of the AI Act and their weaknesses in relation to the objectives of regulatory sandboxes to promote innovation and regulatory learning. Finally, suggestions for improvement and future areas of application are outlined.

Keywords: legal regulation, AI Act, regulatory sandboxes, innovation, regulatory learning, EU law

1 Introduction

1.1 Regulating emerging technologies

Digital technologies, in particular algorithm-based decision-making systems discussed under the term "artificial intelligence" (AI), pose special challenges to the concept of legal regulation. In addition to the characteristic of irreversibility, certain digital technologies represent a qualitative leap in that they lack an analogous counterpart in the physical realm. Predictions generated by algorithms are based on the analysis of vast quantities of data and thus require automated processes exclusive to the digital environment. The need for legal guardrails as a response seems undisputed: the internet is supposedly a legal vacuum,[1] the prospect of deep-fakes and misinformation looms as a dystopic potential outcome,[2] AI-technology is used in war depending on the favour of private companies.[3] On the government side, there is uncertainty and a lack of knowledge about what is regulated and what is not. Consequently, there is a great need for future-oriented and resilient legal frameworks.[4]

The question of whether to regulate does little to answer the difficult details of the "how", which begins with the procedural aspects. The development of technology and legislative processes often exhibit stark differences since the law in the form of legislation and technology move at different paces. On the one hand, the development of digital technology is flexible, dynamic, and progressing at a rapid pace. Law, on the other hand, especially in the form of legislation, is slow, reflecting the fact that negotiating compromises and thus majorities in democratic processes takes time. On the contrary, technical innovations are a "moving target", which challenges the often reactive functioning of the law.[5] The declared aim of political initiatives is therefore to create

"robust", "future-proof" and "flexible" legal requirements.[1] Finding the right way forward is complex: abandoning regulation by law is not an option; entire laws cannot be changed in anticipatory obedience without closely looking at potential impacts. This dilemma creates a need for flexible but effective regulation, which could be met by experimental regulatory instruments, such as experimental clauses, regulatory sandboxes and temporary laws and evaluations.[6] For example, the discussion about banning Bitcoin at the European level,[7] shows how difficult the legal struggle for the right answers to disruptive technologies is. Regulatory projects are often (unfairly) accused of stifling innovation.[8,9] This is very present in the political debate, and many believe there is a risk of disproportionate requirements that will stifle innovation, pushing it out into other less stringent jurisdictions.[10]

Instruments of experimental regulation are supposed to contribute solutions to these conflicts. In general, experimental regulation combines empirical evidence and legal requirements in a more flexible model than "traditional" legislation.[11] It foresees the testing of innovative products and services are to be tested in a supervised environment in close cooperation with the competent supervisory authority within a time-limited framework, often with the application of substantive legal exceptions or no enforcement letters.[12] The aim is to promote innovation, as well as to generate expert knowledge on the government side in order to adapt the regulatory framework or to gain insights for new administrative and legislative processes. Experimental regulation is not new, but remains relatively unexplored from a legal perspective in the field of regulating digital technologies.[13–15] Instruments of experimental regulation, which have been underresearched in the context of law and new technologies,[12,16,17] could contribute to a solution by providing a flexible framework for the generation of state regulatory knowledge while also providing processes to promote innovation. Regulators can use real-world laboratories to create a testing ground for new technologies in which real-world legal requirements do not need to be enforced during an experimental phase, thus providing insights into the object of regulation.[18]

1.2 Experimental regulation: Regulatory sandboxes

Regulatory sandboxes are part of the toolbox of new epistemic methods.[12,16] The AI Act now explicitly calls for the establishment of regulatory sandboxes as a measure to promote innovation in every member state on a national level, Articles 57 et seq.[19] The AI Act defines regulatory sandboxes as a controlled framework set up by a competent authority which offers providers or prospective providers of AI systems the possibility to develop, train, validate, and test innovative AI systems, where appropriate in real-world conditions, pursuant to a sandbox plan for a limited time under regulatory supervision, article 3 (55). Recital 138 explicitly mentions the idea of regulatory sandboxes as part of a legal framework that promotes innovation and is future-proof and resilient.

[1] Cf. Recital 138 EU Regulation 2024/0138 European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)) (AI Act).

Furthermore, Recital 139 states that the aim is to support innovation and regulatory learning and counteract the information deficit in relation to the regulatory subject of "AI systems", Article 3 (1).[2] Sandboxes could provide for a controlled environment that fosters innovation and facilitates the development, training, testing, and validation of innovative AI systems for a limited time before release onto the market or put into service pursuant to a specific sandbox plan agreed between the prospective providers and the competent authority.[3] Pursuant to Recital 139, participation in the regulatory sandbox should focus on issues that raise legal uncertainty for providers and prospective providers, allowing them to innovate, experiment with AI in the Union, and contribute to evidence-based regulatory learning.

In highly complex, technical areas, there is often a knowledge gap between the regulatory authorities and the specifics of the field to be regulated, specifically because ex ante, the unpredictability of these highly complex, dynamic technologies often eludes traditional forecasting schemes for hazard prevention, risk regulation, prohibition with reservation of authorisation, and ex ante risk assessment. In these fields there is a structural lack of information, as seen in the fintech-sector.[20,21] In addition, as evolving and disruptive technologies have a cross-sectional effect and touch on different areas and fields of law, their impact often simply cannot be assessed.[22] This is why there is a growing call for new legal instruments that can cope with digital transformation. Given the negative effects already apparent, a failure to react until all effects of new technologies are fully known or making wholesale changes to existing law is clearly undesirable.

Nevertheless, regulatory sandboxes cannot and should not replace effective regulation or create large-scale exemptions from it. The focus at the regulatory level should therefore continue to be on creating effective regulatory requirements in the digital sector that protect fundamental rights and promote the public good. The AI Act is a first step in the right direction, but it still contains many gaps and inadequate provisions.[23]

This paper examines regulatory sandboxes and related concepts from a public law regulatory perspective and adds to the ongoing debate, which primarily discusses regulatory sandboxes from the perspective of promoting innovation and the economy[16,21,17]. The potential for regulatory learning on the part of the supervisory authorities is another

[2] The AI Act defines AI-System as "a machine-based system designed to operate with varyin levels of autonomy, that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments".

[3] Recital 139 states: The objectives of the AI regulatory sandboxes should be to foster AI innovation by establishing a controlled experimentation and testing environment in the development and pre-marketing phase with a view to ensuring compliance of the innovative AI systems with this Regulation and other relevant Union and national law, to enhance legal certainty for innovators and the competent authorities' oversight and understanding of the opportunities, emerging risks and the impacts of AI use, to facilitate regulatory learning for authorities and undertakings, including with a view to future adaptions of the legal framework, to support cooperation and the sharing of best practices with the authorities involved in the AI regulatory sandbox, and to accelerate access to markets, including by removing barriers for SMEs, including start-ups.

important but underrepresented consideration and an explicit goal of the AI Act, Article 57 (9 d).

1.3 Concept of regulatory sandboxes

The term 'sandbox' used in the AI Act originates from the field of computer science. In that field, sandboxes are generally referred to as isolated or quarantined areas where the effects are contained and do not affect the infrastructure, platform or business-critical networks on which they run.[24] The idea of regulatory sandboxes has similarities and differences. On the one hand, sandboxes can be specifically designed to provide a framework for promoting innovation by creating a 'safe space' for testing new business ideas within an existing or newly planned regulatory regime.[25] On the other hand, the aim of regulatory sandboxes is precisely to generate external effects and consequences for observation. At an abstract level, the overall aim of the sandbox is to gain knowledge of unknown facts, consequences and risks in all areas.

The broader concept of sandboxes, where different scenarios are tested in a real environment, comes from the social sciences. Similar to the idea of sandboxes in computer science, they act for example as test spaces for sustainable urban development.[26] However, not every real-world lab is a regulatory sandbox. Many real-world labs simply combine science and practice without necessarily leading to legal issues or the involvement of a regulator.[26] These sandboxes are not included in the following analysis.

Regulatory sandboxes are also referred to as 'real laboratories', or rather derisively as 'sandbox playgrounds'. Compared to the sandbox concept in social science, regulatory sandboxes are narrower in scope and mostly aim at improving innovation and regulation. Regulatory sandboxes therefore involve a regulator and a regulated party. They act as test rooms for innovation and regulation at the same time, allowing the testing of technologies, services, products or approaches that are only partially compliant with existing legal and regulatory frameworks, or where compliance is unclear. The first sandbox-style framework was established by the US Consumer Financial Protection Bureau (CFPB) in 2012, dubbed 'Project Catalyst'. This programme aimed to foster the development and expansion of innovative consumer financial products and services. The objective was to collaborate with the community of innovators to ensure that superior financial products and services were accessible to American consumers.[25] Even if the CFPB did not formally announce their initiatives as a sandbox, it followed policies such as No Action Letters[4] that fulfilled many functions of a sandbox.[27]

Outside the US, the UK's Financial Conduct Authority (FCA) has been a pioneer of the specific concept of regulatory sandboxes, testing hundreds of applications in its sandbox since 2016.[28] In its sandbox, the FCA follows the concept of a market-driven regulatory sandbox with annual 'cohorts' selected from a general pool of aspiring innovators. Different forms of regulatory sandboxes are now being discussed. In contrast to general purpose sandboxes, thematic sandboxes are designed to pursue specific

[4] No-action letters are designed to recognise the value of innovative financial technologies by committing the CFPB to take no enforcement or supervisory action with respect to the subject matter of the letter, Information Collection, 81 Fed. Reg. 8686, 8692 (Feb. 22, 20169).

policy objectives that are thematically limited.[25] This is usually done by limiting them to specific technologies, products, or business models.

1.4 Legal context

By creating a normative framework for new regulatory approaches, the law can make a decisive contribution to proactively shaping developments in digital transformation. Experimental regulation can be an important aspect of the development of the law. Future-proof regulation requires flexibility on the one hand, and legal certainty and resilience on the other. Digital technologies are limitless, which makes it all the more important to make the best use of existing legal possibilities in the interests of democracy, the rule of law, and the common good.

In general, regulatory requirements, like all government action, are based on the law and the constitutional principle of proportionality (in the EU and all Member States); restrictions and requirements must be set in relation to the potential damage to legitimate interests worth protecting, e.g. consumer protection, security, etc. However, there is already an information deficit in many digital technologies. As a result, the potential damage, or more precisely the potential risk, is not known. Different risk levels are reflected in national and European legislation: pharmaceuticals have to undergo an approval procedure, while operating a permanent business only requires notification. The idea of adapting traditional regulatory structures is obvious, as the "analogue" effects of digital technologies are more difficult to grasp. For example, when Facebook was launched, few people foresaw its development into one of the most powerful companies in the digital economy with an impact on political processes, elections and other public opinion-forming activities.

Significantly, the matter at hand concerns a de facto regulation of something whose evolution is not yet known. This poses challenges for both sides, the product owner or developer and the relevant regulator. Structural uncertainty is not new to law, especially public law, most recently gaining public awareness during the pandemic. Unlike other related disciplines such as social and political sciences, experiments are not part of the tradition of administrative science and administrative law. In addition, highly regulated areas of law have often been developed for the interface of market access and thus do not correspond to the reality of digital products, which continue to evolve dynamically once they are already on the market. Regulatory sandboxes therefore offer a first opportunity to gain information and knowledge for regulators and participants at the same time. On the other hand, it is important to ensure that regulatory sandboxes should not become a carte blanche for risky products of dubious legality. While it should not be forgotten that product providers can design their products to comply with the applicable law, regulatory sandboxes do provide an opportunity for supervisors to gain valuable knowledge about the products they regulate while also allowing for the legal framework to be adapted to current developments. In all innovation efforts, the limits of the separation of powers must be kept in mind. Although the legislator has the democratically legitimacy to decide on changes to the legal framework, the knowledge gained from a sandbox can only be provided by the administration. The protective purpose of the regulation is crucial: if this is not achieved de lege lata, the compatibility with the existing requirements is of little use. A combination of the two approaches therefore seems promising.

Regulatory sandboxes are intended to mitigate the the danger of regulation consistently trying to catch up with current developments and new legal requirements being outdated even before they are adopted. Their application to matters involving an information deficit is less clear. Information deficits exist, for example, in decision-making systems where the path from input to output is not 100% traceable, either because of the volume of data or the complexity of the calculation. This can be addressed in a number of ways, most commonly through a general ban or general authorisation of such technologies, through an ex-ante authorisation procedure as in pharmaceutical law, or through a ban subject to authorisation. Regulatory sandboxes could come into play to reduce the information deficit in all variations, but should not aim at deregulation or at reducing protection standards. Instead, sandboxes should serve to create an appropriate regulatory framework to bridge the gap between new technologies and regulations that may have been developed prior to the invention of the smart phone, let alone the app being developed.

In other countries, regulatory sandboxes have already established themselves as an instrument in highly regulated areas such as the financial sector. So far, real-world labs have primarily been discussed in the context of financial market regulation, fintechs, blockchain, or crypto-assets. In Germany, they are a comparatively new phenomenon and not particularly popular: the Federal Financial Supervisory Authority has so far rejected regulatory sandboxes, citing a lack of mandate from the administration,[29] while the German Finance Committee rejected a proposal for regulatory sandboxes based on the UK proposal in mid-2020, citing consumer protection and conflicting European law.[30,31]

Regulatory sandboxes are now explicitly addressed at European level by the AI Act in Article 57 et seq. The AI Act aims to classify potential risk into different risk classes. It provides for prohibited AI systems (Article 5), high-risk systems (Article 6 et seq.), and low-risk systems to which only general obligations apply (Article 50). Risk is defined as the combination of the likelihood of harm and the severity of that harm, Article 3 (2). The group of high-risk systems will have the greatest practical relevance, as they are defined according to a dual regulatory approach of product safety law and the protection of fundamental rights. AI systems are considered high-risk under the AI Act if they constitute a product or a safety component of a product according to the harmonised provisions of Annex I, and are also subject to conformity assessment by third parties or pose a threat to fundamental rights in the case of the examples of use in Annex III. The danger of AI systems may be obvious in some cases, such as social scoring. Fundamentally, however, the problem remains that law alone cannot regulate the multidimensional nature of risks, especially in the case of AI. An “all or nothing” approach does not seem promising either, but it is important to be aware of the limits of legal regulation from the outset.

2 Examples

Most examples of regulatory sandboxes can be found in the fintech sector. This is due to the fact that this is a highly regulated market, which is at the same time highly innovative and has undergone significant changes on the supply side in recent years.

In Norway, the federal data protection authority established a "regulatory privacy sandbox" in 2020 to establish and stimulate privacy-enhancing innovation and digitalisation.[32] Every year 3-4 applicants were selected, resulting in 12 projects in total having been tested in the sandbox by 2023. Participants came from the private and public sectors in the fields of health, transport, environment, and digital consumer services. In addition, generative AI applications were tested in a targeted manner.[33] The DPA documented the selection of participants in a transparent and comprehensible manner, and the sandbox has already been evaluated. Its hosting by the Norwegian DPA means it has retained the clearly defined objective of privacy enhancing technologies as central. The documentation, evaluation, and time limit of the sandbox, as well as the clear competences and focus on the DPA's objectives promote ethical and responsible application without limiting official supervision. The guiding principles rely on the ethics guidelines for trustworthy AI from the High-Level-Expert Group on AI set up by the European Commission.[34] After a pre-defined timeframe, the DPA publishes a detailed exit report. In a comment on the Commission's draft AI law, the Norwegian DPA echoed the general criticism of the lack of precision in the Commission's draft AI law in its concerns about the concept of regulatory sandboxes, stating: "However, we also see a need for some guidance on how competent authorities can strike a good balance between being a supervisory authority on the one hand and giving detailed guidance through a sandbox on the other. We propose that the AI Act specifies that participation in a sandbox does not constitute a stamp of approval, and that the organization/controller is still accountable for its processing of personal data."[35]

France has also set up an "Edtech sandbox" under the French data protection authority. This aims to help participants to develop and include data protection by design as required by the GDPR into their products. So far, the French approach does not provide for enforcement exceptions and is limited to legal and technical assistance to participants by the authorities.

Spain quickly followed suit in 2020, launching its regulatory sandbox on AI even before the AI Act came into force. The regulatory sandbox is housed in the newly established AI supervisory authority, the Spanish Agency for the Supervision of AI, located at the Ministry of Digital Transformation.[36] Thus far, Spain has established two sandboxes, one for the financial system, pursuant to the Law 7/2020 of 13 November on digital transformation of the financial system, and another for the electricity sector as a result of Royal Decree 568/2022 of 11 July establishing the regulatory framework for research and innovation in the electricity sector. In contrast to other sandboxes, the Spanish sandbox will be set up essentially to serve as a vehicle for studying the operability of the requirements of the AI Act, It is expected to result in reports on best practices and the compilation of technical guidelines for execution and supervision based on the evidence obtained, rather than collaborating with the authorities in defining and developing an adequate regulatory framework.[37]

There are currently no regulatory sandboxes at federal level in Germany. In the political debate, the term "real-world laboratories" is used more often than "regulatory sandboxes". However, North Rhine-Westphalia (NRW) was the first federal state to launch the "Digi-Sandbox.NRW" project. So far however, these projects do not specifically focus on AI. There are also no plans for cooperation with a specific supervisory authority, but there is some support for the establishment of real-world laboratories.

At the federal level, a Real World Laboratory Act is planned, which provides for a one-stop-shop principle with a competent authority. The Federal Ministry of Economics and Technology (BMWi) has presented a concept for a Real-World Laboratory Act, which is intended to enshrine overarching standards for real-world laboratories and experimental clauses in law, and also to enable new real-world laboratories in important areas of digital innovation.[38] A public consultation phase was completed in 2023. There has so far been a lack of legal standards for real-world laboratories. Potential areas of application under discussion include AI applications in the field of modern mobility or Industry 4.0, innovative digital identification procedures, e.g. for digital driving licences, and digital legal services and procedures. The planned mandatory review of the experimental clause in the legislation is particularly interesting, as it is intended to allow the continuous identification of new areas of application. The main aim of the concept is to create innovation-friendly and thus economically favourable prospects for companies and strengthen Germany as a business location. However, another persistent problem is that the legislator is lagging behind digital transformation without actively shaping it.

3 Regulatory Sandboxes and the AI Act

Regulatory sandboxes, experimental clauses, and experimental regulation in general are relatively unknown in EU law. This is partly due to the fact that experimental regulatory approaches are viewed with suspicion, as a legal vacuum incompatible with legal certainty or the unity of the legal order.[39] The AI Act now creates the first horizontal regulatory regime for sandboxes on the Union-level. The following is an outline of what can be expected for regulatory sandboxes based on the AI Act as released

3.1 Establishment of regulatory sandboxes

This requirement for at least one AI regulatory sandbox as established in Article 57 AI Act can also be satisfied by establishing this sandbox jointly with the authorities of other member states. The obligation can also be fulfilled by participating in an existing sandbox as long as that participation provides an equivalent level of national coverage for the Member States, Article 57 (1). Article 52 (2) provides for additional regulatory sandboxes to be established at a local and regional level, thereby showing that the requirements are not intended to be exhaustive. The goal is to provide a controlled environment for the development, testing, and validation of innovative AI systems under the direct supervision and guidance of the competent authorities. Article 57 (9) explicitly states that the establishment of regulatory sandboxes shall follow the objectives of improving legal certainty and compliance with the AI Act and other applicable law,

supporting best practices, fostering innovation, contributing to regulatory learning and facilitating access to the Union market for start-ups and SMEs.

The European Data Protection Officer may establish a regulatory sandbox at the Union level for Union institutions. Article 57 (4) also requires competent authorities be adequately resourced for regulatory sandbox tasks, as well as cooperation and collaboration between the authorities and with the AI Office where appropriate.

3.2 Procedural aspect and governance

In terms of procedure, participants and the supervisor must agree on a specific sandbox plan. Instructions to participants are optional, but the supervisor is required to provide guidance on supervisory expectations, Article 57 (5-7). The publication of final reports should have been mandatory rather than subject to the consent of participants, Article 57 (8). This would, for example, improve the situation for consumers who are better able to verify the "sandbox" label. The involvement of data protection authorities and other competent authorities is declaratory. Article 57 (11) states that any significant risks to health, safety and fundamental rights identified during the review of the sandbox will lead to immediate risk mitigation, including the temporary or permanent suspension of the testing process. National competent authorities must submit annual reports to the AI Office and to the Board, from one year after the establishment of the AI regulatory sandbox and every year thereafter until its termination and the issue of a final report. Annual reports or abstracts thereof will be made available to the public online.

3.3 Substantial design and exceptions

The substantial design elements of the regulatory sandboxes are laid down in Articles 57 and 58 of the AI Act. In temporal terms, the regulatory sandboxes as envisioned only apply to AI systems before entering the market or service, Article 57 (5). The AI Act does not stipulate legal exemptions or obligations for non-enforcement. However, Article 57(12) specifies that authorities will not impose fines for any infringements of the AI Act itself, provided prospective providers adhere to the outlined plan and conditions for participation, and faithfully follow the guidance of the competent national authorities. There are therefore no exceptions for other areas of law.

Unsurprisingly, given that these decisions are the responsibility of Member States, the AI Act makes it clear that participants in the sandbox will remain liable to third parties for any damage caused as a result of testing in the sandbox, Article 57 (12).

The specific design of real-life-testing conditions remains unclear as, according to Article 58 (1), the exact modalities of real-world laboratories will be defined in the implementing acts of the Commission. As a result, the AI Act does not regulate the important issues of selection criteria and eligibility, procedures and applications, or requirements and conditions for sandbox participation. This is unfortunate, as these are precisely the factors that will determine whether a regulatory sandbox will succeed. Article 58 (2) provides a list of goals that delegated acts must pursue, such as transparent and fair criteria, equal access, flexibility of the national authorities, the inclusion of other actors in the "AI ecosystem" such as standardisation bodies or start-ups, clear and simple communication of sandbox entry and exit conditions, time limits, facilitation of regulatory learning factors such as accuracy, robustness, security and risk mitigation measures for fundamental rights and society at large. These objectives, combined with the new instrument of sandboxes and the new requirements of the AI Act, are too vague

and undefined to derive a specific design. Moreover, the important decision on these criteria should not have been left to the Commission alone.
The only reference to the actual design of the sandboxes is to be found outside the binding part of the Regulation: Recital 138 mentions that regulatory sandboxes could be established in physical, digital, or hybrid form and may accommodate physical as well as digital products. The reference to physical products is consistent with the dualistic regulatory approach of Article 6 AI Act which classifies AI-Systems as high-risk-systems when they are subject under product safety law. Product safety law partly includes software systems, such as under the medical device regulation,[5] but strongly follows a physical understanding of products which characterizes the entire AI Act.[40]

Article 59 of the Regulation provides legal topics for discussion by providing an exception to the basic principle of purpose limitation of data protection under Article 6 (1a, 4) GDPR if the conditions set out in paragraph 1 are met. For this purpose, the AI systems covered must be developed to safeguard a substantial public interest falling within the areas listed in Art. 59 (1) (a i-v). These include public safety and health, a high level of environmental protection, sustainable energy, safety and resilience of transport systems and mobility, and the efficiency and quality of public administration. This breach of the principle of purpose limitation for collective public interests is a novelty in data protection law, which continues to focus on protecting the fundamental rights of individuals.[41] In line with the objective of promoting innovation, SMEs should be given priority access to the AI regulatory sandboxes However, the participation of other companies is not excluded under the conditions of Article 62(1). This approach is not convincing, as providing the benefit of regulatory sandboxes to big tech companies will further increase their market power. It would have been preferable to restrict access to regulatory sandboxes to small, SMEs, start-ups and public organisations, which are not even mentioned in the AI Act.

3.4 Testing outside of regulatory sandboxes

The final version of the AI Act significantly weakened the concept of regulatory sandboxes for regulatory learning. In addition to the regulatory sandboxes, Article 60 now provides the opportunity to test high-risk AI systems under real-life conditions outside of regulatory sandboxes. This is problematic for several reasons. First, the AI Act is silent on what the actual real-life conditions should be. We do know natural persons may be involved, as Article 60 (4 g) states that participants must be adequately protected where required by their age or physical or mental disability. Additionally, Article 61 requires informed consent from participants and Article 60 (5) stipulates that test participants under real conditions may terminate their participation in the test at any time by withdrawing their informed consent without justification and request the immediate and permanent deletion of their personal data without incurring any disadvantage.

Second, the possibility of testing under real-life conditions outside regulatory sandboxes massively weakens the potential for regulatory learning. This is not explicitly standardised as an objective of the Article 60 procedure, but the authority retains its traditional authorisation power. As a result, no expert knowledge and, above all, no new

[5] Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC.

knowledge about the products in question is generated. Article 76 merely stipulates that competent authorities must supervise tests carried out under real-life conditions in a sandbox. Also of concern is the notional authorisation in Article 60(4)(b) which deems testing under real-life conditions to be authorised 30 days after application to the market surveillance authority. This puts considerable pressure on the competent authorities, but the rule should remain that authorisation is mandatory.

Third, Article 76 provides for the possibility of carrying out the tests under real-life conditions in a regulatory sandbox, which makes it much more difficult to distinguish between the two procedures. The conditions under which tests can be carried out under real conditions in the regulatory sandbox is unclear. In case of doubt, firms will opt for the simpler procedure outside a regulatory sandbox, with no benefit for supervisors.

4 Criticism

The proliferation of regulatory sandboxes is not without controversy. The recent explosion in the number of sandboxes suggests that some regulators are simply setting up a sandbox to take advantage of this trend. However, sandboxes are only effective tools if they are set up with the financial and human resources to run sustainably. A report commissioned by the United Nations Secretary General's Special Advocate for Inclusive Finance for Development (UNSGSA) found that "around a quarter of regulators have launched sandbox initiatives without first evaluating feasibility, demand, potential outcomes, or collateral effects."[42]

Furthermore, criticism has focused on the potential for inadequate consumer protection, dilution of regulation, overly generous exemptions, and unequal treatment. Germany's Federal Financial Supervisory Authority (BaFin) has long rejected regulatory sandboxes, arguing that promoting innovation is a market matter that does not require official involvement. Consumer groups and others have warned that regulatory sandboxes are bad for consumers.[43] The result could be a 'race to the bottom' for the least possible regulation to attract start-ups and other businesses.[44] There are also concerns that there will be no real exchange between authorities and companies, but that the latter will use the sandboxes as a regulatory discount and only for PR purposes. As sandboxes become more popular and more companies advertise with sandbox validation, it may become harder for consumers to distinguish between sandbox testing and real validation, resulting in companies and their new products being wrongly perceived as more trustworthy. The AI Act also provides for official confirmation of participation in the sandbox, which companies can then use for their own purposes, Article 57 (7).

This could be countered, for example, by supervisory authorities such as the FCA no longer publishing the list of sandbox participants, but instead offering confidential advice. This is prevented however by the transparency of such administrative action required to allow a large number of applications for sandboxes, rather than providing confidential advice to a few selected firms. Transparency is also the only way to ensure effective monitoring, for example, whether criteria for approval are consistent with the requirements of non-discrimination. In the area of fintech, some argue regulatory sandboxes can lead to "riskwashing". *Brown* and *Piroska* argue that sandboxes ease the introduction of fintech into society and finance to the extent that sandboxes themselves

become a part of a fintech-financialization apparatus that intensifies penetration into typically non-financialised social relations with potentially socially disruptive effects.[39] The critical analysis of fintech and regulatory sandboxes as solutionism, particularly in relation to people who do not have access to banking services,[39] cannot be transferred seamlessly to AI and other technologies which cover a much wider range of applications. Nevertheless, the tendency towards solutionism is particularly clear in the field of AI, without first critically questioning what problems can be solved by technological innovation.[45–47] The one-sided promotion and creation of sandboxes can lead to the uncritical adoption of narratives and discourses from individual industries.

All these objections are understandable and realistic. Especially in the area of digital technologies, it is not the promotion of innovation that has been underrepresented so far, but rather the lack of effective regulation. Under no circumstances should regulatory sandboxes be used to undermine often new regulatory requirements. The success of a sandbox in promoting innovation and regulatory learning, without creating loopholes that undermine consumer interests and other regulatory objectives, depends on its specific design. Well-designed sandboxes can address these criticisms and still successfully achieve their objectives, not only focusing on specific test runs, but also contributing their findings to the regulatory debate in a targeted way. Additionally, the current uncertainty surrounding regulations contributes to only a minor fraction of the challenges encountered by newcomers on the digital market. It is therefore not judicious for regulatory bodies to allocate excessive resources to formulating exemption policies.[27] Instead, their efforts should be concentrated on dismantling barriers to the debut of new products that are not only compliant but also have the potential to enhance welfare.

Sandboxes should therefore, where they provide legal benefits such as exemptions from data protection rules, be located exclusively with the competent supervisory authorities, with a focus on regulatory learning and on targeting applications that promote the public good.

5 Conclusion and Outlook

The basic idea of regulatory sandboxes is a good approach to meeting the challenges of digitalisation with legal means. However, the AI Act should have been bolder in its specific design. Regulatory sandboxes should not provide legal rebates and should be located exclusively within the competent authorities. The resulting benefits for firms and the associated costs are only justified if supervisors also benefit. Nevertheless, the effectiveness of the AI Act heavily relies on its enforcement and Member State implementation of regulatory sandboxes. The requirement for all Member States to have or be involved in an AI specific regulatory sandbox thus creates a unique opportunity to develop and compare best practices. Ongoing evaluation should lead to the harmonisation of standards and the adoption of sound designs. Among the many other tasks involved in digital supervision, there is also a need for dialogue between competent authorities. Sandboxes should not be used as a fig leaf to blindly promote innovation, encourage lax regulation, or create a privacy discount. Not every innovation or development is desirable, but should be in the public interest. Sandboxes should therefore

focus more on regulatory learning, particularly for the benefit of regulators, rather than smoothing the path for companies. There is still much potential for a key focus on the protection of fundamental rights and individuals, e.g. with regard to stakeholder participation.

Finally, well-functioning sandboxes can be expanded in their scope, allowing benefits beyond the administration. New technologies can also be tested in a targeted way, so that the lessons learnt from the sandboxes can feed not only into administrative implementation, but also into future legislative processes or evaluations. Hence, legislators can use the results sandboxes to inform future legislation on digital topics and technologies.

Acknowledgments. This contribution is part of the project "Trial and Error. Experimentelle Regulierung im Mehrebenensystem" funded by the Daimler und Benz Stiftung.

Disclosure of Interests. The author reports no conflicting interests.

References

1. Newman, N. The Costs of Lost Privacy: Consumer Harm and Rising Economic Inequality in the Age of Google. *Wm. Mitchell L. Rev.* **40**, 849 (2013).
2. Barman, D., Guo, Z. & Conlan, O. The Dark Side of Language Models: Exploring the Potential of LLMs in Multimedia Disinformation Generation and Dissemination. *Machine Learning with Applications* **16**, 100545 (2024).
3. Frenkel, S. Israel Deploys Expansive Facial Recognition Program in Gaza. *The New York Times* (2024).
4. Ranchordás, S. & Roznaˀi, Y. Future-Proofing Legislation for the Digital Age. in *Time, law, and change: An interdisciplinary study* (eds. Ranchordás, S. & Roznaˀi, Y.) 347–366 (Hart Publishing; Bloomsbury Publishing, [Oxford]; [London], 2020).
5. Beutel, F. K. The Lag between Scientific Discoveries and Legal Procedures. *Neb. L. Rev.* **33**, 1 (1953).
6. Ranchordás, S. Experimental Regulations for AI: Sandboxes for Morals and Mores. *Morals & Machines* **1**, 86–100 (2021).
7. Fanta, A. Climate measures: Behind closed doors, EU officials talk about banning Bitcoin. https://netzpolitik.org/2022/climate-measures-behind-closed-doors-eu-officials-talk-about-banning-bitcoin/.
8. Renda, A. & Pelkmans, J. EU regulation: hindering or stimulating innovation? in *Handbook of Innovation and Regulation* 263–293 (Edward Elgar Publishing).
9. Smuha, N. A. From a 'race to AI' to a 'race to AI regulation': regulatory competition for artificial intelligence. *Law, Innovation and Technology* **13**, 57–84 (2021).
10. Müller, R. KI: Kann die EU künstliche Intelligenz nur regulieren? *FAZ.NET* (2024).
11. Ranchordás, S. Experimental Regulations and Regulatory Sandboxes: Law without Order? *Law and Method* **2021**, (2021).
12. Makarov, V. O. & Davydova, M. L. On the Concept of Regulatory Sandboxes. in *'Smart Technologies' for Society, State and Economy* (eds. Popkova, E. G. & Sergi, B. S.) 1014–1020 (Springer International Publishing, Cham, 2021).
13. Ranchordás, S. Innovation-Friendly Regulation: The Sunset of Regulation, the Sunrise of Innovation. *Jurimetrics* **55**, 201–224 (2015).
14. Ranchordás, S. *Experimental Regulations and Regulatory Sandboxes: Law without Order? SSRN Electronic Journal* (2021). doi:10.2139/ssrn.3934075.

15. Ranchordás, S. & van 't Schip, M. *'Future-Proofing Legislation for the Digital Age'*. *SSRN Electronic Journal* (2019). doi:10.2139/ssrn.3466161.
16. Buocz, T., Pfotenhauer, S. & Eisenberger, I. Regulatory sandboxes in the AI Act: reconciling innovation and safety? *Law, Innovation and Technology* **15**, 357–389 (2023).
17. Zetzsche, D. A., Buckley, R. P., Arner, D. W. & Barberis, J. N. Regulating a Revolution: From Regulatory Sandboxes to Smart Regulation. SSRN Scholarly Paper at https://doi.org/10.2139/ssrn.3018534 (2017).
18. Krönke, C. Sandkastenspiele – »Regulatory Sandboxes« aus der Perspektive des Allgemeinen Verwaltungsrechts. *JZ* **76**, 434–443 (2021).
19. Botta, J. Die Förderung innovativer KI-Systeme in der EU. *ZfDR* 391–412 (2022).
20. Ahern, D. Regulatory Lag, Regulatory Friction and Regulatory Transition as FinTech Disenablers: Calibrating an EU Response to the Regulatory Sandbox Phenomenon. *Eur Bus Org Law Rev* **22**, 395–432 (2021).
21. Bromberg, L., Godwin, A. & Ramsay, I. Fintech Sandboxes: Achieving a Balance between Regulation and Innovation. SSRN Scholarly Paper at https://papers.ssrn.com/abstract=3090844 (2017).
22. Ruschemeier, H. AI as a challenge for legal regulation – the scope of application of the artificial intelligence act proposal. *ERA Forum* **23**, 361–376 (2023).
23. Smuha, N. A. *et al.* How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act. *SSRN Electronic Journal* 33899991 (2021) doi:10.2139/ssrn.3899991.
24. Prevelakis, V. & Spinellis, D. Sandboxing Applications. in *USENIX Annual Technical Conference, FREENIX Track* 119–126 (Citeseer, 2001).
25. Donelan, E. *Regulatory Governance: Policy Making, Legislative Drafting and Law Reform*. (Springer International Publishing, Cham, 2022). doi:10.1007/978-3-030-96351-4.
26. Wagner, F. Reallabore als kreative Arenen der Transformation zu einer Kultur der Nachhaltigkeit. in *Die Experimentalstadt: Kreativität und die kulturelle Dimension der Nachhaltigen Entwicklung* (eds. Reinermann, J.-L. & Behr, F.) 79–94 (Springer Fachmedien, Wiesbaden, 2017). doi:10.1007/978-3-658-14981-9_5.
27. Quan. A Few Thoughts on Regulatory Sandboxes. *Stanford PACS* https://pacscenter.stanford.edu/a-few-thoughts-on-regulatory-sandboxes/.
28. Regulatory Sandbox. *FCA* https://www.fca.org.uk/firms/innovation/regulatory-sandbox (2022).
29. Hufeld, F. Neujahrspresseempfang der BaFin. (2016).
30. BT-Drucks. 19/19506. Regulatory Sandboxes – Für mehr Innovationen im Finanzmarkt.
31. Eberle, N. Die „Regulatory Sandbox". 175–179 (2020).
32. 20 applied for the sandbox. *Datatilsynet* https://www.datatilsynet.no/en/news/aktuelle-nyheter-2023/20-applied-for-the-sandbox/.
33. Time for generative AI in the sandbox. *Datatilsynet* https://www.datatilsynet.no/en/news/aktuelle-nyheter-2023/time-for-generative-ai-in-the-sandbox/.
34. European Commission. *Ethics Guidelines for Trustworthy AI | Shaping Europe's Digital Future*. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (2019).

35. Norwegian Data Protection Authority. Norwegian Position Paper on the European Commission's Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (COM(2021) 206).
36. Agencia Española de Supervisión de la Inteligencia Artificial | España Digital 2026. https://espanadigital.gob.es/lineas-de-actuacion/agencia-espanola-de-supervision-de-la-inteligencia-artificial.
37. Rivaya, G.-J. F. & Vidal, A. Spain: The artificial intelligence regulatory 'sandbox' has arrived. *Lexology* https://www.lexology.com/library/detail.aspx?g=99939c25-d7bb-4d06-b154-4a972eb71e9b (2023).
38. Bundesministerium für Wirtschaft und Energie & BMWi. *Neue Räume, Um Innovationen Zu Erproben. Konzept Für Ein Reallabore-Gesetz*. (2021).
39. Brown, E. & Piroska, D. Governing Fintech and Fintech as Governance: The Regulatory Sandbox, Riskwashing, and Disruptive Social Classification. *New Political Economy* **27**, 19–32 (2022).
40. Almada, M. & Petit, N. *The EU AI Act : A Medley of Product Safety and Fundamental Rights?* https://cadmus.eui.eu/handle/1814/75982 (2023).
41. Mühlhoff, R. & Ruschemeier, H. Predictive analytics and the collective dimensions of data protection. *Law, Innovation and Technology* **0**, 1–32 (2024).
42. UNSGSA FinTech Working Group and CCAF. *Early Lessons on Regulatory Innovations to Enable Inclusive FinTech: Innovation Offices, Regulatory Sandboxes, and RegTech*. https://www.jbs.cam.ac.uk/wp-content/uploads/2020/08/2019-summary-earlylessonsregulatoryinnovations.pdf (2019).
43. team. Joint Letter: 80 groups oppose CFPB's no-action letter and sandbox proposal. *Americans for Financial Reform* https://ourfinancialsecurity.org/2019/02/joint-letter-80-groups-oppose-cfpbs-no-action-letter-sandbox-proposal/ (2019).
44. Kelly, J. A "fintech sandbox" might sound like a harmless idea. It's not. https://www.ft.com/content/3d551ae2-9691-3dd8-901f-c22c22667e3b (2018).
45. Morozov, E. *To Save Everything, Click Here: Technology, Solutionism and the Urge to Fix Problems That Don't Exist*. (Allen Lane, London, UK, 2013).
46. Paquet, G. Governance as subversive bricolage in the 21st Century. *Governance: Canada/Ireland. Canadian Embassy, Dublin Craig Dobbin Chair of Canadian Studies (University College Dublin) and the Association for Canadian Studies in Ireland* (2003).
47. Sætra, H. S. *Technology and Sustainable Development: The Promise and Pitfalls of Techno-Solutionism*. (Taylor & Francis, 2023).