Cambridge Forum on AI: Law and Governance (2025), **1**, e36, 1–17
doi:10.1017/cfl.2025.10013

Cambridge Forum

RESEARCH ARTICLE

Exploring the boundaries of AI regulatory sandboxes under the AI Act: Flexibility and real-world testing

Nathan Genicot[1,2] and Thiago Guimaraes Moraes[2,3]

[1]Perelman Centre for Legal Philosophy, Université Libre de Bruxelles, Brussels, Belgium; [2]Law, Science, Technology & Society (LSTS) Research Group, Vrije Universiteit Brussel, Ixelles, Belgium and [3]Law School, Universidade de Brasilia (UnB), Brasília, Brazil
Corresponding author: Nathan Genicot; Email: nathan.genicot@ulb.be

(Received 13 March 2025; revised 24 April 2025; accepted 19 May 2025)

Abstract
The Artificial Intelligence (AI) Act introduces AI regulatory sandboxes and testing in real-world conditions as experimentation tools to support innovation in AI development. While regulatory sandboxes have been widely adopted across various sectors, their precise function and scope under the AI Act remain complex and multifaceted. This paper examines the legal framework governing AI regulatory sandboxes, analysing the degree of regulatory flexibility they provide and how they interact with other experimentation tools established by the AI Act. In particular, the study assesses whether these instruments enable AI systems to be tested under reduced regulatory constraints or whether they primarily serve as forums for regulatory dialogue. It then explores the differences between AI regulatory sandboxes and the testing in real-world conditions mechanism, highlighting how sandboxes focus on legal guidance, while testing in real-world conditions aims to remove legal barriers and facilitate market entry. The analysis also identifies key interpretative challenges, including the interplay between Article 57 (AI regulatory sandboxes) and Article 60 (testing in real-world conditions), the compliance requirements for each mechanism, and the role of regulatory authorities. The paper concludes that further clarification from the European Commission is necessary to address inconsistencies in the AI Act's provisions, ensuring effective implementation of AI regulatory sandboxes and testing in real-world conditions.

Keywords: AI Act; AI regulatory sandboxes; testing in real-world conditions; regulatory flexibility; experimental legislation

1. Introduction

Artificial intelligence (AI) regulatory sandboxes are a new instrument introduced by the Artificial Intelligence Act (AI Act).[1] Recent years have seen an impressive boom in this policy tool which aims to strike a balance between regulation and innovation, particularly digital technologies, by enabling economic players to test new products under the supervision of a competent regulator. Initially deployed in the field of finance, regulatory sandboxes have flourished around the world in a variety of sectors, including energy, healthcare, transport, data protection, and now AI (Ranchordás, 2021a).

[1]Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) [2024] OJ L.

The European Union (EU) itself has been actively promoting the use of this instrument. The European Commission describes regulatory sandboxes as enabling 'innovative solutions not already foreseen in regulations or guidelines to be live-tested with supervisors and regulators'[2] and included them as an *emergent policy instrument* in its 2021 Better Regulation Toolbox.[3] Numerous EU regulations adopted in 2024 – such as the Interoperable Europe Act,[4] the Net-Zero Industry Act,[5] and the Cyber Resilience Act[6] – provide for the possibility of creating regulatory sandboxes. However, perhaps the most ambitious form of regulatory sandboxes at the EU level is the AI regulatory sandboxes introduced by the AI Act, which not only encourages Member States to establish such sandboxes but obligates them to create at least one at the national level by 2 August 2026.

Despite the popularity of the concept of regulatory sandboxes, it appears to cover a great diversity of practices. This is confirmed by the existence of related notions – such as testbeds, pilot projects, testing environments, or real-world laboratories – which are sometimes used with a meaning very close to that of regulatory sandboxes (Arntzen, Wilcox, Lee, Hadfield & Rae, 2019, p. 14). While some regulatory sandboxes offer technical infrastructure or allow economic actors to effectively test their products in dedicated environments, others do not provide any form of infrastructure, but do offer regulatory leeway by, for example, allowing participants to operate without a licence. Other regulatory sandboxes merely provide legal guidance from the regulator without products being effectively tested (Genicot, 2024).

The question therefore arises as to where AI regulatory sandboxes introduced by the AI Act fit into this landscape. The provisions of the AI Act setting out the rules governing AI regulatory sandboxes are indeed very setting out and complex and require close scrutiny. In fact, the AI Act speaks not of one but two distinct instruments – *AI regulatory sandboxes* and *Testing in real-world conditions* – which are interrelated in a complex way. This article seeks to explore the contours of these two mechanisms by addressing the following questions: Do AI regulatory sandboxes under the AI Act enable AI systems to be tested or are they merely a forum for regulatory dialogue? Does participation in the sandbox imply the granting of regulatory leeway and a regulation-free space? What form of experimentation is allowed by the real-world testing mechanism, and how does it relate to AI regulatory sandboxes?

The paper is divided into five sections. The next section discusses the concept of regulatory sandboxes by highlighting how it navigates between three poles: testing, legal guidance, and regulatory flexibility. The third section examines the rules governing AI regulatory sandboxes under the AI Act and the form of regulatory flexibility that these rules allow for. The fourth section discusses the testing in real-world conditions mechanism that has been introduced alongside regulatory sandboxes in the AI Act, and reveals some lack of clarity and inconsistencies. The last section concludes.

[2]European Commission, 'An SME Strategy for a Sustainable and Digital Europe' (Communication) COM(2020) 103 final, p. 9.

[3]Accessible here: https://commission.europa.eu/law/law-making-process/planning-and-proposing-law/better-regulation/better-regulation-guidelines-and-toolbox_en.

[4]Regulation (EU) 2024/903 of the European Parliament and of the Council of 13 March 2024 laying down measures for a high level of public sector interoperability across the Union (Interoperable Europe Act) [2024] OJ L.

[5]Regulation (EU) 2024/1735 of the European Parliament and of the Council of 13 June 2024 on establishing a framework of measures for strengthening Europe's net-zero technology manufacturing ecosystem and amending Regulation (EU) 2018/1724 [2024] OJ L.

[6]Regulation (EU) 2024/2847 of the European Parliament and of the Council of 23 October 2024 on horizontal cybersecurity requirements for products with digital elements and amending Regulations (EU) No 168/2013 and (EU) No 2019/1020 and Directive (EU) 2020/1828 (Cyber Resilience Act) [2024] OJ L.

https://doi.org/10.1017/cfl.2025.10013 Published online by Cambridge University Press

2. Regulatory sandboxes: between regulatory flexibility, legal guidance, and testing environment

Regulatory sandboxes are generally defined as a controlled environment in which companies can test innovative products under the guidance of a competent regulator, and with a relaxation of regulatory requirements (notably through the granting of individual exemptions) (Allen, 2019; Zetzsche, Buckley, Barberis & Arner, 2017). They may therefore present three key features: (i) the opportunity to experiment with new technologies while benefiting from regulatory flexibility, such as temporary waivers; (ii) the possibility to do so in a testing environment that is close to the real world; and (iii) the benefit of guidance from a regulator, who tells sandbox participants how to comply with the law. However, not all regulatory sandboxes share these three characteristics to the same degree.

Regulatory flexibility is typically presented as a definitional feature of regulatory sandboxes. Ranchordás considers, for example, that regulatory sandboxes 'are types of legal experiments that either waive or modify national rules on a temporary basis in order to promote innovation' (Ranchordás, 2021b, p. 92). Similarly, a policy note of the Organisation for Economic Cooperation and Development (OECD) defines regulatory sandboxes as 'a limited form of regulatory waiver or flexibility for firms, enabling them to test new business models with reduced regulatory requirements' (Attrey, Lesher & Lomax, 2020, p. 7). Typically, in the financial sector, where they were first introduced, regulatory sandboxes allow certain regulatory constraints to be relaxed to facilitate the testing of fintech products with real customers (Allen, 2019, p. 592).

For instance, in the Philippines, a regulatory sandbox approach was adopted as early as 2004, when its central bank allowed two major telecommunications firms to pilot mobile money services in a regulatory environment that temporarily relaxed conventional rules (World Bank Group, 2020). At the time, there were no established regulations governing mobile money, and the financial authority allowed the firms to test innovative financial service models through non-bank entities. The central bank maintained close oversight throughout the process, ensuring consumer protection while fostering innovation. This controlled experimentation ultimately resulted in the publication of formal 'Guidelines on the Use of Electronic Money' five years later (Schellhase & Garcia, 2009).

However, not all regulatory sandboxes allow participants to benefit from exemptions from the normally applicable rules. Some consist mainly of a space for dialogue between the competent regulator and the regulated parties. This is the case of the regulatory sandboxes created in the field of data protection. Indeed, several countries – such as France, Norway, Denmark, and others – have created data protection sandboxes that aim to help participants develop innovative products (often using AI) that involve the processing of personal data (Genicot, 2024; Moraes, 2024b). These regulatory sandboxes do not allow participants to derogate from the law for a simple reason: it is not allowed under EU data protection law. The French *Commission nationale de l'informatique et des libertés* (CNIL) even states that it has not launched a 'regulatory sandbox' but only a 'sandbox' as participating in it 'does not allow the removal of legal constraints' (*Commission nationale de l'informatique et des libertés*, 2023).

Similarly, the recently adopted EU Interoperable Europe Act and EU Cyber Resilience Act, which both encourage Member States to create regulatory sandboxes, do not provide for the possibility to lift regulatory requirements.[7] Although it is often presented as an inherent characteristic, the possibility of derogating from the law is thus not systematically present in all regulatory sandboxes. In this sense, the European Commission stresses in a staff working document that 'it is important to note that the presence of a derogation is not a necessary element of regulatory sandboxes but that the involvement of a competent authority is necessary'.[8] The main advantage of taking part in these types of regulatory

[7]Both regulations state that regulatory sandboxes shall not affect the supervisory and corrective powers of the competent authorities (see Article 12(4) of the Interoperable Europe Act and Article 33(2) of the Cyber Resilience Act).

[8]European Commission, 'Regulatory learning in the EU Guidance on regulatory sandboxes, testbeds, and living labs in the EU, with a focus section on energy', SWD(2023) 277/2 final, 28 August 2023, p. 11.

https://doi.org/10.1017/cfl.2025.10013 Published online by Cambridge University Press

sandboxes lies in the close relationship that can be established with the regulator, who can guide the participant through the sometimes difficult interpretation of the applicable law.

The nature of the 'controlled environment' (or 'safe space') in which an innovation can be tested also varies from one regulatory sandbox to another. In many cases, the notion of a controlled environment solely refers to the fact that participants are allowed to test a product with (a limited number of) real customers, under the supervision of the regulator, in compliance with appropriate safeguards, and for a limited period. However, regulatory sandboxes sometimes provide real technical infrastructure, which may include physical facilities or intangible resources like datasets.

In such cases, regulatory sandboxes are akin to other notions such as testbeds or real-world laboratories. The Swedish Innovation Agency, Vinnova, proposed three levels to categorise test and demonstration environments (Arntzen et al., 2019, p. 17): (1) laboratories which correspond to 'strictly controlled test sites where innovators can test specific technical properties in isolated, artificial and heavily controlled circumstances'; (2) simulated environments which offer 'a simulated or constructed version of reality, still closed off and able to control by the testers'; and finally, (3) real-world testbeds which refer to 'controlled or bounded environments for testing innovation in real-world, or close to real-world, conditions in the manner (or close to the manner) in which they will be used or operated'.

While (contrary to regulatory sandboxes) real-world testbeds are not necessarily operated by regulators,[9] they are often involved. Moreover, they sometimes necessitate modifications to local laws (Engels, Wentland & Pfotenhauer, 2019; Laurent, Doganova, Gasull & Muniesa, 2021), such as modifying traffic rules to enable autonomous driving, which shows that the distinction between regulatory sandboxes and real-world testbeds is far from clear-cut.

Against this backdrop, the question arises as to what scope for regulatory flexibility is provided by the AI Act and to what extent AI systems can be tested in real-world conditions. To this end, the next section examines the rules governing AI regulatory sandboxes, while the subsequent section explores the mechanisms for real-world testing.

3. Regulatory flexibility in AI regulatory sandboxes under the AI Act

Before examining the contours of AI regulatory sandboxes, it is worth recalling the main rationale behind the AI Act. As it is now well known, this regulation is risk-based, meaning that the level of requirements and obligations imposed on the AI system depends on the risk the system poses to health, safety, and fundamental rights (European Commission, 2024). AI systems that present a risk deemed too substantial are banned, those that present a high risk are subject to a series of legal requirements, and those with a low risk are subject to minimal or no requirements. Most of the requirements contained in the AI Act relate to high-risk AI systems and fall on the providers of such systems.[10] These include establishing a risk management system, drafting technical documentation that demonstrates compliance with the AI Act, maintaining a data governance framework (which notably aims to control the quality and representativeness of the data used to feed the AI model), ensuring that a human oversees the AI system and its outputs, etc.[11]

The AI Act aligns with the New Legislative Framework, which guides EU legislation on product safety.[12] In this sense, the AI Act requires providers of high-risk AI systems a conformity assessment before placing their systems on the market or putting them into service. In some limited cases, the conformity assessment will have to be carried out by a third-party body, referred to in the AI Act

[9] European Commission, 'Regulatory learning in the EU Guidance on regulatory sandboxes, testbeds, and living labs in the EU, with a focus section on energy', SWD(2023) 277/2 final, 28 August 2023, p. 13.

[10] There are also specific rules contained in Chapter V of the AI Act which concern general-purpose AI models.

[11] See Chapter III of the AI Act.

[12] Recital 9 of the AI Act.

https://doi.org/10.1017/cfl.2025.10013 Published online by Cambridge University Press

as a notified body. In most cases, however, this conformity assessment will take the form of a self-assessment: providers shall evaluate themselves whether their systems comply with the requirements of the AI Act.[13] Once this self-assessment is successfully completed, the provider must affix a CE mark to the system and issue an EU declaration of conformity.[14] This logic of self-assessment differs from that observed in highly regulated sectors such as pharmaceuticals or financial services, where economic operators must receive a licence to operate and/or obtain *ex ante* authorisation before placing a product on the market. This is particularly relevant in relation to regulatory sandboxes, as in both of these highly regulated sectors a regulatory sandbox makes it possible to avoid such prior approval.[15] In the case of the AI Act, the situation is different and more akin to data protection regulation, given the limited role of regulators at the *ex ante* stage. In most cases, a company that wants to bring an AI system to market does not need any authorisation but only needs to ensure that its AI system complies with the regulation.

Although the AI Act introduces a whole range of new rules, it is not intended to hinder the development of AI, but on the contrary to encourage innovation in this field. One of its primary aims is to improve the functioning of the internal market by imposing uniform rules throughout the EU and preventing Member States from adopting stricter rules.[16] In this respect, regulatory sandboxes are supposed to play a key role in fostering a smooth implementation of the regulation and thus promoting AI innovation. According to the AI Act, the introduction of regulatory sandboxes should make it possible to stimulate the development of AI systems that comply with the AI Act (and any other applicable legislation), thereby accelerating the process of bringing AI systems to market. They should also promote legal certainty, enable the sharing of best practices between authorities, and stimulate regulatory learning (in particular to anticipate possible future adaptations of the legal framework).[17]

Article 3(55) of the AI Act defines AI regulatory sandboxes as

> […] a concrete and 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, where appropriate in real world conditions, an innovative AI system, pursuant to a sandbox plan for a limited time under regulatory supervision.

All the rules governing the sandbox process, including eligibility and selection criteria, are detailed in Articles 57 and 58 of the AI Act and will be further specified in an implementing act to be adopted by the European Commission. As already mentioned, Member States must ensure that their national competent authority set up at least one national regulatory sandbox, to be operational by 2 August 2026.[18] Additional AI regulatory sandboxes may also be established at regional or local levels or jointly with the competent authorities of other Member States.[19]

The definition of AI regulatory sandboxes contains various elements which echo our previous discussion: the notion of controlled environment (in this case 'framework'), the possibility to temporarily

[13] The subtleties of conformity assessment are contained in Article 43 of the AI Act, a key principle being that conformity assessment for all the high-risk AI systems listed in Annex III can be carried out through the internal control procedure (with the exception of biometric systems described in point 1, which are subject to slightly more demanding rules).

[14] Articles 16(h) and 48 of the AI Act. CE marking is a cornerstone of the EU's harmonisation legislation. Affixing the CE mark to a product indicates that it complies with all applicable legal requirements under EU law, allowing the product to circulate freely within the EU. In accordance with the New Legislative Framework approach, providers will have the option of following harmonised standards as part of their conformity assessment, which will lead to a presumption of conformity with the AI Act requirements. Harmonised standards are technical standards that are adopted by European standardisation organisations on the basis of a request made by the Commission for the application of Union harmonisation legislation. They are currently being negotiated. On the role of harmonised standards in the AI Act, see Gornet and Maxwell (2024).

[15] On the distinction between regulatory sandboxes in licensing and in non-licensing regimes, see Moraes (2024a).

[16] Recital 1 of the AI Act.

[17] See Recital 139 and Article 57(9) of the AI Act.

[18] Article 57(1) of the AI Act.

[19] Articles 57(1) and (2) of the AI Act.

https://doi.org/10.1017/cfl.2025.10013 Published online by Cambridge University Press

develop and test an AI system (in some cases in real world), and the supervision by a regulatory authority. The 'sandbox plan' refers to a 'document agreed between the participating provider and the competent authority describing the objectives, conditions, timeframe, methodology, and requirements for the activities carried out within the sandbox'.[20] This plan acts as a contract between the regulator and the regulated party with regard to experimentation. At the same time, it serves as a roadmap for what will be carried out during testing. Although it is not explicitly specified in the AI Act, it can be assumed high-risk AI systems and of general-purpose AI models, in particular, are targeted, given that one of the main aims of AI regulatory sandboxes is to improve regulatory compliance and that these categories of AI systems are primarily subject to the obligations outlined in the AI Act. This view was embraced in Spain, where an AI regulatory sandbox pilot was launched in 2022. The royal decree adopted for this purpose states that regulatory sandboxes were open to high-risk AI systems.[21] The European Commission will further define the selection criteria for participating in the sandbox and the whole process of testing in its implementing acts.[22]

Regulatory sandboxes must be operated by national competent authorities. According to Article 3(48) of the AI Act, national competent authorities refer either to notifying authorities or market surveillance authorities. Notifying authorities are responsible for designating and monitoring the bodies that carry out third-party conformity assessments (notified bodies).[23] Market surveillance authorities are responsible for the post-market surveillance of AI systems covered by the AI Act,[24] in accordance with the regime set out in Regulation 2019/1020 on market surveillance and compliance of products.[25] Market surveillance authorities seem more suited to managing regulatory sandboxes, as this task involves providing legal guidance to (potential) providers, which is a role closer to that of market surveillance authorities.[26] Indeed, the AI Act mandates the national competent authority to provide 'guidance, supervision and support within the sandbox with a view to identifying risks, in particular to health, safety and fundamental rights, testing, mitigation measures, and their effectiveness in relation to the obligations and requirements of the AI Act and, where relevant, other Union and Member States legislation supervised within the sandbox'.[27] The national competent authority must offer guidance to providers on regulatory expectations, fulfilling the requirements and obligations of the AI Act, and supporting compliance with conformity assessment obligations.[28] At the conclusion of the testing phase, and upon request of the provider of the AI system, the national competent authority issues a written proof of the activities successfully completed in the sandbox, and prepares an exit report detailing the activities carried out, along with the related results and learning outcomes. Providers may use these documents – the exit report and the written proof – to demonstrate compliance with the AI Act during the conformity assessment process or other market surveillance activities. The exit reports and written proofs issued by the national competent authority must be positively taken into account by market surveillance authorities and notified bodies, with the aim of

[20] Article 3(54) of the AI Act.

[21] As well as to general-purpose AI models. Real Decreto 817/2023, de 8 de noviembre, que establece un entorno controlado de pruebas para el ensayo del cumplimiento de la propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial.

[22] Article 58(1) of the AI Act.

[23] Articles 3(19) and 28 of the AI Act.

[24] Articles 3(26) and 74 of the AI Act.

[25] Regulation (EU) 2019/1020 of the European Parliament and of the Council of 20 June 2019 on Market Surveillance and Compliance of Products and Amending Directive 2004/42/EC and Regulations (EC) No 765/2008 and (EU) No 305/2011 OJ L 169/1.

[26] Regulation (EU) 2019/1020 of the European Parliament and of the Council of 20 June 2019 on Market Surveillance and Compliance of Products and Amending Directive 2004/42/EC and Regulations (EC) No 765/2008 and (EU) No 305/2011 OJ L 169/1.

[27] Article 57(6) of the AI Act.

[28] Articles 57(6–7) and 58(2)(e) of the AI Act.

https://doi.org/10.1017/cfl.2025.10013 Published online by Cambridge University Press

reasonably accelerating conformity assessment procedures.[29] In other words, while a successful participation in a sandbox will be a positive element in assessing the conformity of the AI system with the AI Act, it does not automatically lead to full compliance.

Under Article 57(11), such experimentations must not affect the supervisory and correctional powers of the competent authorities supervising the sandbox. Significant risks to health and safety or fundamental rights identified during the development and testing of AI systems in the controlled environment must be mitigated. If this is not possible, the testing may be suspended temporarily or permanently. To illustrate, consider a developer testing an AI-powered recruitment tool within a regulatory sandbox. During the testing, the national competent authority observes that the system systematically disadvantages candidates based on sensitive attributes such as gender and ethnicity, in violation of EU anti-discrimination law.[30] Despite guidance from the authority, the developer fails to implement effective mitigation measures to eliminate discriminatory outcomes. Given the significant and ongoing risk to the fundamental right to non-discrimination,[31] the national competent authority may decide to temporarily suspend the testing until appropriate safeguards are in place – or, if such risks cannot be mitigated, to permanently terminate the sandbox experimentation to prevent harm.

In the initial version of the AI Act proposed by the European Commission, the regime applicable to AI regulatory sandboxes did not provide any form of regulatory flexibility.[32] This has been modified during the negotiation process. Indeed, although the possibility of exempting participants from certain normally applicable rules is not explicitly provided for, participants are granted a form of regulatory flexibility through limited sanctions: as long as participants respect the sandbox plan and the terms and conditions for their participation and follow in good faith the guidance given by the national competent authority, no administrative fines shall be imposed by the competent authority.[33] While this measure is intended to allow providers to develop their products without fear of penalties, the question arises as to its effect in practice. Indeed, the requirements and obligations set out in the AI Act (along with the fines in case of non-compliance) only apply *after the AI systems have been placed on the market or put into service*.[34] However, participation in a sandbox takes place precisely before AI systems reach this stage.[35]

Another important provision regarding regulatory flexibility, Article 57(12), states that if other authorities responsible for European or national legislation are involved in the supervision of an AI system in the sandbox and provide advice on compliance, no administrative fine shall be imposed

[29] Article 57(7) para. 2 of the AI Act.

[30] Article 14(1)(a) of the Directive (EU) 2006/54/EC of the European Parliament and of the Council of 5 July 2006 on the implementation of the principle of equal opportunities and equal treatment of men and women in matters of employment and occupation states that 'there shall be no direct or indirect discrimination on grounds of sex in the public or private sectors, including public bodies, in relation to […] conditions for access to employment, to self-employment or to occupation, including selection criteria and recruitment conditions, whatever the branch of activity and at all levels of the professional hierarchy, including promotion'. A similar statement can be found in Article 3(1)(a) of Council Directive 2000/43/EC, the 'Racial Equality Directive'.

[31] Article 21 of the EU Charter.

[32] Commission, '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 final, Article 53 ('Proposed AI Act').

[33] Article 57(12) of the AI Act. See also Buocz et al. (2023, pp. 368–369).

[34] Article 2(8) of the AI Act states that 'This Regulation does not apply to any research, testing or development activity regarding AI systems or AI models *prior to their being placed on the market or put into service*. Such activities shall be conducted in accordance with applicable Union law. Testing in real world conditions shall not be covered by that exclusion.' The issue of testing in real-world conditions is the subject of specific provisions, which are discussed below. This view is shared by Burden and Stenberg (2023, p. 13).

[35] 'AI regulatory sandboxes established under paragraph 1 shall 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 their being placed on the market or put into service* pursuant to a specific sandbox plan agreed between the providers or prospective providers and the competent authority' (Article 57(5) of the AI Act).

https://doi.org/10.1017/cfl.2025.10013 Published online by Cambridge University Press

for that legislation either. The rationale of this provision is that if a participant in an AI regulatory sandbox complies with the sandbox plan but infringes another EU or national law, this participant should not be penalised.

The legality of this provision may be questionable in some cases, as it is doubtful that the AI Act has the authority to limit the supervisory powers of national authorities exercised under other national or EU laws. In the case of the General Data Protection Regulation (GDPR),[36] for example, the AI Act explicitly stipulates that 'this Regulation does not seek to affect the application of existing Union law governing the processing of personal data, including the tasks and powers of the independent supervisory authorities competent to monitor compliance with those instruments'.[37] Therefore, this AI Act provision may conflict with other European or national rules that do not allow regulators to refrain from imposing a fine despite finding a violation. One of the particularities of the AI Act is that its scope is extremely broad, since AI systems can be integrated into toys, medical devices, or used by banks, insurance companies, public services, etc. All these areas are governed by various specific rules. To remove any ambiguity, the European Commission could specify in its implementing acts that such an exemption from fines is only possible insofar as it is authorised by other applicable national and European laws. Member States could also consider the different regulators likely to be involved in a regulatory sandbox and adopt a national law that would specify the conditions under which they should or should not impose a fine.

That being said, one might wonder what legal requirements will be applicable to the sandbox testing of AI systems when this testing takes place at the 'laboratory' or 'simulated environment' levels (for real-world testing, see the next section). This depends, for each area of law, on the criterion that determines when a law begins to apply. In many areas, such as consumer law or product safety, most rules will only apply once a product is placed on the market or a service is provided to someone.[38] For this reason, we can assume that sandbox testing of AI systems (when not carried out in real-world conditions) will not yet be subject to most of the rules that will apply once they are commercialised.

This is not the case, however, for the processing of personal data, as European data protection legislation applies as soon as personal data are processed, even when it is only for the experimental development of a product. A company that develops an AI system and feeds it with personal data is required to comply with the GDPR. Therefore, according to Article 57(12), if personal data are processed in the sandbox in breach of the GDPR but the participant complies with the sandbox plan and follows in good faith the guidance provided, the data protection authority (DPA) should not fine the participant.

In addition, Article 59 of the AI Act provides for an important exception to the GDPR.[39] This provision introduces a specific regime for the processing of personal data within the regulatory sandboxes: in some circumstances, personal data lawfully collected for other purposes may be processed solely for the purposes of developing, training, and testing certain AI systems in the sandbox. A number of conditions must be met to do this. First, AI systems shall be developed for safeguarding substantial public interest in one or more of the following areas: public safety and public health, protection of the environment, energy sustainability, transport systems and mobility, critical infrastructure and networks, public administration, and public services. Second, the data processed are necessary for

[36] Regulation (EU) 2016/679 of the European Parliament and Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L 119/1.

[37] Recital 10 of the AI Act.

[38] For example, consumer protection and product safety rules applicable to an AI-powered toy typically only come into effect when the toy is offered for sale or made available to consumers.

[39] There is a similar provision in the Interoperable Europe Act regarding interoperability regulatory sandboxes.

https://doi.org/10.1017/cfl.2025.10013 Published online by Cambridge University Press

complying with the requirements for high-risk AI systems where those requirements cannot effectively be fulfilled by processing anonymised, synthetic, or other non-personal data. Third, different measures aimed at safeguarding the rights of data subjects must be taken.[40]

Another relevant aspect is the principle of good faith. Originating in contract law, it can be understood in two ways: objectively, as a tool to introduce fairness and mitigate imbalances in legal relationships; and subjectively, as a person's genuine belief that they are acting lawfully or in situations where third parties deserve protection based on that belief (Gjoni & Peto, 2017). The AI Act's reference to good faith suggests a quasi-public contractual relationship between authorities and providers within the AI regulatory sandbox. This is reinforced by the authority's power to suspend or terminate testing if appropriate mitigation measures are not adopted (van der Valk, 2023). However, such actions do not automatically result in sanctions. Here, good faith offers a form of 'comfort zone' for providers: as long as they can demonstrate a sincere intent to comply, they should not fear punishment. This fosters a collaborative environment, where experimentation is not treated as a trap to gather evidence for enforcement, but as a space for honest testing and learning.

With regard to damage that may be caused to a third party during participation in a sandbox, the provider remains liable under applicable EU and Member States liability legislation.[41] Although it is common practice not to exempt participants from civil liability in regulatory sandboxes, this provision has been criticised for potentially deterring developers of AI systems from joining an AI regulatory sandbox (Buocz, Pfotenhauer & Eisenberger, 2023; Gromova & Stamhuis, 2023; Truby, Brown, Ibrahim & Parellada, 2022). An important question is whether compliance with the testing plan will prevent a participant from being deemed to be in breach of duty (and therefore liable under a fault-based liability regime). Compensation schemes, such as insurance, could be implemented to limit the risk to participants while protecting affected individuals from potential harm (Buocz et al., 2023; Truby et al., 2022).

4. Differentiating real-world testing: inside vs. outside sandboxes

In the initial version of the AI Act proposed by the European Commission, it was not possible to test AI systems in real-world conditions.[42] However, this has been modified during the legislative process. The AI Act now provides that testing may take place 'where appropriate in real-world conditions'.[43] However, understanding the rules applying to testing in real-world conditions (TRWC) and how this mechanism interacts with regulatory sandboxes is not an easy task. A careful reading of the AI Act shows that the provisions relating to TRWC contain certain inconsistencies that the European Commission will have to address in its implementing acts.

Article 3(57) of the AI Act defines TRWC as

> […] the temporary testing of an AI system for its intended purpose outside a laboratory or other simulated environment, with a view to gathering reliable and robust data and to assessing and verifying the conformity of the AI system with the requirements of this Regulation and it does not qualify as placing the AI system on the market or putting it into service within the meaning of this Regulation, provided that all the conditions laid down in Article 57 or 60 are fulfilled.

[40] More details can be found in Article 59 of the AI Act. This provision is not supposed to contradict the GDPR which contains strict rules regarding the further processing of personal data. However, as discussed by A. Papageorgiou, the legality of this provision is doubtful as it does not fully comply with Article 23(2) of the GDPR (Papageorgiou, 2024).

[41] Article 57(12) of the AI Act.

[42] In the original draft, the European Commission defined regulatory sandboxes as a 'controlled environment facilitating the development, testing and validation of innovative AI systems for a limited period prior to their placing on the market or putting into service in accordance with a specific plan'. See European Commission, Proposed AI Act.

[43] See also Article 5(57) which states that 'regulatory sandboxes may include testing under supervised real-world conditions'.

https://doi.org/10.1017/cfl.2025.10013 Published online by Cambridge University Press

Table 1. Provisions governing testing in real-world conditions

	High-risk AI systems	Non-high-risk AI systems
Inside the regulatory sandbox	Articles 57–61 and 76 of the AI Act	Article 57–59 of the AI Act
Outside the regulatory sandbox	Articles 60–61 and 76 of the AI Act	Unclear

An important feature of TRWC is that it is considered to occur before the AI system is placed on the market or put into service. As discussed earlier, this entails that most requirements arising from the AI Act are not applicable. The TRWC mechanism therefore offers a form of regulatory flexibility in that it allows an AI system to be tested in a real-life setting with no need for the provider to carry out a conformity assessment beforehand. However, the prohibition of certain AI practices under Article 5 of the AI Act continues to apply during TRWC.[44]

According to this definition, TRWC must meet either the conditions of Article 57 or those of Article 60. Whereas Article 57 concerns regulatory sandboxes and paragraph 5 mentions the possibility of TRWC within regulatory sandboxes, Article 60 concerns TRWC of high-risk AI systems outside regulatory sandboxes. It is therefore possible to test an AI system both within a sandbox process and outside a regulatory sandbox. At first glance, one might think that when a high-risk AI system is tested in real-world conditions *within* a sandbox, the rules set out in Articles 60 and 61 would not apply (since these provisions concern TRWC '*outside* regulatory sandboxes'). However, Article 76(2) of the AI Act states that testing of a high-risk AI system within a sandbox must comply with Article 60. Another question that arises is whether it is possible to test in real-world conditions an AI system that is not high-risk – for example, a chatbot that is subject to transparency obligations under Article 50 – outside a regulatory sandbox. The AI Act says nothing about this. Table 1 summarises the different forms of TRWC and the interpretation that seems most coherent.

As a result, a provider of a high-risk AI system that participates in a regulatory sandbox and wishes to test its AI system in real-world conditions will need to comply with both Articles 57–59, which relate to regulatory sandboxes, and Articles 60–61 and 76, which relate to the TRWC. In order to verify the potential overlaps and differences between these provisions, we provide Table 2 to guide the analysis.

As discussed in the previous section, while the AI Act mandates national competent authorities to operate the regulatory sandbox, market surveillance authorities seem more suited than notifying authorities to undertake this role. As for TRWC, the AI Act explicitly requires market surveillance authorities to monitor it. It is therefore likely that market surveillance authorities will be responsible for both mechanisms.

For TRWC to be conducted within a regulatory sandbox, the national competent authority will have to start by establishing the sandbox,[45] and it must engage all relevant regulators for a given testing.[46] As for TRWC outside sandboxes, the request has to be done by the (prospective) provider to the market surveillance authority,[47] and there is no requirement to involve other regulators. This is noteworthy, as real-world conditions are inherently more complex scenarios, which increases the relevance of involving additional regulators. Such testing could provide valuable insights for regulators, helping them to build the necessary knowledge to regulate and get answers to their questions regarding the innovation.[48] For example, the EU has already established more than 90 smart cities projects, with tailor-made solutions in areas such as energy, transport, and ICT (*Projects|Smart Cities Marketplace*, n.d.). Such real-world environments could attract the interest of many regulators at once,

[44] Article 60(1) of the AI Act.
[45] Article 57(1) of the AI Act.
[46] Articles 57(4) c/c (10) of the AI Act.
[47] Article 60(4)(a) of the AI Act.
[48] See Arntzen et al. (2019, p. 52).

https://doi.org/10.1017/cfl.2025.10013 Published online by Cambridge University Press

Table 2. Comparison of rules on AI regulatory sandboxes and TRWC outside them

Guiding question	AI regulatory sandboxes (Articles 57–59)	TRWC outside AI regulatory sandboxes (Articles 60–61)
Who has the initiative?	Established by the national competent authority – Article 57(1)	Proposed by the (prospective) provider – Article 60(4)(a)
Who must be involved?	Must engage other relevant authorities, including DPAs – Article 57(4) c/c (10)	Only requires involvement of the market surveillance authority – Article 60(4)(b)
What is the application assessment?	Sandbox plan must be selected, according to criteria determined by the national competent authority – Article 58(1)	Real-world testing plan must be approved by the market surveillance authority – Article 60(4)(b)
What is the role of the authority?	National competent authority must provide guidance and support to risk assessment, in particular to fundamental rights, health, and safety, as well as mitigation measures and their effectiveness – Article 57(6)	No obligation on guidance, but market surveillance authorities may carry out unannounced remote or on-site inspections and checks on the conduction of tests – Article 60(6)
What is the duration?	No specific duration, but experimentation should be conducted for a limited time before market deployment – Article 57(5)	No longer than necessary, for a maximum of 6 months, which can be extended for the same period subject to prior notification to the market surveillance authority – Article 60(4)(f)
What are the reporting activities?	National competent authority must draft an exit report after the testing – Article 57(7) – and also annual reports – Article 57(16)	Does not require an exit report, but serious incidents must be reported by the (proposed) provider – Article 60(7)
What are the transparency mechanisms to the wider public?	Exit report may be transparent if both national competent authority and participant agree – Article 57(8) National competent authority's annual reports or their abstracts must be made available to the public, online – Article 57(16)	Testing must be registered in the EU database for high-risk AI systems, save some exceptions – Article 60(4)(c)
Are there liability exemptions?	Participants remain liable for damage inflicted on third parties, but may be exempt from administrative fines under certain conditions – Article 57(12)	(Prospective) Providers remain liable for damage inflicted on third parties, no information regarding exemption from administrative fines[49] – Article 60(9)
What are the specific safeguards to protection of (data) subjects?	DPAs must be involved in any AI regulatory sandbox project that involves the processing of personal data – Article 57(10) Specific rules for further processing of personal data for developing certain AI systems in the public interest in the AI regulatory sandbox – Article 59	Specific rules for obtaining informed consent of subjects interested in participating in TRWC outside AI regulatory sandboxes – Article 61

and although the AI Act only requires one market surveillance authority per testing, it may be the case that several authorities may wish to be involved and there are no clear rules on how these kinds of 'consortiums' could be established outside sandboxes.

[49] However, as already discussed, if all conditions of Article 60(4) are met, testing will be considered a pre-market stage, thus avoiding administrative fines.

https://doi.org/10.1017/cfl.2025.10013 Published online by Cambridge University Press

Another essential element of experimental regulation frameworks is for the regulator to analyse the interested party's readiness to test: it must develop a plan with clear goals, suitable risk assessment, and envisioned mitigation measures.[50] The chosen documents in the AI Act are the sandbox plan (for AI regulatory sandboxes) and the real-world testing plan (for TRWC outside AI regulatory sandboxes). While some elements, such as objectives, scope, methodology, and timeframe, are common to both, the differences between them will only be clearer with the Commission's implementing acts.[51] A currently open question is whether, when TRWC is conducted within AI regulatory sandboxes, both plans would be required or if a merged version would suffice.[52]

The role of the authority within or outside sandboxes may also be substantially different: in the former, the AI Act explicitly states that the national competent authority must provide guidance and support to risk assessment. In fact, the goal to establish a strong collaboration between regulators and innovators is a key feature of regulatory sandboxes.[53] Nevertheless, this collaborative approach must be taken with care to avoid risks of regulatory capture, in which influence is exercised over regulators only for the benefit of industry and in detriment of the public interest (Ranchordás & Vinci, 2024).

One countermeasure to this risk is to establish transparent processes (Ranchordás & Vinci, 2024), and within AI regulatory sandboxes, the AI Act seems to go in the right direction, by requiring that annual reports or their abstracts are made available to the public.[54] These reports shall include best practices, incidents, lessons learnt, and recommendations on their set-up, as well as, where relevant, on the implementation and possible review of this regulation. That being said, it is unlikely that sharing only abstracts will be enough for proper public scrutiny. Furthermore, for each project, exit reports detailing the activities carried out in the sandbox are to be provided by the national competent authority.[55] Nevertheless, there is no obligation to make these reports public, since the participant must agree to do so. While this limitation aims to address confidentiality concerns, it may undermine the openness of the process. These transparency mechanisms also apply to TRWC inside AI regulatory sandboxes.

On the other hand, the role of the market surveillance authorities in TRWC outside AI regulatory sandboxes seems to be much more related to their 'market surveillance' powers, such as requiring (prospective) providers to provide information, carrying out unannounced remote or on-site inspections and performing checks on the conduct of the testing.[56] The language used in the AI Act seems to depart from the collaborative approach of the sandboxes, which may disincentive regulatees looking for regulatory dialogue to apply for testing in real world outside regulatory sandboxes. There is also no obligation of annual reports or exit reports, and the main transparency mechanism is the registration of the real-world testing in the EU database for high-risk AI systems, save some exceptions.[57] Looking at the list of Annex VIII, the information to be submitted does not include any of the elements required by the AI regulatory sandbox annual reports.

[50] See European Securities and Markets Authority et al. (2018, p. 23).

[51] The Commission is specifically required to do that according to Articles 58(1)(b) and 60(1).

[52] Besides, regarding TRWC within sandboxes, Article 58(4) of the AI Act requires that national competent authorities specifically agree the terms and conditions of such testing, in particular, the appropriate safeguards with the participants, with a view to protecting fundamental rights, health, and safety.

[53] In that sense, Ranchordás and Vinci define regulatory sandboxes as 'collaborative regulatory instruments where regulators interact closely with a selected group of market actors (usually startups) to create a safe testbed to understand how to best regulate new types of services or products' (Ranchordás & Vinci, 2024). Also, the Datasphere Initiative defines regulatory sandboxes as 'time-limited collaborative endeavours involving regulators, service-providers and other relevant stakeholders to test innovative technology and data practices against regulatory frameworks' (The Datasphere Initiative, 2022).

[54] Article 57(16) of the AI Act.

[55] Article 57(7) of the AI Act.

[56] Article 60(6) of the AI Act.

[57] According to Article 60(4)(c), the exemptions to register are for (prospective) providers of high-risk AI systems referred to in points 1, 6, and 7 of Annex III in the areas of law enforcement, migration, asylum, and border control management, and high-risk AI systems referred to in point 2 of Annex III (critical infrastructure).

https://doi.org/10.1017/cfl.2025.10013 Published online by Cambridge University Press

Therefore, testing conducted outside sandboxes does not foster regulatory learning and appears to be much more oriented to serve as a 'last step' that providers may opt for before deploying their systems on the market. The wording of Recital 141 also goes in this direction, by saying that these stakeholders may benefit from a specific regime for testing high-risk AI systems in real-world conditions, without participating in an AI regulatory sandbox, with the goal to accelerate the process of development and the placing on the market of these systems.

As for duration, the rules for AI regulatory sandboxes are quite flexible since the AI Act did not fix a specific term. A study conducted by the World Bank Group in 73 fintech sandboxes revealed that testing periods ranged from three months to two years, with more than two-thirds lasting at least one year (World Bank Group, 2020, p. 22). Therefore, this flexibility seems fortunate, giving more freedom to Member States or national competent authorities to determine the testing duration, according to the experimentation context. The same cannot be said for TWRC outside regulatory sandboxes, which have a fixed 6-month term, renewable once for the same period if authorised by the market surveillance authority. While it is understandable that this constraint is to avoid a scenario in which providers keep extending testing indefinitely, it also makes more challenging to match some experimentations in this timeframe. Testbeds usually take more than one year to conduct their experiments. One example are smart cities testbeds, such as SmartSantander, which lasted four years, and the Ruggedised project, which involved six European cities and experimented for five years under the same umbrella.[58] By imposing this time limitation, it becomes much harder for such consortiums to test high-risk AI systems, since they will need to apply for each system separately, raising complexities for the experimentation. At least, within sandboxes, the fixed term of Article 60(4)(f) of the AI Act may be derogated, as stated in Article 76(2).[59]

The discussion on liability and exemptions from administrative fines was addressed in Section 3. It should be noted that in both cases (within or outside AI regulatory sandboxes), there are no civil liability exemptions.

Finally, it is also important to note that different safeguards are provided for the protection of individuals in each setting. In the case of AI regulatory sandboxes (including TRWC within them), whenever the innovative AI system involves the processing of personal data, DPAs should be involved. Of course, if the DPA is also designated as the market surveillance authority under the AI Act, it will be engaged in AI regulatory sandboxes regardless of personal data being processed.[60] Besides, as discussed in the previous section, further processing of personal data lawfully collected for other purposes will only be allowed in certain circumstances.

As for TRWC outside sandboxes, there is no requirement to involve the DPA (unless it is designated by the AI Act or the Member State as the market surveillance authority), nor is it possible to process personal data for other purposes. Instead, rules for obtaining freely given informed consent from the subjects of testing are in place, as per Article 61. These rules also apply for TRWC inside AI regulatory sandboxes.[61] In fact, the rules for informed consent should be considered a good practice for any testing which involves affected subjects, either inside or outside sandboxes, and regardless of involving high-risk AI systems or not: their participation in the experimentation could prove essential as they can provide feedback for developing systems that will ultimately affect them when deployed in the market (Gonzalez Torres & Sawhney, 2023, p. 309).

[58] SmartSantander (European Commission, 2010); RUGGEDISED (European Commission, 2017).

[59] Another condition that Article 76(2) allows to be derogated is the protection of vulnerable groups of the elderly and persons with disabilities (Article 60(4)(g)). Neither the article nor the Recitals provide any explanation for why this safeguard is prone to derogation. The derogation is particularly concerning, given that this condition serves as a fundamental rights safeguard for these vulnerable groups.

[60] See Genicot (2024). Also, note that for high-risk systems used for law enforcement purposes, border management, and justice and democracy, the market surveillance authority must be the data protection, as stated by Article 74(8) of the AI Act.

[61] Article 76(2) c/c Article 60(4)(i) of the AI Act.

https://doi.org/10.1017/cfl.2025.10013 Published online by Cambridge University Press

Imagine the hypothetical scenario of a regulatory sandbox established between a hospital, in collaboration with AI developers, and a competent authority willing to conduct alpha testing.[62] This project takes place in a controlled, simulated environment, where AI systems designed to interact with patients and medical staff are tested internally before broader deployment.[63] AI tools are evaluated in key scenarios, such as a chatbot triaging symptoms and an AI-assisted diagnosis system analysing medical images. At this stage, only a small group of testers – primarily developers, researchers, hospital staff, and a small number of individuals who would represent patients' perspectives – engage with the system to identify potential transparency gaps, usability issues, and compliance challenges. While these AI systems do not involve testing in real-world conditions, requiring informed consent[64] from the participating subjects – the hospital staff and the patients' representatives – should still be considered a good practice. Nevertheless, the AI Act neither mandates this requirement in its main Articles nor recommends it in the Recitals.

As already discussed, the GDPR still applies in this context, including the provisions regarding further processing of personal data.[65] Hence, while it is not explicitly stated in the AI Act, the enforcement of the GDPR provisions may ultimately require the engagement of the DPA when testing in real world outside sandboxes, even when this authority is not the designated market surveillance authority for certain AI systems.

Therefore, it seems to be a good practice to involve DPAs in real-world testing outside AI regulatory sandboxes, whenever personal data are processed. Surely, the engagement of these authorities in so many testing instances both within and outside sandboxes comes with the risk of overburdening them and Member States must develop strategies or raise their capacity so they can deal with their growing roles under the AI Act. In this sense, to counter that challenge, not only in relation to sandboxes but also to the new tasks that DPAs have to implement due to this Regulation as a whole, the European Data Protection Board stated that DPAs should have their capacity increased, including adequate additional human and financial resources (European Data Protection Board, 2024).

To close this section, we illustrate the challenges faced by providers and authorities on weighing the benefits of conducting testing in real-world conditions within or outside AI regulatory sandboxes. Imagine a start-up developing a high-risk AI system for medical devices. If the provider chooses to test the system *within* an AI regulatory sandbox, they benefit from direct guidance and supervision by the national competent authority and, most probably, the health authority, which will help them to identify and mitigate risks early and align the system with regulatory expectations. This collaborative environment fosters regulatory learning and can strengthen the provider's confidence in achieving conformity later on. However, it also requires a more structured and possibly longer process, including closer oversight. On the other hand, *testing in real-world conditions outside* the sandbox offers a potentially faster route to market, since the AI Act frames this type of testing as a kind of transitional phase prior to full compliance. While this route lacks the same level of regulatory support, it allows for testing with real users at a faster pace – given that the procedures under Article 60 seem to be more expeditious than those of Article 57. This could be ideal for providers who believe their systems are already close to meeting the AI Act's requirements. Ultimately, the choice depends on the provider's priorities: legal clarity and collaborative development within a sandbox, or greater speed and autonomy outside it. Of course, this also presents a challenge for authorities, who must design

[62] Alpha testing is an early-stage user acceptance test, part of the software development life cycle, conducted in a controlled environment by the internal development team before a system is released to external users. It involves limited user engagement, typically including developers, testers, and a small number of potentially affected subjects who are closely connected to the project. The goal is to identify bugs and usability issues, while refining the system before broader external user testing or deployment. Unlike beta testing, which involves real users or external testers, alpha testing remains a preliminary phase focused on internal evaluation. For more differences between alpha testing and beta testing, see GeeksforGeeks (2024).

[63] These obligations are established in Article 50 of the AI Act.

[64] As outlined in Article 61 of the AI Act.

[65] Article 6(4) of the GDPR.

https://doi.org/10.1017/cfl.2025.10013 Published online by Cambridge University Press

incentives compelling enough to encourage providers to conduct real-world testing within regulatory sandboxes – rather than defaulting to the less supervised, faster paths outside them.

5. Conclusion

In this paper, we examined the concept of regulatory sandboxes and their role in balancing testing, legal guidance, and regulatory flexibility under the AI Act. We have analysed the provisions governing AI regulatory sandboxes, particularly the legal flexibility afforded to participants, and explored how the testing in real-world conditions (TRWC) mechanism interacts with AI regulatory sandboxes.

Based on our analysis of the AI Act provisions, it seems that AI regulatory sandboxes primarily aim to provide legal guidance, whereas TRWC focuses on removing legal barriers to accelerate market entry. Indeed, while Article 57(12) of the AI Act provides for an exemption from fines for breaches of the AI Act for sandbox participants acting in good faith, it remains an open question whether this rule will have any effect when it comes to tests carried out solely in a laboratory environment. It therefore seems that the value of AI regulatory sandboxes lies more in the regulatory dialogue that participants will engage in with regulators than in benefiting from exemptions. The situation is different with TRWC since this mechanism allows testing to be carried out in real conditions and with real subjects while being considered as taking place at a pre-marketing stage, which implies that most of the provisions of the AI Act are not yet applicable to the AI system being tested.

Another key finding, however, is that Article 57(12) – which also shields sandbox participants from administrative fines that would result from the violation of other national and EU laws – may face legal challenges due to potential conflicts with other regulatory frameworks. Further clarification from the European Commission would be recommended.

Furthermore, the study has highlighted the complex relationship between AI regulatory sandboxes and TRWC. The AI Act sets out distinct provisions for TRWC inside and outside regulatory sandboxes, yet there are inconsistencies in how these provisions interact. The requirement for compliance with both Article 57 and Article 60 when TRWC takes place within a sandbox creates interpretative challenges. The process for initiating TRWC also differs significantly: inside sandboxes, multiple regulators must be involved from the outset, whereas outside sandboxes, the prospective provider submits a request directly to the market surveillance authority without a mandatory engagement of other regulators.

Another critical issue relates to transparency and accountability. The AI Act introduces reporting requirements for AI regulatory sandboxes, such as annual and exit reports, yet these obligations do not extend to TRWC outside sandboxes. The only transparency mechanism in this case is the registration of high-risk AI system testing in the EU database. Additionally, while DPAs are not formally required to oversee TRWC outside sandboxes (unless they are the designated market surveillance authority), their involvement could enhance data protection oversight, especially when personal data are processed. That being said, the growing number of AI testing instances may risk overburdening DPAs, necessitating capacity-building measures at the national level.

Finally, while not mandatory for AI regulatory sandboxes that do not conduct real-world testing, the informed consent provisions in Article 61 should be regarded as a good practice whenever experimentation involves participating subjects.

In light of these findings, future regulatory efforts should focus on resolving ambiguities in the rules for testing environments. Thankfully, the AI Act mandates the Commission to develop implementing acts on AI Regulatory Sandboxes and testing in real-world conditions.[66] Hopefully, it may address the issues highlighted in this study.

[66]See Articles 58(1) and (2) c/c 60(1) of the AI Act.

https://doi.org/10.1017/cfl.2025.10013 Published online by Cambridge University Press

Funding statement. The authors declare none.

Competing interests. The authors declare none.

References

Allen, H. J. (2019). Regulatory sandboxes. *The George Washington Law Review*, *87*(3), 579.

Arntzen, S., Wilcox, Z., Lee, N., Hadfield, C., & Rae, J. (2019). *Testing innovation in the real world—Real-world testbeds*. NESTA. https://www.nesta.org.uk/report/testing-innovation-real-world/

Attrey, A., Lesher, M., & Lomax, C. (2020). *The role of sandboxes in promoting flexibility and innovation in the digital age*. OECD Going Digital Toolkit Notes. OECD. doi: 10.1787/cdf5ed45-en

Buocz, T., Pfotenhauer, S., & Eisenberger, I. (2023). Regulatory sandboxes in the AI Act: Reconciling innovation and safety? *Law, Innovation and Technology*, *15*(2), 357.

Burden, H., & Stenberg, S. (2023). *Sustainable AI and disruptive policy – AI regulatory sandboxes*. RISE Research Institutes of Sweden AB. https://diva-portal.org/smash/record.jsf?pid=diva2%3A1835556&dswid=−7488

Commission nationale de l'informatique et des libertés. (2023, July 28). *"Sandbox": CNIL launches call for projects on artificial intelligence in public services*. https://www.cnil.fr/en/sandbox-cnil-launches-call-projects-artificial-intelligence-public-services

The Datasphere Initiative. (2022). *Sandboxes for data: Creating spaces for agile solutions across borders*. https://www.thedatasphere.org/datasphere-publish/sandboxes-for-data/

Engels, F., Wentland, A., & Pfotenhauer, S. M. (2019). Testing future societies? Developing a framework for test beds and living labs as instruments of innovation governance. *Research Policy*, *48*(9), 103826. doi:10.1016/j.respol.2019.103826

European Commission. (2010). *SmartSantander|SmartSantander Project|Fact Sheet|FP7*. https://cordis.europa.eu/project/id/257992

European Commission. (2017, March 1). *RUGGEDISED|Smart Cities Marketplace*. https://smart-cities-marketplace.ec.europa.eu/projects-and-sites/projects/ruggedised

European Commission. (2024, August 1). *AI Act enters into force*. https://commission.europa.eu/news/ai-act-enters-force-2024-08-01_en

European Data Protection Board. (2024). *Statement 3/2024 on data protection authorities' role in the Artificial Intelligence Act framework*. EDPB. https://www.edpb.europa.eu/our-work-tools/our-documents/statements/statement-32024-data-protection-authorities-role-artificial_en

European Securities and Markets Authority, European Bank Authority, & European Insurance and Occupational Pensions Authority. (2018). *FinTech: Regulatory sandboxes and innovation hubs*. https://www.esma.europa.eu/sites/default/files/library/jc_2018_74_joint_report_on_regulatory_sandboxes_and_innovation_hubs.pdf

GeeksforGeeks. (2024, May 16). *Difference between alpha and beta testing*. GeeksforGeeks. https://www.geeksforgeeks.org/difference-between-alpha-and-beta-testing/

Genicot, N. (2024). *From blueprint to reality: Implementing AI regulatory sandboxes under the AI Act*. FARI & LSTS Research Group (VUB). https://www.fari.brussels/research-and-innovation/publication/ai-regulatory-sandboxes

Gjoni, G., & Peto, Z. (2017). An Overview of Good Faith as a Principle of Contractual Interpretation with Special References to the Albanian Law. *European Scientific Journal, ESJ*, 13(25), Article 25. https://doi.org/10.19044/esj.2017.v13n25p288

Gonzalez Torres, A. P., & Sawhney, N. (2023). Role of regulatory sandboxes and MLOps for AI-enabled public sector services. *The Review of Socionetwork Strategies*, *17*(2), 297–318. doi:10.1007/s12626-023-00146-y

Gornet, M., & Maxwell, W. (2024). The European approach to regulating AI through technical standards. *Internet Policy Review*, *13*(3), 1.

Gromova, E., & Stamhuis, E. (2023). Real-life experimentation with artificial intelligence. In J. Temperman, and A. Quintavalla (Eds.), *Artificial intelligence and human rights*. Oxford University Press, 551.

Laurent, B., Doganova, L., Gasull, C., & Muniesa, F. (2021). The test bed island: Tech business experimentalism and exception in Singapore. *Science as Culture*, *30*(3), 367–390.

Moraes, T. (2024a). Regulatory sandboxes as tools for ethical and responsible innovation of artificial intelligence and their synergies with responsive regulation. In *The quest for AI sovereignty, transparency and accountability*, Fundação Getulio Vargas - FGV. FGV – Direito Rio, 303. https://vlex.com.br/vid/regulatory-sandboxes-as-tools-1034960669

Moraes, T. (2024b). Regulatory sandboxes for trustworthy artificial intelligence – Global and Latin American experiences. *International Review of Law, Computers & Technology*. doi:10.1080/13600869.2024.2351674

Papageorgiou, A. (2024). *Addressing the challenges arising from the implementation of regulatory sandboxes under the AI Act*. [Master Thesis, KU Leuven]. https://repository.teneo.libis.be/delivery/DeliveryManagerServlet?dps_pid=IE21041849&

Projects|Smart Cities Marketplace. (n.d.). Retrieved 14 January 2025, from https://smart-cities-marketplace.ec.europa.eu/projects-and-sites/projects

Ranchordás, S. (2021a). Experimental lawmaking in the EU: Regulatory sandboxes. *EU Law Live. Weekend Edition*, *76*, 1.
Ranchordás, S. (2021b). Experimental regulations for AI: Sandboxes for morals and mores. *Morals & Machines*, *1*(1), 86.
Ranchordás, S., & Vinci, V. (2024). Regulatory sandboxes and innovation-friendly regulation: Between collaboration and capture. *Italian Journal of Public Law*, *1*. doi:10.2139/ssrn.4696442
Schellhase, J., & Garcia, A. (2009). FinTech in the Philippines: Assessing the State of Play. Milken Institute. https://milkeninstitute.org/sites/default/files/reports-pdf/FinTech-in-the-Philippines-Update%20%281%29.pdf
Truby, J., Brown, R. D., Ibrahim, I. A., & Parellada, O. C. (2022). A sandbox approach to regulating high-risk artificial intelligence applications. *European Journal of Risk Regulation*, *13*(2), 270.
van der Valk, W. (2023). Public Sector Contracting. In J. Grandia, and L. Volker (Eds.), *Public Procurement: Theory, Practices and Tools* (pp. 121–135). Springer International Publishing. https://doi.org/10.1007/978-3-031-18490-1_7
World Bank Group. (2020). *Global experiences from regulatory sandboxes*. https://openknowledge.worldbank.org/handle/10986/34789
Zetzsche, D., Buckley, R., Barberis, J., & Arner, D. (2017). Regulating a revolution: From regulatory sandboxes to smart regulation. *Fordham Journal of Corporate & Financial Law*, *23*(1), 31.

Nathan Genicot is a researcher at the Perelman Centre for Legal Philosophy (ULB) and an affiliated researcher at LSTS. He completed a PhD in Law on the history of algorithmic profiling and subsequently conducted postdoctoral research on AI regulatory sandboxes.

Thiago G. Moraes is a joint-degree Ph.D. Candidate in Law at University of Brasilia (UnB) and Vrije Universiteit Brussels (VUB) and a PhD fellow of the Digital Governance cluster at the United Nations University's Institute on Comparative Regional Integrated Studies (UNU-CRIS). His research focuses on participatory approaches for AI regulatory sandboxes.

Cite this article: Genicot, N., & Moraes, T.G. (2025). Exploring the boundaries of AI regulatory sandboxes under the AI Act: Flexibility and real-world testing. *Cambridge Forum on AI: Law and Governance*, *1*, e36, 1–17. https://doi.org/10.1017/cfl.2025.10013