CYBERSECURITY NATIONAL LAB

SERICS
SECURITY AND RIGHTS IN THE CYBERSPACE

REGULATORY SANDBOXES FOR AI AND CYBERSECURITY

Questions and answers for stakeholders

edited by
Filippo Bagni
Fabio Seferi

The volume has been produced by:

In collaboration with:

Work partially supported by the project SERICS–Eraclito and CybeRights (PE00000014) under the MUR National Recovery and Resilience Plan funded by the European Union NextGenerationEU.

ISBN 9788894137378

Title: Regulatory sandboxes for AI and Cybersecurity. Questions and answers for stakeholders

February 2025

TABLE OF CONTENTS

TABLE OF CONTENTS

Introduction
THE WHITE PAPER ON REGULATORY SANDBOXES FORAIANDCYBERSECURITYANDITSSIGNIFICANCE FOR THE MISSION OF THE CINI CYBERSECURITY NATIONAL LAB

The Cybersecurity National Lab is the primary laboratory in the CINI network. The Lab is organized as a network of interconnected Nodes located in 74 major Italian Universities and research Institutions. The Nodes include 986 professors and researchers worldwide and span the whole country.

The Lab works towards fostering the Italian national cybersecurity ecosystem, via the promotion of a continuous process of aggregation of research and training structures in a multi- and inter-disciplinary perspective, pushing synergy and joint activities between public and private research entities.

This aggregation process includes a variety of initiatives:

- definition of common languages via National Frameworks in continuous interaction with key stakeholders, such as the Agenzia per la Cybersicurezza Nazionale (ACN), Accredia (the national accreditation body), and UNI (the national standardization body);
- community building through the organization of annual events, like ITASEC: the Italian Conference on CyberSecurity;
- creation of a cybersecurity technical workforce and the selection and training of cybersecurity young talents (such as CyberChallenge and OliCyber);
- cooperation to the development of the European Cyber Security Challenge ECSC; in particular, the Cybersecurity National Lab is a member of the ECSC steering committee and it has hosted and organized the 2024 edition of ECSC in Italy;
- participation in collaborative national and international research projects, such as the national SERICS Extended Partnership and the European SPARTA Project;
- involvement in several R&D projects in collaboration with public and private

entities aimed at improving the resilience of the Nation to cyber-attacks.

The Cybersecurity National Laboratory is also an active member of the European Cyber Security Organization (ECSO) and strongly cooperates with NIST (USA).

The transitioning of Cybersecurity and AI from technical disciplines for ICT specialists to societal imperatives due to their impact on critical infrastructures, economies, national security, and individual lives, has spurred in the EU several important regulations in many key areas:

- **Data Protection and Privacy:** the General Data Protection Regulation (GDPR) is a cornerstone of data privacy, enforcing strict rules on how organizations handle personal data while giving individuals greater control over their information;
- **Critical Infrastructure Protection:** the NIS2 Directive (Directive on Security of Network and Information Systems) aims to ensure the resilience and security of essential services, including energy, healthcare, transport, and financial services;
- **Artificial Intelligence Regulation:** the AI Act aims to establish clear rules for the development, deployment, and use of AI systems, particularly those impacting safety, privacy, and fundamental rights, such as high-risk AI applications;
- **Cybersecurity Certification Frameworks:** the Cybersecurity Act introduced an EU-wide cybersecurity certification framework to ensure trust and security in digital products, services, and processes;

  Financial Sector and Cyber Resilience: the Digital Operational Resilience Act (DORA) targets financial institutions, requiring robust measures to ensure resilience against cybersecurity threats.

Furthermore, given the fast-paced evolution of the technological landscape, it is very likely that more regulations and normative guidelines —let alone updates of existing ones— will be produced in the years to come.

In this context, ***regulatory sandboxes*** have emerged as an effective means of supporting the development of regulations in key areas by providing a controlled environment to test cutting-edge technologies under the supervision of regulatory authorities. Additionally, regulatory sandboxes can help strike a balance between encouraging innovation and protecting societal interests, such as privacy, safety, and

cybersecurity. They can also play a key role in enhancing trust in technology by anticipating possible weaknesses in regulations and roadblocks in their application.

By contributing to consolidating best practices, legal frameworks, and lessons learned, this white paper offers guidance to policymakers and stakeholders in implementing effective sandboxes. This will ultimately help bridge the gap between academic research and practical application, advancing Italy's cybersecurity capabilities.

This white paper is perfectly aligned with the mission of the CINI Cybersecurity National Lab which aims at aggregating research, fostering synergies within the Cybersecurity research community, and promoting collaboration with experts from different disciplines. With over 980 researchers across 74 Nodes, the CINI Cybersecurity National Lab can integrate diverse knowledge into sandbox experiments, ensuring robust and innovative solutions, and promoting the wide dissemination of the results, thereby contributing to the implementation of broader national and international Cybersecurity strategies.

This white paper has been conceived and developed in the context of **Eraclito** and **CybeRights**, collaborative research projects devoted to enhancing Cyber-risk management methodologies and best practices for national critical infrastructures. The Eraclito and CybeRights projects are two of the 27 research projects carried out by the SERICS Extended Partnership (PE00000014), a national initiative under the MUR National Recovery and Resilience Plan funded by the European Union NextGenerationEU.

Alessandro Armando
CINI Cybersecurity National Lab, Director
SERICS Foundation, Chairman of the Scientific Committee

Foreword
THE CHALLENGE OF "REGULATORY SANDBOXES" AND THE FUTURE OF EUROPEAN POLICY ON TECHNOLOGICAL INNOVATION

ANDREA SIMONCINI*

1.

Today, the European Union has undoubtedly taken world leadership in the regulation of technology and, in particular, of new and emerging digital technologies. In the global arena, we usually see three main actors: the US, China, and Europe. Of course, this is largely a simplification, as there are other major players besides the three mentioned, such as Taiwan, South Korea, or India. But, in any case, of all the key players in the race for Artificial Intelligence, the European Union is unquestionably the only one today that is systematically and analytically addressing the problem of regulation and setting legal limits in the development and application of new digital technologies.

This puts Europe at the forefront.

In the last 10 years, we have seen the emergence of what we may call the *EU digital acquis* (Bogucki et al., 2022), a truly comprehensive body of law regulating digital technology, with a large number of directives and regulations of great impact.

In the field of *Cybersecurity*, one thinks of the NIS1 (Directive (EU) No 2016/1148) and NIS2 (Directive (EU) No 2022/2555) or the Cyber Resilience Act ('CRA', Regulation (EU) No 2024/2847); in the field of *Data Protection*, in addition to the well-known GDPR, the Data Governance Act ('DGA', Regulation (EU) No 2022/868) and the Data Act (Regulation (EU) No 2023/2854); in the field of the regulation of *Digital Platforms*, the Digital Services Act ('DSA', Regulation (EU) No 2022/2065) and the Digital Markets Act, Regulation (EU) No 2022/1925); and finally,

* Full Professor of Constitutional Law at the University of Florence, Department of Legal Sciences. Principal Investigator of the Project CybeRights (SERICS).

for artificial intelligence systems and models, the Artificial Intelligence Act ('AI Act', Regulation (EU) No 2024/1689).

Europe is therefore continuing to build on its identity and on its model that distinguishes it from the Chinese or American context: a common market area that promotes and stimulates economic and industrial growth while ensuring a high level of protection of fundamental rights and the rule of law. As President Ursula von der Leyen has clearly stated[1]:

"The digital transition needs clear rules. People need to know that they can trust the technology in their hands. Businesses need predictability to plan their investment. And this is exactly why we have come up with the most ambitious agenda for digital reforms and investment in our Union's history."

2.

But there is also a flip side.

The risk is that uncoordinated and unclear over-regulation can become an undue burden and ultimately a competitive disadvantage for citizens and especially for European or foreign companies wishing to operate in Europe.

We must be very clear on this point: the problem is not the regulation, as such.

We do not intend to advance libertarian economic-legal ideologies that advocate a minimal (i.e. "zero") (Nozick, 1974) role for the state, or policies of - more or less - absolute deregulation in industrial production.

The question is not *whether* to regulate, but *how*.

In other words, what's increasingly important today is the *quality* of regulation, not the *quantity*. As a matter of fact, rulemaking in the field of technology is an age-old issue that has always challenged legal theory and the practice of the institutions and administrations in charge of such rules.

Technology, in fact, requires its own regulation, a 'technical' regulation, that takes into account, at least, three crucial factors.

---

[1] Keynote speech by President von der Leyen at the 'Masters of Digital 2022' event, available at: https://ec.europa.eu/commission/presscorner/detail/en/speech_22_746.

Firstly, there is the 'time' factor.

The revolution associated with the transition from analog to digital technologies has enormously amplified this dimension, especially because of the rapidity with which scientific knowledge and its technical applications evolve; a rapidity that risks rendering any attempt at discipline 'obsolete' and out-of-date a very short time after its adoption.

Then there is the problem of 'language'.

The recipients of these rules usually use their own 'professional' language; if the legal rules do not speak this professional language, they risk remaining unintelligible and therefore unworkable.

Finally, new technologies pose a problem of 'testing'; a problem probably unknown to the law-making procedures invented in nineteenth-century parliamentary democracies and left essentially unchanged in post-World War II constitutions. Legal norms are still conceived of as general and abstract rules governing the future, to which we are all bound.

In recent decades, the need to provide a space and time to test these prescriptions has become increasingly urgent. That is, there is a growing need to 'experiment' the new rules in a controlled environment before they become generally binding, to see if they actually succeed in having the hoped efficacy.

In fact, in many cases, regulators do not have a complete and accurate idea of how their rules will affect the intended recipients (e.g., the production system). Some rules, although they may be acceptable in the abstract, may turn out - in practice - to be inapplicable, or to be applicable only at excessive costs, or to outplace the product from the market altogether: well, all this vital information, despite the fact that deliberative procedures have been equipped by including in the pre-decisional phase various consultations and technical expertise, very often cannot be understood 'before' the concrete experimentation of the rule.

It is therefore necessary for the rule maker to be able to 'learn' from the applied experimentation and possibly supplement or correct the regulation itself in an adaptive-recursive process.

One might argue that today the European model faces a critical challenge to its identity in the face of the impressive expansion of new digital technologies.

The possibility of preserving its specificity - a flourishing and growing market in which an effective system of protection of civil, social, and political rights is guaranteed -

depends on 'good regulation', i.e. regulations that do not multiply administrative burdens for no reason and that do not irrationally overlap, creating irrational duplication or doubt and uncertainty for operators.

3.

A very clear signal of this need for simplification and clarification of technology law comes from the report on the future of European competitiveness commissioned by the European Commission from Mario Draghi in September 2024.

The report is straightforward on this point:

"While the ambitions of the EU's GDPR and AI Act are commendable, their complexity and risk of overlaps and inconsistencies can undermine developments in the field of AI by EU industry actors. The differences among Member States in the implementation and enforcement of the GDPR (...), as well as overlaps and areas of potential inconsistency with the provisions of the AI Act create the risk of European companies being excluded from early AI innovations because of uncertainty of regulatory frameworks as well as higher burdens for EU researchers and innovators to develop homegrown AI. As in global AI competition 'winner takes most' dynamics are already prevailing, the EU faces now an unavoidable trade-off between stronger ex ante regulatory safeguards for fundamental rights and product safety, and more regulatory light-handed rules to promote EU investment and innovation, e.g. through sandboxing, without lowering consumer standards. This calls for developing simplified rules and enforcing harmonised implementation of the GDPR in the Member States, while removing regulatory overlaps with the AI Act. This would ensure that EU companies are not penalised in the development and adoption of frontier AI." (European Commission, 2024)[2].

[2] M. Draghi, The future of European competitiveness, Part B | In-depth analysis and recommendations, September 2024, p. 79.

4.

This major regulatory challenge posed by new technologies is not new to European institutions.

Indeed, as early as 2001, the need for a new appropriate and responsive European regulation was explicitly recognized (Garben et al., 2018) by promoting a sharp regulatory quality strategy, the *Better Regulation Strategy*[3] (Redaelli, 2006).

The main objectives of this strategy are: a) ensuring EU policy making is based on evidence; b) making EU laws simpler and better, and avoiding unnecessary burdens c) involving citizens, businesses and stakeholders in the decision-making process.

The strategy has been articulated through a series of initiatives and documents[4]; in particular, two regulatory policy tools have been developed (*Better Regulation Guidelines* [5] and *Better Regulation Toolbox*[6]).

The term 'regulatory sandbox' - to which this *White Paper* is dedicated - appears in one of these two tools (the *Toolbox*[7]).

In the *Better Regulation Toolbox*, 'sandboxes' are included among the possible tools for improving regulation.

"*A range of regulatory and non-regulatory instruments or combinations of instruments may be used to reach the objectives of the intervention*[8]".

And *Regulatory Sandboxes* are introduced as follows:

---

[3] The first appearance in official European documents of the topic of 'Better policies, regulation and delivery' dates back to the 2001 White Paper on European Governance (European governance - A white paper, COM/2001/0428 final, Official Journal 287 , 12/10/2001) which, in turn, originated as a product of the 'Lisbon Strategy' (Presidency Conclusions of the Lisbon European Council 23 and 24 March 2000 in www.europarl.europa.eu/summits/previous.htm), the first relevant political act of the newly created Prodi Commission (1999-2004).

[4] Available at: https://commission.europa.eu/law/law-making-process/planning-and-proposing-law/better-regulation_en#simplifying-eu-laws

[5] Available at: https://commission.europa.eu/document/download/d0bbd77f-bee5-4ee5-b5c4-6110c7605476_en?filename=swd2021_305_en.pdf

[6] Available at: https://commission.europa.eu/document/download/9c8d2189-8abd-4f29-84e9-abc843cc68e0_en?filename=BR%20toolbox%20-%20Jul%202023%20-%20FINAL.pdf

[7] Cfr. Better Regulation Toolbox, cit., p.131.

[8] Cfr. Better Regulation Toolbox, cit., p.122.

"Technological transformation, the emergence of new products, services, and business models can be quite challenging from a regulatory perspective. To enable firms to test innovations in a controlled real-world environment, under a specific plan developed and monitored by a competent authority, a relatively new policy instrument – a 'regulatory sandbox' – can be set up". [9]

"Although no commonly agreed definition exists, regulatory sandboxes can be broadly described as schemes enabling producers to test innovations in a controlled real-world environment, under a specific plan developed and monitored by a competent authority".[10]

In the Draghi *Report* itself, already mentioned, the reference to regulatory sandboxes appears 17 times as a key tool to achieve in practice the two fundamental objectives: on the one hand, to ensure a simplified and effective single regulatory framework and, on the other hand, to promote the competitiveness and innovation of European industry.

"Experimentation should be encouraged via the opening up, EU-wide coordination and harmonisation of national 'AI Sandbox regimes' to companies participating in the plan. These experimental 'sandboxes' would enable regular assessments of regulatory hindrances deriving from EU or national legislation and provide feedback from private companies and research centres to regulators"[11].

5.

The contributions collected in this White Paper, edited by Filippo Bagni and Fabio Seferi, focus on this regulatory tool, which is quite new and, in some ways, still "unknown", but in other ways already experimented and increasingly used.

The word 'sandbox' originates from the technical language of software programmers as protected environments in which to test new codes while avoiding to produce relevant effects in the "real world"; and, as often happens within the new digital regulations, the term has moved from the computer-science lexicon to the legal one, becoming *'regulatory' sandboxes.*

---

[9] Cfr. Better Regulation Toolbox, cit., p.599.

[10] Cfr. Better Regulation Toolbox, cit., p.131.

[11] M. Draghi, The future of European competitiveness, Part A, p. 34.

The attention on the sandboxes then grew considerably when some of the most important regulations of the new EU digital acquis - the AI Act, the Cyber Resilience Act, and the Interoperable Europe Act - explicitly refer to these tools.

All European member states will therefore have to address this issue, and as we will see in this volume, they are already doing so.

The volume edited by Bagni and Seferi, and all the authors contributing to it, represents the first comprehensive study on the topic of regulatory sandboxes in Europe in the field of AI and Cybersecurity and is therefore particularly valuable and timely.

In the White Paper, one can find condensed both fundamental reflections on the European regulatory framework and on the legal basis for European action, analyzing also the regulations that explicitly deal with sandboxes, as well as the interplay with existing regulations - such as the GDPR -; on the other hand it is examined the enforcement dimension of regulatory sandboxes, understood as tools to support innovation, especially in the startup and SMEs sector, which is particularly crucial in Europe.

Finally, also the practical problems of regulatory sandbox implementation, comparatively examining some relevant cases that exist and operate today, up to considering the ethical profile of sandboxing, are comprehensively addressed in the volume.

In doing so, this White Paper fills an important gap in an area where both European and national institutions will soon be called upon to act. Regulatory sandboxes in perspective may prove to be a very important, if not indispensable, tool to ensure that the challenge of European regulation of technological innovation can be met.

FROM LEGAL EXPERIMENTATION TO REGULATORY SANDBOXES: THE EU'S PIONEERING APPROACH TO DIGITAL INNOVATION AND REGULATION

ERIK LONGO*, FILIPPO BAGNI**

SUMMARY

ABSTRACT

This contribution explores the concept of regulatory experimentation as a strategic response to the challenges of regulating emerging technologies within conventional regulatory frameworks. The rapid pace of technological innovation often outpaces traditional regulatory approaches, necessitating a more agile and adaptive framework. Regulatory experimentation, particularly through the use of regulatory sandboxes, has emerged as a promising solution. These controlled environments allow companies to develop, test and refine innovative products and services under regulatory oversight, while regulators refine their frameworks in real time to ensure adaptability, market safety and compliance with public interest objectives. This contribution constitutes the introduction to a broader study that examines the theoretical underpinnings, design and practical implementation of regulatory sandboxes, highlighting their dual role in fostering innovation and ensuring compliance with public interest objectives. By examining the interplay between innovation, regulation and societal safeguards, this research contributes to the ongoing discourse on the governance of emerging technologies. It highlights the potential of regulatory sandboxes to strike a balance between technological progress and a safe, well-

* Full Professor of Constitutional Law at the University of Florence. Contact email: erik.longo@unifi.it.
** PhD Candidate at IMT School for Advanced Studies Lucca and Legal officer at European Commission (DG Connect). The information and views set out in this article belong to the author and do not necessarily reflect the official opinion of the European Commission. Contact email: filippo.bagni@imtlucca.it.

regulated internal market, ultimately informing the development of more effective and responsive regulatory frameworks.

---

1. – INTRODUCTION

What occurs when technological innovation outstrips the ability of laws and regulations to keep pace? How can legislation and regulation adapt to the rapid evolution of digital technologies without compromising the protection of public and private interests? This challenge is known as the 'pace problem' and has significant implications for the governance of emerging technologies (Downes 2009). While there is broad agreement in policy circles that promoting innovation should be a priority, there is considerably less consensus on what this means and how to achieve it (Butenko 2015).

In recent years, the fields of law and technology have discovered new ways to make legal regulation more efficient and adaptable than ever before. Scholars and practitioners are focused on creating experimental tools for law and regulation to address the problem presented by Collingridge (1982). This dilemma highlights regulators' challenges with new technology, and specifically whether to act early on when circumstances are more flexible or to wait until the technology is more established. Delaying action can often lead to risks and increased costs for implementing any changes regulators may request.

One key element of this quandary is to reference regulations and technology correctly and the socio-technical environment in which they will operate (Dizon 2012).

This contribution seeks to address these questions by examining the challenges legislators and regulators encounter when attempting to govern emerging technologies. It begins with an introduction to the concept and practice of experimentation in law and regulation. Subsequently, it explores how the core conceptual and practical elements of experimentation align with the defining features of regulatory sandboxes. These sandboxes, established by EU legislators, serve as mechanisms for crafting preliminary regulatory frameworks that enable regulators to monitor the innovation-driven digital sector. Simultaneously, they provide companies with a controlled environment to develop, test, and validate innovative products over a limited period, prior to their deployment in the market or operational use.

2. – WHAT IS EXPERIMENTATION IN LAW

In classical terms, the law is created to command, direct, and change the reality through the power of authority (Hart 1961). At every level, legal provisions should be just, valid, and certain. Certainty is one of the most important objectives of the entire empire of the law. Legislatures create rules that must last, and judges, as well as administrations, interpret those rules by sticking closely to the text. Only when the reality changes and the legislature acknowledge it, can the law change through a formal procedure. Of course, the domain of law does not forget the experimental element, but this is just a prerequisite of the act of ruling. Experimentation comes before the law, also to understand the values to be balanced in legislation. It is therefore odd to speak of experimentation in law.

As we have seen, modernity, the acceleration of reality, science and technology have disrupted this geometrical form (Leibniz 2020) and made the regulatory landscape more complex than in the past. Accelerated social relationships have made it challenging for legislatures and interpreters to adhere to a single statutory conduct as just, valid, and effective. In a world dominated by informatics, all entities, natural and artificial, exchange information in a relentless dynamic flux of input and output. The subjection of society to the logic of 'control' (cybernetics) requires that law can also be seen as a technological device, that is, as a tool to be used to overcome the limits inherent in human experience. In response to this complexity, firstly in social security and then in finance and tax law, different forms of rules have appeared, which have in their same structure the possibility to change rapidly as the subject matter regulated changes.

This is in short, the reason why legislatures have started to enact rules that are different from the 'ordinary' form. They experiment a certain configuration of social relations, enact rules in a non-final form, measure the extent and performance of the rules, and reshape them through further regulatory intervention that depends on the outcome of monitoring. This endeavour is carried out in the name of the pursuit of economic, social, and cultural optimisation.

In legal terms, the whole phenomenon pertains to the domain of 'experimental rule', because the legislature enacts it on an experimental basis, i.e., as an attempted or trial regulation of the subject matter regulated. Experimental rules serve two additional purposes compared to non-experimental ones, often explicitly stated by the legislature:

(1) monitoring the effects of the rules, and (2) allowing for potential further adjustments to the framework. In summary, an experimental norm incorporates three elements: (1) a rule of conduct, (2) monitoring of the effects produced by the discipline, and (3) a second-degree norm that modifies, repeals or stabilises the experimental one.

On a theoretical level, we must acknowledge three elements. Firstly, for an experimental rule to be effective and valid, it is necessary that the legislature decides expressly to enact it (certainty); neither the judge nor the administration can determine the experimentality of the rule. Secondly, experimentation may include any type of rule. Thirdly, the experimental rule also serves to harvest information for changing reality through law; legislation exercises a kind of control in a 'cybernetic' sense (Wiener 1948), since regulatory activity generates retroactive feedback that is used to introduce other instructions, experimental and otherwise, into the system (Luhmann 2008).

Experimental rules as evidence-based can be applied to two different realities. The first case is the 'legislative experiment'. Two scenarios are conceivable in this regard. We might speak of a means test and an ends test of the law. In one case, one merely ascertains whether the choice of a means to achieve a particular goal was appropriate. In the other case, the experimentation is more profound and can lead to a rethinking of the aims of the entire rule.

On the other side, they can be applied to create and regulate an entire context of relationships. The main purpose of adopting an experimental rule in this case is to create a context (sandbox) in which certain conduct is subject to legal effects that are partially or completely different from the usual ones. Even these laws are subject to a test and should be reconsidered by the legislature at a certain point in time.

As we will see below, the concept of regulatory sandbox is as fascinating as it is complex and tricky. Thus, before analysing how regulatory experimentation plays out in the context of regulatory sandboxes, three critical considerations about the phenomenon must be made.

First consideration: While it is natural for legislators to exercise caution in regulating entirely new phenomena with unpredictable consequences and exponential effects, it seems problematic that human relations and economic interests are so readily subjected to experimentation. In fact, one could provocatively suggest that the legislator is incapable of half-measures: on one extreme, the play and spontaneity of the child; on

the other, the sterile indifference of an algorithm developed by a software developer. One could go so far as to say that when normative activity becomes unreflective creativity, like plastic matter in the hands of an infant, or mere technological innovation, as it is functional to the introduction of new devices or applications, law ceases to be prudentia juris and politics ceases to be just measure. And it is for this reason that in 'experimental norms' one can read against the light the problem of the relationship between legality and information technology (Lessig 2006).

Second consideration: Underlying the phenomenon of regulatory sandboxes is a narrative in which technological progress is posited as a political goal. Legislatures and governments promote progress because science has replaced politics, which is dominated by the arbitrariness of power. This gives rise to the notion of the state as a 'business' and society as a laboratory (which is basically behind the concept of regulatory sandboxes).

Third consideration: The manipulation of social relations occurs through the incorporation of 'control', which adopts the model of cybernetic feedback and infuses it into law. Social relations are malleable matter in the hands of the sovereign, who, as the holder of absolute power in principle, can also afford to experiment. In this context, while regulatory sandboxes are likely to become a defining feature of our time and secure a stable place in jurisprudence, greater legal reflection is needed on the notion of rules as 'miserable' entities."

3. – REGULATORY EXPERIMENTATION

As the shortcomings and obstacles of the main state and market approaches to regulation have become increasingly apparent, legislators and policymakers have begun to explore a much wider range of policy mechanisms than simple legislation and regulation. As for the latter, new forms of policy instruments include economic instruments, self-regulation, information-based strategies, and voluntarism (Baldwin 2010). However, while these instruments offer a much wider range of policy options than traditional regulation, they have yet to be completely effective when used in markets closely linked to technology.

The accelerated pace of technological advancement affecting society and the increasing complexity of contemporary problems contribute to destabilising political and administrative systems based on rigid regulation dynamics. Complex technologies, like

Artificial Intelligence, cannot be tackled with permanent solutions or in isolation. On the contrary, they require constant adaptability and a systemic vision to be used by people.

The range of policy and governance solutions that allow regulators to adapt regulation to a rapidly changing world and be keen and supportive of innovation is broad. Today, the most important are those that allow the embracement of experimental policies to promote a fair and good balance between innovation and the protection of public interests. Policies promoting regulatory experimentation are sensitive and reactive, involving direct engagement and interaction between regulators and market participants (OECD 2023). Benefits include avoiding the potential regulatory-market gaps that sometimes accompany hard law. At the same time, regulatory experimentation impacts market and innovation dynamics. Information and regulatory learning are at the core of experimental policies (Hofmann et al. 2022).

The most important development in this process has been the move from delayed to real-time supervision of firms. As in the case of financial regulation, supervision was initially delayed, with regulators increasingly demanding more information in order to supervise financial firms in real time. This has led to demands for tailored forms of (corporate technology) governance in the private sector. The consequences of this development at a higher level of supervision should also be noted. Supervision, which cannot be human-based but must be data-driven, must use the most advanced and sophisticated computational tools to ensure the highest levels of accuracy and appropriateness. This requires technical resources and appropriate adjustments to the administrative process, which, as we know from the literature on financial regulation, are far from guaranteed. The channelling of technological expertise to regulators is the rationale for the most advanced form of experimental regulation, in addition to facilitating contact between innovators and regulators: the "regulatory sandboxes".

These tools generally refer to mechanisms through which authorities collaborate with companies to pilot innovative products or services that challenge established legal frameworks. As such, regulatory sandboxes foster the development and testing of innovations, in some cases also in a real-world environment (business learning) and support the formulation of experimental legal regimes to guide and support businesses in their innovation activities under the supervision of a regulatory authority (regulatory learning). This approach aims to enable experimental innovation within a framework

of controlled risks and supervision and to improve regulators' understanding of new technologies.

As will be further analysed in the next section, over the past years the sandbox approach has gained considerable traction across the EU as a means of helping regulators address the development and use of emerging technologies – such as AI and blockchain technologies – in a wide range of sectors. Today, a new momentum has emerged, with the regulation of digital technologies being one of the main objectives of the EU in 2030. EU legislatures increasingly favour a more agile approach to innovation and regulation in the digital sectors, mainly promoting regulatory tools to support start-ups in getting challenging technologies to the market and enabling cross-border testing.

With the globalisation of the economy and the renovation of social and economic structures even the part of the law that is closer to business and markets has needed to include forms of smart, agile, and responsive governance structures. Regulatory sandboxes stem from a new regulatory approach that even bypasses de-regulation and laissez-faire by embracing regulatory experimentation (Gromova 2023).

By relaxing specific regulatory requirements and anticipating conformity controls, sandboxes lower barriers to entry for innovative products, encouraging startups and established firms to pursue ambitious projects that might otherwise be hindered by regulatory complexity (Zetzsche 2017). With innovative collaborations with stakeholders – including across borders – companies can conduct live testing with real customers, gathering authentic feedback and performance data. This accelerates the product development cycle and helps refine offerings before a full launch. Participants work closely with regulators to gain insight into compliance expectations. In addition, the controlled environment enables identifying and managing potential risks associated with new products. Companies can proactively address issues and improve their innovations' safety and reliability, helping to develop clarity and certainty and communicate them to market actors.

4. – REGULATORY SANDBOXES IN THE EUROPEAN DIGITAL REGULATION

Regulatory sandboxes offer the flexibility needed to align with the EU's evolving approach to digital regulation. As one of the key pillars of the European Commission's strategy over

the last decade, digital regulation aims to foster innovation while ensuring compliance with fundamental rights and legal norms.

The expression 'digital regulation' is not a legal one. It attempts to grasp the virtues of the European legislators and regulators in creating a new environment for the EU market to grow by respecting fundamental rights and creating a balance between innovation and respect for the law. In its important scholarly work on the effects of EU law worldwide, Anu Bradford (2023) depicts digital regulation as a form of power that thoroughly influences the production and market of digital services and, as such, the digital economy.

As we have seen, the rapid pace of technological progress and the introduction of new products and services have created a new era of regulatory complexity. The inherent flexibility of technological progress has tested the traditional rigidity of law- and policymaking. As a result, innovative regulatory approaches have been devised (some referred to as 'experimental lawmaking') (Ranchordas 2021), including the addition of 'experimental clauses' (van Gestel-van Dick 2011; Mousmouti 2018). These legal provisions allow enforcement authorities flexibility in dealing with emerging technologies, products or strategies, even if they do not fully comply with existing legal norms (Attrey – Lesher - Lomax 2020).

These clauses serve as a basis for innovative regulatory experimentation. In this context the tool of regulatory sandboxes has prospered (European Commission 2023, 131). The term 'sandbox' traditionally evokes two images - one of playgrounds, where children can play without restrictions, and the other of computing, where it represents a safe testing environment that protects the system from malicious programs (Yordanova 2019). The addition of 'regulatory' refers to a tool designed to test new services and products in a simulated regulatory environment.

As already mentioned, a regulatory sandbox provides a controlled space for companies operating in regulated sectors (such as banking, finance, and insurance) or high-tech areas (such as AI systems, digital products) to test their innovative products and services for a limited period of time. During this time, the testing is carried out in continuous communication with the supervisors responsible for ensuring the compliance of the innovative product/service before it enters the market, potentially benefiting from a simplified transitional regime. There's no one-size-fits-all sandbox model, and it depends

on factors such as the technology used, the sector, the supervising authority, and others.

The main advantage of the regulatory sandbox is the opportunity to experiment and possibly make mistakes with a product that may not yet comply with existing rules, under the careful supervision of regulators. The goal is to develop an innovative product/service that complies with European market rules by the end of the testing period. Specifically, regulatory sandboxes have a dual purpose: (a) to promote business learning, development and testing of innovations in a real-world environment; and (b) to support regulatory learning by creating experimental regulatory frameworks to guide and support businesses in their innovative activities under the supervision of regulators.

The essence of the sandbox is based on a traditional win-win scenario. On the one hand, it facilitates market growth and development by enabling rather than hindering the introduction of technologically innovative products and services. At the same time, it ensures an appropriate level of consumer protection and competition through ongoing dialogue with regulators. While the company develops a product within an environment that provides guidance and, under certain conditions, possible regulatory exemptions, the regulator gains insight into the operator's activities and, through continuous dialogue, acquires new technical expertise.

5. – MAIN ELEMENTS OF REGULATORY SANDBOXING

Although the concept of regulatory sandboxes varies widely, certain common elements can be identified (Ranchordas 2021a). First and foremost, the regulatory sandbox concerns innovative products or services that are not yet available on the market and that offer added value to consumers or society (for example, by contributing to public policy objectives such as environmental protection). In addition, the product or service must be developed to a stage that allows immediate testing (neither too preliminary nor too advanced to allow changes), and the activity to be tested must be economically viable throughout the testing period. Finally, it is necessary to identify the applicable legislation, the legal hurdle against which the product or service is to be tested, and the appropriate institutional authority (Bagni 2023).

Ensuring legal predictability is crucial for the participating companies. The boundaries and conditions of the sandbox should be defined in advance, preferably

through legislation or memoranda of understanding with market surveillance authorities. It is important to clearly define the legislation and sectors included in the test, the planned exemptions, the rules for entry, the duration, and the conditions for exit, in order to facilitate the measurement and evaluation of the results of the sandbox. In addition, even in a controlled environment, appropriate safeguards need to be in place (e.g. security measures for autonomous vehicle testing).

Participation in the regulatory sandbox is typically subject to approval, monitoring, and evaluation by the competent authority, and is limited to a certain number of participants. The authority usually opens temporary windows (open calls) inviting interested operators to participate by presenting their projects. After the window closes, a selection and interview process take place, leading to the definition of approved projects and the launch of the experimental project.

This framework has both advantages and disadvantages. On the positive side, it allows companies to test their innovations under specific and (sometimes also) real-world conditions and to gain a better understanding of the relevant regulations. Participation in a regulatory sandbox can also facilitate access to funding and reduce time-to-market. Furthermore, from a regulatory perspective, sandboxes allow for a degree of flexibility without compromising regulatory standards, thus facilitating learning in highly complex and difficult to regulate sectors.

However, there are also downsides. Firstly, regulatory sandboxes could potentially increase the risk of fragmentation of market regulation in the absence of a common approach, leading to different outcomes across the EU. Secondly, these tools require dedicated resources, time, and expertise from both parties (companies and regulators), which may not always be feasible for smaller companies. Thirdly, normally a participation in a sandbox does not automatically guarantee product or service compliance and risk-free market entry. Finally, operationally, sandboxes present a number of complexities, apart from the specific technical complications related to each reference sector (banking, insurance, finance, technology, digital) (European Commission 2023, 600).

For a long time, there was no precise institutional definition of regulatory sandboxes. Recently, however, three different European regulations have included such a definition, highlighting the rising importance of the regulatory sandbox tool in the eyes of European lawmakers. Specifically, we refer to the Artificial Intelligence Act ('AI Act',

Regulation (EU) No 2024/1689), which defines the 'AI regulatory sandboxes' in Article 3(1)(55), the Interoperable Europe Act (Regulation (EU) No 2024/903), which defines the 'interoperability regulatory sandboxes' in Article 2(1)(14), and the Cyber Resilience Act (Regulation (EU) No 2024/2847), which (indirectly) defines the 'cyber resilience regulatory sandboxes' in Article 33(2).

The intersections between these three regulations are analysed in more datil in the following sections of this publication. To conclude this first contribution, it is crucial to emphasise that all three definitions share some common elements with the core concept of a regulatory sandbox as discussed above. Specifically, they all refer to a controlled framework established by a competent authority in which participants - whether public or private - can develop, validate, and test innovative products under regulatory supervision for a limited period of time.

This is merely the starting point of the discussion. In subsequent contributions, this publication will provide a more detailed examination of the convergence between the legislative intent and the technical features that define the regulatory sandbox instrument. Future sections will address many of the critical issues surrounding regulatory sandboxes and provide an in-depth analysis of their potential, challenges, and wider implications. This introductory contribution sets the stage for a comprehensive exploration of this evolving regulatory framework.

LEGAL BASIS FOR REGULATORY SANDBOXES: KEY ASPECTS FOR A COHERENT THEORETICAL AND PRACTICAL FRAMEWORK

GIUSEPPE MOBILIO*, MATTEO GIANNELLI**

SUMMARY

ABSTRACT

The legal basis for the regulatory sandbox sets up and regulates this advanced kind of experimentation tool used to find better regulation for new technologies. When defining the legal basis for a regulatory sandbox, any policymaker needs to consider several issues to ensure legality and the rule of law, to protect fundamental rights and democratic systems, but also to adopt appropriate and effective regulation for new technologies. The aim of this paper is to address some of these issues in relation to the legal basis of the RS, in terms of the choice of the regulatory instrument, its content and its effectiveness. The paper provides an overview of the legal basis of relevant regulatory sandboxes in the European area, touching upon the aspects that most directly affect the characteristics of this scheme. In particular, the analysis focuses on the features that define the models of regulatory sandboxes, such as purpose, scope, context, and the subjects involved (public authorities entrusted with management and supervision), their powers and tasks, the possibility of granting a relaxation of existing legal

* Associate Professor in Constitutional Law at University of Florence. Contact: giuseppe.mobilio@unifi.it
** Assistant Professor in Constitutional Law at University of Florence. Contact: matteo.giannelli@unifi.it

rules, the existing limits and constraints on their discretionary powers, and the consequences of the experimentation.

---

1. – INTRODUCTION

The aim of this contribution is to address some of the topics related to the legal basis of the regulatory sandboxes, i.e. the norm that sets up and regulates a regulatory sandbox, with regard to the choice of the legal instrument, its content and its effectiveness. The paper will touch upon some of the aspects that most directly affect the characteristics of the regulatory sandbox and that any policymaker must consider, both in theory and in practice.

There is no single definition or model for regulatory sandboxes. There are many norms at the EU level that provide for regulatory sandboxes, but without a homogeneous approach (Ranchordás 2021, 2; Ranchordás and Vinci 2024, 108-109; Bagni 2023, 4). For the purposes of this contribution, it is sufficient to consider a number of profiles that any legal basis for a regulatory sandbox should address, as already highlighted in a recent Commission Staff Working Paper (European Commission 2023, 10 ss.).

2. – THE CHOICE OF THE REGULATORY INSTRUMENT FOR THE LEGAL BASIS: AN OVERVIEW

The legal basis can be formulated through a variety of regulatory instruments with different characteristics, e.g. in terms of democratic legitimacy, freedom of content definition or degree of specificity (Pagallo et al. 2019, 3; Costantini 2021, 169).

2.1. – Law

The regulation of the regulatory sandbox is best established through legislation, which represents the highest expression of public bodies with the strongest democratic legitimacy.

Germany is a particularly interesting experience in this respect. Since 2019, a government-wide strategy has been defined with a series of documents adopted by the Federal Ministry for Economic Affairs and Energy (BMWi), in which the issue of legal

basis and legislative exemptions is clearly highlighted. Among the clauses provided for at the legislative level are, in particular, those on mobility, which were introduced on the basis of the experimental clause pursuant to Art. 2(7) of the Carriage of Passengers Act (PBefG).

Another interesting example is the Swiss Energy Regulatory Sandbox. From 1 January 2023, Art. 23a of the Federal Act of 23 March 2007 on the Supply of Electricity (ESA) allows the Federal Department of the Environment, Transport, Energy and Communications (DETEC) to approve sandbox projects and their implementation. The aim is to support innovation and to facilitate the continued development of draft legislation on electricity supply.

The French regulatory sandbox for telecoms experimentation is similar. Section 92 of Law No. 2016-1321 of 7 October 2016 for a Digital Republic amended the 'Post and Electronic Communications Code' (Art. L. 42-1.IV and Art. 44.IV) in order to allow companies that so request to be exempted from all or part of the obligations related to the use of frequencies, numbers or even the status of network operator to allow start-ups and entrepreneurs to test their innovative technology or service.

In Italy, it is worth mentioning the case of 'Experimentation Italy', established by Art. 36 of Decree-Law no. 76 of 16 July 2020 (Simplification and Digital Innovation) converted by Law no. 120 of 11 September 2020, which is characterised by a 'multi-sectoral' regulatory sandbox aimed at the digital transformation of the public administration. Here, companies and public administrations can submit their projects to the government, while at the same time requesting a temporary exemption from government regulations.

A very specific case is the Portuguese Technology Free Zones ('Zonas Livres Tecnológicas'; ZLTs), where the government has decided to adopt a framework law. Decree-Law No. 67/2021 of 30 July establishes the legal basis for the creation of ZLTs, following on from Council of Ministers Resolution No. 29/2020 of 21 April. The Decree-Law does not create the ZLTs per se, but, as stated in the Resolution, it establishes the general principles for their creation and regulation. Each ZLT will be subject to specific rules defined by the relevant member of the Government in collaboration with the relevant supervisory authorities, depending on the specific characteristics of each sector.

2.2. – Government regulations

Various legal bases are then specified in government regulations or decisions. In these cases, the law provides the general discipline and leaves the task of further specification to lower sources.

This is the case with Austria's experimentation with automated vehicles in real traffic conditions. Here the Motor Vehicle Act 1967 was amended in 2016, and the Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK) accordingly adopted the 'Ordinance on Automated Driving', which defines the specific legal requirements for the different automated mobility use-cases.

A similar situation has arisen with the Spanish AI Regulatory Sandbox. Art. 16 of Law no. 28 of 21 December 2022, on the Promotion of the Start-up Ecosystem, provides for the creation of controlled environments for limited periods of time, with a view to operationalising the requirements of the European AI Act (Regulation (EU) No 2024/1689), which was at that moment under discussion. Accordingly, the Royal Decree no. 817/2023 was adopted, defining the specific rules for the creation of the regulatory sandbox.

In Italy, one of the first experiences of regulatory sandbox was the Financial Services Regulatory Sandbox, established by the Decree Law no. 34 of 30 April 2019 ('Growth Decree'), which sets out the 'FinTech Committee rules and experimentation', i.e. the regulatory sandbox for FinTech activities, then implemented by Decree no. 100/2021 of the Ministry of Economy and Finance.

2.3. – Deliberations of regulatory authorities

Many other regulatory sandboxes see the protagonism of various regulatory authorities. Therefore, the deliberations of the regulatory authorities contribute to the construction of the legal basis. In this way, the policymaker chooses a different circuit from the government administration (Buocz et al. 2023, 362).

Consider the Maltese Sandbox Regulatory Framework on the experimentation of distributed ledger technology (DLT) within the gaming industry. In accordance with Art. 7 of the Gaming Act (Chapter 583 of the Laws of Malta), the Malta Gaming Authority (MGA) offers licensed operators the opportunity to apply for this kind of technology

in order to exploit the potential of digital currency without the intermediation of other operators.

Another notable example is the Austrian Financial Market Authority Sandbox, which was established by Art. 23a of Financial Market Authority Act (FMABG), as amended in Federal Law Gazette I No. 89/2020. Here the regulatory sandbox allows FinTechs or licensed entities to test an innovative ICT-based business model that is currently under development in a controlled environment.

In all these cases, the legal basis establishes the sandbox scheme and regulates the resulting powers of the regulatory authority. In some cases, however, there is no general rule covering both of these contents, so the regulatory sandbox is activated autonomously by the regulatory authority by means of a very broad framework regulating its powers.

This is what happens in Italian experimentations in the energy sector, where the 'Regulatory Authority for Energy Networks and Environment' (ARERA) has been granted wide regulatory powers by law (Law no. 481 of 14 November 1995). Since 2010, ARERA has been carrying out many experimentations in the field of energy system innovation, without relying on a sandbox scheme. In these cases, the legal basis is provided by ARERA's constitutive law and its deliberation which activates the experimentation.

Similarly in the case of the Bank of Greece Regulatory Sandbox, the Executive Committee Act 189/1/14.05.2021 was adopted with the aim of defining the terms and conditions for the establishment of regulatory sandboxes, and facilitating financial innovation, enhancing legal certainty, and promoting knowledge.

In the case of the Danish Regulatory Test Zones for energy technologies, the regulatory authority in the field of energy is promoting a number of initiatives based on a policy document adopted by the main political forces of 2018 ('Energy Agreement') to develop the electricity market and determine the potential role of the gas system in the green transition.

Finally, the Danish Civil Aviation and Railway Authority may grant an operational authorisation in the specific category referred to in Regulation (EU) 2019/947 on the rules and procedures for the operation of unmanned aircraft.

2.4. – Other types of acts

Lastly, individual regulatory sandboxes may find timely regulation in specific acts that are

not necessarily of a regulatory nature.

For example, in the case of the Portuguese ZLT mentioned above, each of them shall have internal regulations, drawn up by the respective management entity, subject to the opinion of the competent regulatory authority and the approval of the testing Authority (Art. 6.3 Decree Law).

Finally, the case of the Austrian Framework Conditions for Automated Driving also deserves consideration for reference to additional specific sources, such as the 'Code of Practices'. These regulatory instruments contain the guidelines for defining the measures to be taken to ensure safety during tests of automated vehicles on public roads. All test organisations should comply with the Codes, which are formally qualified as not being legally binding, but rather intended to promote responsible testing.

3. – ELEMENTS OF THE LEGAL BASIS

Certain elements defined by the legal basis help to qualify the regulatory sandbox model. The definition of these elements is not completely open but depends on the regulatory instrument chosen for the legal basis and its content (Ranchordás 2021, 6).

3.1. – Purposes of regulatory sandboxes

The legal basis is called upon to position the regulatory sandbox according to a first alternative that conditions the whole model and purpose of testing. In fact, a regulatory sandbox can be aimed at regulatory testing, i.e. the possibility of experimenting with the modification, interpretation and application of a regulation for the purpose of testing and providing lessons to policymakers. Differently, the regulatory sandbox can be aimed at product testing, for the benefit of the economic operator wishing to place a new product or system on the market, as long as it complies with the regulation in force.

In the first case (regulatory testing), the legal basis will need to consider the type of regulation being tested, so it will need to be at a higher level than the more specific rules being tested. In the second case (product testing), the legal basis is likely to be more detailed, or a government or regulatory authority will be in charge of adopting specific rules, as the technical specifications of a particular technology need to be addressed.

Regulatory experimentation is one of the main objectives of regulatory sandboxes.

This is evident in the case of the Portuguese ZLTs, where the founding Decree-Law also provides for the possibility of creating specific legal and regulatory instruments aimed at facilitating the testing of technologies, products, services and processes. This is in line with the Council of Ministers Resolution 29/2020 of 21 April, which states that the legal framework to be approved must consider not only mechanisms to encourage experimentation, but also mechanisms to make the law more flexible. In the case of Experimentation Italy, it is planned to update existing regulations every time an experimentation proves successful, thus opening up opportunities and simplifying procedures for the benefit of the national economic system.

Many regulatory sandboxes are primarily designed for product testing. This is the case of the Austrian Framework Conditions for Automated Driving, which aims to test automated vehicles thanks to rules defined at the government level; or the Maltese Sandbox Regulatory Framework on the experimentation of DLT within the gaming industry, which offers licensed operators the opportunity to apply for the use of Innovative Technology Arrangements.

But in many cases the purpose of the regulatory sandbox lies in the middle of this alternative. Consider the Spanish AI Regulatory Sandbox, whose main purpose is to test in advance the regime defined by the European AI Act, both to support compliance efforts of new products and to provide European and national legislators with insights on how to interpret legal requirements. In the case of the Italian Financial Services Regulatory Sandbox, the aim is also to move from product testing to regulatory testing. Through the regulatory sandbox, the supervisory authorities aim to support the growth and development of the Italian FinTech ecosystem, while monitoring the latest technological developments and identifying the most appropriate and effective regulatory interventions to facilitate the development of FinTech.

Beyond the approach that formally emerges from the legal basis, it is always necessary to look at practice to understand what the objectives of individual experiences are. In the German Passengers Transportation Act Regulatory Sandbox, the main objective does not seem to be changing the regulations but developing new products with high technological content, limiting possible risks. However, a closer look at the content and approach of the periodic reports of the Ministry of Economic Affairs reveals that the ultimate goal is still to 'improve the innovation-friendliness and adaptability of the legal framework by

creating greater legal flexibility and regulatory learning' (German Federal Ministry for Economic Affairs and Climate Action (BMWK), 2022).

3.2. – Scope of regulatory sandboxes

The legal basis also plays a crucial role for the scope of experimentation. The regulatory sandbox may be context-specific, as the experimentation relates to a limited scope or a specific technology, or it may be more broadly open to innovations proposed by individual participants, without any specific restrictions (Bromberg et al. 2017, 8; Brown and Piroska 2021, 5).

In these different scenarios, the legal basis must consider the level of maturity of existing regulations and the discipline governing the type of technology being experimented with. Thus, if there is no regulation of the technology, policymakers will have a greater opportunity to choose the regulatory instrument and decide the degree of specificity of the content of the legal basis. On the other hand, if there is a regulation, policymakers will be constrained in its choice and will have to choose the law in the face of broader experimentation, which may then be specified by other public authorities.

A context-specific example is the French regulatory sandbox for telecoms experimentation, which allows quick, easy and temporary access to numbers and frequencies authorised by Arcep for experimental purposes. In this case, the framework is regulated by a source (the law) that includes the ordinary frequency and number management regime.

Instead, broadly open examples include the Spanish AI Regulatory Sandbox, which, in the absence of prior regulation, is governed by a Royal decree that allows applicants to experiment with 'high risk' AI systems, under a legislation (the AI Act) that is not yet applicable. Another case is the Portuguese ZLTs, which aim to develop innovative technology-based products, services and processes. Here the regulatory sandbox framework lies in the law. The Italian Financial Services Regulatory Sandbox is also open to innovations proposed by individual participants, in collaboration with the authorities according to a process detailed by Ministerial Decree no. 100/2021.

3.3. – Environment of regulatory sandboxes

The legal basis must be adapted to the controlled context in which the regulatory sandbox

takes place, i.e. whether in the real world or in strictly controlled environments (Omarova 2020, 76).

In the first case ('real-world conditions'), the experimentation may involve unintended third parties and thus have consequences for their fundamental rights and freedoms. In the second case, this risk does not occur. For this reason, the legal basis must bear the burden of providing sufficient guarantees and protection for those involved, especially if the regulatory sandbox is performed in the real world.

For regulatory sandboxes in real world conditions, we can consider the Austrian Framework Conditions for Automated Driving, which has a legal basis providing a long list of requirements for testing on public roads (Art. 1, par. 1, p. 3, of the Ordinance on Automated Driving). This is similar to the German Passengers Transportation Act Regulatory Sandbox, where the Carriage of Passengers Act (PBefG) provides for the possibility of testing new modes or means of transport on public roads, insofar as they do not conflict with public transport interests.

For regulatory sandboxes in strictly controlled environments, we can recall the Portuguese ZLTs, where the framework law mentions regulatory sandboxes in physical environments, geographically located, in a real or quasi-real environment, used for testing and experimentation. In the case of the Maltese Sandbox Regulatory Framework on the experimentation of DLT within the gaming industry, virtual tokens will be assessed by the MGA on a case-by-case basis and must be kept within a closed loop ecosystem. Similarly in the case of the Bank of Greece Regulatory Sandbox, the mechanism allows participants to carry out small-scale testing of innovations in a controlled regulatory environment, within specified parameters and timeframes, while formally engaging and cooperating directly with the Bank of Greece and preserving the financial stability and efficiency of the financial system.

4. – SUBJECTS

The legal basis is also conditioned by its author, due to its legitimacy, powers, framework of competences and freedom to design a regulatory sandbox.

4.1. – Government and/or regulatory authorities

A further feature established by the legal basis is the nature of the public authorities

involved and the role they play. In the main EU examples of regulatory sandboxes, it is possible to distinguish between experiences that rely on political/administrative authorities, mainly belonging to the government, and experiences that rely on regulatory authorities.

Regulatory sandboxes based on government authorities include the Austrian Framework Conditions for Automated Driving, which assigns the Austrian BMK the task of managing the scheme; or the Spanish AI Regulatory Sandbox, which involves bodies within the Ministry of Economy, Trade and Enterprise and the Ministry of Economic Affairs and Digital Transformation, such as the State Secretariat for Digitalisation and Artificial Intelligence.

Among regulatory sandboxes governed by regulatory authorities we need to distinguish when the legal basis attributes new powers to pre-existing authorities, as in the case of the Austrian FMA Sandbox with the Financial Market Authority, or the French regulatory sandbox for telecoms experimentation, with the French Arcep, or the Bank of Greece Regulatory Sandbox, with the Central Bank of Greece.

In other cases, the legal basis establishes new authorities, such as the Portuguese ZLTs, where the constitutive Decree-law also defines the governance model for the ZLTs, creating a 'testing authority' which is responsible for centrally managing and boosting the network of ZLTs, without prejudice to the powers of other entities.

Regulatory sandboxes may also provide for the involvement of a variety of public authorities. Regulatory authorities may also be joined by other political or administrative authorities. This is the case, for example, of the Austrian FMA Sandbox, where the legal basis establishes an advisory board (Regulatory Sandbox Beirat) at the Federal Ministry of Finance to assess the impact of sandbox business models and to contribute to the case-by-case evaluation, in particular the assessment of public interest objectives. Also, in the case of the Italian Financial Services Regulatory Sandbox, a new ad hoc body has been created at the ministerial level, namely the 'Fintech Committee', composed of members of the government and many regulatory authorities (in the fields of banking, stock exchange, insurance, data protection, digital services) involved in FinTech. The Committee is responsible for identifying objectives and defining programmes and actions to promote the development of techno-finance, as well as making regulatory proposals.

4.2. – Powers and tasks

However, it is obviously not possible to limit the analysis to the type of public authority involved. What really matters are the powers and tasks that the legal basis assigns to these actors.

The legal basis must be consistent with the general rule governing the powers of the public authority. If the general rule is contained in a law, it must be carefully considered whether the legal basis of the regulatory sandbox has a similar level or whether it provides mechanisms for linking different sources of law.

In the Austrian FMA Sandbox, the law establishing the regulatory sandbox empowers the FMA to grant a restricted licence, which allows experimentation to commence, in accordance with the listed applicable Federal laws which generally grant supervisory powers to the FMA. Similarly, in the French regulatory sandbox for telecoms experimentation, the founding law allows experimentation to commence subject to authorisation by Arcep, which is the authority normally competent to grant licences for their use under the 'Post and Electronic Communications Code', which is a law. The same applies to the Maltese Sandbox Regulatory Framework on the experimentation of DLT within the gaming industry. In the case of the Swiss Energy Regulatory Sandbox, the framework for each regulatory sandbox, as well as the rights and obligations of project participants, are defined in an ad-hoc ordinance adopted by the Federal Department of the Environment, Transport, Energy and Communications (DETEC). The Swiss case is particularly interesting because the law designates the government authority responsible for the regulatory sandbox, defines its powers and the possibility of activating the regulatory sandbox (Art. 23a Federal Electricity Supply Act), but leaves the concrete definition of such crucial aspects to DETEC and the Swiss Federal Office of Energy (SFOE), with wide room for manoeuvre.

4.3. – Legal flexibility

One of the most characteristic aspects of a regulatory sandbox is that it allows experimentation with new technologies by relaxing the constraints of current regulations (Yefremov 2019, 84). This is not a mandatory element, but this feature has many advantages: on the one hand, it allows the participant to test a system or a product without

immediately having to worry about compliance with the rules in force, which can be numerous and complex; on the other hand, it allows the policymaker not to discourage innovation and to refine these legal rules once the technology being tested is understood.

The choice of regulatory instrument for the legal basis, its content, its degree of specificity and other aspects are determined by the nature of the legislation to be derogated. At the same time, the legal basis will have to decide whether to specify directly which rules can be derogated from, or to entrust another authority responsible for the regulatory sandbox with the task of identifying them.

There are many examples of regulatory sandboxes that do not allow participants to derogate from the applicable rules. In the Austrian FMA Sandbox, the FMA may grant a restricted licence, approval, authorisation or registration in accordance with the respective applicable Federal Acts that regulate the same powers of the same FMA listed in the legal basis (Art. 2 paras. 1 to 4). In this context, the supervisory requirements are only adapted within the scope of the principle of proportionality for supervision depending on the business model, where supervisory laws permit this. Likewise in the Austrian Framework Conditions for Automated Driving, the testing of autonomous vehicles must in any case be carried out in accordance with the provisions of the applicable legislation, such as the general Austrian road traffic regulations (StVO 1960). In the case of the Maltese Sandbox Regulatory Framework, approval to participate in the sandbox is conditional on the applicant holding the relevant licence issued by the MGA, without prejudice to regulatory requirements stemming from other applicable legislation.

In other hypotheses, participation in the regulatory sandbox allows for a derogation from the regulations in force. This is the case in the French regulatory sandbox for telecoms experimentation, which allows companies that request it to be exempt from all or part of the obligations related to the use of frequencies, numbers or even the status of network operator, for a maximum period of two years. Similarly, in the case of the Italian Financial Services Regulatory Sandbox, numerous exceptions are provided for by Ministerial Decree No. 100 of 30 April 2021, implementing Art. 36 of Decree-Law No. 34 of 2019 (Growth Decree). In fact, the latter only provides for the adoption of one or more regulations to define the conditions and modalities for carrying out the experimentation, leaving a wide discretion to the public entities involved, in particular for the issuance of the authorisation act. As a result, the ministerial decree provides for numerous derogations

that the competent authority may grant for the purposes of the experimentation (e.g. to the laws on banking and credit) 'in compliance with European Union law and the principles of proportionality and equal competition between operators'. In the case of the Swiss Energy Regulatory Sandbox, the approval of projects that partially deviate from the current legal framework is based on Art. 23a of StromVG and the possibility to approve such sandbox projects. The same Article specifies that derogations may occur regarding the universal service, the duties of the grid operators, and its use, to be determined on a case-by-case basis by DETEC, which also deals with costs.

Regulatory sandboxes often do not fit neatly into one category but fall into a grey area. This is also the case for legal flexibility. For example, in the Portuguese ZLTs, the legal basis (Art. 4 Decree-Law No. 67/2021) allows for the creations of ZLTs that do not entail a derogation from the existing legal framework, in which case it can be created by order of the members of the Government responsible for the areas of economy, science and the field of activity in which the ZLT is located. By contrast, if the ZLT does entail a derogation from the existing legal framework, it must be created by a legislative act, after a prior hearing of the competent regulatory body. Quite peculiar is the case of the Italian experimentations in the energy sector, where exceptions to rules and constraints are allowed, but only for those that fall within the competence of the Italian regulator ARERA. This is because the regulatory sandboxes are activated directly by ARERA without a legally defined sandbox scheme.

4.4. – Discretionary power and constraints

The public authority managing the regulatory sandbox has a certain margin of discretion in order to define its regulation. The legal basis itself may directly indicate the rules that are to be applied, thus binding the public authorities and reducing their margin of discretion.

In the Austrian FMA Sandbox, the FMA does not have wide discretion to decide on a participant in the regulatory sandbox, as the parameters for admission are defined by law. Another example is the Spanish AI Regulatory Sandbox, where the constitutive Royal decree sets out a long list of criteria for the evaluation of applications (Art. 8, par. 2) that the government authorities must consider, and a set of requirements to be assessed during the development of the tests (Art. 11). Similarly in the cases of the Italian Financial

Services Regulatory Sandbox and Experimentation Italy, where the competent bodies (see Section 3.1 for the Fintech Committee) decide on the basis of very detailed regulations.

On the contrary, the public authority responsible for managing the scheme may not be subject to rigid constraints and enjoys a wide margin of manoeuvre. This is the case of the Swiss energy regulatory sandbox, where the competent authorities (DETEC and SFOE) decide on the basis of their own acts, which supplement the scarce legal framework (see above, Section 3.2). Similarly, in the case of the Italian experimentations in the energy sector (see above, Section 3.3).

4.5. – Consequences and benefits

Finally, one of the aspects most affected by the legal basis concerns the consequences for participants once the trial period has been successfully completed.

Usually, a positive evaluation does not automatically imply a certificate of compliance with the regulation. However, participants: a) enjoy the development of a product or business model in a controlled environment; b) benefit from the evidence of experience gained under an authority's supervision, often the same authority that will then be responsible for conducting compliance audits; and c) can gain, if not a presumption, at least an expectation of compliance with regulation, thanks to the ongoing dialogue with the authority and the transition from the close supervision phase.

This is the case with the Austrian FMA Sandbox, where the test phase is evaluated by the FMA and, if successful, the business model leaves the regulatory sandbox and is transferred to regular supervision.

This can be done with existing legislation, as in the case of the Experimentation Italy, where, at the end of the experiment, the applicant submits a final report on the results and the economic and social benefits generated. The government authority then certifies the success of the initiative and gives an opinion to the President of the Council of Ministers and the relevant Minister on the desirability of legislative changes. The effectiveness and attractiveness of the experimentation mechanism is also ensured by a strict timeframe in the start-up phase, which provides certainty about the timing of the whole process.

But this can also happen with future regulations, as in the case of the Spanish AI Regulatory Sandbox, where participants can test how to implement the requirements

applicable to high-risk AI systems before the AI Act becomes applicable. The aim is to provide evidence-based guidance and experimentation to help companies adapt to the new Regulation.

5. – KEY FINDINGS

The development of regulatory sandboxes lacks standardized models and definitions, necessitating careful consideration of theoretical and practical implications when designing them. A key challenge lies in selecting the appropriate regulatory instrument to establish their legal basis. Policymakers must balance innovation-friendly approaches with rule-of-law principles, ensuring that public authority powers remain grounded in democratic legitimacy while safeguarding individual rights.

The legal basis must consider the regulatory sandbox's purpose, whether it focuses on regulatory testing, requiring alignment with existing legislation, or product testing, which demands specific technical rules. Additionally, the scope influences its legal framework: narrower contexts face constraints from sector-specific regulations, while broader frameworks may allow more flexibility, especially in the absence of national laws.

The regulatory sandbox environment also matters. Experiments in real-world settings necessitate legal provisions to protect non-participating individuals' rights, creating a trade-off between regulatory burdens and realistic testing conditions. Authorities managing regulatory sandboxes (whether political, administrative, or independent) require powers and responsibilities defined in the legal basis. These include granting experimental freedoms, particularly in cases of derogation from existing regulations, where the legal basis must match or exceed the legal status of the relaxed rules.

Finally, the legal basis must limit discretionary authority to uphold legality and avoid discrimination while preserving flexibility to enable innovation. It must also define the outcomes of experimentation, ensuring benefits for participants and informing policymakers to improve legislation and public policies, thereby enhancing the sandbox's overall appeal and effectiveness.

TORT LIABILITY AND REGULATORY SANDBOXES

GIOVANNI MARIA RICCIO*

SUMMARY

ABSTRACT

In the legal framework governing regulatory sandboxes, the AI Act devotes only a single paragraph to non-contractual liability, stating that European and national rules apply to providers. However, this choice opens potential scenarios for diverging applications by national courts on the liability of the conducts within the sandbox, which should be considered in light of the different legal models of the single Member States, and which could jeopardize the harmonization process. Furthermore, in light of the applicable regulations, consideration should also be given to the proposal to amend the Product Liability Directive, which includes AI systems within its scope. This proposal leans toward a strict liability model or a framework where, through the reversal of the burden of proof, fault plays a limited role, as seen in the proposed AI Liability Directive. The present paper suggests that, in the case of regulatory sandboxes, fault-based liability should apply during the experimental phase in order to not to discourage a developing market and to allow newcos without significant financial investments to access the experimental phase.

1. – TORT LIABILITY IN THE CONTEXT OF REGULATORY SANDBOXES

One of the most intriguing aspects introduced by the Artificial Intelligence Act ('AI Act') is arguably the establishment of regulatory sandboxes.

Therefore, regulatory sandboxes represent an innovative approach to balancing innovation and regulation, particularly in the field of artificial intelligence, fostering

* Full Professor of Private Comparative Law at University of Salerno. E-mail: gmriccio@unisa.it.

technical development while ensuring that new technologies adhere to regulatory standards. More specifically, these tools allow for the testing of new technologies before they are brought to the market, in order to monitor both their potential risks (i.e., their propensity to cause damage to third parties) and their compliance with the fundamental principles outlined in the AI Act (Simoncini, 2023, 1).

However, this is not an absolute novelty, considering that we are dealing with a legal model that has been legislatively introduced in sectors other than artificial intelligence, albeit in different legal systems. Some elements, however, are common across the various national experiences.

Indeed, according to a study commissioned by the German government, conducted on regulatory sandboxes implemented in various sectors including energy, transport, and logistics infrastructure, these solutions share the following characteristics: a) they are experimental areas created for a limited period, focused on a specific sector, where innovative technologies and business models can be tested and made available to the public; b) they rely on regulatory flexibility, not imposing administrative penalties for non-compliance with existing regulations; c) and they allow regulators (legislative power and administrative authorities) to gain knowledge for developing future regulations and public policies (BMWi, 2021).

However, it would be reductive to limit the discussion to these aspects. AI systems, especially high-risk ones, can be tested under regulatory supervision and through a constant dialogue between the industry and the competent oversight and regulatory authorities. This is an important issue, considering the risk of nullifying often substantial investments and developing a product or service that may not meet regulatory standards (Moraes, 2023).

This highlights that, in the case of the AI Act and other forthcoming regulations permitting the use of controlled regulatory spaces, there is a shift from a regulatory approach centered on ex-post control by administrative authorities and national courts to a model emphasizing pre-market cooperation among various stakeholders. This shift aims to address regulatory concerns before the launch of a product, service, or new technology. An approach which reminds that which is applied by the GDPR (General Data Protection Regulation of 2016) regarding data protection (Bagni, 2023, 206).

This opens the field to a potentially broader discussion which, for obvious reasons

of brevity, can only be briefly mentioned in this contribution. Reference is made to the extensive, often excessive, discretion granted to independent administrative authorities: we should take into account the recent events in the field of data protection, where the arbitrariness of the national data protection authorities, in assessing regulatory compliance, is raising uncertainty among private and public operators (Riccio, 2024, 17). For example, it is worth noting that some national courts (e.g., in Italy and Poland) have invalidated sanctions imposed by the national Data Protection Authorities because the provisions on which these sanctions were based were too generic, including with reference to the principles of privacy by design and privacy by default.

Moreover, such an attitude is favored and incentivized by the vagueness of regulatory formulations which, in several sectors, not least in the AI Act, generate instability and complexity in determining investments (Mantelero, 2020, 1).

In this context, the effort to reduce uncertainty for businesses by establishing standards and protocols for demonstrating compliance with imposed obligations is certainly commendable, albeit with certain limitations that will be addressed. As correctly noted, regulatory sandboxes can also be a useful tool for public authorities to better understand the solutions adopted by regulated entities, following the technological development of such solutions step-by-step and allowing an in-vivo rather than in-vitro understanding of the services and products over which these authorities exercise control (Parenti, 2020, 24).

Regarding this aspect, it is also necessary to specify that many of the issues addressed in this contribution should be specified and (hopefully) clarified in the next future by the European Commission in the Delegated and Implementing Acts pursuant to Article 97 of the AI Act, which grants the Commission a period, expected to be five years from the entry into force of the regulation, for further regulatory interventions, which should affect not only regulatory experimentation spaces but also protocols and guidelines.

As mentioned, at the present moment, one of the major concerns associated with regulatory sandboxes relates to the potential discrepancy between determinations made by national supervisory authorities, which could significantly affect the process of European legal harmonization. Moreover, the discretionary power granted to national authorities could, at least in theory, also result in a risk of 'forum shopping' (if we may say so): some companies, especially those with a transnational expansion, could choose to operate in

countries whose authorities, to facilitate the national economy, are more permissive and open to particularly daring and potentially dangerous technological solutions (Truby - Brown - Caudevilla Parellada, 2022, 278).

The EU legislator seems to be aware of this risk, which is why, in Article 58 of the AI Act, to avoid legislative fragmentation within the Union, it has provided that the Commission may adopt implementing acts, pursuant to Article 291 TFEU, that can specify 'the detailed arrangements for the establishment, development, implementation, operation and supervision of the AI regulatory sandboxes' (Bagni - Seferi, 2024).

2. – NATIONAL INFLUENCES ON AI LIABILITY

An aspect of undeniable interest pertains to civil liabilities associated with the regulatory experimentation phase. As highlighted by Recital 138 of the AI Act, the rapid development of AI technologies requires 'responsible innovation and integration of appropriate safeguards and risk mitigation measures'.

Article 57(12) of the AI Act holds that providers are liable under national and Union law 'for any damage inflicted on third parties as a result of the experimentation taking place in the sandbox'. Therefore, the provision states that these parties are liable for any injuries according to the ordinary rules of tort liability. However, it would be preferable to expand the scope of the provision by including not only damages after the experimentation, as a literal interpretation would suggest, but also damage occurring during such phase.

The provision included in the AI Act, however, is laconic, as it refers to EU and national tort law rules. It is well known that, in the single legal systems of the European Union, the definitions of liability and damage are different, especially for the interpretations provided through the years by the courts.

National legal traditions could significantly influence judges, jeopardizing the harmonization of the AI liability rules in the member States, due to different case-law trends in the single legal systems. In other words, the highlighted risk is that, without full European harmonization of the rules of extracontractual liability, discrepancies could arise at the national level.

The first key point of the debate concerns the nature of liability. In this sense, as

noted, 'the modern evolution of tort law shows a long-term tendency to distance tort law from the stance that moral or religious wrongdoing is, by itself, enough to establish civil liability' (Koziol – Steininger, 2016).

Over time, the European legal system has undergone a shift from a 'personal' to a more 'functional/utilitarian' notion of liability. This evolution has progressively moved away from fault-based rules (i.e., unintentional breaches of duty of care) and toward the adoption of strict liability models aimed at ensuring compensation for victims of wrongful conduct. If indeed personal responsibility remains a key concept in the discourse over the structure of tort law, still 'the advent of vicarious liability, strict liability and the diffusion of no-fault, collective compensation schemes (...) have surely cast doubts on the meaning of the notion' (Koziol - Steininger, 2016).

Inevitably, describing this evolution requires a quick analysis of three main legal systems: (i) the French open-list model; (ii) the German closed-list model; (iii) the UK rules governing tort liability.

Under the first model, adopted for the first time by Article 1382 of the French Civil Code of 1804, 'Tout fait quelconque de l'homme, qui cause à autrui un dommage, oblige celui per la faut duquel il est arrivé, à le réparer' (Any act of a person that causes damage to another obliges the one through whose fault it occurred to make reparation) (Viney, 1982, 10).

In other words, a liability claim can be brought, subject to the proof of (i) the damage, (ii) the causal relationship and (iii) the fault of the defendant. A similar model was adopted in Italy where, in addition to the three mentioned elements, further evidence concerning the 'injustice' of the damage also had to be submitted (Article 2043 of the Italian Civil Code of 1942).

Separately, according to the closed list system established in Germany under the BGB, protection had to be granted in a fixed number of specific circumstances, expressly listed by the German codification (Article 823): 'A person who, intentionally or negligently, unlawfully injures the life, body, health, freedom, property or another right of another person is liable to make compensation to the other party for the damage arising from this'. As for the elements to be proved, evidence also had to be given in relation to the (allegedly) 'unlawful' nature of the act.

A totally different approach is historically adopted in the UK, where the original

system for regulating liability was based on the forms of action. Thus, in this legal system, liability was not confined to a single form but encompassed several possible actions (torts).

However, this hybrid scenario was not meant to remain unchanged. And in fact, over time, the apparently strict borders among the three national options gradually blurred in legal interpretation. More in detail, if on the one side under the French/Italian model, the notion of liability and recoverable damage was extended - and so it happened in the UK with the introduction of the tort of negligence -, on the other side the German closed-list model was subject to a broader judicial interpretation, aimed at compensating the damage of any third party right (Gordley, 2015, 173).

The described evolution, apart from reducing the distance among the three systems, also contributed, in general terms, to the extension of the notion of recoverable damages and, specifically, to the adoption of a strict liability approach (Zweigert - Kotz, 1998).

Similarly, there is no EU legal framework for tort liability, despite the attempts made, through the last decades, by several research projects. Furthermore, it is not possible to reconstruct this legal framework from individual decisions of the Court of Justice (Vaquer, 2008, 30), although some guidance has been provided over the years in the definition of damage. The issue of harmonizing extracontractual liability rules (and civil liability in general) at the EU level has been debated for a long time.

In summary, while the lack of harmonization in these rules could undermine efforts to unify the single market, it may also necessitate amendments to national criminal law regulations affected by civil liability frameworks (van Dam, 2013, 20).

The problem is not solely with the formal rules (i.e., the text of the law) but primarily with the application of these rules by the courts, as the interpretation could vary critically because judges are inevitably influenced by their university education (which precedes any harmonization rules) and by jurisprudential interpretations, which also precede any harmonization rules (Resta, 2024; Banakas, 2002, 179).

Few areas, like extracontractual liability, have undergone such radical transformations not through changes to written rules but through their interpretation by the courts.

3. – FAULT OR STRICT LIABILITY FOR AI?

Furthermore, Article 57(12) of the AI Act must necessarily be integrated with the

proposed 'AI Liability Directive' (2022/303/COD), whose objective is to harmonize national regulations, in light of the observation that said regulations, especially concerning fault, are inadequate for addressing liability claims related to damages caused by AI-based products and services as well as to the 'New Product Liability Directive' (new PLD) (2022/495/COD), on which the Parliament and the Council reached a provisional agreement on 14th December 2023.

While the first proposal is still waiting for its approval and seems to be parked in the EU offices (Novelli - Casolari - Hacker - Spedicato - Floridi, 2024), the new PLD is very close to its final version. In fact, the COREPER confirmed the agreement at the Council on January 24, 2024, and the Parliament officially endorsed the text during its March 2024 Plenary session. The directive now awaits formal approval by the Council, and the new rules should take effect on products placed on the market 24 months after the directive comes into force.

However, the possibility of applying this directive to the regulatory sandboxes is precluded by Article 6 of the new PLD which states that a product, including an AI system, can be 'considered defective when it does not provide the safety which the public at large is entitled to expect'. The same article sets a non-exhaustive list of circumstances which should be taken into account in considering the defectiveness of a product, chief among them being: the presentation of the product, and its instructions for installation, use and maintenance; the reasonably foreseeable use and misuse of the product; the effect on the product of any ability to continue to learn after deployment; the moment in time when the product was placed on the market or put into service; product safety requirements (Spindler, 2023).

In other words, the PLD, where liability rules are essentially based on the defectiveness, is applicable exclusively to products and services which are put on the market, and then it excludes the experimental phase of the regulatory sandboxes.

A different approach could be considered for the latter directive, which is still under the scrutiny of EU institutions.

The goal of the proposed AI Liability Directive is not to harmonize extracontractual liability systems, but solely to establish certain principles regarding the burden of proof. The new regulation, in fact, should apply to all AI systems, mandating disclosure obligations for evidence related to high-risk AI systems, to avoid the risk of a black-

box effect, i.e., to prevent the demonstration of the potential tortfeasor's conduct from becoming a probatio diabolica for the injured party. Before AI systems are placed on the market, there is, on the one hand, the interaction of various operators, significantly affecting the demonstration of the causal link between the tortfeasor's conduct, event, and damage, and, on the other hand, the opacity, autonomous behavior, and complexity that can make it excessively difficult, if not impossible, for the injured party to satisfy the burden of proof.

Specifically, Article 3 of the proposed AI Liability Directive provides for a mechanism that, through judicial authority intervention, can redress the informational asymmetries existing between users and operators by requiring a provider to disclose 'relevant evidence at its disposal about a specific high-risk AI system that is suspected of having caused damage'. Such a judicial order can be granted in case of a refusal by the system operators to disclose such information, provided there are 'facts and evidence sufficient to support the plausibility of a claim for damages' (a sort of fumus boni juris or prima facie evidence) and that the plaintiff 'has undertaken all proportionate attempts at gathering the relevant evidence from the defendant' (paragraphs 1 and 2, Article 3).

The liability exemption set forth in Article 57(12) of the AI Act, on the other hand, exclusively concerns potential administrative penalties, which cannot be imposed by the supervisory authority, when providers 'observe the specific 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 authorities for infringements' of the obligations set by the AI Act. After all, considering that regulatory sandboxes are based on direct communication and interaction between the supervisory authorities and providers, it would still be unlikely to imagine a sanction from such authorities, except in cases of willful misconduct by the provider during the experimentation, subsequently verified after the commission of the related conduct. From a regulatory policy perspective, it is first necessary to determine the purpose assigned to extracontractual liability within the context of regulatory sandboxes.

In summary, the choice is between favoring the injured parties or the companies involved in this process. In the former case, it is evident that a criterion of strict liability would be preferable (as in the new PLD); in the latter, a fault-based criterion would need to be adopted.

Apparently, the AI Act seems to prefer a fault-based criterion. A possible interpretation may be the following: the party participating in the experimentation acts improperly and should therefore be liable for such conduct i.e. it shall be expected to compensate the damage for its wrongful conduct and thus according to a fault liability standard. In our opinion, the rationale behind this choice is related to the fact that, in the case of the PLD, products are already placed on the market, whereas regulatory sandboxes deal with experimental phases, and therefore there should be more tolerance for the possibility of the tested services causing harm to third parties.

However, tort law scholars are aware that the boundaries between the fault and strict liability models are blurred, especially when higher standards of diligence are imposed (Büyüksagis - van Boom, 2013, 609). That is the reason why, in this scenario, it will be crucial to establish clear standards and protocols, creating almost an automaticity in liability and leaving minimal room for discretion to both national courts and national administrative authorities (the latter, of course, would not be competent in matters of liability).

A fault-based liability could also have a positive impact on competitive dynamics. In fact, the application of strict liability would mean that only economically solid companies (the so-called deep pocket parties) would be willing to participate in regulatory sandboxes, considering the costs of accidents. This approach could impact on the eligibility of start-ups and new companies, which might not have sufficient financial resources to cover potential damage.

4. – CONCLUSIONS

It is difficult to draw conclusions in the presence of three regulations that are not yet fully effective. In fact, only the AI Act has been approved, although it will not be fully enforceable until 2026 and will be supplemented by implementing and delegated acts, while the two proposed directives still need to complete their legislative process. However, the alternation between two antithetical liability models, the first based on fault (in the case of regulatory sandboxes), the second on strict liability (or, at least, the reversal of the burden of proof), has a rational explanation when considering the policy underlying the choices of the European legislative bodies.

Civil liability serves functions (compensation for damages and prevention of unlawful acts) that significantly influence economic development. Therefore, in the case of products already on the market, strict liability is chosen with the aim of fully protecting individuals who could be harmed by devices whose potential for harm (or exposure to danger) appears high (for example, autonomous vehicles).

In the case of regulatory sandboxes, however, the liability based on fault should be applied. It should not be forgotten, in fact, that strict liability would allow only the economically stronger companies to enter and remain in the market, also favoring the creation of oligopolistic or monopolistic positions.

Therefore, there is no legislative schizophrenia on the part of EU law in applying different liability criteria to different situations and moments. In fact, in the case of the proposed AI Liability Directive, the product is examined when it is already on the market, after a testing phase. For regulatory sandboxes, however, liability applies at an earlier stage, when the products are still in the experimental phase, and companies do not yet take advantage of any economic benefit from their commercialization (and so the old Latin maxim cuius commoda eius et incommoda does not apply).

From this perspective, an interpretation that leads to considering it a case of fault-based liability, rather than strict liability, seems preferable when it comes to civil liability in regulatory sandboxes. In fact, during the development phase of AI products (and thus within the regulatory sandboxes), it is important to ensure a safe harbor for developers, and that the prospect of compensating damages, to be paid regardless of unlawful conduct, does not hinder investments and, above all, the testing of innovative products, whose development could benefit the European economy.

REGULATORY SANDBOXES AS A BRIDGE BETWEEN AI AND CYBERSECURITY: EXPLORING THE INTERPLAY BETWEEN THE AI ACT AND THE CYBER RESILIENCE ACT

FILIPPO BAGNI*

SUMMARY

ABSTRACT

The contribution examines the role of regulatory sandboxes in the context of the Artificial Intelligence Act (AI Act) and the Cyber Resilience Act (CRA), highlighting their key features, objectives, and potential benefits. Through a comparative analysis, the paper explores the interactions between the two pieces of legislation, with a particular focus on cybersecurity requirements for AI systems. It argues that regulatory sandboxes can facilitate dialogue and coordination between the AI Act and the CRA, ultimately improving regulatory compliance. The analysis has a dual scope: to identify similarities and differences between the two regulations, and to highlight the critical role of cybersecurity in the context of AI systems and regulatory sandboxes. The findings suggest that regulatory sandboxes have the

* PhD Candidate at IMT School for Advanced Studies Lucca and Legal officer at European Commission (DG Connect). Contact email: filippo.bagni@imtlucca.it. The information and views set out in this article belong to the author and do not necessarily reflect the official opinion of the European Commission.

potential to play a crucial role in promoting a safe, fair, and healthy digital ecosystem in Europe. The contribution highlights the importance of dedicated cyber resilience sandboxes and proposes the development of a comprehensive framework of regulatory sandboxes for AI and cybersecurity, which could foster innovation and experimentation at both European and national levels.

1. – INTRODUCTION

The challenges posed by technological transformation and the emergence of new products and services have brought about new regulatory complexities (Weimer-Marin 2016, 469; European Commission 2023, 131). The flexibility of technological progress has tested the capabilities of lawmakers and their inherent regulatory rigidity (Bennett-Moses 2013). Consequently, new regulatory approaches have been developed, including the concept of specific regulatory experimentation spaces, known as 'regulatory sandboxes' (van Gestel-van Dick 2011; Ranchordas 2015; Heldeweg 2015; Mousmouti 2018; Attrey-Lesher-Lomax 2020).

The contribution analyses the role of regulatory sandboxes under the newly introduced Artificial Intelligence Act ('AI Act', Regulation (EU) No 2024/1689) and the Cyber Resilience Act ('CRA', Regulation (EU) No 2024/2847). It highlights their growing significance as hybrid governance tools for emerging digital technologies, focusing on their impact at the intersection of artificial intelligence (AI) and cybersecurity. The research underscores the potential of regulatory sandboxes to shape the future of AI and cybersecurity in Europe.

This research is divided into four distinct sections. The introductory section provides an in-depth examination of the regulatory sandbox instrument in a broader context, with a particular focus on the 'AI regulatory sandboxes' framework as outlined in the AI Act. The following section provides a comprehensive examination of the CRA Regulation, focusing on the specific provisions envisaged for the 'cyber resilience regulatory sandboxes'. The third section attempts to unpack the multiple interactions between the AI Act and the CRA Regulation, together with a reflective exploration of the potential role of regulatory sandboxes as a common ground for dialogue between the

two legislative frameworks. Finally, the fourth section is dedicated to drawing conclusions and proposing a structured European framework for a harmonized European regulatory sandbox ecosystem embracing both the AI and cybersecurity domains, in order to facilitate seamless interaction and dialogue between these interrelated areas.

2. – REGULATORY SANDBOXES AND THE AI ACT: KEY INSIGHTS

2.1. – Regulatory sandbox national case studies

The notion of regulatory sandboxes is not new in Europe, particularly in highly technical and regulated sectors such as banking, insurance, energy, and data protection (Ranchordas 2021; Ranchordas 2021a).

The fintech space was one of the first areas to adopt sandbox experimentation, given its high level of technicality and sector-specific regulatory oversight (Omarova 2020; Allen 2019; Zetzsche et al. 2020, 55). A notable example is the Bank of Italy regulatory sandbox, which was introduced through explicit legislative provisions (Decreto Legge n. 34/2019) to facilitate dialogue between the competent authority and supervised banks[1]. Notably, Bank of Italy has adopted a comprehensive experimentation scheme for the Fintech sector based on three pillars: the 'Fintech Channel', which consists of an Innovation Hub established in 2017 as regulatory support; the 'Milan Hub', introduced in 2020 as a place for research initiatives, specifically focused on the project development phase of innovative products; and finally, the regulatory sandbox, introduced in 2021. What makes this experience unique is the Bank of Italy's engagement with companies since the early stages of idea development, project implementation and testing of fintech products and services.

Another area of experimentation is the processing of personal data, with national data protection authorities playing a leading role (Malgieri 2019). The United Kingdom (UK) and Norway have developed notable sandboxes in this area. The UK's sandbox, set up by the Information Commissioner's Office (ICO), explores new technologies such as voice biometrics and facial recognition, and provides free support to companies on risk mitigation and data protection integration[2]. Another example is Norway's sandbox,

[1] See https://www.bancaditalia.it/focus/sandbox/index.html.

[2] See https://ico.org.uk/for-organisations/advice-and-services/regulatory-sandbox/the-guide-to-the-sandbox/.

developed by the Norwegian Data Protection Authority, focused on the intersection of privacy and artificial intelligence[3]. This tool is open to public and private companies developing AI systems with significant privacy implications (Fenwick-Vermeulen-Corrales 2018; Smuha 2021).

The German approach to regulatory sandboxes is also notable for its systematic and coordinated approach. The Federal Ministry of Economics and Technology (BMWi) has developed a comprehensive framework for regulatory sandboxes, providing implementation guidelines and experimental clauses that allow individual states to tailor their own rules and exemptions[4].

The list could continue[5]. Despite the variety of sandboxes in today's landscape, some common characteristics can be identified: a regulatory sandbox typically involves innovative products or services that offer added value to consumers or society, are developed to a stage that allows immediate testing and are economically viable throughout the testing period (Bagni 2023).

In order to ensure legal predictability, it is essential that the applicable legislation, legal barriers, boundaries and conditions of the sandbox are clearly defined and communicated in advance. This includes specifying the relevant legislation and sectors involved, outlining exemptions and derogations, establishing rules for entry and exit, and determining the duration of the sandbox. In addition, it is necessary to implement robust safeguards to mitigate potential risks, even within a controlled environment.

Participation in a regulatory sandbox is typically subject to approval, monitoring, and evaluation by the competent authority, with a limited number of places available. The authority usually issues open calls for interested operators to submit their projects, followed by a selection and interview process leading to the launch of the experimental project.

[3] See https://www.datatilsynet.no/en/regulations-and-tools/sandbox-for-artificial-intelligence/.

[4] See https://www.bmwk.de/Redaktion/EN/Dossier/regulatory-sandboxes.html.

[5] Other relevant national use cases include (not limited to) the Maltese Technology Assurance Sandbox (https://tech.mt/mdia-the-technology-regulator/technology-assurance-sandbox/) sector and the Estonian Digital Product Management Sandbox (https://sandbox.cs.ut.ee/).

2.2. – 'AI regulatory sandboxes' under the AI Act

The technology sector that has recently received the most exponential attention in relation to regulatory sandboxes is undoubtedly the field of AI. The debate surrounding AI regulation has intensified, particularly with the entry into force of the AI Act (1 August 2024), a groundbreaking piece of legislation that subjects AI systems to conformity assessment before they can be placed on the market. This regulatory approach makes AI systems an ideal candidate for testing in a controlled environment, such as a regulatory sandbox.

Notably, the AI Act recognises the importance of regulatory sandboxes at the EU regulatory level by classifying them as 'measures in support of innovation' (Chapter VI) and dedicating a comprehensive set of provisions to this tool (Recitals 138-141; Articles 57-59)[6]. In doing so, the AI Act recognises the 'institutional dignity' of regulatory sandboxes and formalises their role in facilitating innovation and experimentation in the AI sector.

The AI Act has the merit of providing a clear definition of the concept of regulatory sandbox, even if specifically tailored to the AI sector (Article 3(1)(55)). This definition incorporates the common elements of European sandboxes, including a controlled framework, the active role of the supervising competent national authority and the possibility for the prospective provider to develop, train, validate and test its innovative product for a limited period of time.

The innovative aspects of this definition are dual: the introduction of a 'sandbox plan' and the explicit possibility of experimentation under 'real world conditions' (Bagni and Seferi).

In particular, the 'sandbox plan' is an agreement between the participating company and the authority that sets out in advance the objectives, conditions, timetable, and methodology of the experiment. This plan enables the parties involved to structure the modalities of the experiment in a concerted and well-defined manner, and it also plays a role in determining the potential liability of the provider for the activities carried out

[6] At the time of its proposal (April 2021), the AI Act was the first European regulation to introduce the concept of regulatory sandboxes. However, the Interoperable Europe Act (Regulation (EU) No 2024/903), which entered into force earlier (April 2024), became the first European regulation to formally establish regulatory sandboxes.

during the experiments (see Article 3(1)(54) and Article 57(12) of the AI Act).

On the other hand, the possibility to conduct experiments under real world conditions offers advantages such as more accurate test results and a less ambiguous assessment of compliance. However, it also increases the risk of harm to users and third parties, as it involves real interests (e.g. the risks associated with the use of real personal data to train the AI system). As a result, there is a need for greater supervision by the competent authority during experimentation and for appropriate safeguards (see Article 3(1)(53) and Article 58(4) of the AI Act).

The legal framework of the AI Act clearly outlines the main features of the AI regulatory sandbox, as set out in Articles 57 and 58. In particular, Article 57(1) requires each Member State to establish at least one national regulatory sandbox for AI and to ensure that it is fully operational within 24 months of the entry into force of the Regulation (august 2026). The provision also encourages the development of additional sandboxes at local and regional level, suggesting a broader objective of creating a comprehensive European system of regulatory sandboxes for AI.

Article 57(9) explicitly outlines the 5 objectives of the AI regulatory sandbox: 1) enhancing legal certainty, emphasising that participation in the sandbox should focus on issues that create legal uncertainty (recital 139); 2) exchanging of best practices through cooperation between stakeholders; 3) fostering innovation and competitiveness in the internal market; 4) contributing to the 'regulatory learning' of providers of AI systems, giving the sandbox a didactic role that goes beyond mere regulation; 5) facilitating market access for small and medium-sized enterprises (SMEs) and start-ups, highlighting the regulator's awareness of the compliance costs associated with the new digital sector rules for companies operating in the AI sector.

In addition, the Regulation stipulates that the provider's path within the AI sandbox must be thoroughly documented in order to prove the activities carried out and the results achieved. At the end of the experimentation period, in fact, the competent national authority is required to issue two types of documents: a 'written proof' of the activities successfully carried out (optional and at the request of the provider) and an 'exit report' (mandatory) detailing all the activities carried out and the results achieved. These documents are crucial in the context of future conformity assessments, as they can be used by the provider to demonstrate the compliance of the AI system with the AI Act and

other relevant regulations on a case-by-case basis.

The details of the operation of the AI regulatory sandboxes are set out in Article 58 of the Regulation. However, it is worth mentioning that the Commission is actively working on the adoption of an implementing act aimed at specifying the key elements for the establishment, development, implementation, operation, and oversight of AI regulatory sandboxes (Article 58(1) AI Act). The objective of this act is to ensure a consistent implementation across the Union and to guarantee that AI regulatory sandboxes are used in a consistent and effective manner to support innovation and regulatory compliance in the field of AI.

3. – REGULATORY SANDBOXES FOR CYBERSECURITY: AN ANALYSIS OF THE CYBER RESILIENCE ACT

3.1. – Main elements of the Cyber Resilience Act

Like artificial intelligence, cybersecurity has also gained significant importance at the European level in the digital decade[7] program. In alignment with the EU Cybersecurity Strategy Digital Decade, several significant new regulations have been proposed in this field, such as the Cybersecurity Act, the new NIS2 Directive, the Cyber Resilience Act, and the Cyber Solidarity Act. Hence, companies find themselves increasingly confronted with numerous new rules and compliance obligations also in the cybersecurity domain (Chiara 2024).

In this context, the proposed 'Cyber Resilience Act' ('CRA'; Regulation (EU) No 2024/2847), published in its final text on 20 November 2024 and entered into force on 10 December 2024, is of particular importance.

The CRA has been deemed necessary due to the cross-border nature of digital products and the risks of cyber-attacks (Shaffique 2024; Jara et al. 2024). Currently, most hardware and software products lack any uniform legislation ensuring their cybersecurity, and no regulation addresses the cybersecurity of non-embedded software, which represents a critical vulnerability in the era of digital products (Nuthi 2022; Chiara 2022). Therefore, the CRA aims to introduce a horizontal regulatory framework at the European level, establishing comprehensive and uniform cybersecurity requirements for

[7] See https://digital-strategy.ec.europa.eu/en/policies/cybersecurity-strategy.

all 'products with digital elements' (defined in Article 3(1) of the CRA) entering the European internal market. Such products include a wide range of hardware and software, for instance consumer internet-connected devices (e.g. smart toys, smart speakers), operating systems (e.g. for computers, smartphones), and applications (apps, e.g. health-monitoring apps).

The proposal seeks to address two key issues: (a) the low level of cybersecurity of digital products in the European single market, and (b) the inadequate understanding and access to information by users, preventing them from choosing products with adequate cybersecurity properties and/or using them securely.

To address these issues, the proposal takes a two-pronged approach: (i) it requires manufacturers to design and develop their products in compliance with certain objective-oriented and technology-neutral essential requirements set out in the Regulation, and to ensure that these requirements are maintained throughout the life-cycle of the product; and (ii) it empowers businesses and consumers to use products with digital elements with confidence by providing them with the necessary information and tools to do so safely.

Like the AI Act, the CRA imposes specific obligations on economic operators throughout the production chain, including manufacturers, distributors, and importers, regarding the placing on the market of products with digital elements. These obligations are tailored to their respective roles and responsibilities, ensuring a comprehensive approach to product security.

All products with digital elements under the CRA are subject to a conformity assessment procedure, which includes several key steps: conformity assessment, registration of the declaration of conformity, CE marking, and maintenance of technical documentation. Only products that successfully complete this process can be placed on the market, provided that they are properly installed, maintained and used for their intended purpose, thereby ensuring that they are considered 'cyber-safe'.

The mandatory conformity assessment under the CRA adopts a risk-based approach, considering the level of criticality of the product. All digital products must meet certain essential requirements and undergo self-assessment, while products classified as 'important' (Article 7) or 'critical' (Article 8) are subject to a more stringent conformity assessment, involving either a self-assessment using harmonized cybersecurity standards (for Class I products such as operating systems) or an independent assessment (for Class

II products such as firewalls, intrusion detection and prevention systems). Unlike the AI Act, which only requires conformity assessment for high-risk products, the CRA applies this requirement to all products.

3.2. – 'Cyber resilience regulatory sandboxes' under the CRA

During the CRA negotiations, the legislator's position on regulatory sandboxes evolved significantly. Initially, the CRA proposal did not contain any reference to sandboxes, which raised concerns, particularly in the European Parliament[8]. However, the call for the introduction of sandboxes was successful and the final text of the CRA explicitly provides for 'cyber resilience regulatory sandboxes'.

Contrary to the AI Act, the CRA does not provide a formal definition of these sandboxes. However, it is possible to infer a definition from certain provisions of the Regulation, in particular Article 33. According to this article, 'cyber resilience regulatory sandboxes' refer to controlled testing environments established by Member States for the development, design, validation and testing of innovative products with digital elements for a limited period of time before their placing on the market, in order to facilitate compliance with the CRA.

This definition confirms the main elements common to AI regulatory sandboxes: controlled framework, validation, and testing activity for a limited period; focus on innovative products; oversight by a public authority and facilitated access for small and medium-sized enterprises and start-ups. In addition, the objectives of the CRA are the traditional ones associated with regulatory sandboxes: promoting innovation and competitiveness and improving legal certainty (Recital 97). However, unlike the AI Act, there is no reference to real-world testing and the sandbox plan.

The discipline associated with cyber resilience sandboxes under the CRA is rather limited. In fact, the CRA only devotes paragraph 2 of Article 33, entitled 'Support

---

[8] During the negotiations, the Parliament made its first comments on their inclusion. In particular, the text presented by the ITRE Parliamentary Committee (May 2023) proposed a new recital (69a) and a new article (49a) encouraging the Commission, the European Union Agency for Cybersecurity (ENISA) and Member States to establish 'European cyber resilience regulatory sandboxes'. This initial text will be followed by a formal report (July 2023), confirming the Parliament's willingness to invest in regulatory sandbox tools in the area of cybersecurity.

measures for microenterprises and small and medium-sized enterprises, including start-ups', to this issue. Four key elements emerge from the analysis of this provision.

First, the establishment of cyber resilience sandboxes at the national level by Member States is not mandatory but optional ('where appropriate'). This may be due to regulators' caution in investing in a sensitive area such as cybersecurity with a hybrid and innovative regulatory tool like regulatory sandboxes. Moreover, the broader scope of the CRA (all products with digital elements) may have led the regulator to avoid requiring a mandatory tool for such a broad category of products.

Second, sandboxes will only be established 'for the purpose of complying with this Regulation'. This suggests that, despite the broader cybersecurity regulatory landscape - including regulations such as the NIS2 Directive, the Cyber Solidarity Act and the Cybersecurity Act - the use of these sandboxes is specifically limited to ensuring compliance with the CRA. The term 'cyber resilience sandboxes' reinforces this narrow focus.

Third, there is a provision for optional coordination ('where appropriate') between the Commission and the European Union Agency for Cybersecurity (ENISA) to provide technical support to these national sandboxes. This highlights the high level of specialisation required for compliance in the cybersecurity sector, and the need for coordinated expert support at European level.

Finally, unlike the AI Act, the CRA omits any provision for the issuance of documentation at the end of the experimentation phase, a crucial aspect for participating companies seeking to demonstrate future compliance. This omission could potentially reduce the attractiveness of engaging in these sandboxes.

4. – THE INTERSECTION OF AI AND CYBERSECURITY: EXPLORING THE SYNERGIES BETWEEN THE AI ACT AND CRA

4.1. – Cybersecurity requirements for AI systems

There are clear similarities between the CRA and the AI Act. Both proposals (i) aim to ensure the safety and reliability of digital technologies in the internal market; (ii) impose compliance requirements and obligations on companies developing digital products through a risk-based approach; (iii) require special attention to the protection of personal

data; (iv) seek to enhance consumer confidence in the use of digital technologies; and (v) devote particular attention to SMEs and their compliance costs.

Cybersecurity is a fundamental pillar of the AI Act and is repeatedly mentioned as a guarantee for the safety and reliability of AI systems. Specifically, it is mentioned in three key areas: (1) risk assessment and threat management for AI systems, (2) implementation of security measures to protect data and information processed by AI systems, and (3) compliance with security standards to minimise the risk of cyberattacks by third parties (as stated in recital 76 of the AI Act). To this end, the legislation explicitly requires providers to implement robust security controls, data poisoning prevention, and other measures to prevent data breaches and hostile attacks.

Notably, most of the references to cybersecurity in the AI Act relate to areas that are considered the riskiest, such as general-purpose AI (GPAI) models with systemic risks and high-risk AI systems (HRAIs). This highlights the importance of ensuring that systems and models are cyber resilient.

With respect to GPAI models with systemic risks, Article 55 of the AI Act requires providers to take measures to ensure the security of the model from both a software and hardware perspective, ensuring an 'adequate level of cybersecurity protection' for the model itself and the security of the model's physical infrastructure.

On the other hand, HRAIs are an area where the two regimes most clearly overlap, particularly where a HRAI under the AI Act is also considered a product with digital elements under the CRA.

The AI Act clearly imposes specific cybersecurity requirements on HRAIs providers. In particular, Article 15 of the AI Act, entitled 'Accuracy, Robustness and Cybersecurity', explicitly states that HRAIs providers must design and develop their products to achieve an 'appropriate level of cybersecurity' throughout their lifecycle, and that technical solutions to ensure the cybersecurity of HRAIs must be 'adequate to the circumstances and relevant risks' (Novelli et al. 2024; Nolte et al. 2024).

In addition, with respect to HRAIs, Article 11(1) requires that the technical documentation demonstrating compliance with the Regulation include a section describing in detail the 'cybersecurity measures adopted' to meet the above requirements (Annex IV, point 2(h)).

In this specific context of the cybersecurity requirements under the AI Act, a general

rule applies, namely the presumption of conformity: if an HRAI system falls within the scope of the CRA and fulfils its cybersecurity requirements, a presumption of conformity applies to the cybersecurity requirements for HRAIs under the AI Act (e.g. resilience against unauthorised use by third parties). This principle is clearly expressed in recitals 77 and 78 of the AI Act[9] and is also mentioned by the CRA in recital 51 and Article 12[10].

The interaction between the two regulations is therefore clear, but not complete. On the one hand, if an HRAI is also considered a product with digital elements under the CRA, the conformity assessment procedure under Article 43 of the AI Act also applies to the CRA. In this case, the interaction is total, as an act provided for in the AI Act is directly relevant in the conformity structure of the CRA. On the other hand, if the product is considered 'important' or 'critical' under the CRA, the conformity assessment procedure provided for in the Regulation is not replaced by that of the AI Act. The reason for this is that in this second hypothesis, the risk-based approach of the two regulations is no longer considered to be fully aligned, and the conformity assessment of the CRA regains its autonomy, without there being a total overlap between the two disciplines in terms of conformity requirements.

In this complex framework, cooperation between the market surveillance authorities designated under the AI Act and the CRA is essential to ensure compliance with both regulations. Not surprisingly, Article 41(10) of the CRA explicitly provides that the market surveillance authorities designated under the AI Act are also responsible for compliance with the CRA for products with digital elements classified as HRAIs.

Finally, there are also clear signs of interaction in the governance aspect of the AI Act. Indeed, Article 66 of the AI Act provides that, among the various tasks of the Board, there shall be cooperation with all European institutions and relevant organisations

---

[9] The cybersecurity requirements of the AI Act will be met if the 'essential cybersecurity requirements set out in that regulation', i.e. the CRA, are met. Furthermore, the principle of presumption of compliance in the following terms: 'When high-risk AI systems fulfil the essential requirements of [CRA], they should be deemed compliant with the cybersecurity requirements set out in this Regulation [AI Act].'

[10] 'Products with digital elements classified as high-risk AI systems [...] which fall within the scope of this Regulation should comply with the essential cybersecurity requirements set out in this Regulation'. 'Where those high-risk AI systems fulfil the essential cybersecurity requirements set out in this Regulation [CRA], they should be deemed to comply with the cybersecurity requirements set out in Article 15 of Regulation (EU) 2024/1689 [AI Act]'.

in the field of cybersecurity. Article 70, on the other hand, requires Member States to designate national competent authorities under the AI Act with specific competences in various areas, including cybersecurity (paragraph 3), and reiterates that these national authorities must take appropriate measures to ensure an adequate level of cybersecurity (paragraph 4). These provisions also explain why some Member States, such as Italy[11], are considering including national cybersecurity authorities among the subjects responsible for implementing the AI Act.

4.2. – Regulatory sandboxes as a common ground for AI Act and CRA implementation

The close link between the AI Act and CRA is particularly evident mostly in the area of regulatory sandboxes. Both regulations enable the use of this experimental tool, and the CRA explicitly acknowledges this link by stating in Article 12(4) that manufacturers of products with digital elements that are classified as HRAIs under the AI Act 'may participate in the AI regulatory sandboxes'.

This crucial provision not only demonstrates the close connection between AI and cybersecurity in the specific case of HRAIs, but also tells us something more: AI regulatory sandboxes, which are mandatory at the national level, can serve as a direct link between the AI Act and the CRA in terms of conformity assessment of products. In this context, sandboxes play an important role in facilitating regulatory coordination, enabling effective communication and cooperation between regulators and companies developing AI technologies, and ensuring compliance with both sets of rules.

This is confirmed also by the text of the AI Act, which stipulates in Article 58(2)(i) that the future implementing act enable AI regulatory sandboxes to facilitate the development of tools and infrastructure for evaluating AI systems, specifically in areas such as accuracy, robustness, and cybersecurity, to support regulatory learning. This reiterates the importance of regulatory learning, and the establishment of AI regulatory sandboxes aims to achieve this outcome as a key objective.

In this way, thanks to regulatory sandboxes, it will be possible to reduce the risks

---

[11] See Article 18 of the draft law (DDL - 20 May 2024) in which Italy proposes the ACN (Autorita' Nazionale per la Cybersecurity) as the national authority for artificial intelligence. The text of the draft law is available at www.senato.it/.

and uncertainties associated with the development and use of AI technologies and to promote trust and security in the cybersecurity market. At the same time, companies will benefit from a better understanding of the rules and how they interact, which will also help the authorities involved to identify potential gaps or uncertainties in a complex regulation that needs to be future proof.

In summary, the close link between the AI Act and the CRA in the context of regulatory sandboxes highlights the importance of a coordinated approach to regulatory innovation. By aligning and harmonising different regulatory frameworks, a level playing field for all stakeholders can be promoted. The effective implementation of regulatory sandboxes will depend on the ability of regulators to work with industry stakeholders to create a supportive ecosystem for innovation, while ensuring the necessary safeguards to protect the public interest.

Regulatory sandboxes can be seen as a critical component of a broader regulatory innovation ecosystem. By fostering a culture of experimentation and collaboration, sandboxes can help promote a more dynamic and responsive regulatory environment, where rules can be adapted and updated in response to changing technological and societal needs. Ultimately, this can help ensure that the benefits of AI are realised in a way that is safe, secure, and beneficial for all members of society.

5. – CONCLUSIONS

The AI Act and the CRA share a common objective: to ensure a secure European internal market through the regulation of technology. Both legislative initiatives focus on the regulation of products, in particular AI systems and products with digital elements, making regulatory sandboxes a valuable tool for companies and authorities to work together and engage in continuous dialogue. This collaboration will facilitate the development of innovative and safe products, ultimately benefiting the market.

An analysis of the two regulations reveals two important points of contact. First, the AI Act requires providers of HRAIs and GPAI models to ensure an adequate level of cybersecurity protection throughout the lifecycle of the system or model. This emphasises the importance of cybersecurity in the design and development of AI systems. Secondly, the overlap between the two regulations is evident in the provision of

a regulatory experimental space aimed at promoting cybersecurity. The CRA establishes 'cyber resilience regulatory sandboxes', while the AI Act stipulates that AI sandboxes must facilitate the development of cybersecurity profiles for AI systems undergoing experimentation.

In essence, both regulations recognise cybersecurity as a priority, highlighting the importance of experimentation and innovation in this area, and identifying AI regulatory sandboxes as a potential tool to promote cybersecurity in the AI sector. Based on these premises, it is likely that AI regulatory sandboxes will become spaces for empirical dialogue between the AI Act and the CRA in the near future, also thanks to their mandatory nature.

It is no coincidence that European regulators are investing in sandbox frameworks in the areas of AI and cybersecurity. Similarly, Mario Draghi's emphasis on regulatory sandboxes in his 2024 report (European Commission 2024, 34), describing them as 'a catalyst for innovation in Europe's digital economy', highlights their strategic value. Potentially, within a few years, a comprehensive framework of regulatory sandboxes for AI and cyber resilience could emerge at national and local levels. This could be an opportunity to develop a framework of interconnected national sandboxes focused on AI and cybersecurity.

Within this transformative regulatory landscape, cybersecurity is emerging as a critical issue that cannot be overlooked. Regulatory sandboxes with a focus on cybersecurity can play a crucial role in fostering a safe, fair, and healthy digital ecosystem. In fact, the safety and security of AI products are inextricably linked to their cybersecurity stance. A robust cybersecurity framework is the foundation upon which safe and secure products are built, and its absence can compromise the integrity of even the most innovative technologies.

With regulatory sandboxes already established in many countries and others preparing to launch AI-focused experimentation spaces, EU Member States have a unique opportunity to respond to European regulators by creating interconnected national experimentation frameworks for AI and cybersecurity. These sandboxes could also be linked to sector-specific initiatives, such as those for medical devices.

By investing in a comprehensive network of national and local regulatory sandboxes, a Member State could position itself as a pioneer in technology experimentation. Such

an initiative would provide valuable opportunities for testing and dialogue for national technology companies, especially SMEs and start-ups, increasing their productivity, fostering the growth of digital markets, and improving their international competitiveness.

Beyond the economic benefits, a structured regulatory framework for AI and cybersecurity sandboxes would also promote product safety, disseminate knowledge on smart innovation, and foster a culture of experimentation. This approach would enable regulators to raise awareness and effectively enforce regulations. Ultimately, prioritising interconnected AI and cybersecurity sandboxes would support safe experimentation, drive innovation, and contribute to a secure and robust digital ecosystem.

LEGISLATIVE INTERSECTION PERSPECTIVES ON REGULATORY SANDBOXES: NAVIGATING THE INTERPLAY BETWEEN THE AI ACT AND THE GDPR

DAVIDE BALDINI*

SUMMARY

ABSTRACT

AI systems often process personal data, implicating the GDPR alongside the AI Act. This paper addresses interpretative issues produced by the cumulative application of the GDPR and the specific rules established by the AI Act in the context of regulatory sandboxes, proposing potential solutions. Article 59 of the AI Act is the central provision in this respect, as it governs the re-use of personal data in regulatory sandboxes. The paper thus explores the relationship between this provision and other applicable data protection rules, especially Article 6(4) of the GDPR. Specific attention is also devoted to the re-use of special category data within regulatory sandboxes, where an interpretative solution is advanced with a view of ensuring sufficient safeguards in light of Article 9 of the GDPR. The use of synthetic data

* PhD Student at University of Florence. Contact email: davide.baldini@unifi.it

within regulatory sandboxes and the performance of DPIAs are then addressed. Finally, the Garante per la Protezione dei Dati Personali's role in supervising regulatory sandboxes is examined, emphasizing the need for coordinated governance and to ensure the authority's independence. The paper concludes that the highlighted issues necessitate detailed guidelines to ensure legal certainty and avoid the risk of contrasts with the GDPR.

1. – AI REGULATORY SANDBOXES: THE RELEVANCE OF DATA PROTECTION LAW

It is widely understood that most AI systems which are currently in use, especially high-risk AI systems and general-purpose AI systems, process personal data either in the context of their production, deployment, or both (Sartor and Lagioia 2020). When this is the case, the material scope of application of Regulation (EU) 2016/679 (General Data Protection Regulation, 'GDPR') is triggered, so that both the GDPR and Regulation (EU) 2024/1689 ('AI Act') apply. In light of the general clause of prevalence of the former over the latter,[1] the GDPR applies in full also to the processing of personal data by AI systems which takes place in the context of regulatory sandboxes.

While much attention has been devoted to the general interplay between the GDPR and the AI Act (Falletta and Marsano 2024), less focus has been placed on the application of the GDPR in the context of AI regulatory sandboxes.

The EU legislator has taken into consideration the likely occurrence of personal data processing in AI regulatory sandboxes, thereby establishing specific rules in case of such occurrence. In particular, Article 59 of the AI Act lays down rules dedicated to the (further) processing of personal data in the sandbox, while Article 57(10) of the AI Act mandates national data protection authorities ('DPAs') to be associated with the operation of regulatory sandboxes and involved in the supervision of aspects which are relevant for data protection.

While seemingly unproblematic at a first glance, these provisions raise interpretative issues, especially considering their relationship with the GDPR, or even with other AI Act provisions addressing personal data processing. The solution to these issues has, as we

[1] Art. 2(7) AI Act.

shall see, relevant practical implications for providers seeking to attend a sandbox.

The following paragraphs will examine these provisions and their relationship with the GDPR, thereby highlighting some critical aspects and suggesting possible interpretations aimed at preventing inconsistencies.

2. – THE RULES ON PROCESSING PERSONAL DATA WITHIN REGULATORY SANDBOXES: THE SCOPE OF ARTICLE 59 OF THE AI ACT AND ITS RELATIONSHIP WITH THE GDPR

Given the nature and function of regulatory sandboxes, the processing of personal data by the attending provider can take place for the purposes of development, training and testing of the AI system. When this is the case the provider which attends the sandbox is considered, in GDPR terms, as a data controller, being the entity that determines the purposes and the essential means of the processing (EDPB 2021-a).

As a consequence, the provider is responsible for complying with all GDPR rules which are aimed at data controllers, except where more specific rules on data protection are provided by the AI Act itself, in accordance with the interpretative principle *lex specialis derogat generali.*[2] In particular, given its specific aim of regulating the re-use of personal data in the context of regulatory sandboxes, we contend that Article 59 of the AI Act is to be considered *lex specialis* compared to Article 6(4) GDPR, which covers all instances of data re-use and is thus *lex generali.*

Aside from the specific prevalence of Article 59 of the AI Act over Article 6(4) of the GDPR – which we will further explore in the following paragraph –, it is reasonable to maintain that the processing of personal data within the sandbox must still respect all other GDPR rules and principles that are not directly intended to regulate data re-use, in light of the general prevalence of the GDPR over the AI Act.[3] For example, the provider which attends the sandbox, as data controller, will need to conform with the data protection principles outlined in Article 5 of the GDPR, such as the principle of transparency by providing an exhaustive privacy notice to data subjects, updating the

---

[2] It should be noted that this principle is part of the General Principles of EU Law (*ex multis:* CJEU 30.4.2014, C-280/13).

[3] See n. 1, supra.

Records of Processing Activities ('ROPA') under Article 30, and so on.

In light of the above, we contend that the concerns raised by some Member States during the AI Act legislative procedure regarding the alleged shortcomings of Article 59 in respecting the right to personal data protection, are misplaced. In particular, Austria has maintained that the provision 'completely disregards the data protection principle of data minimisation pursuant to Article 5(1)(c) GDPR, because neither the scope nor the categories of personal data potentially processed in regulatory sandboxes are limited in any way' (Council of the European Union 2024). The assertion does not seem to be accurate, given that – according to the interpretation proposed above – Article 59 of the AI Act does not prejudice the applicability of Articles 5(1)(c) of the GDPR, which means that the provider must only process the minimum amount of personal data necessary to develop, train and test the AI system, along with the other data protection rules and principles.

2.1. – Focus on the relationship between Article 59 of the AI Act and Article 6(4) of the GDPR

It should be highlighted that the scope of Article 59 of the AI Act is narrower than it appears at first: the provision is not intended to lay down a general framework covering any processing of personal data which takes place in the sandbox; rather, it addresses only situations of so-called 're-purposing' or 're-use' of personal data. The expression refers to instances where personal data collected by the provider/data controller for legitimate purposes is subsequently used for different purposes, in this case for developing, training and testing AI systems in the sandbox.

A primary practical consequence is that when the provider intends to directly collect personal data for one or more specific sandbox-related purposes, rather than repurposing previously acquired data, Article 59 of the AI Act will not apply. Instead, in this case only the lex generali (GDPR) will regulate the gathering and the subsequent processing of personal data within the sandbox.

Therefore, by only applying to instances of re-purposing of personal data in the sandbox, we contend that Article 59 of the AI Act is a specification of the so-called 'compatibility test' laid down in Article 6(4) of the GDPR. This provision applies in

fact to 'the processing for a purpose other than that for which the personal data have been collected'. As a result, and as anticipated above, when the provider intends to re-use personal data for developing, training and testing AI systems in the sandbox, only Article 59 AI Act should apply, as lex specialis vis-à-vis Article 6(4) of the GDPR, despite the general clause of prevalence of the GDPR over the AI Act.

The other possible interpretation is that both provisions cumulatively apply in the case of data re-purposing in the context of regulatory sandboxes, based on the general clause of prevalence of the GDPR over the AI Act.[4] However, this second reading is not only arguably incorrect from a formal perspective, as discussed above, but also less preferable in practice, given that – as it shall be seen in the following paragraph – complying with Article 59 of the AI Act alone can be quite challenging. If adopted, this interpretation would therefore have the effect of making the lawful re-use of personal data in the sandbox quite challenging.

In light of the foregoing, we suggest that official guidance on the relationship between the two provisions be issued either by the European Data Protection Board (EDPB) or by the AI Board, clarifying that only Article 59 of the AI Act applies to instances of personal data re-use within the sandbox. This would reduce legal uncertainty and prevent possible fragmentation between Member States' regulatory sandbox solutions, which could happen in the case where national data protection authorities pick up different interpretations on this issue.

2.2. – The processing of special category data in the sandbox

Another relevant question pertains to whether Article 59 of the AI Act lays down a suitable legal basis for re-using (i.e., processing) a special category of personal data (Art. 9(1) GDPR) in the sandbox. The issue has relevant practical implications, considering the broad interpretation of this category which has recently been advanced by both the Court of Justice of the European Union (CJEU 1.8.2020, C-184/20), and the EDPB (EDPB 2021-b, 32-34), the latter with specific regard to the creation of a special category data through algorithmic inferences, an occurrence which has a huge impact on the development and functioning of AI systems.

---

[4] See n. 1, supra.

In this respect, Recital 140 of the AI Act identifies Article 59 of the AI Act as the appropriate legal basis for processing special category data under the 'substantial public interest' ground outlined in Article 9(2)(g) of the GDPR. It acknowledges that the processing activities within the sandbox serve substantial public interests, as established by Union law.

While this explicit recognition appears to resolve the issue by permitting blanket authorization for the processing of special category data within the sandbox, a systematic analysis of the AI Act provisions reveals a more nuanced reality. In the context of requirements for developing high-risk AI systems, Article 10(5) of the AI Act permits the processing of special category data solely for the purpose of ensuring bias detection and correction.

Again, the question arises as to whether the two provisions should be applied cumulatively or not and, thus, whether the provider who attends the sandbox must comply with both Articles 10(5) and 59 of the AI Act when processing special category data.

While Recital 140 of the AI Act seems to suggest the prevalence of Article 59 of the AI Act, we contend that the two norms should be seen as applying cumulatively, for two main reasons.

Firstly, from a systematic perspective it would appear irrational to establish different requirements concerning data governance when special category data are processed within or outside of the sandbox. It should be noted that Article 10 of the AI Act is placed within Section 2 of Chapter III, which establishes the general requirements for high-risk AI systems, and that the overarching aim of regulatory sandboxes is precisely to facilitate the development of AI systems in a way which complies with the requirements of the Regulation. It would thus appear inconsistent to lessen or waive such requirements when the provider attends the regulatory sandbox, as this would run counter to the very reason that justifies the existence of sandboxes.

Secondly, a different interpretation appears to be at odds with the GDPR and could lead to the invalidity of Article 59 of the AI Act due to a contrast with Article 8 of the EU Charter of Fundamental Rights (the 'Charter'), a risk identified by Austria during

the closing negotiations on the AI Act (Council of the European Union 2024),[5] in light of the insufficient safeguards provided by Article 59 of the AI Act alone. As indicated above, in fact, Recital 140 identifies Article 9(2)(g) of the GDPR as the applicable legal basis for such processing; the latter provision, however, requires that the relevant Union law that authorizes the processing shall be 'proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject'. In this respect, it is questionable whether the application of the measures outlined in Article 59 of the AI Act are, alone, sufficient to safeguard the fundamental rights and interests of data subjects, given the high impact that processing special category data has on the individual, compared to 'common' personal data. Instead, the cumulative application of both Articles provides a much more robust framework of measures to safeguard data subjects' rights. In this respect, while some requirements set forth by the two articles may partially overlap with each other, we contend that any possible inconsistency can be easily solved via interpretation.[6]

In light of the above, we contend that the provider must process special category data within the regulatory sandbox only when strictly necessary to ensure bias detection and correction, and by adhering to the safeguards outlined in both Articles 10(5) and 59 of the AI Act. In this respect, official guidance on the cumulative application of both provisions should be adopted by the EDPB and/or by the AI Board in a timely manner, in order once again to reduce legal uncertainty and possible fragmentation.

---

[5] In particular, Austria held that 'In Austria's view, the processing of special categories of personal data is not permissible on the basis of Article 6(4) of the GDPR and runs counter to the risk assessment underlying the GDPR'.

[6] For example, Art. 9(2)(g) provides that 'data shall be deleted once the participation in the sandbox has terminated or the personal data has reached the end of its retention period', while Art. 10(5)(e) requires that 'the special categories of personal data are deleted once the bias has been corrected or the personal data has reached the end of its retention period'. Given that the specific purpose of processing special category data is to address biases, the latter provision should prevail over the former. The EDPB and AI Office should in any case issue official guidance on these overlaps.

2.3. – The role of synthetic data

Research on synthetic data[7] in the field of AI has been gaining traction in recent years (Lòpez and Elbi 2022), due to their potential for mitigating many risks to fundamental rights and safety of individuals, thereby facilitating compliance with both the GDPR and AI Act, while preserving the attributes and patterns inherent to the original dataset. Synthetic data is thus becoming one of the most commonly used and promising Privacy Enhancing Technologies to be applied in the field of AI (EDPS 2022, 10-11).

The AI Act contains two express references to synthetic data. One of the requirements set forth by Article 59 of the AI Act mandates that any processing of personal data that takes place within the regulatory sandbox must be strictly necessary for complying with the AI Act's provisions on high-risk AI systems, and only where such requirements may not be effectively fulfilled by processing 'anonymised, synthetic or other non-personal data'. Similarly, one of the conditions laid down by Article 10(5) of the AI Act to allow the processing of special category data for bias detection and correction purposes – including within the regulatory sandbox, as argued in the previous paragraph – is that such purposes cannot be effectively fulfilled by processing other data, including synthetic or anonymized data.

In this respect, it should be noted that both requirements seem to be a plain application of the data minimisation principle which, according to the EDPB interpretation, also mandates complete data avoidance when a given purpose can be fulfilled without processing any personal data (EDPB, 2020). Considering this, the provisions do not seem to have a specific legal effect, as they arguably recall already existing and applicable GDPR obligations.

More relevant, however, is the reference to 'synthetic data'. While the AI Act does not explicitly define synthetic data, Article 59 notably appears to classify it as fully non-personal data ('synthetic or other non-personal data'). However, the personal or non-personal nature of synthetic data remains a subject of debate in the literature (López and

[7] According to the European Data Protection Supervisor (EDPS 2022): 'Synthetic data is artificial data that is generated from original data and a model that is trained to reproduce the characteristics and structure of the original data. This means that synthetic data and original data should deliver very similar results when undergoing the same statistical analysis. The degree to which synthetic data is an accurate proxy for the original data is a measure of the utility of the method and the model'.

Elbi, 2022), particularly considering the various techniques used to generate synthetic data. Conversely, Article 10(5) of the AI Act seems much more 'agnostic' and simply refers to 'synthetic or anonymized data', which leaves open the possibility to use synthetic data that are still considered 'personal data' under the broad notion adopted by Article 4(1) of the GDPR.

Considering its growing importance, and given the references contained in the AI Act, data controllers and providers wishing to leverage synthetic data within the regulatory sandbox arguably deserve more clarity over their function within the new Regulation. As seen above, the latter seeks to promote synthetic data as a tool to advance data minimization and data avoidance in the context of regulatory sandboxes, including to address bias in the datasets, but at the same time fails to provide both a definition of the term and adequate indications concerning their use.

Looking outside of the AI Act, the situation does not seem to improve: the last official guidance at EU level on anonymization techniques (Article 29 Working Party 2014) has recently turned ten years old and, thus, does not provide specific indications on – nor even mention – synthetic data, as the technology has emerged only in the last few years.

In light of this, possible use-cases and related benefits for providers that seek to make use of synthetic data remain rather obscure. We thus recommend that both data protection and AI regulators address synthetic data and their function within both the AI Act and the GDPR, with specific reference to their processing in the context of regulatory sandboxes and for purposes of bias detection and correction, as well as to whether and in which conditions they may be considered as non-personal data.

In this respect, the EDPB has announced within its work program for 2023-2024 (EDPB 2023) its plan to issue updated Guidelines on Anonymization. It is reasonable to expect that the updated Guidelines will provide some clarity, at least concerning the relationship between synthetic and non-personal data. As a result, providers of high-risk AI systems that are eager to leverage synthetic data to facilitate compliance with both the GDPR and AI Act should keep an eye out for the issuing of said Guidelines.

2.4. – Data Protection Impact Assessments ('DPIAs') and regulatory sandboxes

A further requirement laid down by Article 59(1)(c) of the AI Act to allow the re-purposing of personal data in the context of regulatory sandboxes is the presence of 'effective monitoring mechanisms to identify if any high risks to the rights and freedoms of the data subjects, as referred to in Article 35 of Regulation (EU) 2016/679 (...), may arise during the sandbox experimentation, as well as response mechanisms to promptly mitigate those risks and, where necessary, stop the processing'.

Looking at the GDPR, it should be noted that the processing of personal data in the context of the development or use of high-risk AI systems typically translates to high risks for the rights and freedoms of data subjects under the GDPR which, in turn, triggers the obligation on the data controller to carry out a DPIA in light of Article 35 of the GDPR (Article 29 Working Party 2017). As a result, the provider that wishes to re-use personal data in the regulatory sandbox is arguably already expected to have carried out a DPIA over the relevant processing activities.

In this respect, the question arises whether Article 59(1)(c) of the AI Act adds further requirements for providers attending the regulatory sandbox, beyond those already stemming from the correct performance of the DPIA. At a first glance, the answer seems to be positive: the reference to the response mechanisms needed to promptly mitigate the risks and, if necessary, stop the processing, are arguably new. However, at a closer inspection, it may be argued that similar response mechanisms should already be identified within the relevant DPIA encompassing the development of a high-risk AI system, be it within or outside the regulatory sandbox. This conclusion is based on two main reasons. On the one hand, Article 35(7)(d) of the GDPR mandates that the DPIA include 'measures envisaged to address the risks, including safeguards, security measures and mechanisms to ensure the protection of personal data (...)'; on the other hand, Article 36 of the GDPR clarifies that any given processing that still presents a high risk for the rights and freedoms of data subjects should not take place. Read together, the two GDPR provisions arguably already require data controllers to implement measures that are aimed at stopping a personal processing which creates high risks, despite any mitigation measure already in place.

In light of the above, we may conclude that, while Article 59(1)(c) of the AI Act

does not create further obligations for providers attending the regulatory sandbox, beyond those already foreseen by Article 35 of the GDPR, it is a useful reminder for providers of high-risk AI systems that they need to carefully consider, and document within the DPIA, the implementation of monitoring systems and response mechanisms to mitigate risks and stop the processing, where necessary.

2.5. – A (provisional) conclusion: is Article 59 of the GDPR an enabler of, or an obstacle to, data-reuse within the regulatory sandbox?

Having specified the scope of application of the provision, as well as its relationship with other AI Act and GDPR provisions, a final and more general question pertains to the role and function of Article 59: is the provision intended to facilitate the further processing of personal data in the regulatory sandbox, as it may seem at first, or is it an obstacle to the possibility of data re-use, by providing more stringent conditions?

The answer can be found in the comparison with Article 6(4) of the GDPR. The latter provides five broad criteria which the data controller is required to consider when they intend to process data for purposes other than those for which they were collected. The provision does not limit in any way which new purposes may be pursued by the controller, which thus remains free to determine the purposes as long as they are legitimate, specific and explicit. Furthermore, the broad and flexible nature of the criteria arguably allows the data controller to perform and document a mostly 'discursive' internal assessment.

On the contrary, Article 59 of the AI Act appears way more specific and stricter. Firstly, it allows the provider to re-use the data solely for the purposes of developing, training and testing AI systems in the regulatory sandbox; secondly, the relevant AI system has to be aimed at safeguarding substantial public interest, pursued by a public authority or another natural or legal person, and only in specific fields; lastly, many specific technical and organisational measures need to be put in place, including but not limited to functional separation, keeping of the logs, and so on.

Considering this, given that the special provision (Article 59 of the AI Act) seems to establish stricter conditions compared to the general provision (Article 6(4) of the GDPR), it is reasonable to maintain that the former is not intended as an enabler to data re-use. As a result, Article 59 of the AI Act will arguably have the – intended or

unintended – effect of discouraging the re-use of data in regulatory sandboxes. A possible effect of this situation could be that providers will be pushed to circumvent this new legal obstacle, by instead collecting data for the specific purpose of developing, training and testing AI systems in the regulatory sandbox. This endeavour, however, is not without many challenges in its own respect.

3. – THE ROLE OF THE ITALIAN DATA PROTECTION AUTHORITY IN THE SANDBOX

Article 57(10) of the AI Act requires national Data Protection Authorities ('DPAs') to be associated to the supervision of regulatory sandboxes together with the competent AI regulators, whenever the relevant AI system carries out personal data processing activities. The competent DPA – that is, for Italy, the Garante per la Protezione dei Dati Personali (GPDP) – will thus have the role of supervising the relevant AI system concerning data protection rules, including those laid down by the AI Act, examined above.

Similarly to AI regulators, the relevant DPA is also required to provide guidance on regulatory expectations to the attending provider (Article 57(7) of the AI Act), while being prevented from issuing sanctions for bona fide failures to comply with the law (Article 57(12) of the AI Act).

Against this background, we will now briefly examine the two main issues arising from the above provisions, with specific reference to the Italian case.

3.1. – AI regulatory governance and its implications for the supervision of regulatory sandboxes: the Italian case

The AI Act leaves considerable leeway to Member States concerning which authority(ies) to appoint as AI Act regulator(s): each Member State may choose to appoint existing authorities, including the national DPA, or create new entities.

In this respect, the provisional decision of the Italian legislator may be taken as an example. Article 18 of the draft Legislative Decree on AI designates Agenzia per l'Italia Digitale (AgID) and Agenzia per la Cybersicurezza Nazionale (ACN) as the relevant authorities, despite the GPDP's express request to be appointed in light of the experience already accrued on the enforcement of data protection rules in the field of AI and given

the GPDP's independent nature (GPDP 2024).

While the purpose of this paper is not to discuss the merits of Italy's decision, the countries which opt to appoint their DPA as the AI Act regulator will be able to streamline 'by design' the governance of regulatory sandboxes, simply by having the same authority supervising both bodies of law at the same time. Conversely, Italy will need to carefully consider measured to ensure coordination between AgID, ACN and the GPDP, by establishing clear operational rules on their roles, competences and functions when jointly supervising regulatory sandboxes. This is especially important considering the heavy overlaps between AI regulation, data protection and cybersecurity aspects of AI. Failure to ensure proper coordination will likely lead to increased costs for authorities, more bureaucracy and costs for participants and, in the end, less attractive regulatory sandboxes.

3.2. – Regulatory sandboxes as a 'stress test' for the independence of DPAs

During the final phases of legislative procedure, Austria has maintained that the exemption for regulatory sandbox participants from the imposition of fines for breaches of GDPR and other data protection rules is at odds with Article 83 of the GDPR, as this norm does not foresee any possible exemption for fines in case of non-compliance with data protection rules (Council of the European Union 2024). Moreover, the exemption allegedly collides with the requirement of independence for the DPAs laid down in Articles 52 of the GDPR and 8(3) of the Charter. According to this interpretation, the exemption foreseen by Article 57(12) of the AI Act would thus be invalid for violation of EU Primary Law.

This interpretative position, however, seems to rest on a misguided reading of the requirement for the DPA's independence. Indeed, when confronted with the matter, the CJEU has adopted a broad interpretation of the notion of 'complete independence' (CJEU 16.10.2012, C-614/10), which includes inter alia independence from political influence such as in the case where the head of the DPA can be removed by the government (CJEU 8.4.2014, C-288/12) and other forms of indirect influence such as in the case of reporting obligations to the government (CJEU 9.3.2010, C-518/07). Despite this, nothing in the Court's jurisprudence, nor in the wording of Article 52 GDPR, seems to

suggest that a specific and time-limited exemption from the possibility to impose fines is at odds with the requirement for the DPA's independence. At any rate, it should be noted that Article 57(11) of the AI Act establishes a sort of 'safety clause', by providing that supervisory authorities are not prevented from making use of their other corrective powers, for example by issuing a ban on the unlawful processing activities, so that their powers are not entirely curtailed even when supervising regulatory sandboxes.

Nonetheless, we suggest that the DPAs' independence may instead be jeopardized by another risk which is inherent to regulatory sandboxes, that is, the risk of regulatory capture, whose possible occurrence in this context is well-documented (Ranchordas and Vinci 2024). Still, this risk materializes because of the close collaboration that takes place between the regulators and regulated, especially when there is a lack of transparency, not as a result of exemptions from imposing fines. While this risk is indeed present, the AI Act seeks to mitigate it by establishing transparency and openness requirements, whose effectiveness is addressed in other sections of this work.

At any rate, the independence requirement for DPAs set forth by the Charter arguably establishes a higher bar when it comes to the avoidance of regulatory capture, compared to the AI Act's less stringent indications on the independence of market surveillance authorities, as can be seen by comparing Article 70(1) of the AI Act and Article 52 of the GDPR. This may also be due to the fact that the DPAs' independence requirement stems directly from EU Primary Law (Articles 8(3) of the Charter and 16(2) TFEU). According to the CJEU, moreover, DPAs shall not only be completely independent, but shall also 'remain above any suspicion of partiality'.

In light of the above, we contend that the requirements set forth by the AI Act to avoid the risk of regulatory capture should be better specified and strengthened when it comes to DPAs association to the supervision of regulatory sandboxes. This could take place by means of issuing an EDPB opinion containing indications for DPAs to preserve their complete independence when attending regulatory sandboxes, with specific reference to their interactions with attending providers, taking into account relevant CJEU case-law and doctrinal analysis of the independence requirement.

4. – CONCLUSIONS

This paper has examined the nuanced interplay between the GDPR and the AI Act in the context of AI regulatory sandboxes, distinguishing between the substantive rules on the processing of personal data and the governance aspects of regulatory sandboxes. The detailed analysis and interpretation of the relevant provisions revealed several critical aspects and suggested pathways to address identified challenges.

In particular, we have seen that the GDPR generally remains applicable within AI regulatory sandboxes, except in specific instances where AI Act provisions address personal data processing. In this respect, we highlighted that Article 59 of the AI Act, despite its intention to facilitate the re-use of personal data in regulatory sandboxes, establishes stringent conditions that may inadvertently discourage data re-use. Providers might resort to collecting new data specifically for sandbox purposes, circumventing the hurdles posed by Article 59 of the AI Act.

Concerning the governance perspective, as the Italian case shows, the decision to appoint authorities other than the DPA as the relevant AI Act regulator gives rise to the necessity of putting in place a clear framework for cooperation of the three authorities within regulatory sandboxes. Moreover, the requirement of complete independence for DPAs, enshrined in EU primary law and reinforced by CJEU case-law, needs to be addressed with specific regard to the well-known risk of regulatory capture, inherent in the operation of regulatory sandboxes.

In conclusion, the integration of GDPR and AI Act provisions within AI regulatory sandboxes presents complex legal and practical challenges, that need to be effectively addressed in order for regulatory sandboxes to become effective instruments of cooperative regulation. While we have suggested possible interpretative solutions to the issues posed by the interplay of the two bodies of law in the context of regulatory sandboxes, we contend that joint guidance from competent authorities at EU level (essentially, the EDPB and AI Board) is essential to ensure legal certainty, avoid inconsistencies and prevent the occurrence of market fragmentation across the EU.

REGULATORY SANDBOXES UNDER THE INTEROPERABLE EUROPE ACT: TOOLS FOR REGULATORY EXPERIMENTATION

ELEONORA BONEL*

SUMMARY

ABSTRACT

In the context of rapid digital transformation, the European Union's regulatory frameworks must balance innovation with legal compliance. The EU's Interoperable Europe Act (IEA) introduces the 'interoperability regulatory sandboxes' as an innovation support measure to enable public sector authorities across the European Union to experiment with innovative solutions in controlled environments before integrating them into broader public systems. This contribution examines how the interoperability sandboxes expect to foster collaboration among public administrations and encourage the involvement of GovTech actors, promoting advancements in digital government innovations. It also explores the legal foundation for the interoperability sandboxes, focusing on Articles 11 and 12 of the IEA, which outline the conditions for their implementation. Additionally, the contribution compares the

* External services provider, DG DIGIT, European Commission. Contact email: eleonora.bonel@sciencespo.fr. This publication is not affiliated with the European Commission and the views expressed herein are solely those of the author.

IEA's interoperability sandboxes with those established by the AI Act, highlighting their complementary roles in supporting innovation while ensuring regulatory compliance. Although the two frameworks target different aspects of digital governance, their intersections — particularly in AI-driven and cross-border initiatives — suggest potential synergies.

---

1. – REGULATORY SANDBOXES: RELEVANCE FOR THE INTEROPERABLE EUROPE ACT

According to the Organisation for Economic Co-operation and Development (OECD), policymakers face a new set of regulatory challenges associated with digital transformation, and are responding to them in different ways, from 'wait and see approaches to outright bans on specific technology uses and applications, as is the case of Regulation (EU) 2024/1689 ('AI Act') (Attrey et al., 2018). Responses to emerging digital business models vary, leading to regulatory uncertainty. This can discourage innovation and delay user adoption (OECD, 2020), with the rapid pace of digital transformation further complicating efforts to predict market trends and public policy concerns.

To adapt, policymakers are moving beyond passive approaches and exploring innovative tools like regulatory sandboxes (Attrey et al., 2018). Indeed, the use of regulatory sandboxes is also foreseen by a new regulation: Regulation (EU) 2024/903 ('Interoperable Europe Act'), entered into force on 11 April 2024. The case of the Interoperable Europe Act (IEA) provides an interesting insight into how regulatory sandboxes can be aimed at public sector organizations specifically, providing insight into their participation as both regulator and innovator.

Overall, the IEA has the objective of strengthening cross-border interoperability and cooperation in the public sector across the European Union (EU). More specifically, cross-border interoperability is defined as the ability of Union entities and public sector bodies of Member States to interact with each other across borders by sharing data, information, and knowledge through digital means, in line with the legal, organisational, semantic, and technical requirements related to such cross-border interaction. In order to facilitate cross-border interoperability, the regulation introduces a cooperation framework for public administrations to exchange information and to stimulate public

sector innovation and public-private 'GovTech' projects. The term is defined in Article 2 of the IEA as 'technology-based cooperation between public and private sector actors supporting public sector digital transformation'. To support these aims, Article 10 of the IEA introduces some innovation and support measures, including setting up the interoperability regulatory sandboxes for policy experimentation, which are then further articulated in Articles 11 and 12.

Article 2 of the IEA defines such interoperability regulatory sandboxes as a 'controlled environment set up by a Union entity or a public sector body for the development, training, testing and validation of innovative interoperability solutions, where appropriate in real world conditions, supporting the cross-border interoperability of trans-European digital public services for a limited period of time under regulatory supervision'. As confirmed by the case of the IEA, experimentation clauses offer the legal basis for setting up regulatory sandboxes. In particular, Article 10 of the IEA outlines the process by which the Interoperable Europe Board, the designated governance body, may propose the creation of an interoperability regulatory sandbox to facilitate the development of innovative measures.

As suggested by the Council Conclusions (2020), we know that regulatory sandboxes could be considered tools for an innovation-friendly, future-proof and evidence-based regulatory framework that tries to combine two seemingly contrasting elements: fostering regulatory learning as well as European technological sovereignty. The present contribution explores how interoperability regulatory sandboxes could do so, while remaining a voluntary tool for Member States' public administrations and Union entities. It acknowledges the existence of other sandboxes with overlapping purposes for public sector use cases. The following chapters seek to clarify the unique value of interoperability regulatory sandboxes while comparing them to the AI Act sandboxes, exploring potential synergies and their service offerings.

2. – THE INTEGRATION OF REGULATORY SANDBOXES IN THE INTEROPERABLE EUROPE ACT: THE SCOPE OF ARTICLES 11 AND 12

Under the IEA, the 'interoperability regulatory sandboxes' as defined above, are a mechanism that is offered to Union entities and public sector bodies within the

Member States to contribute to at least one of seven objectives – as listed in Article 11. These objectives focus on three thematic areas: (i) fostering innovation by facilitating the development and deployment of innovative digital interoperability solutions for public services; (ii) promoting cross-border cooperation among national, regional, and local authorities to enhance synergies in public service delivery; (iii) facilitating the development of an open European GovTech ecosystem through cooperation with Small and Medium-sized Enterprises (SMEs), research and educational institutions, and start-ups. Most crucially, the sandboxes aim to enhance authorities' understanding of opportunities or barriers related to cross-border interoperability, support evidence-based regulatory learning, and improve legal certainty while sharing best practices to ensure compliance with relevant regulations.

The notion of the interoperability regulatory sandbox improving legal certainty is one of the key objectives of Articles 11, one which warrants greater reflection. The concept of 'legal certainty' can be understood as making laws predictable, interpreted as the law that must be certain, foreseeable and easy to understand (Venice Commission, 2011). Taking a step back, this notion can also be identified as an aim of the broader category of 'experimentation spaces' in which regulatory sandboxes can be placed in, alongside test beds and living labs (JRC)[1]. Indeed, these spaces for experimentation seek to offer a structured, collaborative, and evidence-based approach to regulation in industries where innovation outpaces traditional legal frameworks. In this sense, these tools can help to clarify the application of laws and set precedents that contribute to greater legal certainty (De Pasquale, 2023). In the case of the interoperability sandboxes, said objective depends on how well the sandbox's structure and process are designed, both in order to balance innovation with risk management, and it is largely defined on a case-by-case basis.

Article 12 of the IEA, in fact, specifies the conditions for participating in an interoperability regulatory sandbox. As noted in Article 11(4), application to the regulatory sandbox would be allowed through a joint request from at least three participants. Participant 'innovators' requesting access to the regulatory sandbox would be: (i) Public

[1] According to the Joint Research Center (JRC), test beds, living labs and regulatory sandboxes are spaces for experimentation that allow for the co-development of innovative solutions in (near) real world environments (Kert et al. 2022). They differ according to a set of criteria including time frame, context, regulators' scope and mechanisms of regulatory learning.

bodies established in one of the EEA Member States; (ii) A Union entity. Upon receiving the joint request, the European Commission, after consulting the Interoperable Europe Board, can authorise the establishment of such a sandbox. This authorisation applies as a checklist to procede to the sandbox, and it does not apply to Union entities. Participation is allowed also for GovTech actors, upon request of the applicants, following Article 12 of the IEA. The actors of this GovTech ecosystem, which is broadly defined in the IEA, can range from private sector actors to SMEs and standardization organizations, which would play a crucial role in experimenting with and testing new interoperability solutions within the sandbox. Their involvement in the sandbox, in particular in the potential testing of interoperability solutions, is evaluated on a case-by-case basis to ensure the right participants are chosen for each initiative.

2.1. – The focus on public sector authorities

The implementation of the IEA, and the consequent set-up of interoperability regulatory sandboxes, is expected to be a tool to support public sector innovation. Existing literature highlights the importance of beginning each regulatory sandbox initiative with a clearly defined hypothesis, accompanied by a comprehensive strategy for data collection and analysis, which may benefit from third-party assistance. Additionally, it is critical to ensure broad dissemination of the findings, particularly to policymakers involved in the legislative process, as well as to the general public (Armstrong et al., 2020; Nesta, 2020). If the 'ingredients' of the regulatory sandbox are effectively set-up, this could lead to different benefits for the experimentation's participants.

In the context of the IEA, it is possible to identify a number of benefits for the participants setting up the sandbox, i.e. the public sector authorities of the Member States and the EU institutions. Following the formulation of articles 11 and 12 of the IEA, public sector entities participating in the regulatory sandbox can expect to benefit from several key advantages, including:(i) anticipating risks before they are integrated in the network and information systems in a cross-border setting; (ii) increasing innovators' knowledge of the regulatory frameworks and their ability to plan and pre-empt enforcement action; (iii) facilitated coordination in exploring regulatory questions on cross-border innovation issues; (iv) dedicated support from relevant regulators throughout the process, ensuring

that participants receive timely guidance to address regulatory issues related to cross-border interoperability. While these are derived from some of the existing legal texts, the forthcoming implementing act for the IEA will further specify some of the obligations and rights of the participants.

On the other hand, by taking the perspective of the regulators setting up the regulatory experimentation exercise, the sandbox would allow for two key identifieable benefits: regulatory agility of the Interoperable Europe Act and providing a tool to foster collaboration between public sector authorities. Scholars argue that regulatory experimentation plays a vital role in 'reflexive governance', which leverages 'information obtained from experimentation with existing policies and regulations to enhance them' (Bauknecht, Bischoff, et al., 2020, p. 50). This principle may apply also in the case of the interoperability sandbox where, based on Articles 11 and 12 of the IEA, regulators are incentivized to establish interoperability regulatory sandboxes primarily to enhance public sector innovation, improve cross-border digital services, and foster collaboration between public administrations and private actors.

Moreover, establishing an interoperability regulatory sandbox is not a legal obligation for governments to implement, while it is an obligation of the Commission to help establish its overarching framework. As stated in Article 11 IEA, the establishment of interoperability regulatory sandboxes occurs upon a joint request of a minimum of three public sector authorities or Union entities. This indicates that the responsibility for establishing such sandboxes rests with the relevant public sector authorities, supported by the appropriate regulatory bodies, with the primary objective of enabling these authorities to derive benefit from the tool. While the Regulation establishes an interoperability governance structure, it is important to underline that the individual regulatory sandbox projects may explore different thematic areas according to the requests of the establishing participants, subject to the submission of a plan as specified in article 12 of the IEA. An entry point for discussions, for example, could also be any regulatory uncertainty tied to the interoperability assessments.

It is therefore an important opportunity for participants to cooperate and engage in regulatory learning on innovative cross-border issues across the EU. By identifying and addressing issues in a sandbox environment, public authorities can avoid larger-scale problems that might arise during broader implementation or before interoperability

solutions are integrated into the network and information systems of the European public sector.

2.2. – Involvement of the GovTech ecosystem

Participation in the sandbox is allowed also for GovTech actors, upon request of the applicants, following Article 12 of the IEA. The IEA is in fact the first regulation to explicitly promote the deployment and reuse of GovTech solutions, where the public sector engages with private actors, especially SMEs and start-ups, to procure innovative interoperability solutions. This term becomes important in the ecosystem of digital government, yet it is well distinguished. Indeed, following existing research, GovTech is oriented to enlarge the collaboration with innovation stakeholders, with the aim to 'reframe' the way services are designed and delivered and by introducing changes in the established processes. Therefore, the involvement of GovTech actors would be important in experimenting with and testing new interoperability solutions within the sandbox. Their involvement, particularly in testing interoperability solutions, is evaluated on a case-by-case basis to ensure the right participants are chosen for each initiative. While the interoperability sandbox is aimed at public sector entities, the involvement of GovTech entities would be beneficial due to their expertise in digital government innovations, which would contribute to practical insights and ground-level experience to the discussions in the regulatory dialogue. Engaging with the GovTech actors through sandboxes could help build trust and cooperation between regulators and innovators. There is some initial evidence of this type of public-private cooperation from the first cohort of the European Blockchain Sandbox, which underlined how the sandbox 'facilitated the sharing of knowledge and experience between regulators/authorities and with innovators on the basis of concrete use cases resulting in a better understanding of compliance requirements' (European Commission, 2024, p. 78).

Identifying shared innovation needs and priorities, along with focusing common GovTech and experimentation efforts across borders, would help Union public sector bodies to share risks and best practices. While these effects cannot be anticipated, evidence from the first Blockchain regulatory sandbox highlights the benefits of cross-border collaboration in promoting a more unified regulatory approach. From this angle,

successful GovTech projects and innovation measures piloted by Interoperable Europe innovation measures should help scale up GovTech tools and interoperability solutions for reuse.

2.3. – Implementation conditions: how are Interoperability sandboxes expected to support public sector innovation?

By 12 April 2025, an implementing act for the Interoperable Europe Act is expected to be adopted, which will provide crucial guidance on the conditions and implementation of interoperability regulatory sandboxes. The implementing act is crucial for preventing fragmentation across the Union, and in the context of the IEA it will set out the detailed arrangements and functioning of interoperability regulatory sandboxes. In accordance with Regulation (EU) 182/2011, the Commission will adopt the implementing act after consulting with committees of technical experts from EU Member States. In line with Article 12 of the IEA, these should specify the obligations of sandbox participants. This would include the application procedures, the conditions of participation, but also clauses on the termination of the interoperability sandboxes, including the sandbox exit report. As a result, a full understanding of the operating conditions of the sandboxes will only be possible once the implementing act is adopted, which will provide further clarity on how they will operate.

The IEA introduces interoperability regulatory sandboxes as an auxiliary tool to support innovation. As noted above, unlike the regulatory sandboxes mandated by the AI Act, there is no legal requirement for Member States to establish interoperability sandboxes. However, during the trialogue negotiations, their establishment was seen as an opportunity for regulatory experimentation and fostering cross-border cooperation. Given the novelty of the initiative, the European Commission must dedicate significant effort to promoting the interoperability sandboxes and clarifying their scope, especially through the upcoming implementing act and related communications.

According to Article 12 of the IEA, all relevant information regarding the interoperability regulatory sandbox should be accessible via the Interoperable Europe Portal. The portal will serve as a centralized platform aimed at facilitating the sharing and reuse of high-quality, reliable interoperability solutions among public administrations. To

ensure the effectiveness and coherence of this initiative, it would be important not only to adopt a clear implementing act but also to design a strategic and user-friendly experience for innovators navigating the Interoperable Europe portal.

Ultimately, the participation in the regulatory sandboxes shall not affect the supervisory and corrective powers of any authorities supervising the sandbox, which is an integral part of Article 12. The participation in the interoperability regulatory sandbox is set to be according to specific criteria, including a limited period and a specific plan that ought to be elaborated by the participants. This said plan reinstates some conditions for participation, including a risk management mechanism, a system for reporting requirements and an indication of where it is strictly necessary and proportionate to process personal data. The specific provisions relative to personal data processing, and more broadly the formulation of regulation 2012/903 in general, do not constitute a lex specialis, as the GDPR is still the main regulation applicable in the running of the regulatory sandbox.

3. – REGULATORY INTERSECTIONS AND DIVERGENCES: THE ARTIFICIAL INTELLIGENCE ACT SEEN ALONGSIDE THE INTEROPERABLE EUROPE ACT

This section focuses on a comparative analysis of the regulatory sandboxes introduced by the AI Act and the Interoperable Europe Act, exploring their roles in fostering innovation and providing a future-proof legal framework. Both regulations, as highlighted in their recitals (Recitals 138-143 of the AI Act; Recitals 41-42 of the IEA) establish at least one regulatory sandbox designed to create controlled environments where emerging technologies can be tested before wider implementation. The AI Act, through Articles 57-61, focuses specifically on artificial intelligence systems, enabling developers and regulators to interact in a sandbox environment to ensure AI products meet compliance standards. On the other hand, Articles 11 and 12 of the IEA present a broader scope, targeting public sector innovation through digital interoperability, which may include AI-based solutions, but also GovTech actors.

While the AI and the interoperability regulatory sandboxes diverge in some of their peculiarities, they share the common goal to foster regulatory learning for innovative uses cases. The following paragraphs aim to identify areas where these frameworks

complement or diverge from each other in promoting innovation within the European Union's regulatory landscape.

3.1. – Key differences

The regulatory frameworks for the sandboxes under the AI Act and the IEA present some differences, both in their legal obligations and in the scope of their application. One significant distinction lies in the legal mandate established by the AI Act. This regulation sets a clear obligation on Member States to ensure that at least one regulatory sandbox for AI is created at the national level. Each Member State must then guarantee that these regulatory sandboxes are operational by August 2, 2026 (following Article 57). This national focus places a substantial responsibility on each Member State to foster AI innovation while maintaining a robust regulatory framework. It is possible, moreover, that Member States could fulfil their legal obligation by participating in already existing regulatory sandboxes or 'establishing jointly a sandbox with one or more Member States' competent authorities' (Recital 138). Nonetheless, while the sandbox aims to accelerate time-to-market for AI products and ensure they are safe for consumers, the benefits of participation in these sandboxes are likely to vary depending on how they are implemented.

On the other hand, the IEA does not impose a binding obligation on Member States to establish regulatory sandboxes. Rather, the establishment of the interoperability sandboxes is overseen by the Interoperable Europe Board and is not confined to specific Member States, and participation is also allowed to Union entities. The interoperability regulatory sandboxes are cross-border by default, highlighting the importance of collaboration and cooperation between public sector stakeholders, aiming to improve digital infrastructure and public services throughout the European Union. While legal questions surrounding the use of AI systems may be part of the dialogues in the interoperability regulatory sandboxes, their scope is broader and extends to fostering innovation and facilitating the development of digital interoperability solutions for public services (Article 11 IEA).

Another key distinction between the two regulations is their different focus regarding the stakeholders involved in the regulatory sandboxes. The AI regulatory sandbox places a strong emphasis on the private sector, particularly companies developing

AI systems, with a primary focus on ensuring the safety and compliance of certain AI systems before they enter the market. Participation in the AI regulatory sandbox should revolve around addressing issues that raise legal uncertainty for providers and prospective providers seeking to innovate with AI in the EU, and also to contribute to evidence-based regulatory learning (Recital 139 AI Act). On the other hand, the IEA's sandbox emphasizes facilitating public sector authorities' understanding of opportunities or barriers to cross-border interoperability. It aims to contribute to the enhancement of Interoperable Europe solutions, but also more broadly improve legal certainty to ensure compliance with the IEA and, where appropriate, 'with other Union and national law' (Article 11 IEA).

An example of a use case for interoperability regulatory sandboxes is the role of data protection authorities (DPAs). In the interoperability sandbox, DPAs would act as supervisors but could also take on the role of innovators, using the sandbox environment to test new methods of cross-border data exchange and improve their collaboration with other DPAs. This flexibility is particularly relevant in light of the legislative proposal for Regulation no. 2023/202, which aims to strengthen the enforcement of the EU's General Data Protection Regulation (GDPR). Despite the GDPR being in effect since May 2018, reports have highlighted that cross-border issues are complicated by differing administrative procedures. Ultimately, this could be an opportunity to reduce the burden in the long-term and get a head-start in cross-border DPA cooperation. A well-planned initiative means easier supervision.

Overall, we can see how the two types of sandboxes have different scopes, yet they can still complement each other in meaningful ways. This matters because when regulatory frameworks for public sector innovation, such as the AI and interoperability sandboxes, work in harmony, they can unlock greater efficiency and innovation across different technological domains. In fact, the next paragraphs underline the legislative intersections of the two regulations. By fostering synergies between these instruments from the start, we can avoid implementation challenges (especially where use cases overlap) and create a more unified regulatory environment that supports public sector innovation across a variety of technologies.

3.2. – Points of regulatory intersections: implementing acts obligations and fostering legal certainty

Both the IEA (Article 11) and the AI Act (Article 68) propose the adoption of implementing acts on the detailed arrangements for and functioning of regulatory sandboxes. Notably, both articles specify that their respective implementing acts should include: (i) eligibility and selection criteria for participation; (ii) procedures for application, the selection of, and exit conditions from the relevant regulatory sandbox; (iii) the 'rights and obligations' of sandbox participants. While for the IEA, the latter is formulated as rights and obligations, under the AI Act this is specified as 'terms and conditions'.

While both regulations set out the specific details to be elaborated further in their implementing acts, the interoperability sandbox is the first initiative with a legal deadline for the relevant implementing act on 12 April 2025, making it a cornerstone for others to build on. This could potentially help foster synergies between the implementing acts of the AI Act regulatory sandbox, ensuring any auxiliary coordination mechanisms between these experimentation tools.

Crucially, the interoperability regulatory sandbox and the AI Act sandbox can play significant roles in ensuring compliance with other Union legislation, provided that the relevant competent authorities are involved in their supervision. In particular, Recital 139 of the AI Act further suggests that relevant authorities might use other regulatory sandboxes to ensure AI systems comply with this Regulation. The interoperability sandbox (Article 11 IEA) focuses on 'improving legal certainty' and sharing best practices to ensure compliance with its Regulation and applicable laws. This collaborative approach integrates sandbox outcomes into broader regulatory frameworks, enhancing clarity across legal domains. Likewise, the AI sandbox (Article 57 of the AI Act) has among its objectives to 'improve legal certainty' for regulatory compliance. Both sandboxes, therefore, can align innovative practices with existing legal obligations, promoting regulatory coherence.

Moreover, both the interoperability sandboxes and AI sandboxes share similar considerations in terms of data protection, security, and operational structure. In both cases, national and EU data protection regulations remain the primary governing frameworks, ensuring that personal data is handled with the highest standards of security

and oversight. National DPAs continue to play a central role in supervising personal data processing, with binding authority to ensure compliance. Aditionally, both frameworks reinforce the importance of documenting and publicly sharing project objectives and outcomes, promoting transparency while safeguarding sensitive information.

To achieve their relevant objectives, regulatory authorities, in collaboration with participating organizations, must commit significant resources to innovation, involving both human capital and financial investment. A challenge, in this setting, could be that the same authorities already overseeing compliance with the AI Act and other key EU regulations, such as the Digital Operational Resilience Act (DORA) and the Digital Services Act (DSA), could also be responsible for supervising sandbox activities. This may strain the capacity of regulatory bodies, which are often already overburdened, as sandbox programs add to their workload. Therefore, public authorities must absorb the associated operational costs to ensure that experimental environments do not undermine the successful deployment of new technologies or overextend regulatory agencies.

3.2.2. – Points of regulatory intersections: overlapping use cases

An additional point of intersection between the AI and interoperability sandboxes could be the scope of their use cases. Indeed, the use cases of the interoperability sandboxes, while broadly defined on an individual basis, could include questions about AI-based interoperability solutions. The case for this overlap is strongly supported by the Public Sector Tech Watch (PSTW), which tracks the use of emerging technologies in the public sector. Out of over 1,000 use cases mapped out in the observatory, 940 involve AI-based solutions (European Commission, 2024, p.50). This underscores the reality that a large part of the innovation happening in the public sector revolves around AI, making it highly likely that many initiatives might see an overlap in the scope of both regulatory sandboxes.

Take, for example, a cross-border, high-risk AI system developed by public authorities, such as an AI-powered health data exchange platform. In such a case, both the AI Act and Interoperability sandboxes could both play critical roles, individually or, if envisioned in the future, jointly. The AI sandbox would focus on ensuring that the AI system adheres to the safety, transparency, and ethical guidelines set out by the

AI Act. Meanwhile, the Interoperability sandbox would concentrate on solving the technical and legal challenges of exchanging health data between national systems, addressing the complexities of cross-border interoperability. To avoid duplication, early-stage coordination is crucial, perhaps through a shared database that flags overlapping use cases for joint review. This would ensure that AI-specific and interoperability issues are addressed in parallel without overburdening regulators, ultimately streamlining the process, and ensuring smoother implementation of high-risk, cross-border AI systems.

Although this kind of coordination between the sandboxes is theoretically possible, as it stands, both frameworks operate independently. However, more specific details on how these sandboxes might coordinate would be very much welcomed in the future. The upcoming implementing acts for both the AI Act and Interoperable Europe Act are expected to give further clarity on the practical running of the sandboxes, and this could perhaps include how they interact with other running sandboxes. If such coordination is encouraged, it could streamline regulatory processes and prevent duplication of efforts, particularly in projects where AI and cross-border interoperability intersect.

3.2.3. – Interaction of AI sandbox and Interoperability sandbox with other pro-innovation mechanisms

An additional point of synergy is the interaction of both the AI and interoperability regulatory sandboxes with other pro-innovation mechanisms. Both the AI Act and the Interoperable Europe Act emphasize the importance of collaborating between different stakeholders across public and private sectors. This synergy is not only central to advancing technological development but also crucial in the context of legal frameworks governing innovation and experimentation. The involvement of a wide array of actors, such as testing and experimentation facilities (TEFs), European digital innovation hubs (EDIHs), standardization organizations, and research laboratories, is an advantage for innovators, enabling more robust experimentation and governance.

In the case of AI regulatory sandboxes, Article 58 of the AI Act provides a clear complementary role for EU Testing and Experimentation Facilities (TEFs), which could serve as the technical infrastructure supporting the testing of AI applications within sandboxes. Such interaction between TEFs and the interoperability regulatory sandbox

could be further explored from an administrative and public law perspective. This aspect is further expanded on in some recommendations by the OECD's report (2023) on AI regulatory sandboxes, highlighting how 'TEFs' interactions with sandboxes from administrative, public law perspectives are yet to be explored' (OECD, 2023).

Similarly, Article 11 of the IEA encourages the involvement of the European GovTech ecosystem, but also 'research and experimentation labs, innovation hubs, and companies wishing to test innovative interoperability solutions' reinforcing the alignment of innovative solutions with existing tools and legal frameworks. This interconnected approach also reflects broader trends in regulatory experimentation, as seen in the OECD's mapping of regulatory instruments, which treats each tool as part of an integrated system. Through its analysis, each regulatory experimentation tool could be thought of as one node interacting with other nodes in an interconnected system (OECD, 2023). Spain's insight in a presentation to the OECD Working Party on AI Governance (OECD, 2023), also welcomed creating international clusters of AI sandboxes enabling cross-testing. According to Spain, future regulatory trends will be to create international clusters of AI sandboxes enabling cross-testing. To some extent, this could be envisioned beyond AI Sandboxes and including the Interoperability sandboxes.

4. – CONCLUSIONS

Regulatory sandboxes often share fundamental design components, although their implementation can vary significantly based on the context (Jenik and Lauer, 2017). In the case of the IEA, interoperability regulatory sandboxes offer a flexible framework for innovation and regulatory experimentation, particularly for public sector authorities. Articles 11 and 12 of the IEA establish the legal basis for these sandboxes, aiming to promote cross-border collaboration while safeguarding compliance with overarching legal frameworks such as the GDPR.

The forthcoming implementing act, expected by April 2025, will define the specific procedures and obligations for these sandboxes. It will have to strike a balance between fostering innovation with maintaining regulatory oversight, especially as participation in the sandboxes remains voluntary. Neverthless, the potential for these sandboxes to drive cross-border cooperation and innovation is significant and offers the EU an important

chance to cultivate controlled experimentation and regulatory learning.

A comparison of the interoperability sandboxes with the regulatory sandboxes established by the AI Act reveals complementary yet distinct goals. The AI Act focuses on AI system testing, minimizing administrative burdens for SMEs, and supporting the development of tools for regulatory learning in areas like accuracy, robustness, and cybersecurity. The IEA's sandboxes, on the other hand, target the development of interoperability solutions across public administrations. However, both frameworks share a common purpose: they provide controlled environments for innovation that seek to reduce regulatory uncertainty while ensuring compliance. Whether that is achieved, is a different story.

Looking ahead, there is significant potential for closer coordination between these two sandbox frameworks. Regulators should explore how governance can address overlaps and foster practical cooperation, such as harmonizing eligibility criteria or providing clearer guidance for innovators.

The reasoning behind such need for coordination is clear: streamlined regulatory processes would prevent duplication of effort, reduce administrative burdens, and enhance the overall efficiency of governance. Shared resources, such as a joint database to flag overlapping use cases, could ensure that projects falling within the scope of both frameworks are efficiently managed and appropriately supervised. Ultimately, while the possibility of formal coordination between the AI sandboxes and Interoperability sandboxes remains hypothetical at this stage, it represents a promising path forward. The forthcoming implementing acts will provide further clarity, but regulators ought to start thinking now about how cooperation within the Commission could optimise the sandbox frameworks to drive public sector innovation in AI-driven, cross-border scenarios.

OPERATIONALIZING AI REGULATORY SANDBOXES: A LOOK AT THE INCENTIVES FOR PARTICIPATING START-UPS AND SMES BEYOND COMPLIANCE

ANTONELLA ZARRA*

SUMMARY

ABSTRACT

Initiatives of agile, multi-stakeholder and experimental regulation are increasingly being adopted by authorities to govern emerging technologies. Regulatory sandboxes in artificial intelligence (AI) have been established in Europe within the framework of the Artificial Intelligence Act (AI Act) as a pro-innovation tool allowing AI providers to test their products in a controlled environment while understanding their potential risks and improving them. While on paper regulatory sandboxes can boost innovation by offering a safe space for experimentation, in practice their implementation brings about several operational challenges for the involved parties, along with risks of regulatory capture and forum shopping across

* PhD Candidate at Universität Hamburg and Case Handler Officer at the European Commission. The information and views set out in this article belong to the author and do not necessarily reflect the official opinion of the European Commission. Contact email: zarra@law.eur.nl.

jurisdictions. Against this background, this paper outlines the main incentives for innovative firms to take part in AI regulatory sandboxes and highlights the challenges for such an instrument to act as a 'silver bullet' for the innovation-safety dilemma. The paper calls for more streamlined and clear procedures for participating firms and an effective governance structure, concluding with recommendations for policymakers and practitioners willing to embark on a sandbox-journey.

---

1. – FROM EXPERIMENTAL REGULATION TO AI REGULATORY SANDBOXES

Emerging technologies, including AI-powered products and services, are excellent means to promote wider consumer choice and greater productivity. Nevertheless, their benefits pair with equally important risks, which are to be mitigated. In this trade-off, governments are called to play a role and strike the right balance between encouraging innovation and ensuring safety. Such a role may imply traditional regulatory oversight, which spans from ex-ante regulation to careful ex-post monitoring and regulatory vigilance (Scherer 2016).

At the same time, however, due to the inner features of AI systems and the structure of the AI market, standard legal interventions may fall short in maximizing social welfare and correct market failure. For this reason, decision-makers may rely more extensively on adaptive and flexible legislative schemes, which promote a shift in the way rules are designed and enforced and encourage a close collaboration between regulators and regulated entities, reducing information asymmetry and negative externalities. Recently, policymakers have opened the doors to experimental regulation by establishing mechanisms such as regulatory sandboxes and by envisioning experimental-friendly provisions in sectorial legislations such as the EU AI Act (Reg. (EU) 2024/1689) (OECD 2023).

At its origins, the use of experimentalism in regulation represented a way for institutions to channel their power, promoting self-regulation at the local level, while nowadays it is aimed at improving the quality of lawmaking, incentivizing innovation and, ultimately, promoting evidence-based policymaking (Ranchordás, 2021). In the framework of the AI Act, which devotes an entire chapter (Chapter VI) to measures

supporting innovation, regulatory sandboxes are designed to boost the EU market for AI systems, in alignment with Article 114 of the Treaty for the Functioning of the European Union (TFEU) on which the regulation is grounded, that aims to harmonize the internal market for AI. According to Articles 57-61 of the AI Act, firms will be able to test innovative products according to a testing plan that is agreed upon and monitored by national competent authorities. While this mechanism is expected to favor the internal AI market, questions remain about whether experimental tools like regulatory sandboxes will achieve their policy goals of fostering innovation and overcoming the constraints of command-and-control regulation, or if they will merely serve as a superficial exercise in appearances.

This contribution offers a review of the tool of AI regulatory sandboxes in the broader context of the governance of AI systems by assessing the conditions that need to be met for such regulatory sandboxes to achieve their envisioned policy goals. By examining how these sandboxes can balance the inherent trade-offs between innovation and safety, this paper proposes ways forward to mitigate potential agency problems and prevent regulatory capture, thereby enhancing the overall effectiveness and integrity of AI governance through experimentalism.

2. – INCENTIVES FOR AI PROVIDERS TO PARTICIPATE IN REGULATORY SANDBOXES

Regulatory sandboxes may play a crucial role in ensuring that AI systems are safe, ethical, and compliant with legal standards. In a policy prototyping experiment on the AI Act, more than 50 AI providers from Europe, mostly start-ups and small medium enterprises (SMEs), were queried about their willingness and motivation to join an AI regulatory sandbox (Andrade and Zarra, 2022). Almost all participants indicated that a sandbox environment could facilitate more responsible AI innovation and expressed a strong willingness and enthusiasm to engage in a regulatory sandbox. Firms highlighted that their main incentive for participating in such an environment was the opportunity to test their AI systems in settings that closely mimic real-life scenarios. Additionally, the chance to collaborate with regulators, ensure compliance, and help implement technical requirements was also noted as a significant benefit (Andrade and Zarra, 2022, 66). The

policy prototyping results provided practical insights into how firms perceive regulatory sandboxes and what they would seek in such an initiative. Drawing from these inputs and the existing literature, we can identify several incentives for an AI provider to enter a regulatory sandbox, which can benefit their development and market entry processes.

2.1. – Savings in compliance costs for start-ups and SMEs

Regulatory compliance is costly, especially for small businesses. A high-risk AI provider must complete several compliance activities before placing their system on the market. This includes building a quality management system, maintaining detailed technical documentation, conducting a conformity assessment, and registering their system in an EU database. Breaking down these activities in smaller tasks will result in additional responsibilities for employees, such as familiarizing with the new framework, filling out forms, checking data, implementing new review processes, holding internal and external meetings, filing documents, and initiating and attending training courses. The study supporting the AI Act's impact assessment (IA) (SWD (2021) 84 final) provides a stark illustration of the financial challenges that small businesses might face under the proposed AI Act, particularly when developing high-risk AI systems. For a small business with a workforce of up to fifty people and a turnover of EUR 10 million, the total compliance costs associated with deploying just one high-risk AI system could reach up to EUR 300,000 (Renda et al. 2021). This situation mirrors the challenges SMEs faced with the implementation of the GDPR (Freitas et al. 2018), leading to a disproportionate and cumulative burden of digital regulation costs on small market players. The costs of adhering to the requirements for high-risk AI systems could substantially affect SMEs' ability to innovate in the AI domain, potentially impacting their growth and sustainability, thus damaging the European AI market.

The AI Act aims to remedy this by establishing ad hoc safeguards and exceptions for small businesses, in particular SMEs and start-ups that qualify as providers and deployers of AI (Article 62 AI Act). The regulation's Impact Assessment maintains that besides receiving ad hoc guidance and training on the AI Act's regulatory framework, SME providers should have their interests and needs well considered, particularly when it comes to their compliance costs. For this purpose, in case they have a registered branch in

the Union, Member States will grant them with priority and free access to AI regulatory sandboxes, which should act as 'compliance facilitators'. In the sandbox environment, start-ups and SMEs will benefit from technical and legal assistance in adhering to the requirements, thus saving compliance costs. In support of this claim, some studies have estimated significant drops in compliance costs for start-ups and SMEs, stemming from their participation in testing and experimentation facilities (TEFs) and sandboxes supported by the European Digital Innovation Hubs (EDIHs) (Pellegrino et al. 2022). In such a scenario, the number of Full-Time Equivalents (FTEs)[1] needed would also decrease. Compliance costs would decrease if sandboxes were to offer trainings on AI Act requirements and guidance on establishing compliance processes. This would equip employees with the necessary skills to meet the requirements effectively. Furthermore, legal and technical support provided by stakeholders involved in the sandbox, including national competent authorities, notified bodies, and standardization organizations, could assist with tasks related to record-keeping procedures, document filing, and organizational changes.

2.2. – Network and spillover effects from participation

Regulatory sandboxes have the potential to guide the innovation process itself, ensuring it aligns with broader social, economic, and technological goals, as well as regulatory considerations (Ranchordás and Vinci 2024, 129). Participation in regulatory sandboxes affords AI providers a range of opportunities that extend beyond mere regulatory compliance. First, it allows for direct interaction with regulatory authorities and key stakeholders like standardization bodies, providing a chance to influence the development of industry standards and set best practices. Additionally, sandboxes constitute a unique platform for networking, fostering connections with other industry players and sparking new business partnerships. Perhaps most importantly, participation underscores a commitment to safety and responsibility, with consequent reputational gains. Furthermore, the sandbox environment allows for real-time feedback from regulators, which helps AI providers adapt their products to meet regulatory standards

[1] FTE is a unit of measure that indicates the workload of an employed person in a way that makes workloads comparable across various contexts. One FTE is equivalent to one full-time worker.

more efficiently. This iteration not only speeds up the compliance journey but also ensures that the solutions developed are aligned with legal expectations from the outset.

In the EU case, AI sandboxes will be mostly established at the national level, with competent authorities being the main interlocutors of AI providers. To fulfil their role, authorities must be equipped with the adequate multidisciplinary expertise. While at the centralized level the AI Office[2] will employ over 140 experts with diverse technical, legal, and economic skills to carry out its supervisory role, it is paramount that Member States ensure the same level of know-how for their resources, so as to reduce any knowledge gap.

An AI regulatory sandbox can serve not only as a secure platform for the exchange of best practices between industry and authorities. Undoubtedly, it also facilitates the sharing of know-how among participating firms and the establishment of a dialogue that may have positive spillover effects. This could result in new partnerships, acquisitions, and investments, particularly in the context of cross-border regulatory sandboxes with providers from different countries.

Finally, testing AI solutions in a controlled environment allows providers to identify and mitigate risks associated with their technologies before full-scale deployment. Within these environments, tools and processes can be developed, such as AI explainability methods and techniques for documenting the AI development process, contributing to the growth of best practices across the industry. This controlled testing also reduces the potential for unforeseen harms and enhances the reliability of their products. For instance, a proactive approach to risk assessment as part of a sandbox project helps in detecting vulnerabilities, biases, and blind spots early in the development cycle. The ability to pre-emptively identify and rectify issues enhances the reliability of the AI products and the AI provider, reducing the likelihood of post-deployment failures that could lead to costly recalls, legal liabilities, and, ultimately, reputational damage. Hence, as a member of the regulatory cohort, an SME can enhance its credibility in the eyes of

---

[2] Within Directorate A of DG CONNECT, the Office is structured into five units, reflecting the multidimensional nature of its assigned functions: excellence in AI and robotics (Unit A1), regulation and compliance (Unit A2), AI safety (Unit A3), innovation and policy coordination (Unit A4), and AI for social good (Unit A5). The Office will also benefit from the expertise of consultants, including a chief scientific advisor who will focus on overseeing general-purpose AI models, and an international affairs advisor responsible for global AI collaboration. More info available at https://digital-strategy.ec.europa.eu/en/policies/ai-office.

investors, customers, and regulatory bodies. This is particularly beneficial in an industry where safety considerations are increasingly critical. Furthermore, participation to a regulatory sandbox signals to the market that the start-up's technologies have undergone thorough scrutiny and validation, potentially reducing perceived risks associated with adopting their solutions. This can lead to stronger market positioning, attracting high-quality partnerships and investments.

2.3. – Accelerating access to market

One of the main objectives of AI regulatory sandboxes should be to expedite market access for participants, by removing certain barriers. In other sectors, such as fintech pharmaceutical regulation, regulatory sandboxes have reduced the speed of market approval ('time-to-market') for innovations, giving firms increased legal certainty and thus leading to greater overall innovation in society (Ringe 2023). Having been deployed widely across more than fifty countries, regulatory sandboxes in the financial sector have been reportedly quite successful in improving fintech's access to capital and competitiveness. In the UK,[3] for instance, firms participating in the sandbox experienced a 15% increase in capital raised after entry, a 50% higher likelihood of securing funding, and positive impacts on survival rates and patenting activity (Cornelli et al 2024).

The AI Act does not specify the exact types of barriers to be lifted. For instance, financial regulatory sandboxes often temporarily relax licensing requirements, thereby facilitating easier market entry for innovative firms. In the financial industry, offering certain financial products without the necessary financial license is often illegal. Fintech firms, which do not conform to existing regulatory frameworks, find it impossible to obtain these licenses. In such cases, regulators may grant a suspension of enforcement, allowing these firms to test their products within the regulatory sandbox under supervision. In contrast, legal barriers are not alleviated for AI sandboxes. At the same time, according to the AI Act, there is no prohibition on placing an AI system on the EU market if it complies with the regulations' requirements, meaning that successful participation in a regulatory sandbox is not a prerequisite for market entry. Thus, it can be argued that the regulation focuses on mitigating more practical barriers, such as the lack of technical

[3] For more info please see: https://www.fca.org.uk/firms/innovation/regulatory-sandbox.

and legal knowledge necessary for compliance with the provisions (Andrade et al 2023b, 22). However, as will be discussed later, in addition to receiving legal support within a regulatory sandbox, it is crucial for AI firms to have clear deadlines and a well-defined timeframe for these programs, as market and regulatory timing do not always go hand in hand.

In this regard, the AI Act stipulates that participation in the regulatory sandbox should be time-limited, yet it does not provide specific guidance on the duration of these programs. In fact, the timeframe will be defined in the implementing act, which is to ensure that it is 'appropriate' and extendable by national competent authorities. Legal certainty regarding the length of a program is crucial for a successful regulatory sandbox. This guarantees that participants who successfully complete the sandbox phase can transition into the market seamlessly. For instance, Article 60 of the AI Act allows for the possibility of testing high-risk AI systems outside the AI regulatory sandbox (in 'real world conditions'). The testing can be conducted for up to six months, with a possible extension of an additional six months. A detailed testing plan must be defined and submitted to the market surveillance authority in the Member State where the testing is to occur, with tacit approval granted if no response is received within 30 days. Similar terms could be formulated in the context of the implementing act.

The time required to transition an AI system from training to market placement varies based on factors such as the provider's business model, the type of product, and the sector involved. Consequently, the duration of an experimental testing project within a sandbox also varies. Drawing from experiences in other sectors and pilot programs in various jurisdictions, regulatory sandbox projects can range from a minimum of three months to several years. Regardless of the duration, the critical aspect is that the regulatory sandbox conditions should enable a faster deployment of the product compared to outside the sandbox. If this condition is not met, it indicates that the bureaucratic hurdles for the firm are greater within the sandbox, defeating its purpose.

3. – OVERCOMING LIMITATIONS AND AMBIGUITIES OF AI REGULATORY SANDBOXES

A specific challenge for AI regulatory sandboxes lies in making them sufficiently attractive

to AI providers. This will hopefully be addressed with the adoption of implementing acts by the AI Office. AI providers may potentially avoid the regulatory sandbox if the benefits are not clearly defined compared to normal market entry, and if the administrative burdens are perceived as prohibitively high.

3.1. – Lack of regulatory leniency and liability exemption

The AI Act does not envision an official presumption of compliance for participating firms, which may limit their attractiveness for firms. In other words, according to the text of the AI Act, regulatory sandboxes can 'facilitate' compliance, but mere participation does not certify the conformity of AI systems with all regulatory obligations, in line with most sandbox projects in other sectors. Certainly, streamlined regulatory processes reduce administrative hurdles, allowing firms to focus more on innovation and less on bureaucracy. However, the concept of time-to-market in a sector that will not be subject to the same stringent regulatory obligations as the financial or pharmaceutical one, has a different relevance, as firms may have less incentives to apply in absence of barriers to entry and given the dynamism of the competitive environment where they operate (Andrade et al 2023b, 27). Moreover, it should be noted that the AI Act regulatory sandbox does not allow for a liability exemption, hence firms remain liable under the applicable EU and national law for any harm inflicted on third parties as a result of the experimentation and testing taking place in the controlled environment. From this perspective, the regulatory sandbox would not allow for testing the level of exposure of certain AI products to potential liability, thus for instance discouraging providers of high-risk AI systems from applying (Truby et al 2022).

At the same time, however, there is leeway for interpretation of the text and flexibility, which could pave the way for a more lenient approach toward participants. Although the Impact Assessment of the regulation contends that no exemptions from the applicable legislation would be granted, taking into account the high risks to safety and fundamental rights and the need to ensure appropriate safeguards, the same Impact Assessment stresses that competent authorities would be granted certain flexibility in applying the rules within the limits of the law and within their discretionary powers when implementing the legal requirements to concrete AI projects in the regulatory sandbox.

In fact, Article 57(7) of the AI Act posits that, upon request, the competent authority will furnish AI system providers with written proof of their successful activities in the sandbox, along with an exit report detailing the activities, results, and lessons learned. Providers can leverage this documentation to demonstrate compliance with the regulation during conformity assessments or market surveillance activities. These exit reports and written proof should be favorably considered by market surveillance authorities and notified bodies, helping to expedite the conformity assessment process.

3.2. – Competitive disadvantage and risks of forum shopping

Creating conditions that are too favorable for participants in the regulatory sandbox may lead to a distortion of the level playing field (Andrade et al 2023b, 12; Ranchordás and Vinci 2024, 110). This is particularly true for the European AI market, which is highly fragmented. For this reason, the implementing acts proposed by the AI Office will be crucial in ensuring that member states and authorities harmonize eligibility and selection criteria at the union level. These acts will also standardize the procedures for application, participation, monitoring, and exiting the AI regulatory sandbox, including the sandbox plan, the exit report, and the terms and conditions applicable to participants.

Similarly, authorities should cooperate to mitigate the risks of forum shopping or 'sandbox shopping' due to a potential variability in approaches to sandboxes across jurisdictions. Forum shopping refers to the practice whereby firms seek to operate in jurisdictions that offer the most favorable regulatory environment. From this angle, diverging rules within these sandboxes across Member States might create distorted incentives for firms to choose jurisdictions with the most lenient or advantageous selection criteria or exit conditions. For instance, if one Member State's sandbox offers more relaxed compliance requirements or faster market entry processes, AI providers might prefer to base their operations in that jurisdiction, even if their primary market is in another Member State. This can undermine the harmonization efforts of the EU and lead to an uneven playing field, where firms in more lenient jurisdictions gain an unfair competitive advantage. This would in turn go against the very goal of the AI Act, namely the harmonization of the AI single market in the Union. Such a distorted incentive can be avoided through a cohesive centralized coordination by the AI Office at

the EU level as well as by a constant cooperation and communication between national competent authorities. In addition, to further curb any negative effects, it is important to note that the implementing act will define common and harmonized principles for participants across Member States, including eligibility and selection criteria, procedures for application, participation, and termination of the sandbox, along with harmonized terms and conditions applicable to participants.

3.3. – Risks of regulatory capture and window dressing

The regulatory sandbox environment, intended to foster innovation by offering a more flexible regulatory framework for innovative firms, may have counterproductive effects if not well-designed. It can inadvertently create conditions conducive to 'regulatory capture', a well-studied phenomenon in economics and social sciences (Stigler 1971). Regulatory capture can manifest concretely in a regulatory sandbox through the disproportionate influence of industry stakeholders in the decision-making processes. This can occur when representatives from the AI industry occupy key advisory roles within the regulatory sandbox, enabling them to shape the sandbox framework to accommodate their interests. As a result, the participants' profit goals prevail over public interest.

Another manifestation of regulatory capture occurs if certain firms are given a preferential treatment over others, particularly those with high lobbying and financial powers. These firms may receive more favorable terms in terms of conformity processes. Furthermore, there might be a risk of national competent authorities becoming dependent on the technical competences of the AI providers, which can lead to a form of intellectual capture. In this scenario, regulators may adopt the industry's perspective, potentially overlooking broader societal impacts or alternative viewpoints. Furthermore, regulatory capture can result in a lack of stringent enforcement of rules within the regulatory sandbox. This can lead to insufficient oversight of AI systems tested in the sandbox, allowing harmful practices.

Finally, on a related note, it is worth acknowledging that regulatory sandboxes can sometimes be used by firms and governments to project an image of innovation and proactive regulation without genuinely addressing the underlying issues associated with AI development and deployment. This can result in a mere window dressing exercise,

serving primarily as a marketing tool. While regulatory sandboxes are designed to foster innovation and ensure new AI technologies are tested and implemented safely, they might primarily enhance the reputation of those involved. Firms participating in these sandboxes could market themselves as pioneers in ethical and responsible AI, thereby gaining public trust and attracting further investments. Similarly, governments and regulatory bodies can showcase their commitment to modernizing regulations and supporting technological advancements.

To avoid the risks of regulatory capture and window dressing, it is paramount for the AI Office and the national competent authorities to establish robust mechanisms of accountability and transparency that would allow public scrutiny. The AI Act addresses some of these concerns by mandating national competent authorities to submit annual reports to the AI Office and the Board, beginning one year after the establishment of the AI regulatory sandbox and continuing annually until its termination, along with a final report, which would allow for an independent oversight. These reports must provide detailed information on the progress and results of the sandboxes, including best practices, incidents, lessons learned, and recommendations. To enhance transparency, these reports must be made publicly available online, and thus subject to public scrutiny. Additionally, the Commission is tasked with developing an interface containing all relevant information related to AI regulatory sandboxes, enabling stakeholders to raise inquiries with competent authorities. Finally, to further reduce the risk of regulatory capture, the implementing act could define and enforce clear and strict conflict of interest policies, mandating thorough vetting of prospective participants and other stakeholders. It should also ensure that the cohort of participating firms includes a diverse range of use cases, thereby balancing different interests.

4. – RECOMMENDATIONS FOR REGULATORS, COMPETENT AUTHORITIES AND PRACTITIONERS

4.1. – The need for an effective governance structure

Regulatory sandboxes might not be the 'silver bullet' that achieves all the goals set out by legislators, but they do constitute an important piece of the intricate AI regulatory puzzle (Andrade et al. 2023b, 30). To fully reap its benefits, an effective governance

structure is essential. In this setting, national competent authorities do not act as a paternalistic regulatory opponent, but as a partner that accompany the market entry of a new technology (Ringe 2023). In this regard, the adoption of a multi-layered and multi-stakeholder inclusive approach is key for the success of AI regulatory sandboxes in the EU. As emphasized by the AI Act, stakeholders should include providers of AI systems, especially SMEs and start-ups, the AI Office, national competent authorities and the European Data Protection Supervisor, which plays a key role in ensuring that the training and testing of AI models is compliant with the EU data protection regime. Moreover, it is key that also secondary stakeholders participate in the establishment and running of the regulatory sandbox. To this end, a broader range of entities will have to be involved, such as the AI Board, research and experimentation labs, the AI Scientific Panel, and various EU initiatives like the TEFs, the AI factories, the EDIHs, the AI-on-demand-platform (AIoDP) and the AI Pact[4]. Furthermore, involving national and European standardization organizations, notified bodies, and civil society organizations will enrich the governance structure with diverse perspectives and expertise.

To foster collaboration, the AI regulatory sandboxes should establish close ties with the AI Innovation Accelerator preparatory action. This integration can be facilitated by for instance including a representative from the Accelerator in the governance structure of the regulatory sandboxes. Finally, cross-border collaboration is particularly important for addressing the international nature of AI development and deployment and avoiding problems of forum shopping and fragmentation.

4.2. – Establishing clear conditions and communication channels

First, clear communication channels must be established to reduce the burden on SMEs and start-ups with limited legal and administrative capacities. Simplifying and communicating clearly these processes is crucial for encouraging participation. For instance, the call for applications should clearly outline the schedule of meetings, enabling participants to plan their resource allocation accordingly. A participant could report on their progress in the

[4] The AI Pact is a framework launched by the European Commission to prepare the implementation of the AI Act during the transitional period between its entry into force and the date of applicability. It convenes AI organizations who commit on a voluntary basis to implement key obligations of the AI Act ahead of the legal deadlines. More details available at https://digital-strategy.ec.europa.eu/en/policies/ai-pact.

testing on a regular basis, such as in weekly or monthly bilateral meetings. Additionally, it should specify the project deliverables, ranging from the initial testing plan to the exit report. Besides, monitoring and evaluation systems should be in place to continuously assess the performance of the testing. In addition, national competent authorities should be empowered to request additional information if needed as well as to carry out field visits.

To prevent unlimited tenure of firms within the regulatory sandbox, it is imperative to formulate clear exit criteria. AI providers should have the freedom to exit if they deem it necessary. Therefore, ad hoc mechanisms should be in place to allow them to leave the testing exercise, subject to an assessment by the national competent authority. At the same time, the national competent authorities should take a final decision about the success of the testing by a predefined deadline, potentially outlined in the implementing act. An essential exit criterion should be the participants' ability to meet the conformity assessment requirements upon exiting the regulatory sandbox. This assessment could be documented by the national competent authority in an exit report that details all activities conducted within the sandbox and their outcomes. The report could for instance include a 'positive mark' or a 'score' indicating the AI product's performance on each conformity requirement. Since regulators cannot exempt participants from complying with the AI Act requirements, fulfilling these criteria constitutes a necessary condition for a successful exit. This would ensure that participants can transition directly into the market in full compliance with the AI Act, thereby not merely expediting the conformity assessment process but also ensuring readiness for immediate market entry.

Finally, results from the regulatory sandbox should be generalized. If findings are too case- or participant-specific, the utility of the regulatory sandbox for society is limited, potentially favoring participants over non-participants. The regulatory sandbox should not 'pick winners' or provide indirect benefits solely to those within it but should also benefit society at large (Andrade et al 2023b, 25). Therefore, it is crucial that the results of regulatory sandbox activities are published.

5. – CONCLUSIONS

Regulatory sandboxes are instrumental in fostering innovation while ensuring compliance

with regulatory requirements. By providing a controlled environment for testing AI systems, they offer numerous benefits to firms, particularly SMEs and start-ups. In order to fulfil their goals, effective governance, robust technical implementation, and collaboration with various stakeholders are key to the success of these regulatory sandboxes. Based on the insights gained, this contribution has advanced several recommendations for policymakers and practitioners. Enhancing incentives and support mechanisms for SMEs and start-ups will encourage participation. Strengthening the governance framework to include a diverse range of stakeholders will ensure that regulatory sandboxes are effective.

While regulatory sandboxes and similar initiatives based on experimental governance principles are promising tools, they can result in limited outputs and scarce results in practice if not designed adequately. The upcoming implementing act will hopefully address some of the challenges by establishing clear criteria and conditions to facilitate communication and knowledge sharing while streamlining administrative processes, thus reducing the burden on participants. Moreover, to ensure a harmonized application of the AI Act and create a level playing field within the EU, cross-border eligibility for participation in regulatory sandboxes should be promoted. However, the risk of 'sandbox shopping' must be considered. In addition, guided by the implementing act, when contemplating the establishment of AI regulatory sandboxes, Member States should also consider alternative mechanisms to achieve the same goals. In sum, a well-structured regulatory sandbox presents a key avenue for the advancement of responsible AI systems. Although the creation of such a regulatory framework entails major investment in human resources and expertise, the derived advantages far overcome the initial costs. If properly designed and implemented, AI regulatory sandboxes as envisioned by the AI Act will undoubtedly contribute to Europe's competitiveness in the global technology race.

LEVERAGING TECHNICAL STANDARDS WITHIN AI REGULATORY SANDBOXES: CHALLENGES AND OPPORTUNITIES

ALESSIO TARTARO*, ENRICO PANAI**

SUMMARY

ABSTRACT

The AI Act relies on harmonized standards to provide cutting-edge technical solutions for implementing regulatory requirements, streamlining conformity assessment procedures, and offering a presumption of compliance for high-risk AI system providers. This paper examines the multifaceted role of standards within the AI Act, with a particular emphasis on their interaction with regulatory sandboxes. It begins with a comprehensive overview of the standards' role in the regulation, followed by an in-depth analysis of the synergy between regulatory sandboxes and technical standards. The study then explores the challenges and opportunities arising from the use of harmonized standards within the sandbox environment. Finally, it presents targeted recommendations for competent authorities, European standardization organizations (ESOs), and AI providers to maximize the potential of standards in regulatory sandbox environments. This research contributes to the ongoing discourse on effective AI governance and the practical implementation of the AI Act.

---

* PhD student at University of Sassari. Contact email: a.tartaro@phd.uniss.it

** Professor of AI & NLP in Decision Making at University Cattolica del Sacro Cuore of Milan. Contact email: enricopanai@gmail.com

1. – INTRODUCTION

The AI Act leverages the regulatory approach established by the New Approach (Council Resolution of 7 May 1985) and recently refined in the New Legislative Framework, or NLF (Regulation (EC) 765/2008, Decision 768/2008, Regulation (EU) 2019/1020). The NLF is a suite of measures designed to enhance the internal market for goods and streamline the placement of a wide range of products on the EU market. Currently, 26 pieces of legislation, encompassing products from medical devices to toys, are based on the NLF, with the AI Act poised to become the 27th[1].

Under the NLF, the legislation focuses on defining essential requirements that products must meet for accessing and enjoying free movement in the EU market. Technical specifications for implementing these essential requirements and achieving the legislation's objectives are delegated to European standardization organizations (ESOs), i.e., the European Committee for Standardization (CEN), the European Committee for Electrotechnical Standardization (CENELEC), and the European Telecommunications Standards Institute (ETSI). The Regulation 1025/2012 on European standardization and the Vademecum on European standardization[2] detail the process of how technical standards support EU legislation and policies[3].

Harmonized standards serve as a tool to demonstrate and achieve compliance with the relevant legislation. As outlined in the Blue Guide,[4] applying a harmonized standard

---

[1] For a full list of the legislative acts under the NLF, see https://single-market-economy.ec.europa.eu/single-market/goods/new-legislative-framework_en.

[2] Vademecum on European standardisation in support of Union legislation and policies, Commission Staff Working Document, SWD(2015) 205.

[3] This process involves several steps, beginning with EU legislators adopting a piece of legislation, such as a directive or regulation, which delegates the implementation of essential requirements to technical standards. The EU Commission then transmits a standardization request to the relevant ESOs, outlining the objectives, scope, and deadlines for developing these standards (Cucurru 2020). The ESOs take the lead and draft the technical standards, which are then evaluated by the Commission for compliance with requirements and alignment with EU legislation. Upon successful evaluation, the standards are published in the Official Journal of the European Union, marking their transformation into harmonized standards and the completion of the process.

[4] The 'Blue Guide' on the Implementation of EU Product Rules, Commission Notice, 2022/C 247/01, OJ C 247/1.

creates a presumption of conformity for a product or service with the corresponding legal requirements. The presumption of conformity is a legal concept used within the New Approach and the NLF. According to this principle, when manufacturers follow harmonized standards for their products, this creates a presumption that such products are compliant with the requirements of the relevant legislation. Conformity assessment procedures, often involving certification bodies, can further streamline this process.

Following the NLF approach, the AI Act adopts the same regulatory strategy, making harmonized standards crucial for the implementation of the regulation. This close linkage between standards and regulations supports the argument that 'standardization is arguably where the real rule-making in the Draft AI Act will occur' (Veale and Borgesius 2021).

Given the importance of technical standards in the AI Act, it is important to investigate what role they may play within AI regulatory sandboxes. As discussed in this paper, technical standards can indeed play a critical role. In order to support this view, the following sections will: a) identify three key functions of technical standards within the AI Act; b) analyze their interaction with AI regulatory sandboxes; c) explore the challenges and opportunities associated with the use of technical standards in the AI sandbox environment; d) formulate recommendations for leveraging technical standards within the AI regulatory sandbox.

2. – THE ROLE OF TECHNICAL STANDARDS IN THE AI ACT

In the AI Act, harmonized standards fulfill three critical functions (McFadden et al. 2021, Tartaro 2023a).

The first function is to provide state-of-the-art technical solutions to providers to ensure compliance with the requirements of the AI Act. In line with NLF's philosophy, the AI Act itself does not provide specific technical methods for implementing the essential requirements. For instance, Article 10 of the AI Act mandates data and data governance requirements to which high-risk AI systems must conform before entering the EU market. Among these requirements, Article 10(3) of the AI Act establishes that 'training, validation and testing data sets shall be relevant, sufficiently representative, and to the best extent possible, free of errors and complete in view of the intended

purpose'. However, the AI Act does not provide any technical means to implement and assess these characteristics of data sets (Maccabiani 2022). Let us consider the case of 'sufficiently representative' data. While the representativeness of samples is crucial for appropriate inferences in AI systems, there are many distinct interpretations of what data representativeness is as well as many different ways to measure it (Clemmensen and Kjærsgaard 2023). Providers seeking compliance with these requirements will not find an answer to navigate these issues in the AI Act. In other cases, standards are the sole method to characterize states that are beyond reach. For instance, the concept of being 'free of errors,' though operationally unattainable, should be seen as an ideal principle to strive toward, not a categorical one to be strictly enforced.

Here technical standards come into play. As highlighted in the standardization request, Annex I, 2.2(b), technical standards are tasked with providing '[...] specifications on quality aspects of datasets used to train, validate and test AI systems (including representativeness, relevance, completeness and correctness)'. These specifications are expected to offer concrete guidance to providers seeking compliance with the Act's essential requirements, including, in this case, requirements on data representativeness.

The second function of harmonized standards in the AI Act is to streamline conformity assessment processes. Conformity assessment is a procedure carried out under the provider's responsibility to verify an AI system's compliance with the requirements of the AI Act (Thelisson and Verma 2024). Conformity assessment can either be conducted by the provider itself (i.e., conformity assessment procedure based on internal checks according to Article. 43(1a) and Annex VI of the AI Act) or involve a third-party body (Article 43(1b) and Annex VII of the AI Act). Technical standards are central to both conformity assessment processes. Since conformity assessment relies heavily on inspecting the technical documentation compiled by providers (as mandated by Article 11), and this documentation must include the list of the harmonized standards applied by the provider (as per Annex IV(7) of the AI Act), these standards become an essential point of reference for assessing compliance of high-risk AI systems with the requirements of the AI Act. This aligns with common practices in safety legislation. Similar to the Medical Device Regulation (Annex II, 4c-d), which requires the specification of implemented standards to demonstrate compliance, technical documentation and harmonized standards serve as evidence that a product adheres to legislative requirements by adhering to relevant

standards.

The third function of the harmonized standards in the AI Act is to provide a presumption of conformity with the requirements of the regulation (Article 40 AI Act). Such presumption of conformity offers three main benefits to operators. Firstly, by complying with harmonized standards, providers can ensure their AI systems move smoothly within the EU market. These standards streamline the process by reducing administrative burdens. Since AI systems that follow the standards are presumed compliant, they do not face additional conformity assessments for those specific aspects covered by the standards. This translates to faster and more efficient market access within the European Union.

Secondly, the presumption of conformity grants providers additional legal certainty. Authorities are more likely to view AI systems favorably if they adhere to harmonized standards. This recognition of the standards as a reliable compliance measure gives providers greater confidence in their legal standing.

Finally, there is a potential, but debated, shift in the burden of proof (Portalier 2017). In case of challenges to an AI system from regulators, some argue that the presumption of conformity might shift the burden of proof to the authorities. This means that authorities might need to demonstrate an AI system's non-compliance, rather than requiring the provider to prove absolute compliance with every aspect of the regulation. These advantages highlight the crucial role of harmonized standards in ensuring a streamlined rollout of AI within the EU market.

3. – INTERACTIONS BETWEEN AI REGULATORY SANDBOXES AND TECHNICAL STANDARDIZATION

The AI Act does not directly address the connection between AI regulatory sandboxes and technical standards. However, it is important to note that both share two key objectives: fostering innovation and providing legal certainty.

Firstly, technical standards, like regulatory sandboxes, play a crucial role in encouraging innovation. They streamline the development and deployment process, making it easier for providers to bring their AI systems to the market. This streamlined approach fosters investment in AI technologies, leading to market growth and further

innovation. Secondly, like regulatory sandboxes, technical standards help increase legal certainty thanks to the presumption of conformity. They can offer a path for providers to ensure their AI systems align with legal requirements.

There is a key difference between AI regulatory sandboxes and technical standards, however. Using technical standards does not offer the same level of flexibility as regulatory sandboxes. Sandboxes provide a structured but adaptable environment where participants can test their AI systems with some regulatory leeway. In contrast, technical standards establish a predefined set of requirements, specifications, and best practices for AI systems, with limited room for deviation.

Regardless of similarities and differences, the AI Act acknowledges some points of intersection between AI regulatory sandboxes and technical standards. To begin with, the learnings and evidence gathered through sandbox experimentation are expected to inform and improve harmonized technical standards for AI. As stated in the impact assessment accompanying the AI Act proposal, 'The regulatory sandboxes [...] would also provide regulators with new tools for supervision and hands-on experience to detect early on emerging risks and problems or possible need for adaptations to the applicable legal framework or the harmonised technical standards.'[5]

Such learning opportunities can be leveraged as far as harmonized standards are used within the sandbox environment. However, the AI Act remains silent on the use of technical standards within sandboxes. Nevertheless, examples from other fields suggest that technical standards are indeed an important element in regulatory sandboxes. For example, the UK Civil Aviation Authority launched a 'Regulatory Sandbox for Unmanned Aerial Systems (UAS)'[6] in 2024. The validation and testing of equipment and procedures against technical standards within the sandbox environment is a key use case encouraged by the Civil Aviation Authority.

Similarly, the Malta Financial Services Authority requires applicants to their Fintech regulatory sandboxes to consider 'Regulatory technical standards and implementing

---

[5] Commission staff working document: Impact Assessment accompanying the Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts.

[6] Regulatory Sandbox for the development of capabilities to integrate Unmanned Aerial Systems (UAS) in unsegregated airspace.

technical standards'. [7]

Based on these examples, it is reasonable to expect technical standards to play a role within AI regulatory sandboxes as well. This expectation is consistent with the AI Act's requirements for competent authorities to offer, within the AI sandboxes, guidance, supervision, and support aimed at the identification of risks to fundamental rights, health and safety, as well as testing and mitigation measures to minimize these risks (Article 57(6) of the AI Act). Technical standards precisely serve this function. The future harmonized standard on risk management, for instance, will provide guidance on addressing and mitigating risks associated with high-risk AI systems. Similarly, other standards on human oversight, transparency, robustness, etc., are expected to provide guidance on how to implement these mitigation measures to minimize the risks associated with high-risk AI systems. Furthermore, Article 57(7) of the AI Act requires competent authorities to offer guidance on how to fulfill the requirements and obligations set out in the AI Act. Harmonized standards are exactly the type of guidance needed to meet these requirements, as they provide state of the art technical specifications to implement the requirements.

Finally, a last interaction between AI regulatory sandboxes and technical standards concerns the involvement of ESOs in the sandbox. Article 58(2)(f) of the AI Act mandates that the Commission's implementing acts specifying the detailed arrangements for the establishment, development, implementation, operation and supervision of the AI regulatory sandboxes shall ensure that relevant actors within the AI ecosystem are involved in the AI regulatory sandboxes, including standardization organizations.

These provisions leave little doubt that technical standards can and will play a significant role within regulatory sandboxes. The next section will delve deeper into the opportunities and challenges associated with the use of technical standards with AI regulatory sandboxes.

4. – CHALLENGES AND OPPORTUNITIES

While the previous section highlights the potential role of technical standards within AI regulatory sandboxes, it is crucial to address the benefits and challenges associated with

[7] The MFSA Fintech Regulatory Sandbox.

the use of harmonized standards.

One of the main challenges concerns the availability of technical standards before the first regulatory sandboxes are operational. In order to evaluate this challenge, we need to assess timelines to determine if harmonized standards can be developed in time.

According to Article 57 of the AI Act, Member States are obligated to ensure their competent authorities establish, or participate in, at least one AI regulatory sandbox at the national level. These sandboxes must be operational within 24 months of the entry into force of the AI Act, i.e., by August 2, 2026. However, some Member States may launch them sooner. Spain, for example, launched a pilot on AI regulatory sandboxes in partnership with the Commission in November 2023,[8] while the UK has already allocated budgets for establishing sandboxes.[9] Additionally, the Luxembourg Institute of Science and Technology (LIST) has set up an AI sandbox platform with tools for testing large language models (LLM).[10] While not yet full regulatory sandboxes, these initiatives demonstrate the eagerness to explore this space. Despite these promising developments suggesting earlier launch dates, however, let us conservatively assume that regulatory sandboxes will be operational by August 2, 2026. This timeframe needs to be assessed against the availability of technical standards.

The deadline for ESOs to deliver technical standards is set in the standardization request on January 31, 2025. However, additional time will be needed for the assessment of these standards by the Commission and their publication in the Official Journal. Since this can take around 108 days,[11] it is reasonable to assume it will be completed by August 2, 2026, if the January 2025 deadline is met.

Considering these timelines, there is initial optimism that technical standards might be available by the launch of regulatory sandboxes. However, the situation is

---

[8] Available at: https://digital-strategy.ec.europa.eu/en/news/first-regulatory-sandbox-artificial-intelligence-presented.

[9] Available at: https://techmonitor.ai/technology/ai-and-automation/government-backs-ai-regulatory-sandbox.

[10] Available at: https://ai-sandbox.list.lu/.

[11] See Report from the Commission to the European Parliament and the Council on the Implementation of the Regulation (EU) No 1025/2012 from 2015 to 2020, COM(2022) 30. Available at: https://eur-lex.europa.eu/legal-content/.

more nuanced. While the initial deadlines suggest timely standards, there are additional elements to consider.

Firstly, the Commission has only issued a draft standardization request. This is because an official standardization request can only be transmitted to the ESOs after the relevant legislation, the AI Act in this case, enters into force. Consequently, the final standardization request might adjust the deadlines, although it is unlikely that this will cause an 18-month delay, pushing them from January 2025 to August 2026.

Second, and more importantly, the ability of ESOs to meet deadlines should be carefully evaluated. As the following cases suggest, ESOs can struggle to deliver the required harmonized standards in time. For example, on December 1, 2020, the Commission sent a draft standardization request to ESOs regarding the energy labeling of refrigerating appliances to support Delegated Regulation (EU) 2019/2016 and Regulation (EU) 2019/2019. On that occasion, the Commission required the revision of four harmonized standards with a deadline set at December 31, 2022. On August 29, 2022, the Commission sent a final standardization request, confirming the same deadline (C(2022)5637 – Standardisation request M/585). Despite this, the revision of the four standards was completed only in April 2024, and three of them are still pending publication.[12] None of them are yet referenced in the Official Journal, and no harmonized standards are currently available for Delegated Regulation (EU) 2019/2016 and Regulation (EU) 2019/2019. A similar case regarding harmonized standards in support of the Radio Equipment Directive demonstrates the potential for delays.[13]

While these are just examples, they highlight the challenges that ESOs face in delivering harmonized standards according to the deadlines set by the EU Commission. These past experiences raise concerns about the ability of ESOs to meet the January 31, 2025, deadline for AI technical standards, potentially delaying their availability for the

---

[12] The four standards to be revised are: EN 62552-1:2020, EN 62552-2:2020, EN 62552-3:2020 and EN 60704-2-14:2013/A1:2019. Available at: https://standards.cencenelec.eu/.

[13] A standardization request regarding radio equipment in support of Directive 2014/53/EU and Commission Delegated Regulation (EU) 2022/30 was sent by the Commission on August 5, 2022. This request required three new standards on common security requirements for Internet connected radio equipment with a deadline of September 30, 2023. With an amendment on August 23, 2023, this deadline was extended to December 31, 2025. Currently, the three standards (FprEN 18031-1-3) are under approval.

launch of regulatory sandboxes.

Even if harmonized standards are delivered on time, another hurdle exists: will they pass the Commission's assessment in order to be referenced in the Official Journal? Past experiences with standardization for other legislation raise concerns. More often than not, standards proposed by ESOs do not fully align with the corresponding legislative acts. The Report on the Implementation of Regulation (EU) No 1025/2012[14] reveals a concerning trend. In the period between 2016 and 2020, only 57% of proposed harmonized standards were accepted by the Commission and referenced in the Official Journal. Around 36% of standards proposed by the ESOs for harmonization were rejected primarily due to inconsistencies with EU law and misalignment with policy and legal requirements. The remaining 7% of standards were either pending a decision or in the process of being cited. More recent data, however, suggests an even higher rejection rate, with figures approaching 80% (CEN-CENELEC 2023).

These examples highlight potential roadblocks. In addition, the AI Act's requirements extend beyond technical aspects, encompassing societal and ethical considerations like bias. Reaching a consensus on how to translate these multifaceted requirements into technical standards could be complex. This complexity could potentially hinder the standards' ability to fulfill the AI Act's requirements in a timely manner, ultimately impacting their chances of receiving a positive assessment from the Commission.

The availability of harmonized standards is a precondition for their use within AI regulatory sandboxes. However, even with timely delivery and publication, access to harmonized standards remains a challenge. This longstanding issue stems from the fact that these standards can only be purchased through national bodies, despite being recognized as 'part of EU law' in the James Elliot case (Colombo and Eliantonio 2017). Scholars and companies have long advocated for free access to harmonized standards (Ducato 2023). This principle of free access, however, clashes with CEN and CENELEC's claim to intellectual property rights and commercial interests. The European Court recently settled this debate in the Public.Resource.org case (CJEU 5.3.2024, C-588/21 P),

[14] Report from the Commission to the European Parliament and the Council on the Implementation of the Regulation (EU) No 1025/2012 from 2015 to 2020, COM(2022) 30.

recognizing the overriding public interest in freely accessible standards. This necessitates making harmonized standards freely available. In a promising development, the European Commission announced a free readability platform for these standards in May 2024, with an expected launch in June 2024.[15]

A final concern is the potential for a shortage of resources and expertise in specific domains to compromise the quality of relevant standards and hinder their implementation. This is particularly relevant in the field of fundamental rights (Tartaro 2023b). The standardization request highlights this concern as it emphasizes that relevant expertise in the area of fundamental rights should be involved in the standardization process in order to ensure due considerations of the fundamental rights implications of AI. However, technical committees often lack such expertise, being primarily composed of industry representatives focused on the technical aspects of AI systems. This raises concerns about inclusivity and representativeness within standardization bodies, which needs to be addressed to ensure comprehensive harmonized standards.

On the other hand, the shortage of expertise and resources can have a negative impact on the ability to apply technical standards within regulatory sandboxes. This is particularly relevant for small and medium- sized enterprises (SMEs). Unlike larger market players, SMEs often lack the financial resources to navigate the high costs associated with the implementation of technical standards. These costs encompass investments in equipment, testing procedures, and staff training. SMEs, which constitute the majority of European businesses, are particularly vulnerable due to their limited financial resources. Furthermore, the complexity of standards themselves presents a hurdle for SMEs. Lengthy and convoluted standards with unclear structures can be challenging to understand and implement. Technical jargon and overly complex language can further impede comprehension. SMEs may also encounter difficulties if the resources required to meet the standard's specifications are not readily available on the market, especially for those operating within national markets. Staying informed about updates and revisions can be another challenge due to a lack of clear communication. Finally, complex standardization areas such as AI may require additional support that may not be readily available.

---

[15] See the announcement by Director-General for Industry, Internal Market, Entrepreneurship & SMEs at European Commission Kerstin Jorna Available.

Despite the inherent challenges, the incorporation of technical standards into regulatory sandboxes presents a plethora of opportunities. These sandboxes offer a unique proving ground, not only for testing and refining AI technologies, but also for evaluating and advancing technical standards. This enables regulators to identify the need for new standards, recognize gaps in existing standards, and pinpoint areas where they may not adequately address specific AI risks. This aligns with the findings of the impact assessment accompanying the AI Act proposal, which is quoted in the first part of this paper.

The early engagement of ESOs within the sandbox also proves advantageous. On the one hand, ESOs can offer guidance on the implementation of the standards to providers. On the other hand, participants in the sandbox can provide valuable feedback on the standards themselves. This two-way communication gives rise to well-informed standards that cater to the specific needs and challenges of the real-world AI ecosystem. The collaboration between participants in the sandbox significantly enhances the overall relevance and effectiveness of technical standards, ultimately contributing to the development of a more robust and future-proof regulatory framework for AI.

Beyond standard development, the benefits of integrating technical standards into regulatory sandboxes extend to providers and potential providers within the sandbox. By collaborating with ESOs and national standardization bodies, they gain a deeper understanding of how to implement technical standards once they exit the sandbox environment. This acquired knowledge becomes crucial for demonstrating compliance with the requirements of the AI Act in the real world, and so it facilitates a seamless transition from the sandbox environment to the broader market.

5. – RECOMMENDATIONS AND CONCLUSIONS

The integration of technical standards into regulatory sandboxes presents a unique opportunity to foster innovation and mitigate risks associated with AI. However, to fully capitalize on this potential, it is essential to address the challenges that arise from this integration. This last section proposes a series of recommendations to competent authorities, ESOs and national standardization bodies, as well as providers or prospective providers, to ensure the effective utilization of technical standards in regulatory sandboxes.

First, competent authorities play a crucial role in facilitating the integration of

technical standards into regulatory sandboxes. To support this effort, we recommend that they develop clear and concise guidelines on the use of technical standards tailored to the specific needs of sandbox participants. These guidelines should provide explicit instructions on how to identify relevant technical standards, implement them within and outside the sandbox environment, and evaluate their effectiveness in mitigating risks. Furthermore, competent authorities should establish mechanisms for efficient information exchange among sandbox participants, particularly between providers and ESOs. This ensures that providers can seek guidance from ESOs on the implementation of standards and that lessons learned from sandbox projects can inform the standardization process, leading to technical standards that are readily usable for market operators.

Second, ESOs and national standardization bodies also have a critical role to play in the regulatory sandboxes. We recommend that they actively seek out and utilize the valuable insights and lessons learned from sandbox experiments to inform the development and refinement of technical standards. This information can be crucial for ensuring that standards are up-to-date, address real-world challenges, and effectively mitigate risks associated with AI. Moreover, ESOs and national standardization bodies should encourage early and active participation from sandbox participants in the standards development process, in particular regulators and providers or prospective providers, in order to facilitate the development of technical standards that effectively address the needs of the entire AI ecosystem.

Third, providers or prospective providers can benefit the most from the use of technical standards. We recommend that they engage actively in the standardization process, which can provide valuable insights into how technical standards can be applied within and outside the sandbox environment. Furthermore, providers should utilize the technical standards available within the sandbox to ensure that their AI systems comply with regulatory requirements and effectively address any potential risks they might pose. By doing so, they can learn how to achieve compliance with the requirements of the AI Act outside of the sandbox environment.

In conclusion, the timely development and adoption of high-quality technical standards will be crucial for operationalizing the AI Act requirements. The interplay between a traditional regulatory tool (standards) and an innovative one (regulatory sandboxes) holds promise for fostering innovation, ensuring compliance, and addressing

the complex challenges presented by AI. By harnessing the strengths of both, policymakers can create a balanced framework that promotes trustworthy AI while minimizing risks. As the EU continues to shape its AI regulatory framework, recognizing the critical role of technical standards in AI sandboxes is essential for ensuring the safe and responsible deployment of AI systems. Such an intersection of technical standards and regulatory sandboxes in the AI Act represents a pivotal moment in the evolution of AI governance in Europe. Ultimately, the success of this endeavor will depend on the ability of the involved stakeholders to reimagine the boundaries of innovation, regulation, and governance, and to create a future where AI enhances—rather than erodes—the core values that sustain our societies.

FROM THEORY TO PRACTICE - PRACTICAL CHALLENGES FOR BUSINESSES TO IMPLEMENT CYBER SECURITY RISK ASSESSMENTS AND HOW REGULATORY SANDBOXES CAN HELP

NILS BRINKER*

SUMMARY

ABSTRACT

The AI Act (AIA) contains provisions regarding cybersecurity, especially for high-risk AI systems, which usually require some form of risk management. This paper explains the difficulties in the practical implementation from the perspective of businesses, and in particular small and medium-sized enterprises. The paper argues that those difficulties primarily arise from the methodological ambiguity of Article 15 AIA, as well as from the high degree of abstraction of the provisions contained therein. It is argued that small and medium-sized enterprises in particular do not have the know-how to close the methodological inconsistencies and ambiguities conceptually. These consist mainly in the lack of a concrete protected goods as a reference point for a risk management process. The paper illustrates the concern that cybersecurity will be perceived as a pure end in itself, resulting in a laborious implementation, without actually benefiting operational IT security. Furthermore, there are practical problems

* Senior Cybersecurity Expert at intcube GmbH. Contact email: nils.brinker@intcube.io

in obtaining information due to the relative novelty of the technology. Since the majority of these problems stem from a lack of legal clarity, it is argued that regulatory sandboxes offer an excellent opportunity to function as a concretization mechanism, if they are used to develop and articulate practical guidelines and catalogues.

1. – INTRODUCTION

While the European Artificial Intelligence Act (AIA)[1] sets out to regulate AI systems (as it is defined in Article 3(1) AIA) as a relatively new technology, the topic of cybersecurity is also becoming increasingly relevant against the backdrop of an internationally evolving threat landscape. Therefore, it seems only appropriate that the AI Act also considers obligations related to managing cybersecurity risks. However, any regulatory provision is only as good as its practical implementation and vice versa. Ultimately, there is always an actor who must implement the theoretical requirements. This paper will discuss the practical challenges of implementing cybersecurity risk management requirements for the norm addressees, as well as how regulatory sandboxes provide an opportunity to address these issues. The focus here is on the pitfalls that arise from the vaguely formulated requirements of the AIA in their practical application, especially with small and middle-sized enterprises (SMEs) in mind.

To this end, Section 2 will first discuss the role of risk management as a regulatory tool for promoting cybersecurity. To illustrate the actual process, the steps of established risk management methods will be briefly explained. Section 3 will provide a theoretical overview of the cybersecurity regulations set out in the AIA. Section 4 will elaborate on the practical challenges of implementation. The focus will be on illustrating the difficulties for companies, especially SMEs, but also on how this can have a negative impact on the actual purpose of improving cybersecurity. Within this analysis it is argued how those issues can be addressed within regulatory sandboxes to develop practical solutions.

2. – RISK MANAGEMENT AS A REGULATORY MEANS

2.1. – Risk management as a means to foster cybersecurity

[1] Regulation (EU) 2024/1689 (Artificial Intelligence Act) further referred to as AIA

Risk management is an increasingly popular tool to regulate technology (see e.g. Article 21 Directive (EU) 2022/2555 (NIS 2 Directive), Article 10(2) CRA,[2] and Article 32 Regulation (EU) 2016/679 (General Data Protection Regulation)). From an IT security law perspective, the obligations to implement risk management fulfil a dual purpose. On the one hand, risk management is a generally recognized method in the field of IT security for identifying concrete measures to promote operational IT security (see Brinker 2024, p. 5).

From a regulatory perspective, a reference to risk management is also necessary as a proxy to be able to provide regulatory intervention to begin with. Cybersecurity is not a binary state of information technology that can be legally 'prescribed' in this context. If IT, as a subset of cybersecurity, is defined as the absence of events compromising the confidentiality, integrity, availability (CIA), and authenticity of information technology—or, depending on the definition, data as well (see Article 6(1) of the NIS 2 Directive, 'network and information system')—it must be acknowledged that achieving such a state with near-certainty is technically impossible. Even if the phrase '100% is impossible' is used as a cliché, this does not change the fundamental truth of the statement (see e.g. Craigen et. al. 2014; Sigmüller 2023). Information systems are too complex for it to be possible to completely prevent configuration or logic errors that could be exploited by a malicious attacker. However, the legislator can require proportionate measures. Rather than merely prescribing the 'security' of a system, it is necessary to take a more nuanced approach by implementing measures that adequately reduce the probability of risks. Risk management serves both to determine the operationally necessary 'sufficient' level and to demonstrate to the regulatory authorities that the utmost has been done in terms of sufficiency (Ritter 2023, p. 10). The same argument applies to the broader term cybersecurity, in which IT security is included as a crucial part, yet also includes further negative impacts, that do not necessarily violate the classic IT security protection goals (CIA triad).

2.2. – Established risk management methods

Most regulatory requirements that require the integration of risk management do not

[2] Text referred to as CRA is the adopted text by the European Parliament with its legislative resolution of 12 March 2024.

provide any definitive or deterministic specifications regarding methodology or the risks to be considered. In principle, there is a range of regulations that only implicitly require the implementation of risk management by requiring 'appropriate' measures, through to regulations that at least abstractly formulate process steps, risks to be considered, factors to be weighed up, and measures that are potentially necessary (Werner, Brinker, Raabe 2022).

The legislator therefore usually leaves the norm addressee some leeway regarding the methodology to be implemented, as long as it fulfils the established requirements and principles. In addition, however, there are established standards and best practices for the implementation of risk management, whose strict application is usually not mandatory for the fulfilment of legal requirements yet reduces the effort of argumentation when proving compliance. In the area of cybersecurity, these include ISO 27005 (ISO/IEC 2022) and, in Germany, the BSI Standard 200-3 (BSI 2017). To illustrate the practical problems that can be encountered when implementing these standards, the following section will briefly explain the approach taken by the ISO 27005 risk management standard.

As an established and recognized standard for risk management, the ISO 27005 procedure essentially consists of the following steps: a) selection of the risk management methodology and definition of the scope; b) context establishment; c) risk analysis; d) risk evaluation; e) risk treatment; and f) risk assessment.

Before embarking on the actual risk management steps, preparatory work must be done, such as selecting/determining the risk management methodology and defining the scope of application. In the context establishment step, the overall conditions will first be determined. In the risk identification step, an inventory of all potential risks will then be produced. These will be quantified in the risk analysis step, for example by determining values such as the potential amount of damage or the probability of an event occurring. Based on these values, the level/severity of the risks is assessed and prioritized in the risk evaluation step. Within this paper, the steps of risk analysis and evaluation are summarized and collectively referred to as risk assessment. In the risk treatment step, measures are then evaluated to mitigate or at least reduce the identified risks. Any residual risks must be accepted in the following step.

Especially the steps of risk identification and assessment rely on input e.g. of occurring threats or values for risk evaluation. Sources of those input are commonly

standard catalogues, external information sources like threat intelligence or the experience of the risk manager.

3. – RISK MANAGEMENT OBLIGATION OF THE AIA FOR HIGH-RISK AI SYSTEMS

The following sections illustrate the risk management obligations within the AI Act for high-risk AI systems. It should be noted that the requirements governing the otherwise prohibited use of real-time biometric identification systems, as outlined in Article 5(2) of the AIA, could also be interpreted as a form of risk management. However, since those requirements apply specifically to law enforcement, they do not need to be fulfilled by private actors and, therefore, are outside the scope of this paper.

3.1. – General risk management obligations of Article 9 AIA

Providers of high-risk AI systems are obliged to implement risk management for cybersecurity risks, in particular by Article 15(5). However, Article 9 of the AIA provides a framework for a general risk assessment. According to Article 9(2) of the AIA, risk management is to be understood as a 'continuous iterative' process that is 'planned and implemented throughout the entire life cycle of the AI system'. The AIA does not prescribe a conclusive methodology here as well, but merely formulates guidelines with regard to the procedure, the risks to be considered and the risk acceptance level.

Regarding the procedure, the AIA prescribes a 'continuous and iterative process' over the 'entire life cycle of the AI system'. Within established standards such as ISO/IEC 27005, this is usually the case. Nevertheless, clarification within the AIA is to be welcomed, as it is fundamentally based on product safety law. Therefore, this clarification is helpful in understanding that the obligation to carry out risk management does not end when a product is placed on the market (Reusch 2023, p. 156).

With regard to the risks to be considered within the risk identification step, these are, according to Article 9(2)(a) of the AIA, those that are 'reasonably foreseeable' for 'health, safety or fundamental rights'. The AIA formulates specific protected goods here. However, these are formulated at such an abstract level that only a small limitation of the scope of practical risk identification is achieved. Furthermore, the AIA does not specify

the scope of persons potentially affected by a risk. Therefore, not only must the risks for users of an AI system be considered, but also any third party that could be affected by an AI system.

Furthermore, only those risks that arise from the use of the AI system in accordance with its 'intended purpose' or, in accordance with Article 9(2)(b) of the AIA, 'reasonably foreseeable misuse' are to be taken into account. Limiting the potential applications by defining the purpose of the system makes it possible to limit the scope of the analysis, which could otherwise potentially remain uncertain. However, the obligation to consider potential misuse prevents the operator of an AI system from avoiding any liability by skilfully formulating the purpose.

In addition, the AIA requires the consideration of empirical values in accordance with Article 9(3). A test phase is also required by Article 9(6) AIA.

3.2. – Management of cyber risk in Article 15(5) of the AIA

In principle, cybersecurity risks would already be included in the obligation to ensure general risk management in accordance with Article 9 of the AIA. However, Article 15 of the AIA articulates further requirements in terms of accuracy, robustness and cybersecurity, at least with respect to the AIA understanding (Nolte and Schreitmüller 2024). It should be noted here that Article 15 of the AIA explicitly refers to the definition of 'cybersecurity' in the sense of Article 2(1) Regulation (EU) 2019/881 (Cybersecurity Act), as is the case with the Cyber Resilience Act. This definition is broader than, for example, the definition of the 'security of network and information technology', to which the obligations of the NIS 2 Directive for operators of critical infrastructure refers. The definition of 'cybersecurity' does not explicitly refer to the classic protection goals such as confidentiality, integrity and availability (CIA), but rather encompasses the prevention of any risks that may arise when using information technology (see Article 2(1) and (2) Regulation (EU) 2019/881 (Cybersecurity Act)).

In general, Article 15(5) of the AIA is formulated in an unsystematic way. It seems that the intention is not to provide concrete methodological guidelines, but rather to provide a collection of bullet points of factors that should be considered in relation to cybersecurity (Nolte and Schreitmüller 2024; Bomhard, Siglmüller 2024, p. 49).

For example, Article 15(5) of the AIA requires that high-risk AI systems be 'resilient against attempts by unauthorised third parties to alter their use, outputs or performance by exploiting system vulnerabilities'. First of all, it should be noted that a loose reference is made here to the classic protection goals of information security[3] (change of use or output as a violation of integrity, or change of performance as a violation of availability, Nolte and Schreitmüller 2024). However, there is no explicit reference to the classic CIA protection goals, yet the enumeration of scenarios to be avoided could be interpreted as an implicit reference. The enumerative style in conjunction with the absence of an explicit reference in recital 76 of the AIA, suggests that the information security protection goals are actually not intended. The alteration of 'use, outputs, or performance' rather refers directly to the purpose of the AI to be determined by the operator, according to Article 4 (12) of the AIA.

The use of the wording 'exploiting system vulnerabilities' also suggests that the AIA distinguishes between AI applications themselves as technical services and systems as the resources used to provide them (Werner, Brinker, Raabe 2022). However, this understanding is again undermined by the fact that Article 15(5)(3) of the AIA explicitly refers to 'AI specific vulnerabilities'. Threats such as 'data poisoning' or 'model evasion' must be considered here, which in turn relate specifically to AI as a technology. Nevertheless, classic attacks aimed at altering stored training data, as well as AI-specific attacks, are imaginable, in which for example learning systems are constantly manipulated by repeated inputs. This means that both the AI technology itself, the technical resources used to provide it and the development environment are potentially included (Nolte and Schreitmüller 2024). In the end, this inconsistent use of terms leads to uncertainty regarding the scope to be included in a risk assessment.

The term 'AI specific vulnerabilities' itself is also fundamentally imprecisely defined. Even with the listed example threats such as 'data poisoning' in self-learning systems, it can be argued that this is not a vulnerability, but an inherent functionality of AI models (Nolte and Schreitmüller 2024). In this case, the fault might not be seen as the implementation of AI as a means, but rather in the unsuitability of AI for a specific

---

[3] As used within the definition of the 'security of network and information systems' in Article 6(2) NIS2; for a definition see e.g. NIST 2024.

application. Appropriately, the potential risks for the application itself are taken as a reference point here. However, this reference is not systematically outlined. In practice, this creates confusion as to which threats must be considered in concrete terms.

Furthermore, Article 15 does not explicitly prescribe risk management as a methodological obligation. However, the necessity of implementing risk management or a related method arises from the requirement to take measures that are 'appropriate to the circumstances and risks'.

Even though Article 15 formally exhibits systematic shortcomings, some basic ideas are not wrong. It shows that AI-specific threats, such as 'data poisoning', are real threats that can affect the intended functionality of an AI system without necessarily violating the classic CIA protection goals. Nevertheless, Article 15 remains dogmatic and methodologically incoherent and overall insufficient. In the context of the practical implementation of risk management, the main difficulties arise from legal and methodological uncertainties, as a result of these shortcomings.

A large number of AI applications also fall within the scope of the CRA, which also includes risk management for cybersecurity risks (See Article 9(2) of the CRA in conjunction with Annex III). Nevertheless, 'AI-specific cybersecurity risks' must be taken into account within the risk management of the CRA. Although the requirements in the CRA for risk management are much more specific, this does not apply to 'AI-specific risks', which means that the methodological problems of Article 15 are inherited.

4. – CHALLENGES FOR THE PRACTICAL IMPLEMENTATION OF RISK MANAGEMENT AND OPPORTUNITIES OF REGULATORY SANDBOXES

This chapter elaborates on the challenges faced by the norm addressee when attempting to conduct a cyber risk assessment as required by the AIA, and how insights from a regulatory sandbox can help address these challenges. To this end, the significance of regulatory sandboxes, in particular with regard to the concretization of abstract requirements, is first evaluated. The further description of specific challenges is based on the steps of risk management as laid down in Chapter 2.2. Not all of the described challenges occur solely within the risk management step in which they are categorized within this paper. They can also impact other steps of the risk management process in different forms. However,

the formal steps of risk management should serve as a rough outline.

4.1. – General practical challenges and how regulatory sandboxes can help

The fundamental challenge in the practical implementation of risk management requirements is the high degree of abstraction and eclectic methodology requirements of the AIA, which will be discussed in detail in the following sections.

In particular, the high degree of abstraction of individual regulations is due to the deliberate concept of the AIA as a 'principle-driven' regulation. The AIA deliberately refrained from formulating overly specific requirements, but instead sets up abstract categories for fulfilling the requirements (Arjoon 2006, p. 58). This is not a wrong approach, especially in relation to constantly evolving technologies whose concrete form and type of use cannot currently be finally assessed. The AIA also contains a number of mechanisms for concretizing its provisions for specific use cases, in addition to general mechanisms such as case law (Schuett et. Al. 2024, p. 33 ff.).[4] Those are e.g. harmonized norms as laid down in Article 40 of the AIA or common specifications as laid down in Article 41 of the AIA (see also EPRS 2022). However, there is currently a challenge in that these concretization mechanisms have not yet been put into practice (Ebers 2021; Micklitz 2023).

Regulatory sandboxes can also be understood as a type of concretization mechanism, or at least as a source of information for such mechanisms. They can serve as test laboratories for the general requirements of the AI Act itself to be concretized in a use-specific, evidence-based manner, as it is explicitly stated of goals of the sandboxes in Article 57(9) of the AIA. These challenges can serve as touchstones in the context of regulatory sandboxes, to work towards concrete regulatory cost learnings.

4.2. – Challenges of the preparation phase: general legal uncertainty

The AI Act does not prescribe a specific methodology for either general or IT security-specific risk management. For the norm addressee, the consequence is a necessity to develop methodologies, in particular for risk management, which meet the abstract

[4] Cf. a similar development in regard to the GDPR.

requirements of the AIA themselves. However, this is not a trivial matter, especially for SMEs. The specific uncertainties are discussed in the following sections.

4.2.1. – Selection of a risk management method

The difficulty begins with the selection of a basic methodology for implementing risk management. As described above, the AIA does not prescribe a formal methodology, but only eclectic requirements.

In terms of cybersecurity, there are established standards such as ISO 27005 (ISO/IEC 2022), which provide structured guidance on how to implement risk management. The ISO/IEC 42001 Standard is also being developed as a specific standard for AI standardization, and hence remains abstract in terms of cybersecurity risks (For AI Standardisation in general, see Ebers 2021). However, for SMEs in particular, a word-for-word implementation of these standards is costly and elaborate, so it might exceed existing resources. Standards such as ISO 27005 primarily contain organizational requirements for risk management that are more geared towards the structures of large companies, even though they argue to be generally suited for every company size. To a certain extent, implementation requires the establishment of a parallel 'compliance' organization, as the overall effort involved exceeds that which could be handled in the course of day-to-day business operations.

Nevertheless, it is possible to orientate towards the basic approach of these standards, adapted to the existing structures of SMEs. The explicit goal then is not to obtain certification.[5] Certifications can be a tool for compliance verification, but they are not explicitly required by law in all cases.

Nonetheless, modifying existing standards or developing a completely individual approach would mean leaving the predetermined path. Both approaches involve a conceptual effort, whereby a certain degree of legal uncertainty always remains. Assuming that the competence of SMEs in the field of AI development lies primarily in technical development, SMEs face the problem that this conceptual effort may exceed their know-how.

While participating companies also start on an equal playing field, regulatory

[5] See e.g. DIN SPEC 27076 as an approach to tailor elaborate certifications to the needs of SMEs.

sandboxes offer the potential for a closer collaboration between companies and authorities to figure out a method that meets regulatory requirements. For those insights to have a full-scale impact it is necessary that the lessons learned are shared. This can either be done by the participating companies themselves as multipliers (e.g. though professional networks and communities) or by the supervisory authority publishing guidelines. The latter has the added advantage of creating greater clarity regarding the regulatory authority's expectations.

4.2.2. – Scope

As described above, the AIA, like the CRA, relies on the definition of 'cybersecurity', which does not require a strict violation of the classical protection goals (CIA) for a threat to be within the scope.

Depending on the circumstances, e.g. the organisational structure of a company, this broad definition might cause practical difficulties to systematically demarcate the scope. This can be illustrated by using the threat of 'data poisoning' as an example to highlight the issue of demarcating the operation of an AI system as a service from the company's IT infrastructure. In the case of non-continuously learning models, protecting the integrity of the data would inevitably mean expanding the scope to include the development environment in which the training data is stored. If the development environment is not properly separated from the rest of the company's IT, the scope would have to be extended to include the entire company's IT. The necessity of this is also supported by recital 76 of the AIA which states that 'the underlying ICT infrastructure' should be 'taken into account'.

It should be noted that although additional implementation effort represents a difficulty for companies, it does not necessarily indicate an impermissibly harsh regulatory approach. In the case described above, the scope can be reliably reduced by properly separating the development environment from other company IT.

Within regulatory sandboxes therefore, authorities should evaluate which practical scope it recognizes as necessary for a specific use case. Ideally, criteria for delimitation should be evaluated and communicated based on practical experience.

4.3. – Challenges in the risk identification and assessment step

The following section will examine the challenges that occur within the process of risk management itself. Here, the main issue is the lack of legal clarity, which leaves the norm addressees uncertain of the requirements they are expected to fulfil. However, there are also practical issues especially with threat identification, which are rooted in the novelty of AI as a technology.

4.3.1. – AI specific threats as a technical novelty

A further practical challenge with the broad scope in combination with the relative novelty of AI, is the identification of threats to be considered.

This requires a basic knowledge of potentially possible events.[6] However, the nature of a relatively innovative technology such as AI is that the potential attack vectors are only being researched. The scientific field of cybersecurity of AI systems is a relatively new field of research. Although lists of potential threats to AI systems exist (see OWASP 2024), some of these are already listed in Article 15(5) of the AIA. Nevertheless, it can be assumed that these are not exhaustive lists. In contrast to 'classic' IT security risks, there is therefore a lack of experience, while standard catalogues can be assumed to be work in progress,[7] which increases the effort of the practical threat identification.

It therefore makes sense to integrate application-specific threat modelling within regulatory sandboxes, or to support the participating companies conducting such a process.

4.3.2. – The lack of definition of protected goods as a reference point for a risk analysis

A fundamental weakness in IT security law is the widespread lack of explicitly formulated protected goods (Brinker 2024, Werner, Brinker, Raabe 2022). Article 15(5) of the AIA does not clarify this aspect either (Bomhard and Siglmüller 2024, p. 53). It only formulates the alteration in the use, output or performance of an AI system as a negative

---

[6] For which may exist standard catalogues. See for example the standard catalogue of threats within Annex 3 ISO/IEC 27005.

[7] For an example of an attempt to create a comprehensive risk library, see MIT AI Risk Repository 2024.

consequence to be avoided. The omission of specific protected goods - at this point - is not necessarily wrong, since AI systems as means cannot directly affect protected goods anyway. This only happens through their use within a specific application (or service) (Werner, Brinker, Raabe 2022). Nevertheless, an explicit reference to the application or to the affected protected goods as a point of reference is missing. Unless cybersecurity is to be understood as an end in itself, corresponding protected goods must therefore be derived beyond the wording of the law. This can be done, for example, with regard to the purpose of the AIA articulated in Article 1. Likewise, a systematic reference to Article 9 of the AIA can be drawn. Both Article 1 and Article 9(5) of the AIA abstractly formulate the protection of health, safety and fundamental rights as a reference point. The difficulties arising from this high level of abstraction will be explained in Section 4.2.3.

The protected good serves as an essential point of reference for risk identification and risk assessment. A risk can only be identified and assessed if it is clear 'what' it is a risk to. From a practical perspective, the problem is that it is unclear which risks need to be considered at all, which in turn leads to legal uncertainties. Theological or systematic approaches to interpretation are of limited help from a practical point of view, as they are complex, especially for SMEs (Bomhard and Siglmüller 2024, p. 54). As a result, a company conducting a risk assessment will address cybersecurity in some manner but will need to invest significant effort into determining what is actually required by law.

This is also an issue for the effectiveness of the risk management process in terms of promoting operational cybersecurity. Without a protected good, systematic demarcation is not possible, and the risk inventory remains not defined in its practical implementation. The result of such a risk assessment thus remains relatively random. From a regulatory perspective, cybersecurity as an end in itself is therefore an empty exercise.

While this is a systematic weakness within the IT security law, regulatory sandboxes can help to identify and address concrete protection goals, and to communicate them e.g. through guidelines by the authority.

4.3.3. – A formulated protected good: the issue of abstraction

As described above, the weakness regarding the non-formulated protected interests of Article 15 of the AIA can at least partly be circumvented by a systematic inclusion

of Article 9. Although protected interests are articulated here, their high degree of abstraction contains its own particular pitfalls. According to Article 9 of the AIA, risks to health, safety, and fundamental rights must be taken into account, particularly (but not exhaustively).

Health risks can be relatively easily derived from the purpose of the AI system or its functionality in practice and should not pose too great a problem even for legally untrained users. This is more difficult for safety (see e.g. Amedei et. al. 2016).

For the best illustration of the problems of abstractly formulated protected interests, however, it is useful to look at fundamental rights. Although fundamental rights are universally valid, i.e. they affect everyone, the methodology for determining an impermissible infringement is not trivial. Fundamental rights are formulated abstractly by design (Janssen et. Al. 2022, p. 209 f.; compare also the abstraction level of the Ethics guidelines for trustworthy AI, EU 2019). However, if no standard catalogues are available, scenarios in which fundamental rights are infringed upon must be identified by the norm addressee himself. In practice, specific scenarios would have to be devised and examined for an infringement.

For developers of AI systems and private companies, the fundamental rights audit is usually not part of their day-to-day business. Accordingly, SMEs in particular lack the know-how to carry out such an audit properly. As a result, also the formulation of protected goods again leads to an eclectic implementation of the risk management process, if the degree of abstraction is too high.

Parallel to concretization, regulatory sandboxes can help to identify concrete scenarios where the specific AI use case of the sandbox affects fundamental rights. Those could be included in standard catalogues of threats or risks that could support entities in their practical risk assessment. Besides catalogues, practical and fundamental rights assessment methods suitable for the target group could be developed (compare e.g. similar publications by the data protection authorities concerning the data protection impact assessment, DSK 2018).

5. – LONG STORY SHORT: IN PRACTICE CLARITY IS KEY

The problems described above in the context of the practical implementation of risk

management regarding cybersecurity can be summarized as follows:

a) eclectic method requirements and a high degree of abstraction in the text of the AIA lead to uncertainties regarding practical implementation. The major point here is the lack of explicitly formulated protected goods.

b) SMEs in particular lack the resources to fill the gaps in the law conceptually.

c) There is also a risk that the risk management process will remain completely ineffective due to an eclectic and arbitrary execution with regard to the goal of achieving an appropriate level of cybersecurity. This would result in a classic 'compliance before security case', which leads to nothing but too much paperwork and, in particular, undermines the actual protective purpose of the AIA.

Although the AIA was purposely formulated in an 'principle driven' manner, 'technology openness' should not be confused with randomness. The AIA act might cross this fine line especially in regard to cybersecurity obligations, if its integrated measures for concretisation are not used properly.

However, regulatory sandboxes offer an excellent opportunity to fill the void caused by legal uncertainty and gaps described above based on practical experience. In order to be successful, it is, therefore, important to actually use this opportunity to formulate explicit methods for risk assessment, clarify protected goods, and evaluate concrete risks and threats.

A COMPARATIVE ANALYSIS OF REGULATORY SANDBOXES FROM SELECTED USE CASES: INSIGHTS FROM RECURRING OPERATIONAL PRACTICES

FABIO SEFERI*

SUMMARY

ABSTRACT

This contribution presents a comparative analysis of regulatory sandboxes and experimentation practices across various jurisdictions and sectors, focusing on selection criteria, procedural phases, and reporting obligations. By examining 87 use cases from financial services, digital, energy, and other industries, the analysis identifies key patterns and best practices for the operation of regulatory sandboxes. Primary criteria for participation include the innovative value of the proposed solution, societal benefits, and market and testing readiness. The regulatory sandbox participation process typically involves an application phase, testing and experimentation under controlled conditions with continuous monitoring by authorities, and an exit phase to ensure a smooth transition from experimentation to standard market conditions. The findings offer valuable insights into the design and implementation of regulatory sandboxes, providing a framework for policymakers and regulators to enhance the effectiveness of these initiatives.

---

* PhD Candidate in Cybersecurity at the IMT School for Advanced Studies Lucca and the University of Florence, Italy. Contact email: fabio.seferi@imtlucca.it.

1. – INTRODUCTION: RATIONALE AND GOALS OF THE COMPARATIVE ANALYSIS

Given the rapid and widespread adoption of regulatory sandboxes as 'schemes that enable firms to test innovations in a controlled real-world environment, under a specific plan developed and monitored by a competent authority' (European Commission 2023a, 599), it becomes increasingly important to understand how they operate in practice. This is even more the case if we consider that by 2 August 2026, the national competent authorities of EU Member States are required to establish national artificial intelligence (AI) regulatory sandboxes – as mandated by Article 57 of the AI Act (Regulation (EU) No 2024/1689).

In this regard, the primary objective of conducting a comparative analysis of established regulatory sandboxes is to gain relevant insights into several aspects of these frameworks across different jurisdictions and sectors. By systematically examining various initiatives, the overall analysis identifies best practices and emerging trends, thus facilitating informed decision-making and policy formulation, in particular for future rules on the establishment and functioning of regulatory sandboxes.

In this view, the contribution first examines the selection criteria typically used to evaluate applicants for admission to regulatory sandboxes. It categorizes these criteria into three distinct groups based on their frequency of occurrence across the use cases analysed. This approach provides decision-makers with an empirically grounded basis to prioritize specific criteria against others when defining the regulatory sandbox framework.

Secondly, the analysis highlights key insights for the structuring of the operational phases of participation to regulatory sandboxes. This phase encompasses the critical stages applicants undergo, including the application process, the testing or experimentation phase conducted under controlled conditions with regulatory oversight, and the final exit stage.

Finally, the contribution presents a set of concise conclusions, highlighting the most significant findings and implications to offer a better understanding of regulatory sandboxes' operation.

2. – A BRIEF METHODOLOGICAL NOTE

The use-cases selection and data collection for the comparative analysis entailed a multifaceted approach to ensure broad geographical coverage and relevance. Firstly, the selection of relevant use cases was primarily focused on countries within the European Economic Area (EEA) to ensure alignment with regional regulatory frameworks and market dynamics. This also provides a first comprehensive understanding of sandbox initiatives within the European context. Additionally, to capture diverse perspectives and experiences, use cases from other geographies were also included, particularly from regions with advanced regulatory frameworks or notable innovations in tech-enabled sectors.

In general, data collection involved open-source research, analysis of relevant websites, and evaluation of legal documents and guidelines pertaining to each initiative.

The comparative analysis has comprised a total of 87 relevant use cases, pertaining to the following sectors: financial services industry (50 cases, of which 3 with a focus on insurance-related solutions); cross-sectoral (12 cases, of which 6 with a focus on emerging technologies such as AI); transportation (7 cases); energy (6 cases); data protection and management (6 cases); healthcare (3 cases); telecommunications (3 cases). The overall list of the 87 use cases selected can be found in Table 2 in the Annex. However, it is not exhaustive of all regulatory sandboxes and experimentation practices activated across geographies and sectors.

Moreover, depending on the core features evaluated, not all listed use cases are considered as regulatory sandboxes, but more as a regulatory experimentation mechanism.[1] For example, 'France - CNIL's Personal Data Sandbox' clarifies in its governing rules that it is not a 'regulatory' sandbox since there is no derogation from existing rules, considering that its remit falls within data protection and privacy regulations. Nevertheless, these use cases also provide practical insights when it comes to the selection criteria or the operational phases which may be considered when defining the framework of regulatory sandboxes.

---

[1] On this specific point, see also the contribution of E. Longo and F. Bagni included in this work.

3. – IDENTIFICATION OF THE SELECTION CRITERIA

The analysis begins with an examination of the selection criteria, which are the key parameters that determine whether applicants are admitted into regulatory sandboxes. This is particularly important also in view of the future implementing acts on AI regulatory sandboxes, that must also define common principles for their definition (refer to Article 58(1) of the AI Act[2]). By selectively admitting projects based on specific requirements, regulatory sandboxes ensure that only those solutions that support some underlying values or policy goals are given the opportunity to develop under controlled conditions: the flexible scheme is applied only to a set of projects, which better align with the objectives of the competent authority (i.e., the public decision maker).

The analysis initially focused on identifying the selection criteria highlighted in specialized studies and reports, including from the European Commission (2023b), the OECD (Attrey et al. 2020), the European Parliament (Parenti 2020), the Joint Research Centre (Gangale at al. 2023) and the European Supervisory Authorities (2023). Such criteria included the innovative character of the project, public interest or alignment with broader policy objectives, maturity of the project, safeguards for consumers, identification of regulatory barriers, time limit, specific authority mandate, need for testing, and increase of legal certainty. Additional ones were also considered, despite not being explicitly mentioned in the examined reports, as they emerged across multiple use cases. These criteria included reputation, employment of a specific technology, and proposal clarity and completeness.

The selection criteria were refined to ensure they were mutually exclusive and collectively exhaustive. To facilitate a structured analysis, they are assorted into three groups based on frequency of occurrence: (a) more than 50 occurrences, (b) 26 to 50 occurrences, and (c) 1 to 25 occurrences. Table 1 below shows a synthetic view of the criteria analysed, and their respective occurrences.

---

[2] The implementing acts on AI regulatory sandboxes will include common principles on the eligibility and selection criteria for participation.

Table 1. List of the selection criteria analysed and respective number of occurrences

ID	Selection criteria	Occurrences
A	Degree of innovativeness	74
B	Public interest or societal benefit	67
C	Level of maturity	63
D	Risk management mechanisms	43
E	Authority mandate or remit	33
F	Need for testing or experimentation	31
G	Presence of legal barriers	22
H	Existence of clear boundaries and exit strategy	20
I	Reputation or 'fit and proper' principle	19
J	Employment of a specific technology	15
K	Increase of legal certainty	12
L	Proposal clarity and completeness	10

The occurrences of the selection criteria for each use case are shown in Table 2 in the Annex.

3.1. – Core criteria (more than 50 occurrences)

The first group – with more than 50 occurrences – represents the fundamental selection criteria that must be considered when establishing a regulatory sandbox. It comprises three dimensions, i.e. the degree of innovativeness of the solution, the potential for public

or societal benefit, and the level of maturity of the solution to be experimented with.

The most frequent criterion is the degree of innovativeness of the solution, defined by its innovative value or nature, as evidenced by 74 instances overall (see Table 1, reference value [A]). Specifically, this criterion assesses the novelty and innovative character of the project or solution proposed for inclusion in the regulatory sandbox. It evaluates whether the proposed innovation addresses existing market gaps, introduces new functionalities, or offers distinct value propositions. This criterion can be operationalized by examining factors such as whether the product or service employs innovative technologies, the absence of similar offerings in the domestic market, the extent to which the solution diverges from currently available alternatives, or the introduction of a significantly innovative business model (e.g., 'Italy - Financial Services Regulatory Sandbox').[3] The criterion may also be demonstrated through market studies highlighting existing ecosystem gaps and the clear potential of the proposed project (e.g., 'Portugal - Free Zones for Technology'). Ultimately, the criterion encompasses entirely novel products, fresh perspectives on existing ideas, or solutions yet to gain substantial traction in the market (e.g., 'United Kingdom - FCA's Fintech Regulatory Sandbox').

The second most present criterion pertains to the public interest and societal benefits that the solution may provide once implemented, with a total of 67 instances (Table 1, value [B]). This criterion evaluates the broader societal implications of the project or solution, specifically the extent to which it advances the public interest or delivers societal benefits, thereby contributing added value to citizens and/or the market. Regulatory sandboxes often prioritize initiatives with the potential to generate positive societal impacts, such as improving accessibility, promoting sustainability, or enhancing public and private services. This may also include fostering a more dynamic and competitive market. A practical application of this criterion could involve requiring applicants to explain how the admission of their project is expected to result in, or has the potential to result in, a net public benefit, as demonstrated in the 'Australia - Enhanced Regulatory Sandbox (ERS)' case. Eligible projects should therefore incorporate a dimension of public interest or contribute to collective welfare, such as enhancing privacy

[3] In the case of 'Norway - Regulatory Sandbox for Archives, Data, and Public Access,' this is framed as 'innovation height', emphasizing radical innovations, or rethinking processes and solutions.

protections (e.g., 'France - CNIL's Personal Data Sandbox'). Regarding market benefits, the proposed solution might deliver tangible advantages to market users (as in 'Greece - CBG's Financial Services Regulatory Sandbox') or promote market inclusion and awareness, ultimately becoming widely beneficial and adopted (as in 'Indonesia - FSA's Digital Finance Innovation Initiative').

The final criterion pertains to the maturity of the solution, with a total of 63 occurrences (Table 1, value [C]). It assesses the readiness of the project for experimentation upon admission. Key considerations include financial sustainability, development stage, and technical capability of the innovation to endure testing conditions. Additionally, this criterion encompasses the potential for scalability, thus evaluating the solution's ability to be deployed into the broader market while maintaining performance and efficiency under increased demands. Indicators of maturity may include technical readiness, the implementation of appropriate safety measures – particularly for experimentation in real-world conditions – a contingency plan for premature termination, and the identification of risks along with corresponding mitigation strategies (e.g., 'Slovakia - NBS's Financial Services Regulatory Sandbox'). Another factor to be considered is the solution's implementation potential, including the presence of detailed implementation plans (as illustrated in 'Switzerland - Energy Regulatory Sandbox'). The sustainability of the project may also be evaluated through the availability of adequate financial and technical expertise, (e.g., 'United States - Florida's Financial Technology Sandbox Innovator').

3.2. – Suggested criteria (between 26 and 50 occurrences)

The second group, comprising criteria emerged between 26 and 50 times in the analysis, represents the recommended selection criteria to be taken into account when developing a regulatory sandbox. This group includes risk management mechanisms, the remit of a specific market surveillance authority and the need for testing.

The most frequent criterion in this group is represented by the implementation of sound and robust risk management mechanisms, which emerged in 43 entries (Table 1, value [D]). This involves conducting risk assessments for the identification of potential risks (in terms of impact, severity, and likelihood), and the application of strategies to mitigate them, as demonstrated in the 'Malta - MDIA's Technology Assurance Sandbox'

and 'Taiwan - Regulatory Sandbox for Self-Driving Vehicles'. Risk assessment and mitigation strategies should be tailored to the specific solution being considered for the regulatory sandbox. For instance, the 'Austria - Framework Conditions for Automated Driving' requires a detailed route analysis and risk assessment for the proposed test route or area. Additionally, it may be necessary to evaluate potential threats to market stability or consumer protection, as outlined in examples such as 'Austria - FMA's Financial Services Sandbox,' Hong Kong's 'HKIA's Insurtech Sandbox' and 'HKMA's Fintech Supervisory Sandbox,' and the 'Netherlands - DNB & AFM Financial Services Regulatory Sandbox'.

The second suggested criterion concerns the administrative remit and mandate of the relevant authority, pertaining to 33 occurrences (Table 1, value [E]). This criterion emphasizes the necessity for the project and its applicant to fall within the supervisory jurisdiction of the authority responsible for the regulatory sandbox. For instance, the project's activities must be directly or indirectly governed by specific or sectoral legislation, as seen in 'Denmark - FSA's FT Lab' and 'Greece - CBG's Financial Services Regulatory Sandbox.' Alternatively, the product or service entering the regulatory sandbox must have a direct connection to the country's relevant sectors, as exemplified by the 'Estonia - FSA's Financial Services Test Environment.'

Another suggested criterion relates to the necessity for testing or experimentation, emerged in 31 instances (Table 1, value [F]). This criterion evaluates whether applicants would derive a clear benefit from testing their solution in a controlled environment, such as validating its feasibility, functionality, or compliance prior to market deployment. A practical application of this criterion involves the advantage of receiving guidance from the relevant authority. For instance, in projects concerning personal data protection, guidance from a Data Protection Authority can be invaluable, as demonstrated by the 'Iceland - Privacy Sandbox Project' and the 'Norway - Personal Data Regulatory Sandbox'. The regulatory sandbox should therefore serve as an appropriate framework for evaluating the solution's risks, challenges, and potential regulatory or supervisory gaps, as highlighted in the 'Malta - MFSA's FinTech Regulatory Sandbox'.

3.3. – Additional criteria (25 or fewer occurrences)

The final group, consisting of criteria cited 25 or fewer times, represents the additional

selection criteria identified in the analyzed use cases.

One criterion is related to the presence of legal barriers and challenges (22 occurrences, Table 1, value [G]). It specifically requires the applicant to identify any legal barriers or restrictions that may impede market rollout of certain projects or solutions (e.g., 'Malaysia - BNM's Financial Services Regulatory Sandbox' or 'Netherlands - DNB & AFM Financial Services Regulatory Sandbox'). This is particularly important in highly regulated sectors that may thus hinder the entrance of new market players with innovative business ideas.

Another criterion is related to the existence of clear boundaries (in terms of scope of the experiment) and/or of an exit strategy (20 occurrences, Table 1, value [H]). The boundaries often imply a clear identification of the time duration and volume of participants to the testing phase and of the derogation requested. This is coupled with the presence of a sound exit strategy once the derogation or experimentation phase expire, and the solution would need to either enter the market or get discontinued. Examples of this include 'United Kingdom - OFGEM's Energy Regulation Sandbox' and 'Japan - Regulatory Sandbox Framework'.

In some instances, reputation also plays a role when selecting the applicants for admission in the regulatory sandbox, often referred to as the 'fit and proper' principle (19 occurrences, Table 1, value [I]). Evaluating the reputation and prior conduct of the applicant or participating entities, this criterion considers factors such as past regulatory compliance, ethical track record of key personnel involved in the development of the solution, and corporate governance practices. Relevant examples especially concern sandboxes established in the United States.

Fifteen regulatory sandboxes also clarify the need to employ a specific technology (Table 1, value [J]). For example, this criterion could focus on whether the project or solution being considered for the regulatory sandbox is ICT-based or pertains to specific areas of technology regulation. Several sandboxes explicitly require that the solution proposed is technology-enabled, in certain cases giving emphasis even to new technologies (such as distributed ledger technology or artificial intelligence). This is the case for 'United Kingdom - ICO's Privacy Regulatory Sandbox' which requires the alignment of the project proposed with its key focus areas, or 'France - CNIL's Personal Data Sandbox' which shifts the area of investigation with each cohort, however always considering the

value of the project for data protection.

In some cases, it is also important that the project provides an increase of legal certainty, thus a better understanding of the legal framework (12 occurrences, Table 1, value [K]). For example, it may entail the presence of uncertainty on how the rules should be interpreted and implemented in practice ('Iceland - Privacy Sandbox Project') or the presence of grey areas where there is a need for regulatory guidance ('Sweden - Regulatory Sandbox on Data Protection').

Lastly, proposal clarity and completeness may also be required to streamline the selection procedure (10 occurrences, Table 1, value [L]). This criterion specifically requests the presence of a transparent, precise and complete proposal for regulatory sandbox participation (e.g., 'Austria - Exemptions from system usage charges for research and demonstration projects", "France - France Experimentation', and 'United Kingdom - OFGEM's Energy Regulation Sandbox').

4. – OPERATIONAL PHASES OF REGULATORY SANDBOXES

This section delves into the procedural aspects of regulatory sandboxes, focusing on key phases such as application, participation, and exit. It considers the mechanisms that govern the entry of participants, the support and oversight provided during the testing phase, and the conditions under which participants conclude their involvement. By analysing these processes, the chapter highlights the flexibility and regulatory considerations that shape the operation of regulatory sandboxes. This evaluation offers valuable practical insights that can inform the design and implementation of regulatory sandboxes, in particular in the field of AI (pursuant to Article 58(1) point (b) of the AI Act[4]), by providing a framework for optimizing their effectiveness and adaptability in diverse regulatory environments.

4.1. – Application

Regarding the admission procedure, there is no clear-cut indication between application windows (i.e., cohort-based) or permanent on-demand applications (i.e., on a rolling basis).

[4] The implementing acts on AI regulatory sandboxes will include common principles on procedures for their application, participation, monitoring, exiting from and termination.

Application windows are used in the 'European Blockchain Regulatory Sandbox', 'France - CNIL's Personal Data Sandbox' and 'Italy - Financial Services Regulatory Sandbox'. On the other hand, on-demand applications are envisaged in the e.g., 'Mauritius - FSC's Financial Services Regulatory Sandbox License', 'Oman - CBO's Fintech Regulatory Sandbox Framework' and 'United States - Kentucky's Insurance Regulatory Sandbox'. The United Kingdom's FCA Fintech Regulatory Sandbox initially operated as a cohort-based program until 2020. Starting in 2021, it transitioned to an on-demand application model.

It is important to notice how in some use cases a mixed approach has been adopted. For example, the 'France - France Experimentation' foresees on-demand applications for regulatory blockages and cohort-based ones for legislative blockages.[5] In a further case, the 'Malaysia - BNM's Financial Services Regulatory Sandbox' features two distinct and parallel tracks: (i) the Standard Sandbox, established in October 2016, which enables fintech companies to test innovative solutions through the existing standard procedures, foreseeing on-demand applications; and (ii) the Green Lane, introduced in February 2024, an expedited track designed to simplify and accelerate the testing process for financial institutions with a proven track record in risk management, specifically for innovative solutions encountering regulatory barriers and envisaged on a cohort basis. This enhances the flexibility of the regulatory sandbox scheme, by allowing certain types of interested participants to get faster access to regulatory guidance and thus a quicker way to the market.

On the other hand, the application phase presents several common aspects across the analysed use cases. For instance, applicants are required to submit specific documents, such as a letter of intent or a complete application form, an eligibility self-assessment checklist (e.g., explaining how the selection criteria are fulfilled), or a specific sandbox testing/implementation plan. Additional requirements may be requested by

---

[5] Regulatory and legislative blockages refer to different types of barriers that innovative projects may face that prevent their deployment. Regulatory blockages occur when existing rules or standards are an obstacle for the innovation, which can often be addressed through exemptions or adjustments by administrative bodies. Legislative blockages, however, stem from laws requiring changes that should be adopted by the legislative body – this is why such changes are typically facilitated through thematic calls for projects aligned with upcoming legislative updates.

the competent authority to aid in the evaluation process. Some processes require the submission of an expression of interest before the formal application (e.g., 'Portugal - Free Zones for Technology'). This allows potential applicants to receive support, guidance, and recommendations from competent authorities before submitting a formal proposal.

The submitted applications undergo evaluations, either by internal committees or an external pool of experts, to assess their eligibility and potential impact. The assessment thus involves considering the selection criteria defined by the specific regulatory sandbox. Competent authorities aim to communicate their decisions regarding admission within a specified timeframe, typically ranging from thirty to ninety calendar days after receiving the complete application. Some processes, however, aim to expedite reviews, especially for substantially similar projects or previously granted waivers. In certain regulatory sandboxes relevant agencies or stakeholders are also involved in consultations before the decision for admitting applicants. It is also worth noticing that, in the case of 'Philippines - BSP's Financial Services Regulatory Sandbox', applicants that do not meet the selection standards could file a new application after a cooling-off period of 6 months since the notification of the negative result.

The participation to the regulatory sandboxes is generally free of charge, except in some cases where fees are foreseen for the application administrative review or participation (usually between 50 and 500 EUR).

4.2. – Participation and monitoring

After a project has been successfully admitted to the regulatory sandbox, the testing and experimentation phase begins. Upon acceptance, participants may engage in a preliminary preparatory phase to finalize testing parameters, further develop the sandbox testing plans, and negotiate and agree the conditions for participation with the competent authorities. This supports the accomplishment of an adequate level of operational readiness by the participants. This preparation phase often involves consultations, adjustments to ensure compliance with non-waived regulations, and finalisation of the testing protocol.

In the actual participation phase, participants conduct the experiments or trials under regulatory supervision. Testing may be divided into different stages, such as design and implementation phases. During this phase participants receive support and

guidance from the competent authorities to address compliance issues, navigate legal requirements, and refine their innovations before market deployment. Assistance may include consultations and technical support.

The duration of this phase is usually pre-defined, generally lasting between 6 and 12 months, especially with regards to sectors such as financial services, healthcare, data protection and management. Longer periods for experimentation have been found in sectors such as telecommunications (12 to 24 months), energy (24 to 48 months), and transportation (12 to 48, or even up to 60 months). This is also an indicator of the inherent complexities of the different sectors in which the experimentation should be deployed: for energy and transportation, a proper experimentation is longer due to the need for more comprehensive testing, involving larger-scale trials, infrastructure assessments, and the integration of complex systems. These sectors often require prolonged testing periods to ensure that innovations are viable under real-world conditions, considering factors such as safety, scalability, and long-term sustainability. Additionally, the regulatory and technical challenges associated with these industries necessitate extensive monitoring and adjustment throughout the experimentation phase. In many cases, extensions on the testing period may be granted based on specific criteria or evaluation outcomes.

In this phase, it becomes crucial to have a proper monitoring mechanism for supervising the testing activities. In general, participants are required to maintain a continuous communication with the competent authorities running the regulatory sandboxes. This includes providing regular updates on the progress of the experimentation. Reporting requirements vary but generally include submitting periodic reports detailing milestones achieved, consumer interactions (if any), and any modifications made to the solutions under experimentation. Participants are thus required to maintain comprehensive records related to their innovative products or services tested within the regulatory sandbox. On the other hand, competent authorities may issue written instructions or request modifications to the sandbox implementation plan in order to ensure compliance with regulations and the proper conduct of tests. Competent authorities may also provide guidance and support to participants through workshops, consultations, and informal supervision, such as informal 'steers' or advice on various aspects of their innovation projects, including risk mitigation and design considerations.

Another important aspect to consider is the possibility to either suspend or

terminate participation in the regulatory sandbox before its planned end date. There are a number of different situations in which this may occur. First, participants' exemptions or approvals may cease automatically if they breach conditions attached to the exemption, fail to satisfy regulatory requirements, or become licensed to provide the same services that are being tested. On the other hand, competent authorities always reserve the right to cancel or revoke exemptions or approvals if participants fail to comply with conditions, fail to communicate progress, or if there are significant concerns about consumer protection, market integrity, or legal compliance (this may include the occurrence of particularly impactful incidents). Competent authorities may also suspend or terminate sandbox participation if there are material deviations from the approved activities in the sandbox testing plan or for technical or implementational reasons. In addition, participants may also choose to withdraw voluntarily from regulatory sandboxes by notifying the competent authorities. Early termination or withdrawal may require approval from the regulatory body and the development of an exit strategy to minimize potential harm to consumers or data subjects.

Participants may also request an authorization to extend the participation phase. In general, this request is forwarded to the competent authority up until 30 calendar days before the planned testing end date. Competent authorities have discretionary powers in relation to granting or not the requested extension.

4.3. – Exiting

After the participation period has ended, the participants need to exit the regulatory sandbox. The exiting phase involves the evaluation of the testing and experimentation activities and may entail a decision on whether the solution would be ready to transition to regular supervision. Participants are required to develop robust exit strategies as part of the regulatory sandbox framework. These strategies aim to ensure an orderly market exit if the tested service proves unsuccessful, minimizing risks to customers and the market.

Once the testing has ended, there is the need to formalise the results and the main evidence gathered during such testing, into dedicated exit reports. These exit reports are usually prepared by the participants, or in some cases jointly by the participants and the supervisory authorities, within 30 to 60 days after the regulatory sandbox has been

concluded. Some key elements are contained in these exit reports. First and foremost, the reports include the main outcomes, key performance indicators against agreed measures for the success or failure of the testing (timeline, budget, scope, etc.), and findings from the testing conducted. This is conducive to a general evaluation of the project, and to the effectiveness of the sandbox as a regulatory tool. A full account of all incident reports and resolution of (eventual) customer complaints may also be included in exit reports to allow a better refinement of the solution before market deployment.

In case of failed testing, reports may include lessons learnt from the testing conducted. On the other hand, in case of successful tests, it may be useful to outline a plan for the transition of the solution to a commercial scale. It is also important to include in exit reports a description of the key issues identified or faced during the participation, or possible regulatory barriers to the viability of the solution, and how those issues were effectively resolved. The exit report may also include key insights from the regulatory sandbox for possible action points in wider policy formulation.

Final summary reports outlining a summary of the overall work, recommendations, and outcomes of the projects tested are often prepared by competent authorities to disseminate the lesson learnt through the operation of the regulatory sandboxes. Depending on the outcome of the participation phase, various actions may be taken. If successful, operational restrictions may be removed, and participants may proceed to provide the service fully compliant with regulatory requirements. Participants may also choose to terminate the provision of the service.

5. – CONCLUSIONS

This contribution had the aim to provide insights emerging from the comparative analysis of 87 use cases of regulatory sandboxes and regulatory experimentation initiatives. The insights focused on the selection criteria adopted to evaluate applicants for admission to the regulatory sandbox, and the operational phases of the latter.

The analysis demonstrated that the most significant selection criteria adopted across regulatory sandboxes include (i) the innovative value of the project, (ii) the associated public collective benefits stemming from the specific solution, and (iii) the maturity and readiness of the solution for testing and subsequent market entry. This finding is indeed

in line with the aim of regulatory sandboxes to tackle 'regulatory challenges generated by technological transformation, and the emergence of new products, services and business models' (European Commission 2023a, 599). In this sense, the innovative projects need to be at a sufficiently advanced level to permit their testing or evaluation, whilst also possessing enough value for society, companies, citizens and consumers if deployed.

As to the participation process, the application phase is characterized by both cohort-based and on-demand models, reflecting a certain degree of flexibility in the scheme design. Some jurisdictions like how these frameworks can be adapted to fit different needs, ranging from expedited processes for experienced firms to more inclusive approaches that ensure adequate support for emerging players.

The participation and monitoring phase is central to the success of regulatory sandboxes. While regulatory bodies provide essential support, guidance, and supervision during testing, the duration and complexity of these phases vary significantly across sectors. The emphasis on continuous monitoring, reporting, and collaboration between sandbox participants and regulatory authorities ensures that testing remains aligned with legal requirements and market expectations.

On the other hand, the exit phase is a critical determinant of the success of the regulatory sandbox, particularly in terms of providing clear pathways for innovation to transition to full regulatory compliance or market deployment. The preparation of exit reports, which document the outcomes, challenges, and lessons learnt, plays a vital role in refining solutions for broader market adoption and in gathering evidence for subsequent regulatory learning. These reports not only help improve individual products but also provide valuable insights into regulatory barriers and operational issues that can inform future policymaking.

ANNEX

Table 2. Occurrences of the selection criteria analysed for each use case

ID	Title of use case	[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]	[K]	[L]
1	Argentina - CNV's Financial Services Innovation Hub	✓	✓	✓	–	✓	✓	–	–	–	–	–	✓
2	Australia - AEMC's Energy Regulatory Sandboxes	✓	✓	–	✓	–	✓	✓	✓	–	–	✓	–
3	Australia - Enhanced Regulatory Sandbox (ERS)	✓	✓	–	–	✓	–	–	–	✓	–	–	–
4	Austria Exemptions - from system usage charges for research and demonstration projects	✓	–	✓	–	–	–	–	–	–	–	–	✓
5	Austria - FMA's Financial Services Sandbox	✓	✓	✓	✓	✓	–	✓	–	–	✓	–	–

* The table below shows the occurrences of each selection criterion considered with respect to the use cases analysed. The selection criteria are identified in the table with a matching letter, for which the following applies: [A] = degree of innovativeness; [B] = public interest or societal benefit; [C] = level of maturity; [D] = risk management mechanisms; [E] = authority mandate and remit; [F] = need for testing or experimentation; [G] = presence of legal barriers; [H] = existence of clear boundaries and exit strategy; [I] = reputation or 'fit and proper' principle; [J] = employment of a specific technology; [K] = increase of legal certainty; [L] = proposal clarity and completeness.

ID	Title of use case	[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]	[K]	[L]
6	Austria - Framework Conditions for Automated Driving	–	–	✓	✓	–	–	–	–	–	–	–	–
7	Bahrain - CBB's Financial Services Regulatory Sandbox	✓	✓	✓	✓	–	–	–	✓	–	–	–	–
8	Brazil - BCB's Financial Services Regulatory Sandbox	✓	✓	✓	✓	✓	–	–	✓	✓	–	–	–
9	Colombia - LaArenera	✓	✓	–	–	✓	✓	✓	–	–	–	–	–
10	Czechia - Use of radio frequencies for experimental purposes	✓	–	–	–	–	–	–	–	–	–	–	–
11	Denmark - FSA's FT Lab	✓	✓	✓	–	✓	✓	–	–	–	–	–	–

* The table below shows the occurrences of each selection criterion considered with respect to the use cases analysed. The selection criteria are identified in the table with a matching letter, for which the following applies: [A] = degree of innovativeness; [B] = public interest or societal benefit; [C] = level of maturity; [D] = risk management mechanisms; [E] = authority mandate and remit; [F] = need for testing or experimentation; [G] = presence of legal barriers; [H] = existence of clear boundaries and exit strategy; [I] = reputation or 'fit and proper' principle; [J] = employment of a specific technology; [K] = increase of legal certainty; [L] = proposal clarity and completeness.

ID	Title of use case	[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]	[K]	[L]
12	Denmark - Regulatory Test Zones for energy technologies	✓	✓	✓	✓	✓	–	✓	✓	–	–	–	–
13	Egypt - CBE's Financial Services Regulatory Sandbox	✓	✓	✓	–	✓	✓	–	–	–	–	–	–
14	Estonia - FSA's Financial Services Test Environment	✓	✓	✓	–	✓	✓	–	–	✓	–	–	–
15	European Blockchain Regulatory Sandbox	✓	✓	✓	–	–	–	–	–	✓	–	✓	–
16	European Union - EIT's Digital Sandbox Accelerator for Healthcare	✓	✓	✓	–	✓	✓	–	–	–	–	–	✓
17	France - Arcep's Regulatory Sandbox for Telecommunications	✓	✓	–	–	–	–	–	–	–	–	–	–

* The table below shows the occurrences of each selection criterion considered with respect to the use cases analysed. The selection criteria are identified in the table with a matching letter, for which the following applies: [A] = degree of innovativeness; [B] = public interest or societal benefit; [C] = level of maturity; [D] = risk management mechanisms; [E] = authority mandate and remit; [F] = need for testing or experimentation; [G] = presence of legal barriers; [H] = existence of clear boundaries and exit strategy; [I] = reputation or 'fit and proper' principle; [J] = employment of a specific technology; [K] = increase of legal certainty; [L] = proposal clarity and completeness.

ID	Title of use case	[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]	[K]	[L]
18	France - CNIL's Personal Data Sandbox	–	✓	✓	–	–	–	–	–	–	✓	✓	–
19	France - CRE's Energy Sandboxes	–	✓	–	–	✓	–	–	–	–	–	–	–
20	France - France Experimentation	✓	✓	✓	✓	–	–	✓	–	–	–	–	✓
21	Germany - Hub Chain of Osnabrück	✓	–	–	–	–	–	–	–	–	–	–	–
22	Germany - Hub Chain of Osnabrück	✓	–	–	–	–	–	–	–	–	–	–	–
23	Germany - On-demand transportation (Hannover Region)	✓	–	–	–	–	–	–	–	–	–	–	–

* The table below shows the occurrences of each selection criterion considered with respect to the use cases analysed. The selection criteria are identified in the table with a matching letter, for which the following applies: [A] = degree of innovativeness; [B] = public interest or societal benefit; [C] = level of maturity; [D] = risk management mechanisms; [E] = authority mandate and remit; [F] = need for testing or experimentation; [G] = presence of legal barriers; [H] = existence of clear boundaries and exit strategy; [I] = reputation or 'fit and proper' principle; [J] = employment of a specific technology; [K] = increase of legal certainty; [L] = proposal clarity and completeness.

ID	Title of use case	[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]	[K]	[L]
24	Germany - Self-driving public bus (Monheim am Rhein & Kelheim)	✓	–	–	–	–	–	–	–	–	–	–	–
25	Greece - CBG's Financial Services Regulatory Sandbox	✓	✓	✓	–	✓	✓	–	–	–	–	–	–
26	Hong Kong - HKIA's Insurtech Sandbox	–	–	✓	✓	–	–	–	✓	–	✓	–	✓
27	Hong Kong - HKMA's Fintech Supervisory Sandbox	–	–	✓	✓	–	–	–	✓	–	–	–	–
28	Hong Kong - SFC's Financial Services Regulatory Sandbox	✓	✓	–	✓	✓	–	–	–	–	✓	–	–
29	Hungary - MNB Innovation Hub - Financial Innovation Testing Environment	✓	✓	✓	–	–	✓	–	–	–	–	–	–

* The table below shows the occurrences of each selection criterion considered with respect to the use cases analysed. The selection criteria are identified in the table with a matching letter, for which the following applies: [A] = degree of innovativeness; [B] = public interest or societal benefit; [C] = level of maturity; [D] = risk management mechanisms; [E] = authority mandate and remit; [F] = need for testing or experimentation; [G] = presence of legal barriers; [H] = existence of clear boundaries and exit strategy; [I] = reputation or 'fit and proper' principle; [J] = employment of a specific technology; [K] = increase of legal certainty; [L] = proposal clarity and completeness.

ID	Title of use case	[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]	[K]	[L]
30	Iceland - Privacy Sandbox Project	–	✓	✓	–	✓	✓	–	–	–	✓	✓	–
31	India - RBI's Financial Services Regulatory Sandbox	✓	✓	✓	✓	–	–	✓	✓	✓	–	–	✓
32	Indonesia - FSA's Digital Finance Innovation Initiative	✓	✓	✓	✓	–	–	–	–	–	✓	–	–
33	Italy - Financial Services Regulatory Sandbox	✓	✓	✓	✓	–	✓	✓	–	–	–	–	–
34	Italy - Italy Experimentation	–	✓	✓	✓	–	–	–	–	–	–	–	–
35	Japan - Regulatory Sandbox Framework	✓	–	✓	–	–	–	✓	✓	–	–	–	–

* The table below shows the occurrences of each selection criterion considered with respect to the use cases analysed. The selection criteria are identified in the table with a matching letter, for which the following applies: [A] = degree of innovativeness; [B] = public interest or societal benefit; [C] = level of maturity; [D] = risk management mechanisms; [E] = authority mandate and remit; [F] = need for testing or experimentation; [G] = presence of legal barriers; [H] = existence of clear boundaries and exit strategy; [I] = reputation or 'fit and proper' principle; [J] = employment of a specific technology; [K] = increase of legal certainty; [L] = proposal clarity and completeness.

ID	Title of use case	[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]	[K]	[L]
36	Jordan - CBJ's FinTech Regulatory Sandbox	✓	✓	✓	✓	–	–	–	✓	–	–	–	–
37	Kuwait - CBK's Financial Regulatory Sandbox - Innovation Hub	✓	✓	✓	–	–	–	–	–	–	–	–	–
38	Latvia - Financial Services Regulatory Sandbox	✓	✓	–	–	✓	–	–	–	–	–	–	–
39	Lithuania - Financial Services Regulatory Sandbox	✓	✓	✓	–	–	✓	–	–	–	–	–	–
40	Malaysia - BNM's Financial Services Regulatory Sandbox	✓	✓	✓	✓	–	–	✓	–	✓	–	–	–
41	Malta - MDIA's Technology Assurance Sandbox	✓	✓	✓	✓	✓	✓	–	–	–	✓	–	–

* The table below shows the occurrences of each selection criterion considered with respect to the use cases analysed. The selection criteria are identified in the table with a matching letter, for which the following applies: [A] = degree of innovativeness; [B] = public interest or societal benefit; [C] = level of maturity; [D] = risk management mechanisms; [E] = authority mandate and remit; [F] = need for testing or experimentation; [G] = presence of legal barriers; [H] = existence of clear boundaries and exit strategy; [I] = reputation or 'fit and proper' principle; [J] = employment of a specific technology; [K] = increase of legal certainty; [L] = proposal clarity and completeness.

ID	Title of use case	[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]	[K]	[L]
42	Mauritius - FSC's Financial Services Regulatory Sandbox License	✓	–	–	✓	–	✓	–	–	–	–	–	–
43	Mauritius - FSC's Financial Services Regulatory Sandbox License	✓	✓	✓	✓	–	✓	–	–	–	–	–	–
44	Mauritius - MEDB's Regulatory Sandbox License	✓	✓	✓	✓	–	–	✓	✓	–	✓	–	–
45	Netherlands - DNB & AFM Financial Services Regulatory Sandbox	✓	✓	–	–	–	–	✓	–	–	–	–	–
46	Nigeria - CBN's Regulatory Sandbox	✓	✓	✓	✓	–	–	–	✓	–	–	–	–

* The table below shows the occurrences of each selection criterion considered with respect to the use cases analysed. The selection criteria are identified in the table with a matching letter, for which the following applies: [A] = degree of innovativeness; [B] = public interest or societal benefit; [C] = level of maturity; [D] = risk management mechanisms; [E] = authority mandate and remit; [F] = need for testing or experimentation; [G] = presence of legal barriers; [H] = existence of clear boundaries and exit strategy; [I] = reputation or 'fit and proper' principle; [J] = employment of a specific technology; [K] = increase of legal certainty; [L] = proposal clarity and completeness.

ID	Title of use case	[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]	[K]	[L]
47	Norway - FSA's Financial Services Regulatory Sandbox	✓	✓	✓	–	✓	✓	–	–	–	–	–	–
48	Norway - Personal Data Regulatory Sandbox	–	✓	–	–	✓	✓	–	–	–	✓	–	–
49	Norway - Regulatory Sandbox for archives, data and public access	✓	✓	✓	–	–	✓	–	–	–	–	✓	–
50	Oman - CBO's Fintech Regulatory Sandbox Framework	✓	✓	✓	✓	✓	–	–	✓	–	–	–	✓
51	Philippines - BSP's Financial Services Regulatory Sandbox	✓	–	✓	✓	–	✓	–	–	–	–	–	–
52	Portugal - Free Zones for Technology	✓	✓	–	✓	–	–	–	✓	–	–	✓	✓

* The table below shows the occurrences of each selection criterion considered with respect to the use cases analysed. The selection criteria are identified in the table with a matching letter, for which the following applies: [A] = degree of innovativeness; [B] = public interest or societal benefit; [C] = level of maturity; [D] = risk management mechanisms; [E] = authority mandate and remit; [F] = need for testing or experimentation; [G] = presence of legal barriers; [H] = existence of clear boundaries and exit strategy; [I] = reputation or 'fit and proper' principle; [J] = employment of a specific technology; [K] = increase of legal certainty; [L] = proposal clarity and completeness.

ID	Title of use case	[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]	[K]	[L]
53	Qatar - CBQ's FinTech Sandbox & Licensing Registration Platform	✓	✓	✓	✓	–	–	–	✓	–	–	–	–
54	Saudi Arabia - CST's Emerging Technologies Regulatory Sandbox	✓	✓	✓	✓	–	–	✓	✓	–	✓	✓	–
55	Saudi Arabia - SAMA's Open Banking Regulatory Sandbox	✓	✓	✓	✓	–	–	–	✓	–	–	–	–
56	Saudi Arabia - SDAIA's Data and Privacy Regulatory Sandbox	✓	✓	✓	–	–	–	–	✓	–	–	–	–
57	Singapore - MAS's FinTech Regulatory Sandbox	✓		✓	✓	✓	✓		✓	✓	–	–	–

* The table below shows the occurrences of each selection criterion considered with respect to the use cases analysed. The selection criteria are identified in the table with a matching letter, for which the following applies: [A] = degree of innovativeness; [B] = public interest or societal benefit; [C] = level of maturity; [D] = risk management mechanisms; [E] = authority mandate and remit; [F] = need for testing or experimentation; [G] = presence of legal barriers; [H] = existence of clear boundaries and exit strategy; [I] = reputation or 'fit and proper' principle; [J] = employment of a specific technology; [K] = increase of legal certainty; [L] = proposal clarity and completeness.

ID	Title of use case	[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]	[K]	[L]
58	Singapore - MOH's Licensing Experimentation and Adaptation Programme (LEAP)	✓	–	–	–	–	–	–	–	–	✓	–	–
59	Slovakia - NBS's Financial Services Regulatory Sandbox	✓	✓	✓	–	✓	✓	–	–	✓	–	–	–
60	South Africa - IFWG'sFinancial Services Regulatory Sandbox	✓	✓	✓	–	✓	–	✓	–	–	–	–	–
61	South Korea - FSC's Financial Services Regulatory Sandbox	✓	✓	✓	✓	✓	–	✓	–	–	–	–	–
62	South Korea - Smart City Regulatory Sandbox	–	–	–	–	–	–	✓	–	–	–	✓	–
63	Spain - AI Regulatory Sandbox Pilot Scheme	✓	✓	✓	✓	✓	–	–	–	–	–	–	–

* The table below shows the occurrences of each selection criterion considered with respect to the use cases analysed. The selection criteria are identified in the table with a matching letter, for which the following applies: [A] = degree of innovativeness; [B] = public interest or societal benefit; [C] = level of maturity; [D] = risk management mechanisms; [E] = authority mandate and remit; [F] = need for testing or experimentation; [G] = presence of legal barriers; [H] = existence of clear boundaries and exit strategy; [I] = reputation or 'fit and proper' principle; [J] = employment of a specific technology; [K] = increase of legal certainty; [L] = proposal clarity and completeness.

ID	Title of use case	[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]	[K]	[L]
64	Spain - Financial Services Regulatory Sandbox	✓	✓	–	✓	–	✓	–	–	–	–	–	–
65	Sweden - Regulatory Sandbox on Data Protection	–	✓	✓	–	–	–	–	–	–	–	✓	–
66	Switzerland - Energy Regulatory Sandbox	✓	✓	✓	–	–	–	✓	–	–	–	✓	–
67	Switzerland - Zurich's Innovation Sandbox for Artificial Intelligence (AI)	✓	✓	✓	–	✓	✓	–	–	–	–	✓	–
68	Taiwan - FSC's Financial Technology Innovative Experimentation	✓	✓	–	✓	✓	–	–	–	✓	–	–	–
69	Taiwan - Regulatory sandbox for self-driving vehicles	✓	✓	✓	✓	–	–	✓	–	–	–	–	–

* The table below shows the occurrences of each selection criterion considered with respect to the use cases analysed. The selection criteria are identified in the table with a matching letter, for which the following applies: [A] = degree of innovativeness; [B] = public interest or societal benefit; [C] = level of maturity; [D] = risk management mechanisms; [E] = authority mandate and remit; [F] = need for testing or experimentation; [G] = presence of legal barriers; [H] = existence of clear boundaries and exit strategy; [I] = reputation or 'fit and proper' principle; [J] = employment of a specific technology; [K] = increase of legal certainty; [L] = proposal clarity and completeness.

ID	Title of use case	[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]	[K]	[L]
70	Thailand - NBTC's Telecommunications Sandbox Notification	✓	✓	–	–	–	✓	–	–	–	–	–	–
71	United Arab Emirates - ADGM's FinTech RegLab	✓	✓	✓	✓	–	–	–	✓	✓	–	–	✓
72	United Arab Emirates - Dubai's Innovation Testing Licence (ITL) Programme	✓	–	✓	–	✓	✓	–	–	–	–	–	–
73	United Arab Emirates - RegLab	✓	✓	–	✓	–	✓	✓	–	–	✓	✓	–
74	United Arab Emirates - TDRA's ICT Regulatory Sandbox	✓	✓	✓	✓	–	✓	✓	–	–	✓	–	–

* The table below shows the occurrences of each selection criterion considered with respect to the use cases analysed. The selection criteria are identified in the table with a matching letter, for which the following applies: [A] = degree of innovativeness; [B] = public interest or societal benefit; [C] = level of maturity; [D] = risk management mechanisms; [E] = authority mandate and remit; [F] = need for testing or experimentation; [G] = presence of legal barriers; [H] = existence of clear boundaries and exit strategy; [I] = reputation or 'fit and proper' principle; [J] = employment of a specific technology; [K] = increase of legal certainty; [L] = proposal clarity and completeness.

ID	Title of use case	[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]	[K]	[L]
75	United Arab Emirates- CBUAE's Regulatory sandbox for the insurance sector	✓	✓	✓	–	✓	✓	–	–	✓	–	–	–
76	United Kingdom - FCA's Fintech Regulatory Sandbox	✓	✓	✓	–	✓	✓	–	–	–	–	–	–
77	United Kingdom - ICO's Privacy Regulatory Sandbox	✓	✓	✓	–	✓	–	–	–	–	✓	–	–
78	United Kingdom - OFGEM's Energy Regulation Sandbox	✓	✓	✓	–	–	✓	✓	✓	–	–	–	✓
79	United States - Arizona's Financial Services Regulatory Sandbox	–	–	✓	–	✓	–	–	–	–	–	–	–

* The table below shows the occurrences of each selection criterion considered with respect to the use cases analysed. The selection criteria are identified in the table with a matching letter, for which the following applies: [A] = degree of innovativeness; [B] = public interest or societal benefit; [C] = level of maturity; [D] = risk management mechanisms; [E] = authority mandate and remit; [F] = need for testing or experimentation; [G] = presence of legal barriers; [H] = existence of clear boundaries and exit strategy; [I] = reputation or 'fit and proper' principle; [J] = employment of a specific technology; [K] = increase of legal certainty; [L] = proposal clarity and completeness.

ID	Title of use case	[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]	[K]	[L]
80	United States - Florida's Financial Technology Sandbox Innovator	✓	–	✓	✓	–	–	–	–	✓	–	–	–
81	United States - Kentucky's Insurance Regulatory Sandbox	✓	✓	✓	✓	–	–	✓	–	✓	✓	–	–
82	United States - Nevada's Financial Services Sandbox Program	✓	✓	✓	✓	–	–	–	–	–	–	–	–
83	United States - North Carolina's Financial and Insurance Regulatory Sandbox	✓	✓	✓	✓	–	–	–	–	✓	–	–	–
84	United States - Utah's Regulatory Relief Program	–	–	–	–	–	–	–	–	✓	–	–	–

* The table below shows the occurrences of each selection criterion considered with respect to the use cases analysed. The selection criteria are identified in the table with a matching letter, for which the following applies: [A] = degree of innovativeness; [B] = public interest or societal benefit; [C] = level of maturity; [D] = risk management mechanisms; [E] = authority mandate and remit; [F] = need for testing or experimentation; [G] = presence of legal barriers; [H] = existence of clear boundaries and exit strategy; [I] = reputation or 'fit and proper' principle; [J] = employment of a specific technology; [K] = increase of legal certainty; [L] = proposal clarity and completeness.

ID	Title of use case	[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]	[K]	[L]
85	United States - Vermont's Insurance Regulatory Sandbox	–	✓	–	✓	✓	–	✓	–	✓	–	–	–
86	United States - West Virginia's FinTech Sandbox	✓	✓	✓	✓	✓	✓	–	–	✓	–	–	–
87	United States - Wyoming's Financial Technology Sandbox	✓	–	✓	✓	–	–	–	–	✓	–	–	–
TOTAL		74	67	63	43	33	31	22	20	19	15	12	10

* The table below shows the occurrences of each selection criterion considered with respect to the use cases analysed. The selection criteria are identified in the table with a matching letter, for which the following applies: [A] = degree of innovativeness; [B] = public interest or societal benefit; [C] = level of maturity; [D] = risk management mechanisms; [E] = authority mandate and remit; [F] = need for testing or experimentation; [G] = presence of legal barriers; [H] = existence of clear boundaries and exit strategy; [I] = reputation or 'fit and proper' principle; [J] = employment of a specific technology; [K] = increase of legal certainty; [L] = proposal clarity and completeness.

LEARNINGS FROM THE AI SANDBOX IN ZURICH: A PRACTICAL PERSPECTIVE

RAPHAEL VON THIESSEN*

SUMMARY

ABSTRACT

This paper presents insights from the AI sandbox programme in Zurich, highlighting practical learnings. Initiated in 2021, the programme involved multiple stakeholders from public administration and the private sector, and research to establish Zurich as a hub for artificial intelligence. Key objectives included facilitating regulatory learning, fostering innovation, promoting knowledge transfer and providing regulatory input. The AI sandbox served as a unique environment for testing and developing AI technologies, while identifying regulatory gaps and contributing to future regulatory frameworks. Through practical implementation, including projects like smart parking, autonomous systems, and drone inspections, the AI sandbox demonstrated the importance of hands-on testing and collaboration with various partners. Effective communication was identified as a critical success factor, ensuring broad dissemination of insights and fostering a more integrated AI ecosystem. The paper also explores strategic considerations for designing AI sandboxes, such as sector-specific versus sector-agnostic frameworks. The Zurich programme emphasised practical implementation and utilising existing legal frameworks for quicker adaptation. This case study underscores the importance of an iterative approach to AI sandbox development, tailored to specific national and regional contexts, to support sustainable AI innovation and regulatory learning.

---

* Programme manager AI sandbox (Canton of Zurich): raphael.vonthiessen@vd.zh.ch

1. – DEFINING THE GOALS OF THE AI SANDBOX

The rapid advancement of artificial intelligence (AI) technologies presents significant opportunities and challenges for both innovators and regulators. To address these, regulatory sandboxes have emerged as a practical approach to foster innovation while ensuring compliance and safety.

In 2021, various institutions from public administration, the private sector, and research formed a working group to address the growing importance of AI in the Canton of Zurich. At that time, AI did not dominate discussions about technology and innovation as it did after the launch of ChatGPT 3.5 in November 2022, but it was clear that AI would have a significant impact across sectors. The working group designed measures to advance Zurich as a hub for artificial intelligence in a practical and concrete manner. All initiatives focused on a multi-stakeholder approach combining interdisciplinary expertise. Additionally, the network emphasised that the interplay between innovation and regulation should be seen as an opportunity. It became evident that a testing environment for AI would benefit all stakeholders involved. The working group developed the vision of an AI sandbox, inspired by both technical and regulatory interpretations of the term. It is important to note that the term sandbox is used in various contexts, and the approach in Zurich may differ significantly from what has been defined as an 'AI regulatory sandbox' in Article 3(55) of the EU AI Act. To avoid confusion, the term regulatory sandbox is not used in the Canton of Zurich.

Once the vision of creating an AI sandbox was established, the Office for Economy took on a leading role by hiring personnel specifically dedicated to developing a concept and running the programme's operations. This section outlines the strategic objectives of the AI sandbox approach, emphasising regulatory learning, fostering innovation, knowledge transfer, and providing input for future regulation.

One of the main objectives of the Zurich AI sandbox was to facilitate regulatory learning. This involved clarifying regulatory topics and identifying new questions arising from technological advancements and innovation. The AI sandbox experience in Zurich demonstrated the importance of shifting from a strictly compliance-oriented mindset to a more forward-looking approach that anticipates new AI-related questions and topics. This learning process benefited both regulators and AI sandbox participants. Regulators gained

insights into emerging technologies, enabling them to better understand and anticipate future trends. Participants received guidance on navigating regulatory landscapes. Their innovations aligned with existing frameworks and identified areas needing regulatory evolution. It is important to note that the development and commercialisation of AI did not take place in a regulatory vacuum. In most cases, various legal frameworks such as data protection, intellectual property rights, and sector-specific guidelines could and must be applied.

Regulatory learning took place within the constraints of existing legal requirements: one benefit was that the AI sandbox team gained access to a wide range of real-life use cases reflecting current research and market trends. All parties involved benefited from direct exchange. Participating organizations received inputs on shaping their AI products, while public authorities identified future regulatory questions. Most importantly, all regulatory insights were shared broadly so that other private and public institutions could benefit as well.

A second goal was fostering innovation. The AI sandbox programme provided a unique opportunity for participants to test, develop, and validate innovative AI technologies, services, and products. It offered startups, SMEs, and research institutions access to essential resources such as regulatory knowhow and data sources (see '2. - Developing the AI sandbox concept' for more information). By lowering barriers to entry and providing a supportive environment, the AI sandbox encouraged a wide range of entities to explore and scale AI innovations, driving technological progress and economic growth.

Furthermore, the adoption of AI technologies within public administration was also a vital part of the AI sandbox approach. Concrete examples of how the initiative enabled the spread of AI innovation included testing and implementing smart parking solutions using computer vision, drone inspections for infrastructure maintenance, and machine translation for public servants (see '4. - Implementing the AI sandbox use cases' for more information). The insights generated during these projects enabled the adaptation and distribution of new technologies beyond the AI sandbox programme.

Another critical objective of the Zurich AI sandbox approach was promoting knowhow transfer. This environment helped avoid redundant efforts by enabling collaboration and sharing of expertise across public administration, business, and research

domains. This collective knowledge-sharing approach ensured that solutions to common challenges were disseminated widely, fostering a more integrated and efficient ecosystem. By leveraging the insights and experiences of various stakeholders, the AI sandbox accelerated the diffusion of best practices and innovative approaches.

One concrete example of how the programme promoted knowhow transfer was the publication of reports summarizing insights from specific use cases. Additionally, the AI sandbox team frequently held workshops with stakeholders such as AI firms and public authorities to disseminate the acquired knowledge.

Finally, the programme served as a vital source of input for future regulation. By identifying regulatory gaps and areas for action based on real-life experiences and concrete use cases, the sandbox ensured that regulations kept pace with AI advancements. This bottom-up approach complemented top-down frameworks, providing a nuanced understanding of the regulatory needs emerging from actual technological applications. While an AI sandbox cannot cover all relevant AI use cases comprehensively, it contributes to the sustainable development of AI technologies, balancing innovation with safety and ethical considerations.

One notable example emerged in the context of autonomous ground vehicles (see '4. - Implementing the AI sandbox use cases' for further details). The team tested an autonomous tractor for agriculture and a self-driving lawn mower for professional greenkeeping. Both vehicles were designed to operate mainly on private ground and only needed to cross public roads occasionally. Through detailed assessment, the team discovered that these vehicles were being categorized under the same regulatory requirements as traditional vehicles such as cars, trucks, or buses. It also provided regulatory input to better reflect the operational and technological differences of these systems in future regulations.

2. – DEVELOPING THE CONCEPT

The development of the AI sandbox concept was a collaborative and dynamic process. The programme involved eight different institutions, forming an interdisciplinary steering committee. All these institutions were already part of the working group that initiated the programme.

The Office for Economy led the AI sandbox programme with the primary goal of fostering Zurich as an AI hub through collaboration across business, research, and public administration. While not a regulator of AI technologies in the strict sense, the Office for Economy acted as a coordinator, bringing together diverse stakeholders to ensure a holistic approach that covered a wide range of AI-relevant issues. The steering committee included representatives from the Office for Economy, Statistical Office and Division of Digital Government (all Canton of Zurich), the Office for Economy (Canton of Schwyz), Metropolitan Area Zurich Association, ETH AI Center, Center for Information Technology, Society, and Law (University of Zurich) and ZHAW Entrepreneurship. A crucial element was the inter-regional collaboration with the Zurich Metropolitan Area, ensuring that eight different cantons were part of the AI sandbox. This approach was suitable as regulatory challenges are similar across regions and resources can be pooled to overcome fragmentation.

A critical early realization was the absence of specific legal instruments tailored for a regulatory sandbox, such as AI-specific experimentation clauses or no-action letters (Volz, 2022). Instead of waiting to create new regulatory sandbox-specific laws, which would have delayed the programme significantly, it was decided to establish the initiative within the existing legal framework. The programme was based on economic development policies that were already well established in specific sectors (e.g. finance). A novel element of the programme was that it targeted AI as a technology that transforms multiple sectors at the same time. Furthermore, all AI sandbox activities had to comply with current legal requirements without exemptions (e.g. in data protection or public law). This setup enabled a quick programme launch, allowing it to start gaining practical experience and providing valuable insights into the legal and regulatory needs of AI technologies. It took only five months between the formal kick-off and the public launch of the programme through the first call for projects. This highlights the importance of pragmatism and speed in the context of AI sandbox approaches.

Given the lack of a pre-existing AI sandbox template, particularly in the absence of the AI Act definition available today, the team conducted extensive interviews with counterparts in other countries, such as the UK and France, as well as with potential participants, including AI startups. These consultations provided valuable insights and helped shape a pragmatic and adaptable model tailored to Zurich's unique context. For

example, other regulatory sandbox operators mentioned that they had specified too many requirements for project proposals, which became a barrier to entering the programme. The Zurich AI sandbox team avoided this issue by requiring only a high-level outline for potential projects.

A distinctive feature of the Zurich AI sandbox was its focus on data provisioning and the real-world implementation of AI projects. Unlike traditional regulatory sandboxes that primarily emphasised legal assessments, the Zurich model integrated practical, hands-on testing and application of AI technologies. This approach not only facilitated immediate feedback and learning but also enabled the discovery of practical challenges and regulatory issues that would not arise in a theoretical setup.

The regulatory assessment included evaluating legal issues across multiple domains such as data protection, intellectual property rights, and sector-specific regulations in areas like AI in education. The data provisioning involved identifying suitable implementation partners within the public administration network who could support AI sandbox projects as data owners (e.g., public administration, educational institutions, military organisations). It was clear that in certain projects, one of the two main AI sandbox services would dominate based on the specific use case. This openness was intentionally designed to accommodate different types of projects.

A fundamental aspect of the AI sandbox was its sector-agnostic approach, intentionally avoiding limitations to any specific industry to remain flexible and responsive to diverse market needs. This inclusivity encouraged broad participation from startups, SMEs, research institutions, and businesses across various sectors, enriching the learning environment and facilitating cross-sectoral knowledge transfer. The use cases (a total of 45) submitted in the two completed project calls in 2022 and 2024 came from a wide range of sectors such as autonomous systems, sustainability, health, public administration, and education. This confirmed that a sector-agnostic approach was the right decision for Zurich. However, only a few big corporations submitted proposals. A possible explanation is that large organisations already have access to most of the regulatory knowhow and data sources provided within the programme. Furthermore, it was unexpected that out of the 45 submissions, only one project came from the financial services sector, which is traditionally strong in the Canton of Zurich.

The success of the programme was also based on assembling a team with diverse

skills. The team included a generalist programme manager, a legal expert, a data scientist, and a communication expert. This combination ensured that all aspects of the AI sandbox, from legal compliance to technical implementation and stakeholder communication, were effectively managed.

The steering committee identified effective communication as a key success factor. Clear, consistent communication with stakeholders, including potential participants and regulators, helped build trust and transparency. This openness was crucial for encouraging participation and ensuring that the AI sandbox could effectively gather and share insights from the projects. One key element of successful communication was the joint communication activities by all institutions that were part of the steering committee. Leveraging the communication channels of eight different organisations from public administration, the private sector, and research ensured that project calls or new reports reached the right target audience.

3. – SELECTING THE AI SANDBOX USE CASES

The selection of the first cohort of AI sandbox use cases began with an open call for projects lasting two and a half months.

For applications to the AI sandbox, there were five key requirements that needed to be met. Firstly, applicants had to possess their own AI competencies, ensuring they had the necessary technological expertise. Secondly, participation did not include any financial compensation. Thirdly, participants had to be willing to share and publish their findings, although IP-related information such as code was excluded, to contribute to best practices. Additionally, applicants were required to have a presence in Switzerland. Lastly, each organization was allowed to submit only one AI project.

The application process was designed to be straightforward and accessible. Applicants completed a simple online form, providing information about their organisation, general activities in AI, and a detailed description of their proposed use case. The aim was to avoid overly detailed and bureaucratic requirements, which could discourage potential participants. Recognising that projects need to be developed iteratively, the process allowed for flexibility and adaptability. It was nearly impossible to foresee and describe project proposals in exhaustive detail at the application stage.

The selection criteria were designed to ensure a balanced and impactful set of projects that reflected the diverse perspectives of the steering committee. It was important to communicate the selection criteria in advance so that interested organisations could address them in their application.

Firstly, a key criterion was the maturity level of the AI project for concrete implementation. It was not necessary that AI projects use cutting edge technology. The assumption was that well established AI applications could also provide useful insights. Secondly, the potential for building regulatory know-how was assessed. This ensured that projects raised legal questions with potential for clarifications based on a concrete use case. Thirdly, the potential for utilising data sources from the administrative ecosystem was evaluated. Additionally, the potential for providing AI-based services that serve the public interest was considered important.

The potential for strengthening the innovation hub through collaboration among business, research, and administration was another crucial criterion. Furthermore, the potential for applying the results to other AI application areas or sectors was assessed. The relevance of the results for cantons, cities, and municipalities in the Zurich metropolitan area was also a significant factor.

The necessity of participating in the AI sandbox for project implementation was considered. Technical feasibility, based on requirements such as infrastructure, hardware, and models, was evaluated. Lastly, non-technical feasibility, including considerations like data access and political sensitivity, was also an important criterion.

A 30-minute interview with each applicant helped assess motivation and competencies. This online exchange provided deeper insights into the project's potential. The steering committee evaluated each application against the selection criteria, shortlisting the most promising use cases, followed by an in-depth discussion of topics that needed to be covered within a project. Although the evaluation results were not shared with organisations that were not selected, the project team tried to suggest alternative areas for support outside the AI sandbox (e.g., connecting organisations with relevant authorities or complementary technology providers). In cases where similar projects were submitted (e.g., two different projects for machine translation in public administrations), these were bundled to address overarching questions.

4. – IMPLEMENTING THE USE CASES

The implementation involved continuous iterations. The iterative approach stemmed from two main factors. First, the low barrier for entry in the AI sandbox meant that project proposals were very high-level and needed further specification. The complexity of each project varied significantly, from practical implementations requiring the procurement and installation of hardware to those primarily focused on legal assessments. Second, the provision of data sources from implementation partners (e.g., public bodies, military organisations) required that suitable organisations were onboarded after the application had been accepted.

A critical aspect of implementation was identifying and securing the right partners and experts. Implementation partners, such as data owners, were essential for testing AI applications. It was important that they saw the benefits of contributing to an AI sandbox project. This matching of participants and implementation partners created a strong dependency, as the willingness of data owners to participate could not be controlled by the AI sandbox operators themselves. Additionally, involving third-party providers or regulatory experts was often crucial to address specific technical or legal challenges.

One of the practical challenges in implementing the AI sandbox use cases was continuous budgeting. Given the iterative nature of AI project development, funding needs evolved over time. Ensuring flexible and ongoing financial support was essential for addressing unforeseen costs or requirements. While the participating organisations did not receive financial compensation, it was important that the AI sandbox budget covered additional costs (e.g., for hardware, legal expertise, or third-party technology providers). This meant that the implementation partners did not need to contribute their own financial resources, providing further incentives to participate in the programme.

AI use cases often involved regulatory compliance across multiple legal domains, requiring coordination with various regulators. Engaging different regulators early and maintaining ongoing communication was key to managing compliance effectively. Implementation also had to comply with all existing political and administrative processes. In the case of the AI sandbox in Zurich, there were no shortcuts or relaxations of regulatory requirements, and all processes had to be legally and politically fulfilled. This meant that some of the projects lasted for almost one and a half years.

During the project call between March and June 2022, organisations submitted 21 projects across various sectors. Most proposals were submitted by start-ups and SMEs, although big corporations and technology providers also applied. The five projects that were successfully realised were smart parking, autonomous systems, automated infrastructure maintenance, machine translation, and AI in education. These projects are briefly described below.

Urban areas often face challenges with inefficient parking management, leading to congestion, wasted time, and unnecessary emissions. The project on smart parking tested a solution using image recognition technology developed by ETH spin-off Parquery AG to optimise parking space utilisation and guide drivers to available spots. A focus of the collaboration was on ensuring data protection of image recognition in the public sphere through privacy-by-design measures. The project resulted in best practices for other Swiss cities and municipalities (Volz & von Thiessen, 2023a). A benefit of this project was its tangibility. The smart parking system has been alive since November 2023 and provided access to an AI-based service for all citizens in this area. Parking management is also a topic of high public interest, highlighting the importance of non-legal questions such as communication to the public, transparency, and public accountability throughout this project. A challenge was the complexity of procuring and installing camera systems in public spaces. The operational requirements for scaling the solution —such as ensuring 24-hour access to electricity grids and coordinating with real estate owners— took a lot of time and effort. Notably, most of these challenges were not AI-related.

As for autonomous systems, regulation and standardisation for autonomous systems are lagging behind technological developments, creating an unclear legal framework for manufacturers. This project tested autonomous ground vehicles with two startups, including an autonomous tractor for agriculture by Lonomy and a self-driving lawn mower by Ronovatec for professional greenkeeping. The project clarified legal questions for manufacturers across multiple legal domains such as product safety, autonomous driving on public roads, and data protection. The output was a high-level regulatory guideline that helps manufacturers comply with the current and future regulatory environment (Volz & von Thiessen, 2023b). One benefit of this project was the bundling of two similar use cases. This ensured that the AI sandbox team clarified regulatory questions relevant for other manufacturers as well (e.g., autonomous cleaning robots). The project

provided a high-level entry point before manufacturers may conduct company-specific legal assessments. One challenge was that one of the participants went out of business during the collaboration. Even though the reasons were completely unrelated to any AI sandbox activities, it required adaptability. The AI sandbox team was able to finish the regulatory guidelines without the remaining contribution of the firm.

Visual inspections of roads, bridges, and dams are largely performed manually, which is time-consuming and potentially hazardous. The project on automated infrastructure maintenance used drones to create high-quality imagery of an airstrip (pixmap gmbh) and automatically detect cracks and damages for visual inspections (IBM Research). This collaboration took place at a military air base. The AI sandbox created a high-quality data set that can be leveraged by further innovators to test and validate visual inspection algorithms. The AI sandbox programme also published a best practice report that shared the technical and operational findings of this AI application (Scheidegger, von Thiessen & Weiss, 2023). One opportunity of this project was engaging a specific third-party drone provider for capturing high-quality data of the airstrip. To find the right partner, the AI sandbox team compared multiple drone providers regarding their capabilities. This highlighted the importance of a strong partner network, as the project wanted to test the limits of what is currently possible from a technological perspective (sub-millimetre resolution). Furthermore, the data set is now shared with other AI developers working in this area. After the successful completion of the project, the operational challenges of integrating drone inspections into daily operations became apparent. This showed that even a practical AI sandbox environment can be far removed from the day-to-day operations of most organisations.

Language barriers hinder communication and efficiency in public administration. The project on machine translation conducted two different case studies with the Commercial Register of the Canton of Schwyz and the Integration Unit of the Canton of Zurich to evaluate AI-based machine translation tools. The project resulted in guidelines for legal questions regarding the use of AI translation in public administration. Furthermore, the team developed a best practice approach to benchmark different AI translation services to better reflect the specific needs of public institutions (Volz & von Thiessen, 2024). The bundling of two cases enabled the project team to analyse different approaches to customising machine translation tools for the needs of public administration.

Specifically, the team compared fine-tuning with human expert feedback versus manual customisation based on the preferences of subject-matter experts. Furthermore, machine translation is also a topic of high interest among various public administration bodies. It helped to contrast specific AI tools that are customised based on local needs against generic online tools. A challenge was benchmarking the different AI translation services. The blind testing was based on feedback from certified professional translators who compared different machine translation outputs with human translations. This proved to be very complex and time-consuming. Due to the resource and time constraints within the AI sandbox, the blind testing was also not representative.

Teachers are increasingly using AI tools in their education methods. Often, there is uncertainty regarding the legal requirements for the safe use of AI tools among school officials, teachers, pupils, and parents. Another project admitted to the AI sandbox tested an AI application that allows students to use a smartphone scan to automatically correct handwritten math and spelling exercises. The project resulted in a legal best practices guideline that helps the education sector address challenges in data protection and copyright issues (Volz & von Thiessen, 2023c). A success of this project was that the provision of legal expertise resulted in a strategic realignment of the participant. The startup was able to focus on its core functionalities in automated correction. The founders made a trade-off and deprioritised non-core features that were increasing the regulatory burden for the processing of personal data. A clear challenge was the matchmaking with implementation partners, in this case elementary schools. The project team and partner network were too far removed from actual classrooms. It took only two months for the AI sandbox team to identify a suitable partner that specialised in the testing of edtech solutions. Furthermore, the report that summarised the insights on AI in education targeted edtech providers – and not school representatives. It became clear that messaging would need to be different to reach teachers and school principals as a target audience.

5. – ENSURING KNOWHOW TRANSFER

Ensuring effective know-how transfer was a critical component of the AI sandbox approach. Each use case within the AI sandbox served as a means to an end. The aim was always to generate insights that could be applied beyond the specific project. The project

team focused on identifying and extracting relevant learnings that could benefit a wider audience. This process involved continuously monitoring and documenting key findings throughout the implementation phase.

The primary target audiences for these insights often included market participants such as technology providers, as well as public administration entities. Identifying the relevant questions and concerns for these audiences was a key responsibility of the AI sandbox team. The five projects showed that many technology providers were so focused on product development that regulatory questions were not considered a top priority. Even the clarification of basic regulatory concepts across different legal areas generated added value for the participants.

To ensure that insights were effectively communicated to diverse target audiences, the project team created various materials during the implementation phase. These included videos, interviews, and detailed reports that captured and conveyed learnings from each use case. Documenting the process and outcomes in multiple formats helped ensure that the knowledge was accessible and useful to different stakeholders. Graphic design, with illustrations of the AI applications and a professional report layout, was important in conveying the innovative character of the sandbox as a testing environment.

The AI sandbox team published all findings online in the form of best practice reports available in German and English. Translating all findings was crucial as the AI community in Switzerland is very international. To reach diverse target audiences, the AI sandbox employed various formats for know-how transfer beyond the reports. These included blog posts, workshops, keynotes, academic papers, and other relevant formats. Tailoring the delivery method to the specific needs and preferences of each audience ensured that the knowledge was effectively disseminated and utilised.

The know-how transfer also aimed to identify opportunities to contribute to future regulatory frameworks. Depending on the suitability of current legislative processes, the AI sandbox provided input for ongoing regulatory initiatives or highlighted the need for new administrative and political efforts. This proactive approach ensured that insights gained from the projects informed and improved regulatory practices. Input for future regulation occurred on two levels: findings from specific projects, like regulatory gaps in autonomous systems, helped shape sector-specific regulations, and insights from the AI sandbox approach guided policymakers regarding the necessary legal frameworks and

tools for AI testing environments. Swiss policymakers had already adopted Zurich's AI sandbox programme as a model and used it to launch a national political initiative for AI testing environments.

Sharing know-how across borders was also a priority. An example was the AI Sandbox Summit held in Zurich in January 2024 with seven European regulatory sandbox operators. This event provided a platform for cross-border knowledge transfer, facilitating the exchange of insights and best practices between different regions and countries. This international collaboration enhanced the overall impact of the AI sandbox by integrating a broader range of experiences and perspectives from other operators with different maturity levels. An international database of AI sandbox projects could be one way of collaborating across borders to maximise impact.

6. – STRATEGIC CONSIDERATIONS FOR AI SANDBOX DESIGN

The previous chapters described the experiences and learnings from the programme in Zurich. Designing an AI sandbox involves navigating strategic trade-offs, each with its own advantages and disadvantages. The following considerations shape the AI sandbox's effectiveness and suitability for different contexts on a more strategic level.

A thematic AI sandbox concentrates on a specific AI-related issue, such as privacy, which allows for in-depth exploration and specialised guidance. However, Zurich chose a holistic approach, addressing diverse aspects of AI solutions, including legal, technical, and communication dimensions. This aimed to offer a comprehensive understanding of AI's impacts. While it may dilute focus and resources, it ensures a more rounded and practically relevant framework for AI development.

A sector-specific approach targets an industry like healthcare, addressing its unique challenges and regulations to drive significant advancements in this area. Conversely, Zurich chose a sector-agnostic approach, accommodating applications across various sectors. This promotes flexibility and diversity, reflecting the cross-sectoral needs of the AI community, though it may require more generalised solutions.

A competent authority, like a data protection office, provides clear guidance within its area of competence, beneficial for specific regulatory issues but limited in scope. As almost all AI applications are subject to multiple legal frameworks and regulators, Zurich

adopted a coordinative role, addressing multiple legal frameworks and levels (regional, national, supranational) to integrate diverse regulatory perspectives.

Legal assessments clarify regulatory requirements, providing clear guidelines for AI development but missing practical insights. Zurich opted for practical implementation, engaging in hands-on AI deployment, including data provisioning, hardware installation, and public communication. This approach highlights real-world challenges and opportunities, ensuring solutions are practically viable. However, it requires more resources and coordination and creates dependencies with implementation partners.

Developing AI sandbox instruments like experimentation clauses or no action letters can provide clear benefits but can be time-consuming to establish. Zurich utilised existing legal frameworks for quicker implementation and adaptability. While this might not support some AI projects, it allows for immediate application. A parallel approach could combine both options by launching the AI sandbox under the current legal framework and developing new instruments based on practical insights during the programme.

In conclusion, designing an AI sandbox involves balancing strategic considerations and practical approaches to create an environment that enables AI innovation and regulatory learning. An iterative approach to its development is crucial, allowing for continuous optimisation based on practical experiences and evolving needs. Additionally, adapting the AI sandbox to national culture and regional conditions ensures that it remains relevant and effective within its unique context. By thoughtfully navigating these trade-offs, AI sandbox designers can create a robust framework that supports AI innovation while addressing regulatory learning and practical challenges.

THE NEED FOR AN ETHICAL APPROACH TO REGULATORY SANDBOXES

KATE E. FRANCIS*

SUMMARY

ABSTRACT

Regulatory sandboxes are promising tools which have the potential to promote privacy-enhancing innovation and enable trust in technologies, such as artificial intelligence (AI). They promote regulatory learning and provide an opportunity for embedding legal and ethical compliance directly into technologies. Despite their many potential benefits, sandboxes also present several ethical risks and challenges which will undermine their usefulness as a regulatory tool if not properly mitigated. Challenges and risks posed by regulatory sandboxes include possible risks stemming from the Collingridge dilemma, a lack of resources, ethics washing, abuse and misuse, a lack of independence of the supervisory and regulatory capture, regulatory fragmentation, and market fragmentation. It is therefore necessary that further work is done to delineate a set of ethical principles which should be agreed upon at the European Union (EU) level for implementation by the actors involved in regulatory sandboxes. By giving due consideration to ethics in the context of regulatory sandbox initiatives, the EU will be in the position to better ensure that AI products involved in regulatory sandboxes not only do not harm the fundamental rights and freedoms of individuals, but even produce benefits for society.

---

* PhD Student at Maastricht University. Contact email: k.francis@maastrichtuniversity.nl

1. – ETHICS IN REGULATORY SANDBOXES

1.1. – Background

To ensure that regulatory sandboxes promote privacy-enhancing innovation, enable trust, and positively contribute to society through regulatory learning, which is fundamental for their success, there is a need to limit and mitigate both legal and ethical risks. This is particularly relevant in relation to artificial intelligence (AI) which has been shown to perpetuate bias and discrimination, threaten the fundamental rights and freedoms of individuals, and even lead to financial ruin or the loss of life (European Union Agency for Fundamental Rights 2022). This contribution focuses on the need for ethics to be considered in the design and functioning of regulatory sandboxes, the ultimate aim being that of ensuring that technologies such as AI tested within regulatory sandboxes benefit society and do not facilitate injustices or violate the principle of fairness embedded in Article 5(1)(a) of the General Data Protection Regulation (GDPR).

The GDPR calls for personal data processing to be fair, fairness being understood as

> an overarching principle which requires that personal data should not be processed in a way that is unjustifiably detrimental, unlawfully discriminatory, unexpected or misleading to the data subject (European Data Protection Board 2019, 17; Palumbo 2023).

However, such definition of fairness arguably fails to set the bar high enough to ensure that data processing positively contributes to society. At the same time, violations of fairness under the GDPR are poorly enforced (Palumbo 2023). These shortcomings in relation to fairness illustrate the need for taking ethics into greater consideration when it comes to regulating technology and more generally, the processing of personal data.

This section provides an overview of ethical issues and considerations that should be made in relation to regulatory sandboxes. While it makes specific references to AI regulatory sandboxes envisaged in the AI Act, the observations and suggested actions may apply more generally to any regulatory sandbox developed in the European Union (EU).

1.2. – Ethics vs. the law

Before delving into the necessity of embedding ethics into regulatory sandboxes, it is necessary to clarify the difference between ethics and law. Ethics and the law, while related, are distinct. The European Data Protection Supervisor (EDPS), Mr. Wojciech Wiewiórowski, in his speech at the 2024 CPDP conference in Brussels, noted that 'For as much as the law can provide for regarding permitted vs. banned uses of a certain technology, the law will not always draw a specific line of "what is right" from "what is wrong"' (Wiewiórowski 2024). This statement highlights the fundamental necessity of going beyond what is strictly codified in the law to also ensure that AI does 'what is right'. Essentially, the law alone is insufficient for ensuring that new AI technologies and data processing which takes place through them are 'good', 'ethical', and create a societal benefit (Balboni and Francis 2023).

Adhering to ethical principles, however, is no easy feat. This is because ethics is inherently cultural, context-dependent, complex and fluid (Balboni and Francis 2024). Because of its subjective nature and lack of legal consequences for failure to comply, ethics is difficult to enforce (Balboni and Francis 2024). In response to this, in recent years, a plethora of frameworks for AI ethics have been developed.

Examples of ethical frameworks for AI are found in the work of the High-Level Expert Group on Artificial Intelligence (AI HLEG), which drafted, among others, the Ethics Guidelines for Trustworthy AI and an Assessment List for Trustworthy AI (ALTAI). Other examples include UNESCO's Recommendation on the Ethics of Artificial Intelligence, the OECD Principles for Trustworthy AI updated in May 2024, and the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. Floridi and Cowls suggest that such proliferation of ethical principles for AI represents a problem and may even lead to the creation of 'a "market for principles" where stakeholders may be tempted to "shop" for the most appealing ones' (Floridi and Cowls 2019, 2).

As suggested by Prem in analyzing over 100 such frameworks, '[w]hile the frameworks excel in the identification of ethical issues, they are less convincing in providing practical recommendations for implementation and practice' (Prem 2023, 699). While some frameworks are more practical, such as the one developed by Balboni and Francis (2023) which entails a set of five principles, 25 rules, and 44 controls which

can be followed by organizations to engage in ethical data processing activities with the aim of creating a better data-driven world, there is a pressing need to identify and find consensus on European ethical principles and their concrete application.

Fundamentally, it is not enough to merely identify ethical principles to be followed. Instead, clear guidance on how they can be complied with in the context of AI regulatory sandboxes and more generally, in the development and deployment of AI and in the context of any regulatory sandbox must also be agreed upon. The ethical principles agreed upon at the EU level, which may be based on, e.g., those identified by the AI HLEG, and how they are to be complied with, should be made available in written form to the public and to all entities participating in European regulatory sandboxes in order to ensure that the latter contribute to the development of ethical technologies, tools, and services and produce a positive outcome for society.

1.3. – Regulatory sandboxes as a platform for exploring ethics

In the context of regulatory sandboxes, competent authorities should consider their mandate as one aimed at both protecting the fundamental rights of individuals and promoting the ethical development of technology. The EDPS' Opinion 4/2015 Towards a New Digital Ethics Data, Dignity and Technology confirms this in stating that 'In today's digital environment, adherence to the law is not enough; we have to consider the ethical dimension of data processing' (EDPS 2015, 4). Undheim et al. have suggested that regulatory sandboxes represent an ideal place 'for exploring the boundaries of ethics, exploring hypothetical risk and uncertainties, or more fundamentally, fostering a moral imagination' (Undheim et al. 2023, 999).

On a macro level, practical ethical 'dilemmas' to be considered within AI regulatory sandboxes include the extent to which and how risks can be mitigated in the context of regulatory sandboxes. While AI regulatory sandboxes provide an opportunity for competent authorities to contribute to risk mitigation, significant attention must be paid to transparently evaluating risks, identifying effective mitigation measures, and ensuring that such measures are effectively implemented. The accurate identification and subsequent mitigation of risks is highly dependent on transparency, i.e., the quality of communications and the flow of relevant information between the participating

organization and the competent authority. Additionally, the interactions between authorities and organizations participating in sandboxes should be scrutinized, and authorities should be attuned to identifying attempts to engage in regulatory arbitrage, explored in greater detail below. The effectiveness of regulatory sandboxes as a regulatory tool permitting trustworthy innovation should also be explored. This is because, despite being widely considered as a promising tool, large-scale impacts may be limited to a small number of projects. Moreover, if legal, cybersecurity, and ethical risks are not adequately dealt with in regulatory sandbox projects, they will not prove successful in promoting trustworthy innovation in practice.

On a micro level, potential misuses and unintended uses of the technology under development and consequences for the labor market should be taken into consideration. In AI regulatory sandboxes, it is important to analyze the environmental impact of AI, including its energy and resource consumption, as well as the actual benefits and returns of specific AI technologies. This analysis is especially crucial given the significant hype surrounding AI's potential to revolutionize processes and drive growth. The extent to which the benefits of a given technology outweigh the risks, including those arising from potential mis- and untended uses, should furthermore be carefully evaluated from the ethical and fundamental rights perspectives in regulatory sandbox projects.

Ethical principles, some of which are also closely related to legal data protection principles, that should be actively considered within regulatory sandboxes include human agency and oversight, technical robustness and safety (including reproducibility, accuracy, and reliability), privacy and data governance, transparency (including explainability and traceability), diversity, non-discrimination and fairness (including accessibility, stakeholder participation, and the avoidance of bias), societal and environmental wellbeing (including sustainability), and accountability (High-Level Expert Group on AI 2019). The principle of beneficence or the potential for technologies within sandboxes to improve society, solidarity (e.g., solidarity with vulnerable groups), and empowerment are additional ethical AI principles which deserve evaluation within sandboxes (Jobin et al. 2019). To this end, the clear ethical principles and guidelines on how to implement them mentioned in Section 1.2 should be accompanied by a process or procedure for verifying compliance with them. The functioning of the process or verification procedure should also be made available to sandbox participants and enforced by the competent authorities

and overseen at the EU level. It is paramount that the guidelines on how to implement ethical principles are drafted in a way that allows for their implementation to be verified or audited. At the very least, evidence of compliance with the ethical principles should be documented and ideally, should be made available to the public.

Under Article 57(7) of the AI Act, competent authorities are intended to provide participants in AI regulatory sandboxes with guidance on regulatory expectations and information on how they can comply with relevant requirements and obligations under the Act. Additionally, upon request, the authority is meant to provide written proof in relation to the successful completion of sandbox activities as well as an exit report. Authorities should avoid rubber-stamping technologies as being ethical in the context of, e.g., exit reports 'detailing the activities carried out in the sandbox and the related results and learning outcomes' under Article 57(7) of the AI Act, merely because a statement is made by the organization that they comply with ethical principles.

This is particularly relevant insofar as providers are entitled to use documentation provided by the competent authorities to demonstrate their adherence to the requirements of the Act, and such written proof is meant to 'be taken positively into account by market surveillance authorities and notified bodies, with a view to accelerating conformity assessment procedures to a reasonable extent' (see Article 57(7) AI Act).

2. – CHALLENGES AND RISKS IN THE CONTEXT OF REGULATORY SANDBOXES

2.1. – Overview of challenges and risks in AI regulatory sandboxes

Regulatory sandboxes present a number of challenges and risks which are relevant from the ethics perspective. Challenges include those which arise from the Collingridge dilemma, a lack of resources available to regulators, and potential resource shortages within the industry. Risks include ethics-washing, abuse and misuse that could harm fundamental rights and freedoms, lack of independence of the supervisory authority, regulatory capture, regulatory and market fragmentation, placing trust in organizations and technologies that are not truly trustworthy, and insufficient resources to effectively participate in and oversee sandboxes (Baldini and Francis 2024).

Failure to adequately consider and mitigate the aforementioned risks may

undermine the objectives of the EU legislator. For example, if a regulatory sandbox project results in a technology which is later found to not be trustworthy, harms individuals, or has a negative impact on the environment, trust in both the company and in the competent authority will be irrevocably damaged. For this reason, due attention should be paid to them to ensure that regulatory sandboxes contribute to regulating new technologies and fostering innovation and competition in line with the objectives of Article 57(9) of the AI Act. The following subsections deal with these challenges and risks at a high level and propose possible solutions to be taken into consideration by Member States and relevant competent authorities, among others.

2.2. – Collingridge dilemma and regulatory capture

The 'Collingridge dilemma' can be understood as a dilemma in which, on the one hand, it is difficult to regulate technology in an environment where information asymmetry persists and on the other, once technology is developed and its social risks are clear, it is increasingly difficult to regulate (Moraes 2023; Demos Helsinki 2022). AI regulatory sandboxes provide regulators with the opportunity to understand how new AI applications are developed and thus, potentially identify individual, social, and even political risks resulting from them, known as regulatory learning (Kert et al. 2022). As such, they represent an attractive solution in a world where technological developments outpace regulation. They also represent an opportunity to ensure that specific technologies are ethical and produce benefits for citizens, in line with the GDPR's Recital 4 which states that '[t]he processing of personal data should be designed to serve mankind'.

Despite the promising potential of sandboxes to help address the Collingridge dilemma by providing insights into how AI products are developed and tested through interactions between developers, regulators, experts, and consumers—sometimes leading to a more consensual approach to defining applicable rules (Madiega and Van De Pol 2022, 2)—caution should be exercised to avoid overzealous optimism and attempts to exert inappropriate influence over regulators. In particular, it should be ensured that companies participating in sandboxes are truly transparent about how their technologies work and that no misrepresentations are intentionally made to manipulate the guidance of regulators. Similarly, concerned authorities should exercise due caution when 'defining

the applicable rule' to account for attempts to manipulate outcomes. This is because regulatory sandboxes provide ample fuel for regulatory capture, defined as 'the process through which special interests affect state intervention in any of its forms' (Dal Bò 2006, 203).

In the context of regulatory sandboxes, regulatory capture entails the prioritization of innovation and the interests of participating organizations over safeguards which could ensure the protection of rights for individuals and society. This may also lead to a 'race to the bottom' situation where safeguards are lowered with the objective of attracting innovators (Parenti 2020). To avoid regulatory capture and a potential race to the bottom, it is necessary that transparent rules and guidelines for selecting participating organizations and for interacting with such organizations during the sandbox are established (Baldini and Francis 2024).

Article 58(1) of the AI Act, concerning the detailed arrangements and functioning of AI regulatory sandboxes, mandates the Commission to adopt implementing acts for the establishment, implementation, operation, and supervision of AI regulatory sandboxes. These acts will specifically outline principles for the selection and eligibility criteria for participants, as well as procedures for application, participation, monitoring, exiting, and termination of sandboxes, along with the terms and conditions for participants. Such rules and guidelines should be strictly adhered to, and adherence should be verified from time to time thanks to oversight at the EU level. Eligibility requirements should take into consideration the potential for the technology to present ethical and societal risks as well as the technology's potential to produce social or environmental benefits. The implementing acts to be developed should also specifically aim to combat regulatory capture through appropriately detailed rules on the terms and conditions, application, participation, and monitoring processes.

Article 57(12) of the AI Act, insofar as it provides for no administrative fines to be imposed by the authorities for infringements of the AI Act where (i) 'providers observe the specific plan and the terms and conditions for their participation and follow in good faith the guidance given by the national competent authority' and (ii) '[w]here other competent authorities responsible for other Union and national law were actively involved in the supervision of the AI system in the sandbox and provided guidance for compliance', further exemplifies the necessity for the ethical impacts of technologies to

be thoroughly and transparently considered by authorities to reduce potential negative consequences of AI products which may not be subjected to administrative fines.

The implementation of a review process to evaluate the effectiveness of sandboxes in promoting ethical technologies is therefore highly relevant (Parenti 2020, 9; Baldini and Francis 2024). The functioning of the ethics review process should be transparent to allow the public to comprehend potential risks to individuals and society which may be posed as a result of technologies which have gone through the sandbox.

2.3. – Resources

Regulatory sandboxes are costly initiatives which require significant resources, energy, and desire to collaborate (Dutheillet de Lamothe 2024). To ensure regulatory learning which 'enable[s] regulators to gain better regulatory knowledge and to find the best means to regulate innovations based on real-world evidence' (European Commission 2023), both regulators and organizations must have adequate resources at their disposal. Article 57(4) of the AI Act requires Member States to ensure that competent authorities allocate sufficient resources to regulatory sandboxes. However, it is well-known that oversight bodies frequently lack sufficient resources to carry out their activities, as exemplified in the case of the GDPR which has seen relatively limited enforcement. It is thus paramount that budgets for competent authorities are regularly evaluated to ensure that they are sufficient to meaningfully oversee regulatory sandboxes and acquire the necessary technological skills to properly identify and evaluate potential risks as a result of the project.

At the same time, it should be ensured that participating organizations have adequate financial and human resources to meaningfully participate in the regulatory sandbox. This is fundamental to permitting the implementation of legal and ethical measures as suggested by the competent authorities. The participatory criteria for involvement in AI regulatory sandboxes should therefore consider the availability and resources of the organization.

Careful attention should be paid by Member States to facilitate the participation of SMEs and start-ups in AI sandboxes, as suggested by Recital 143 and Article 62(1) of the AI Act. This is highly relevant as such organizations may have more limited access to the legal expertise needed to develop, e.g., AI in a compliant and ethical manner. SMEs and

start-ups may furthermore benefit from targeted guidance on how ethics can be embedded into AI given their relative lack of resources in comparison to larger organizations.

2.4. – Regulatory fragmentation, market fragmentation, and regulatory arbitrage

Regulatory fragmentation, market fragmentation, and regulatory arbitrage represent additional risks in AI regulatory sandboxes. The EU single market risks both market and regulatory fragmentation if the participatory and testing parameters of national AI regulatory sandboxes substantially deviate from one another (Madiega and Van De Pol 2022, 3). Cooperation among competent authorities is central in the text of the AI Act, with the legislator's aim to combat fragmentation. Article 58(4) of the AI Act states that:

> Where national competent authorities consider authorising testing in real world conditions supervised within the framework of an AI regulatory sandbox [...] they shall specifically agree the terms and conditions of such testing and, in particular, the appropriate safeguards with the participants, with a view to protecting fundamental rights, health and safety. Where appropriate, they shall cooperate with other national competent authorities with a view to ensuring consistent practices across the Union.

Article 57(13) of the AI Act calls for AI sandboxes to 'be designed and implemented in such a way that, where relevant, they facilitate cross-border cooperation between national competent authorities'; Article 57(14) requires that national competent authorities 'coordinate their activities and cooperate within the framework of the Board'; and Article 57(15) promotes cross-border cooperation and interaction through the publication of a list of planned and existing sandboxes by the AI Office. Effective actuation of these articles through good cooperation and communication is fundamental to ensure that a common minimum standard of what is acceptable is respected across the EU, also in terms of compliance with ethical principles. At the EU level, it is therefore paramount that national authorities work together and share best practices, including in relation to ethics. By effectively collaborating and communicating, potential negative impacts of fragmentation which go against the spirit of the AI Act, and which would thus

undermine the objectives of the European legislator can be better avoided.

A significant risk which may arise from fragmentation in approaches and failure to appropriately coordinate is that of regulatory arbitrage. Regulatory arbitrage is defined as a strategy employed by organizations 'that can be used to achieve an economically equivalent outcome to a regulated activity while avoiding the legal constraints (colloquially, complying with the letter but avoiding the spirit of the law)' (Allen 2020, 309). If baseline rules for what is acceptable and what is not within AI regulatory sandboxes are not established or uniformly respected across the EU, also in terms of ethics principles, organizations may engage in 'forum-shopping', i.e., seeking out 'more lenient' sandboxes that fit their objectives (OECD 2023, 18). Such a situation would undermine the objectives of sandboxes, the AI Act itself, and more generally, the single market.

To avoid fragmentation, Member States should take the opportunity that Article 57 of the AI Act provides to jointly establish sandboxes with the competent authorities of other states. Best practices and the transparent sharing of best practices within sandboxes should be shared in a systematic and effective manner. The AI Office will also likely play an important role in avoiding fragmentation and regulatory arbitrage and should be highly attuned to such risks, taking relevant actions in a timely manner to mitigate potential instances of diverging application standards.

2.5. – Risk - and ethics - washing

To combat risk-washing and ethics-washing, the ethics guidelines identified under Section 1.2 should be carefully drafted by individuals with competencies in ethics and updated as necessary according to societal and technological developments. Risk-washing is defined as a 'regulatory institution's making products or processes of a company seem to involve less risk for stakeholders by engaging in activities that mimic in a superficial or narrow way genuine attempts to assess and reduce risk' (Brown and Piroska 2021, 20). AI ethics washing is defined as 'the phenomenon of instrumentalising ethics by misleading communication, creating the impression of ethical Artificial Intelligence (AI), while no substantive ethical theory, argument, or application is in place or ethicists involved' (Schultz et al. 2024, 1).

To limit ethical risks and promote truly ethical technologies, competent authorities

should involve not only legal and technical experts and economists, but also ethicists and sociologists in the teams tasked with regulatory sandbox activities. Ethicists and sociologists should also be involved in the process of developing the ethical principles and guidelines to implement them. This is because ethicists and sociologists are better suited to identify potential ethical and societal risks of technologies that legal and technical professionals may not be attuned to.

The importance of multidisciplinary teams in evaluating risks stems from the wide range of potential use cases of new technologies and their resulting impacts. AI, for example, has a complex value chain, impacting a wide array of processes and people, and requires thinking 'beyond status quo ideas' (IEEE Standards Association 2023, 7). By including ethicists and sociologists in their regulatory sandbox teams, competent authorities can benefit from social and philosophical knowledge and best practices that fall outside of their scope of expertise. Furthermore, because sociologists and ethicists are attuned to different kinds of risks, which are more social and societal in their nature as opposed to legal or technical, their involvement in regulatory sandboxes will allow for risks to be identified and evaluated in a more comprehensive manner, thus also allowing for more impactful mitigation measures to be identified. Such an approach is in line with the objective of the European Commission in fostering the development of trustworthy technologies and may lead to increased trust from the public in relation to new technologies, regulatory sandboxes, and more generally, promote trust in the competent authorities.

Along these lines, the Norwegian Data Protection Authority, which has run a regulatory sandbox for 'privacy-enhancing innovation and digitalization' which focused on AI in recent years, has adopted a best practice which involves making use of external experts (Datatilsynet n.d.); Baldini and Francis 2024, 11). According to Markussen (2023, 17), the Norwegian authority's sandbox selection committee is comprised of

> an internal, interdisciplinary group that conducts interviews with all applicants. An external reference group, comprising members from Innovation Norway, the Norwegian Computing Centre, the Equality and Anti-Discrimination Ombud and Tekna, will assist in assessing the public benefit of the potential projects. The final selection of

> projects accepted into the sandbox will be made by the steering committee, made up by the Authority's management.

National competent authorities should be encouraged to involve an array of individuals with different backgrounds to better facilitate the identification of risks posed by technologies being tested within regulatory sandboxes. For example, the Norwegian Data Protection Authority also involves communications consultants in its sandbox activities when their expertise is deemed relevant, which may also be considered a best practice in addition to involving social scientists and ethicists (Datatilsynet 2021). Concretely, as suggested above, the teams which competent authorities task with working on each sandbox project should include ethicists and sociologists in addition to engineers and legal experts. Such individuals should be involved in all phases of regulatory sandbox projects, starting from the selection phase, and their expertise should be duly considered through the exit report phase to ensure coherence and the comprehensive implementation of measures and safeguards which if not dealt with, could undermine both the sandbox project and the objectives of the legislator identified in Article 57(9) of the AI Act.

Furthermore, clear guidelines should be drafted at the EU level and made available to sandbox participants in relation to communicating their compliance, both legal and ethical, after the conclusion of their participation in the sandbox to avoid making claims which may lead the public to think that certain technologies are without risk or are fully compliant with applicable law.

3. – CONCLUSION

Regulatory sandboxes represent a promising regulatory innovation. They have great potential because they facilitate regulatory learning and may contribute to mitigating the Collingridge dilemma. In relation to the potential for technologies to 'do good', sandboxes are especially promising because they allow authorities to encourage the adoption of ethical practices for the benefit of society. Despite offering many possible benefits, regulatory sandboxes also present risks which must be dealt with to ensure that the objectives of the EU legislator are met.

At the EU level, a set of ethical principles should be agreed upon for implementation

by the actors involved in regulatory sandboxes to better ensure that products involved in sandboxes not only do not cause harm, but even produce benefits for society. It should also be made clear to organizations how they can comply with the ethical principles agreed upon, i.e., how they can implement ethical principles in practice.

Competent authorities should take ethics into consideration as opposed to merely considering the letter of the law due to the potential for them to enhance fundamental rights protections through a multifaceted approach that also includes ethical considerations. Among others, the rules and guidelines for selecting participating organizations and for interacting with such organizations should be transparent. Eligibility requirements for organizations to participate in sandboxes should also include an evaluation of the specific technology's potential to present ethical and societal risks as well as the technology's potential to produce social or environmental benefits.

The AI Office may play an important role in contrasting regulatory capture and fragmentation through the support it will provide to Member State governance bodies. Specifically, the Regulation and Compliance Unit of the AI Office is charged with coordinating the regulatory approach to permit consistent enforcement and a uniform application of the AI Act throughout the EU (European Commission 2024). However, the potential for fragmentation to be realized should not be underestimated and should actively be contrasted through specific measures and good communication and transparency.

To ensure their effectiveness, the budgets of competent authorities should be evaluated on a regular basis to permit adequate oversight. Competent authorities should also work closely with ethicists and sociologists who may be better suited to identify ethical and societal risks in the context of sandboxes, as should the AI Office. Clear guidelines for how organizations may represent their participation in sandboxes should be made available to organizations participating in sandboxes to avoid ethics washing on the part of organizations. Finally, transparent and effective cooperation and communication between authorities is paramount to ensuring consistency and avoiding fragmentation across the EU.

This contribution has focused on the need for ethics to be considered in the design and functioning of sandboxes. By carefully considering ethics in the context of regulatory sandboxes and taking concrete steps towards embedding ethics in their operation, the

EU will be able to facilitate the development of technologies which are 'fair' in the sense of Article 5(1)(a) GDPR. Ethics and the law go hand-in-hand and if effectively operationalized together, have the power to build a better and safer digital future for Europe.

BIBLIOGRAPHY

ALLEN H.J. (2019), *Regulatory Sandboxes*, in George Washington Law Review, vol. 87, n. 3.

ALLEN H.J. (2020), *Sandbox Boundaries*, Vanderbilt Journal of Entertainment & Technology Law, 22(2), p. 299-321, available here.

AMODEI D., OLAH C., STEINHARD J., CHRISTIANO P., SCHULMAN J., MANÉ D. (2016), *Concrete problems in AI safety*, available here.

ANDRADE N., ANTTILA J., GALINDO L., GRONCHI J., SCOTT S., SIGORA J., CHA WEI QUAN R., ZARRA A. (2022), *Towards an Experimental Governance Framework for Emerging Technologies,* Chapter 1: Learning from the past, available here.

ANDRADE N., ANTTILA J., GALINDO L., GRONCHI J., SCOTT S., SIGORA J., CHA WEI QUAN R., ZARRA A. (2023A), *Towards an Experimental Governance Framework for Emerging Technologies.* Chapter 2: Interpreting The Present, available here.

ANDRADE N., GALINDO L., ZARRA A. (2023B), *Artificial Intelligence Act: A Policy Prototyping Experiment EU AI Regulatory Sandboxes*, available here

ANDRADE N., ZARRA A. (2022). *Artificial Intelligence Act: A Policy Prototyping Experiment: Operationalizing the Requirements for AI Systems–Part I*, available here

ARJOON S. (2006), *Striking a Balance Between Rules and Principles-based Approaches for Effective Governance: A Risks-based Approach*, in Journal of Business Ethics 2024, p. 53 ff.

ARMSTRONG H., BÁRD I., & ENGSTRÖM E. (2020), *Regulator approaches to facilitate, support, and enable innovation*, UK Department for Business, Energy & Industrial Strategy (BEIS).

ARTICLE 29 WORKING PARTY (2014), *Opinion 05/2014 on Anonymisation Techniques*, available here.

ARTICLE 29 WORKING PARTY (2017), *Guidelines on Data Protection Impact Assessment (DPIA)*, available here.

ATTREY A., LESHER M., LOMAX C. (2020), *The role of sandboxes in promoting flexibility and innovation in the digital age,* OECD Going Digital Toolkit Policy, Note N. 2, available here.

BAGNI F. (2023), *The Regulatory Sandbox and the Cybersecurity Challenge: from the Artificial Intelligence Act to the Cyber Resilience Act*, in Rivista Italiana di Informatica e Diritto, n. 2, available here.

BAGNI F., SEFERI F. (in the process of publication), *Commentario Regolamento sull'Intelligenza Artificiale (Artificial Intelligence Act)* - Articles 57 and 58, a cura di Riccio-Resta-Mantelero.

BALBONI P., FRANCIS K. (2023), *Data Protection as a Corporate Social Responsibility*, Edward Elgar.

BALBONI P., FRANCIS K. (2024 – Forthcoming), *Data ethics and digital sustainability: Bridging legal data protection compliance and ESG for a responsible data-driven future*, Unpublished 2024.

BALDINI D., FRANCIS K. (2024), *AI Regulatory Sandboxes between the AI Act and the GDPR: the role of Data Protection as a Corporate Social Responsibility*, ITASEC 2024: The Italian Conference on CyberSecurity, April 08–11, 2024, Salerno, Italy, available here.

BALDWIN R., CAVE M., & LODGE M. (2010), *Introduction: Regulation - The Field and the Developing Agenda*, In Id. (Ed.), The Oxford Handbook of Regulation, p. 3-16, Oxford: Oxford University Press.

BANAKAS E. (2002), *The Contribution of Comparative Law to the Harmonisation of European Private Law*, in Comparative Law in the 21st Century, (A. Harding -E. Orucu, eds.), Kluwer Law International, London/The Hague/New York, 2002, p. 179 ff.

BAUKNECHT D., BISCHOFF T. S., BIZER K., FÜHR M., GAILHOFER P., HEYEN D. A., PROEGER T., & VON DER LEYEN K. (2020), *Exploring the pathways: Regulatory experiments for*

sustainable development – an interdisciplinary approach, Journal of Governance and Regulation, 9(3), p. 49–71.

BENNETT MOSES L. (2013), *How to Think About Law, Regulation and Technology: Problems with 'Technology' as a Regulatory Target*, in "Law, Innovation & Technology", vol. 5, n. 1.

BMWI - FEDERAL MINISTRY FOR ECONOMIC AFFAIRS AND ENERGY (2021), *Making space for innovation: The handbook for regulatory sandboxes*, 2021.

BOGUCKI A., ENGLER A., PERARNAUD C., RENDA A., T*he AI act and the emerging EU Digital Acquis. Overlaps, gaps and inconsistencies*, September 2022, Cesps, available here.

BOMHARD D., SIGLMÜLLER J. (2024), *AI Act – das Trilogergebnis, RDi 2024*, p. 45 ff.

BRADFORD A. (2023), *Digital empires: The global battle to regulate technology* (Oxford University Press 2023).

BRINKER N. (2024), *Identification and demarcation - A general definition and method to address information technology in European IT security law*, Computer Law and Security Review April 2024.

BROMBERG L., GODWIN A., RAMSAY I. (2017), *Fintech sandboxes: Achieving a balance between regulation and innovation,* in Journal of Banking and Finance Law and Practice, 4, 2017, p. 314-336.

BROWN E., PIROSKA D. (2021), *Governing Fintech and Fintech as Governance: The Regulatory Sandbox, Riskwashing, and Disruptive Social Classification*, in New Political Economy, p. 1-13.

BROWN E., PIROSKA D. (2021), *Governing Fintech and Fintech as Governance: The Regulatory Sandbox, Riskwashing, and Disruptive Social Classification, New Political Economy*, 27(1), p. 19–32, available here.

BUOCZ T., PFOTENHAUER S., EISENBERGER I. (2023), *Regulatory sandboxes in the AI Act: reconciling innovation and safety? in Law, Innovation and Technology*, n.2, p. 357-389.

BUTENKO A., & LAROUCHE P. (2015), *Regulation for innovativeness or regulation of innovation? in Law, Innovation and Technology*, 7(1), p. 52-82.

BÜYÜKSAGIS E., VAN BOOM W.H. (2013), *Strict Liability in Contemporary European Codification: Torn between Objects, Activities, and Their Risks*, Georgetown Journal of International Law, 2013, p. 609 ff.

CEN-CENELEC (2023), *Innovative Process for homegrown Harmonized Standards*, available here.

CHIARA P.G. (2022), *The Cyber Resilience Act: the EU Commission's proposal for a horizontal regulation on cybersecurity for products with digital elements*, an introduction, in "International Cybersecurity Law Review", n. 3.

CHIARA P.G. (2024), *Towards a right to cybersecurity in EU law? The challenges ahead, in Computer Law & Security Review: The International Journal of Technology Law and Practice*, available here.

CLEMMENSEN L. H., KJÆRSGAARD R. D. (2023), *Data Representativity for Machine Learning and AI Systems,* available here.

COLLINGRIDGE, D. (1982), *The social control of technology*, London: Francesc Pinter.

COLOMBO C., ELIANTONIO M. (2017), *Harmonized technical standards as part of EU law: Juridification with a number of unresolved legitimacy concerns?*, in Maastricht Journal of European and Comparative Law, 24(2), p. 323–340, available here.

CORNELLI G., DOERR S., GAMBACORTA L., MERROUCHE O. (2024), *Regulatory Sandboxes and Fintech Funding: Evidence from the UK*, Review of Finance, Volume 28, Issue 1, January 2024, p. 203–233, available here.

COSTANTINI F. (2021), *L'esperienza giuridica nella società dell'informazione contemporanea: il problema delle «norme sperimentali*, in A. Scerbo (ed.), Teoria e prassi dell'esperienza giuridica. In ricordo di Francesco Gentile, Napoli, Edizioni Scientifiche Italiane, p. 165-184.

COUNCIL DIRECTIVE (EU), 2022/2555 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2Directive) [2022] OJ L 333/80, further referred to as NIS 2.

COUNCIL OF THE EUROPEAN UNION (2024), *Draft minutes Council of the European Union (Transport, Telecommunications and Energy)* 21 May 2024, available here.

COUNCIL OF THE EUROPEAN UNION (2024), *Data protection: Council agrees position on GDPR enforcement rules*, Consilium, available here.

COUNCIL REGULATION (EU), 2016/679 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 L119/1, further referred to as GDPR.

COUNCIL REGULATION (EU), 2024/1689 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).

COUNCIL REGULATION (EU), of 17 April 2019 on ENISA (the EuropeanUnion Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/ 2013 (Cybersecurity Act) [2019] OJ L151/15 further referred to as CSA.

CUCCURU P. (2020), *Regulating by Request: On the Role and Status of the Standardisation Mandate under the New Approach*, in M. Eliantonio & C. Cauffman (Eds.), The Legitimacy of Standardisation as a Regulatory Technique, Edward Elgar Publishing, available here.

DAL BO' E. (2006), *Regulatory Capture: A Review, Oxford Review of Economic Policy*, 22(2), p. 203-225, available here.

DATATILSYNET (2021), *Framework for the Regulatory Sandbox*, available here.

DATATILSYNET (n.d.), *Regulatory privacy sandbox*, available here.

DATENSCHUTZKONFERENZ (DSK) (2018), *Kurzpapier Nr. 5: Datenschutz-Folgenabschätzung nach Article 35 DS-GVO*, available here.

DE PASQUALE P. (2023), T*he principle of legal certainty in EU law. Diritto pubblico comparato ed europeo*, Rivista trimestrale, 2(2023), p. 405–422.

DEMOS HELSINKI (2022), *What is the Collingridge dilemma and why is it important for tech policy?*, available here.

DIZON M. A. C. (2012), *From regulating technologies to governing society: Towards a plural, social and interactive conception of law*, in H. M. Morgan & R. Morris (Eds.), Moving Forward: Tradition and Transformation p. 115-139. Cambridge: Cambridge University Press.

DOWNES L. (2009), *The laws of disruption: Harnessing the new forces that govern life and business in the digital age*, New York: Basic Books.

DUCATO R. (2023), *Why Harmonised Standards Should Be Open, in IIC - International Review of Intellectual Property and Competition Law*, 54(8), p. 1173–1178, available here.

DUTHEILLET DE LAMOTHE L. (2024), *How Can Compliance Sandboxes Help You panel, Speech given at the Privacy Symposium in Venice on 12 June 2024*, available here.

ESA – EUROPEAN SUPERVISORY AUTHORITIES (2023), *Update on the functioning of innovation facilitators – innovation hubs and regulatory sandboxes*, Europe, available here.

EUROPEAN COMMISSION (2019), *Directorate-General for Communications Networks, Content and Technology, Ethics guidelines for trustworthy AI*, Publications Office, 2019, available here.

EUROPEAN COMMISSION (2023a), *Better Regulation Toolbox*, Europe, available here.

EUROPEAN COMMISSION (2023b), *Commission Staff Working Document, 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 final, available here.

EUROPEAN COMMISSION (2024), *The future of European competitiveness*, Report by Mario Draghi, September 2024.

EUROPEAN COMMISSION (2024), *Public sector tech watch: Mapping innovation in EU public services*, publications Office of the European Union, Luxembourg, ISBN 978-92-68-11627-2, available here.

EUROPEAN COMMISSION (2024), *Commission establishes AI Office to strengthen EU leadership in safe and trustworthy Artificial Intelligence*, available here.

EUROPEAN DATA PROTECTION SUPERVISOR (2015), *Opinion 4/2015 Towards a new digital ethics Data, dignity and technology,* available here.

EUROPEAN DATA PROTECTION BOARD (2019), *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default Version 2.0*, Adopted on 20 October 2020, available here.

EUROPEAN DATA PROTECTION BOARD - EDPB (2020), *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default*, available here.

EUROPEAN DATA PROTECTION BOARD - EDPB (2021-a), *Guidelines 07/2020 on the concepts of controller and processor in the GDPR*, available here.

EUROPEAN DATA PROTECTION BOARD - EDPB (2021-b), *Guidelines 08/2020 on the targeting of social media users*, available here.

EUROPEAN DATA PROTECTION SUPERVISOR – EDPS (2022), *TechSonar 2021-2022 Report*, available here.

EUROPEAN DATA PROTECTION BOARD - EDPB (2023), *EDPB Work Programme for 2023-2024*, available here.

EUROPEAN PARLIAMENT, legislative resolution of 12 March 2024 on the proposal for a regulation of the European Parliament and of the Council on horizontal cybersecurity requirements for products with digital elements and amending Regulation (EU) 2019/1020 (COM(2022)0454 – C9-0308/2022 – 2022/0272(COD)), further referred to as CRA.

EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS (2022), *Bias in Algorithms: Artificial Intelligence and Discrimination*, available here.

FALLETTA P., MARSANO A. (2024), *Intelligenza artificiale e protezione dei dati personali: il rapporto tra Regolamento europeo sull'intelligenza artificiale e GDPR*, in Rivista Italiana di Informatica e Diritto, 1/2024, available here.

FEDERAL OFFICE FOR INFORMATION SECURITY (BSI) (2017), BSI Standard 200-3: Risk Analysis based on IT Grundschutz, Bonn 2017.

FENWICK M., VERMEULEN E.P.M., CORRALES M. (2018), *Business and Regulatory Responses to Artificial Intelligence: Dynamic Regulation, Innovation Ecosystems and the Strategic Management of Disruptive Technology*, in M. Corrales, M. Fenwick, N. Forgó (eds.), "Robotics, AI and the Future of Law", Springer.

FLORIDI L., COWLS J. (2019), A *Unified Framework of Five Principles for AI in Society*, Harvard Data Science Review, 1(1), available here.

FREITAS M. C., MIRA DA SILVA M. (2018). *GDPR Compliance in SMEs: There is much to be done*, Journal of Information Systems Engineering & Management, 3(4), p. 30 ff., available here.

GANGALE F., MENGOLINI A.M., COVRIG L., CHONDROGIANNIS S., SHORTALL R. (2023), *Making energy regulation fit for purpose. State of play of regulatory experimentation in the EU*, Publications Office of the European Union, Luxembourg, doi:10.2760/32253, available here.

GARANTE PER LA PROTEZIONE DEI DATI PERSONALI – GPDP (2024), *Artificial intelligence: the Italian Data Protection Authority writes to Parliament and Government. Need for independent and impartial supervisory authorities*, available here.

GARBEN S., GOVAERE I., *The EU Better Regulation Agenda: A critical assessment*, Oxford, Hart Publishing, 2018.

GERMAN FEDERAL MINISTRY FOR ECONOMIC AFFAIRS AND CLIMATE ACTION (BMWK) (2022), *Regulatory Sandboxes – Enabling Innovation and Advancing Regulation*, available here.

GORDLEY J. (2015), The Architecture of the Common and Civil Law of Torts: An Historical Survey, in M. Bussani and A. J. Sebok (eds.) Comparative Tort Law: Global Perspectives, Elgar Publishing, available here.

GORYWODA L. (2009), *The New European Legislative Framework for the Marketing of Goods*, available here.

GROMOVA E. A., STAMHUIS E. (2023), *Real-Life Experimentation with Artificial Intelligence, in A. Quintavalla & J. Temperman (Eds.), Artificial intelligence and human rights*, Oxford: Oxford University Press.

HACKER P., BERZ A. (2023), *Der AI Act der Europäischen Union – Überblick*, Kritik und Ausblick, ZPR 2023, p. 226 ff.

HACKER, P. (2023), *Die Regulierung von ChatGPT et al. – ein europäisches Trauerspiel*, GRUR 2023, p. 289 ff.

HAMMON C., BUYERS J., DOQUER B., SCHEFZIG J., VICKERY K., SHARPE T., KIRSCHKE BILLER J. (2023), I*s the proposed European AI Act innovation-friendly?*, available here.

HART H. L. A. (1961), *The Concept of Law (I ed.)*. Oxford: Oxford University Press.

HELDEWEG M.A. (2015), *Experimental legislation concerning technological & governance innovation – An analytical approach*, in "The Theory and Practice of Legislation", vol. 3, n. 2.

HIGH-LEVEL EXPERT GROUP ON AI. (2019), *Ethics Guidelines for Trustworthy Artificial Intelligence*, available here.

HIGH-LEVEL EXPERT GROUP ON AI (2020), *Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment*, available here.

HOFMANN H. C., ZETZSCHE D. A., & PFLÜCKE F. (2022), *The changing nature of 'Regulation by Information': Towards real-time regulation?*, in European Law Journal, 28(4-6), p. 172-186.

IEEE STANDARDS ASSOCIATION (2023), *The Benefits of a Multidisciplinary Lens for Artificial Intelligence Systems Ethics*, available here.

INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS (n.d.), *The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems*, available here.

ISO/IEC (2018), ISO/IEC 27005:2018(E) *Information technology - Security techniques - Information security risk management*, Switzerland 2018.

ISO/IEC (2022), *Information security, cybersecurity and privacy protection - Guidance on managing information security risks*, Switzerland.

JANSSEN H., SENG AH LEE M., SINGH J. (2022), *Practical fundamental rights impact assesments,* in International Journal of Law and Informations Technology 2022, p. 200 ff.

JARA A., MARTINEZ I., SANCHEZ J. (2024), *Cybersecurity Resilience Act (CRA) in practice for IoT devices: Getting ready for the NIS2*, available here.

JENIK I., LAUER K. (2017), *Regulatory sandboxes and financial inclusion*, working Paper, Washington, D.C.: CGAP. available here.

JOBIN A., IENCA M., VAYENA E. (2019), *The global landscape of AI ethics guidelines. Nature Machine Intelligence*, 1, available here.

KERT K., VEBROVA M., SCHADE S. (2022), *Regulatory learning in experimentation spaces*, European Commission, available here.

KOZIOL H., B.C. STEININGER B. C. (2016), *European Tort Law Yearbook, Institute for European Tort Law*, ECTIL, De Gruyter, 2016.

LEIBNIZ G. W. (2020), *Discourse on metaphysics (New edition Translated with Introduction and commentary by Rodriguez-Pereyra*, G.). Oxford: Oxford University Press.

LESSIG L. (2006), *Code. Version 2.0. New York*, Basic Books.

LÓPEZ C.A.F., ELBI A. (2022), *On the Legal Nature of Synthetic Data*, available here.

LUHMANN N. (2008), *Law as a social system (translated by Ziegert, K. A.)*, Oxford: Oxford University Press.

MACCABIANI N. (2022), *The European path towards Data Quality and its standardisation in AI: A legal perspective*, in BioLaw Journal - Rivista Di BioDiritto, 4, Article 4, available here.

MADIEGA T., VAN DE POL A.L. (2022), *Artificial intelligence act and regulatory sandboxes*, European Parliamentary Research Service, available here.

MALGIERI G. (2019), *Automated Decision-Making in the EU Member States. The right to Explanation and other "suitable safeguards" for Algorithmic Decisions in the EU National Legislations*, in Computer Law & Security, vol. 35, n. 5.

MANTELERO A. (2020), T*he future of data protection: Gold standard vs. global standard*, 40 Computer Law & Security Report 2020, n. 1.

MARKUSSEN T. (2023), *Evaluation of the Norwegian Data Protection Authority's Regulatory Sandbox for Artificial Intelligence*, Datatilsynet, available here.

MCFADDEN M., JONES K., TAYLOR E., OSBORN G. (2021), *Harmonising Artificial Intelligence:*

The role of standards in the EU AI Regulation, Oxford Information Labs, available here.

MFSA (2023), *Fintech Regulatory Sandbox*, available here.

MICKLITZ H. (2023), *AI Standards*, EU Digital Policy Legislation and Stakeholder Participation, EuCML 2023, p. 212 ff.

MIT AI RISK REPOSITORY (2024), *AI Risk Repository – A Comprehensive Database of Risks from AI Systems*, available here.

MORAES T. (2023), *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 - Official Outcome of the UN IGF Data and Artificial Intelligence Governance Coalition, 2023.

MOUSMOUTI M. (2018), *Making Legislative Effectiveness an Operational Concept: Unfolding the Effectiveness Test as a Conceptual Tool for Lawmaking*, in "European Journal of Risk Regulation", vol. 9, 2018, n. 3, available here.

NESTA (2020), *Regulators' experimentation toolkit. Centre for Regulatory Innovation*, available here.

NOLTE H., RATEIKE M., FINCK M. (2024), *Robustness and Cybersecurity in the EU Artificial Intelligence Act*, available here.

NOVELLI C., CASOLARI F., HACKER P., SPEDICATO G., FLORIDI L. (2024), *Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity*, available here.

NOZICK R., *Anarchy, State, and Utopia*, Basic Books, 1974.

NUTHI K. (2022), *An Overview of the EU's Cyber Resilience Act*, Center for data and innovation, available here.

OECD (2018), *OECD Regulatory Policy Outlook 2018*, OECD Publishing, Paris, available here.

OECD. (2019), *Principles for trustworthy AI*, available here.

OECD (2023), *Regulatory sandboxes in artificial intelligence*, OECD Digital Economy Papers, No. 356, OECD Publishing, Paris, available here.

OMAROVA S.T. (2020), T*echnology v Technocracy: Fintech as a Regulatory Challenge*, in Journal of Financial Regulation, vol. 6, n. 1.

OWASP (2024), *OWASP Top 10 for Large Language Model Applications*, available here.

PAGALLO U., CASANOVAS P., MADELIN P. (2019), *The middle-out approach: assessing models of legal governance in data protection, artificial intelligence, and the Web of Data*, in The Theory and Practice of Legislation, 1, p. 1–25.

PALUMBO A. (2023), *The unexplored potential of the fairness principle under the GDPR: lessons from the recent TikTok case*, KU Leuven Centre for IT & IP Law blog post, 10 October 2023, available here.

PARENTI R. (2020), *Regulatory Sandboxes and Innovation Hubs for FinTech*, study for the committee on Economic and Monetary Affairs, Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament, Luxembourg, available here.

PELLEGRINO M. (2022), *The AI Act: help or hindrance for SMEs? An analysis of the cost of compliance with the AI Act for SMEs,* available here.

PELKMANS J. (1987), T*he New Approach to Technical Harmonization and Standardization*, in JCMS: Journal of Common Market Studies, 25(3), p. 249–269, available here.

PORTALIER P. (2017), *Myths and realities of the presumption of conformity,* available here.

PREM E. (2023), *From ethical AI frameworks to tools: a review of approaches, AI and Ethics*, 3, p. 699–716, available here.

EUROPEAN GROUP ON TORT LAW (n.d.), Principles of European Tort Law, available here.

RAMARINE E. (2015), *Understanding Problems of Subjectivity and Uncertainty in Quality Risk Management*, In: Journal of Validation Technology, p. 21 ff.

RANCHORDAS S. (2015), *Innovation-Friendly Regulation: The Sunset of Regulation, the Sunrise of Innovation*, in Jurimetrics, vol. 55, n. 2.

RANCHORDAS S. (2021a), *Experimental Regulations for AI: Sandboxes for Morals and Mores*, University of Groningen Faculty of Law Research Paper No. 7/2021, 2021, available here.

RANCHORDAS S. (2021b), *Experimental lawmaking in the EU: Regulatory Sandboxes,* in University of Groningen Faculty of Law Research Paper Series, 12, p. 1-10, available here.

RANCHORDAS S. (2021c), *Experimental Regulations and Regulatory Sandboxes*, available here.

RANCHORDAS S., VINCI V. (2024), *Regulatory sandboxes and innovation - Friendly regulation: between collaboration and capture*, Italian Journal of Public Law, Issue 1/2024.

RAUDLA R., JUUSE E., KUOKŠTIS V., CEPILOVS A., CIPINYS V., YLÖNEN M. (2024), *To sandbox or not to sandbox? Diverging strategies of regulatory responses to FinTech*, in Regulation & Governance, 2024.

REDAELLI C.M., *Better regulation and the Lisbon Agenda*, Londra, 2006, available here.

RENDA A., et al (2021), *Study to support an impact assessment of regulatory requirements for artificial intelligence in Europe*, European Commission: Brussels, Belgium.

RESTA G. (2024), *"So Lonely": Comparative Law and the Quest for Interdisciplinary Legal Education*, in International Journal For The Semiotics Of Law, Vol. 37, p. 1569-1586, available here.

REUSCH P. (2023), *KI und Software im Kontext von Produkthaftung und Produktsicherheit*, RDi 2023, p. 152 ff.

RICCIO G.M. (2024), *Data protection and appropriate measures: too many uncertainties in the judicial*

applications?, in UNIO EU Law Journal, 2024, p. 17 ff.

RINGE W.G. (2023), *So ließe sich Künstliche Intelligenz intelligent regulieren*, available here.

RITTER S. (2023), § 2 BSIG, in KIPKTER K., REUSCH P., RITTER, S., *Recht der Informationssicherheit*, C.H. Beck (2023), München.

SARTOR G., LAGIOIA F. (2020), *The Impact Of The General Data Protection Regulation (GDPR) On Artificial Intelligence, European Parliament Research Service*, available here.

SCHEIDEGGER F., VON THIESSEN R., WEISS R. (2023), *Automated infrastructure maintenance – Drone inspections with computer vision,* available here.

SCHERER M.U. (2016), *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, Harvard Journal of Law & Technology, Vol. 29, No. 2, Spring 2016, available here.

SCHUETT J., ANDERLJUNG M., CARLIER A., KOESSLER L., GARFINKEL B. (2024), F*rom Principles to Rules: A Regulatory Approach for Frontier AI, forthcoming in HACKER P, ENGEL A., HAMMER S., MITTELSTADT (eds)*, The Oxford Handbook on the Foundations and Regulation of Generative AI, Oxford 2024.

SCHULTZ M.D., CONTI L.G., SEELE P. (2024), *Digital ethicswashing: a systematic review and a process-perception-outcome framework*, AI Ethics, available here.

SHAFFIQUE M. R. (2024), *Cyber Resilience Act 2022: A silver bullet for cybersecurity of IoT devices or a shot in the dark?*, in Computer Law & Security Review: The International Journal of Technology Law and Practice, available here.

SIGLMÜLLER J. (2023), *Cyber Resilience Act und Digital Operational Resilience Act – lässt sich IT-Sicherheit rechtlich erzwingen?*, in Tagungsband Herbstakademie 2023, Oldenburg, OLWIR, 2023, p. 339 ff.

SIMONCINI A. (2023), *Il linguaggio dell'intelligenza artificiale e la tutela costituzionale dei diritti*, in Riv. Associazione italiana dei costituzionalisti, 2023, n. 2.

SMUHA N.A. (2021), *From a 'Race to AI' to a 'Race to AI Regulation' - Regulatory Competition for Artificial Intelligence*, in Law, Innovation and Technology, vol. 13, n. 1.

SPINDLER G. (2023), *Different Approaches for Liability of Artificial Intelligence – Pros and Cons – the New Proposal of the EU Commission on Liability for Defective Products and AI Systems*, available here.

STIGLER G. J. (1971), *The Theory of Economic Regulation*, The Bell Journal of Economics and Management Science, vol. 2, no. 1, 1971, p. 3–21, available here.

STUDY FOR THE COMMITTEE ON ECONOMIC AND MONETARY AFFAIRS (2020), *Policy Department for Economic, Scientific and Quality of Life Policies,* European Parliament, Luxembourg, 2020, p. 24 ff.

TARTARO A. (2023a), *Regulating by standards: Current progress and main challenges in the standardisation of Artificial Intelligence in support of the AI Act, in | European Journal of Privacy Law & Technologies*, 1, Article 1, available here.

TARTARO A. (2023b), *Towards European Standards supporting the AI Act: Alignment challenges on the path to Trustworthy AI*, in Proceedings of the AISB Convention 2023, p. 98–106, available here.

THELISSON E., VERMA H. (2024), *Conformity assessment under the EU AI act general approach, in AI and Ethics*, 4(1), p. 113–121, available here.

TRUBY J., BROWN R.D., IBRAHIM I.A., CAUDEVILLA PARELLADA O. (2022), *A Sandbox Approach to Regulating High-Risk Artificial Intelligence Applications*, in 13 European Journal of Risk Regulation, p. 270 ff.

UK CIVIL AVIATION AUTHORITY (2024), *Regulatory Sandbox for the development of capabilities to integrate Unmanned Aerial Systems (UAS) in unsegregated airspace*, available here.

UNDHEIM K., ERIKSON T., TIMMERMANS B. (2023), *True uncertainty and ethical AI: regulatory sandboxes as a policy tool for moral imagination, AI and Ethics*, 3, p. 997–1002, available here.

UNESCO (2023), *Recommendation on the Ethics of Artificial Intelligence*, available here.

VAN DAM C. (2013), *European Tort Law*, 2nd ed., Oxford Univ. Press, 2013.

VAN GESTEL R., VAN DICK G. (2011), *Better Regulation through Experimental Legislation*, in "European Public Law", vol. 17, n. 3.

VAQUER A. (2008), *Damage, in Tort Law of the European Community. Tort and Insurance Law*, Koziol H. – Schulze R. (eds.), vol. 23, Springer, 2008, p. 30 ff.

VEALE M., BORGESIUS F. Z. (2021), *Demystifying the Draft EU Artificial Intelligence Act*, in Computer Law Review International, 22(4), p. 97–112, available here.

VENICE COMMISSION (2011), *Report on the rule of law (CDL-AD (2011)003rev-e*, Council of Europe, available here.

VINEY G. (1982), *La responsabilité: conditions*, in J. Ghestin (ed.), Traité de droit civil, Paris, 1982.

VOLZ S. (2022), *KI Sandboxen für die Schweiz*, Zürich, Schulthess Juristische Medien AG, 2022, p. 51 ff.

VOLZ S., VON THIESSEN R. (2023a), *Smart parking – Best practices for image recognition*, available here.

VOLZ S., VON THIESSEN R. (2023b), *Autonomous systems – Guidelines for regulatory questions*, available here.

VOLZ S., VON THIESSEN R. (2023c), *Artificial Intelligence in education – Legal best practices*, available here.

VOLZ S., VON THIESSEN R. (2024), *Machine translation – Recommendations for public administration*, available here.

VRABEK H.U., CUSTERS B. (2016), *Legal Barriers and Enablers to Big Data Reuse - A Critical Assessment of the Challenges for the EU Law*, in European Data Protection Law Review, 2/2016, available here.

WEIMER M., MARIN L. (2016), *The Role of Law in Managing the Tension between Risk and Innovation: Introduction to the Special Issue on Regulating New and Emerging Technologies*, in European Journal of Risk Regulation, n. 3.

WERNER C., BRINKER N., RAABE O. (2022), *Grundlagen für ein gesetzliches IT-Sicherheitsrisikomanagement — Ansätze zur Vereinheitlichung von Rollenmodell*, Risikomanagement und Definitionen für das IT-Sicherheitsrecht, CR 2022, p. 817 ff.

WIENER N. (1948), *Cybernetics or Control and Communication in the Animal and the Machine*, Cambridge, MA: MIT Press.

WIEWIOROWSKI W. (2024), *Devising a trajectory towards a just and fair future: the identity of data protection in times of AI*, Speech given at the CPDP Conference in Brussels on 24 May 2024, available here.

YEFREMOV A. (2019), *Regulatory Sandboxes and Experimental Legislation as the Main Instruments of Regulation in the Digital Transformation*, in D. Alexandrov et al (eds.), Digital Transformation and Global Society, 4th International Conference, DTGS 2019, p. 82-91.

YORDANOVA K. (2019), *The Shifting Sands of Regulatory Sandboxes for AI*, KU Leuven CiTiP Blog, 18 July, 2019, 217, available here.

ZETZSCHE D. A., BUCKLEY R. P., BARBERIS J.N., ARNER D.W. (2017), *Regulating a revolution: From regulatory sandboxes to smart regulation*, in "Fordham J. Corp. & Fin. L.", vol. 23, n. 1.

ZWEIGERT K., KOTZ H. (1998), *Introduction to Comparative Law*, Oxford Univ. Press, 1998.

AUTHORS

Alessandro Armando received his PhD in Computer Engineering at the University of Genoa. His appointments include positions as research fellow at the University of Edinburgh and at INRIA-Lorraine (France). He is professor at the University of Genoa where he teaches Computer Security. He has contributed to founding and directing the Master in Cybersecurity and Data Protection of the University of Genoa and the Master in Digital Forensic and Cyber Technologies at the School of Telecommunications of the Armed Forces (STELMILIT). He has served as director of the PhD Program in Security, Risk and Vulnerability of the University of Genoa. He founded and led the Security & Trust Research Unit of the Bruno Kessler Foundation in Trento. He has been coordinator and/or team leader in several national and EU research projects, including a European Industrial Doctorate in partnership with SAP. He contributed to the discovery of authentication flaws in Single Sign-On standards and implementations, including a serious man-in-the-middle attack on the SAML-based SSO for Google Apps. He is currently serving as director of the Cybersecurity National Laboratory of the National Inter-university Consortium on Informatics (CINI) and as chairman of the Scientific Committee of the SERICS Foundation.

▪

Filippo Bagni* is a lawyer and researcher specialising in the regulation of the digital ecosystem, particularly in the field of artificial intelligence and cybersecurity. He is currently pursuing a PhD as part of the Italian National PhD Programme in Cybersecurity at the IMT School for Advanced Studies in Lucca. Filippo serves as a Legal Officer at the European Commission's Directorate-General for Communications Networks, Content and Technology (DG Connect) in Brussels. Within the Digital Services Act Enforcement team, he focuses on critical risks posed by online platforms, with a particular focus on generative AI, recommender systems, the protection of minors online, and data access for researchers. Filippo's career bridges law, economics, and technology, reflecting a multidisciplinary expertise that shapes both his professional and academic contributions.

---

* The opinions and views expressed are solely those of the author and do not reflect or represent the official policy or position of the European Commission.

He is also the author of numerous publications exploring the legal challenges posed by emerging technologies.

▪

Davide Baldini is an attorney and researcher, pursuing a double PhD degree at the Florence and Maastricht Universities, where he focuses his research on algorithmic discrimination. He is a Partner at the international law firm ICT Legal Consulting, advising multinational corporations and innovative start-ups on personal data protection, telecommunications, media, technology (TMT), and AI law. He is a certified Data Protection Officer (ECPC-B), Lead Auditor and Consultant for the CSA Code of Conduct for GDPR Compliance, and Implementer and Auditor for the Europrivacy™ Certification Scheme. Davide also participates in national and international research projects and he previously held a research fellowship at the University of Florence, where he focused on profiling and automated decision-making under the GDPR.

▪

Nils Brinker is a Cybersecurity Expert specializing in regulatory aspects of cybersecurity and data protection. He holds a position as senior cybersecurity consultant at intcube where he advises small and middle-sized companies on regulatory questions in the application of new technology and cybersecurity in general. Before his current position, he held different position within industry (insurance) and academia (e.g., Digital Society Institute of the ESMT Berlin) working on practical and scientific solutions for the development and application of technology law.

Eleonora Bonel* is a digital policy expert, specialising in governance and implementation challenges related to digital government transformation. She is currently working as an external services provider to the European Commission, in the Directorate-General for Digital Services (DG DIGIT). Her professional experience ranges from digital government research in different academic institutions, to cybersecurity risk consulting, to working in civil society in the field of privacy and digital rights. She holds a Master's degree in International Public Management from Erasmus Rotterdam and a Master's

* The opinions and views expressed are solely those of the author and do not reflect or represent the official policy or position of the European Commission.

degree in Public Policy and Digital, New Technologies from Sciences Po (Paris). She has authored publications on digital government and the intersections of digitalization challenges with society, and the use of Artificial Intelligence in the public sector.

▪

Kate Francis is a PhD candidate at the Maastricht University Faculty of Law and the Head of Research, Data Ethics and Digital Sustainability at ICT Legal Consulting (ICTLC), an international law firm. Kate's research topics include transparency in the EU legal framework, Data Protection as a Corporate Social Responsibility, and GDPR enforcement. She is the co-author of "Data Protection as a Corporate Social Responsibility" (Edward Elgar, 2023), which introduces the first auditable framework for integrating data protection within the CSR and ESG domains. Kate is also a certified Data Protection Officer (ECPC-B DPO, European Centre of Privacy and Cybersecurity, Maastricht University), a certified Cloud Security Alliance GDPR Compliance Code of Conduct Consultant, and a certified Europrivacy - GDPR compliance, audit and certification Auditor.

▪

Matteo Giannelli is Assistant Professor of Constitutional Law at the University of Florence (position activated within the Extended Partnership "Security and Rights in CyberSpace" (SERICS) funded within line 4.2. of the NRRP). He graduated in Law (2016) and later obtained his Ph.D. at the University of Florence (2021). In 2023 he obtained the qualification of Associate Professor of Constitutional Law. His current research interests include the patterns of foreign power in the Italian legal system, the system of sources of law, and constitutional issues related to the use of new technologies, particularly in the field of cybersecurity.

▪

Erik Longo is Full Professor of Constitutional Law at the University of Florence (Italy) where he teaches Public Law, Rights and Rules for Artificial Intelligence and Data Protection Law. Previously he worked as a Constitutional Law Professor at the University of Macerata (Italy). Professor Longo's scholarly interests are deeply rooted in the legal challenges presented by digital technologies, alongside a profound commitment to analysing the foundational elements of Italian and European constitutional law. He has been visiting at the Center for Civil and Human Rights of the Notre Dame University-

Program for Law in 2012, at the University of Sussex-School of Global Studies in 2014, and at the Queen's University (Belfast) in 2018.

▪

Giuseppe Mobilio is Associate Professor of Constitutional Law at the University of Florence. He graduated in Law at the University of Florence (2009) and later obtained his Ph.D. at the University of Pisa (2014). In 2022 he obtained the qualification of Full Professor of Constitutional Law. His current research interests include biometric surveillance, the regulation of technology and the protection of fundamental rights. His areas of research also include the dynamics of forms of government, sources of law, territorial and metropolitan government and constitutional reform.

▪

Enrico Panai is an AI ethicist with a background in philosophy and extensive consulting experience in Italy. He spent seven years as an adjunct professor of Digital Humanities at the University of Sassari. He is professor of Responsible AI and AI Ethics at EMlyon Business School, ISEP in Paris, and La Cattolica in Milan. Additionally, he is the president of the Association of AI Ethicists. Currently, his main role is as an officer of the French Standardization Committee for AI and convenor of the working group on fundamental and societal aspects of AI at the European CEN-CENELEC JTC21—the European standardization body focused on producing deliverables that address European market and societal needs.

▪

Giovanni Maria Riccio is full professor of Comparative Law (University of Salerno, Italy) where he teaches Art Law and Cultural Heritage Law. He has worked in several European Universities and he has been a consultant of the European Commission. Professor Riccio has written and edited 6 books and more than 130 articles in Italian and international reviews. His research area is mostly focused on data protection, IT law, artificial intelligence law, copyright, and civil liability.

▪

Fabio Seferi is a PhD candidate in the Italian National PhD Program in Cybersecurity at the IMT School for Advanced Studies Lucca and the University of Florence, Italy. His research is focused on cyber resilience regulatory sandboxes, both from a legal framework

and a practical implementation perspective. Previously, he has worked as a consultant in the private sector, mainly in cybersecurity governance, risk and compliance. He holds several postgraduate diplomas and industry certifications in areas such as intelligence, information security, and artificial intelligence.

▪

Andrea Simoncini is Full Professor of Constitutional Law at the Department of Legal Sciences of the University of Florence. He is PI of CybeRights Project within the SERICS Foundation. He has been Director of the DSG and teaches general constitutional law and constitutional law of technology; he teaches Rights and Rules for Artificial Intelligence at the School of Engeneering. He was a "Fulbright Distinguished Fellow" at the Notre Dame Law School and a visiting professor at the Nanovic Institute for European Studies. He is currently a Permanent Research Fellow at the Center for Ethics and Culture and Distinguished Research Affiliate at the Law School at the University of Notre Dame (USA). He clerked for Professor Ugo De Siervo at the Constitutional Court between 2002 and 2003. His main research interests range from issues related to Italian and European constitutional law to the sources of law, the relationship between technology and constitutional rights, environmental law, social rights and the relationship between natural law and positive law.

▪

Alessio Tartaro is a PhD candidate at the University of Sassari, Italy. Trained as a philosopher at the Scuola Normale Superiore in Pisa, his research focuses on the role and limitations of technical standards as regulatory tools for artificial intelligence. Alessio's research interest in standardization is complemented by a practical involvement in the field, reflected in his active participation in ISO/IEC, ETSI, and CEN-CENELEC technical committees on AI. He gained international research experience at the Budapest University of Technology and Economics, the European Commission's Joint Research Centre, and the Tilburg Institute for Law, Technology, and Society.

▪

Raphael von Thiessen leads strategic initiatives in technology and innovation. He is the Programme Manager of the Innovation Sandbox for Artificial Intelligence at the Canton of Zurich. As an author with a focus on strategic foresight, Raphael has published several

books, including Decoding Artificial Intelligence (NZZ Libro, 2020). In addition to his role at the Innovation Sandbox, he is the Co-Founder of Momentum Collaboration, a strategy and innovation consulting firm.

▪

Antonella Zarra* is an expert in AI policy and platform regulation. She is currently serving as a Case Handler Officer at the European Commission. Within the Digital Services Act Enforcement team, she focuses on the risks posed by online platforms' algorithms (including automated content moderation, generative AI and recommender systems). She is pursuing a PhD in Law and Economics at the University of Hamburg, exploring the role of experimental governance in AI regulation. Her research extends to the impact of technologies on vulnerable groups, gender equality and women's representation in policy-making.

* The opinions and views expressed are solely those of the author and do not reflect or represent the official policy or position of the European Commission.

◆ ACKNOWLEDGEMENTS ◆

This White Paper would not have been possible without the outstanding contributions of all the authors involved in the project. We would like to express our sincere thanks to each of them for the time, effort and expertise they shared with us. If we are proud of our final product, it is largely because of them.

We would also like to thank our academic supervisors who have guided and supported us along the way. Their guidance, support and expertise were instrumental in the completion of this project.

Finally, we would like to thank our universities and research centres, as well as CINI and SERICS for providing us with resources and infrastructure.

A special thanks goes to all those who have contributed to making this project possible. Their dedication and commitment have been essential to the completion of this White Paper.

———

Filippo and Fabio