Responsible AI UK Response to the DSIT's Call for Evidence on the AI Growth Lab

Date: 07 January 2026

We are submitting this response to the Department for Science, Innovation and Technology (DSIT) open call for evidence regarding the AI Growth Lab[1] on behalf of Responsible AI UK (RAi UK), an open and multidisciplinary network that brings together experts from across the four nations of the UK to understand how we should shape the development of AI to benefit people, communities and society. To arrive at this response, we sent out a call to Principal Investigators and Co-Investigators of all RAi UK-funded projects[2], inviting them to contribute to this consultation. What follows is a synthesis of the responses we received from our research community.

Executive Summary

Overall, the RAi UK research community welcomes the government's efforts to create spaces that balance innovation with appropriate guardrails, ensuring responsible AI development and deployment. When properly designed and implemented, regulatory sandboxes can be a helpful mechanism for addressing regulatory barriers and enhancing regulators' learning. The comments herein aim to inform the government's endeavours to establish a cross-economy sandbox, providing recommendations regarding its design so that the whole of the UK society can enjoy the benefits of this legal experimentation, while being attentive to and mitigating potential risks.

Potential Benefits

- Cross-economy sandboxes can facilitate coordination among regulators, avoiding fragmentation and regulatory forum shopping.

- Multi-regulator approaches enable holistic assessment of AI products and services.

- Evidence shows sandboxes can catalyse investment and innovation when properly implemented.

Critical Concerns

- Poorly designed sandboxes risk exploitation, limited public benefits, and regulatory costs exceeding value.

- Multi-regulator models may cause implementation delays and institutional overlap.

- The proposal lacks meaningful involvement of end users and impacted communities, particularly given that safeguards may be relaxed during experiments.

---

[1] https://www.gov.uk/government/calls-for-evidence/ai-growth-lab> Accessed on 16/12/2025
[2] https://rai.ac.uk/all-projects/ > Accessed on 16/12/2025

Priority Recommendations

1. **Public Participation:** Implement participatory governance models, co-design methods, and people's panels to centre the voices of those most impacted by AI systems.

2. **Protected Rights:** Never modify or disapply legislation protecting human rights, equality, worker protections, consumer safety, data protection principles, environmental standards, and intellectual property rights. Where irreversible harm is possible, safety standards must remain intact.

3. **Regulatory Clarity Over Deregulation:** Focus on supporting understanding of existing regulations rather than suspending safeguards. Launch innovation hubs and hotlines alongside the sandbox to avoid "risk washing".

4. **Sector Prioritisation:** Focus on areas where the UK has competitive strengths (life sciences, financial services, health) and where AI capabilities are mature enough for time-constrained pilots. Such a focus can be coupled with efforts to address the AI sector beyond sector-specific regulated areas, so that core AI development is not left unattended or without safeguards.

5. **Robust Oversight:** Establish statutory oversight committees with sectoral regulators and independent experts, coupled with public transparency, participatory auditing, and cross-national evidence bases

6. **Academic Inclusion:** Include university researchers and academics as eligible participants to leverage cutting-edge research insights. Explicitly mention university-borne/researcher-led startups in the range of organisations invited to participate to attract cutting-edge ideas.

7. **Democratic Safeguards:** Successful pilots should not justify streamlined legislative processes that bypass democratic checks and balances, particularly without public input in determining success.

8. **Assurance cases:** the sandboxes should be geared to produce clearly defined, methodologically sound assurance cases for key AI technologies in specific applications to calibrate the confidence that practitioners, regulators, and end-users have in such technologies.

Our Answers to DSIT's Specified Questions

[Question 6] To what extent would an AI Growth Lab make it easier to develop or adopt AI?

The evidence on how an AI Growth Lab would impact the development and adoption of AI is mixed. RAi UK researchers held different views on this, with opinions ranging from not having an effect to making development and adoption somewhat easier, and published works show varied findings.

This lack of agreement may reflect two aspects. First, it can reflect scepticism or confusion about the term "Growth Lab" being used to refer to sandboxing, when the alternative "regulatory sandbox" is better known and already associated with the UK's leading role in its implementation. Second, it can reflect concerns regarding the trade-offs of regulatory sandboxes, which show promise in innovation contexts but also present risks and may not

always be successfully implemented – as debated in the literature about the topic for which we present a short summary below.

Several papers point to sandboxes as an effective mechanism for driving innovation while promoting regulatory learning, and balancing policy goals, including consumer protection, economic stability, and growth.[3] In the FinTech sector, research shows that sandboxes can be powerful catalysts for investments.[4] The OECD (2025) corroborates this view by placing regulatory sandboxes as:[5]

> a conducive environment for innovation by allowing businesses to test new products, services, and business models in a controlled and supervised setting. This approach fosters experimentation without the immediate pressure of full regulatory compliance and immediate generalised consequences, thereby encouraging creativity and technological advancement.

Regulatory sandboxes can help to facilitate the development and deployment of innovative technologies and business models, but their success is not guaranteed, depending on factors such as regulatory flexibility, resource availability, and alignment with national energy priorities.[6]-[7]

Ranchordas (2021) points out that sandboxes, as a type of experimental legal regime, are criticised for their loose methodology, casuistic nature, and limited validity of their results.[8] The author adds that studies have unveiled these experiments as flawed, given their frequent politicisation, their premature termination, dissatisfactory evaluations, and the general absence of methodological preoccupations.

Kwok and Taeihagh (2025) also warn that poorly designed sandboxes, with lenient entry requirements and inadequate oversight, risk exploitation by private firms, leading to low-quality experiments, limited public benefits, and regulatory costs that exceed the value the sandbox provides.[9]

[Question 7] What advantages do you see in establishing a cross-economy AI Growth Lab, particularly in comparison with single-regulator sandboxes?

Considering that the AI Growth Lab intends to work as a regulatory sandbox – i.e., a temporary regulatory waiver or flexibility, allowing new products, services, or business models to be

---

[3] Gabriel Kwok Hui Chen & Araz Taeihagh (11 Oct 2025): Designing regulatory sandboxes: a comprehensive framework for aligning functionalities and objectives, Policy Design and Practice, DOI: 10.1080/25741292.2025.2570954; Gromova E., Ivanc T. Regulatory Sandboxes (Experimental Legal Regimes) for Digital Innovations in BRICS. *BRICS Law Journal*. 2020;7(2):10-36. https://doi.org/10.21684/2412-2343-2020-7-2-10-36; Lilian Gumbo & Uche A. K. Chude-Okonkwo (2025) Regulatory sandbox as a frontier for innovation and sustainability: a systematic review, Cogent Business & Management, 12:1, 2510555, DOI: 10.1080/23311975.2025.2510555.

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

[5] OECD, Regulatory Sandbox Toolkit: A Comprehensive Guide for Regulators to Establish and Manage Regulatory Sandboxes Effectively (Technical Paper), July 2025, https://digital.gob.es/content/dam/portal-mtdfp/funcion-publica/gobernanza-publica/simplificacion/doc-referencia/RegulatorySandboxToolkit_OECD-en.pdf

[6] Inter-American Development Bank, Regulatory Sandboxes and Innovation Testbeds: A Look at International Experiences and Lessons for Latin America and the Caribbean (Final Report), 2020, https://www.technopolis-group.com/wp-content/uploads/2020/09/Regulatory-Sandboxes-and-Innovation-Testbeds-A-Look-at-International-Experience-in-Latin-America-and-the-Caribbean.pdf

[7] Zühre Aydın, Okan Yardımcı, Regulatory sandboxes and pilot projects: Trials, regulations, and insights in energy transition, Engineering Science and Technology, an International Journal, Volume 56,2024, 101792, https://doi.org/10.1016/j.jestch.2024.101792.

[8] Ranchordás, S. (2021). Experimental Regulations and Regulatory Sandboxes: Law without Order? Law and Method, 2021. https://doi.org/10.5553/REM/.000064.

[9] Gabriel Kwok Hui Chen & Araz Taeihagh (11 Oct 2025): Designing regulatory sandboxes: a comprehensive framework for aligning functionalities and objectives, Policy Design and Practice, DOI: 10.1080/25741292.2025.2570954.

tested in real market conditions with fewer regulatory constraints, one potential benefit of a cross-economy experiment is to facilitate coordination among diverse regulators, avoiding fragmentation in standards and requirements setting. This encourages the simultaneous exploration of multiple stakeholders' points of view, providing an opportunity for (i) a holistic check on the product, service, or model at hand, and (ii) allowing designers and developers to understand the different regulatory trade-offs and improve their products accordingly.

When additional sectoral requirements for specific applications (e.g., medical devices) must be considered (presumably, regulated services would remain regulated as a matter of good governance), they may be added on top of the established common baseline, thereby eliminating any overlap with the baseline standards.

A cross-economy AI Growth Lab could guarantee that responsible AI development is approached in the same way across sectors and different types of regulators, by ensuring cohesiveness. Such an approach could avoid not only overlapping regulatory burdens – when agencies enact different compliance requirements to achieve the same policy goal or ethical behaviour, but also "regulatory forum shopping" – when market participants search for the regulatory sandbox that is most favourable to their goals with the least stringent regulation. It also means that later updates on standards and regulations (in the UK or internationally) can be applied consistently and in a timely manner. Here, engaging with the Digital Regulators Cooperation Forum is encouraged since this is an already established cooperation venue in the UK's digital services industry.

A cross-economy sandbox can facilitate the exit process when the product, service, or business at hand requires licensing (or a similar form of prior authorisation) from multiple regulators. As Kwok and Taeihagh (2025) highlight, regulators responsible for the transition plan and market entry at the end of the sandbox duration should align with those involved in the initial entry and authorisation process.[10]

Finally, a multi-regulator approach, depending on its design and implementation, may lead to institutional multiplicity, in which multiple institutions are assigned to perform the same function. For example, different regulators oversee the sandboxing experiment simultaneously. Even though they share the exact baseline requirements, they operate independently within their respective remits. This situation is particularly relevant if the Growth Lab is operated in a more decentralised manner, i.e., with less central government control. Studies suggest that when agencies compete with each other, this competition pushes them to improve and make more effective institutional changes to stay ahead.[11]

Such parallel institutional structures can also enhance the benefits of specialisation, enabling an institution to perform its specific role more effectively. Additionally, if one institution fails to fulfil its responsibilities, another can step in to fill the gap. This arrangement may help reduce the risk of failures at each stage of the accountability process.

[10] Gabriel Kwok Hui Chen & Araz Taeihagh (11 Oct 2025): Designing regulatory sandboxes: a comprehensive framework for aligning functionalities and objectives, Policy Design and Practice, DOI: 10.1080/25741292.2025.2570954.
[11] Carson, Lindsey, and Mariana Mota Prado. *Brazilian anti-corruption legislation and its enforcement: Potential lessons for institutional design*. IRIBA Working Paper: 09. Manchester. UK, 2014.

[Question 8] What disadvantages do you see in establishing a cross-economy AI Growth Lab, particularly in comparison with single-regulator sandboxes?

Multi-regulator sandboxes can lead to implementation delays compared with single-regulator sandboxes. This is primarily due to the need for greater alignment among the involved regulators during the pre-launch phase. Regulators must agree on various aspects, including the design, execution plan, eligibility and testing criteria, and monitoring protocols from the outset. This process can be particularly challenging when regulators have differing goals and compete to promote their own agendas.[12]

As noted in the response to question 7, a cross-economy AI Growth Lab can create institutional multiplicity, which also comes with certain drawbacks.

First, this multiplicity can lead to institutional overlap, where multiple agencies perform the same task (e.g., monitoring a live testing experiment based on shared requirements). This situation may be viewed as inefficient resource use. Secondly, competition among multiple institutions performing similar functions can be counterproductive and create unnecessary tensions among them.[13]

Regardless of whether it is a single or multi-regulator experiment, sandboxes may be disadvantageous for establishing public trust in the innovation under test and/or the government's role in regulating it when the involvement of decision-subjects/end users is not thoroughly considered from the outset.

This means that the targeted consumers, those who use the technology but are not involved in its design, should be given a role in the risk assessment process throughout the lifecycle, especially during early parts involving purpose identification and data collection prior to deployment.

Research from RAI UK's projects (i) Public Voices in AI,[14] and (ii) Participatory Harm Auditing Workbenches and Methodologies[15] highlights that AI risk management should be balanced with respect to other stakeholders (not only AI developers) and, crucially, attends to historically excluded or marginalised voices in AI development/use.[16]

In the current AI Growth Lab proposal, the voice of decision-subjects/impacted communities seems to be missing, as there is no reference to their role in the sandbox design and execution.

The RAi UK research community encourages the government to engage with those who will be directly affected by the AI systems under testing, especially given that safeguards may be relaxed during experiments, which could increase users' vulnerability. If something goes wrong with the experiment and users are harmed, it may undermine public trust. That is why it is crucial to plan for participatory auditing mechanisms and to involve impacted communities in assessing the risks and trade-offs of AI applications from the outset.

---

[12] Gabriel Kwok Hui Chen & Araz Taeihagh (11 Oct 2025): Designing regulatory sandboxes: a comprehensive framework for aligning functionalities and objectives, Policy Design and Practice, DOI: 10.1080/25741292.2025.2570954.

[13] Carson, Lindsey, and Mariana Mota Prado. *Brazilian anti-corruption legislation and its enforcement: Potential lessons for institutional design*. IRIBA Working Paper: 09. Manchester. UK, 2014.

[14] https://digitalgood.net/dg-research/public-voices-in-ai/ > Accessed on 07/01/226

[15] https://phawm.org > Accessed on 21/12/2025

[16] Quyoum, Aunam and Wong, Mark (2024) Valuing lived experience and co-design solutions to counter racial inequality in data and algorithmic systems in UK's digital services. *Information Communication and Society*, 27(9). https://www.tandfonline.com/doi/full/10.1080/1369118X.2024.2331781; Beatrice Vincenzi, Simone Stumpf, Alex S. Taylor, and Yuri Nakao (2024) Lay User Involvement in Developing Human-centric Responsible AI Systems: When and How? *ACM J. Responsib. Comput.* 1, 2, Article 14 (June 2024). https://doi.org/10.1145/3652592

[Question 9] What, if any, specific regulatory barriers (particularly provisions of law) are there that should be addressed through the AI Growth Lab? If there are, why are these barriers to innovation?

As the OECD points out,[17] there are manly three types of regulatory barriers: (i) costly compliance, (ii) uncertainty created by regulation, and (iii) innovation prohibited by regulation. In terms of uncertainty created by regulation, one example can be found in the responses to the Bank of England and the FCA's 2023 Discussion Paper, which emphasized the need for greater regulatory clarity, including defining bias and fairness under the Equality Act 2010 and the FCA's Consumer Duty.[18] The 2024 Bank of England and FCA survey reaffirmed these concerns.[19]

Duff and Jeník (2020) note that another regulatory barrier is a limited understanding of rules and regulations.[20] In many cases, regulatory barriers are not pieces of legislation in themselves, but rather the perception of their restrictive nature. This may result from companies and organisations misunderstanding the interpretation and application of the legal provisions. For example, under the UK GDPR, truly anonymised data is not protected; nonetheless, organisations often apply the same level of protection as to identifiable data and deny public access to it under a heightened-care approach that is not required by law.

Thus, the government's focus should be on supporting understanding of the ethical and regulatory aspects of innovation and on fostering an overall innovation environment that is responsible, ethical, and trustworthy, rather than suspending regulatory safeguards. To that end, it is recommended that the launch of the regulatory sandbox be preceded by and coupled with other innovation facilitators – e.g., innovation hubs, hotlines, or similar initiatives.[21] Such a measure can avoid risk washing, whereby sandboxes are used as a stamp of marketability to reduce perceived risk around innovations, which can leave consumers in a vulnerable position.[22]

This is especially relevant to sustaining the UK's regulatory credibility amid capture threats from influential players in the AI market (already discussed in the EU context),[23] and to avoid

---

[17] OECD, Regulatory Sandbox Toolkit: A Comprehensive Guide for Regulators to Establish and Manage Regulatory Sandboxes Effectively (Technical Paper), July 2025, https://digital.gob.es/content/dam/portal-mtdfp/funcion-publica/gobernanza-publica/simplificacion/doc-referencia/RegulatorySandboxToolkit_OECD-en.pdf

[18] Bank of England. 2023. FS2/23 – Artificial Intelligence and Machine Learning. https://www.bankofengland.co.uk/prudential-regulation/publication/2023/october/artificial-intelligence-and-machin e-learning (Accessed: 27/05/2025); and Bank of England. 2022. DP5/22 - Artificial Intelligence and Machine Learning. https://www.bankofengland.co.uk/prudential-regulation/publication/2022/october/artificial-intelligence (Accessed: 27/05/2025)

[19] Bank of England. 2024. Artificial intelligence in UK financial services - 2024. https://www.bankofengland.co.uk/report/2024/artificial-intelligence-in-uk-financial-services-2024 (Accessed: 27/05/2025)

[20] Duff, Schan; Jeník, Ivo. *How to Build a Regulatory Sandbox : A Practical Guide for Policy Makers.* Washington, D.C.: World Bank Group, 2020. http://documents.worldbank.org/curated/en/126281625136122935

[21] Duff, Schan; Jeník, Ivo. *How to Build a Regulatory Sandbox : A Practical Guide for Policy Makers.* Washington, D.C.: World Bank Group, 2020. http://documents.worldbank.org/curated/en/126281625136122935; Gumbo, L., & Chude-Okonkwo, U. A. K. (2025). Regulatory sandbox as a frontier for innovation and sustainability: a systematic review. *Cogent Business & Management*, *12*(1). https://doi.org/10.1080/23311975.2025.2510555

[22] Brown, E., and D. Piroska. 2022. "Governing Fintech and Fintech as Governance: The Regulatory Sandbox, Riskwashing, and Disruptive Social Classification." *New Political Economy* 27 (1): 19–32. https://doi.org/10.1080/13563467.2021.1910645

[23] Bram Vranken, "Big Tech Lobbying Is Derailing the AI Act," Corporate Europe Observatory, November 24, 2023, https://corporateeurope.org/en/2023/11/big-tech-lobbying-derailing-ai-act; Cynthia Kroet, "Industry Flags 'Serious Concerns' With Latest Draft of EU AI Code of Practice," Euronews, March 12, 2025, https://www.euronews.com/next/2025/03/12/industry-flags-serious-concerns-with-latest-draft-of-eu-ai-code-of-practice; and Alexandre Piquard, "France Keeps Up Pressure on EU's AI Act, Despite Mounting Criticism," *Le Monde*, January 27, 2024, https://www.lemonde.fr/en/economy/article/2024/01/27/france-keeps-up-its-pressure-on-the-eu-s-ai-act-despite-mounting-criticism_6471038_19.html; Martin Greenacre, "EU Is 'Losing the Narrative Battle' Over AI Act, Says UN Adviser," Science Business, December 5, 2024, https://sciencebusiness.net/news/ai/eu-losing-narrative-battle-over-ai-act-says-un-adviser.

the transformation of sandboxes into disguised deregulation, which poses significant risks to democratic oversight, strategic autonomy, and technological sovereignty.[24]

To that end, there are significant concerns among Minoritised Ethnic people in England and Scotland about the privacy and security of adopting AI and digital systems in essential services, including health, housing, and energy. Concerns are centred around reinforcement of racism by AI and digital essential services, particularly due to the lack of enforcement of regulations and protection from racial discrimination. In addition, the fear of repercussion creating high-stake dilemmas and potential withdrawal from use of digital services when racism and privacy concerns are not safeguarded in AI and digital innovation.[25]

[Question 10] Which sectors or AI applications should the AI Growth Lab prioritise?

The AI Growth Lab can prioritise sectors where the UK has existing competitive strengths, such as life sciences, financial services, professional services, health and care, social sciences, and creative industries. Prioritising existing competitive strengths can maximise the likelihood of translating pilots into sustained UK-based economic activity rather than innovations that migrate elsewhere for commercialisation. Another approach that can be considered is prioritising high-reward areas and/or regulatory innovation, despite their heightened risk.

In addition, the Lab could prioritise sectors with significant potential for public benefit (e.g., health and care) and economic growth, where there are clear regulatory barriers that specifically obstruct AI development and/or adoption, and where AI capabilities are mature enough to deliver measurable outcomes in time-constrained pilots. For example, our work within the Responsible AI UK Health and Care Working group has surfaced a number of barriers to AI adoption in the health and care sector that include but are not limited to: (i) government procurement rules (ii) multiplicity of regulatory requirements (iv) lack of accepted standards to work with (iii) lack of access to capital (due to investor fear of regulation) to startups to progress from early stage ideas to fully tested solutions. Sandboxes could potentially alleviate some of these issues.

The proposed approach aligns with OECD recommendations,[26] which suggests that early-stage testing offers limited benefits if the application remains unclear or impacts are highly uncertain. The OECD states that, for early-stage innovations, sandboxing is beneficial in facing regulatory uncertainty, helping to develop technology and regulation in parallel so that regulations are ready when the innovation reaches the market. If products are market-ready, a sandbox with minimal regulatory uncertainty is preferable, allowing focus on specific technical aspects.

Nonetheless, focusing on already regulated sectors with defined frameworks may not be a sufficient basis to overlook AI development outside them, allowing core AI development to proceed without safeguards. It is encouraged that such an approach be coupled with

---

[24] Raluca Csernatoni (May, 2025), The EU's AI Power Play: Between Deregulation and Innovation, Carnegie Europe. https://carnegieendowment.org/research/2025/05/the-eus-ai-power-play-between-deregulation-and-innovation?lang=en

[25] Quyoum, A., Wong, M., Ghosh, S. and Shahandashti, S. (2025) Minoritised Ethnic People's Security and Privacy Concerns and Responses towards Essential Online Services. In: SOUPS '25 Proceedings of the Twenty-First USENIX Conference on Usable Privacy and Security, Seattle, United States, 10 -12 Aug 2025, pp. 259-278. https://dl.acm.org/doi/10.5555/3767870.3767885

[26] OECD, Regulatory Sandbox Toolkit: A Comprehensive Guide for Regulators to Establish and Manage Regulatory Sandboxes Effectively (Technical Paper), July 2025, https://digital.gob.es/content/dam/portal-mtdfp/funcion-publica/gobernanza-publica/simplificacion/doc-referencia/RegulatorySandboxToolkit_OECD-en.pdf

mechanisms that allow a holistic view to avoid risking gaps in offerings and ensure proper scrutiny is evenly applied.

Other criteria for prioritisation may include: (i) focusing on applications which face public scepticism when employed, especially unsupervised by humans; and (ii) testing applications that have the potential to be adopted in people's everyday lives, which can benefit from a clearer understanding of their risks to avoid exposing users to unpleasant situations (e.g., with them revealing data unintentionally).

[Question 11] What could be potential impacts of participating in the AI Growth Lab on your company/organisation?

Following the literature on regulatory sandboxes,[27] when looking into private entities and market participants who are developers or deployers of AI, the AI Growth Lab could allow them to test products that they wouldn't otherwise be able to test, make their companies more internationally competitive, and bring their products to market quickly than otherwise.

The proposal states that the Lab participants will be sought from "*start-up innovators, established FTSE companies and global AI developers alike, as well as innovators in the public sector*". RAi UK is a research and innovation programme that could benefit from this approach as it opens up opportunities for startups born out of the programme (through our accelerator and enterprise fellowships programme) to trial out their technology at an early stage. It would help to include explicitly university-borne/research-led startups in the construction of such sandboxes. Otherwise, the government risks missing the opportunity to engage with those who are at the forefront of innovation, bringing the latest research insights, methodologies, and deep subject expertise that offer a holistic understanding of responsible AI development.

[Question 14] What types of regulation (particularly legislative provisions), if any, should be eligible for temporary modification or disapplication within the Lab? Could you give specific examples and why these should be eligible?

Any temporary modification or disapplication of regulations in the form of guidance or code of practice must be carefully considered since it poses the risk not only of exposing users to harm, but also creating market distortions that hinder competition, for example, by creating an uneven playing field, acting as an imperfect substitute for other regulatory enablers, or acting as a de facto gatekeeper or substitute for interactions with traditional licensing or supervisory processes.[28] Such types of modifications should be used only to identify benefits (or gaps) in adherence to responsible and ethical AI principles.

---

[27] Lilian Gumbo & Uche A. K. Chude-Okonkwo (2025) Regulatory sandbox as a frontier for innovation and sustainability: a systematic review, Cogent Business & Management, 12:1, 2510555, DOI: 10.1080/23311975.2025.2510555; Gabriel Kwok Hui Chen & Araz Taeihagh (11 Oct 2025): Designing regulatory sandboxes: a comprehensive framework for aligning functionalities and objectives, Policy Design and Practice, DOI: 10.1080/25741292.2025.2570954; OECD, Regulatory Sandbox Toolkit: A Comprehensive Guide for Regulators to Establish and Manage Regulatory Sandboxes Effectively (Technical Paper), July 2025, https://digital.gob.es/content/dam/portal-mtdfp/funcion-publica/gobernanza-publica/simplificacion/doc-referencia/RegulatorySandboxToolkit_OECD-en.pdf; Duff, Schan; Jeník, Ivo. *How to Build a Regulatory Sandbox : A Practical Guide for Policy Makers.* Washington, D.C.: World Bank Group, 2020. http://documents.worldbank.org/curated/en/126281625136122935
[28] Duff, Schan; Jeník, Ivo. *How to Build a Regulatory Sandbox : A Practical Guide for Policy Makers.* Washington, D.C.: World Bank Group, 2020. http://documents.worldbank.org/curated/en/126281625136122935

Potential provisions that could be good candidates for experimentation within the Lab include those imposing prescriptive human-actor requirements, requiring activity to be undertaken by a person rather than AI. For example, when AI can match or exceed human performance, temporary modifications could enable supervised testing of autonomous operations.

Note that any modification or disapplication should be accompanied by alternative safeguards, clear outcome metrics, and the ability to terminate pilots immediately if concerns arise. According to the OECD,[29] sandbox terms can define conditions for pausing, changing, or ending trials to prevent or reduce harm. This may involve thresholds based on indicators like consumer complaints, environmental damage, or market distortions, which might not be known initially.

Furthermore, experimental legislation ought to establish mechanisms for consumer redress and compensation for any harm caused during experiments. Therefore, legal exemptions should guarantee that participating firms are liable and accountable for damages from such experimental activities.[30]

[Question 15] We propose that certain types of rules and obligations, such as those relating to human rights, consumer rights and redress mechanisms, and workers' protection and intellectual property rights, could never be modified or disapplied during a pilot. What types of regulation (particularly legislative provisions) should not be eligible for temporary modification or disapplication within the Lab (e.g. to maintain public trust)?

The RAi UK research community agrees that some types of legislation should not be eligible for modification or disapplication, including those listed in the question, and more specifically:

- The Equality Act 2010, the Worker Protection Act 2023, the Employment Relations Act 2004 and other associated protections against hate speech, harassment, and discrimination in the workplace and in society overall.
- The Copyright, Designs and Patents Act 1988.
- The Human Rights Act 1998, and all related legislation.
- The Freedom of Information Act 200.
- Consumer rights, including the 1987 Consumer Protection Act (CPA) and the 2005 General Product Safety Regulations (GPSR). The GPSR, however, may introduce regulatory uncertainty regarding its application to the AI value chain and the concept of "safe product", which could benefit from learning experiences of regulators within sandboxes. In such cases, attention should be given to ensure consumers are not left without safety mechanisms, and it would be appropriate to test modified requirements when harm **is not irreversible**, with the primary goal of providing safety rather than enabling innovation at all costs. **Where there is potential for irreversible harm, for example in robotics, safety standards should not be relaxed**.
- Financial crime prevention, including anti-money laundering and counter-terrorism financing obligations – e.g., Proceeds of Crime Act 2002, the Money Laundering Regulations 2017, the Terrorism Act 2000, the Sanctions and Anti-Money Laundering Act 2018. Such legislation should not be eligible for modification or disapplication to

---

[29] OECD, Regulatory Sandbox Toolkit: A Comprehensive Guide for Regulators to Establish and Manage Regulatory Sandboxes Effectively (Technical Paper), July 2025, https://digital.gob.es/content/dam/portal-mtdfp/funcion-publica/gobernanza-publica/simplificacion/doc-referencia/RegulatorySandboxToolkit_OECD-en.pdf
[30] Gabriel Kwok Hui Chen & Araz Taeihagh (11 Oct 2025): Designing regulatory sandboxes: a comprehensive framework for aligning functionalities and objectives, Policy Design and Practice, DOI: 10.1080/25741292.2025.2570954.

ensure that any financial system under the AI Growth Lab will not commit (or enable) any financial crimes, which can cause reputational, legal liability, and public trust damage.

- Environmental standards should not be eligible for modification since environmental harm is irreversible, and it affects the health and sustainability of current and future generations.
- Data protection principles under UK GDPR, including lawfulness, fairness, transparency, purpose limitation, and data subject rights, should not be eligible for modifications. **While specific procedural requirements might be streamlined, the protections that maintain individual autonomy and dignity must remain intact**. For example, private and sensitive personal information, such as medical records or biometric identifiers, if compromised, constitutes an irreversible harm.

[Question 16] What oversight do you think is needed for the Lab?

The RAi UK community involved in this consultation agrees that both: (i) a statutory oversight committee composed of sectoral regulators and independent experts, and (ii) public transparency and reporting are essential for effective sandboxing oversight. This approach provides some degree of independence from the political environment and ensures that oversight considers not only technical matters but also the importance of disclosure and information sharing within a democratic society.

Additionally, drawing from the project Participatory Harm Auditing Workbenches and Methodologies[31], the following measures could be considered, aiming at increasing public participation and, as a result, increasing public trust. The recommendations below recognise that public voices are central to the development, use, and evaluation of AI.[32]

- **Implementing participatory AI governance models, which centre the needs and priorities of people who are most negatively impacted by AI, using co-design methods and principles.** Co-design is a methodological tool that ensures public and marginalised groups, such as minoritised ethnic communities, have meaningful participation in the oversight and scrutiny of decisions on how AI is developed and tested.
- **Developing people's panels** at the UK and devolved levels, which govern and develop principles and public-led governance of the activities and standards upheld by the Lab. Informed by deliberative democracy and citizen assembly methods, people's panels support members of the public in providing expertise through lived experience and/or learning to inform oversight and scrutiny.
- **Creating a robust cross-national evidence base to record potential harms and risks tested (or allowed) via the AI Growth Lab environment.** It needs clear documentation of how AI is tested, for what purpose, and which areas require AI to be safer and more ethical.

---

[31] https://rai.ac.uk/new_projects/participatory-harm-auditing-workbenches-and-methodologies-phawm/ > Accessed on 16/12/2025
[32] Wong, M., Quyoum, A. and Mishra, A. (2024) Minoritised Ethnic People's Code of Practice for Equitable Digital Services. PRIME Protecting Minority Ethnic Communities Online. https://www.primecommunities.online/outputs/code-of-practice/

[Question 18] What criteria should determine which organisations or projects are eligible to participate in the Lab?

Following the OECD recommendations,[33] the RAi UK community understands that the eligibility criteria for determining participation in the AI Growth Lab should include: (i) innovations directly connected to AI, (ii) the presence of a well-defined regulatory barrier, which the Lab would help to overcome, and (iii) the existence of significant regulatory compliance resources, which the Lab could facilitate for testing.

[Questions 19 and 20] Which institutional model for operating the Lab is preferable? What are your reasons for selecting this model?

Each of the proposed models, whether run by the central government or the lead regulator, has pros and cons depending on the policy goals. Central government involvement may be necessary to coordinate regulators, provide a platform for institutional dialogue, and facilitate the exchange of experience. It can also help align goals and standardise practices, ensuring all sectoral regulators follow the same best practices and adopt the same regulatory language. However, the central government may lack sufficient knowledge of sector-specific regulations and market dynamics; therefore, regulators should be equally involved, especially when their expertise is mobilised.

Another option endorsed by some RAi UK researchers is to appoint an independent, non-governmental body to operate the Lab, with relevant expertise to advise the central government and regulators on their activities. Such a model would be welcomed to help avoid political manipulation.

An example of such an institutional setup is the European Blockchain Sandbox for Distributed Ledger Technologies. Managed by Bird & Bird, an international law firm chosen by the European Commission, the DLT sandbox requires applicants to submit pilot project proposals for review by a team of blockchain specialists. This team, appointed by Bird & Bird, recommends use cases for approval by the European Commission. Subsequently, Bird & Bird identifies the appropriate regulators for each use case, averaging about 1.5 regulators per case.[34]

[Question 21] What supervision, monitoring and controls should there be on companies taking part in the Lab?

Adequate supervision and monitoring should balance enabling genuine innovation with maintaining public trust, which is essential to the Lab's legitimacy and long-term success. Controls should be proportionate to risk, with higher-risk applications, such as AI in healthcare or public services, facing more intensive oversight.

Some key performance indicators that can be adopted to control the experiments include:

- **Fairness and bias metrics**, which are better addressed when technology and regulation develop in tandem, as trained AI models cannot be easily modified to comply

[33] OECD, Regulatory Sandbox Toolkit: A Comprehensive Guide for Regulators to Establish and Manage Regulatory Sandboxes Effectively (Technical Paper), July 2025, https://digital.gob.es/content/dam/portal-mtdfp/funcion-publica/gobernanza-publica/simplificacion/doc-referencia/RegulatorySandboxToolkit_OECD-en.pdf
[34] Gabriel Kwok Hui Chen & Araz Taeihagh (11 Oct 2025): Designing regulatory sandboxes: a comprehensive framework for aligning functionalities and objectives, Policy Design and Practice, DOI: 10.1080/25741292.2025.2570954

with new regulations. Post-facto checks are technically almost impossible, given the difficulty in locating where and how decision-making occurs within the model. AI's unintended consequences should be addressed through regulations at the process's beginning, not as an afterthought.

- **Incident detection and response rates**: how quickly and effectively failures, biases, or data drift issues are identified and mitigated within the controlled environment.
- **Explainability vis à vis accuracy**
- **User feedback scores**
- **Human oversight and override rate**: how often human reviewers intervene to correct or reverse automated decisions.
- **Data governance measures, documentation and record-keeping** for AI model design and development, lifecycle changes, regulatory adherence levels, and logging of AI incidents with full traceability.
- **Safeguarding mechanisms**, such as fine-tuning, red-teaming, and privacy-preserving methods, which play crucial roles in mitigating potential risks.

Depending on the technology's market readiness, requirements related to the dataset, model design, and training may be challenging to implement, especially if the effort targets ready-to-sell AI solutions that have passed the proof-of-concept phase.

If the sandboxing is designed more like an auditing process, it is crucial that end users be involved throughout, not only as passive subjects of innovation but as active participants in ensuring pilots serve the public interest.

Participatory auditing of AI enables end users and public/impacted communities to contribute to ongoing evaluation and maximise the application's success. This might include structured feedback collection using participatory AI auditing tools (e.g., developed by the PHAWM project) and accessible reporting tools. Independent user-led audits should be conducted to complement technical assessments, where end-users help identify systemic issues and 'blind spots' with new potential harms undetectable or unknowable to developers or regulators.

In this sense, the AI Growth Lab can encourage the use of co-designed tools and processes to inform the development and decision-making for ethical, trustworthy AI applications. One example of such tools can be found in the work of the project Participatory Harm Auditing Workbenches and Methodologies[35], which notes that these methods can serve as a 'gold standard' to supplement traditional AI auditing.

Finally, monitoring should be designed to account for its frequency and intensity, so that regulators can use the experiment as a learning experience to help address informational asymmetries between them and those controlling technology development. Kwok and Taeihagh (2025) suggest that monitoring can be dynamic, with its frequency and intensity changing over time.[36] During low-risk stages such as planning experiments, less frequent and intensive monitoring might be adequate. Conversely, during experimental phases, regulators may escalate both the frequency and intensity to better manage potential risks. When regulators lack familiarity with the technology, they tend to start with higher monitoring levels to gather more information, then decrease them as their understanding improves.

---

[35] https://rai.ac.uk/new_projects/participatory-harm-auditing-workbenches-and-methodologies-phawm/ > Accessed on 16/12/2025

[36] Gabriel Kwok Hui Chen & Araz Taeihagh (11 Oct 2025): Designing regulatory sandboxes: a comprehensive framework for aligning functionalities and objectives, Policy Design and Practice, DOI: 10.1080/25741292.2025.2570954

[Question 22] Do you think a successful pilot in the AI Growth Lab would justify streamlined powers for making changes permanent, as opposed to following existing legislative processes, which would take considerably longer?

The use of Lab results to streamline legislative processes for making changes permanent risks overriding democratic checks and balances, undermining the rule of law and the constitutional order – especially if the Lab is designed without accounting for public voices, and end users/affected communities' views in determining its success. A streamlined process would be more easily justified for specific sectoral regulations (stricto sensu) than for laws (Acts of Parliament), when there is enough evidence to support the change (which would preferably be gathered through a series of experiments rather than just one).

As Ranchordas (2021) points out, experimental legal regimes, such as sandboxes, often exhibit methodological deficiencies that hinder assessing their scientific and legal validity, which are essential for making any changes permanent through an expedited process. According to Ranchordas, employing an experiment in lawmaking can create an illusion of objectivity. This scientific appearance is likely to be manipulated by political reasoning and may undermine the original aims of experimental regulations, which are to promote evidence-based, iterative, and innovation-friendly regulatory responses.[37]

[Question 24] Would there be value in extending the AI Growth Lab to other high-potential technologies?

As Mittelsteadt (2025) argues,[38] permanent sandbox programs are not only tools for innovation, but also for crisis-ready regulatory flexibility. In emergencies such as pandemics or cyberattacks, governments often need to rapidly adjust regulations to allow testing and deployment of critical technologies. Extending the AI Growth Lab model to other high-potential or essential technologies could create pre-established, trusted mechanisms for temporarily relaxing specific rules, enabling faster responses when time is critical.

Having these sandboxes in place ahead of crises would prevent regulators and companies from having to design ad hoc regulatory workarounds under pressure. Instead, they could quickly identify which rules can be safely adjusted to meet urgent societal needs.

About Responsible AI UK

Responsible AI UK is a research and innovation programme focused on addressing the most pressing challenges for the UK and the global community. We draw from our multi-disciplinary research programme to deliver novel frameworks, tools, and policy advice for the development and deployment of safe and responsible AI so that it benefits everyone in society.

The organisations and research teams involved in this response to this consultation include:

RAi UK Leadership and Researchers

- Professor Sarvapali (Gopal) Ramchurn, FIET, CEO, Responsible AI UK, and Professor of Artificial Intelligence at the University of Southampton

---

[37] Ranchordás, S. (2021). Experiment(septal Regulations and Regulatory Sandboxes: Law without Order? Law and Method, 2021. https://doi.org/10.5553/REM/.000064
[38] Matt Mittelsteadt (September 2025), Digging into AI Sandboxes: Benefits, Risks, and the Senate SANDBOX Act Framework. CATO Institute, https://www.cato.org/blog/digging-ai-sandboxes-benefits-risks-senate-sandbox-act-framework

- Professor Gina Neff, Deputy CEO, Responsible AI UK; Executive Director of the Minderoo Centre for Technology and Democracy at the University of Cambridge, and Professor of Responsible AI at Queen Mary University of London.

- Dr Isabela Parisio, Research Associate, King's College London

- Dr Sarah Kiden, Research Fellow, University of Southampton

- Dr Athina Georgara, Research Fellow, University of Southampton

RAi UK Project Teams

[PHAWM: Participatory Harm Auditing Workbenches and Methodologies[39]]

- Professor Simone Stumpf, Professor of Responsible and Interactive Artificial Intelligence (School of Computing Science) at the University of Glasgow

- Dr Mark Wong, Senior Lecturer (Urban Studies & Social Policy) at the University of Glasgow

- Dr Siamak F. Shahandashti, Senior Lecturer (Department of Computer Science) at the University of York

- Dr Marios Aristodemou, Research Associate (Department of Computer Science) at the University of York

Please send any queries and comments about this response to info@rai.ac.uk.

---

[39] https://rai.ac.uk/new_projects/participatory-harm-auditing-workbenches-and-methodologies-phawm/ > Accessed on 16/12/2025