

Bridging Liability Gaps in the Age of AI: The Case for No-fault Compensation Schemes

Ha-Chi Tran

Independent Researcher *
chih.tran1@gmail.com

Abstract

Emerging technologies, including artificial intelligence (AI), are rapidly outpacing traditional legal frameworks, exposing regulatory gaps and weakening the effectiveness of conventional governance mechanisms. This study examines liability gaps stemming from algorithmic opacity, the distributed architecture of AI systems, and the systemic and diffuse nature of AI-related harms. These characteristics highlight the inadequacy of both fault-based and strict liability regimes in addressing AI-specific risks. The analysis is situated within the context of regulatory sandboxes, adaptive governance instruments designed to balance innovation and risk management by permitting firms to test new technologies under regulatory supervision with temporary legal exemptions. While such models effectively foster innovation, they leave unresolved questions of liability when harms arise from compliant experimentation. By placing primary civil and criminal responsibility on participating firms, sandbox frameworks fail to account for the distinctive nature of AI-related harms, resulting in compensation mechanisms that are often insufficient, delayed, and burdensome for affected parties. To address this gap, the study advances the proposal of a state-backed no-fault compensation scheme, modeled on the Vaccine Injury Compensation Programs (VICPs).

Introduction

The rapid advancement of artificial intelligence (AI) increasingly outpaces the capacity of traditional legal and institutional frameworks, creating regulatory lag and diminishing the effectiveness of conventional governance mechanisms (Dixon 2023; Ahern 2025). The unpredictability and opacity of complex AI systems, particularly sub-symbolic models with black-box, non-linear architectures, pose significant challenges to comprehensive risk assessment (Arrieta et al. 2020; Karnow 2016). Under such conditions, conventional ex-ante regulatory approaches often prove inadequate, either failing to provide sufficient protection (Zhang et al. 2025; Scherer 2016; Rangone and Megale 2025) or becoming

ing overly restrictive (Carvão et al. 2025), thereby stifling innovation (Abiri and Huang 2023; Arnal 2025).

In this context, the emergence of regulatory sandboxes has attracted global attention as a means of fostering innovation in emerging technologies such as AI (Cornelli et al. 2024; Qiu et al. 2025; Moraes 2025). By allowing selected entities to test products, services, or business models in controlled real-world environments, sandboxes temporarily relax certain compliance obligations while maintaining regulatory oversight (Kilian, Jäck, and Ebel 2025; Morgan 2023). Continuous supervision and real-time evaluation generate empirical insights that help regulators detect latent risks, emergent behaviors, systemic biases, and complex socio-technical dynamics that often elude theoretical models and ex-ante assessments (Zhang et al. 2025; Scherer 2016).

Despite these advantages, sandbox frameworks face inherent tensions between promoting innovation and safeguarding public and consumer interests. The temporary relaxation of legal obligations exposes third parties to heightened risks that would otherwise be mitigated under fully regulated conditions (Allen 2020; Knight and Mitchell 2020; Ranchordas 2021). Moreover, existing liability and compensation models remain inadequate, offering limited mechanisms to ensure that individuals harmed by activities conducted under experimental or exempted regimes receive timely and appropriate redress.

Research Question and Objectives

The central research question guiding this study is:

In the context of regulatory sandboxes, how and why do existing compensation mechanisms, whether fault-based or strict liability, fail to address harms arising from AI systems, and in what ways might a state-backed no-fault compensation scheme provide a more effective alternative?

To investigate this overarching question, the study sets out the following specific research objectives:

- **Critically examine** the limitations of existing liability regimes, both fault-based and strict liability, in addressing compensation for harms arising from AI systems.
- **Evaluate** the potential of state-backed no-fault compensation schemes to address liability gaps within regulatory sandbox frameworks.

*This research was conducted independently prior to submission. At the time of the conference, the author is affiliated with the London School of Economics and Political Science, UK (c.h.tran@lse.ac.uk).
Copyright © 2025, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

- **Develop and assess** institutional design options, supported by comparative case studies, to explore the feasibility and adaptability of such schemes across diverse legal and regulatory contexts.

Current Progress

The present research has generated several key insights, primarily corresponding to Objectives 1 and 2. By contrast, the aspects related to Objective 3 remain less developed and are reflected in the “Future Work” section.

Compensation Liability Gaps in AI Regulatory Sandboxes

Regulatory sandboxes provide a controlled environment for testing emerging technologies, including AI, under regulatory oversight. Yet they cannot fully mitigate the inherent risks of AI systems (Yoshikawa 2019; Kolt 2023). As a result, participating companies typically remain liable for third-party harms, reflecting a broader consensus that firms should bear civil or criminal responsibility, though enforcement varies across jurisdictions (Carneiro 2024; OECD 2025). The challenge is not only whether companies should be liable, but that they are often the sole bearers of responsibility, both during sandbox trials and after regulatory approval. Whether public authorities should share this responsibility remains unresolved, a concern compounded by the inadequacy of traditional compensation mechanisms, which are ill-suited to address the opacity, diffuse accountability, and causation challenges characteristic of AI-related harms.

This research has highlighted the limitations of both strict liability and fault-based liability regimes in addressing the distinctive risks posed by AI systems within leading regulatory sandbox programs. It argues that current models, which primarily assign civil and criminal liability to participating firms, are inadequate for managing AI-related harms during experimental deployment and after market approval. Three principal shortcomings underpin this inadequacy. First, legal exemptions granted within sandbox programs may restrict access to traditional fault-based compensation, thereby narrowing legal remedies for affected individuals. Second, the opaque and decentralized nature of AI systems, often described as “black box” technologies, hinders the clear attribution of liability, even when causal links are evident. Third, many AI-related harms are systemic, diffuse, and socioeconomically far-reaching, making company-level liability insufficient to capture their structural dimensions or ensure accountability. These challenges are particularly acute for small and medium-sized enterprises (SMEs) in low- and middle-income countries (LMICs), whose limited resources constrain their ability to internalize or remediate such harms.

State-Backed No-Fault Compensation

This research also advanced a relatively underexplored yet promising proposal in AI governance: the establishment of a state-backed, no-fault compensation program for harms arising from experimentation within AI regulatory sandboxes. Such a mechanism might operate in parallel with, and complementary to, existing strict liability frameworks. It ac-

knowledges the limitations of traditional fault-based liability principles, particularly the unrealistic expectation that compensation must hinge on proving negligence, and mitigates the potential chilling effect of imposing strict liability solely on private enterprises. This concern is especially salient in the AI context, where harms may stem from inherent risks in system design and operation, even absent identifiable error or negligence.

Drawing on precedents such as the Vaccine Injury Compensation Programs (VICPs), the proposed model advocates a publicly funded, non-adversarial compensation scheme that eliminates the need to prove fault, negligence, or product defect (Yang and Reiss 2023). Its design reflects a deliberate policy choice to balance two core imperatives: ensuring timely and fair compensation for parties harmed by socially valuable technologies, and preserving the conditions necessary for sustainable innovation and public welfare (Yang and Reiss 2023; Mungwira et al. 2020). In this respect, it offers a persuasive model for AI sandbox governance, which also requires mechanisms to manage the tension between experimental innovation and potential harm.

The rationale for this approach rests on four interrelated grounds. First, by actively designing and supervising sandbox environments, the state becomes a co-producer of legal risk and therefore bears a principled duty to provide compensation mechanisms. Second, a no-fault program would overcome the evidentiary and legal challenges posed by the opacity and distributed agency of AI systems, which often disadvantage claimants under traditional liability regimes. Third, the model would reduce compliance burdens on SMEs, particularly in LMICs, thereby promoting more inclusive and equitable innovation ecosystems. Fourth, by offering predictable and accessible avenues for redress, such a framework could enhance public trust and regulatory legitimacy, ultimately supporting the responsible development and deployment of AI technologies in sandbox settings.

Future Work

Future research should pursue two complementary directions. The first is empirical: In-depth case studies of jurisdictions that have implemented regulatory sandboxes for disruptive technologies would provide valuable insights into how liability regimes function in practice and how they intersect with AI-related risks. The second is normative and institutional: further work should examine the practical pathways for implementing a state-backed no-fault compensation scheme. This includes clarifying whether such a mechanism should operate alongside and supplement existing litigation and civil compensation processes or instead replace them. Equally important is evaluating the feasibility of such a scheme, including the state’s financial capacity to share compensation responsibilities with private actors and the appropriate allocation of liability between governments and industry participants in sandbox environments.

Acknowledgments

I am grateful to Mr. Duc-Lam Nguyen and Ms. Anh-Tho Tran for their thoughtful reviews and comments.

References

- Abiri, G.; and Huang, Y. 2023. A Red Flag? China's Generative AI Dilemma. *Harvard Journal of Law & Technology*, 37.
- Ahern, D. 2025. The New Anticipatory Governance Culture for Innovation: Regulatory Foresight, Regulatory Experimentation and Regulatory Learning. *European Business Organization Law Review*, 26(2): 241–283.
- Allen, H. J. 2020. Sandbox Boundaries. *Vanderbilt Journal of Entertainment and Technology Law*, 22(2): 299–321.
- Arnal, J. 2025. AI at Risk in the EU: It's Not Regulation, It's Implementation. *European Journal of Risk Regulation*, 1–10.
- Arrieta, A.; Díaz-Rodríguez, N.; Del Ser, J.; Bennetot, A.; Tabik, S.; Barbado, A.; Garcia, S.; Gil-Lopez, S.; Molina, D.; Benjamins, R.; Chatila, R.; and Herrera, F. 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58: 82–115.
- Carneiro, G. 2024. *Moving Beyond the Innovation-Regulation Dichotomy: Regulatory Sandboxes for Artificial Intelligence Regulation*. LL.M. diss., Harvard School of Law, Harvard University, Boston, MA.
- Carvão, P.; Ancheva, S.; Atir, Y.; Jeloka, S.; and Zhou, B. 2025. Governance at a Crossroads: Artificial Intelligence and the Future of Innovation in America. Technical report, Mossavar-Rahmani Center for Business & Government, Harvard University.
- Cornelli, G.; Doerr, S.; Gambacorta, L.; and Merrouche, O. 2024. Regulatory Sandboxes and the FinTech Funding: Evidence from The UK. *Review of Finance*, 28(1): 203–233.
- Dixon, R. B. L. 2023. A Principled Governance for Emerging AI Regimes: Lessons From China, The European Union, and The United States. *AI and Ethics*, 3(3): 793–810.
- Karnow, C. E. A. 2016. The application of traditional tort theory to embodied machine intelligence. In *Robot Law*, 51–77. Edward Elgar Publishing.
- Kilian, R.; Jäck, L.; and Ebel, D. 2025. European AI Standards - Technical Standardisation and Implementation Challenges under the EU AI Act. *European Journal of Risk Regulation*, 1–25.
- Knight, B.; and Mitchell, T. 2020. The Sandbox Paradox: Balancing the Need to Facilitate Innovation with the Risk of Regulatory Privilege. *South Carolina Law Review*, 72(2): 445–475.
- Kolt, N. 2023. Algorithmic Black Swans. *Washington University Law Review*, 101(4).
- Moraes, T. 2025. Regulatory Sandboxes For Trustworthy Artificial Intelligence - Global and Latin American Experiences. *International Review of Law, Computers & Technology*, 39(1): 55–74.
- Morgan, D. 2023. Anticipatory regulatory instruments for AI systems: A comparative study of regulatory sandbox schemes. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, 980–981. AAAI Press.
- Mungwira, R. G.; Guillard, C.; Saldaña, A.; Okabe, N.; Petousis-Harris, H.; Agbenu, E.; Rodewald, L.; and Zuber, P. L. F. 2020. Global Landscape Analysis of No-fault Compensation Programmes for Vaccine Injuries: A Review and Survey of Implementing Countries. *PloS one*, 15(5).
- OECD. 2025. Regulatory Sandbox Toolkit: A Comprehensive Guide for Regulators to Establish and Manage Regulatory Sandboxes Effectively. Technical report, The Organization for Economic Co-operation and Development.
- Qiu, Y.; Yao, H.; Ren, P.; Tian, X.; and You, M. 2025. Regulatory Sandbox Expansion: Exploring The Leap from Fintech to Medical Artificial Intelligence. *Intelligent Oncology*, 1(2): 120–127.
- Ranchordas, S. 2021. Experimental Regulations for AI: Sandboxes for Morals and Mores. *Morals & Machines*, 1(1): 86–100.
- Rangone, N.; and Megale, L. 2025. Risks Without Rights? The EU AI Act's Approach to AI in Law and Rule-Making. *European Journal of Risk Regulation*, 1–16.
- Scherer, M. U. 2016. Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. *Harvard Journal of Law & Technology*, 29(2): 353.
- Yang, Y. T.; and Reiss, D. R. 2023. Vaccine Injury Compensation in the United States. In *Vaccine Law and Policy*, 43–54. Springer International Publishing.
- Yoshikawa, J. 2019. Sharing the Costs of Artificial Intelligence: Universal No-Fault Social Insurance for Personal Injuries. *Vanderbilt Journal of Entertainment and Technology Law*, 21(4).
- Zhang, X.; Zhou, W.; Hu, W.; Zhou, S.; Hu, X.; and Yang, L. 2025. Can Ex Ante Conformity Assessment Regulations Contribute to Trustworthy Foundation Models? An Evolutionary Game Analysis from an Innovation Ecosystem Perspective. *Technology In Society*, 82.