

The Institutional Logic and Regulatory Architecture of Financial Data Regulatory Sandboxes

Ye Ju, Meiting Zhan *

Faculty of Law, College of Applied Arts and Science, Beijing China

*Corresponding Author: Meiting Zhan (Email: 18959928425@163.com)

Abstract

The core challenge in financial data governance has shifted from mere rights protection to the construction of trustworthy circulation of data. Traditional regulation, which relies on ex-ante rules and ex-post accountability, struggles to adapt to the compound risks generated by continuous data flows and iterative model development. It often finds itself trapped between regulatory lag and stifled innovation. The institutional significance of the financial data regulatory sandbox lies in embedding data processing and algorithmic operations within an observable and verifiable regulatory process, transforming uncertain risks into empirical knowledge necessary for rule recalibration. Although extraterritorial practices vary across jurisdictions, they collectively point toward a transformation of regulatory capacity—from case-by-case approvals to infrastructure development. Whether through data infrastructure supporting regulatory learning, model validation tools embedded in regulatory workflows, or real-world scenarios testing data authorization and consumer protection, the essence is to ground regulatory judgment in empirical evidence through controlled testing. China's construction of a financial data regulatory sandbox should transcend the mindset of temporary exemptions and project-based pilots, moving toward an institutionalized regulatory mechanism anchored in lawful data entry, model impact assessments, process auditing, consumer remedies, and rule feedback.

Keywords

Financial Data, Regulatory Sandbox, Digital Finance, Data Governance, AI Regulation.

1. The Problem Statement: Why Financial Data Governance Requires an Experimental Institution

The deep integration of finance and technology is reshaping the operational logic of financial markets [1]. Data is no longer merely a business record or a risk-control auxiliary; it has become a foundational condition for financial service generation, risk identification, and value allocation. Its value does not stem from static possession but arises from the dynamic processes of continuous circulation and model feedback. The resulting institutional challenge lies in the fact that increased data liquidity simultaneously delivers efficiency dividends and risk spillovers, creating a structural feature where value creation and risk accumulation expand in tandem. The difficulty of financial data governance is rooted precisely in this inherent contradiction.

Traditional financial regulation, which relies primarily on institutional admission and ex-post sanction, is suited to relatively stable financial activities but falls short in adequately addressing data-driven innovation [2]. The risks in digital finance are no longer concentrated in transaction outcomes; they are endogenous to the successive technical links of data processing and model operation. Accordingly, the regulatory target has expanded from licensed financial institutions to various technology service providers along the data processing chain. In this context, relying

solely on document review and outcome-based accountability is no longer sufficient to determine, in a timely manner, whether an innovative solution is genuinely controllable [3]. The regulatory sandbox thus carries deeper institutional significance. It is not merely a general tolerance for innovators, nor is it simply a facilitation arrangement to shorten product time-to-market. Rather, it should be understood as an experimental regulatory technique: within legal boundaries, the regulatory authority permits innovators to conduct controlled testing, accumulating judgmental bases through process records and risk feedback. When existing rules cannot provide predetermined answers to novel data processing issues, the sandbox can provide an empirical foundation for rule generation.

On this basis, three fundamental propositions regarding the financial data regulatory sandbox can be formulated: its core task is to transform uncertain risks into observable, auditable, and correctable objects of regulation; its institutional unit should shift from individual projects to the life cycle of data and models; and its effectiveness ultimately depends on the capacity for regulatory learning itself. If test results cannot be crystallized into rules, standards, or regulatory guidelines, the sandbox will remain at the level of case-by-case support and will fail to evolve into a stable data governance institution.

2. The Boundaries, Institutional Functions, and Normative Limits of the Financial Data Regulatory Sandbox

2.1. From Fintech Sandbox to Financial Data Sandbox

The fintech regulatory sandbox primarily revolves around products, services, and business models, focusing on whether innovative activities can obtain regulatory flexibility within a limited scope. The financial data regulatory sandbox has a more focused and penetrating concern: it examines whether the data processing underpinning the relevant business is itself legitimate, and how model outputs affect financial consumer clients and market order [4].

Accordingly, the financial data regulatory sandbox can be defined as a controlled testing arrangement organized by financial regulatory authorities or through cooperative regulatory mechanisms [5]. Its scope covers the entire life cycle of financial data processing, coordinating the relationship between data security, consumer rights, and innovation efficiency through process auditing and rule feedback. Data, models, scenarios, and accountability together constitute the fundamental control boundaries of this institution.

2.2. Institutional Function: From Compliance Confirmation to Evidence Generation

The distinctive value of the financial data regulatory sandbox lies in its function of evidence generation. General compliance review determines whether a solution conforms to existing rules, whereas the sandbox observes risks that are not yet fully covered by rules. These risks mostly arise from the specific operational processes of data processing and algorithmic model behavior. For instance, whether de-identification can genuinely block re-identification risks is difficult to determine based on abstract provisions alone; it requires empirical evidence obtained from actual trajectories within a controlled environment.

The financial data regulatory sandbox thus performs three simultaneous institutional functions: providing a communicable compliance pathway for innovators, thereby reducing the cost of uncertainty faced by lawful innovation; supplying regulatory authorities with technical facts and behavioral evidence, thereby enhancing the precision of rule-making and enforcement decisions; and embedding necessary procedural safeguards for data subjects and financial consumer clients, ensuring that risk allocation is no longer determined unilaterally by the technical system.

2.3. Normative Limits: Guarding Against Technological Solutionism and Sandbox Opacity

The financial data regulatory sandbox must not reduce complex rights-configuration issues to mere technical engineering problems. Financial data processing implicates multiple legal interests, and no single technical tool can automatically complete authorization judgments, responsibility allocation, or public oversight. If the sandbox degenerates into a showcase for technical solutions, regulatory authorities are likely to be led by technical narratives, thereby weakening their scrutiny of procedural fairness. The sandbox may also invite regulatory capture and competitive unfairness. Participating firms often gain access to more intensive regulatory resources, placing non-participants at a relative disadvantage. To prevent the sandbox from devolving into selective favoritism, entry criteria and exit conclusions should maintain necessary transparency. Trade secrets may be protected through desensitization, but the operation of the institution itself should not become an unobservable black box.

3. Extraterritorial Rules and Operational Lessons: Four Institutional Forms of Financial Data Sandboxes

3.1. The United Kingdom: Shaping the Financial Digital Sandbox through Compliant Datasets and Secure Testing Environments

The UK experience presents a dual-track approach: the data protection sandbox and the financial digital sandbox. The ICO's Data Protection Regulatory Sandbox translates principled privacy protection requirements into enforceable rules in specific contexts through case-specific guidance and process monitoring. Projects such as FutureFlow and the Heathrow Airport initiative demonstrate that many complex compliance determinations can be brought within the regulatory purview at the testing stage.

The FCA Digital Sandbox more directly embodies the institutional form of a financial data sandbox [6]. It builds upon synthetic data and secure development environments, incorporating data accessibility itself into regulatory infrastructure development. For data-intensive financial innovation, the lack of shareable and verifiable compliant data resources often constrains innovation more than the lack of algorithms themselves. By reducing exposure to real sensitive data, the digital sandbox enables model validation and regulatory observation to be completed within the same institutional space.

The implication of the UK model is that the financial data regulatory sandbox should not remain a mere regulatory window or application procedure; it should pivot toward data infrastructure development itself [7]. Only with this infrastructure in place can regulatory authorities continuously observe risks, compare model performance, and refine regulatory standards. Without it, the sandbox can only handle isolated cases and cannot generate public governance capacity.

3.2. Singapore: Bridging Data Development and Model Validation through APIX and Veritas

Singapore has not established a single institution bearing the name "financial data regulatory sandbox." Its distinctive feature lies in combining the regulatory sandbox, open-interface platforms, and model evaluation mechanisms to form a composite governance framework. The key to the Singaporean experience is not the institutional nomenclature but the connection of data development, model validation, and regulatory communication into a continuous process. This model demonstrates that the financial data sandbox can move from a procedural arrangement to platform-based governance. The open-interface platform itself does not generate regulatory exemptions, but it provides necessary conditions for development and connectivity for data-driven innovation. Similarly, model evaluation tools do not substitute for

legal compliance, but they help financial institutions identify model biases and responsibility allocation issues. Together, these two elements illustrate that financial data regulation cannot stop at whether data can enter the system; it must also focus on how models absorb data and affect client interests.

The Singapore experience particularly addresses a deep-seated challenge in financial AI governance: the risks of machine learning models often lurk in the technical details of training and fine-tuning, making after-the-fact review insufficient to detect deviations in a timely manner. This combination suggests that the sandbox should provide reusable model validation tools, enabling principles such as fairness and transparency to move from declaratory statements to testable evaluation metrics.

3.3. Hong Kong SAR: Advancing Real-Scenario Testing through Authorized Data Sharing and Generative AI Sandboxes

The distinctive feature of Hong Kong's fintech regulatory sandbox is its use of banks' real business processes as a limited testing field [8]. Its significance for financial data extends beyond technical trials; it lies in testing, through the cycle between real business data and compliance measures, whether innovative solutions can withstand constraints in client protection. The authorized data-sharing infrastructure provides the scenario conditions for Hong Kong's sandbox testing, creating a genuine, auditable, and authorized data channel. Regulatory authorities can thereby observe the effectiveness of authorization mechanisms and whether financial institutions risk using data beyond the authorized scope.

The Generative AI Sandbox advances financial data governance into new application scenarios. Its value is not merely to encourage financial institutions to adopt large models, but to incorporate model safety and interpretability responsibilities into controlled testing. Generative AI magnifies the uncertainty of data processing: unstructured inputs may expand privacy risks, and the unreliability of model outputs may mislead client decisions. The Hong Kong experience demonstrates that real-scenario testing helps expose risks, but must be accompanied by authorization and remedy mechanisms to limit risk diffusion.

3.4. The European Union: The DLT Pilot Regime as an Indicator of the Infrastructure-Oriented Sandbox Path

The EU DLT Pilot Regime is not a narrow regulatory sandbox but a market infrastructure pilot system with sandbox functionality [9], providing a controlled space for innovation in capital market infrastructure. The EU model expands the institutional imagination of the financial data sandbox: distributed ledger technology changes not only data storage methods but also may reshape fundamental institutional links such as financial transaction and settlement. Traditional product-based sandboxes are ill-equipped to accommodate such systemic effects, hence the need for infrastructure-oriented pilot mechanisms [10]. The coordination arrangements between Member States and the EU level establish regulatory balance between local discretion and uniform standards.

This experience suggests that China's financial data regulatory sandbox should be advanced in a tiered manner. General financial data applications are suited to project-based sandboxes; open finance and credit-reference collaborations more urgently require data-infrastructure sandboxes; and scenarios involving systemic financial infrastructure should enter infrastructure-oriented pilots [11]. Different tiers correspond to different risk intensities, and a single model cannot cover all forms of innovation. Extraterritorial financial data regulatory sandboxes, through differentiated instruments adapted to local contexts, have established collaborative mechanisms that provide diverse pathway references for improving China's financial data governance system (see Table 1).

Table 1. Instrument Types and Institutional Functions of Extraterritorial Financial Data Regulatory Sandboxes

Country/Region	Primary Instruments	Financial-Data-Related Mechanisms	Institutional Functions	Implications for China
United Kingdom	ICO Sandbox, FCA Digital Sandbox	Compliant datasets, secure development environments, synthetic data, and APIs	Supporting regulatory learning and model validation through data infrastructure	Establish a national-level synthetic financial data and secure testing environment
Singapore	MAS Sandbox, APIX, Veritas	Open interfaces, data development environments, and AI model validation tools	Translating principled AI governance into testable evaluation procedures	Develop sector-wide model validation templates for financial AI
Hong Kong SAR	FSS, CDI, IADS, Gen. A.I. Sandbox	Authorized data sharing, real-business feedback, and GenAI scenario testing	Testing data authorization, model outputs, and client protection in real scenarios	Strengthen limited-client testing, cross-agency coordination, and consumer remedies
European Union	DLT Pilot Regime	DLT transaction, registration, settlement, and regulatory data coordination	Observing financial market infrastructure changes through pilot regimes	Explore infrastructure-oriented financial data sandboxes and cross-sectoral coordination

4. The Regulatory Architecture of China's Financial Data Regulatory Sandbox

4.1. Institutional Positioning: From Innovation Facilitation to Auditable Governance

China has already established a parallel framework consisting of fintech innovation regulatory tools, data governance rules, and personal information protection systems. The next critical step is to clarify the positioning of the financial data regulatory sandbox within the existing institutional system. It should be understood as an auditable governance instrument. It does not replace administrative licensing or prudential regulation, nor does it weaken baseline obligations such as data security. Its function is to address novel data processing activities that existing rules cannot readily anticipate, and to generate regulatory evidence through testing. Under this positioning, the implementation of China's financial data regulatory sandbox should follow several common principles: the intensity of testing should match the degree of risk; participating projects must not expand the scope of data processing under the pretext of testing; the entire testing process should generate traceable records; and test results should serve rule formation. These principles point to the same logic: the legitimacy of the sandbox derives from risk control and evidence accumulation, not from regulatory exemption itself.

4.2. Entry Review: Centering on Public Interest and Regulatory Uncertainty

The financial data sandbox should not be open to all innovative projects. The core of entry review lies in assessing whether a project possesses substantive innovative value, whether existing rules genuinely fail to fully address its compliance boundaries, and whether the applicant has demonstrated a baseline level of credible risk-control capability. Technical novelty can only constitute a necessary condition; it cannot substitute for public value and compliance fundamentals. Particular caution is needed against projects that use technical

sophistication to mask commercial motives: if data processing primarily serves user-profile expansion or regulatory avoidance, it should not enter the sandbox even if novel technologies are employed. The regulatory authority should require applicants to submit adequate risk assessments, ensuring that entry decisions rest on evidence rather than on technical narratives or commercial promises.

4.3. Operational Rules: Dual-Track Governance Based on the Data Life Cycle and the Model Life Cycle

The operational rules of the financial data regulatory sandbox should take the data life cycle as the primary thread, covering the continuous process of data entry, processing, output, and exit, with corresponding review and control requirements at each stage. This continuous control from entry to exit constitutes the institutional feature that distinguishes the sandbox from ordinary project filings. Model life-cycle governance is equally indispensable. For machine learning and generative AI projects, regulatory authorities should focus on the entire process from training to deployment. Testing standards should not pursue accuracy alone; they should also examine the fairness and interpretability of algorithmic model behavior. Any automated decision that directly affects client rights should be accompanied by procedures for explanation, review, and remedy.

4.4. Coordination Mechanisms: From a Single Regulatory Window to Pluralistic Co-Governance

Financial data processing naturally crosses multiple regulatory domains. If a single department operates the sandbox in isolation, it risks producing inconsistent rule interpretations and delayed risk transmission. China may establish a composite governance structure led by a primary regulatory authority, with the participation of relevant departments and technical institutions, enabling division of labor and coordination in entry judgments, domain-specific constraints, and technical verification. Transparency is the precondition for effective coordination mechanisms. The sandbox should not disclose trade secrets, but it should disclose the purposes of testing, principal risks, and summaries of exit conclusions. Test results of general significance should be transformed into public rules. The precipitation of case-specific experience into institutional resources helps transform the sandbox from an innovation-facilitation tool into an institution-generating mechanism.

5. China's Solution for the Financial Data Regulatory Sandbox: Tiered Classification and Closed-loop regulatory cycle

Drawing on extraterritorial experience and China's institutional foundations, the financial data regulatory sandbox can be structured into three tiers: application-oriented, platform-oriented, and infrastructure-oriented. These correspond to different risk intensities and governance needs, and a single institutional model should not be applied indiscriminately to all scenarios. Its operational process can be summarized as five stages: entry review, boundary setting, process auditing, exit assessment, and rule transformation. These five stages together constitute a closed loop from case-specific testing to public rules. This solution must particularly avoid two misunderstandings: the sandbox is not a regulatory preferential policy; its flexibility derives from risk control and evidence generation. Nor is the sandbox a purely technical experiment; financial data governance always involves rights allocation and public trust [12]. Only by combining technical testing with institutional feedback can the financial data regulatory sandbox become an effective instrument for China's digital financial governance.

6. Conclusion: From Sandbox Experimentation to Intelligent, Prudent, and Trustworthy Financial Data Governance

The financial data regulatory sandbox addresses how regulation should manage highly uncertain data behaviors in the digital finance era. Its institutional significance can be summarized as risk isolation, capacity building, and rule learning. At the individual level, it protects data subjects; at the institutional level, it enhances compliance capabilities; and at the systemic level, it accumulates verifiable regulatory experience. Comparative experience indicates that the regulatory sandbox is evolving from traditional fintech experimentation into a more sophisticated data governance instrument, and will increasingly depend on auditable and reusable regulatory capacity. China's construction of a financial data regulatory sandbox should take the life-cycle governance of data and models as its main thread [13], striking a dynamic balance between financial stability and innovation efficiency. As regulatory technology and privacy-preserving computation continue to mature, financial data governance has the potential to move from empirical pilots toward an explainable and accountable institutional system.

References

- [1] Huang, Y., & Huang, Z. (2018). Digital finance in China: present and future. *China Economic Quarterly*, 17(4), 1492. [In Chinese]
- [2] Xu, D. (2018). The disruptive innovation nature of FinTech and new approach to RegTech. *Oriental Law*, (2), 7. [In Chinese]
- [3] Liao, F. (2019). Analysis of the theory and practice of regulatory sandbox in the context of FinTech. *Journal of Xiamen University (Philosophy and Social Sciences)*, (2), 14. [In Chinese]
- [4] Li, W., & Jiang, Z. (2017). FinTech development and regulation: a regulator's perspective. *Financial Regulation Research*, (3), 7. [In Chinese]
- [5] Li, S., & Li, M. (2021). Reflection and prospect of FinTech regulation: taking sandbox regulation as an example. *Gansu Social Sciences*, (3), 112. [In Chinese]
- [6] Li, W., & Jiang, Z. (2017). FinTech development and regulation: a regulator's perspective. *Financial Regulation Research*, (3), 10. [In Chinese]
- [7] Cornelli, G., Doerr, S., Gambacorta, L., & Merrouche, O. (2024). Regulatory sandboxes and fintech funding: evidence from the UK. *Review of Finance*, 28(1), 203.
- [8] Tan, Y. (2024). FinTech innovation and regulation: taking Hong Kong regulatory sandbox as an example. *FinTech Era*, 32(4), 46. [In Chinese]
- [9] Wang, D., & Zhai, R. (2021). International experience and policy implications of FinTech regulatory sandbox. *Contemporary Financial Research*, (Z4), 101. [In Chinese]
- [10] Dardykina, A. (2025). Is there enough sand in the sandbox? *European Business Law Review*, 36(7), 1083.
- [11] Lehmann, M. (2023). Regulating innovation through digital platforms: the sandbox tool. *European Company and Financial Law Review*, 19(5), 828.
- [12] Zhang, Y., Qu, J., & Zhang, X. (2025). AI regulatory sandbox: concept definition, practical experience and implications for China. *World Sci-Tech R&D*, 47(S1), 54. [In Chinese]
- [13] Ye, X. (2024). Regarding the Regulatory Sandbox Route and Mechanism for Governance of Artificial Intelligence. *Science Technology and Law (Chinese-English Version)*, (5), 136. [In Chinese]