FinTech-Enabled Small and Medium-sized Enterprises Credit: Data Governance Bottlenecks in Big Data Applications and Institutional-Technological Synergy Solutions

Yurun Wang

College of Arts and Information Engineering, Dalian Polytechnic University, Dalian, China
wang_yurun@outlook.com

Abstract. For the stable development of small and medium-sized enterprises, the credit issue has always been an important constraint. With the steady development of financial technology, big data technology has been applied in the credit field for small and medium-sized enterprises, significantly enhancing the efficiency of bank credit approval and risk identification capabilities. However, the existence of data governance issues such as data silos, privacy leaks, and unclear ownership has restricted the full release of the application potential of big data technology. This paper systematically examines the current application status of big data technology in the credit of small and medium-sized enterprises, identifies key governance bottlenecks such as data acquisition, data quality, security and compliance, and proposes a feasible path for the collaborative governance of system and technology. By constructing a four-dimensional collaborative model of "regulatory framework - data standards - technical tools - enterprise collaboration", further suggestions are provided for promoting the full empowerment of financial technology in credit support for small and medium-sized enterprises.

Keywords: Financial technology, credit for small and medium-sized enterprises, big data, institutional and technological synergy.

1. Introduction

For some developing countries, small and medium-sized enterprises are an important pillar of national economic growth. However, small and medium-sized enterprises have long been confronted with the problems of difficult and expensive financing. The constraints of funds make it hard for Small and Medium-sized Enterprise (SME) to continuously expand their scale and improve their business quality. The traditional credit model in the banking industry relies heavily on asset collateral and corporate financial statements, making it difficult to effectively assess the credit risks of small and medium-sized enterprises with light assets. In recent years, fintech has provided new ideas for solving this difficult problem through new technological means such as big data and artificial intelligence.

Big data technology can help financial institutions such as banks comprehensively integrate various important data of enterprises, including multi-dimensional unstructured data such as various assets, cash flows, taxes, supply chains, and social behaviors, thereby enabling the construction of a more comprehensive credit profile [1]. However, the extensive collection and application of enterprise data by financial institutions have also given rise to governance challenges such as unclear data ownership, privacy infringement, and algorithmic discrimination. If there is a lack of an effective data governance framework, the support of new technologies will instead inhibit the healthy development of the factor market and accumulate systemic risks.

Against the backdrop of fintech empowering small and medium-sized enterprise credit, this article focuses on analyzing the data governance bottlenecks faced by the application of big data and how to achieve scientific data governance through the collaborative path of systems and technologies. Under the problem-oriented approach, a four-dimensional collaborative governance model is necessary.

2. The current application status of big data in credit for small and medium-sized enterprises

Table 1. Types of data in SME credit, their value, and governance challenges

Data Type	Source	Credit Value	Governance Challenges
Transaction Records	Banks, Payment Platforms	Reflects operational stability	Data silos, inconsistent interface standards
Tax Data	Tax Authorities	Verifies revenue authenticity	Limited government data openness
Supply Chain Data	Enterprise Resource Planning (ERP) Systems of Core Firms	Assesses upstream/downstream risks	Protection of trade secrets
Online Public Opinion	Social Media, News	Identifies reputation risks	High data noise, subjectivity
Owner's Personal Data	Credit Reports, Consumption Records	Supplements enterprise credit assessment	Blurred privacy boundaries

The traditional credit model relies one-sidedly on financial statements and collateral assets, making it difficult to accurately assess the credit of small and medium-sized enterprises with light assets. Based on the original model, big data technology, by integrating multi-dimensional non-financial data, is conducive to enhancing the coverage breadth of credit services provided by financial institutions and the accuracy of enterprise risk identification. At present, big data technology can be applied in many scenarios of credit assessment for small and medium-sized enterprises. During the process of transaction data analysis, the daily income and expenditure data of enterprises can be obtained through bank transactions or payment platforms, thereby reflecting the stability of the enterprise's operation and cash flow status. During the process of tax data verification, the authenticity of a company's profits and financial status can be effectively verified by using value-added tax invoices and the company's tax payment records. In the process of e-commerce and supply chain information analysis, relying on data such as order size, revenue, and user reviews from e-commerce platforms, the market performance of enterprises can be evaluated. In the process of behavioral public opinion data analysis, credit records in aspects such as enterprise management, consumer behavior, and online public opinion can be utilized to supplement the enterprise's credit archives and avoid potential reputation risks that the enterprise may face.

For instance, Alibaba's "310" model (3-minute application, 1-second loan disbursement, and zero human intervention) relies on its internal ecosystem and the massive transaction data of its platform to achieve full-process automated credit services. The "Huidongni" platform of China Construction Bank has provided credit lines to over 4 million small and medium-sized enterprises by integrating more than 10 types of data, including tax, industry, and commerce, and credit reporting, with an average approval time of less than 3 minutes.

However, since enterprise financial data is mainly held by various platforms, there are widespread problems such as chaotic sources, complex forms, and sensitive privacy issues, which pose challenges to data governance and market improvement, as shown in Table 1.

3. The core bottleneck of data governance

Although big data technology has potential in the application of credit for small and medium-sized enterprises, it also faces deep-seated governance challenges. These issues are not merely technical problems; they also encompass aspects such as institutional absence, market boundaries, and ethical dilemmas.

3.1. Data silos and the absence of sharing mechanisms

Data silos are the primary obstacle restricting the empowerment of credit by big data technology. A lot of key data is separately held by entities such as the government, financial institutions, and e-commerce platforms, and there is a lack of necessary sharing mechanisms among these entities. For instance, the tax data held by the tax authorities can truly reflect the revenue of enterprises, but the application of financial institutions requires approval, with a complex process and a long cycle. Meanwhile, there is a lack of an incentive mechanism for data sharing. Data holders generally worry that sharing will weaken their competitive edge or trigger security responsibilities, leading to a fragmented situation in the data element market. This "data segmentation" pattern may intensify market monopolies, and the interests of small and medium-sized enterprises will be affected due to data segmentation.

3.2. The data quality and standards are inconsistent

Data from small and medium-sized enterprises generally suffers from problems such as fragmentation, lack of standardization, and lagging updates. For instance, if the payment platform data of individual business owners is mixed with personal consumption and business income and used directly without processing, it will lead to deviations in credit assessment levels and an increase in risks. Research shows that the misjudgment rate of risk control models trained on unstandardized datasets will rise sharply [2]. In addition, various data standards are not uniform. In terms of the definition standards for "active merchants" on different e-commerce platforms, Taobao takes a monthly transaction volume of more than 5,000 yuan as the standard, while JD.com requires a monthly order number of more than 10. This standard differentiation makes it difficult to aggregate and analyze cross-platform data. Although some international standards for data quality have been published, they have not been effectively implemented in the field of fintech [3].

3.3. Privacy protection and compliance risks

With the implementation of regulations such as the General Data Protection Regulation and China's Personal Information Protection Law, data privacy and protection have become a key focus [4]. In

the credit risk assessment of small and medium-sized enterprises, data analysis may involve sensitive information such as the consumption records and preferences of business owners [5]. Once the legal bottom line is crossed, it may lead to problems such as data leakage. Furthermore, the boundary between enterprise and personal data is not clear. Small and micro business owners also use individual users during their business operations [6]. Their personal credit and enterprise credit are highly coupled and difficult to distinguish. How to make reasonable use of mixed data while protecting personal privacy is not only a difficult problem for regulatory authorities, but also leads to compliance issues for financial institutions.

3.4. Data ownership and benefit distribution are unclear

The issue of data ownership is the root cause of the governance predicament. Who owns the financial and other data generated by small and medium-sized enterprises when they conduct transactions on e-commerce platforms? Is it a platform or an enterprise? Current laws recognize the platform as the data owner, but they overlook the contributions of data producers, namely, small and medium-sized enterprises and consumers [7]. As the main producers of data elements, small and medium-sized enterprises are unable to share the value-added benefits of data, resulting in the phenomenon of "platform data exploitation". Furthermore, when property rights are not clearly defined, it is impossible to quantify the contribution data of each data subject, and thus, it is difficult to establish a data revenue distribution mechanism. In the past, pricing was mainly based on the volume of data, but the quality and scarcity of the data were ignored, making it difficult to reflect the contributions of different entities. Without fair rules for revenue distribution, the marketization level of data elements simply cannot be improved [8].

From the above analysis, it can be seen that the bottleneck of data governance is the result of an imbalance in multiple dimensions such as technology, system, market, and ethics. To break through the governance predicament, it is necessary to establish a systematic and collaborative mechanism.

4. Institutional and technological collaborative governance path

In the face of complex data governance challenges, traditional regulatory methods or purely technical solutions are difficult to achieve the expected results. Starting from a single path inevitably has limitations: relying solely on institutional restrictions is not conducive to innovation, while relying solely on technology will be difficult to control [9]. Therefore, it is necessary to shift to a collaborative governance paradigm that combines systems and technologies, playing an incentive role in the process of risk control and emphasizing the dynamic interaction between system design and technological empowerment. This article constructs a framework of "regulatory system - data standards - technical tools - enterprise collaboration" for analysis.

Figure 1 is a basic explanation of this framework. Under a reasonable regulatory design, promoting data standardization, technological empowerment and enterprise cooperation is conducive to the formation of a well-functioning data element market mechanism and culture.
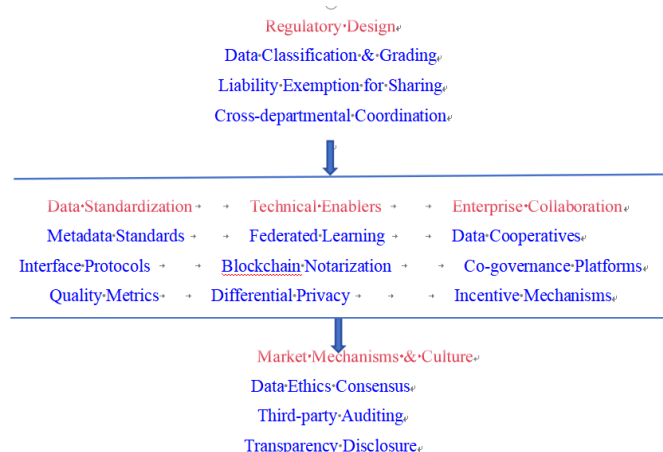
Regulatory·Design

Data·Classification·&·Grading

Liability·Exemption·for·Sharing

Cross-departmental·Coordination

Data·Standardization	Technical·Enablers	Enterprise·Collaboration
Metadata·Standards	Federated·Learning	Data·Cooperatives
Interface·Protocols	Blockchain·Notarization	Co-governance·Platforms
Quality·Metrics	Differential·Privacy	Incentive·Mechanisms

Market·Mechanisms·&·Culture

Data·Ethics·Consensus

Third-party·Auditing

Transparency·Disclosure

Figure 1. Institutional-technical collaborative governance framework (picture credit: original)

4.1. Regulatory system design: building an inclusive legal framework

An effective system is the prerequisite for collaborative governance. At present, supervision mostly adopts "post-event penalties", but what is more important are "pre-event guidance" and "process supervision". This article suggests establishing a hierarchical supervision system. The first step is data classification and ownership definition. In terms of data classification, data such as business registration and taxation are public data and can be made available to the public [10]. Corporate cash transactions, supply chain information, etc., are all commercial data, and enterprises can share them according to the granted permissions. The personal information and other restricted data of business owners can only be held by the owners. At the same time, small and medium-sized enterprises should be granted access rights, control rights and benefit-sharing rights to the data generated by their business operations, so as to break the monopoly of data and data trading platforms. It is necessary to coordinate the design of data opening systems in areas such as finance, industry and information technology, and taxation, and moderately explore new data sharing models.

4.2. Data standard system: promoting the construction of "interoperability"

The standard system is the fundamental rule for data circulation. The current standard system is fragmented and can be addressed through the coordination of mandatory basic standards and voluntary industry standards [11]. By formulating basic standards such as the "Specification for Credit Data Elements of Small and Medium-sized Enterprises", the unification of the data format, coding and transmission standard system should be strengthened. Encourage responsible industry associations to formulate industry-specific standards, clearly defining transaction data of e-commerce platforms and supply chain finance data. In addition, third-party institutions can be invited to rate the database from multiple dimensions, forming a market-oriented mechanism for improving the quality of data elements.

4.3. Technology empowerment tools: from "usable" to "trustworthy"

Technology is not only used as a tool to enhance efficiency, but also as a guarantee to promote the implementation of systems, especially the application of privacy-enhancing technologies. In response to the common problems such as high noise and complex formats in the credit data of small and medium-sized enterprises, financial institutions can utilize data cleaning and

standardization technologies, combined with new algorithms, to correct missing and contradictory data, thereby enhancing the accuracy and consistency of credit data [12]. At the same time, by leveraging artificial intelligence technology, the credit decision-making process can be made transparent, allowing small and medium-sized enterprises to clearly understand the approval rules and enhancing the fairness of the algorithm and user trust. At the data interaction level, monitor the calling behavior to prevent excessive data collection and abuse. By collaboratively applying privacy-enhancing technologies, the transparency and security of data processing can be fully guaranteed, providing reliable technical support for the implementation of the system.

4.4. Enterprise collaboration mechanism: building a data co-governance ecosystem

To reduce credit risks and capital constraints, small and medium-sized enterprises should not merely be one-way providers of credit data in the credit field, but should become participants in data governance. It is suggested that small and medium-sized enterprises can voluntarily unite and negotiate collectively with platform-based enterprises to clarify the terms of credit data authorization and uniformly manage the distribution of data revenue [13]. It is also possible to establish a multilateral governance organization composed of representatives from the financial institutions, platform enterprises, and small and medium-sized enterprises to participate in the formulation of data rules and the definition of rights/ownership, thereby enhancing the accuracy of financial institutions' judgment on the credit level of small and medium-sized enterprises.

5. Conclusion

This article analyzes the application potential and governance challenges of big data in the process of fintech development for financial institutions in handling credit for small and medium-sized enterprises. Research shows that the comprehensive utilization of multi-dimensional data, such as transaction flow, taxation, and supply chain, can enhance the credit availability of small and medium-sized enterprises and reduce financial risks. However, the existence of problems such as data silos, inconsistent quality, privacy leakage, and unclear ownership is not conducive to further development. Under the collaborative governance model of system and technology, this paper holds that a four-dimensional framework of "regulatory design - standard system - technical tools - subject collaboration" can be constructed. In the governance path, it is necessary to take data classification and grading as the foundation, interpretable technology as the support, and enterprise empowerment as the ultimate goal. It has practical significance for achieving the safe application and value manifestation of data elements in the field of small and medium-sized enterprise credit handled by financial institutions.

In the future, regional data sharing platforms for small and medium-sized enterprises can be established, and the adoption of lightweight technologies can be encouraged to enhance the compliance of credit for small and medium-sized enterprises. At the same time, mechanisms such as data sharing and credit incentives should be explored to enhance the say of small and medium-sized enterprises in the data ecosystem. With the development of new technologies such as generative artificial intelligence, future research can further explore cutting-edge issues like dynamic credit systems and cross-domain data sharing, promoting the construction of a fairer and more sustainable digital financial system.

References

[1] Ansell, C. & Gash, A. (2008) Collaborative Governance in Theory and Practice. Journal of Public Administration Research and Theory, 18, 543-571.

[2] Zhang, Y. & Liu, H. (2023) Federated Learning for SME Credit Scoring: A Privacy-Preserving Approach. Journal of Financial Innovation, 9, 112-130.

[3] Abowd, J. M. & Schmutte, I. M. (2019) An Economic Analysis of Privacy Protection and Statistical Accuracy as Information Goods. American Economic Review, 109, 175-206.

[4] Zheng, Z., Xie, S., Dai, H., Chen, X. & Wang, H. (2017) An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends. 2017 IEEE International Congress on Big Data, 19, 557-564.

[5] Zeng, J. & Church, J. (2020) Data Cooperatives: A New Model for Data Governance. Telecommunications Policy, 44, 101-109.

[6] Arner, D. W., Barberis, J. N. & Buckley, R. P. (2020) The Evolution of Fintech: A New Post-Crisis Paradigm?. University of New South Wales Law Journal, 43, 1-44.

[7] Buchak, G., Matvos, G., Piskorski, T. & Seru, A. (2018) Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks. Journal of Financial Economics, 130, 453-483.

[8] Jagtiani, J. & Lemieux, C. (2019) The Roles of Alternative Data and Machine Learning in Fintech Lending: Evidence From the LendingClub Consumer Platform. Financial Management, 48, 1009-1029.

[9] Fuster, A., Plosser, M., Schnabl, P. & Vickery, J. (2019) The Role of Technology in Mortgage Lending. The Review of Financial Studies, 32, 1854-1899.

[10] Zhou, L., Wang, Y. & Li, Q. (2021) Big Data Analytics in FinTech: A Survey of Methodologies and Applications. Financial Innovation, 7, 1-22.

[11] Lin, M., Prabhala, N. R. & Viswanathan, S. (2023) Judging Borrowers by the Company They Keep: Friendship Networks and Information Diffusion in Online Peer-to-Peer Lending. The Review of Financial Studies, 36, 1-35.

[12] Chen, M. A., Wu, L. & Zhang, X. (2022) Explainable AI in Credit Scoring: A Comparative Study of Interpretability Methods. Decision Support Systems, 153, 113-116.

[13] Begenau, J., Palacios, M. & Silva, A. (2023) Big Tech Lending: The Role of Technology and Data in Credit Markets. Journal of Financial Economics, 147, 307-332.