MINDWELL SCIENCE PRESS SDN. BHD.

Journal of Systemic Risk
Volume 1, Issue 1

RESEARCH ARTICLE

Risk Control in Data Sharing and the Improvement of the Governance System

Zengmao Yang[1], Changfeng Wang[1,*], Hongjun Yang[1]

[1] *School of Economics and Management, Beijing University of Posts and Telecommunications, Beijing 100876, China*

Corresponding author: Changfeng Wang, E-mail: wangcf@bupt.edu.cn

Published: 28 May 2026

Abstract

With the rapid development of big data technology, data have become a new factor of production. Data now empower many industries, and they deliver greater value when shared among different actors. Data sharing, however, currently carries such risk factors as data leakage, uncertain data ownership and poor data quality. In the context of data sharing, data are closely bound up with multiple actors, so the mode of governance has to shift from single-actor governance to multi-actor collaborative governance. To resolve the risks of data sharing, a funnel-style dynamic data openness system should be built around multi-actor collaboration and liberal entry with strict exit; the construction of data infrastructure and the transformation of data technology should be accelerated, and high-level personnel trained; data quality should be raised by building a complete data quality indicator system; and legislation on data property rights should be completed to regulate and guide data sharing.

Keywords: data governance; data sharing; security risk; risk control

1 Literature Review

1.1 Data Governance

Data governance has its origins in data management. Data management refers to the use of computer hardware and software technology for the efficient collection, storage, cleansing and use of data, with the aim of realising as much of the value of data as possible. The business of data management is carried out by data management organisations. The purpose of governance is to arrive at a state of good governance. Governance is not an outcome but a continuous process of overall planning and scientific control, and that process also requires the participation of multiple actors. In a broad sense the actors in data governance include government, enterprises, citizens and other organisations. Data are ubiquitous; multiple actors generate data in the ordinary course of their operation and each of them should be one of the actors in data governance. Collaborative governance of data is also the trend of the future. The object of data governance is enormous quantities of big data, which usually reside in government, in enterprises and on the various network platforms. How to make good use of data has always been one of the problems confronting data governance; viewed across the full life cycle of data, data collection, data cleansing, data storage, data circulation and data use all fall within it.

1.2 Data Sharing and Security Risk

Some scholars have proposed that data quality is the core factor in whether data serve their purpose: although the volume of data today is enormous, it is characterised by high volume and low value, and a considerable quantity of data is of such poor quality that it cannot be used directly.[2] On the one hand, cleansing low-quality data requires a relatively large investment of cost; on the other hand, low-quality data must still be stored on media, which not only wastes resources further but also increases the pressure of data governance on government and enterprises. Other scholars have pointed out that data cleansing is a precondition for data to serve their purpose, and that data must be cleansed and processed after collection before they reach a usable state. Yet other scholars have argued that the core factor in realising the value of data is data sharing: because of the complexity and flexibility of data, one actor's acquisition of data does not presuppose the data owner's loss of it, and once data reach a considerable volume they can generate greater value, so the fusion of multi-source heterogeneous data is one of the goals pursued by data governance. Data sharing nonetheless still faces a number of challenges. The first is uncertain data ownership, which leaves the disposal of data without a clear basis and makes it difficult to safeguard the rights and interests of the parties that genuinely have a stake in the data. The second is data quality,

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

where the absence of an industry-recognised standard for the application of data in industry leaves quality uneven and seriously affects the use that the recipient of data can make of it. The third is the absence of sharing rules: without rules on data sharing it is difficult for the holder of data to assess whether the outflow of the data will create problems — government is concerned about public information security and national security, enterprises worry about the disclosure of their trade secrets and intellectual property, and other social organisations have their own misgivings. The practice of data sharing thus still faces a good many problems, and resolving them and eliminating the risks in the process of sharing is the focus of data governance; both the starting point and the methods of data governance should change accordingly.

2 Theoretical Basis and Analysis of Risk Factors

Stakeholders are regarded as the important group that sustains an organisation: they influence the accomplishment of organisational goals and act upon the whole process of accomplishment. Stakeholders differ in their role and standing within an organisation and are accordingly treated differently. Against this background data are likewise one of the shared vehicles of the interests of multiple stakeholders; they come from multiple actors and multiple channels and can bring returns to many stakeholders. Whether government is pursuing the efficient completion of administrative affairs or an enterprise is pursuing the maximisation of its own interests, both should consider the collaboration and expansion of data governance from multiple sides. Multi-actor collaborative governance of data is not only a requirement of social ethics but a task of overall significance, and it helps to realise the goal of good governance in data governance.

Data governance is a process of multi-actor, multi-stage and multi-mode collaborative improvement centred on data resources, and not merely an outcome. In this continuing dynamic process data sharing is the core link that connects multiple actors and runs through multiple stages, but it also generates certain data security risks. The first is the risk of data leakage. Leakage arising in the course of data sharing from problems of technology, personnel and institutions will place personal privacy, enterprise trade secrets and national security in a relatively dangerous position. The second is the loss of the data rights and interests of the rights holder caused by uncertain data ownership. The third is the risk of data quality, where inconsistent standards and incomplete content render data unusable, and falsified data content leads to errors in the content derived from it.

Drawing on the existing literature, this study divides the stakeholders in data sharing into six actors: government and its functional departments, the public, non-governmental organisations, enterprises, the media and research institutions. Combining this with existing research findings, and classifying stakeholders along the three dimensions of relevance, urgency and influence, it distinguishes core stakeholders, secondary stakeholders and peripheral stakeholders.[3] The resulting stakeholder matrix for data sharing is given in Table 1.

Table 1 Matrix of Stakeholders in Data Sharing

Stakeholder Type	Stakeholder	Relevance	Urgency	Influence
Core stakeholder	Government and functional departments	High	High	High
	The public	High	High	High
Secondary stakeholder	Non-governmental organisations	Low → medium	Medium → high	Low → medium
	Enterprises	Low → medium	Medium → high	Low → medium
Peripheral stakeholder	The media	Low	Low	Low
	Research institutions	Low	Low	Low

2.1 Risk of Data Leakage

Data leakage has always been a relatively serious problem in data governance: it not only occurs frequently but affects multiple stakeholders — government, enterprises, the public and other organisations — and causes them direct or indirect loss. The causes of data leakage are various. The first is system vulnerability. In the field of data management a vulnerability refers specifically to a defect in an enterprise's data management system, usually understood as a weakness in software, hardware or protocols arising from actual operation and management policy. Meta (formerly Facebook) was reported to have suffered three data leakage incidents in 2018, two of which were caused by system vulnerabilities, the content including names, contact details, usage records and private photographs and other sensitive personal information. The second is organisational management: IBM statistics show that more than 25% of data leakage incidents arise from human factors. The third is hacking: 48% of leakage incidents are produced by hacking attacks.

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

2.2 Risk to Data Property Rights

Three elements are required for data resources to become data assets: first, the data resources must have a certain volume and be stored in electronic form; second, there must be substantive control over the data resources, which includes not only data internal to the organisation but also external data acquired by various lawful means; third, the data resources must be capable of generating economic value or public service value, displaying economic attributes and producing economic benefit in the future. China is in the process of the capitalisation of data assets, but theoretical research on that capitalisation is still at an early stage and cannot meet the needs of market practice. The current market situation already fully demonstrates the value that data resources contain as an asset: government and enterprises both hold data resources of enormous volume, but the absence of a clear system of data property rights leaves data ownership uncertain and impedes the conduct of data transactions. A transaction presupposes that buyer and seller hold ownership of and the right to dispose of the object of the transaction; to sell data without property rights at will will produce serious consequences, and much of the illegal data trading on the "dark web" is of this kind.

Data trading in China has already reached a certain scale and is in the process of expansion, and there is an urgent need for official policies and regulations to define data property rights clearly and authoritatively. On the one hand, the data held by government departments are affected by China's vertically compartmentalised, hierarchical administrative system, in which each region or department administers regional affairs according to its departmental rank and the region to which it belongs; the data systems of different regions and departments are consequently for the most part relatively independent, and some regions or departments treat internal data as a symbol of their own authority or as an expression of departmental interest, showing considerable resistance to data trading and openness. On the other hand, constrained by the lack of rules, data trading carries a certain risk, which leads government departments and enterprises to take a resistant attitude towards it.

2.3 Risk of Poor Data Quality

Whether data are shared within government or between government and enterprises or other social organisations, the purpose is to expand the volume of data and enrich its content through circulation, and thereby improve the effect of its use. This raises a situation that cannot be overlooked: if data quality is poor, the use value of the data declines. In data quality management, which covers the entire life cycle of data from planning, collection, storage, circulation, operation and application to obsolescence, a response to and resolution of various data quality problems may be required at every stage. The risk of poor data quality is manifest chiefly in the following respects. First, a low level of quality leads to low usability, non-standardised data lead to a narrow range of uses, and falsified data are wholly unusable. Standardisation of data formats and clarification of data quality levels are therefore urgent. Government data are a case in point: the standards for the data held by local governments are not currently unified, large quantities of structured and semi-structured data are mixed together, and data interfaces are of many different kinds. Second, an important defect in data quality is data falsification: for some purpose the falsifier tampers with or deletes parts of the data content. Building on the above analysis and combining it with stakeholder theory, we bring multiple actors and collaborative governance of data together in a single framework, as shown in Figure 1.

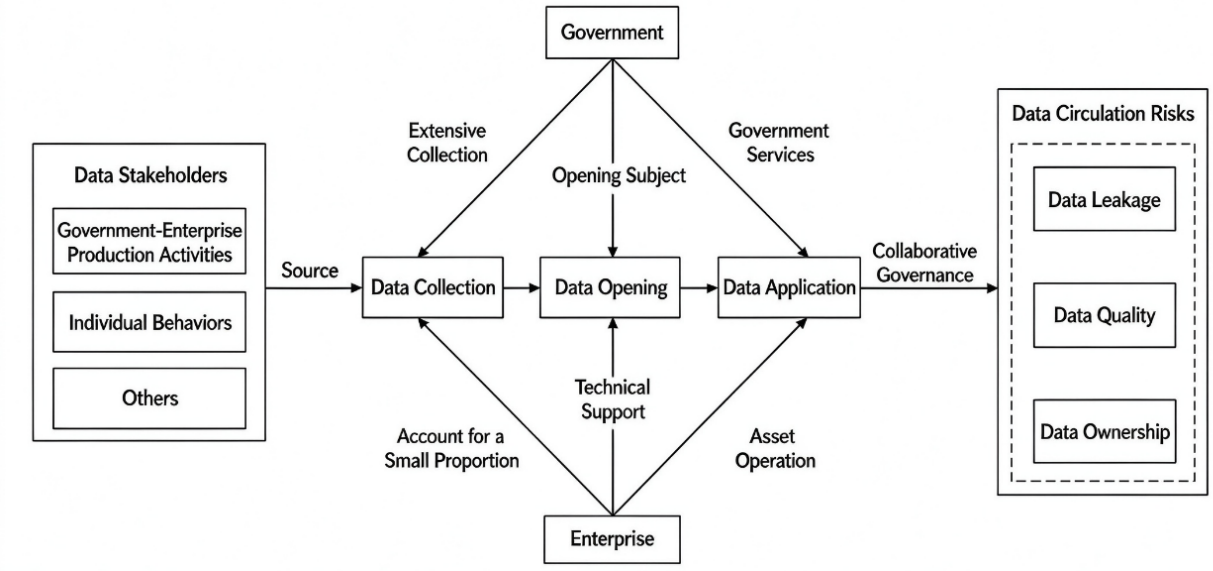


Figure 1 Framework for Collaborative Governance of Data Sharing Risk Based on Stakeholder Theory

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

3 Mechanisms for Resolving the Risks of Data Sharing

3.1 Build a Funnel-Style Dynamic Data Openness System Based on Multi-Actor Collaboration with Liberal Entry and Strict Exit

First, multi-actor collaborative data governance refers to a mode of governance in which multiple stakeholders take part together in the process of governance. Given the contest of interests among multiple actors in data sharing, the objects and objectives of data sharing must be defined and regulated so that every core, peripheral and potential stakeholder can obtain the corresponding benefit and protection according to its relevance, urgency and influence. This is an activity involving multiple actors: government, the public, enterprises and other actors must all be mobilised to take part in the process of governance and to play their proper part. It is, moreover, a dynamic process in which different problems may arise at every stage; every stakeholder should also take part in each stage of governance, exercising its rights and performing its obligations at that stage.

Second, regulate data collection and control data circulation strictly, applying liberal entry and strict exit. Data collection is the precondition for producing data and for making data valuable, but it cannot be unlimited, and the objects and scope of collection should be regulated. The outflow of data should, in addition, be strictly regulated while still giving play to the function of sharing: the entire chain of the source and destination of data should be subject to tracing review and supervision, so that the function of data sharing is realised on the basis that the rights and interests of all stakeholders are protected as far as possible.

3.2 Accelerate the Construction of Data Infrastructure and the Transformation of Data Technology, and Train High-Level Personnel

Data governance technology comprises the sum of the data technologies across the full data life cycle, from collection through processing, cleansing, storage and application to obsolescence. At present government attaches great importance to the introduction and development of data technology in its functional departments and stresses the building of digital government, but the results of the transformation are not yet satisfactory. On the one hand, the government's initial technological strength was weak and was not effectively replenished in the course of the subsequent transformation; on the other hand, although the introduction of technology has gone well, it faces the difficulty that technology and operating personnel do not mesh well — the technical level of operating personnel is relatively low, or officials have a mistaken conception of data technology and feel resistance and rejection towards the changes in work that it brings — and these circumstances all affect the improvement of technological capacity.

First, raising the level of government data governance technology requires policy support. Raising the technical level requires not only the introduction of new technology but also, and just as importantly, an improvement in the level of technical personnel. Once policy and strategic decisions have been made and funding has been secured, the completion of data infrastructure is assured, and the recruitment and training of professional technical personnel can then be carried out. The emerging concepts of big data technology should be combined with the content of traditional personnel training, building in personnel an overall awareness of the understanding, integration, application and security of data resources, so as to deepen the value system of data governance step by step and train more high-level personnel in the field.

Second, improve the construction of data governance infrastructure. Besides completing data infrastructure within its own institutions, government must also build the social public data service facilities that go with it. The new infrastructure that government has been advancing, for example, is new infrastructure construction extending from within departments out into the social public sphere, and such infrastructure can enhance the driving role of data in administrative affairs, public services and economic development. Because of the many new technologies applied, both government and enterprises can effectively save on labour costs and raise efficiency, and once the infrastructure is complete the allocation of resources in data governance can be well optimised and an all-round improvement in data governance technology achieved.

Third, define the strategic positioning of the application of data technology. Once support at the policy and strategic level has been obtained and infrastructure has been built, the application of technology must be strengthened. Access restriction technology should be introduced to prevent leakage in the course of the circulation and use of data. Blockchain technology may also be introduced, its distinctive properties of being "tamper-proof" and "collectively maintained" serving to safeguard data security.

3.3 Improve Data Quality and Build a Complete Data Quality Indicator System

Improving data quality starts at the source. The source of data is the stage of data collection: standards of collection that are relatively well regulated should be formulated at that stage and then promoted so that they can apply across a relatively wide range of fields and industries and gradually become standards common to all industries. Collection personnel must be

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

given definite industry standards to guide their work, and the form that collected data should take must be laid down explicitly in the quality standards. Next, falsification of data must be constrained and supervised. On the one hand, traces of modification in data should be detected and the capacity to discriminate between authentic and false data raised; on the other hand, those who falsify data should be held to account, and anyone who uses false data for gain and causes economic or other loss to others should, in accordance with law, be held responsible. Third, the competence of collection personnel should be raised: on the basis of clear standards, their understanding of data and their capacity to collect it should be improved. Finally, public understanding of data collection should be raised. Whether the data are governmental or corporate, a considerable part consists of private information and information about everyday behaviour collected from the public and then rendered symbolic and formatted; public understanding of the provenance of data must therefore be raised, so as to improve data quality more effectively.

3.4 Complete Legislation on Data Property Rights and Regulate and Guide Data Sharing

Resolving the question of the ownership of data resources is the core content of a system of data property rights. This article divides data by actor into public data, enterprise data and personal data, and different types of rights confirmation are required for different data.

First, confirming rights over public data resources. Public data are chiefly lawfully generated, collected and integrated by government agencies and are held by them; they include both data produced by government departments themselves and data relating to the economy, social management and social public services. At present, both in China and internationally, the tendency is to vest ownership of such data in the collector, with the state holding the right of disposal, and the part originating from individuals requiring that the source be given a corresponding right to be informed about its application and a right to claim relief where privacy is infringed.

Second, confirming rights over enterprise data resources. Enterprise data refer chiefly to the data an enterprise obtains in the course of production and operation: part is internal data on the enterprise's own products, personnel and organisational activities, and part is user data collected by the enterprise, the latter accounting for a relatively large proportion. The internal data held by an enterprise belong to the enterprise and it may dispose of them as it wishes; as for the user data it has collected, once the data have undergone a series of cleansing processes such as de-identification the enterprise should dispose of them, but data collected that are linked to the user's identity and behavioural attributes should belong to the individual user, and the enterprise may not dispose of them at will.

Third, confirming rights over personal data. Legislation protecting personal data should be completed as soon as possible, clearly providing for and protecting the individual's right to be informed about the movements of his or her own data and right to claim relief, so as to ensure that the lawful rights and interests of the individual are not infringed and that the individual can obtain the returns that are his or her due from data over which he or she holds ownership. With the promulgation and entry into force of the Personal Information Protection Law of the People's Republic of China, this has to some extent already been put into effect.

References

[1] Hao WQ. Stakeholders in government data openness: definition, classification and management strategies[J]. Modern Information, 2021, 41(7): 137-145. [in Chinese]

[2] Huang QQ, Zhao Z, Liu ZY. Research on a comprehensive data quality management system and technical framework in the scenario of data circulation and trading[J]. Data Analysis and Knowledge Discovery, 2022, 6(1): 22-34. [in Chinese]

[3] Li F, Zhen L, Huang HQ, et al. Changes in land use functions and stakeholders' willingness to accept compensation and economic compensation: a case study of the ecologically fragile area of Poyang Lake[J]. Resources Science, 2009, 31(4): 580-589. [in Chinese]

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.