Large Model-Empowered Trusted Data Spaces: Data Security Governance and Trust Mechanism Construction (Postprint)

Authors: Pei Lei, Chen Xiaoyu

Date: 2025-06-24T00:00:00+00:00

Abstract

[Purpose/Significance] This study explores the key constituent elements and implementation pathways of trusted data spaces, with a focused analysis on the role and impact of large model technology in data security governance and the construction of trust mechanism frameworks. [Method/Process] Adopting a methodology that combines theoretical analysis with practical case studies, this paper elaborates on three dimensions: data security governance, large model technology pathways, and trust mechanism construction. [Results/Conclusion] The construction of trusted data spaces necessitates deep integration among institutions, technology, and multi-stakeholder collaboration. Large models provide robust support for data security governance, demonstrating significant advantages particularly in real-time monitoring, anomaly detection, intelligent decision-making, and automated compliance management. The refinement of data security governance relies on the dynamic optimization of trust assessment systems. Through classified and graded protection, cross-entity collaboration mechanisms, and explainability enhancement, the transparency and controllability of trusted data circulation can be improved. The development of trusted data spaces is not merely a process of technological innovation, but also a transformation of the data governance paradigm. Its implementation pathway must balance compliance, interoperability, and alignment with international standards to promote the efficient allocation of data element markets.

Full Text

Large Models Empowering Trusted Data Spaces: Data Security Governance and Trust Mechanism Construction

Pei Lei[1], Chen Xiaoyu[2]
[1]Data Intelligence and Interdisciplinary Innovation Laboratory, Nanjing Uni-

versity, Nanjing 210023, China
[2]School of Cultural Heritage and Information Management, Shanghai University, Shanghai 200444, China

Abstract:
[Purpose/Significance] This article explores the key components and implementation pathways of trusted data spaces, with a particular focus on the role and impact of large language model (LLM) technologies in data security governance and the construction of trust frameworks. [Method/Process] A combination of theoretical analysis and practical case studies is adopted to examine three core areas: data security governance, technical pathways of LLMs, and the establishment of trust mechanisms. [Result/Conclusion] The development of trusted data spaces requires deep integration across institutional frameworks, technological infrastructures, and multi-stakeholder collaboration. LLMs offer powerful support for data security governance, particularly excelling in real-time monitoring, anomaly detection, intelligent decision-making, and automated compliance management. The improvement of data security governance hinges on the dynamic optimization of trust evaluation systems, which can enhance the transparency and controllability of trusted data circulation through tiered protection strategies, cross-actor collaboration mechanisms, and enhanced explainability. Ultimately, building trusted data spaces is not only a matter of technological innovation but also a paradigm shift in data governance. Its implementation must align with regulatory compliance, interoperability, and international standards to facilitate the efficient allocation of data as a production factor.

Keywords: Trusted data space; Large language model (LLM); Data security governance; Trust mechanism; Data governance framework

Data security governance and trust mechanism construction have become central issues in global digital development. As a key production factor in the era of data intelligence, the trustworthy circulation and effective governance of data are not only critical for driving high-quality development of the digital economy but also constitute important support for enhancing national competitiveness. The Third Plenary Session of the 20th CPC Central Committee explicitly proposed “accelerating the construction of a unified national market”and “improving the market-oriented allocation system for data elements,” providing policy guidance for the standardized development of China’ s data factor market while imposing higher requirements on the data security governance system. Against this backdrop, how to ensure data security while building a cross-industry, cross-domain, multi-stakeholder collaborative trusted data circulation mechanism has become a core issue in current data governance.

With the rapid advancement of LLM technology, the paradigm of data circulation and governance is undergoing profound transformation. LLMs, with their powerful data processing capabilities, knowledge reasoning abilities, and intelligent decision-making capacities, provide new technical support for data security governance and trust mechanism construction in complex information environments. Particularly in areas such as trusted data sharing, privacy-preserving

computation, and cross-domain collaboration, LLMs can enhance data usage transparency and reduce security risks and compliance risks in data circulation. However, data security is the foundation of trusted data spaces. The opacity of technology, the compliance of data aggregation, and the complexity of data ownership remain critical issues that LLM technology must address in data governance practice. It is essential to ensure the secure and trustworthy circulation of data elements while fully leveraging the intelligent advantages of LLMs.

The *Trusted Data Space Development Action Plan (2024-2028)* (hereinafter referred to as the "Action Plan") proposes that China will build over 100 trusted data spaces by 2028, forming an extensive, interconnected, and value-co-creating data circulation ecosystem. This policy goal not only emphasizes the importance of market-oriented allocation of data elements but also imposes higher requirements on the technical architecture and governance system of trusted data spaces. The construction of trusted data spaces is not merely a technical exploration but also an innovation in institutional and governance models. Integrating technical pathways with governance frameworks to build an efficient, secure, and transparent trust mechanism is a crucial task for promoting the healthy development of the data factor market. However, existing policy frameworks still lack targeted solutions for security and trust mechanisms in large-scale data circulation and cross-industry data sharing.

This article discusses the construction pathways of trusted data spaces, focusing on analyzing the role of LLMs in data security governance and trust mechanism construction, and systematically exploring the institutional design and technical architecture for trusted data sharing. It first analyzes the requirements for data security governance in trusted data space construction; then examines how LLMs can function in privacy protection, data governance, and cross-domain collaboration; and finally constructs a relevant trust mechanism framework, proposes future development pathways for trusted data spaces, and explores their role in the data factor market and high-quality digital economic development.

1.1 Connotation, Characteristics, and Current Status of Trusted Data Spaces

A trusted data space is a data governance ecosystem based on consensus mechanisms and multi-stakeholder collaborative participation. Its key characteristic is ensuring the security, trustworthiness, and controllability of data during circulation, sharing, and transactions through technical means and institutional design. Essentially, a trusted data space is not only a technical architecture but also an infrastructure for the market-oriented allocation of data elements, with its core mission being to resolve issues of "data sovereignty," "data security," and "trust mechanisms" in data circulation.

The construction of trusted data spaces has multiple objectives. First, it emphasizes data sovereignty and autonomy, granting data providers autonomous control over data usage processes through technology and institutions to ensure

that ownership and privacy are not violated during sharing and circulation. Second, trusted data spaces focus on multi-party collaboration and interoperability, breaking down "data silos" through unified technical standards and interface specifications to promote efficient circulation and utilization of data resources across different platforms and systems. Third, security and compliance are fundamental requirements, with cutting-edge technologies such as blockchain, digital signatures, and smart contracts ensuring data transmission integrity, immutability, and traceability while guaranteeing compliance with national and industry laws and regulations. Fourth, in terms of economic efficiency, trusted data spaces significantly enhance data factor transaction efficiency and unlock the potential value of the data economy through reasonable incentive mechanisms and optimized circulation pathways.

Currently, domestic and international trusted data space construction has achieved initial progress and is becoming a critical pillar for global digital economic development. Internationally, Europe has conducted systematic explorations earlier, relying on organizations such as the International Data Spaces Association (IDSA) to build a representative general architecture and standard system for trusted data spaces, such as the IDSA 4.0 reference architecture model, and has promoted multiple demonstration projects in automotive manufacturing (e.g., Catena-X data space) and supply chain management [e.g., Smart Connected Supplier Network (SCSN)]. In contrast, the United States primarily drives cross-enterprise data sharing and intelligent analysis through cloud service providers like Google Cloud, with relatively mature market mechanisms and strong platform capabilities, though it still faces challenges in cross-industry interoperability and data security assurance. In Asia, Japan and South Korea have also shown active momentum. Japan released the Connected Industries Open Framework (CIOF), emphasizing the standardization and collaborative governance of cross-actor trusted circulation of industrial data. South Korea, based on MyData personal data services, explores the construction of personalized data spaces oriented toward data sovereignty.

China' s trusted data space construction is also advancing rapidly. At the policy level, the promulgation of the *Data Security Law of the People' s Republic of China* and the *Personal Information Protection Law of the People' s Republic of China* has provided solid legal guarantees for data circulation and governance. In practice, based on technologies such as blockchain and privacy-preserving computation, China is actively exploring data security frameworks and gradually achieving full lifecycle management of trusted data circulation. Shanghai, Shenzhen, and other cities have begun piloting industry-specific data spaces covering key sectors such as finance, healthcare, and industry, providing innovative application scenarios for data circulation and value mining. These practices have laid an important foundation for building a unified data governance model, with major cases and current status summarized in .

Despite progress, several pressing issues remain in the development of trusted

data spaces. First, global data governance frameworks have not yet achieved coordination and unification, with divergences among countries and regions in data protection standards, technical specifications, and compliance requirements creating obstacles for cross-border data circulation and global promotion of trusted data spaces. Second, technical practices within data spaces require optimization, as existing technologies have not fully resolved the balance between performance, reliability, and scalability when addressing large-scale, multi-scenario, multi-actor data collaboration. Third, insufficient awareness and practical capabilities in data governance in some industries have created resistance to widespread application of trusted data spaces. Finally, trusted data space construction must address the complexity of multi-stakeholder interest coordination, as trust mechanisms and collaboration models among data providers, users, and intermediary service providers need further improvement, with these interactions often profoundly influenced by different governance structures, legal environments, and business models.

1.2 Key Challenges in Data Security Governance

Data security is a cornerstone of digital economic development and a core guarantee for the market-oriented allocation of data elements. Amid the global wave of digital transformation, the secure and trustworthy circulation of data elements not only concerns the effective integration and utilization of resources but also relates to the stability and sustainability of socio-economic operations. In recent years, with the rapid growth of data scale and the increasing complexity of usage scenarios, the connotation of data security governance has continuously expanded, and its importance has far exceeded traditional technical protection, gradually becoming a critical factor affecting national governance, economic development, and social trust.

Data security directly impacts the trust foundation for multi-stakeholder collaboration. The essence of data circulation lies in data sharing and collaborative utilization among multiple parties, with trust mechanisms being key to achieving efficient circulation. However, frequent issues such as data breaches, privacy violations, and illegal misuse have created a trust deficit. Data providers worry about privacy leaks and misuse, while data users face difficulties in ensuring data quality and legality. This trust deficit not only weakens the liquidity of data transactions but also increases the cost and risk of data usage, thereby hindering the release of data value in economic activities. In fields involving sensitive data, security issues trigger particularly concerning chain reactions. In finance, data breaches can lead not only to customer privacy exposure but also to large-scale financial fraud, threatening financial system stability. In healthcare, the misuse or leakage of patient health data not only violates personal privacy but also undermines public trust in medical services, affecting willingness to share data and the realization of research value. In government affairs, government data serves as a crucial resource for national governance, and security incidents causing data loss or improper dissemination can affect public service quality,

weaken public trust in government institutions, and even trigger social stability issues. At a deeper level, trust deficits often create a "data silo" effect. Lacking unified security guarantees and trust mechanisms, data stakeholders tend to protect their own data assets rather than share them, limiting further data value excavation and causing waste and redundancy of data resources. This fragmented state hinders cross-industry, cross-regional collaborative innovation, making the circulation efficiency and economic benefits of data elements far lower than expected.

The complexity and uncertainty of the international data governance landscape have significantly exacerbated the urgency of data security issues, with cross-border data circulation becoming a core element of economic and technological cooperation. International disagreements on data sovereignty, privacy protection, and compliance requirements have created a fragmented state of global data governance. This situation not only increases potential data security risks but also raises the compliance costs and operational difficulties of cross-border data circulation. Differences in interpreting data sovereignty among countries constitute a primary reason for fragmented data circulation rules. For instance, the EU' s General Data Protection Regulation (GDPR) establishes strict data protection and cross-border transmission standards, emphasizing privacy protection as a core principle, though its stringent requirements limit its applicability in other regions. In contrast, the United States tends to adopt a market-oriented approach to data circulation with relatively relaxed privacy protection regulation. China, meanwhile, has built a legal framework centered on the national security attributes of data based on the *Data Security Law* and *Personal Information Protection Law*. This rule system, constructed based on different national interests and governance philosophies, lacks unified standards, leading to differentiated security norms and technical compliance requirements and making legal reviews of cross-border circulation more complex.

Insufficient technical interoperability and the absence of a standardized system are also major bottlenecks in current cross-border data circulation. The diversity of regulatory requirements across countries and regions leads to different technical implementation paths for data encryption, anonymization, and access control mechanisms. Incompatibilities in implementation details further hinder efficient data circulation and create hidden security risks. Particularly between different technical systems, fragmented standards can create interface vulnerabilities that become potential entry points for malicious attacks, further amplifying security risks in cross-border data circulation.

The core of addressing data security issues lies in building a sustainable and trustworthy security ecosystem. A security ecosystem is not only about protecting data resources themselves but also a critical pathway for releasing data element value and enhancing socio-economic effectiveness. Under the security ecosystem framework, data resource circulation is no longer limited to privacy protection and risk prevention but provides a solid trust foundation for multi-stakeholder collaboration and value co-creation. Only under secure premises can

large-scale trusted data circulation be realized, promoting efficient integration of data elements in socio-economic systems, fostering collaborative innovation, and achieving dynamic balance. This security ecosystem requires building a dynamic protection system covering the entire data lifecycle, including legality review during data collection, encryption and protection during storage, path control during transmission, and real-time monitoring and risk response during usage. Simultaneously, security ecosystem construction must be grounded in multi-stakeholder collaboration, promoting coordinated cooperation among enterprises, governments, and social organizations through improved institutional guarantee mechanisms and transparent governance models to resolve trust barriers and enhance data sharing efficiency and governance effectiveness. More importantly, establishing a security ecosystem should not focus solely on current technical issues but must adopt a long-term perspective, continuously unlocking data resource potential in a cost-effective manner to inject strong momentum into the sustainable development of the digital economy.

2.1 Advantages and Challenges of LLM-Driven Data Security Governance

As a significant breakthrough in artificial intelligence technology, LLMs are becoming an important support force for data security governance, demonstrating notable advantages in data analysis, anomaly detection, and risk prediction. Their powerful computational capabilities and deep learning models enable them to efficiently address security issues in complex data environments, not only enhancing data security monitoring and response capabilities but also playing a key role in promoting the marketization of data elements and trusted data circulation.

The advantages of LLMs in data security governance are mainly reflected in three aspects. First, in dynamic security monitoring and anomaly detection, LLMs can analyze massive datasets through deep learning algorithms to accurately identify potential security threats. For example, in cybersecurity, LLMs can detect abnormal access, malicious attacks, and data leakage risks by analyzing data traffic in real time and generate automated alerts to improve security management response efficiency. Second, the self-learning capability of LLMs enables continuous optimization of security strategies and dynamic adaptation to new threat environments. Through deep training on attack patterns and data circulation rules, LLMs can continuously improve detection accuracy, demonstrating stronger adaptability in dynamic risk environments. Third, LLMs also show positive effects in data protection and privacy-preserving computation, such as through differential privacy and federated learning methods, achieving controllable data sharing while ensuring security, thereby laying a technical foundation for trusted data space construction.

However, the widespread application of LLMs may also introduce new security challenges. Their training process relies on large volumes of diverse data, often involving personal privacy, commercial secrets, and even national strategic

information. Under such extensive data aggregation, without rigorous security management measures, they may not only become primary targets for attackers but also trigger large-scale information leaks due to weak links in the data chain, causing immeasurable consequences. In cross-industry, cross-domain data usage scenarios, these risks are further amplified, imposing higher technical and management requirements on data security governance.

The algorithmic complexity and "black box" characteristics of LLMs increase the complexity of data security regulation. Due to the difficulty in explaining their reasoning processes, data users and regulatory agencies struggle to determine whether their decision logic complies with requirements, thereby affecting trust in model predictions. In scenarios with high privacy protection requirements, such as healthcare, finance, or cross-border data circulation, algorithmic opacity may trigger data compliance disputes and even affect the legality of cross-border data sharing. Additionally, LLMs require continuous iteration and optimization, yet vulnerabilities remain in the legality, trustworthiness, and source control of new data. How to ensure compliance in data usage during model optimization has become a key challenge for data regulation and governance.

In distributed computing environments, data security governance becomes even more challenging. While multi-node collaboration models improve computational efficiency, they also increase security risks during data transmission and processing. Permission management, vulnerability remediation, and access control for computing nodes become issues requiring resolution, as security failures in any single node may affect entire system stability through chain reactions and even trigger systemic security incidents. In application scenarios involving cross-border data circulation and multi-industry collaboration, data needs to flow under different legal and technical standards, with countries having varying standards for data privacy protection and cybersecurity compliance, further complicating data governance. Balancing the security and legality of computing nodes, compliance requirements for cross-border circulation, and the computational efficiency and response speed of LLMs presents a multi-objective optimization problem that imposes higher requirements on the technical design and governance system of trusted data spaces.

2.2 Privacy Protection, Data Governance, and Algorithm Transparency

The widespread application of LLM technology in data security governance has made privacy protection, data governance, and algorithm transparency core elements of trusted data space construction. These three components collectively form the trust foundation in data circulation processes, not only affecting the effectiveness of data security and privacy protection mechanisms but also directly relating to the stability and compliance of multi-stakeholder collaboration.

In trusted data spaces, privacy protection is the core task for ensuring secure circulation of data elements, particularly in scenarios of cross-actor collaboration

and cross-border data circulation where the need to protect sensitive information is especially prominent. LLMs rely on large volumes of data during training and inference, often involving personal privacy, commercial secrets, and even national security information. Centralized storage and computation patterns make them vulnerable targets, with increased risks of information leakage. To mitigate privacy risks, technical means such as differential privacy and federated learning are widely adopted to achieve trustworthy computing where data is "usable but invisible."

Differential privacy technology ensures the non-identifiability of individual data points by introducing random noise during data analysis, thereby protecting individual privacy without affecting the effectiveness of overall data analysis. This method has been widely used in personalized recommendations, statistical analysis, and public data sharing. In contrast, federated learning technology employs a decentralized distributed training approach, keeping data stored locally and transmitting only encrypted model parameters to avoid risks of centralized data exposure. In cross-institutional data collaboration scenarios such as healthcare, finance, and intelligent manufacturing, federated learning enables collaborative computation among different actors while ensuring privacy protection compliance. These technical means provide solid support for privacy protection in trusted data spaces, maximizing data value while guaranteeing security.

Data governance plays a dual role of rule-setting and security management in trusted data spaces, serving as both an important mechanism for ensuring data compliance and a key link for optimizing data resource circulation efficiency. Data governance in trusted data spaces mainly involves data storage, permission management, access control, and compliance supervision to ensure legal usage and effective circulation of data.

In terms of data permission management and access control, role-based access control (RBAC) and attribute-based access control (ABAC) have become mainstream governance models. RBAC ensures data access behavior complies with established rules by defining different user roles and their permissions, while ABAC achieves more refined permission management by incorporating factors such as user behavior and environmental conditions. These mechanisms not only improve data access security but also establish dynamic authorization mechanisms among different actors, enhancing the flexibility and adaptability of trusted data spaces.

Furthermore, data governance involves cross-border data compliance issues. Due to significant differences among countries in data sovereignty, privacy protection, and data circulation rules, trusted data spaces need to provide compliance guarantees at the technical level, such as reviewing the legality of data sources, encrypting cross-border data transmission, and monitoring data usage purposes based on smart contracts. For example, under the constraints of the EU' s GDPR, enterprises must ensure user data complies with the "data minimization" principle during cross-border circulation, while China' s *Data Security Law* emphasizes tiered data classification and protection,

setting stricter review standards for the outbound transfer of specific data types. Therefore, the governance system of trusted data spaces must adapt to multi-level data security and compliance requirements to ensure secure circulation under different legal systems.

The "black box" characteristic of LLM reasoning processes poses a significant challenge to algorithmic transparency. Traditional deep learning models often struggle to explain their decision logic, causing doubts about their trustworthiness among users, regulatory agencies, and the public. In data governance scenarios, this lack of explainability may trigger issues such as data misuse, bias amplification, and opaque decision-making, particularly in highly sensitive fields involving public decision-making, medical diagnosis, and financial risk control, where the absence of algorithmic transparency may affect the credibility of data spaces.

Explainable artificial intelligence (XAI) has become an important solution for enhancing algorithmic transparency. XAI makes model reasoning logic clearer through methods such as feature importance analysis, counterfactual reasoning, and model visualization, helping users understand decision processes. For example, in financial risk control, XAI can be used to explain credit scoring criteria, enabling data users to understand the rationality of loan rejection decisions and ensuring algorithmic fairness and compliance. Additionally, algorithm auditing and third-party evaluation are important means to enhance transparency in trusted data spaces. By introducing independent agencies for algorithm review, model bias can be effectively reduced, algorithmic credibility improved, and trust among all actors in data circulation processes enhanced.

In technical practice, the synergistic effect of privacy protection, data governance, and algorithmic transparency is mainly reflected in ensuring data sharing in trusted environments, reducing data leakage risks, achieving effective allocation of data resources through tiered management, access control, and compliance supervision, and enhancing the rationality, transparency, and traceability of data decisions through explainability methods and auditing systems. Building trusted data spaces requires the close integration of these three technical dimensions to optimize data value release while ensuring data security, making the data governance system more stable and sustainable.

2.3 Security Strategies for Cross-Border Data Circulation and Distributed Computing

With increasing international data collaboration and deep development of distributed computing, data circulation and processing have become issues that extend beyond the technical level to involve complex differences in laws and regulations, inconsistent technical standards, and lack of trust mechanisms in multi-stakeholder collaboration. Although the Action Plan proposes building a unified national framework for data circulation and governance, detailed implementation measures are still lacking. The application of cross-border data

circulation and distributed computing in trusted data spaces requires not only innovative technical pathways but also the organic integration of institutional design and collaborative governance.

The core challenge of cross-border data circulation lies in achieving a balance between security and compliance across multiple legal systems. Against the backdrop of incomplete international rule unification, countries have significantly different requirements for data sovereignty and privacy protection, requiring cross-border data circulation to simultaneously meet diverse regulatory requirements. Taking the EU' s GDPR as an example, its strict restrictions on data outbound transfer require detailed compliance explanations, while China' s *Data Security Law* places greater emphasis on the national security attributes of data circulation. In this highly fragmented legal environment, blockchain technology and zero-knowledge proof (ZKP) provide innovative solutions for cross-border data circulation. The distributed ledger characteristics of blockchain ensure transparency and security in data circulation processes through immutable records, while combined with smart contracts to achieve automated permission control and compliance verification. ZKP can complete legality verification without revealing data itself, providing technical guarantees for cross-border use of sensitive data. The combination of these technical means not only lays a technical foundation for international data circulation but also provides possibilities for interoperability between different legal systems.

However, the application of technical means still requires policy and institutional guarantees. Cross-border governance of trusted data spaces needs dual support from trust evaluation systems and smart contract mechanisms to achieve trustworthy data circulation among multiple actors. Establishing multi-dimensional trust evaluation models based on reputation scores, data usage records, and compliance reviews ensures transparency and security in cross-border data circulation. Data providers and users must undergo regular compliance evaluations by independent agencies, with access permissions adjusted based on historical data circulation records to enhance trust mechanisms in international data sharing. Smart contracts can automatically execute cross-actor data circulation rules under different legal systems, ensuring that data usage scope, access permissions, and purpose supervision meet compliance requirements. In international supply chain data sharing scenarios, smart contracts can preset access permissions, allowing only enterprises meeting specific conditions to access critical data, with dynamic adjustments during circulation to ensure legal data usage.

Data security issues in distributed computing environments focus more on the complexity of technical implementation. In distributed training or inference of LLMs, data needs to be transmitted and processed among multiple computing nodes. While this multi-node collaboration model improves computational efficiency, it also increases security risks. On one hand, single-point failures in distributed computing can lead to system-wide vulnerabilities; on the other hand, participation by untrusted nodes may trigger data leakage or tampering with computation results. To address these issues, a more secure distributed com-

puting environment can be built through fine-grained permission control and dynamic risk monitoring technologies. Fine-grained permission control strictly limits the operational scope of each computing node by finely dividing access permissions. Dynamic risk monitoring uses real-time analysis to quickly identify abnormal behavior and intervene in potential risks promptly. Additionally, the introduction of distributed consensus algorithms can improve fault tolerance, ensuring system stability and data integrity even when some nodes fail or exhibit malicious behavior.

In cross-border data circulation and distributed computing environments, achieving data "usable but invisible" is the core objective for ensuring data security and privacy. This concept requires that while data value is fully utilized, the original data content is not leaked or misused. LLMs play an important role in this process, particularly in real-time data monitoring and abnormal behavior identification, where their powerful analytical and reasoning capabilities can dynamically warn of anomalies in data circulation and quickly locate the source and responsible parties of security incidents through traceability mechanisms. Additionally, LLM-based multi-party data permission control and dynamic authentication technology can ensure the legality and transparency of data usage through refined management and real-time adjustment of data access permissions. The technical pathway is illustrated in [Figure 1: see original paper].

[Figure 1: see original paper] Technical Pathway for Cross-Border Data Circulation and Distributed Computing in Data Security Governance

3 Trust Mechanism Construction for Trusted Data Spaces

The effective operation of trusted data spaces depends on the construction of comprehensive trust mechanisms that build transparent, secure, and traceable data circulation systems through institutional guarantees, technical support, and multi-stakeholder collaboration. In the circulation process, trust involves not only secure interaction between data providers and users but also collaborative governance among government regulatory agencies, industry alliances, and technology platforms. How to build a robust trust framework at the institutional, technical, and collaborative levels is key to achieving trusted data circulation.

The foundation of trust first stems from the normative constraints of laws, regulations, and policy frameworks. The cross-actor circulation and transaction of data elements require legal definitions of data ownership, circulation boundaries, and compliance requirements. For example, the *Data Security Law* and *Personal Information Protection Law* of China propose explicit requirements for tiered data classification, cross-border circulation, and usage supervision, ensuring legal and compliant data flow. Additionally, regulations such as the EU' s GDPR provide different compliance pathways for cross-border data circulation. Against this backdrop, the trust mechanisms of trusted data spaces

need to adapt to multi-level data governance rules and establish cross-regional, cross-industry standardized certification systems to ensure secure interoperability across different legal systems.

Technical trust serves as the core support for trusted data space construction, with LLMs playing an important role. LLM-based dynamic access control and intelligent authentication technologies enable data access permissions to be dynamically adjusted based on user behavior, historical records, and risk assessments, improving the transparency and security of data usage. In data security monitoring and anomaly detection, LLMs can accurately identify abnormal access and potential attacks by analyzing data flows in real time, and combine traceability mechanisms to quickly locate responsible parties. Additionally, the introduction of explainable artificial intelligence technology helps enhance transparency in data processing and sharing, enabling data providers, users, and regulatory agencies to understand model decision logic and reducing trust barriers caused by algorithmic "black box" effects. Smart contract technology provides automated execution guarantees for data transactions, ensuring security and immutability in data circulation through preset condition-triggered mechanisms while reducing manual intervention and improving data governance efficiency.

Trusted data space trust mechanisms need to accommodate multi-stakeholder collaborative governance, building dynamic trust systems based on rule constraints, behavior supervision, and incentive mechanisms. Governments play a role in regulation and standard-setting in the data governance system, while industry alliances and data trading platforms are responsible for implementing specific rules. To enhance trust operability, multi-stakeholder data circulation trust evaluation systems can be established, introducing indicators such as reputation scores, data circulation records, and compliance audits to conduct credit ratings for data users and adjust data access permissions based on trust levels. With smart contract support, this trust evaluation can be embedded into data transaction rules to achieve automatic execution of data sharing agreements and improve credibility in cross-actor collaboration. Additionally, regulatory agencies can build transparent and traceable regulatory chains using blockchain technology to conduct real-time reviews of data usage and ensure compliance in data circulation.

The governance framework for trusted data spaces requires integrated advancement of institutions, technology, and multi-stakeholder collaboration. At the institutional level, data classification, compliance supervision, and standardized certification systems constitute the foundational constraints of trust mechanisms. At the technical level, LLMs, privacy-preserving computation, and smart contracts provide core support for data security and circulation. At the collaborative governance level, governments, industry institutions, and data platforms jointly participate in building dynamic trust evaluation and incentive mechanisms to ensure stable operation of data spaces. The establishment of this trust mechanism will provide more robust support for the market-oriented allocation of data elements and create a transparent, secure, and efficient operating en-

vironment for data sharing and transactions. The trust mechanism framework and corresponding levels for trusted data spaces are illustrated in [Figure 2: see original paper].

[Figure 2: see original paper] Trust Mechanism Framework for Trusted Data Spaces

The construction of trusted data spaces is crucial for the market-oriented allocation of data elements and high-quality digital economic development. Amid the deep interplay of technology, institutions, and multi-stakeholder collaboration, data security governance and trust mechanism construction have become important topics in this field. Based on the Action Plan policy framework, this article proposes an LLM-centered data security governance system and trust mechanism framework, addressing current policy deficiencies in technical adaptability, privacy protection, and trusted data circulation. In the future, as policy implementation deepens, the intelligent contract supervision and secure, controllable data circulation models proposed in this article are expected to provide technical and governance support for the Action Plan' s implementation.

Currently, LLM technology provides strong technical support for building trusted data spaces, particularly in real-time monitoring, permission management, and smart contract execution, laying the foundation for achieving data "usable but invisible." Meanwhile, data security governance plays a core role in tiered protection, dynamic risk management, and rule system design, enhancing the security of data circulation and the efficiency of multi-stakeholder collaboration. However, the lack of unified global data governance rules, missing technical standards, and compliance challenges in cross-border circulation still constrain the widespread promotion and application of trusted data spaces. Solving these problems requires comprehensive responses in technological innovation, legal system improvement, and multi-party cooperation models.

Countries have adopted different development paths for trusted data space construction. In comparison, China' s trusted data space construction exhibits clear policy guidance characteristics. In recent years, a legal system centered on the *Data Security Law* and *Personal Information Protection Law* has been gradually improved, with local governments piloting trusted data space construction to form multi-industry, multi-level data governance models. China' s approach emphasizes a comprehensive governance framework of "data classification + technological innovation + government supervision," building cross-departmental, cross-regional data governance systems under government leadership, such as the Shanghai Data Exchange and Shenzhen Data Factor Circulation Pilot. Through government-enterprise collaboration, China is exploring new data transaction mechanisms of "data ownership confirmation—compliant circulation—value realization" to enhance the liquidity and controllability of the data factor market.

This article begins with the connotation of trusted data spaces, focusing on core issues in data security governance and trust mechanism construction, and

systematically analyzes key requirements and technical challenges in current construction processes. Based on this analysis, it proposes an intelligent technology pathway represented by LLMs, exploring their adaptation mechanisms in privacy protection, real-time monitoring, and smart contract execution, and further constructs a multi-dimensional framework covering institutional trust, technical trust, and relational trust, aiming to provide systematic solutions and theoretical support for the sustainable development of trusted data spaces.

Looking forward, trusted data space construction needs to continuously evolve through the deep integration of technical systems and institutional norms, particularly by accelerating the coordination and standardization of international rules to gradually establish a data circulation architecture compatible with multi-national legal systems and industry standards. The next step should involve strengthening technological innovation capabilities while building multi-stakeholder collaborative governance mechanisms to promote trusted data spaces from pilot implementation to systematic rollout and standardized promotion. As the global data ecosystem continues to evolve, trusted data spaces are expected to become important infrastructure supporting high-quality digital economic development and global data collaboration.

References

[1] Fu Shaoxiong, Sun Jianjun. Data Circulation and Security: Standards and Guarantee Systems[J]. Library and Information, 2023(4): 20-28. [2] Wang Yan, Yang Da. Transformation of Chinese Management Accounting System: From Data Elements to Data Assets[J]. Management World, 2024, 40(10): [3] Adhere to Promoting Market-oriented Allocation Reform of Data Elements—National Data Bureau Introduces Progress and Achievements in Data Field Reform[EB/OL]. [2025-04-22]. https://www.gov.cn/lianbo/bumen/202407/content_{6964034}.htm. [4] Liu Jiangfeng, Zhang Ran, Zhang Jundong, et al. Empowering Ideological History Computing Research with Generative AI: Model Construction and Application Exploration[J]. Library Journal, 2025, 44(3): 113-127. [5] Tang Xinglong, Zhang Yu, Zeng Wen. Research on the Construction of Technical Foundation for Scientific and Technological Intelligence in Complex Information Environments[J]. Journal of the China Society for Scientific and Technical Information, 2024, 43(7): 761-772. [6] Zhang Dun. Analysis of the Connotation and Framework Construction of Data Resource Holding Rights under the "Data Twenty Articles" [J]. Journal of Information Resources Management, 2024, 14(2): 54-67. [7] Notice of the National Data Bureau on Issuing the "Trusted Data Space Development Action Plan (2024-2028)" [EB/OL]. [2025-05-06]. https://www.gov.cn/zhengce/zhengceku/202411/content_{6996363}.htm. [8] Xia Yikun, Jiang Jie, Zhang Xiaheng, et al. Response and Reflection of Information Resources Management Discipline on Developing New Quality Productive Forces[J]. Journal of Library and Information Science in Agriculture, 2024, 36(1): 4-32. [9] Wang Xue, Xia Yikun, Pei Lei. Research

Progress on Data Factor Markets at Home and Abroad: A Systematic Literature Review[J]. Library and Information Knowledge, 2023, 40(6): 117-128. [10] Zang Guoquan, Xiao Yang, Zhang Kailiang. Research on Privacy Risk Measurement and Tiered Protection Mechanism for Government Data[J/OL]. Journal of Library Science in China, 2024: 1-14[2024-12-25]. http://kns.cnki.net/kcms/detail/11.2746.G2.20241118.1338.002.html. [11] Fan Ruguo. Platform Technology Empowerment, Public Game, and Complex Adaptive Governance[J]. Social Sciences in China, 2021(12): 131-152, [12] Liu Taoxiong, Rong Ke, Zhang Yadi. Data Capital Estimation and Its Contribution to China' s Economic Growth—From the Perspective of Data Value Chain[J]. Social Sciences in China, 2023(10): 44-64, 205. [13] Cui Wenbo, Zhang Tao, Ma Haiqun, et al. EU Data and Algorithm Security Governance: Characteristics and Implications[J]. Journal of Information Resources Management, 2023, 13(2): 30-41. [14] Sun Jianjun, Pei Lei, Fu Shaoxiong. Inclusive Integration: Data Management in the Context of Information Resources Management Discipline Construction[J]. Journal of Information Resources Management, 2023, 13(1): 9-17. [15] Wang Yan, Yang Da. Transformation of Chinese Management Accounting System: From Data Elements to Data Assets[J]. Management World, 2024, 40(10): [16] Yang Xinya, Wang Ying, Yin Weihong. Research on Data-Driven New Intelligence Services[J]. Journal of Literature and Data, 2019, 1(1): 32-41, 117. [17] Li Yuhai, Wang Rui. Ten-Year Practice and Future Prospects of Government Data Opening[J]. Journal of Literature and Data, 2022, 4(4): 12-14. [18] Ma Haiqun, Zhang Tao. Theoretical Interpretation of Data Productivity and Its "New Quality" Power[J/OL]. Journal of Library Science in China, 2024: 1-16[2024-12-25]. http://kns.cnki.net/kcms/detail/11.2746.g2.20241126.1531.002.html. [19] Liu Zhaoge, Li Xiangyang, Qiao Limin, et al. Analysis Method for Big Data Governance Model in Urban Disaster Risk Response with Case Support[J]. Journal of the China Society for Scientific and Technical Information, 2024, 43(6): 672-684. [20] Liu Mingda, Chen Zuoning, Shi Yijuan, et al. Research Progress of Blockchain in Data Security[J]. Chinese Journal of Computers, 2021, 44(1): [21] Wu Yikai, Li Guoan. Technology Governance and Governance Technology: A Perspective from Blockchain Digital Asset Regulation Research[J]. Bulletin of Chinese Academy of Sciences, 2024, 39(8): 1375-1388.

Note: Figure translations are in progress. See original paper for figures.