International Journal of Ethical AI Application
eISSN: 3093-7736
https://ijeaa.cultechpub.com/index.php/ijeaa

Legal Logic of AI Data Governance Based on Federated Learning: Institutional Evolution from Privacy Protection to Rights Distribution

Wenzhou Shu

French School, Sichuan International Studies University, Chongqing, China; School of International Law, Southwest University of Political Science and Law, Chongqing, China

Email: 20202402110033@stu.sisu.edu.cn

Abstract

This study focuses on the legal logic of data governance under federated learning technology in the field of artificial intelligence (AI). Through legal reasoning and literature analysis, it delves into the importance of federated learning as a key institutional approach that balances privacy and data utilization in the face of real-world challenges such as conflicts between data silos and privacy protection, disputes over the ownership of data rights in AI training works, and compliance pressures from the EU AI Act and GDPR. Starting from a legal logic inference framework and the legal value foundation of federated learning, this study reviews existing research findings and shortcomings, and constructs a logical chain of institutional evolution from privacy protection to rights allocation. The study finds that federated learning is not merely a technical tool but also an opportunity to drive legal institutional design innovation. The institutional chain linking privacy to interest allocation constitutes a new paradigm for AI data governance. By integrating legal logical inference into technical literature and fusing the three legal logics of privacy, ownership, and benefits, this research provides an innovative institutional evolution model and operational governance recommendations for AI data governance, demonstrating significant theoretical and practical significance in interdisciplinary research.

Keywords

Federated Learning, Data Privacy, Legal Logic, AI Regulation, Benefit Sharing, Trade Secrets

1. Introduction: Problem Statement and Significance of Research

1.1 Real-Life Predicament

In the digital age, data has become the core driving force behind AI development. However, the contradiction between data silos and privacy protection requirements is becoming increasingly acute. On the one hand, various institutions and enterprises lock data within their own systems for their own interests, preventing data from flowing and integrating effectively. This severely hinders the aggregation of the massive amounts of data required for AI model training and limits the development potential of AI technology. On the other hand, as AI technology is widely adopted, large amounts of personal data are collected, stored, and processed, significantly increasing the risk of data privacy breaches and sparking public concerns about the security of personal information. The introduction of the European Union's General Data Protection Regulation (GDPR) underscores the stringent requirements for data privacy protection, with any activities involving the processing of personal data now facing unprecedented compliance challenges [1].

During AI training, disputes over the ownership of work data have become increasingly prominent. AI model training often relies on large amounts of copyright-protected works, but the boundaries of rights between data sources, data processors, and AI model developers are unclear during use. For example, language models trained on text data lack clear legal definitions regarding the copyright ownership and authorization of the text used, which not only easily leads to copyright disputes but also affects the healthy development of the AI industry [2].

At the same time, the formal implementation of the EU AI Act has imposed comprehensive regulatory pressures on the compliance of AI models. The Act adopts a risk-based tiered management approach, categorizing AI systems into four risk levels: unacceptable risk, high risk, limited risk, and minimal risk [3], and establishes corresponding stringent compliance requirements and regulatory measures for each level. This means that AI service providers must invest significant resources to meet the bill's requirements during the research, development, deployment, and application of AI systems. Failure to comply could result in hefty fines and other severe penalties, which undoubtedly increases operational costs and management complexity for businesses and has far-reaching implications for the global AI industry landscape.

1.2 Theoretical Value

The rapid development of artificial intelligence (AI) technology has made interdisciplinary research between AI and the field of law increasingly necessary. The application of AI technology must be subject to legal regulations to ensure that technological development complies with social ethics and legal order. At the same time, the design of legal systems also needs to fully consider the special scenarios and needs of AI technology to achieve organic compatibility between the two. As an emerging technology model, federated learning, through decentralized data processing, balances the dual goals of privacy protection and data utilization to a certain extent, providing an institutional entry point for solving the current challenges facing AI development. In-depth research on the legal logic of federated learning in AI data governance from a theoretical perspective [4,5] will help build a comprehensive AI legal regulatory system, fill the gap in interdisciplinary research, and promote the innovative development of legal theory in emerging technology fields.

2. Theoretical Basis and Research Approach

2.1 Basic Framework of Legal Logical Reasoning

In the field of AI governance, legal logic plays a crucial role. Deductive reasoning applies general legal rules to specific AI scenarios. For example, based on the provisions of the GDPR regarding the legality of personal data processing, it determines whether data collection and use during AI model training are compliant. Inductive reasoning, on the other hand, analyzes a large number of AI-related cases to summarize universal legal principles and patterns. There is a close interactive relationship between normative logic and technical facts. Legal norms set boundaries and standards for the development of AI technology, while the practical application of AI technology in turn drives the continuous improvement and updating of legal norms. In the issue of AI model explainability, the "black box" nature of technology conflicts with legal requirements for transparency and explainability, thereby driving the formulation of relevant legal norms to clarify the obligations and responsibilities of AI developers in explaining model decision-making processes.

2.2 The Legal Value Foundation of Federated Learning

Federated learning prioritizes privacy and data minimization as core principles, aligning with modern legal requirements for data protection. Under its distributed governance architecture, data is stored in a decentralized manner across participating nodes, reducing the risk of privacy leaks associated with centralized data storage. Each node uses data locally for model training and achieves collaborative model optimization through parameter exchange under encryption mechanisms, ensuring that "data remains in the database and is usable but not visible," thereby maximizing the protection of data subjects' privacy rights. At the same time, this decentralized data processing model reshapes the boundaries of rights and responsibilities in traditional legal relationships. In traditional data processing models, data controllers often have greater power and responsibility, while in federated learning scenarios, multiple parties participate in data processing and model training, and the rights and obligations of each party need to be redefined, including data ownership, privacy protection responsibilities, and data usage permissions. This presents new challenges and opportunities for the design of legal systems.

2.3 Achievements and Shortcomings of Existing Research

At the technical level, research on federated learning has achieved significant results. Numerous literature has explored the architectural design of federated learning, privacy protection technologies (such as differential privacy and secure multi-party computation), and practical applications in various fields, providing a solid foundation for the technical implementation and optimization of federated learning. However, in terms of legal research, although the importance of federated learning in AI data governance has been recognized, there are still many shortcomings. There is a lack of systematic analysis of the compatibility of federated learning with existing legal systems, and a comprehensive legal framework to regulate the application of federated learning has not yet been established. In terms of data ownership, the ownership of data generated through multi-party collaboration in federated learning scenarios and the rules for rights allocation are unclear, which may easily lead to legal disputes. Regarding the interest allocation mechanism, how to reasonably determine the interest allocation relationship among data originators, data processors, and model developers to incentivize all parties to actively participate in federated learning while safeguarding the legitimate rights and interests of data originators remains a weak area in current legal research, with a lack of operational institutional designs and logical analyses [6].

3. The Logical Chain of Institutional Evolution

3.1 Privacy Protection Logic

Within the global legal framework for data governance, the EU's legislative framework plays a significant leading role. The General Data Protection Regulation (GDPR) and the Artificial Intelligence Act (AI Act) together form a rigorous compliance framework for AI data processing. As the foundational legislation in the field of data protection, the GDPR explicitly establishes the legality of personal data processing on the principles of informed consent and data minimization. It requires data controllers to clearly inform data subjects of the purpose, scope, and methods of data use, and to collect and process only the data necessary to achieve specific purposes. The AI Act further refines the risk assessment mechanisms for AI systems, classifying them based on the potential harm they may cause in different

application scenarios, and imposing strict compliance requirements such as transparency, explainability, and human oversight for high-risk AI systems.

Against this backdrop, federated learning technology has demonstrated strong institutional adaptability. Relying on encryption technology and a distributed architecture, it has established a new data processing paradigm in which "models move, data stays put." Under the federated learning framework, raw data is always stored locally by data providers, and only encrypted model parameters or intermediate calculation results are exchanged between nodes, fundamentally avoiding the remote transmission of sensitive data, which is highly consistent with the data minimization principle advocated by the GDPR. Additionally, federated learning exhibits strong technical compatibility, enabling seamless integration with cutting-edge privacy-preserving technologies such as differential privacy and secure multi-party computation (SMC). Differential privacy achieves data anonymization by adding controlled noise to data without compromising the usability of data analysis; secure multi-party computation allows multiple parties to collaboratively perform computational tasks without disclosing their original data. The combined application of these technologies significantly enhances the depth of data privacy protection.

In terms of trade secret protection, the rapid development of AI technology poses new challenges to traditional trade secret protection mechanisms. With the emergence of massive amounts of intermediate data, model parameters, and optimization algorithms generated during AI model training, these information carriers with economic value and subject to reasonable confidentiality measures should be included in the scope of trade secret protection if they meet the three criteria of "secrecy, value, and confidentiality." In the context of federated learning, due to the involvement of data fusion and collaborative modeling among multiple parties, the intermediate results generated during model training often contain unique commercial value and technical advantages of each party. Clarifying their trade secret attributes can effectively prevent data leakage and unfair competition, providing participating parties with stable expectations for rights protection.

3.2 Equity Attribution Logic

In the AI industry ecosystem, defining the rights of data providers is a core issue in data governance systems. Although the current legal framework generally assigns copyright of AI-generated results to model developers or users, data providers, as the providers of basic resources for model training, cannot be ignored for their contributions. From a legal perspective, the doctrine of unjust enrichment provides a crucial theoretical foundation for data source providers to assert their rights. This doctrine holds that if data processors or model developers obtain excessive benefits by using data source providers' information resources without paying reasonable compensation, they have unjustly enriched themselves, and the data source providers are entitled to demand the return of such benefits in accordance with the law.

In the context of federated learning for cross-platform data collaboration, contract mechanisms have become an important legal tool for rights allocation. By signing detailed data cooperation agreements, all parties can clearly define key issues such as data usage rights, cooperation terms, ownership of results, and profit distribution. For example, in the application of federated learning in the medical field, multiple medical institutions can improve diagnosis and treatment levels by jointly training disease diagnosis models. Contracts can precisely define the scope of patient data provided by each institution, the specific rights to use the model, and the distribution ratio of profits generated from the application of the model. This contractual arrangement not only effectively prevents disputes over rights and interests but also stimulates the enthusiasm of all parties to participate in data collaboration through a clear incentive mechanism, promoting the healthy development of the AI industry ecosystem.

3.3 Trends in Institutional Evolution

The iterative development of AI technology and its deepening application in various industries are driving profound changes in global data governance legal systems. Early data protection legislation focused on preventing data breaches and protecting personal privacy, with an emphasis on restrictive regulations on data processing activities. However, as the AI industry accelerates its transition from technology R&D to commercial application, the economic value of data as a new type of production factor has become increasingly prominent. Pure privacy protection is no longer sufficient to balance the dual demands of data security and industrial development.

Currently, legal systems are gradually evolving toward a composite governance model that combines privacy protection with interest distribution. While strictly adhering to the bottom line of data security, this model places greater emphasis on establishing fair and reasonable mechanisms for the distribution of data value. Through legislation, the rights and obligations of multiple parties, including data subjects, processors, and users, are clarified to promote the orderly circulation and efficient utilization of data elements.

In the three-dimensional interaction between technological innovation, legal regulation, and ethical constraints, AI data governance models are undergoing a paradigm shift. At the technological level, emerging technologies such as federated learning, blockchain, and homomorphic encryption provide powerful technical capabilities for data governance, enabling cross-domain collaboration while ensuring data security and controllability. At the legal level, countries are accelerating the process of data legislation and establishing a set of rules covering the entire life cycle of data, including collection, storage, use, and sharing. At the ethical level, data ethics principles centered on fairness, transparency, and accountability provide value guidance for technology application and institutional design. Future AI

data governance requires the construction of a three-dimensional governance framework that integrates technology, law, and ethics. Through the dual drivers of institutional innovation and technological reform, a dynamic balance between data security and development can be achieved, laying a solid foundation for the sustainable development of the AI industry.

4. Legal Reasoning and Institutional Design

4.1 Legal Inferences Regarding Privacy Protection

The EU's Artificial Intelligence Act (AI Act) and General Data Protection Regulation (GDPR) establish a rigorous legal framework for data processing, with the principles of informed consent and data minimization as its cornerstones. This requires data controllers to strictly fulfill dual obligations throughout the entire lifecycle of AI data processing: on the one hand, they must obtain explicit consent from data subjects through clear, unambiguous notification procedures; on the other hand, they must precisely limit the scope of data collection and use to what is strictly necessary to achieve specific purposes. In the healthcare sector, for example, when using personal medical data to train AI diagnostic models, it is not only necessary to obtain explicit written or electronic authorization from patients but also to strictly screen data types, using only key information directly relevant to disease diagnosis to minimize the risk of overexposure of sensitive information such as patients' health histories or genetic data.

As an innovative distributed machine learning paradigm, federated learning is naturally suited to data minimization requirements. With features such as localized data processing and encrypted parameter transmission, federated learning effectively reduces the risk of privacy leaks caused by cross-domain data flow and can serve as an important technical tool for meeting compliance requirements. However, in the face of increasingly sophisticated attack methods, relying solely on the basic infrastructure of federated learning still poses security risks. Therefore, it is necessary to build a composite solution of "federated learning + privacy-enhancing technology": introduce differential privacy technology to inject controllable noise into the model parameter update stage, making it difficult for attackers to reverse engineer the original data from gradient information; combine secure multi-party computation (SMC) technology to securely aggregate encrypted models on the central server, ensuring the confidentiality of data throughout the entire chain of transmission, storage, and computation, and forming a double protection mechanism at the technical level.

To unlock the innovative potential of federated learning, the legal framework must establish a flexible compliance exemption system. It is recommended that legislation explicitly stipulate that federated learning systems may be exempted from certain requirements when they meet the following technical standards: the use of differential privacy algorithm libraries certified by the National Institute of Standards and Technology (NIST); the deployment of secure multi-party computation protocols based on homomorphic encryption or secret sharing technology; and include a security audit module for continuous monitoring of data flows, with audit logs retained for a period no shorter than the statutory retention period. This "technical compliance-based exemption" framework ensures the protection of data security while reserving room for technological innovation, thereby achieving a dynamic balance between privacy protection and industrial development.

4.2 Data Ownership and Trade Secret Protection

According to the constituent elements of the trade secret legal system, AI-generated information may be legally protected as trade secrets if it meets the three criteria of "secrecy, value, and confidentiality." Secrecy requires that the information is not known to the public; value is reflected in the ability to bring actual or potential economic benefits to the rights holder; and confidentiality requires the adoption of reasonable technical or management measures to prevent information leakage. During AI training, innovative outcomes such as optimized neural network architectures, unique algorithm parameter combinations, and intermediate data reflecting market trends may be included within the scope of trade secret protection if they meet the aforementioned criteria.

In multi-party collaboration scenarios involving federated learning, data ownership and trade secret protection issues are particularly complex. Since model training involves the integration of data from multiple parties and the co-creation of knowledge, the model parameters and intermediate results ultimately generated often embody the core data assets and technical expertise of all participating parties and have extremely high commercial value. For example, the parameter configuration of credit risk assessment models jointly trained by financial institutions through federated learning not only reflects the risk control experience of all parties but also contains potential competitive advantages in the market. For such outcomes, a comprehensive protection mechanism combining "pre-agreement + post-remedy" should be established: during the project initiation phase, legal agreements should clearly define data ownership, usage permissions, and confidentiality obligations; in the event of infringement disputes, civil liability for compensation may be pursued against the infringing party in accordance with relevant regulations such as the Anti-Unfair Competition Law.

From the perspective of institutional improvement, it is imperative to establish a legal framework for data ownership that is adaptable to multi-party collaboration. On the one hand, under the existing legal framework of the Data Security Law and the Personal Information Protection Law, special provisions should be added for federated learning scenarios to clarify the rights and obligations of data contributors, model developers, and application providers. On the other hand, drawing on the intellectual property co-ownership system, we should explore the establishment of data asset co-

ownership models based on shares or joint ownership, and clarify through legislation the rules for determining the ownership of trade secrets in multi-party collaboration scenarios, providing clear legal guidance for practice.

4.3 Institutional Evolution of the Benefit-Sharing Mechanism

As the original providers of data for AI training, data sources typically lack absolute copyright control under current legal frameworks. However, based on the principle of "who contributes, who benefits," they should be entitled to corresponding economic returns. The theory of unjust enrichment provides a solid legal basis for this: when data processors or model developers obtain excessive benefits from using others' data, data sources have the right to assert their rights through legal means. In practice, contractual arrangements are typically used to clarify the mechanism for distributing benefits, with data usage agreements specifying key terms such as methods for assessing data value, profit-sharing ratios, and payment methods.

Within the framework of federated learning technology, smart contracts provide an innovative solution for benefit distribution. Based on blockchain technology, smart contracts are immutable and self-executing, effectively overcoming the high trust costs and fulfillment risks associated with traditional contract execution processes. All parties involved can pre-code the rules for benefit distribution into smart contracts and set trigger conditions (such as model accuracy reaching 90% or completion of the third stage of training). When the conditions are met, the smart contract automatically executes the benefit distribution operation, ensuring that the rights and interests of all parties are realized in a timely manner. This decentralized automated execution mechanism significantly improves collaboration efficiency and reduces dispute resolution costs.

To establish a long-term incentive mechanism, it is imperative to establish the legal right of data originators to receive benefits through legislation. It is recommended that normative documents such as the "Guiding Opinions on the Cultivation of the Data Element Market" clearly stipulate that data originators are entitled to receive no less than 15% of the revenue generated from the sale of data products or services (the specific proportion may be adjusted based on factors such as data type and usage scenarios). Additionally, the revenue rights system must be deeply integrated with privacy protection mechanisms: under the premise of ensuring data anonymization and de-identification, a secure channel for data value assessment and revenue distribution should be established to achieve a virtuous cycle where "privacy protection is guaranteed and data contributions are rewarded," thereby promoting the formation of a more fair, transparent, and sustainable data element market ecosystem.

5. Legal Logic Framework for AI Data Governance Based on FL

Throughout the entire data governance lifecycle, safeguarding the rights and interests of data originators is the top priority, necessitating the establishment of rigorous mechanisms for ownership verification and informed consent. Specifically, during the **data input stage**, the primary task is to clearly define the boundaries of data ownership in legal documents, using rigorous legal instruments to precisely delineate the scope of rights that data sources hold over the data, including ownership, usage rights, and the right to derive benefits. In this process, legal tools such as contracts and declarations should be fully utilized to clearly delineate the rights and obligations of data sources and data processors through detailed clauses. At the same time, the strict enforcement of informed consent mechanisms is indispensable. Data processors have the responsibility to disclose key information such as the purpose of data use, storage period, sharing recipients, and potential risks to data sources in a clear, straightforward, and easy-to-understand manner, and to obtain explicit authorization from data subjects through written or electronic confirmation in a compliant manner. For example, when using personal user consumption data for AI marketing model training, it is not only necessary to provide a detailed explanation of the specific methods of data use and analysis dimensions but also to inform users of the potential derivative value and risks associated with the data, ensuring that users make autonomous decisions based on full disclosure.

In the **data processing stage**, the application of federated learning (FL) technology framework and the deep integration of privacy protection technologies form a double line of defense for data security. In practice, federated learning technology avoids security risks associated with data concentration at the architectural level by enabling distributed storage and local computing of data. At the same time, cutting-edge privacy protection technologies such as differential privacy, secure multi-party computation, and homomorphic encryption are embedded throughout the entire federated learning process. During the model parameter update and transmission process, differential privacy technology is used to add controllable noise, so that even if data is leaked, attackers cannot restore the original data. With the help of secure multi-party computation technology, data aggregation and computation are performed in an encrypted state, ensuring that data remains secure throughout the processing process. This combination of technologies not only guarantees data availability and computing efficiency but also strictly protects data privacy, meeting legal compliance requirements for data processing security.

When AI models have been trained and are ready for **output**, the core tasks are to establish a rights distribution mechanism and a trade secret protection system. On the one hand, it is necessary to establish a scientific and reasonable rights distribution mechanism that comprehensively considers factors such as the value of the data provided by the data source, the technical investment of the data processor, and the intellectual contribution of the model developer. Through contractual agreements or smart contracts that are automatically enforced, the benefits of model application can be distributed fairly. On the other hand, for key information with trade secret attributes, such as model parameters, unique

algorithms, and intermediate data, a multi-layered confidentiality system should be established. This system should address personnel management, technical safeguards, and institutional constraints from multiple dimensions, and establish strict confidentiality agreements, access control strategies, and encryption storage measures to prevent the leakage of trade secrets and protect the legitimate commercial interests of all parties within the AI industry ecosystem.

The entire AI data governance activity based on federated learning is subject to strict **external legal and regulatory oversight**. The EU's AI Act and GDPR have established a world-leading data governance standard system, implementing risk-based management of AI data processing activities and establishing strict privacy protection principles. China's policies and regulations, such as the "Data Twenty Measures," are tailored to domestic conditions and provide clear guidelines for key areas such as data security management and the market-based allocation of data elements. These laws and regulations provide a robust institutional framework and behavioral guidelines for AI data governance at the macro level, ensuring that federated learning technology operates within a legal and compliant framework and safeguarding the healthy and sustainable development of the AI industry.

6. Conclusion

Federated learning is not merely a technical tool in the field of artificial intelligence data governance; it also represents a critical opportunity for reforming legal and institutional frameworks. Its decentralized data processing logic not only responds to the policy direction of "exploring ways to share data value and benefits" outlined in the "Data Twenty Measures" but also aligns with the strict privacy protection requirements of the GDPR and the EU AI Act through its "data remains within its domain" technical characteristics. This ultimately forms a new AI data governance paradigm characterized by the bidirectional coupling of privacy protection and rights distribution.

From the perspective of institutional evolution, the legal logic in federated learning scenarios has broken through the traditional "data controller-data subject" binary framework: on the one hand, by embedding technologies such as differential privacy and secure multi-party computation (SMC), federated learning can meet the requirements of "confidentiality" and "secrecy" in commercial secret protection, incorporating model parameters and intermediate data generated through multi-party collaboration into the scope of protection, thereby addressing the ambiguity in rights attribution during AI training; On the other hand, based on the unjust enrichment system and collective governance model, federated learning provides a technical implementation path for data source stakeholders to share benefits. For example, smart contracts can automatically enforce the "Shapley value method" distribution rules, which not only aligns with Zhang Jiaxin's principle of "data value increment allocation" but also avoids the high costs of interest negotiation in traditional centralized governance.

References

[1] Kairouz P, McMahan H B, Avent B, et al. Advances and open problems in federated learning[J]. Foundations and Trends® in Machine Learning, 2021, 14(1-2): 1-210.

[2] Zhang Jiaxin. Research on the Mechanism for Sharing the Benefits of Work Data Sources in Artificial Intelligence Training [J]. Intellectual Property Rights, 2025(05): 111-126.

[3] Ma C, Li J, Ding M, et al. On safeguarding privacy and security in the framework of federated learning[J]. IEEE Communications Surveys & Tutorials, 2020, 22(3): 2018-2064.

[4] Li Q, Wen Z, Wu Z, et al. A survey on federated learning systems: Vision, hype and reality for data privacy and protection[J]. IEEE Transactions on Knowledge and Data Engineering, 2021, 35(4): 3340-3365.

[5] Yang Q, Liu Y, Chen T, et al. Federated machine learning: Concept and applications[J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2019, 10(2): 1-19.

[6] Xu R, Baracaldo N, Zhou Y, et al. HybridAlpha: An Efficient Approach for Privacy-Preserving Federated Learning[C]//Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security. ACM, 2019: 13-24.