

Research on the Concept and Ownership Issues of Data Assets

Yuhan Zhang*

School of Management, Xi'an Polytechnic University, Xi'an, 710600, China

*Corresponding author: Yuhan Zhang, 2298245530@qq.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: The digital economy's rapid growth has made data a key driver of economic and social development. However, the replicable and shareable nature of data assets complicates the clear definition of ownership, hindering efficient circulation, market allocation, and technological innovation. Defining data ownership involves balancing the interests of multiple stakeholders—individuals, businesses, and governments—amid challenges such as non-exclusivity and replication. Current legal frameworks, including China's Data Security Law and Personal Information Protection Law, remain inadequate in addressing these issues systematically, while international disagreements over data sovereignty add further complexity. This study examines theoretical and practical challenges related to data asset ownership and proposes an ownership framework aimed at balancing efficiency and fairness. Theoretically, it explores innovations in data property rights and accounting standards. Practically, it analyzes ownership disputes between individuals and enterprises, offering insights for improving data infrastructure, facilitating data circulation, and reconciling individual rights with public interests.

Keywords: Data; Ownership; Digital Economy; Legal; Frameworks

Published: Mar 1, 2026

DOI: <https://doi.org/10.62177/apemr.v3i1.1144>

1. Introduction

Globally, the development of data assets is accelerating, with their economic and social benefits becoming increasingly prominent. However, due to characteristics such as replicability and ease of sharing, data faces numerous practical challenges in defining property rights, which serve as a major bottleneck restricting the market-oriented allocation, orderly circulation, and efficient utilization of data factors. Currently, how to scientifically clarify the ownership relationships of data assets, reasonably balance the rights and interests of multiple stakeholders, and establish a property rights rule system aligned with the digital era has become a cutting-edge issue at the intersection of law, economics, and information science. It also represents a pressing practical challenge for China to deepen the market-oriented reform of data factors and actively participate in global digital governance. The issue of data ownership is highly complex, stemming not only from the non-exclusivity and fluidity of data itself but also from the involvement of multiple stakeholders in its generation, processing, and circulation, making it difficult to clearly delineate the boundaries of ownership, usage rights, and profit rights. Meanwhile, existing legal frameworks still lag behind in addressing data ownership issues. Although laws such as the *Data Security Law* and the *Personal Information Protection Law* have laid the groundwork for governance, systemic arrangements for ownership remain insufficient. In light of this, this study aims to systematically analyze the theoretical disputes and practical

challenges surrounding data asset ownership, explore the construction of a property rights definition framework that balances efficiency and fairness, and provide theoretical references and practical guidance for improving China's data foundation system, promoting the orderly circulation of data factors, and balancing individual rights with public interests.

2. Conceptual (Sorting out) of Data Assets

Studying data asset ownership first requires clarifying what constitutes a data asset. This section (sorts out) the concept of data assets based on definitions provided by domestic and international scholars, clarifies which assets are data assets, and uses this as a foundation for researching data asset ownership.

2.1 Domestic and International Definitions of Data Assets

International scholars' research on defining the concept of data assets currently focuses primarily on the core question of "whether data constitutes an asset." The concept of data assets originates from information assets, which themselves are an extension of the Big Data concept. The term "Big Data" was first proposed by American scholar Toffler in 1980 and gradually evolved into a consensus definition based on characteristics such as volume, velocity, and value through deepening research^[1]. With technological advancements and hardware iterations, data volume has grown exponentially, and the connotation and extension of information assets have gradually gained attention. Meyer first explicitly proposed the concept of "digital assets" in his article "A Methodology for Protecting Data Assets"^[2]. This concept sparked ongoing discussions among subsequent scholars. Gregory & Hunter's research on companies benefiting from information technology showed that data reliability is a key factor in whether data can become an asset driving enterprise development; incorrect data is not an asset and may even lead to decision-making errors^[3]. While data volume continued to surge, the research focus shifted to "how data becomes an asset." Building on data reliability, Viktor summarized four characteristics of data and pointed out that data will inevitably become a component of the balance sheet; however, although the accounting industry has focused on data protection and reliability, organizational levels still face challenges in data evaluation and management^[4]. Based on the need for data asset rights confirmation, research has also reversely deduced the definition of data assets, increasingly from a human perspective. For example, Havur G, Steyskal S used the ODRL policy expression language to build a data license clearing center, allowing users to write arbitrary licenses for the legal and secure reuse of third-party data assets and supporting rights attribution issues by providing relevant data information^[5]. Sabina Leonelli emphasized that data is the cornerstone of science, and science is the foundation for understanding the world. She argued that when data is regarded as a stable object, its scientific significance is determined by a few professionals, and illustrated through examples that data is an important asset. Leonelli reviewed how research data was reconfigured as a political-economic asset in the 20th century, pointed out the trend of rising status for data curators, and criticized certain foreign social media companies for abusing customer privacy data to manipulate markets. She finally clarified that data has value at the time of generation, but its importance depends on its source and subsequent processing^[6]. Hannila, while studying data assets and data-driven PPM (Project Portfolio Management), pointed out that although much data currently has low utilization rates, it should still be treated like other assets because data assets themselves are a strategic resource^[7]. Changyu Hu studied the role of data assets in enterprise operations, showing that they positively enhance company efficiency, but the degree of effect is influenced by the company's nature^[8]. Despite numerous related studies, the concept and definition of data assets have only become clearer in recent years. Yuan Zhang proposed a relatively clear definition, stating that data assets are collections of data actively or passively collected by enterprises that can generate profit after processing^[9]. International scholars' research on defining the concept of data assets has evolved from discussing whether data constitutes an asset (originating from the evolution of information assets and big data, initially laid by Toffler, Glazer, Meyer, etc.), to focusing on how data becomes an asset (emphasizing reliability, characteristics as described by Viktor, and recognizing challenges in accounting and management), and then to (focusing on) the confirmation of rights, value sources, and applications of data assets (e.g., Havur et al. addressing rights attribution, Leonelli emphasizing scientific and political-economic value and source/processing dependency, Hannila viewing it as a strategic resource, Changyu Hu empirically proving its role in enhancing operational efficiency). Although research is abundant, its conceptual definition has only recently become clearer, with current consensus tending to view it as datasets owned by enterprises that can generate profit after processing.

Chinese scholars' research on defining the concept of data assets mainly focuses on "what a data asset is," primarily through comparing with intangible assets to (summarize/induce) characteristics, and mostly from an enterprise perspective. Lv Yuqin defined it as a resource realized through electronic form, without physical form, (believing) it widely exists in enterprises and the Internet, can bring economic benefits to enterprises, and its proportion will continue to rise ^[10]. Liu Yu emphasized the difference in sources of data assets, pointing out that they include data actively collected by the holder (with defined purpose) and passively acquired (not fully mined) data ^[11]. Zou Zhaoju proposed that data assets originate from daily business operations and must meet three conditions: enterprise ownership/control, existence in physical or electronic form, and ability to generate expected future (benefits/returns) ^[12]. Li Zehong, based on the intangible asset framework, advocated for incorporating data assets into the accounting scope of intangible assets and systematically studied their initial measurement, subsequent measurement, disposal, and information disclosure ^[13]. Li Maohao, when using real options models for evaluation, defined them as quantifiable data resources generated by historical activities of an enterprise and controlled by the enterprise ^[14]. Ma Dan, Yu Xi, by migrating OECD rules through the DIKW model, clarified that the statistical boundary of data assets requires a usage period exceeding one year, specific quality, profitability, and market tradability ^[15]. In discussing characteristics, Li Yuan believed that data is generated from human activities, stored in generalized databases or physical forms, and belongs to tangible fixed assets that require continuous maintenance, lack stable physical units, are reusable, and do not require revaluation ^[16]. Hu Yaru, (combining) national economic accounting standards, excluded non-digital data (e.g., paper archives) and digital content products (e.g., online music), emphasizing that enterprises need to obtain benefits through holding/using data to possess its economic ownership ^[17]. Gao Hua, Jiang Chaofan integrated previous research and achievements of CAICT (China Academy of Information and Communications Technology), defining data assets as: information resources with clear ownership and value, measurable and readable, after processes of collection, mining, cleaning, labeling, and analysis ^[18].

Sun Yongyao, Yang Jiayu compared domestic and international accounting perspectives, pointing out that China adopts a broader approach, considering data assets as intangible assets but with differing characteristics: their evaluation serves production and operation needs (e.g., user positioning, risk control), possessing scenario specificity; while accounting measurement needs to reflect universal fair value ^[19]. Wang Manlin classified data assets according to subordinate value and designed accounting treatment methods for initial measurement and subsequent measurement respectively ^[20].

2.2 Shooting Accuracy

The user did not explicitly request a citation format, but in an academic context, sources must be cited. Decided to retain scholar names and years in the distinction dimension table, saving space and facilitating user tracing. Finally, using Deng Jiandi's identification decision tree case is very appropriate—the five characteristics she proposed the conversion threshold from resource to asset.

Table 1: Core Definition and Conversion Condition Diagram

Dimension	Data Resource	Data Asset
Essence	Raw, unprocessed foundational data collection, ownership and economic value not yet clear.	Data resources with confirmed ownership, processed, and possessing quantifiable economic value.
Ownership & Control	May be in a dispersed or unconfirmed state (e.g., public data, user behavior data).	Owned or controlled by enterprises/institutions, clear ownership (e.g., authorized, purchased, or self-produced).
Value Attribute	Potential value needs realization through development; cannot be	Value transformation achieved; can be included in accounting measurement or strategic decisions (e.g., customer profiles for precision marketing).
Accounting Treatment	Usually not included in financial statements (costs difficult to aggregate, value not explicit).	Can be included in the scope of intangible assets for initial measurement, amortization, or impairment.
Conversion Conditions	Requires: 1 Processing (cleaning, labeling, etc.); 2 Clear ownership; 3 Ability to generate (benefits/returns).	Already possesses: 1 Measurability; 2 Tradability.

3.Policies Related to Data Asset Ownership

“The Opinions of the Central Committee of the Communist Party of China and the State Council on Constructing Data Basic Systems to Better Leverage the Role of Data Elements,” known as the “Data Twenty Articles,” was released on December 19, 2022. This opinion constructs the data basic system from aspects such as data property rights, circulation and trading, (benefit distribution), and security governance, proposing 20 policy measures aimed at fully leveraging China’s advantages in massive data scale and rich application scenarios to activate the potential of data elements.

3.1 Establishing a Data Property Rights System

Explore a structurally separated data property rights system, proposing a framework for the “separation of three rights”: data resource (holding right), data (processing and usage right), and data (product operation right). This is a major innovation, as the interests of relevant entities involved in data production, circulation, and usage are complex, making traditional property rights systems difficult to apply. The “Data Twenty Articles” innovates the concept of data property rights, downplays ownership, emphasizes usage rights, focuses on the circulation of data usage rights, and constructs a data property rights system with Chinese characteristics.

Data circulation and trading are key to the transformation of data resources into data assets but face challenges such as difficult rights confirmation and pricing. The “Data Twenty Articles” proposes strengthening top-level design from aspects such as circulation rules, trading markets, and service ecosystems. Specifically, it includes establishing standard rules for data circulation access, exploring the construction of a standardized system for data quality; coordinating and optimizing the planning layout of national data trading venues, issuing management measures, and constructing a multi-level market trading system; cultivating two types of entities: data brokers and third-party professional service agencies.

At the primary distribution stage, according to the principle of “whoever invests, contributes, benefits,” promote the rational (tilting) of data element (benefits/returns) towards creators of data value and use value. At the secondary and tertiary distribution stages, focus on public interests and relatively disadvantaged groups, preventing and legally regulating risks such as market monopoly formed by (disorderly expansion) of capital in the data field. This helps improve the system where production factors participate in distribution according to contribution, allowing all people to better share the fruits of digital economic development and promoting common prosperity.

Construct a governance model involving multi-party collaboration between government, enterprises, and society, emphasizing security, controllability, resilience, and inclusiveness. Data, as a new type of production factor, has characteristics such as intangibility and non-consumption, can be replicated infinitely at (near-zero cost), and easily brings security risks. Through multi-party collaborative governance, it can both ensure data security and create a favorable environment for the data industry’s development, promoting compliant and efficient data circulation and use.

3.2 Analysis Related to the Separation of Three Rights

The data property rights “separation of three rights” proposed by the “Data Twenty Articles” refers to a property rights operational mechanism separating the data resource (holding right), data (processing and usage right), and data (product operation right). This is an innovative institutional framework that downplays ownership, emphasizes usage rights, and focuses on the circulation of data usage rights. According to the data source and data generation characteristics, it respectively defines the legal rights enjoyed by various participants in the process of data production, circulation, and use, establishing a property rights operational mechanism that separates the data resource holding right, data processing and usage right, and data product operation right. Based on the policy’s (proposition/advocacy) of the separation of three rights for data assets, we derive the following correspondence (Table 2)

Table2: Entity Actions and Corresponding Rights

Entity Action	Corresponding Right
Production	(Holding Right)
Use	(Processing and Usage Right)
Circulation	(Operation Right)

3.3 Hierarchical and Classified Management of Data Assets

As known from the second part of this paper, forming data assets from data resources requires multiple steps of processing, the core of which is the hierarchical and classified management of data assets. This is also a strategic cornerstone for enterprises to build core competitiveness in the digital era, driven by security and compliance. Increasingly strict domestic and international laws and regulations (such as GDPR, Data Security Law) require the grading of data, especially personal information and important data, and the application of corresponding protective measures. Lack of classification and grading forces security policies to be “one-size-fits-all,” leading to resource misallocation; with scientific grading, enterprises can achieve precise protection of core sensitive data, focus security investments on high-risk areas, clarify management responsibilities, and avoid compliance risks and huge fines. It is a prerequisite for releasing data value and promoting efficient utilization. Classification is like drawing a “map” for data, enabling business personnel to quickly discover, understand, and trust the data they need, effectively breaking down data silos. Grading lays the foundation for the assetization operation and valuation of core high-value data, empowering data-driven decision-making. It enables refined management of risk and cost. By identifying high-risk data and focusing monitoring, it significantly reduces the risk of leakage and misuse, protecting corporate reputation. Meanwhile, formulating different storage and computing strategies based on data levels can effectively optimize IT costs and avoid resource waste. Classification and grading are key (leverage points/starting points) for ensuring the effective implementation of the data governance system. They provide a clear basis for formulating all governance strategies such as data quality, lifecycle management, and sharing processes, making the originally abstract data governance executable, measurable, and auditable. In short, data classification and grading are the hub connecting strategy, security, and value, an indispensable step in transforming data from disordered resources into controllable assets.

According to the definition in “GB/T 38667-2020 Information Technology - Big Data - Guidelines for Data Classification,” data classification is the distinction and categorization of data based on its attributes or characteristics, according to certain principles and methods, to better manage and use data. There is no unique way to classify data; various classification systems can be formed based on enterprise management objectives, protection measures, classification dimensions, etc. Some domestic institutions such as the People’s Bank of China (PBOC), China Securities Regulatory Commission (CSRC), and Ministry of Industry and Information Technology (MIIT) have already issued standards related to data classification and grading, which also provide references to a certain extent.

Table 3 Data Classification Standards Issued by Relevant Domestic Institutions

Standard/Guideline Name	Issuing Institution	Main Reference Content
Guidelines for Financial Data Security Grading (JR/T 0197-2020)	People’s Bank of China (PBOC)	Objectives, principles, and scope of financial data security grading, as well as elements, rules, and process of data security classification.
Technical Specification for Protection of Personal Financial Information (JR/T 0171-2020)	PBOC	Mainly focuses on security protection in the collection, storage, processing, and other links of personal financial information.
Guidelines for Data Classification and Grading in the Securities and Futures Industry (JR/T 0158-2018)	China Securities Regulatory Commission (CSRC)	Classifies data into different levels based on the impact of data leakage or damage, providing a grading method for data security in the securities and futures industry.
Data Classification and Grading Methods for Basic Telecom Enterprises (YD/T 3813-2020)	Ministry of Industry and Information Technology (MIIT)	Data classification and grading in the telecommunications industry involve communication security, user privacy protection, etc.
Personal Information Security Specification (GB/T 35273-2020)	Standardization Administration of China (SAC)	Specifies security requirements for the collection, storage, use, sharing, etc., of personal information to protect it from illegal acquisition and use.

In practical implementation, the path for data classification and grading is often summarized into five steps: Step 1, Consultation, Investigation, and Analysis. Based on industry-related regulatory policies and standard specifications, conduct a comprehensive investigation and analysis of business systems, current state of data assets, and data security status, to have a “clear understanding” of the enterprise’s business, data, and security status. Step 2, Data Asset Inventory. Automatically identify data assets, inventory and tag data assets, build a good data asset catalog and inventory, laying the foundation for enterprise data classification and grading. Step 3, Data Classification Scheme. Design the data classification system based on the data asset inventory, complete the implementation of data classification tagging. After tagging, optimize the classification and grading rules to improve the proportion and accuracy of automated classification. Step 4, Data Grading Scheme. First, design the data grading system, then optimize the data grading rules, strive to improve the coverage and accuracy of automated grading, reduce labor costs, then establish data level change maintenance mechanisms and tool platform settings. Step 5, Data Classification and Grading Panorama. Construct a data classification and grading inventory, achieve visualization of data classification and grading. Simultaneously produce some data classification and grading operation mechanisms to lay a good foundation and prepare for data security graded protection.

From the perspective of data classification, classification methods can be (summarized/grouped) into three basic types: linear classification (method), facet classification (method), and hybrid classification (method).

Linear classification refers to dividing the classified object into multiple levels according to selected attributes or characteristics, with each level further divided into several categories. Under the same branch, categories at the same level are parallel, while categories at different levels constitute a subordinate relationship. All categories are mutually exclusive and non-overlapping in the vertical hierarchy.

Facet classification divides the data object into several independent, non-subordinate “facets” based on its inherent attributes or characteristics, with each facet containing a set of categories. Combining categories from different facets can form (compound categories). This method is a parallel classification approach, supporting multiple classification dimensions at the same level.

Hybrid classification combines both linear and facet classification methods to overcome the limitations of a single method and achieve more reasonable classification results. Its characteristic is using one method as the main approach and the other as auxiliary, typically applicable to scenarios where broad categories are divided first according to one dimension and then subdivided into smaller categories according to another dimension.

Data classification dimensions are diverse and can include aspects such as source, content, and purpose, often using a combination of multiple dimensions. For example, from a personal information perspective, it can be divided into personal information and non-personal information; from a business perspective, it can be divided into financial data, business data, operational data, etc.

4. Protection of Data Asset Ownership Rights and Interests

The “Data Twenty Articles” does not (get entangled in) who owns the data, but innovatively proposes a “separation of three rights” property rights system framework, separating the data resource holding right, data processing and usage right, and data product operation right. This aims to solve the ownership confirmation challenges brought by the unique properties of data elements (such as non-exclusivity, replicability). The fundamental purpose of this structurally separated property rights operational mechanism is to clarify the rights, responsibilities, and benefits of various links and participants in data processing and circulation, providing an institutional basis for the compliant and efficient circulation and use of data, thereby protecting relevant rights and interests. This paper argues that the entities forming data assets are divided into two types: sources and processors, and each should receive protection of their rights and interests.

Table 4: Data Asset Formation Methods and Ownership Protection

Entity Type	Actions Involved	Protection Principle
Source	Providing Data	Fully protect the legitimate rights and interests of the data source provider.
Processor	Collection	Reasonably protect the control rights of data assets held legally and compliantly by data processors.
	Storage	

5. Conclusion

Although domestic and international scholars have provided extensive elaboration and research on the definition of data assets, due to the special nature of data, the current conceptual definition of data assets remains imperfect. Regarding the ownership of data assets, there are currently no specific laws and regulations in China to standardize it. This paper analyzed the definition of data assets and, based on policies already issued by the state, analyzed the ownership definition of data assets, attempting to provide paths and methods for defining data asset ownership, to offer some help for the establishment of data asset classification, grading management, and rights protection systems. Clearly defining the form and ownership of data assets can standardize and advance the asset confirmation, compliance review, and registration work during their process of being (brought onto the balance sheet), thereby effectively reducing legal risks while enhancing the stability of enterprise operations

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

- [1] Graham-Rowe, D., Goldston, D., Doctorow, C., et al. (2008). Big data: Science in the petabyte era. *Nature*, 455(7209), 8-9.
- [2] Glazer, R. (1993). Measuring the value of information: The information-intensive organization. *IBM Systems Journal*, 32(1), 99-110.
- [3] Gregory, A. (2011). Data governance—Protecting and unleashing the value of your customer data assets. *Journal of Direct, Data and Digital Marketing Practice*, 12(3), 230-248.
- [4] Viktor, M. J., Rose, S. H., Van Zyl, W. H., et al. (2013). Raw starch conversion by *Saccharomyces cerevisiae* expressing *Aspergillus tubingensis* amylases. *Biotechnology for Biofuels*, 6(1), 1-11.
- [5] Havur, G., Steyskal, S., Panasiuk, O., et al. (2018). DALICC: A framework for publishing and consuming data assets legally. In *SEMANTICS Posters & Demos*.
- [6] Leonelli, S. (2019). Data—From objects to assets.
- [7] Hannila, H., Silvola, R., Harkonen, J., et al. (2022). Data-driven begins with DATA; potential of data assets. *Journal of Computer Information Systems*, 62(1), 29-38.
- [8] Hu, C., Li, Y., & Zheng, X. (2022). Data assets, information uses, and operational efficiency. *Applied Economics*, 1-14.
- [9] Zhang, Y., Padmanabhan, A., Gross, J. J., et al. (2019). Development of human emotion circuits investigated using a big-data analytic approach: Stability, reliability, and robustness. *Journal of Neuroscience*, 39(36), 7155-7172.
- [10] Lyu, Y., Yuan, H., & Shu, P. (2003). On the accounting recognition and measurement of digital assets. *Journal of Central University of Finance & Economics*, (11), 62-65.
- [11] Liu, Y. (2014). A brief discussion on the recognition and measurement of big data assets. *Commercial Accounting*, (18), 3-4.
- [12] Zou, Z. (2018). Several thoughts on the valuation of big data assets. *Communication of Finance and Accounting*, (28), 35-39.
- [13] Li, Z., & Tan, X. (2018). Accounting recognition, measurement and reporting of big data assets. *Communication of Finance and Accounting*, (10), 58-59+129.
- [14] Li, M. (2019). Research on the valuation method and application of big data assets based on real options [Doctoral dissertation]. Yunnan University.
- [15] Ma, D., & Yu, X. (2020). Data assets: Conceptual evolution and measurement methods. *Journal of Statistics*, 1(2), 15-24.
- [16] Li, Y., Liu, Y., & Li, B. (2022). Analysis of several theoretical issues in data asset accounting. *Statistical Research*, 39(9), 19-28.

- [17] Hu, Y., & Xu, X. (2022). Research on the statistical measurement of enterprise data asset value. *Statistical Research*, 39(9), 3-18.
- [18] Gao, H., & Jiang, C. (2022). Valuation of data assets from the perspective of application scenarios. *Finance and Accounting Monthly*, (17), 99-104.
- [19] Sun, Y., & Yang, J. (2022). Research on accounting issues of data assets. *Friends of Accounting*, (16), 153-160.
- [20] Wang, M. (2022). Research on accounting treatment of data assets. *International Business Accounting*, (13), 55-58+63.