Proceedings of Business and Economic Studies, 2022, Volume 5, Issue 3
http://ojs.bbwpublisher.com/index.php/PBES
ISSN Online: 2209-265X
ISSN Print: 2209-2641

Conflicts in Defining Data Ownership in China's Digital Transformation and the Regulatory Path

Zihan Si[1]*, Nuowen Mai[1], Yifei Quan[1], Yihuan Zhou[1], Xiaoyue Zhang[2]

[1]Jiangnan University, Wuxi 214122, Jiangsu Province, China

[2]Nanjing University of Information Science and Technology, Nanjing 210044, China

***Corresponding author:** Zihan Si, violetsi0111@gmail.com

Abstract: China is aggressively pursuing digital transformation, and data, alike labor, technology, capital, and knowledge, has become as a crucial factor of production. Digital transformation is accelerating the emergence of a data-intensive society, and the ensuing difficulties of balancing freedom and responsibility, openness and security, as well as free sharing and legal regulation are posing new challenges to national and social governance. Among these challenges, defining data ownership, the social disorder and anomie brought about by the unclear definition of data ownership, and data ownership regulatory path are new propositions that need to be urgently addressed in this data-intensive society. This paper systematically explains the theoretical meaning and practical value of data ownership through a literature review on the analysis of domestic and foreign scholars as well as research think tanks, compares the differences and inherent conflicts between the definition of data ownership by the government, enterprises, and society in China, thoroughly compares the definition standards of the European Union, the United States, and Japan, and on this basis, discusses the formation of a definition of data ownership that meets the requirements of China's digital transformation.

Keywords: Digital transformation; Data ownership; Smart cities; Risk regulation

Online publication: June 20, 2022

1. Introduction

The Opinions of the State Council of the Central Committee of the Chinese Communist Party on Building a More Comprehensive Institutional Mechanism for Market-Based Allocation of Production Factors, released on April 9, 2020, specifies data as a new type of production factor, which constitutes a factor market along with land, labor, capital, and technology. The Opinions proposed to build a data factor market for data trading and sharing as well as data property rights protection, promote open sharing of government data, enhance social data resource value, and strengthen the integration and security protection of data resources [1]. As revealed in the past decade, the Chinese society is inexorably entering a data-intensive society, and data has become a necessary, fundamental, and ubiquitous factor of production. The ensuing difficulties of balancing freedom and responsibility, openness and security, as well as free sharing and legal regulation are posing new challenges to national and social governance. In the current environment where legislation is lacking and data ethics are urgently needed, it is critical to investigate the ongoing and potential misconduct and regulation of big data. Among them, how to regulate data ownership and thus safely share personal data to promote the development of national digital economy is a new proposition in the era of big data. These challenges necessitate continuous updates to big data and relevant legal systems. As the pace of development of both entities is not fully synchronized, it is crucial to continuously improve

the data literacy of citizens, enterprises, and the government, as well as raise the awareness of data ethics.

2. Connotation of data ownership in a data-intensive society

2.1. Data-intensive society and data

From the digitization of information to the digitization of production and business models, more and more companies and government agencies in China are taking an active part in the trend of digital transformation. At the same time, they are exploring the mode and path to achieve the digitization of industry and industrial chain as well as accelerating social transition toward a data-intensive society. This paper argues that the data-intensive society, as a new form of society, possesses several characteristics: (1) digital technologies activate application scenarios; (2) data have become a necessary, basic, and ubiquitous factor of production; (3) machines are intelligent; (4) data have become a kind of power. This shows that data play multiple roles in a data-intensive society, which indicates that the meaning of data has changed.

In the internet era, people tend to define data in terms of the process of interpreting and applying data, considering data as the expression and carrier of information. In this case, information is the connotation of data, and data themselves carry no meaning. Data are merely computer symbols that can be regarded as information when they influence the behavior of entities [2]. However, viewing data from the perspective of data generation and dissemination, data originate from the conscious collection of information and are "non-neutral" [3]; information can only become data by combining with other information [4]. Although the above arguments are reasonable, they still fall far short of encapsulating the unique role of data in a data-intensive society. In digital transformation, data are not only carriers of information but also vital means of production, which carry value by participating in all aspects of social production and reproduction. When data are considered manifestations of information or a collection of information, then data will be endless; if there is no scarcity, then the discussion of data ownership is meaningless [5].

Therefore, this paper argues that in the context of digital transformation, which accelerates the formation of a data-intensive society, we should analyze and explore the meaning of data from the political economy perspective. In labor production, data are products of labor, forms of information, and the results of summarizing, calculating, and processing fragmented information. In social reproduction, data can promote value addition in the reproduction process and combine with various aspects of product creation and knowledge development as the outcome of labor. Meanwhile, enterprises can allocate resources and target promotion based on market data, and then promote capital circulation [6]. In terms of importance, it is not an exaggeration to compare data to the "oil of the new era." However, unlike oil, data do not exist naturally and can be extracted at will. The producers of data include not only the government, enterprises, social organizations, and so on, but also individual citizens. At present, personal data are not well-protected and are at risk of being misused. In the long run, such a reality will lead to a decline in the autonomy and motivation of data producers to build and improve the data-intensive society, thus posing a potential threat to the establishment and maintenance of the data-intensive society. This paper believes that regulating data ownership can prevent such problems to a certain extent.

2.2. Definition of data ownership

There is no unified conclusion on the definition of data ownership in existing literatures and reports, in which literatures and reports from Mainland China and other countries hold distinct perspectives. This section discusses the meaning of data ownership based on existing literatures.

Relevant studies from China mainly focus on data ownership from the perspectives of individual data rights, public data rights, and multi-level data ownership. Among them, the most mentioned view is defining data ownership in terms of its property attributes from the perspective of the individual data rights theory. Personal data have the attributes of property rights; individual data ownership means that the user

has the right to possess, use, benefit, and dispose of personal data according to legislation [7]. This view argues that data are generated by individuals' online activities, while companies collect and process these data, in which they are analyzed secondarily by companies using algorithms to produce marketing effects for users themselves. These data then exist in the network and have a certain level of interactivity. Eventually, these data have real-world property attributes. The first three steps reflect the basic features of personal data as property, while the last step reflects the value attribute of personal data as property. Personal data are essentially data codes, and the core of data code is personal information, which generates value due to its interaction with the real world. Therefore, personal data are valuable data codes. In view of their value attributes, these data codes can be regarded as property.

In addition to the property rights theory, the scope of individual data ownership also includes the personality rights theory and the privacy rights theory [8]. The personality rights theory holds that data contain personal information, and the utilization of personal information is related to a person's dignity; thus, data ownership should be discussed and protected under the range of personality rights. The privacy rights theory also holds that data contain personal information, but it emphasizes that personal information is part of a person's privacy and should be discussed in the context of privacy protection. Contrary to the explanation offered by the individual data rights theory, the public data rights theory holds that personal information in data not only has personality and property characteristics, but also public and social characteristics. Hence, personal information protection should transition from individual control to social control. In the era of big data, personal information should be regarded as public goods and protected by public law rather than private law. Besides the definition of private rights theory and public rights theory, there exists the definition of multi-level data ownership. This view holds that data ownership should be determined based on specific events and conditions. **Table 1** summarizes the three main doctrines discussed above.

Table 1. Doctrines related to data ownership

Individual data rights	(1) The personality rights theory holds that individuals have a "right to information self-determination" and a right to decide whether personal information can be collected and used by others. (2) The privacy rights theory states that individuals should have control over and access to data. (3) The property rights theory holds that individuals have the right to possess, use, benefit, and dispose of data, including the right to dispose of the commercial use value embedded in their personal information.
Public data rights	Personal information protection should transition from individual control to social control; personal information should be regarded as public goods; public law governance instead of private law governance.
Multi-level data ownership	The various configurations of personal information ownership will be implemented based on specific events and conditions.

Compared with domestic studies that explored the connotation of data ownership from the perspective of private rights and public rights, foreign literatures tended to determine data rights owned by corresponding subjects through the division of data-involved subjects. **Table 2** offers a brief view on several representative research topics from foreign literatures. The subjects involved can be divided into owners, agents, managers, and operators. Different subjects enjoy different rights to process data, thus reflecting the different degrees of data ownership.

Table 2. Foreign studies on the classification of data ownership

Research topics	Data ownership subjects	Data types	Usages
Web data ownership study	End-users, enterprises, and public service organizations	Personal data from devices and software	Store, analyze, integrate, productize, and consume
A study of big data, databases, and ownership issues in the cloud	Cloud providers and end-users	Consumer data, biological data, public cloud data, and private cloud	To help the software work properly, and not to instruct or inform someone
First Nations Principles of OCAP (Canada)	The sample is a continuing property of the donor or community and is on loan to the researcher, who is obligated to be a steward of the sample and may hold it only if the sample provider agrees to the purpose of its use	Data of indigenous people	Possession, access, and control
IBM Multiple Virtual Storage (MVS) system	Owner, agent, and custodian	Environment development program library and environment development system	Operating environment parameters
An empirical study of data access, ownership, and control of access practices	Journals, funding agencies, and professional or industry associations	Scientific data, including scientific priorities, civil litigation, and commercial or national security interests	Translations and conversions

Since there is no clear definition of data ownership, several concepts similar to data ownership exist. These concepts are often confused with data ownership or are variants of data ownership, as listed in **Table 3**.

Table 3. Concepts related to data ownership

Data industry rights	Based on operations such as collection and analytical processing, the owner or long-term user of the equipment (such as a lessee) has the right to use and license non-personal data to others as well as to prevent unauthorized use of and access to the data by others. When it comes to public interests such as traffic management and environmental management, data producers should not have exclusive access to data.
Data personality rights	It is a new type of personality right. Individuals can refuse to provide personal data, they can monitor the processing of personal data by personal data processors, or they can even stop the in-depth analysis and processing of personal data. Individuals have the right to informed consent to personal data and the right to rectification as well as to be forgotten (deleted) when personal data may negatively affect them due to inaccuracy, incompleteness, lagging, and other issues.
Right to informed consent for data	Service providers (or governments) are required to inform data subjects and obtain their consent before collecting or processing personal data. It consists of the right to data knowledge and the right to data consent.
Right to modify data	The right of the data subject to enjoy or authorize others to modify his or her data.

(Continued on next page)

(Continued from previous page)

Right to be forgotten	Ambrose Meg Leta and other researchers argued that the right to be forgotten in the EU Data Protection Regulation (draft) includes both the traditional right to be forgotten, which refers to the right to the complete deletion of certain publicly available data, and the right to erasure, which addresses the problems posed by search engines and stems from the protection of reputation, identity, and personality; the latter refers to the right to the deletion of personal data used for automated processing and aims to offer individuals more effective control over their data [9].
Data property rights	It is a new type of property right that belongs collectively to property rights, and it is a property right alongside intellectual property rights, property rights, claims, and others.
Data collection rights	The data subject has the right to consent to or prohibit the collection of his or her data.
Data portability rights	The data subject has the right to request that the relevant party (service provider, operator government, and other parties) in possession of his or her data assist him or her in the migration and preservation of personal data between different systems or carriers, such as equipment.
Data access rights	The right of data subjects to use their data. The ability to access, create, standardize, and modify data, as well as all intervening privileges.
Data revenue rights	The right of the data subject to receive benefits based on his or her data.
Data ownership	The right to dominate, dispose of, and benefit from the property of the relevant data.

In short, this paper defines data ownership as the possession, access, storage, analysis, control, productization, and even consumption of relevant data for multiple subjects, such as suppliers and owners, and divides it into private, public, and multi-level ownership in terms of power.

2.3. Necessity of defining data ownership

Firstly, there is a lack of systematic laws. Currently, there is no systematic regulation of ownership in big data transactions at both, national level and sector legislation level in China. In practice, it mainly relies on the regulations of data trading platforms and industry self-regulation to guarantee transactions. First, under the existing legal system, it is difficult for property rights, intellectual property rights, and claims to serve as the basic rights of big data transactions. Second, data have the characteristics of intangibility and low reproduction cost, and their exchange is different from the general commodity transaction method. For this new way of transaction, the law does not clarify provisions on the subject qualification of the trading platform and the transaction process. Objectively, this has caused the low threshold of platforms and a non-programmed transaction process.

Secondly, the fundamental regulations are lagging. China currently lacks specific legal provisions on data ownership and does not legally clarify the attribution of data ownership. For rulings involving personal data protection, judicial practices often invoke relevant provisions on portrait rights, reputation rights or privacy rights, personality rights, and others. Enterprises and other organizations mainly rely on the terms of intellectual property rights, including copyrights and trade secrets, to enjoy the corresponding rights. These may result in multiple rights attributes on the same data, and these rights may contradict. In case of contradictions, the law does not explicitly stipulate the priority of the rights subjects, which may lead to increased risks of data transactions and hinder the process. Moreover, under the current legal system, privacy rights, personality rights, property rights, intellectual property rights, and claims cannot be used as a comprehensive source of fundamental rights for data transactions, nor can they independently solve the problem of data ownership in data transactions. In addition, the use of big data products in these cases lacks legality. Most individuals readily give the necessary consent to use digital services without giving much thought to their legally defended interests in data protection. Given this willingness, and especially the fact that such rights can be enforced by any company capable of providing the services required by individuals

in the digital economy, there are concerns about the practical consequences of creating a new "data ownership." In particular, market players capable of influencing user behavior may become even more powerful in view of this new right, such as global social networks and internet search engines.

Thirdly, poorly defined data ownership can lead to data silos. According to the American Health Information Management Association, 10% of hospital medical records in the United States is repetitive. Data will eventually become isolated in respective data silos carefully guarded by individual business lines or departments. Information technology (IT) departments have pointed out that some data held by marketing departments overlap those in sales, and this silo mentality has been the impetus for many major IT initiatives. Enterprise resource planning (ERP) is supposed to be a powerful tool for eliminating data silos, but the ERP systems in many companies have instead become silos themselves. At the same time, it is increasingly difficult to create data warehouses that keep up with the times and respond to changes in businesses; additionally, data management initiatives are hampered by acute internal political issues and the reluctance to relinquish control over data by members of the organization. Data silos make it difficult for organizations to grasp a full picture of their overall operations and thus to make effective decisions.

Fourth, data ownership boundaries are dynamic. Mainland China does not provide independent data rights and information property rights. Individuals often enjoy the right to control data based on legal rights, such as privacy rights. Organizations mainly have corresponding rights to data based on compilation rights, copyrights, and trade secrets. This problem leads to the possibility of multiple rights conflicting with each other on the same data set. Which right to prioritize is not stipulated in the law, and which right will increase the risk of data transactions and hinder the data transaction process is unclear. The personality rights attribute of personal information makes it possible to retain its rights after the data have been exchanged several times, and data buyers have to bear higher transaction risks.

Fifth, it is difficult to determine the infringement liability, leading to a high cost of data trading. In the age of the internet and big data, it is difficult to obtain effective and timely remedies for the corresponding rights because of technical reasons, which raises the transaction risks and costs. It is difficult to obtain adequate protection for the corresponding rights in big data transactions. From the perspective of data sellers, data sources and data collection subjects show diversity. This diversity makes it challenging to monitor and regulate the sources and subjects; additionally, the boundary between private and public spheres begins to blur. From the perspective of data buyers, big data mining is highly likely to violate the privacy of individuals, and even the infringers themselves do not know whether their privacy has been violated or not. One example is predictive advertising based on big data. The subjects may not know that advertisers push specific ads to them based on their previous data. Even if reverse identification technology is adopted, it is difficult to fully protect the privacy of individuals. In addition, in view of the intangible nature of data, it is easy for data to be copied and disseminated on the internet; it is difficult to eliminate infringement, and even more challenging to protect the ownership, which will undermine the value of data transactions. Finally, existing big data storage and analysis are mostly carried out in the cloud. Cloud services or network providers may apply the "haven" principle and take no responsibility on the ground of ignorance.

Sixth, the single ownership division cannot adapt to the development of the big data era. If data ownership is uniformly attributed to individuals, it will result in excessive transaction costs and bring risks to the processors of the data industry for collection and use. This is because after the anonymization process, a database collection will be formed from the raw data initially collected for data analysis. When these data generate enough commercial value, it will still be considered belonging to those individuals and thus increase the costs of data transactions. At the same time, the anonymized data formed through splitting, reorganization, and integration still belong to respective individuals, indicating that the re-conduct of the data processing act needs to solicit the ownership issue arising from the same transaction of the data subject.

3. Conflicts in defining data ownership

3.1. Defining data ownership at the government level

Since government-controlled data are associated with national interests and public interests, aiming at increasing social welfare, the corresponding property rights should be owned by all people. However, "universal" is not a definite subject, and the government bears the cost of data property rights that individuals do not have to bear. Therefore, it is necessary to regulate the use of public data by market players through relevant laws to ensure the utilization of public data.

Meanwhile, local governments have similar definitions for "public data." Article 3, Paragraph 1 of the Interim Measures for the Opening of Public Data in Shanghai [10] stipulates, "Public data referred to in these measures include all kinds of data resources collected and generated by administrative organs at all levels and institutions performing public management and service functions (hereinafter, collectively referred to as public management and service institutions) in the course of performing their duties by the law." Article 3, Paragraph 1 of the Interim Measures for the Open Management of Public Data in Chongqing [11] also stipulates, "Public data refer to all kinds of data resources generated, collected, and produced by public management and service institutions in the course of performing their duties under the law, and recorded and preserved in a certain form." It can be seen that when personal data are collected by the government, they are converted into public data, and the government has the right to use these data.

The state defines individual data ownership by law with property attribute as the most central and essential attribute. Individual ownership, with property attribute as its essential, is as configurable as traditional factors of production. Although personhood is a special attribute of data compared to traditional factors of production, it, being a relatively secondary attribute, should not be a hindrance to the utilization of property attributes of data. That is to say, the right to property interests should not be restricted on the grounds of protecting personality interests in terms of data [12]. Therefore, on the issue of the initial arrangement of data property rights, the law cannot simply divide data property rights to natural persons as data subjects only because of the consideration of personality interest protection.

3.2. Defining data ownership at the enterprise level

Enterprises, as data industrialists, are crucial investors in the data industry and also the subjects of industrial practice that convert data from massive disorderly resources into data assets with commercial value. Thus, enterprises will definitely consider data as a factor of production. When setting industry norms, enterprise data property rights should balance personality interests and property interests, in order to make a reasonable division of personal data ownership.

The Self-Regulation Convention on Data Circulation Industry (hereinafter, referred to as "the Convention"), jointly drafted by the Internet Law Research Center of China Academy of Information and Communications Technology and the Data Center Alliance, stipulates in Article 7, of Chapter 3 "Data Circulation" that enterprises "shall obtain the consent of users or conduct necessary desensitization if personal privacy is involved, and the data shall be desensitized as necessary; data related to national security and public security shall not be shared without legal authorization"; Article 8 emphasizes that parties of data circulation shall enter into contracts to clarify the rights and obligations of each party and ensure the orderly transfer of responsibilities of different entities in data circulation, thus encouraging the use of model data circulation contracts [13]. The introduction of similar norms clarifies the ownership of big data resources, which can effectively protect the value of data assets, establish a penalty mechanism for unfavorable data security protection and leakage of privacy, effectively promote industry alliances and associations at all levels for industry self-regulation, guide enterprises in data circulation to maintain data quality, improve data availability, maintain good order, as well as enhance user trust.

3.3. Defining data ownership at the individual level

Private data aim to protect the freedom of individuals to withdraw from the public arena, which involves private life unrelated to the public interests of society and the interests of social groups. The distinction between public and private spheres is the core of the construction of privacy laws. Personal information carries the personality interests of the information subject, which is related to human freedom and dignity, and the information subject requires the law to protect such personality interests, instead of considering personal information in a purely materialistic way. This view of information subject conflicts with the way the government and enterprises define the property of personal data.

According to an industry survey conducted by iiMedia Consulting, a leading data mining and integrated marketing agency in China, 65.3% of the interviewed netizens believe that reading call records is an invasion of privacy, while 55.4% believe that reading contact information is an invasion of privacy. The public's perception is that call records, contact information, and text messages contain personal information. Additionally, they pay more attention to the security of information while using applications. However, the same report shows that 57.8% of the interviewed netizens agreed to grant part of the permissions to applications after discerning; 30.3% of those interviewed would stop using an application because they refuse to comply to the application's forced permissions, indicating that netizens tend to be proactive in discerning application permissions, but in the face of excessive application invocation, they still need to improve their awareness of active prevention and further identify as well as firmly reject the illegal invocation of application permissions [14].

3.4. Conflicts in defining data ownership

The first is the conflict between commercial, public, and private use of data, which is the conflict between the data dependence of enterprises and governments and the rights of individuals. In order to function properly, enterprises need to expand their business scope into offline physical space and connect more devices, so they have absolute reliance on user data. For example, in data-driven businesses, the General Data Protection Regulation (GDPR)'s tight protection of the nature of personal data can lead to incomplete access to data and the inability to provide more accurate services, thus reducing the attractiveness to users and the return on capital of the industry, hindering the development of the industry, limiting technological innovation, as well as adversely affecting market competition.

The second is the conflict between the right to leave data and the right to delete, as well as the conflict of the right to delete data of each subject. Personal data are generated by users, but their carrier expressions and even storage are subjected to the interference of enterprise platforms, and their absoluteness is compromised. In addition, most of the personal data are stored in the servers of enterprise platforms; even if users' data in its tools are deleted, the data on the terminal may not be deleted, thus compromising the right of disposal. Although the ownership of personal data lacks the relevant features of traditional ownership, the right to know and the right to delete under relevant laws such as the EU's GDPR can help complete the property attributes of personal data ownership. The concept of personal data ownership focuses on users' rights to possess, use, benefit, and dispose of personal data according to the law [7].

The third is the conflict between data valuation and data identification. Differences in the valuation of personal data lead to conflicts in the definition of personal data ownership among subjects. Data in use add to the difficulty of valuation because of their uniqueness: (1) the use of data does not lead to data consumption, and data exploitation is not exclusive, in which its behavior may even be altruistic; (2) aggregate data are more valuable than individual data and should be more expensive; (3) different data and data from different sources have different values, and this is especially true for medical data; (4) the same data has different values for different users [7]. Most companies will classify data as core assets and adopt confidentiality measures, so the method of data property valuation by negotiated bargaining lacks feasibility.

4. Normative path to data ownership

At present, many countries and organizations have introduced relevant laws and regulations to define data ownership more strictly, reflecting the real needs and value orientation of different countries and regions. Among them, the more representative definition models include the EU's human rights priority model, the United States' data marketization model, and Japan's public-private consultation model. With the rapid development of data circulation and digital economy, China needs to learn from the experiences of existing models and explore its own model of data ownership, while taking into account its national conditions.

4.1. Data ownership definition models: a country-by-country comparison

The EU has defined data ownership from the perspective of "adequate protection of the interests of data subjects," which ensures that the data of citizens are respected and protected, reduces the risk of data misconduct, as well as ensures the security of digital economy within the country, but on the other hand, in today's data-enabled economic environment, it remains to be investigated whether giving data subjects full ownership and imposing strict restrictions on data controllers, processors, and users will affect the sharing, use, and circulation of data as well as the vitality of the data industry [15].

Concerning the use of data, the U.S. has adopted a very different model than the EU. The U.S. is more concerned about the "economic characteristics and personal values of personal data" and actively promotes the marketization of data while ensuring the right to privacy. If the openness and ambiguity of personality rights impose a greater burden on enterprises, which is not conducive to the innovation of micro and small enterprises, then the transaction rules based on consent and property rights give enterprises more confidence in mining and using data, allowing data and information to play a greater role in market allocation [17]. However, the issue of personal privacy protection may arise when data proselytization becomes the norm. Although the law gives the option of data property rights to individual citizens, it is often difficult for individuals to attain protection of their privacy rights and interests when the power of a single data subject and that of oligopolistic data companies differ. For example, in 2018, Facebook Inc.'s third-party company Cambridge Analytica was revealed to have stolen the personal information of over 50 million users in order to push personal news based on algorithmic technology to influence citizens' judgments on the U.S. election and the United Kingdom's exit from the European Union, sparking a public debate on the boundaries of personal information utilization [18].

Under the influence of the EU model and the U.S. model, the Japanese legislative model can be described as a compromise between the former two models, with its characteristics. It adopts a combination of unification and division as well as encourages industry self-regulation while unifying legislation, finding a balance between personal information protection and personal information flow. The Personal Information Protection Law adopts a rule similar to the "opt-out" policy, where personal information is collected and used by default, but if the individual objects, the former processes are discontinued. Without doubt, it also has disadvantages and loopholes, such as the lack of legal protection of the right to informed consent of personal information [16]. **Table 4** compares the definitions of data ownership by the EU, U.S., and Japan.

Table 4. Comparison of data ownership definitions by the EU, U.S., and Japan

Country	Representative laws	Features	Core of data ownership definition criteria
EU	GDPR	Five major rights are granted to data subjects: the right to know, the right to access, the right to object, the right to portability of personal data, and the right to be forgotten A harmonized legislative model was adopted with the establishment of the EU Data Protection Board	"Human Rights Priority Model" Strict control of data controllers, processors, and users with "personality protection" as the core
U.S.	California Consumer Privacy Act (CCPA) and others	Adopting a decentralized legislative model with federal legislation by industry and no specific code to regulate on the issue; adopting an industry self-regulatory model in regulation	"Data Marketization Model" Based on the right to privacy, with the protection of "the economic characteristics and personal values of personal data" as the core, pursuing the principle of combining self-regulation of data users with individual self-help remedies
Japan	Personal Information Protection Act and others	Uniform legislation while encouraging industry self-regulation; third-party independent agency monitoring	"Public-Private Consultation Model" Incorporates the EU model and the U.S. model

4.2. A model for defining data ownership that fits the Chinese context

Three typical data protection legislative models in the international arena have certain implications for defining data ownership in China. The definition of data ownership in China is not an option between "human rights" and "data marketization," but rather a question about how to better integrate personal data rights and data market circulation. The logic behind it should be based on protecting the privacy and personality rights of individual citizens, responding to the development trend of data technology and industry, taking into account certain rights and interests of data users and processors, such as enterprises, as well as meeting the needs of national political security and economic development. Therefore, this paper argues that the model of data ownership that meets China's national conditions should be a government-led multi-party governance model. **Figure 1** gives a clear illustration of how different parties interact under the "Data Rights and Responsibilities Alliance" scheme.

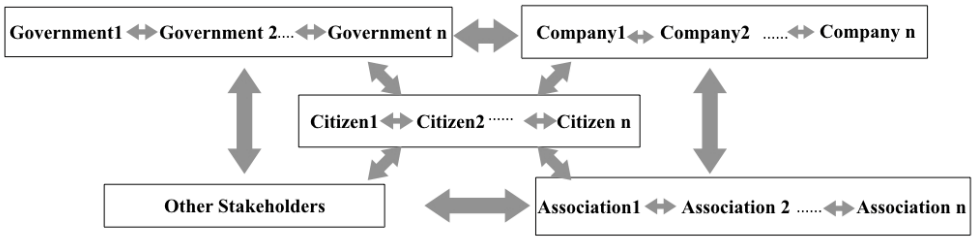

Figure 1. "Data Rights and Responsibilities Alliance" regulatory relationship

At present, China's digital industry regulations are still in the early stages of development; thus, the government must take the lead to ensure that the management, circulation, transaction, and application of data are in accordance with the regulations. The complex intertwining and potential conflicts of interests of data-related subjects necessitate a multi-party participatory process in defining data ownership. Governance

is different from management and ownership, thus reflecting the diversity of participating subjects and the central idea that the definition of data ownership in China should be people-oriented. At the same time, governance reflects that the definition model of data ownership in China should not be static, but rather a dynamic and evolving governance process. The ultimate goal of governance, that is, the ultimate goal of data ownership definition, is to create an environment conducive to the healthy flow of data and the healthy development of digital economy as well as to build a digital-intensive society that serves people.

In a nutshell, this government-led multi-party governance model for defining data ownership needs to follow the following principles: (1) the people-oriented principle; (2) the principle of building a relevant legal system as a basis; (3) the principle of multiple-subject participation (such as national government agencies, local government agencies, enterprise groups, commercial organizations, supervisory bodies, individuals, and other subjects); (4) the principle of respecting the specificity of different industry norms. This paper provides suggestions on how to construct this model.

The first step is to clarify the unity of rights and responsibilities. Before formally proposing this paper on the regulatory path of conflicting data ownership definitions, it is important to clarify that in the process of regulating data ownership definitions, it is not only necessary to define the ownership of rights, but also the responsibilities and obligations of the rights holders. The majority of current studies on data rights focus on defining which subjects are entitled to which rights of data, but little is said about the responsibilities of data rights holders. This paper argues that the basic responsibilities of rights holders include data care and digital literacy enhancement. The former refers to the necessary protection and maintenance of data to which the rights are entitled and the storage devices for such data, and the right holder should be allowed to inform and consult with other stakeholders if the disposal of the data may have a significant social impact or profoundly affect the data. In addition, whether they are governments, companies, organizations, or individuals, rights holders should also improve their digital literacy.

Secondly, data ownership should be regulated through multiple paths. The regulation of data ownership involves multiple stakeholders and touches all aspects of production and life in the era of big data, so its regulatory path cannot be linear, but rather a grid-based multi-path. The multi-path data ownership regulation consists of three parts.

First, building a coalition of data rights and responsibilities norms. It is important to ensure that stakeholders such as government, enterprises, public organizations, industry associations, and citizens can participate equally and in a balanced manner in the formulation of data ownership definition norms. A coalition can be established to conduct extensive, in-depth, and focused research and opinion collection, and the organization of the coalition can take the form of an online-offline combination. By establishing a coalition for data ownership regulation, the channels for feedback by stakeholders can be made more accessible. In the past, public organizations, industry associations, citizens, and other subjects often faced problems, such as numerous steps, complicated processes, and inability to attain feedback, which discouraged their enthusiasm and engagement in feedback. On the other hand, forming an alliance can help establish a multi-party supervision mechanism for data regulation. At present, the main regulatory bodies for data are the government and third-party regulators, both of which have their advantages and disadvantages. The reason we believe a multi-party mutual supervision mechanism should be established is that the supervision of a single subject is likely to cause trust issues, and these issues will in turn lead to the resistance of the regulated subject toward the supervision model. This paper believes that a mutual supervision mechanism can minimize trust issues and ensure fairness and justice to the greatest extent. The multiple parties here include both the mutual supervision of multiple subjects and that within the subjects. The scope covered by the regulation will be wider and more efficient with more participants. It should be noted that the establishment of such a monitoring mechanism also requires a sound legal system and a high level of citizen data literacy as prerequisite guarantees.

Second, refining data ownership based on scenarios. Refinement here has two meanings: one is the refinement of the scenario, and the other is the refinement of data ownership. In the context of digital transformation, data is circulating in the market as a production factor, so a focused definition of data ownership paired with specific market behaviors makes the definition criteria more objective and biased toward neutrality. The scenarios here can be divided into two main categories: those in which personal data are the objects of transaction, and those in which the data are anonymized on this basis [19]. In this context, different industries, such as medical, agricultural, and public administration, can further refine the scenarios in view of the specificity of the industry. Based on the refinement of the scenario, a certain split of data ownership is required, when necessary, where the right holder only has partial ownership, while the remaining can be held by other subjects. In Europe and the United States, the GDPR does not assign specific data ownership, but rather refine it into specific data rights, the definition, and regulation of it. It can be seen that the refinement of data ownership based on scenarios needs to be supported by ideal laws and industry regulations. Different industries form their own data ownership and use norms according to their industrial characteristics. A sound legal system of data ownership should have both, a unified and divided structure, with overall norms and standards at the national level, while leaving some room for industries and localities to further improve norms and standards in their industries or regions according to their characteristics. In this way, it also gives the market flexibility and promotes a more efficient flow of data.

Thirdly, developing algorithms to decide the corresponding rights and responsibilities attribution. Alstyne and other researchers have used mathematical formulas based on the theory of incomplete contracts to calculate how defining the ownership of database data can maximize data utilization and organizational benefits [20]. Several local scholars have also proposed that the protection of data ownership can be achieved using centralized modeling and block chain algorithms [21]. The starting point of both is the maximization of organizational benefits. It can be seen that the specification of data ownership employing algorithms is feasible and optimal decisions can be made by rigorous logic and calculation. The development of relevant algorithms should take into account the refinement scenario, the refinement of data ownership, how to assign refined rights in refined scenarios, as well as the tangible and intangible elements of influence. An example of such a formula is demonstrated in **Figure 2**.

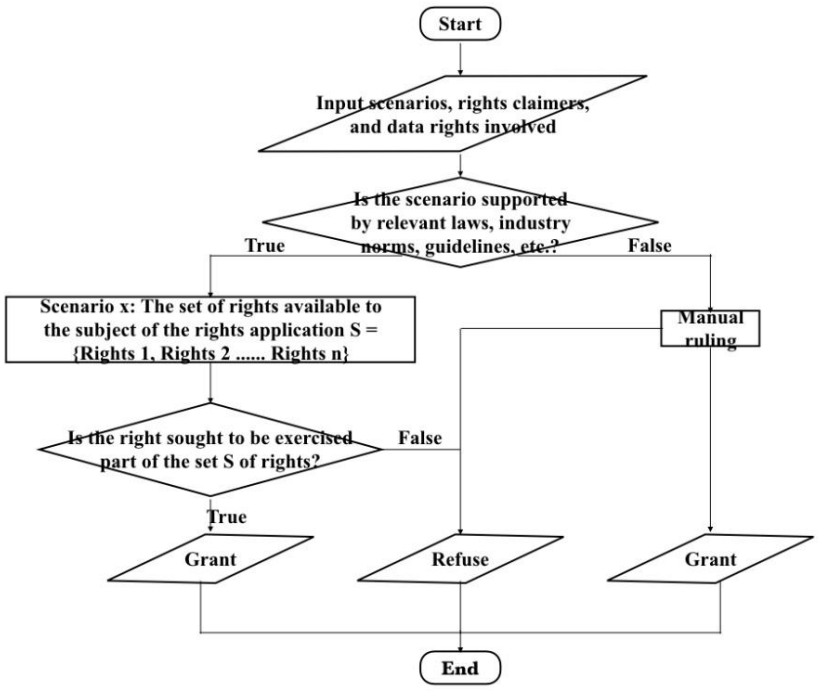

Figure 2. Flow chart on whether to grant the corresponding data rights

5. Conclusion

In the era of big data, the relationship between "people" and data in digital transformation is a cliché and is slow to achieve true harmony. Improving the definition and regulation of data ownership is a crucial step in achieving a harmonious relationship between the two, but this process of improvement is complicated, thus requiring the joint efforts of enterprises, individuals, and the government to strengthen the awareness of people as well as a sound legal system based on different scenarios as a guarantee. While paying attention to data technological innovation, creating a better social environment for big data to provide a better breeding ground for technological innovation is a concern that should be addressed. We hope that in the future, the government, enterprises, and citizens can shift their focus from data to people, strike a balance between the two, and use the soft power of big data to further promote digital transformation.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Pang B, 2020, Opinions of the State Council of the Central Committee of the Communist Party of China on Building a More Perfect Institutional Mechanism for Market-Based Allocation of Factors. www.gov.cn. http://www.gov.cn/zhengce/2020-04/09/content_5500622.htm

[2] Zhou Y, Li Y, Cui K, et al., 2013, Database Principles and Development Applications (2nd Edition), Tsinghua University Press, Beijing, China, 1–2.

[3] Scassa T, 2018, Data Ownership. CIGI Papers No. 187, Ottawa Faculty of Law Working Paper No. 2018-26. http://dx.doi.org/10.2139/ssrn.3251542

[4] Rees C, 2014, Who Owns Our Data?. Computer Law & Security Review, 30(1): 75–79.

[5] Bastani A, 2019, Fully Automated Luxury Communism, Verso Books, London, UK, 29–31.

[6] Liu H, 2021, Data Hegemony and the New Type of Digital Imperialist Plunder. Contemporary Economic Studies, 306(2): 25–32.

[7] Chen L, 2020, A Study on the Concept of Personal Data Ownership System and the Basic Establishment Framework. China Business Journal, 2020(20): 176–178.

[8] Lee J, 2019, A Legal and Economic Analysis of the Ownership of Personal Data. Journal of Chongqing College of Arts and Sciences (Social Science Edition), 38(6): 114–122.

[9] Ambrose ML, Ausloos J, 2013, The Right to be Forgotten Across the Pond. Journal of Information Policy, 3(1): 1–23.

[10] Ying Y, 2019, Interim Measures for the Opening of Public Data in Shanghai (Shanghai Government Order No. 21). Shanghai Municipal People's Government. https://www.shanghai.gov.cn/nw48156/20200825/0001-48156_62825.html

[11] Chongqing Municipal People's Government General Office, 2020, Interim Measures for the Open Management of Public Data in Chongqing (Chongqing Government Order No. 111). Chongqing Municipal People's Government. http://www.cq.gov.cn/zwgk/zfxxgkml/szfwj/xzgfxwj/szfbgt/202009/t20200918_8837781.html

[12] Huang Y, 2020, Legal Protection and Restriction of Data Property Rights from the Positioning of Factors of Production, South China University of Technology.

[13] 2016, Data Circulation Industry Self-Regulation Convention Version 2.0 was Officially Released in Beijing. Sohu. https://www.sohu.com/a/103057203_162886

[14] iiMedia Report, 2020, 2020 China Mobile App Privacy Permissions Measurement Report. iiMedia Research. https://www.iimedia.cn/c1061/69301.html

[15] 2018, European Commission General Data Protection. Intersoft Consulting. https://gdpr-info.eu/

[16] Galvin HK, DeMuro PR, 2020, Developments in Privacy and Data Ownership in Mobile Health Technologies, 2016-2019. Yearbook of Medical Informatics, 29(1): 32–43.

[17] Zhao H, 2021, A Study of the Differences in Data Governance Paths Between the US and Europe in the TikTok Controversy. Journal of Intelligence, 40(5): 104–110 + 131.

[18] U.S. Department of Commerce, EU-U.S. Privacy Shield Framework Principles Issued by The U.S. Department of Commerce. Privacy Shield Framework. https://www.privacyshield.gov/servlet/servlet.FileDownload?file=015t00000004qAg

[19] Wang R, 2015, An Exploration of Data Ownership: The Core Legal Issue of Big Data Transactions. Big Data, 1(2): 41–47.

[20] Van Alstyne M, Brynjolfsson E, Madnick S, 1995, Why Not One Big Database? Principles for Data Ownership. Decision Support Systems, 15(4): 267–284.

[21] Wu C, Yu J, 2020, Exploration of Data Ownership Protection, Pricing and Distributed Application Mechanism for Public Management. E-Government, 2020(1): 29–38.