Advances in Politics and Economics
ISSN 2576-1382 (Print) ISSN 2576-1390 (Online)
Vol. 7, No. 2, 2024
www.scholink.org/ojs/index.php/ape

Original Paper

Research on the Gatekeeper Mechanism of Data Platforms under the Perspective of Data

Haozhe Chen[1a] & Boning Guo[1b]

[1] Southwest Minzu University, Chengdu, Sichuan, China

[a] 2740556305@qq.com, [b] 19801987684@163.com

Received: April 19, 2024 Accepted: May 12, 2024 Online Published: May 28, 2024

doi:10.22158/ape.v7n2p141 URL: http://dx.doi.org/10.22158/ape.v7n2p141

Abstract

Since the fifth technological revolution, represented by artificial intelligence, emerged, data has become an essential driving force for social operation and an important engine for economic growth. Data analysis activities based on artificial intelligence big data analysis models have significantly improved social productivity, especially in promoting the tertiary industry. Data analysis enables precise user targeting, thereby making services more rational and refined. However, while these platforms apply and analyze data, non-compliant usage, illegal usage, and even criminal activities are prevalent. Consequently, the gatekeeper mechanism has been introduced in China as a safeguard mechanism for data compliance, serving as a baseline system to protect people's fundamental rights. This paper aims to explore the rationality of the "gatekeeper mechanism" and how to ensure the smooth implementation of this measure.

Keywords

Data Compliance, Personal Information Protection, Artificial Intelligence

1. Introduction

Data compliance refers to a series of regulations and measures that enterprises, organizations, or individuals follow to comply with laws, industry standards, and government requirements when processing and managing data. The goal of data compliance is to protect the rights of data subjects, prevent data leakage, misuse, and improper use, and establish a trustworthy data environment. Since the advent of the data era, nations have taken various measures to protect citizens' personal information security. According to the Personal Information Protection Law of the People's Republic of China, the country has taken multiple measures to protect citizens' personal information. This clause is the "gatekeeper" provision in China's legal system. However, as a new academic frontier, the theoretical

system construction of the "gatekeeper" system is still not perfect and has many debatable aspects. In China's Personal Information Protection Law, "large quantities" and "complex business types" are used as industry judgment standards for recognizing gatekeepers. This standard is somewhat reasonable and aligns with the current situation in China. However, in the long run, the criteria for identifying the gatekeeper mechanism need further clarification, and the interpretation of these criteria needs more detail. The gatekeeper system is a significant innovation in China's legislative practice, and its exploration of a new era power architecture system is an innovative attempt at the distribution of power and rights in the network era.

The gatekeeper system is not only established in China but also has similar legislative practices in the European Union and the United States. However, there are still certain detailed differences in specific provisions among countries. However, a close examination of its essence reveals that the legislative design of the gatekeeper system is to maintain the stability of the data market and control "data power." Therefore, this paper uses data compliance as the writing perspective to explore the gatekeeper mechanism of data platforms in China.

2. Data Platform

When referring to a data platform, it typically denotes an integrated system or framework utilized for consolidating and managing various data assets both within and outside an organization. A data platform encompasses a range of components including data warehouses, data lakes, data integration tools, data processing and analytics tools, as well as data visualization tools. Through seamless integration of these components, a data platform provides a unified environment for data management and utilization.

2.1 What is a Data Platform

In general, a comprehensive data platform encompasses several essential components, each serving a distinct function:

Data Collection: At the core of a data platform lies its ability to gather information from diverse origins, ranging from databases to APIs. This initial stage is crucial for assembling a robust dataset.Data Storage: An integral aspect of any data platform is its storage capability. This involves housing collected data within appropriate systems like data warehouses or NoSQL databases. Such systems are engineered to uphold high standards of performance, scalability, and security.Data Processing: Following data acquisition, the platform undertakes processing tasks to refine and prepare the data for subsequent analysis and application. This phase encompasses operations such as data cleansing, integration, and transformation, ensuring the integrity and coherence of the dataset.Data Analysis:The ultimate purpose of a data platform is to unlock insights from the amassed data. To achieve this, it offers an array of analytical tools and methodologies, including data mining, machine learning, and statistical analysis. These techniques empower users to discern patterns, trends, and correlations within the dataset.

Moreover, a data platform transcends the realm of a solitary tool or database; rather, it embodies a sophisticated amalgamation of interconnected components. These components seamlessly integrate to facilitate various functions such as storage, processing, analysis, and visualization. This cohesive framework provides users with a centralized environment for effective data management and utilization.

Furthermore, a robust data platform possesses the capacity to harmonize disparate data sources, encompassing internal systems like ERP and CRM, as well as external repositories such as social media and market data. By consolidating these diverse data reservoirs, organizations gain a holistic understanding of their operational landscape.

Efficiency is paramount in the realm of data collection. Hence, a data platform incorporates multiple mechanisms, including real-time data streaming, batch ingestion, and API integrations, to ensure swift and accurate data acquisition.

Additionally, data platforms leverage optimized storage solutions, such as cloud storage, to efficiently manage large datasets while maintaining scalability and performance. This strategic utilization of storage resources lays the foundation for sustained data-driven insights and decision-making.

2.2 The Data Platform in This Paper

This paper supports the data processor theory from previous research, i.e., a data platform is a data business that profits by using data and enterprises and institutions that generate or obtain data during the operation and organization of transactions. (Yu & Liu, 2021) The unique aspect of a data platform compared to other platforms is its ability to process and analyze data. Major platforms supported by cloud computing and big data algorithms are unique from other network platforms. Additionally, as China has set restrictive requirements for gatekeepers in the Personal Information Protection Law, the data platform referred to in this paper refers to network platforms with independent algorithm processing capabilities and a certain dominance in the industry.

3. Issues in Data Compliance in Reality

3.1 Monopolization of Platform Data

In the digital era, data has become a crucial engine driving economic growth. The fierce competition among major internet platforms for the domestic market in recent years is a testament to this. If data platform companies possess or monopolize market power, they may engage in the following illegal activities: First, collecting personal information without the user's knowledge, which severely violates the principles of equality and voluntariness outlined in the Civil Code. Second, collecting personal information in a deceptive manner, using incentives like discounts or free first orders to "trick" users into providing their information, which is then sold. Third, including unnecessary personal information in the collection list, thereby gathering personal data that users do not need to disclose. Dominant online platform companies often infringe on users' personal information rights through improper use or abuse, such as price discrimination based on big data, or selling personal information as a commodity.

Internet giants leverage their market advantage to restrict competitors' access to personal information or limit users' ability to transfer information to other platforms. For instance, the talent wars between Meituan and Ele.me, or the user battles between Kuaishou and Douyin. These actions infringe on users' right to choose between online platforms and create market entry barriers. Healthy market competition is an important driver of market economy development, while monopolistic and unfair competition not only disrupts social order but also harms citizens' rights.

3.2 Challenges in National Regulation

The data processing by data platforms is characterized by its secrecy and internal nature, with the processing procedures often regarded as highly confidential or even trade secrets within the platform. The commercial data collected and analyzed is intertwined with trade secrets, significantly increasing the difficulty of national regulation. As a result, the state can typically only intervene retroactively, conducting examinations of responsible parties long after issues have arisen and had a wide impact. To address these challenges, China has embedded compliance requirements into laws, regulations, and national standards, such as the 2015 Amendment (IX) to the Criminal Law of the People's Republic of China and the 2017 Draft Revision of the Anti-Unfair Competition Law. By leveraging the normative and mandatory functions of the law, these measures promote compliance practices among social entities and actively urge these entities to undertake more extensive regulatory obligations than before. (Wang, 2023) The high complexity of information science presents enormous challenges to national regulation, both in terms of the workload and the complexity of the tasks involved.

4. Current State of Gatekeeper Mechanism Development

4.1 Theoretical Foundation of the Gatekeeper Mechanism

The gatekeeper system refers to the special obligations and quasi-public "gatekeeper rights" granted by the state to online data platforms with significant algorithmic capabilities. This system has been implemented in the legislation of multiple countries, and China formally established it in the 2021 revision of the Personal Information Protection Law. Research indicates that the logical foundation of the gatekeeper mechanism lies in the "theory of premises management obligations" within the Chinese civil law system. That is, premises managers have the obligation to manage the premises and ensure their safety. Since state-level supervision and management are mostly retrospective, acting as remedial punitive measures to prevent the escalation or loss of control of adverse situations, there is a significant insufficiency in preventing ongoing or imminent adverse events. Hence, the importance of establishing a "gatekeeper system" is highlighted, as timely intervention before or during the occurrence of adverse behaviors can effectively curb them in their infancy.

Historically, powerful industry entities have collaborated to control entire industries, as seen in the monopolistic organizations represented by trusts in the era of monopoly capitalism in the United States. While there are certain similarities between monopolistic organizations and gatekeepers in terms of their dominant entities, there are substantial differences in their specific details. The former are

commercial alliances with an industry monopoly nature, whereas the latter are "quasi-official" entities granted special status by the state, bearing more extensive social and national supervisory responsibilities over the industry.

4.2 Legislative Status of the "Gatekeeper Mechanism" in Various Countries

On July 18, 2022, the Council of the European Union adopted the Digital Markets Act (DMA), which came into effect in November of the same year. As a significant legislative attempt at preemptive regulation, the DMA established the "gatekeeper" system. Specifically, Article 6 of the DMA stipulates that, upon the request of end users or third parties authorized by end users, gatekeepers must provide data supplied by or generated through the activities of end users on core platform services, in a manner that ensures effective data portability, free of charge. This includes providing tools to facilitate effective data portability and ensuring continuous and real-time access to such data. (Wang, 2023)

On June 11, 2021, the U.S. House Judiciary Committee introduced five antitrust bills, including the American Choice and Innovation Online Act and the Ending Platform Monopolies Act. These bills introduced a concept similar to the EU's "gatekeeper" - the "covered platform." According to these bills, once a platform qualifies as a "covered platform," it must adhere to specific obligations. For instance, the Augmenting Compatibility and Competition by Enabling Service Switching Act requires covered platforms to ensure data portability for users, provide transparent interfaces accessible to third parties, and enable users to securely transfer their data to other platforms, ensuring interoperability with other platforms.

China's Personal Information Protection Law, Article 58, has already established gatekeeper obligations, similarly imposing specific duties on internet platforms with large user bases and complex business operations. Although the obligations outlined in the aforementioned laws (and drafts) vary, their core essence remains the same: relevant online platforms are endowed with various special responsibilities to maintain the stable and compliant operation of the data market.

5. Future Prospects of the Gatekeeper Mechanism

Regarding the practice of the gatekeeper mechanism, it is currently in the early trial stage both internationally and domestically. The state may further improve this series of systems in the future. Firstly, the concept and criteria for identifying digital gatekeepers need to be further clarified to ensure the accurate grasp of the regulatory boundaries of digital gatekeepers. Subsequently, a corresponding legal framework should be established to ensure that the gatekeeper system can exert its true effectiveness. At the same time, industry supervision needs to be strengthened to ensure that digital gatekeepers comply with rules and fulfill their corresponding responsibilities. This article will explore the future system of the gatekeeper mechanism from the perspective of data compliance.

5.1 Clarifying the Identification Criteria for Gatekeepers

Firstly, defining the identity of gatekeepers is crucial. What kind of enterprises should be given the title of "gatekeeper"? Is the standard for gatekeepers merely having a data platform with strong algorithmic

capabilities? Or does it mean that the platform has significant control over the lives of the public? What specifically does the complexity of business mentioned in Chinese law refer to? Many of these questions need to be confirmed by facts and further research. To solve the above problems, we must first understand the fundamental concept of the gatekeeper system—controlling data power and ensuring data compliance. Algorithm platforms collect and analyze large amounts of data to form their own "data power" or "algorithm power." Although "data power" and "algorithm power" are not public powers in the strictest sense, the public's dependence on large online data platforms is becoming increasingly deep. A significant portion of the population has almost reached a level of dependence that is hard to separate, similar to the public's "inseparability" from government agencies. Based on the above analysis, network data platforms form a quasi-public power in their own domains—a "power of data." This quasi-public power of "data power" must be restricted from all sides; otherwise, data compliance becomes mere empty talk, which fully explains the rationale behind establishing the gatekeeper mechanism.

With the concept of system design clarified, we can better determine the identification criteria for gatekeepers, i.e., the platform has already formed "data power" in its domain. First, the platform should have a relatively large user base and deep social dependency. Second, the platform's use should exert strong control over its users. Third, if the platform abuses its power, it could cause serious negative consequences for society.

5.2 Regulating the Operation System of "Power"

Articles 2 and 3 of the Personal Information Protection Law essentially grant significant personal information processors the "power" of enforcement, legislation, and supervision. However, the Personal Information Protection Law does not regulate this "power," which can easily lead to its abuse. (Li, 2023) The abuse of "power" presents enormous challenges to data compliance. The "power" of gatekeepers in data platforms originates from the delegation of state power. However, with the technological evolution of data platforms and the deepening consolidation of their dominant positions, it is easy for the state's role and functions in this field to weaken, causing the gatekeeper's "power" to expand and encroach upon normal state power. Currently, China's gatekeeper system is still in the practice and exploration stage. Therefore, regulating the operation system of gatekeeper "power" at the initial stage of this system is particularly important. The state should not only act as a guide but also fulfill its legislative role by issuing relevant judicial interpretations and normative documents to improve the full process of "power" usage, ensuring that "power" is exercised under the sun.

5.3 Establishing an Efficient Supervision System

While regulating the operation system of "power," the state must also establish a strong supporting supervision system. Otherwise, relying solely on some normative documents and the industry self-awareness and social responsibility of business entities will hardly effectively regulate an industry.

Traditional administrative industry supervision is strongly ex-post. However, given the hidden and rapid nature of network platform data, a new supervision system must be created for network data

platforms; otherwise, it is difficult for the supervision system to exert its true effectiveness. Establishing a new supervision system also needs to avoid excessive use of social resources, making the inclusion of public power in the new supervision system particularly necessary. From the perspective of service recipients and providers, the public are the first to enjoy the services and have a better response to problems encountered during service acceptance. Although the state has public law enforcement power and can most effectively combat illegal activities, in reality, state agencies often spend a large amount of time on evidence collection and preliminary investigations, which is not conducive to maintaining social order and protecting public rights.

6. Conclusion

At this time, establishing a multi-party supervision system with public participation becomes particularly necessary. Once network platforms are given the "power" of gatekeepers, the public naturally gains the corresponding right of supervision. The law grants Chinese citizens the right to supervise public power, and it is even more necessary to supervise quasi-public power exercised by private entities. Facts have proven that public supervision is an important guarantee for ensuring that power is exercised under the sun. The obligation of information disclosure specified in the realistic legislation of China's gatekeeper system confirms this viewpoint. For the same reason, we cannot rely on the industry self-awareness and social responsibility of related enterprises for the reasonable operation of the industry. This requires a tripartite supervision system led by the state, involving public participation, and coordinated with the platform. The public is responsible for disclosing adverse phenomena within the platform, the state is responsible for quickly addressing and combating illegal phenomena, and the platforms acting as gatekeepers must bear dual identities. On one hand, they are the owners and builders of data platforms; on the other hand, they also bear the obligation to supervise and manage their own platforms. The three parties cooperate with each other, with the state and the public conducting secondary supervision of the platform's self-supervision results, allowing for a more efficient control of the gatekeeper "power" of data platforms. At the same time, legitimate business secrets of data platforms must be reasonably protected, strictly prohibiting any subjects from engaging in suppression or illegal acquisition of platform business secrets and key information under the guise of industry supervision. Otherwise, the establishment of the supervision system will turn into an arena for power struggles.

References

Li, A. J. (2023). Special Supervision of Important Personal Information Processors: A Study on Article 58 of the Personal Information Protection Law of the People's Republic of China. *Journal of China University of Political Science and Law*, *2023*(04), 56-67.

Wang, L. (2023). Triple Path to Promoting Data Openness of Large Platforms from an Antitrust Perspective. *Local Legislation Research*, *8*(01), 33-50.

Wang, P. (2023). Compliance Obligations and Path Implementation for Personal Information Protection on Data Platforms. *Jiangsu Sociology*, (03), 140-147.

Yu, X. L., & Liu, D. Y. (2021). Legal Boundaries for the Commercial Use of Personal Data on Big Data Platforms. *Legal Forum*, *2021*(01), 313-324.