Practical Dilemma and Cracking Path of the Free Flow of Commercial Data Circulation

Chen Jun, Liu Yue

Law School, Anhui University, Hefei, Anhui, 230601, China

Keywords: Data circulation; data scraping; digital economy; Data silos

Abstract: The value of data lies in circulation and the vitality of data also lies in circulation. However, China's legislation in the field of data has long been "Emphasizes Protection but Lights on Sharing ". The legal norms lack of systematic, independent, low-level and principle-based expression. As the legal norms to regulate data has natural limitations, the discriminatory determination of the data capture behavior in judicial practice has led to the difficulty of data sharing and releasing the value of the dividends. In view of this, it is important to draw on the useful experience of data regulation in the world, adhere to the principle of data sharing, improve and standardize the legal rules and informal norms of data circulation. Clarify the idea of regulating data scraping and introduce the theory of necessary facilities for data monopoly to promote the use and circulation of data.

1. Introduction

Unlike traditional factors of production, commercial data are endowed with reproducibility, shareability, unlimited growth and supply\(^\text{[2]}\), which means that the value of data is positively correlated with data openness, data circulation\(^\text{[3]}\) and releasing the value of data also requires data circulation. Therefore, only by guaranteeing the free flow of data to a greater extent, promoting the opening and sharing of data to a greater extent can the value of data be released to a greater extent. The value of data lies in circulation and the vitality of data also lies in circulation\(^\text{[4]}\). In other words, without high-value data supply, data flow and data sharing, the further development of the digital economy will be unsustainable. It means that whether data flow is smooth has become a key point for the high-quality development of China's data economy\(^\text{[5]}\).

In December 2022, the Central Committee of the Communist Party of China and the State Council issued the Opinions on Building a Data Base System to Better Play the Role of Data Elements, stating, "Adhere to the sharing of common use and release the value dividend. Reasonably reduce the threshold for market players to access data, enhance the shareability and universality of data elements, incentivize innovation and entrepreneurship creation, strengthen anti-monopoly and anti-unfair competition and form a development model that is regulated by law, jointly participated and shared dividends." To a certain extent, this also provides a fundamental guideline and action guide for the construction of the data circulation system in the coming period.

However, from the perspective of relevant research in the academic community, the legal research on data circulation in China is still in its infancy\(^\text{[6]}\). The academia mainly focuses its research on the basic concept of data circulation, the legitimacy and necessity of data circulation. No scholars have
systematically elaborated the dilemma of the rule of law regulation of commercial data and its cracking path based on the value position of promoting the flow of commercial data, from the top-level design, current legislative situation, judicial application of the dilemma and other aspects.

In view of this, this paper starts from the most important way of commercial data circulation, i.e., data crawling, analyzes the lack of legislative provision and judicial dilemma of commercial data circulation. Draws on the useful experience of commercial data protection in other countries. On this basis, we establish the concept of "Sharing as the Principle and Protection as the Bottom Line", to improve the legislation of data circulation promotion elements. Introduce the "Interest Measurement Method" to guide the "Determination of Legitimacy" of data crawling behavior. Introducing the theory of necessary facilities for data monopolization, the status quo of commercial data crawling is changed. The data silos are opened to promote data sharing, fully realizing the value of data elements and promoting the sharing of the dividends of the development of the digital economy by all people.

2. Rule of Law Dilemmas in the Flow of Commercial Data

2.1. Data Legislation "Emphasizes Protection but Lights on Sharing"

From the viewpoint of domestic legislation in the field of data, China's data legislation tends to focus on personal information security and privacy protection, as well as the security and protection of cross-border data circulation.[7] There are fewer legal regulations on data circulation and the legislation pays weak attention to data sharing, which overall shows a legislative concept of "Emphasizes Protection but Lights on Sharing".[8]

After combing and studying the relevant legislation in the field of data, comprehensively considering the level, influence and relevance of the legislation. It is not difficult to find that the existing laws on the governance of information and cyberspace are all laws that guarantee security and protect rights and interests. Specifically, these laws reflect the fact that the concept of "Data Protection" is deeply rooted in the existing legal norms for data, while the concept of "Data Opening and Sharing" is relatively weak. For example, Article 10 of the Cybersecurity Law states: "The integrity, confidentiality and availability of network data shall be maintained" and Article 21: "Network data shall be prevented from being leaked, stolen or tampered with"; Article 25 of the National Security Law: "Realize the security and control of information systems and data in important areas"; the Law on Public Security Administration Punishments mentions: "Anyone who violates state regulations by deleting, modifying, adding to data and applications stored, processed or transmitted in a computer information system shall be sentenced to detention "; the Criminal Law Amendment (IX) provides the crime of violating citizens' personal information. Specialized legislation for the legislative purpose of opening data silos and promoting data sharing is rare. At present, there are only a few normative documents in the country, such as the Regulations on the Promotion of the Development and Application of Big Data in Guizhou Province and the Measures for the Management of Shanghai Municipal Governmental Affairs Data, which reflect the legislative philosophy of lawmakers attempting to promote data sharing. There are no clear legislative provisions on the distribution of ownership, measurement of interests, infringement boundaries and other issues in the process of the economic value of data sharing data, only the regulatory model of the Anti-Unfair Competition Law generally stipulates that data circulation, such as data capture and other circulation behaviors, shall not harm competition and be of an unfair nature. However, the vague application standards of the legislative norms of the Anti-Unfair Competition Law and even the lack of regulation have caused many problems in judicial practice. Such as difficulties in application, inconsistency of standards, large limits of discretion and polarization of sentences and penalties.

This series of phenomena shows that China's law on the use of data and the lack of knowledge of the market value of the issue. Neither at the level of legal regulation to encourage and guide the
development of the data market. In the data capture unfair competition, data capture may lead to monopoly and other damage to the market behavior occurs cannot be timely and accurate adjustments. It is difficult to grasp the implementation of the data behavior on the market from the macro level and the flow of data elements. With the development of the data industry and the advancement of the Internet, big data, cloud computing and other technologies. The demand for data protection in the field of data, although still important, is no longer dominant. The demand for breaking up data silos and realizing data sharing has become more urgent. Putting the data legislation concept of "Focusing on Protection but not on Sharing" in the context of the artificial intelligence era, it will be found that the concept has a certain degree of value-oriented bias, which has lagged the needs of practical development and cannot provide the best value guidance for cracking the data island, nor can it provide sufficient legal protection for data sharing.

2.2. Existing Data-Related Legislation Lacks Independence and Systematization

When analyzing the current legislative situation in the field of data in China. It can be found that the existing legal normative system to promote the flow of data is still "Overstretched" in the face of the booming digital economy. On the one hand, China's legislative system is scattered in several laws and regulations of different legal departments, as well as policy documents issued by various ministries and commissions of the State Council, involving civil relations, data security, network security, personal information protection, market regulation, intellectual property rights and so on. There is no single line of legislation specifically targeting the issue of data field. On the other hand, the legislative level is not high, the legislation of the principle of expression is dominant. For the data problem, what kind of theoretical basis should be followed, what kind of regulatory principles should be adopted, how to different data behavior should be regulated by chapter and so on, are in the process of groping for exploration. For this reason, it is urgent to formulate a basic law reflecting the characteristics of the digital economy at the central level, to systematically regulate the basic and key issues in the development of the digital economy. To take the characteristics and key elements of the digital economy as the starting point. To do a good job of legislative coordination, subject synergy and empowerment. To promote accurate and effective regulation through scientific and detailed institutional arrangements.

2.3. The Legal Norms Regulating Data Have Limitations

The legal norms of the circulation of data elements is a "Hard" formal organizational and institutional framework, a kind of mandatory rules closely linked with the state power or organization. Once established, legal norms have relatively high stability. But there are some limitations in the circulation of data elements.

First, the supply of legal norms is relatively limited. Data circulation involves a wide range of legal norms and cannot consider all aspects. There may be a vacuum of rules, resulting in a temporary lack of system supply.

Secondly, the supply of legal norms has a relative lag. It is difficult to adapt to the increasingly diversified forms of data circulation. With the development of cloud computing, big data, artificial intelligence and other technologies, the processing and circulation of data has undergone tremendous changes. The emergence of blockchain technology, decentralized storage and other new types of data management methods has also brought new opportunities and challenges to data circulation. As it takes time to formulate and update legal norms. It is difficult to keep up with the rapid development of technology and business models. This may lead to a disconnect between legal norms and the reality. Also fail to effectively regulate the behavior of data circulation.

Finally, the cost of supplying legal norms is high, especially for an area like data circulation that
involves complex technical, commercial and ethical issues. Due to the complexity and rapid change of data circulation, it is often difficult for the legal norms formulated to adapt to new changes. Thus increasing the cost of updating and adjusting the legal norms. In addition, as data circulation involves a number of fields and aspects, such as technology, commerce and ethics, it is often difficult for the legal norms formulated to harmonize the interests of these fields and aspects. There are costs associated with the formulation of rules. When the existing rules are dysfunctional or inefficiency occurs, both the revision of rules and the reformulation of rules inevitably face the problem of uneconomical.

2.4. Misconceptions of Data Crawling in Judicial Practice

Data crawling, as one of the main ways of data circulation, plays a significant role in promoting data circulation. As some scholars have pointed out "As a program that automatically extracts web pages, the data crawler used in data crawling is an indispensable tool for search engines and it can even be said that, without data crawling, users would be lost in the vast ocean of information on the Internet. Similarly, for small and medium-sized enterprises, data crawling helps them to access the data resources that are difficult to obtain by themselves, to cross the ceiling of the scale effect". However, there are misunderstandings in judicial practice regarding the determination of data crawling, which seriously restricts the effectiveness of the application of this behavior. Some scholars have conducted empirical research on 20 cases involving data crawling disputes over unfair competition between January 2010 and May 2021. Found that the application rate of Article 2 of the Anti-Unfair Competition Law is as high as 80% and in terms of the judgment results, the data crawling party's defeat rate is as high as 85%. By analyzing a series of typical cases, such as "Public Comments v. Ai Bang", "Taobao v. Anhui Meijing Company", "Gumi v. Che Lai Le ". We found that China's judicial system is currently in the process of The analysis of a series of typical cases. In Taobao v. Anhui Meijing Company and Gumi v. Che Lai Le, we found that China's judiciary currently relies excessively on moral intuitive judgments such as "Hitchhiking" and "Getting something for nothing" in the trial of disputes arising from data crawling. Considering that the party against whom the case is filed violates the principle of honesty and good faith, lacks the consideration of business ethics and the creation of social value. Secondly, the determination of unfair competition was rigid and limited to the logic of private law protection, favoring the tort law idea of "Behavior - Damage - Causation", starting from the perspective of the data control enterprise, protecting the competition interest as the exclusive right of the data control enterprise and rigidly determining that the defendant's behavior damaged the legitimate rights and interests of the data control enterprise, constituting unfair competition.

Confined to the abstraction and ambiguity of the general provisions, the courts often have their own views in the application, inevitably with a certain degree of subjective color, resulting in inconsistent application of the law. The embarrassing situation of "Different Judgments in the Same Case", which undermines the fairness and authority of the judiciary. At the same time, the data industry and the traditional industry are significant difference. In order to protect the vitality of the Internet industry innovation, Certain freedom of action should be given to the data field of market players. In summary, the abuse of general provisions and the trial mode of the solidification of the data industry is undoubtedly in the obliteration of commercial innovation that contrary to the legislative purpose of the Unfair Competition Law.
3. Other Stones: Extraterritorial Experience In Legal Regulation Of Data Circulation

3.1. EU: "Gatekeeper's Rule"

On July 18, 2022, the European Commission adopted the Digital Markets Act[^12] (DMA), which sets out to regulate digital markets across Europe. The Digital Markets Act introduces the "Gatekeeper" rule, which is a major change in the regulation of commercial data and the regulation of the crawling behavior of Internet subjects. The provision stipulates that "Gatekeepers" with strong economic power, control over many users in the field and long-lasting influence should fulfill higher standards of data openness. Such as providing data to third-party search engines on fair, reasonable and non-discriminatory terms. This will, to some extent, enhance the various reuse attributes of data on powerful platforms, increase the efficiency of data use as much as possible while ensuring data security and weaken the impact of data monopolies to enhance competition.

3.2. United States: Shift in the Concept of "Shared Circulation" of Data

Under the federal Computer Fraud and Abuse Act (CFAA), for computers that have adopted technical protection measures, it is illegal to maliciously cause damage to the computer information system or maliciously modify or delete the original data when carrying out the act of capturing. It is also illegal to capture data without the authorization of the right holder or beyond the scope of authorization. The legal provisions do not refine and clarify the elements and scope of ‘authorisation’. In practice, the authorisation methods adopted by the subject of the right are varied, which makes the determination standard of ‘authorisation’ in the early years of the United States judicial practice too broad and the legality requirements for data capture too strict. For example, a large number of United States courts have held in their jurisprudence that data harvesting constitutes an offence as long as it violates a contract. There is even jurisprudence that has found a breach of a right holder’s authorisation on the basis of a lawsuit filed by the data controller in relation to the scraping. Such intense legal regulation has led to an inappropriate expansion of the rights and status of data controllers and a greater restriction on the free flow of data. Instead of promoting the development of the data economy, it has been met with many voices opposing strong power. Therefore, in recent years, the U.S. legislation and judicial level began to limit the definition and requirements of "Authorization" interpretation, to give greater freedom of data capture behavior and accelerate the flow and use of data. The legal regulation of data capture behavior, no longer to obtain the full authorization of the rights as the basic requirements, but only punish the law prohibits the implementation of the act. Such as hacking, destruction of information systems, to avoid the technical barriers set by the system and so on.[^14]

The above judicial attitude change reflects the United States for the social utility of data capture behavior, the basic concept and regulatory mode are changing. The United States gradually began to pay attention to data as a market resource which has the value of openness, sharing, circulation and utilization. The appropriate relaxation of the legal boundaries of data capture can accelerate the flow of data, which is more in line with the basic needs of the digital economy driven development in the U.S. market today. It is also conducive to the development and progress of the data industry and even society. Moreover, it can provides a lesson for our country's attitude towards data: data protection is not the goal. Overprotection of data or even the practice of data blocking and monopolization in the name of data protection, is even more unlawful.
4. The Legal Regulation Of Data Circulation Behavior Of The Way Forward

4.1. Establish the Concept of “Sharing as the Principle and Protection as the Bottom Line”

From the level of real demand, domestic industry, scientific and technological development are in urgent need of data "Raw Materials" as a support. However, the current data industry shows an imbalance in supply and demand, closed transactions, insufficient government data openness and other shortcomings, which seriously restricts the economic development. In the early stage of the data industry, the legal norms pay more attention to data security is understandable. However, with the development of social practice, the security concept alone is no longer sufficient to support the current legal system of data regulation. The concept of benefit will inevitably become the law in the regulation of the data industry that should focus on the factors to consider.

In recent years, the CPC Central Committee and the State Council issued a series of policy documents such as The CPC Central Committee and the State Council on the Construction of a More Perfect Institutional Mechanism for the Market-Oriented Allocation of Elements of the Opinions, Digital China Construction Overall Layout Planning, have proposed to accelerate the establishment of data base system, market rules and standards and norms, promote the use of public data aggregation and utilization, smooth the general circulation of data resources, effectively release the value of the data elements. It has clearly shown the state's interest in opening up the flow of data. Demonstrated the state's attitude towards opening the barriers to data circulation and smoothing the flow of commercial data elements.

Therefore, combining the actual situation in China and the top-level planning of the country, the revision of data legislation in China in the future should adopt the legislative concept of “Sharing as the Principle and Protection as the Bottom Line”. That is to say, while vigorously promoting data sharing in the era of artificial intelligence and breaking down data silos, the protection of citizens' privacy and data security should be the “bottom line” of data sharing. It is strictly prohibited to violate citizens' privacy in the name of data sharing, so as to draw the necessary no-go zones and red lines for data sharing. This will not only help balance the conflict of interest between data sharing and privacy protection, but also better suit the long-term interests of the development of the data industry.

4.2. Improvement of Legislation on Data Flow Promotion Elements

First, consider the introduction of the Data Circulation Promotion Law, which clarifies the purpose and principles of data circulation, regulates the channels and methods of data circulation, protects individual privacy, autonomy, data opening and sharing and strengthens the regulatory and self-regulatory mechanisms. It also provides strong protection and support for the legality, standardization and fairness of data circulation. With this law as the core and overarching principle, legislation or norms on the circulation of different data types should be introduced to provide norms for the circulation of data, including commercial data, in the legislative mode of “Overarching + Diversified”.

Second, the minimum necessary data sharing requirements should be maintained to facilitate the flow of commercial data. Relevant departments or industry associations should set different types of data circulation, circulation methods and circulation rights for different types of enterprises. Establish minimum standards for commercial data circulation, so that the circulation of commercial data can be carried out within the scope of minimum necessity, problems such as data misuse and leakage can be avoided.

Thirdly, improve the proactive legislation, incentive mechanism for business data circulation and set up a reasonable cost-sharing mechanism. While actively promoting proactive legislation to promote business data circulation such as business data opening, data sharing and data trading, etc., it stipulates the legal form and legal type of business data circulation, introduces favorable policies
to encourage and guide enterprises to carry out data circulation. Especially business data circulation that promote the conditional release of data to the society, make it profitable for enterprises to stimulate their proactiveness and enthusiasm in data circulation.

4.3. Construction of Informal Norms for the Circulation of Data Elements

Historically, before the creation of legal norms, the relationship between people mainly relied on informal norms to constrain and the formal norms today only account for a small part of the total constraints that determine people's choices. Indeed, informal norms serve the important function of expanding, clarifying, and amending formal legal norms. It has become socially accepted codes of conduct and internally enforced standards of behavior and most of people's behaviors are also governed by informal institutions.

In the field of data circulation, ethical norms and moral codes are of great significance in promoting the reasonable circulation of data elements. Problems such as data monopoly, data silos and imbalance in data utilization in data circulation are to a large extent closely related to the absence of open sharing ethics. Therefore, the author suggests introducing ethical principles of data sharing to strike a balance between data abuse and data silos.

In conclusion, an effective institutional arrangement for data elements is inevitably the organic unity of formal legal norms and informal norms. It is necessary for us to incorporate the informal norms of data element circulation into the data circulation and trading system. In order to realize the optimal cost-benefit, the highest efficiency.

4.4. The Introduction of the "Measure of Interests" to Guide the Data Capture Behavior Justification Determination

The Anti-Unfair Competition Law is modest, the core of which is to safeguard free competition in the market economy. Should not be preemptive, focusing on the protection of the interests of data holders, but should be weighed after the interests of all parties for a comprehensive consideration. As far as possible, to broaden the boundaries of the legitimacy of the act of data capture. So that the flow of useful data to better serve the public.

On the one hand, the principle of “Non-Interference Without Public Interest Necessity” and the standard of business ethics should be integrated to broaden the boundaries of the legitimacy of data scraping behavior. The core issue for the court in applying the general provisions is to determine whether the operator has violated business ethics. In previous cases of unfair competition on the Internet, the principle of “Non-Interference Without Public Interest Necessity” has been recognized as a useful judicial attempt to interpret business ethics. It means that Internet products or services shall not interfere with each other in principle. It is true that out of the need to protect the interests of the public such as network users, network service operators may also interfere with the operation of others' Internet products or services without the knowledge and active choice of network users and the consent of other Internet product or service providers under specific circumstances, but the necessity and reasonableness of the means of interference should be ensured and proved. However, in practice, the principle is applied with different standards, the premise is not clear and there are obvious drawbacks. In addition, since March 20, 2022, the Interpretation of the Supreme People's Court on the Application of the Law of the People's Republic of China Against Unfair Competition on a Number of Issues (hereinafter referred to as the Interpretation) came into force. The main highlight of the Interpretation is that it refines the business ethics stipulated in the general provisions of the Competition Law (Article II). Article 3 of the Interpretation states that it shall consider the rights and interests of consumers, order of market competition, impact of the public interest and other factors to determine whether the operator violates business ethics in accordance with the law. Due to
the rapid changes in the development of the Internet, new technologies and new models are emerging. Whether it is “Non-Interference Without Public Interest Necessity” principle or the Interpretation, are not an exhaustive list of new types of unfair competition on the network of the manifestation of the form and determination of the standard. The prudent attitude of the Supreme Court is to reserve space for market self-evolution, technological innovation and encourage operators to make breakthrough creations in line with business ethics. This concept should be adhered to in judicial practice.

On the other hand, it builds an analytical framework for judging improper behavior based on the principle of proportionality. Under the perspective of the principle of proportionality, no interests are absolute. On the contrary, various legally recognized interests should be compromised, coordinated and compatible as far as possible. So that they can be realized to the greatest extent possible and even get greater development in coordination with each other. It is generally recognized that the principle of proportionality requires the satisfaction of “Appropriateness - Necessity - Proportionality (the principle of proportionality in the narrow sense)”. In today's judicial practice, the application of the principle of proportionality is not structured in accordance with the analytical framework and is often improperly applied which reflected in the cases of “Oneness v. Kingsoft” and “Tencent v. Qihu Unfair Competition”. In this regard, some scholars have proposed the principle of proportionality analysis framework. Briefly summarized as follows: the plaintiff proves that the rights and interests of the damaged - the defendant proves that the damage to the legitimacy of the act - the plaintiff can be questioned on the appropriateness and necessity, the defendant needs to respond - the plaintiff The plaintiff may challenge proportionality, i.e., the harm caused is greater than the benefit realized. Regarding the principle of proportionality, it is usually considered that only when the harm caused is significantly greater than the benefits realized is it considered to be inconsistent with the principle. Judicial practice can be integrated with relevant cases to form a process of determining acts of unfair competition for cases of data capture, which is considered in combination with the concept of modesty, business ethics and anti-monopoly factors after a preliminary judgment is made.

4.5. Introducing the Theory of Necessary Facilities for Data Monopolization to Regulate Data Monopolization

With the increasing importance of data as a factor of innovation in the era of digital economy and the dominant position of Internet companies in the market by virtue of the huge amount of data in their possession. The emergence of data barriers due to the problem of data attribution not only stifles the innovativeness of small and medium-sized enterprises, but also does not contribute to the socio-economic development. The latest legislation and judicial practice of various countries have also begun to pay attention to the antitrust issue in the field of data, such as the determination of corporate data monopoly in the U.S. “LinkedIn Case” and the “Gatekeeper Rule” in the EU's Digital Market Act. However, China's current Antimonopoly Law is not yet ready to deal with data monopoly legislation. The author suggests that the theory of necessary facilities can be introduced when regulating the data monopoly that may be caused by data crawling behavior. For the part of data that cannot be captured to participate in market competition or difficult to complete the original collection by itself can be recognized as essential facilities. The implementation of compulsory licensing mode, prohibiting data controllers to implement anti-scraping behavior on the grounds of infringing on the interests of the data and undermining the competitive advantage. However, it should be noted that, since the purpose of the essential facilities theory is to protect the order of competition and safeguard the overall interests of society, the introduction of this theory can limit the current data monopoly behavior to a certain extent, but it should still be strictly limited in its application.
5. Conclusion

In the era of the knowledge economy, data has risen to become a factor of production after technology. From the metaphor of "Oil", which emphasizes its scarcity, to the metaphor of "Air", which emphasizes its circulation attributes. It reflects our deepening understanding of data and the laws of the data economy. Not only should we protect data at a high level, but we should also promote high-efficiency data circulation and sharing in order to realize more efficient data utilization, activate the potential of data elements, strengthen, optimize, and expand the digital economy. Enhance the new impetus for economic development and build a new competitive advantage for China.

References