Journal of Commercial Law Review
2025, VOL. 9, NO. 4, 38-41
DOI: 10.47297/wspjclrWSP2516-249708.20250904

Legal Regulation Dilemmas and Path Optimization of Data Crawling: From the Perspective of the 2025 Revision of the Anti-Unfair Competition Law

Maocuo Zhou

Minzu University of China, Beijing, 100081, China

ABSTRACT
In the era of the digital economy, data crawling has become an important means for enterprises to gain competitive advantages, but it has also triggered a large number of unfair competition disputes. For a long time, judicial practice has relied on the general clauses of the Anti-Unfair Competition Law for adjudication, which has problems such as inconsistent standards and vague scope of protection. The newly revised Anti-Unfair Competition Law in 2025 adds a special clause on data protection, clarifies the constituent elements of unfair competition acts in data crawling, and provides a special normative basis for regulating data crawling acts. Combining academic theories and judicial cases, this paper starts from the legal characterization dilemmas of data crawling, interprets the normative connotation of the special data clause in the new law, analyzes the key controversial issues in its application, and puts forward specific suggestions for improving the legal regulation of data crawling, aiming to realize the balance of interests between data circulation and rights protection.

KEYWORDS
Data crawling; Anti-unfair competition law; Special clause on data protection; Technical management measures

1 Legal Characterization Dilemmas and Theoretical Controversies of Data Crawling

1.1 Behavioral Characteristics and Classification of Data Crawling

Data crawling refers to the act of traversing online content and acquiring data in accordance with preset rules through automated programs such as web crawlers. According to data attributes and usage purposes, it can be classified into two categories: based on whether the data is public or not, it is divided into public data crawling and non-public data crawling; based on the purpose of use, it is divided into competitive use crawling and non-competitive use crawling [1].

The object of public data crawling is data without access restrictions, and its legality dispute focuses on the legitimacy of crawling methods and usage methods; non-public data crawling usually requires authorization as a legal prerequisite because it involves the control right of the data holder. Competitive use crawling refers to the act of crawling data for business activities that directly compete with the original data holder, such as crawling commodity data from e-commerce platforms for the operation of similar platforms; non-competitive use crawling includes academic research, public welfare purposes and other situations that do not directly impact the market position of the original data holder. This classification provides a basis for legal characterization, and different types of data crawling acts should apply differentiated regulatory rules.

1.2 Focus of Theoretical Controversies on Data Crawling

There are many controversies in the academic circle regarding the legal regulation path of data crawling, focusing on property right allocation and legal interest balance.

The property right allocation approach advocates defining the legality boundary of crawling acts by endowing data holders with exclusive rights. The costs invested by enterprises in data collection and processing should be protected, and endowing exclusive rights can stimulate the production and circulation of data elements [2]. However, opposing scholars point out that data has the characteristics of non-competitiveness and fluidity. Excessive confirmation of rights will lead to "data silos", hinder the free circulation of data, and due to the diverse sources of data, it is difficult to define the ownership of property rights [3]. Commercial platforms do not have ownership or control rights over the collected original data, which cannot meet the prerequisites for trade secret protection, making it difficult to achieve effective protection through property right allocation.

The legal interest balance approach advocates identifying the competitive advantages worthy of protection on the premise of ensuring the free circulation of data, and defining the legality of crawling acts through interest measurement. This approach holds that the legality judgment of data crawling needs to balance the labor input of data holders, the competitive rights and interests of crawlers, and consumer welfare. The "dual supplementary" function of the Anti-Unfair Competition Law makes it an appropriate path for data protection, which can not only supplement the deficiencies of special intellectual property laws but also provide protection for unformed rights and interests [4].

1.3 Qualitative Differences in Judicial Practice

There are significant differences in the judicial practice regarding the legality judgment of data crawling, mainly

focusing on two issues: the effectiveness of user consent and the identification of technical measures.

Regarding the effectiveness of user consent, there are three judicial positions: first, completely denying the defense effect of user consent. For example, in the case of "Weimeng Company v. ByteDance Company", the court held that the scope of user authorization does not cover the legitimate rights and interests of data holders; second, recognizing the obviating effect of user consent. For example, in the case of "Qianjin Company v. Yicheng Company", the court determined that the crawling act was legal on the grounds of voluntary user authorization; third, regarding user consent as an authorization arrangement for data rights and interests. For example, in the case of "Tencent Company v. Zhejiang Soudao Company", the court held that the crawling of a single original data with user consent does not constitute unfair competition.

Regarding the identification of technical measures, the dispute lies in whether a public statement prohibiting crawling alone constitutes a legitimate technical management measure. In the case of "Alibaba v. Mazhu '1688 Data' Case", the court held that the legal statement of the original data holder can restrict crawling acts; while in the case of "Weibo v. Yifang Case", the court argued that for public data without substantive protection measures, crawler crawling is as legitimate as user browsing. This divergence reflects the ambiguity of the standard for judging the legality of data crawling in judicial practice, which urgently needs to be clarified by legislation.

2 Normative Interpretation of the Special Data Protection Clause in the 2025 Anti-Unfair Competition Law

2.1 Provisions Structure and Legislative Intent of the Special Data Protection Clause

The special data protection clause added in Paragraph 3 of Article 13 of the 2025 Anti-Unfair Competition Law adopts a normative structure of "conduct elements + consequence elements" and clarifies the constituent elements of unfair competition acts in data crawling. From the perspective of legislative intent, this clause aims to solve the long-standing problem of lack of special norms for the regulation of data crawling. By typifying improper means and clarifying harmful consequences, it enhances the certainty of legal application.

2.2 Normative Definition of Core Concepts

2.2.1 "Legally Held Data"

"Legally held data" is the object of protection under the special data protection clause, and its identification must meet the dual requirements of legality and controllability. The legality requirement means that the collection and processing of data comply with the provisions of laws such as the Personal Information Protection Law and the Data Security Law. For example, collecting personal information requires obtaining user consent, and processing public data must follow the rules for opening public data; the controllability requirement means that the data holder has actual control over the data and can determine the scope and method of data use.

If the original data collected by commercial platforms meets the constituent elements of unfair competition, it can be protected by the Anti-Unfair Competition Law, but it is necessary to distinguish between original data and data products: original data is difficult to constitute trade secrets due to the lack of control rights, while data products formed through processing can be protected as compiled works if they meet the requirements of originality [5].

2.2.2 "Technical Management Measures"

"Technical management measures" is a core concept for judging the impropriety of crawling methods, referring to the technical protection means adopted by data holders to protect data. The academic circle generally believes that technical management measures should have substantive protective effects, and a mere public statement prohibiting crawling without taking substantive technical means should not be recognized as "technical management measures" as referred to in this clause. The intensity of technical measures should be commensurate with the data value and acquisition difficulty, and a simple pop-up warning on the web page cannot prevent data crawling acts. In judicial practice, common technical management measures include password authentication, IP restrictions, Robots Exclusion Protocol, encrypted codes, etc. In the "first criminal case involving crawler programs in Shanghai", the defendant developed a crawler program to bypass the encrypted code protection measures of the Dewu APP for data crawling, which was determined to have sabotaged technical management measures. This case indicates that the core of technical management measures lies in their ability to effectively restrict unauthorized access and possess identifiable protective functions.

2.2.3 "Dual Harmful Consequences"

The special data protection clause requires that improper data crawling acts must simultaneously cause dual consequences of "harming the legitimate rights and interests of other operators" and "disrupting the market competition order". Regarding the relationship between the two, there are different views in the academic circle: one view holds that the dual consequences are concurrent elements, which must be satisfied simultaneously to constitute unfair competition; the other view holds that "harming the legitimate rights and interests of other operators" is the core consequence, and

"disrupting the market competition order" is a natural result. If a data crawling act only harms the interests of individual operators but improves the total social welfare, it should not be recognized as disrupting the market competition order [6]. This view is supported by judicial practice. In the case of "Weibo v. Yunzhilian Case", the court held that platform operators should have a certain obligation of tolerance for the legitimate collection and utilization of public data.

2.3 Connection with Relevant Laws

The special data protection clause does not exist in isolation, and its application needs to be connected with laws such as the Personal Information Protection Law and the Data Security Law. The connection with the Personal Information Protection Law is mainly reflected in the regulation of crawling personal information data. According to Article 45 of the Personal Information Protection Law, the right to data portability of personal information endows users with the right to transfer personal information. If a third party crawls personal information data with user authorization and meets the conditions for exercising the right to portability, it should not be recognized as unfair competition.

The connection with the Data Security Law is reflected in the obligation to ensure data security. Data crawling acts shall not endanger data security. If a crawling act leads to data leakage or damage, it may not only constitute unfair competition but also violate the relevant provisions of the Data Security Law, and shall bear corresponding administrative liabilities or even criminal liabilities.

3 Key Controversial Issues in the Application of the New Law

3.1 Identification Standards for "Technical Management Measures"

The identification of "technical management measures" is a core difficulty in the application of the special data protection clause, and its controversial focus lies in the intensity requirements and form scope of technical measures. Regarding the intensity requirements, the academic circle generally advocates the adoption of a "reasonableness standard", that is, technical measures should be commensurate with the commercial value of the data and the difficulty of independent acquisition. Technical management measures do not need to reach the level of absolute protection; they only need to reflect the data holder's intention of confidentiality and achieve a certain protective effect. Regarding the form scope, technical management measures include both positive protective measures, such as encryption and firewalls, and negative warning measures, such as the Robots Exclusion Protocol and clauses prohibiting crawling in user agreements. However, it should be noted that the obviating effect of the Robots Exclusion Protocol needs to be distinguished by scenarios: the Robots Exclusion Protocol in the search engine field should comply with the non-discrimination principle, while in other fields, the Robots Exclusion Protocol can be freely set by enterprises on the premise of not constituting a monopoly. In the case of "Weibo v. ByteDance Case", the court held that the Robots Exclusion Protocol in non-search engine scenarios can be regarded as a technical management measure, but it must have legitimate reasons.

3.2 Burden of Proof and Identification of Dual Harmful Consequences

For the "dual harmful consequences" required by the special data protection clause, the burden of proof and identification standards directly affect the effect of legal application.

Regarding the burden of proof, according to the principle of "who claims, who proves", the data holder must prove that the crawling act has simultaneously harmed its legitimate rights and interests and the market competition order. However, in practice, the harm to the market competition order is difficult to prove directly and usually needs to be presumed through indirect evidence. Regarding the identification standards, "harming the legitimate rights and interests of other operators" is mainly manifested in the loss of competitive advantages of data holders; "disrupting the market competition order" is mainly manifested in "bad money driving out good money", that is, improper crawling acts reduce industry innovation incentives and lead to market competition disorder. In the case of "Shanghai Ganglian Company v. Shanghai Zongheng Jinri Company Case", the court held that the defendant's crawling of the plaintiff's steel industry market data through improper means constituted a "free-riding" act, which not only harmed the plaintiff's legitimate rights and interests but also disrupted the market competition order, and should be identified as unfair competition.

3.3 Legitimate Exceptions to Data Crawling

The application of the special data protection clause is not absolute, and there are various legitimate exceptions, mainly including fair use, exercise of the right to data portability, and access to essential facilities.

Fair use refers to the act of crawling data in a small amount and non-competitively for purposes such as public interests and academic research without harming the legitimate rights and interests of data holders.

Data crawling resulting from the exercise of the right to data portability refers to the act of a third party crawling personal information data for the purpose of data transfer with user authorization. This kind of crawling act conforms to the legislative purpose of the Personal Information Protection Law, can break data blockades, and promote market competition, so it should not be identified as unfair competition. However, it should be noted that the exercise of the

right to data portability shall not adversely affect the rights of others. If the crawling act involves the personal information or trade secrets of other users, it may still constitute unfair competition.

Access to essential facilities means that when data constitutes an essential facility, the data holder shall not refuse reasonable crawling requests through technical management measures, otherwise it may constitute a monopolistic act.

4 Suggestions for Improving the Legal Regulation of Data Crawling

First, it is suggested to clarify the specific types and intensity standards of "technical management measures" through judicial interpretations or guiding cases. Specifically, technical management measures can be divided into three categories: first, access control measures, such as password authentication and IP restrictions; second, protective measures, such as encrypted codes and firewalls; third, warning measures, such as the Robots Exclusion Protocol and clauses prohibiting crawling in user agreements. Regarding the intensity standard, a "reasonable protection" standard should be adopted, that is, technical measures only need to effectively prevent general unauthorized access and reflect the data holder's intention of confidentiality, without reaching the level of absolute protection. For those who only issue a statement prohibiting crawling without taking any technical protective measures, it should not be recognized as "technical management measures"; for those who have taken certain technical measures but with low intensity, it is necessary to comprehensively judge whether they are worthy of protection in combination with data value and usage purposes.

Second, it is suggested to refine the identification rules for "harming the legitimate rights and interests of other operators" and "disrupting the market competition order" to reduce the difficulty of proof. For "harming the legitimate rights and interests of other operators", specific situations can be enumerated; for "disrupting the market competition order", a presumption rule can be adopted, that is, when the crawling act meets the elements of improper means and leads to market competition disorder, it is presumed that the market competition order has been disrupted, and the defendant is allowed to adduce evidence to refute. At the same time, the quantitative standards for harmful consequences should also be clarified.

Third, construct a diversified dispute resolution mechanism. Data crawling disputes are characterized by strong technicality and complex controversies. A single judicial remedy cannot meet the practical needs, so a diversified dispute resolution mechanism of "administrative regulation + judicial adjudication + industry self-regulation" should be constructed. In terms of administrative regulation, market supervision departments should strengthen daily supervision over data crawling acts, promptly investigate and deal with large-scale improper crawling acts, and give play to the efficiency advantage of administrative regulation; in terms of judicial adjudication, guiding cases should be issued to unify adjudication standards and clarify the application rules of the special data protection clause; in terms of industry self-regulation, Internet industry associations should be encouraged to formulate norms for data crawling acts, clarify the procedures and boundaries of legal crawling, and guide market entities to operate in compliance with the law.

Finally, strengthen cross-departmental legal coordination. The legal regulation of data crawling involves multiple departmental laws. It is necessary to strengthen the coordinated application of the Anti-Unfair Competition Law, the Personal Information Protection Law, the Data Security Law, and the Anti-Monopoly Law, and establish a cross-departmental coordination mechanism. At the same time, the connection between judicial and administrative law enforcement should be strengthened, and a case transfer mechanism should be established. For acts that not only constitute unfair competition but also violate administrative law provisions, they should be promptly transferred to the relevant administrative departments for handling; for those suspected of crimes, they should be transferred to judicial organs in accordance with the law for criminal liability, forming a regulatory synergy.

About the Author

Maocuo Zhou, female, born in April 2003, is of Tibetan ethnicity. She is from Qinghai Province and works at Minzu University of China. Her research focuses on law (legal studies), and she holds a master's degree.

References

[1] Wu Shuang, Qiu Siyu. Anti-Unfair Competition Law Regulation of Public Data Crawling from the Perspective of Behavior Differentiation[J]. Science Technology and Law, 2025(3): 1-13.

[2] Guo Chuankai. Research on the Legitimate Boundary of Data Crawling[J]. Law Review, 2024(2): 122-132.

[3] Gao Jiancheng. The Legitimacy of Data Crawling Behavior from the Perspective of Competition Law — The Evolution of Ideas from User Consent to the Right to Data Portability[J]. Journal of Jiangxi University of Finance and Economics, 2024(1): 112-123.

[4] Kong Xiangjun. Construction of Data Rights within the Framework of the Anti-Unfair Competition Law — Specific Design Plan of the "Special Data Protection Clause"[J]. Journal of Comparative Law, 2025(1): 50-74.

[5] Huang Wushuang. The Boundary of Legal Protection for Commercial Platform Data[J]. China Applied Jurisprudence, 2025(5): 123-135.

[6] Feng Bo, Liu Long, Fang Lin. Competitive Compliance Paths for Data Crawling and Anti-Crawling — Taking User Public Data as an Example[J]. Journal of Intelligence, 2023, 42(9): 51-56.