DOI: 10.23977/law.2026.050101
ISSN 2616-2296 Vol. 5 Num. 1

Criminal Regulation of Scraping Corporate Data in the Era of Digital Economy

Xingyu Zhao

Southwest Petroleum University, Chengdu, China

Keywords: Web crawler, data security interests, access rights, public data

Abstract: In the era of digital economy, due to the lack of mature and effective data access regulations in our country, the large-scale scraping and utilization of corporate public data have led to a severe expansion of criminal liability. To balance corporate data control and circulation, this can be approached by focusing on the legal interests protected by data crime legislation, shifting the legal interests safeguarded by criminal law from "data confidentiality" to "data utilization security", thereby assigning new connotations to the legal interest of "data security" in the digital economy. The criminal regulation of corporate data scraping should establish a "permission-centric" data access framework, decriminalizing non-malicious scraping of corporate public data to align with the developmental needs of the digital economy.

1. Introduction

Data, as a new type of production factor, serves as the foundation for digitalization, networking, and intelligence. It has swiftly integrated into various aspects such as production, distribution, circulation, consumption, and social service management, profoundly transforming production methods, lifestyles, and social governance. Data is a crucial strategic resource for industrial development, a key support for entrepreneurship across various industries, and the "oil" of the digital economy era. However, web crawling, as an efficient and rapid data acquisition method in the digital economy era, poses significant challenges to corporate data protection.

Web crawlers, also known as web robots, web ants, or web spiders, are programs or scripts that automatically crawl and store specific information from the internet based on rules set by humans in advance. With technological advancements, web crawlers have primarily evolved into four types: general web crawlers, focused web crawlers, incremental web crawlers, and deep web crawlers.

Web crawlers exhibit characteristics of automation, efficiency, and low technical barriers. By simply setting predefined rules in advance, they can automatically execute scripts to efficiently gather large amounts of data needed by the user, then extract and store it. The entry threshold for utilizing web crawlers is relatively low—individuals without software development expertise can directly employ crawler software to collect data. This leads to a situation where publicly available corporate data may be scraped by private or small businesses, enabling potential misuse. Current technical countermeasures typically include robots.txt protocols and anti-scraping measures, but these are often insufficient to effectively curb malicious scraping of corporate public data.

2. Criminal risks of crawling corporate data

In the era of digital economy, our country's legal regulation of web crawlers has become increasingly "strict" under the concept of strengthening data security, gradually shifting from the civil to the criminal field. [1] Based on the protection of enterprise data security, crawling enterprise data poses a risk of constituting a criminal offense. The types and methods of crawling enterprise data differ, and the criminal conviction also varies accordingly. When studying malicious crawling of enterprise data, the main focus is on whether data crawling behaviors that violate crawler protocols and break anti-crawling measures are malicious.

2.1 Analysis of crawled data types

Based on the types of data crawled, this article categorizes data into public data, open data, and completely non-open data according to varying degrees of openness. In academia, the meanings of "public" and "open" are often confused, but in essence, they have different meanings. Public data is not equal to open data, nor is it the same as public data. It has data security legal interests that deserve protection under criminal law. [2] "Public" focuses on whether the information content presented after data processing can be known by others; whereas "open" emphasizes whether data can be accessed and controlled by others. Open data refers to data resources provided by the government or other organizations to the public free of charge, which the public can freely access and use. [3] Open data relinquishes exclusive rights to the data, not only allowing others to access the data but also authorizing them to freely copy and crawl the data. [4] For example, government agencies may disclose some policy content on their official websites, but this does not mean that visitors can copy it at will. The relationship between the two can be summarized as follows: openness necessarily implies publicity, but publicity does not necessarily imply openness. Downloading and using open data does not constitute a crime, but crawling and using data disclosed by other enterprises may constitute unfair competition.

In 2017, Wuhan Yuanguang Company instructed its employees to use web crawlers to crawl information such as bus routes of Shenzhen Gumi Company and its competitors through means such as "changing IP addresses" and "cracking encryption systems". Although Gumi Company publicly provides data for users to use, it does not mean that other companies can crawl this information for their own company's profit. Later, the Nanshan District People's Court of Shenzhen City determined that Yuanguang Company constituted the crime of illegally obtaining data from computer information systems (hereinafter referred to as the Yuanguang case). [5]

2.2 Analysis from the perspective of data crawling methods

From the perspective of data crawling methods, web crawling behavior can be divided into the following three categories: First, crawling behavior that obtains consent from the data website; second, crawling that violates the authorization intention of the data website, typically by violating web crawler exclusion protocols, service agreements, etc.; third, crawling that deliberately avoids or forcibly breaks through the security measures of the data website, manifested as using forged device_id, UA, and IP, or decoding data and other means. The risk of crawling behavior that obtains consent from the data website mainly lies in whether there are issues with the scope of authorization for crawling data and the processing methods for the data.

In 2017, a verdict was handed down in the high-profile case of Shanghai Shengpin Company's illegal collection of network data through technical means. This was regarded as the first typical case in our country where criminal liability was pursued for using web crawlers to breach corporate data protection. The company bypassed system authentication by forging device identification codes and tampered with user identifiers and IP addresses to circumvent access restrictions, thereby scraping video catalogs, categorized content, and related user comments from the "Toutiao" platform owned by Beijing ByteDance. The court determined that this behavior violated Article 285, Paragraph 2, of Criminal Law, constituting the crime of illegally obtaining data from computer information systems. The involved company and its primary responsible persons were sentenced accordingly (hereinafter referred to as the "Shengpin case").[6] This case marked the first time it was clearly established that obtaining corporate data by breaking through anti-crawling measures may also constitute a crime, and that the partially open data of enterprises are also protected by criminal law. For the use of crawler technology to obtain unauthorized but publicly available information within a certain scope, it is suggested that the different protection values of information and data should be distinguished, and whether there is a need to protect data security should be discussed. Only then is it necessary to evaluate the crime of illegally obtaining data from computer information systems. [7]

2.3 Crawler protocols, anti crawling measures, and malicious identification

2.3.1 Web crawler protocol and maliciousness

As a subjective intention reference benchmark for enterprise data crawling behavior, the connotation and nature of the crawler protocol need to be clarified. This protocol is embodied in the robots.txt file located in the root directory of a website, and is used by data subject enterprises to indicate accessible or prohibited access path rules, thus providing clear guidance for data acquirers. According to the "Internet Search Engine Service Self-Regulation Convention", enterprises adhering to this convention must comply with international conventions and business norms, including the robot protocol. This protocol is placed on a par with industry conventions and business rules, indicating its corresponding normative value. It can be seen that the crawler protocol essentially belongs to the self-regulatory norms in the field of the Internet, aiming to guide the compliance boundaries of network data crawling behavior.

Although web crawler agreements do not possess formal legal effect, significant violations of such agreements can serve as a reference for determining whether the perpetrator possesses subjective malice. [8] In practice, there is a tendency to regard compliance with agreements as good faith and violation of agreements as malicious intent, but this viewpoint needs to be comprehensively evaluated in conjunction with the actual harm caused to legal interests by the behavior. For example, in the case of 360 v. Baidu, Baidu imposed discriminatory restrictions on a specific search engine through web crawler agreements. The court determined that this action hindered information flow, damaged competitive order, and constituted unfair competition. This indicates that if the web crawler agreement itself contains unreasonable restrictions, violating it may not necessarily constitute malicious behavior.

On the contrary, if data scraping does not violate the crawler agreement but essentially exceeds the boundaries of reasonable data usage, it may still be deemed as having subjective malice. For example, in the case of Zhao Haichang, although his behavior did not violate the agreement, the court did not accept his claim that it did not constitute a crime because it exceeded the reasonable limits of data disclosure. Furthermore, if a company does not have a clear agreement in place, but the data scraping behavior clearly harms its legitimate rights and interests, the determination of malicious behavior cannot be ruled out. For instance, in the case of Shengpin Company, the court did not deny its subjective malice on the grounds of a lack of agreement. In summary, the crawler agreement should be considered as one of the reference factors in determining malicious behavior, and substantive judgment needs to be made based on the specific circumstances of each case.

2.3.2 Anti-crawling measures and maliciousness

When criminal legislation is enacted to regulate acts involving the illegal acquisition of corporate data, whether the perpetrator employs technical means to bypass anti-crawler mechanisms is often an important criterion for determining their subjective malice. Therefore, it is necessary to clarify the basic concept and legal attributes of anti-crawler measures. The evolution of anti-crawler technology is closely related to the development of web crawler technology. To safeguard their own data rights and interests, internet companies continuously strengthen protective measures against crawlers. Currently, they mainly adopt methods such as identifying user agents (User-Agent), monitoring website traffic, deploying dynamic verification codes, and detecting cookie information to identify and block abnormal access behaviors. [9] The core purpose of such anti-crawler measures is to identify and intercept visits with abnormal behavioral characteristics, thereby preventing website data from being maliciously crawled. It can be seen that anti-crawler measures reflect the clear will of enterprises to protect data and have been transformed into specific technical protection practices.

Although the anti-crawler technology adopted by most enterprises is relatively mature, the corresponding evasion techniques of data crawlers are also constantly upgrading, often able to break through the protections set by enterprises. Such behavior of bypassing anti-crawling measures clearly violates the subjective willingness of enterprises to protect data, and breaking through protective measures itself may bring certain economic losses to enterprises. Therefore, it can be determined that such behavior of bypassing anti-crawling mechanisms to obtain enterprise data is significantly malicious. Considering the potential social harm caused by this behavior, it can be determined that it is criminally illegal.

3. Practical dilemmas reflected by the current judicial situation

3.1 The difficulty of determining the legality of crawling enterprise data

Relevant laws in our country stipulate that obtaining data requires the consent of the data subject, but they do not specify specific situations. The boundary between normal data acquisition and unauthorized data scraping is also blurred due to the lack of legal provisions. In terms of scraping corporate data, there is a legislative gap in criminal law, making it difficult to accurately combat the act of scraping corporate data.

Some scholars argue that "forcibly breaking through anti-crawling technology, invading computer information systems outside the fields of national affairs, national defense construction, and cutting-edge science and technology, and using crawler technology to obtain data within the system, constitutes the crime of illegally obtaining data from computer information systems according to the provisions of Article 285, Paragraph 2 of Criminal Law." Other scholars believe that the act of breaking through anti-crawling measures in the Shengpin case infringes upon the security legal interests of computer information systems, and punishing it is justified. Some viewpoints hold that anti-crawler measures and identity verification mechanisms belong to computer technology barriers, and when web crawlers circumvent or bypass technical barriers, it constitutes an invasion of computer information systems in the sense of criminal law.

In fact, both the Criminal Law and its judicial interpretations lack clear definitions of "intrusion into computer information systems" and "employing other technical means". Courts in different regions have inconsistent discretionary standards for anti-crawler evasion behaviors, which may lead to different judgments for the same case. In the Yuanguang case, the court deemed anti-crawler evasion behaviors as "employing other technical means", while in the Shengpin case, the court held that anti-crawler evasion behaviors constituted intrusion into computer information systems.

Judicial experiences both domestically and internationally indicate that judicial authorities tend to criminalize rather than decriminalize emerging technologies with risks, and then gradually adjust and refine evaluation criteria while protecting individual behaviors. In the digital economy era, such lagging approach not only deals a devastating blow to internet companies but also hinders the development of innovative technologies, severely impeding the growth of the data economy. Therefore, timely solutions must be proposed for the act of crawling corporate data, with accurate qualitative conviction, avoiding both blind criminalization and blind decriminalization.

3.2 The difficulty in defining the legal interests of enterprise data

The legal interests in criminal law refer to the personal life interests protected by criminal law. These interests not only include personal life, body, freedom, reputation, property, etc., but also encompass the national and social interests established on the basis of protecting personal interests. [10]

In existing Data Security Law, Civil Code, and Criminal Law, provisions on data protection are made, but the legal interest nature of data is not clearly defined. Identifying the legal interest nature of enterprise data plays a crucial role in determining what kind of crime constitutes the act of crawling enterprise data. When defining the legal interest nature of enterprise data, not only should the individual interests of the enterprise be considered, but also the impact of enterprise data on the public interests of the country and society should be taken into account, even given priority. In past judicial practice, judicial authorities have tended to identify malicious crawling of enterprise data as the crime of illegally obtaining data from computer information systems. This criterion, which simply criminalizes data crawling technology, can make the crime of illegally obtaining data from computer information systems a "pocket crime," greatly restricting people's access to and utilization of data and hindering the development of the digital economy. Furthermore, data possesses multiple composite legal interest characteristics such as data security and economic value. [11] Professor Zhang Mingkai once pointed out that the object of the act should not be directly regarded as the protected legal interest merely because it may have multiple attributes. Therefore, the determination of protected legal interests should be based on the role played by the data in different scenarios or the nature of the rights involved. [12] Specifically, a comprehensive judgment can be made by considering the purpose, motivation, and consequences of the perpetrator. If malicious crawling of enterprise data is aimed at obtaining personal information, copyright rights, or trade secrets, priority should be given to considering the legal interests of personal information and intellectual property rights, with property interests coming second. Based on the infringement of legal interests such as personal information, intellectual property rights, and trade secrets by the act, it is identified as the crime of illegally obtaining citizens' personal information, infringing copyright rights, or infringing trade secrets. For the identification of legal interests infringed by other enterprise data beyond the aforementioned data, one can start from its property nature.

3.3 Ambiguity in the determination of illegal intrusion

The risk of web scraping behavior constituting a criminal offense is primarily manifested in obtaining data without permission, breaking through or circumventing technical protection measures, or violating technical service agreements. To effectively protect corporate data, such behavior can be regulated through the crime of illegally obtaining data from computer information systems as stipulated in the Criminal Law. The core of determining whether this crime is constituted lies in confirming whether the perpetrator's act of breaking through or circumventing anti-scraping technical measures or violating technical agreements for the purpose of obtaining data constitutes "illegal intrusion" in the sense of criminal law. According to current legal provisions, the objective act of the crime of illegally obtaining data from computer information systems is manifested as "intruding into computer information systems" to obtain data. However, the legislative provisions define "illegal intrusion" in a relatively principled manner, mainly for the purpose of meeting the clarity and simplicity of the principle of legality in criminal law through typified expression, and its specific legal connotation has not been fully elaborated. The relevant provisions of Article 2 of the "Explanations on Several Issues Concerning the Application of Law in Handling Criminal Cases Endangering the Security of Computer Information Systems" provide important guidance for understanding "illegal intrusion". According to this judicial interpretation, "illegal intrusion" can be preliminarily interpreted as the act of "entering a computer information system without authorization or beyond authorization".

Due to the lack of judicial experience in this area , we can draw on the connotation of "unauthorized" in the Computer Fraud and Abuse Act (hereinafter referred to as CFAA) of the United States to explore the appropriate definition of "unauthorized" in our country judicial system. Initially, CFAA uses whether there is contract authorization as the judgment criterion, and websites can define the scope of data access through contracts. However, this approach essentially gives the power to judge criminality to the website in disguise, and the risk of criminalization for internet users is influenced by the website. At this point, some scholars have proposed the "code regulation theory", which believes that the act of bypassing the code barriers set by websites to obtain data is "unauthorized". Therefore, the essence of "unauthorized" can be understood as accessing computer information systems by avoiding or breaking through security protection measures. "Beyond authorization" is easier to understand, and it can be applied to our country local theories.

4. Regulatory Path for Crawling Corporate Data

4.1 Shift of focus in regulation

If obtaining publicly available corporate data from the internet is legal, then the focus of regulation should be placed on the use of such data. Crawling data that companies voluntarily make accessible does not constitute a crime, and there are many cases in practice where only the illegal use of data is penalized. For companies, opening up data not only helps establish connections with external customers and partners, but also reduces data integration costs and enhances efficiency and competitiveness through external innovation. When companies adopt data openness as a business strategy and actively choose to share the data they possess, crawling such data should not be considered a crime. However, there is a difference between publicly available data and open data. Even though they are similar in content and nature, the purpose of publicly available data is usually to inform users about relevant information, which users often need to access by visiting the company's website. This login and browsing behavior helps companies increase their visibility and enhance their competitiveness. If competitors are allowed to freely crawl a company's publicly available data and use it to develop similar products, it may lead to customer loss and serious damage to the data-owning company. Therefore, this article argues that crawling open corporate data does not constitute a crime; however, whether crawling publicly available data constitutes a crime needs to be comprehensively judged from both the illegality of the behavior and the illegality of the results. Malicious crawling of publicly available data may still constitute a crime.

4.2 The protection of criminal law is shifting from data confidentiality to data utilization security

In traditional theory, Article 286 of the Criminal Law is typically interpreted as protecting the legal interests pertaining to computer information system security. However, some scholars hold differing views, advocating that the legal interests protected by criminal law in data-related crimes should be data security. They emphasize that this security should refer to the security of the data content itself, rather than merely the security of the medium (computer system) in which the data is stored or transmitted. Only by directly protecting the data content can it meet the essential requirement of data security as a legal interest. Specifically, unauthorized access can compromise the confidentiality of data. The crime of illegally obtaining data from computer information systems precisely infringes upon the data subject's need for data confidentiality. This crime defines "illegal obtaining" as "intrusion into computer information systems or the use of other technical means," which directly undermines the confidentiality attribute of data. Therefore, "obtaining" behaviors that do not violate data confidentiality should not be evaluated as data crimes within the meaning of this crime. At the same time, other scholars add that this crime should also protect interests related to data utilization, in addition to those related to confidentiality. The utilizable value of data content cannot be ignored and should be included in the scope of data security legal interests. Whether the perpetrator reduces the value of data content by obtaining data through replication means, or causes the data rights holder to lose control and management of the data through other technical means, both constitute infringements on data utilization.

With the promulgation and implementation of the "Data Security Law", there have been new developments in academic discussions on the legal interests protected by the crime of illegally obtaining data from computer information systems. Although specific viewpoints have evolved, the core definition of the legal interests of this crime remains "data security". When discussing the criminal regulation of corporate data crawling behavior, it is necessary to base it on the basic goal of data governance established by the "Data Security Law", which is to promote the rational use and orderly flow of data while ensuring data security, achieving a balance between data security and data sharing.

4.3 Clarify the requirements for data classification

Crawling data at different levels of enterprises can have varying impacts on public interests, making it particularly important to clarify data classification requirements. According to Article 21 of the Data Security Law, it explicitly requires the establishment of a data classification and graded protection system based on the importance of data, addressing behaviors such as tampering, destruction, leakage, illegal acquisition, and illegal utilization. It emphasizes the protection of key data, with stricter protection for national core data. Meanwhile, Article 24 stipulates the establishment of a data security review system to conduct national security reviews on data that poses potential threats to national security.

Although the "Data Security Law" sets forth the requirement for tiered data protection, it does not provide detailed regulations and does not clearly define what constitutes important data. In 2020, the financial industry issued the "Guidelines for Tiered Data Security Classification of Financial Data Security", which categorizes data into five levels based on data type and the degree of harm caused to the data, from high to low: Level 5, Level 4, Level 3, Level 2, and Level 1. Level 5 data is primarily used for critical business purposes and is only accessible to certain personnel with specific internal permissions. If compromised, it could pose a threat to national security and seriously harm public interests. Level 4 data is primarily used for important business purposes and is only accessible to certain personnel with specific internal permissions. If compromised, it does not pose a threat to national security but may have a general impact on public interests or have a serious impact on individual citizens. Level 3 data is primarily used for important business purposes and is only accessible to certain personnel with specific internal permissions. If compromised, it does not affect national security but may have a minor impact on public interests or have a general impact on individual interests. Level 2 data is primarily used for general business purposes and is accessible to internal personnel. It does not affect national security or public interests and only has a minor impact on individual interests. Level 5 data is generally known and used by the public, without affecting national security, public interests, or generally causing adverse effects on individual interests.

The "Practical Guide to Cybersecurity Standards - Guidelines for Classification and Grading of Network Data" issued in 2021 categorizes data from high to low as core data, important data, and general data. Core data refers to national core data, which is defined in Article 21 of the Data Security Law as "data related to national security, national economic lifelines, important livelihood issues, major public interests, etc.", mainly including data that generally infringes national interests and severely infringes public interests; important data, if infringed, may damage national and social interests, mainly including data involving national interests that are slightly infringed and data involving public interests that are generally infringed; general data, if infringed, will not endanger national and social interests, but may have varying degrees of impact on personal interests depending on the method of infringement.

4.4 Decriminalization of non-malicious crawling of publicly available corporate data

Whether the norms in the field of web scraping constitute a crime cannot be judged solely based on form, but also on substance. [13] The act of scraping publicly available corporate data is generally difficult to identify as a crime, and non-malicious scraping of publicly available corporate data should not be punished as a crime. Decriminalizing non-malicious scraping of publicly available corporate data is conducive to reducing the burden on judicial practice and creating a harmonious and stable business environment. The key to distinguishing between goodwill and malice lies in determining whether the scraping behavior is legitimate. Since scraping is based on illegal access to computer information systems, how to define the "illegality" of access is an important basis for judging the legitimacy of scraping behavior. [14] Corporate business interfaces usually default to data disclosure, and anyone, including competing companies, has the right to access them. Accessing corporate data interfaces that have not set any anti-scraping measures is not "illegal", and the scraping behavior is legitimate and not malicious.

The criminal regulation of malicious crawling of publicly available corporate data faces challenges such as determining the boundaries between "crime and non-crime" and the ambiguity of the concept of data in judicial practice. On the one hand, the criminal law system needs to broaden the scope of data protection to ensure consistency and clarity in the connotation and denotation of the data concept; on the other hand, it should clarify the criteria for criminalizing malicious web crawling behavior based on the principle of restraint in criminal law. [15] Malicious crawling of publicly available corporate data should be subject to criminal law constraints, but decriminalizing non-malicious crawling of such data can effectively conserve judicial resources, avoid unnecessary disputes, and safeguard market vitality. Severe punishment for crawling publicly available corporate data is an inevitable trend in the future, but restrictions cannot be ignored. Excluding non-malicious crawling behavior can protect the flow of data between enterprises while adhering to the principle of restraint in criminal law.

5. Conclusion

In the era of digital economy, penalties for crawling public data of enterprises should be differentiated based on specific circumstances, thereby avoiding the severe expansion of criminalization. The protection of legal interests of public data of enterprises by criminal law should shift from "data confidentiality" to "data utilization security", and new data access rules should be established. For the act of crawling public data of enterprises, not only should the legal interests that may be infringed during the acquisition stage be considered, but also those infringed during the utilization stage. More focus should be shifted from the acquisition stage to the utilization stage, avoiding excessive expansion of the application of criminal law, effectively utilizing judicial resources, and thus accurately cracking down on illegal and criminal acts, adapting to the development needs of the digital economy.

References

[1] Yang Zhiqiong. Criminal Regulation of Web Crawlers in the Data Age [J]. Comparative Law Study, 2020, (04): 185-200.

[2] Liu Xianquan. Criminal Regulation of Illegal Access to Public Data [J]. Law Application, 2025, (08): 97-113.

[3] Bi Qiuling. Application of Open Data in Data Journalism [J]. Hubei Social Sciences. 2016(07):190-194.

[4] Tong Yunfeng. Research on the Limits of Criminal Regulation of Web Crawler Behavior in the Era of Big Data [J]. Journal of Dalian University of Technology. 2022(02):88-97.

[5] Case of Wuhan Yuanguang Company Illegally Obtaining Data from Computer Information System, Criminal Judgment of Nanshan District People's Court of Shenzhen, (2017) Yue 0305 Xing Chu No. 153.

[6] Illegal acquisition of computer information system data by Shanghai Shengpin Co., Ltd., criminal judgment of the People's Court of Haidian District, Beijing, (2017) Jing 1080 Xing Chu No. 2384.

[7] You Tao, Ji Lihui. Determining Criminal Liability for Data Collection Using Web Crawlers: A Perspective from the Case of "Shengpin Company" Illegally Obtaining Data from Computer Information Systems [J]. Law Application, 2019, (10): 3-10.

[8] Cao Lanxin. The Boundaries of Criminal Regulation for Malicious Data Crawling Behavior: From the Perspective of the Crime of Illegally Obtaining Data from Computer Information Systems [J]. Journal of China University of Petroleum (Social Science Edition), 2024, 40(02): 130-138.

[9] Huang Zihao, Zhang Shu. Research on the Impact of Web Crawlers on Internet Security and Anti-Crawling Strategies [J]. Science and Technology Innovation. 2021(10):120-121.

[10] Zhang Mingkai. Criminal Law (Sixth Edition) [M]. Beijing: Law Press. 2021: 78-393.

[11] Li Can. Basis for Determining Legal Interests Protected from Malicious Crimes [J]. Science of Law. 2023(06):43-57.

[12] Zhang Mingkai. Basis for Determining the Legal Interests Protected by Specific Crimes [J]. Science of Law. 2023(06):43-57.

[13] Liu Yanhong. Research on Criminal Regulation of Web Scraping Behavior: From the Perspective of Infringing on Citizens' Personal Information [J]. Politics and Law, 2019, (11): 16-29.

[14] Sun Jie. Criminal Regulation of Data Crawling [J]. Journal of Politics and Law, 2021, (03): 115-125.

[15] Sun Yongxing. Security Risks of Web Crawler Technology and Criminal Law Response [J]. Information Security and Communication Security, 2022, (12): 62-72.