

The Fault Determination of Platform Service Providers in Web Crawler Infringement

Xinyue Hu

*School of Law, Jilin University, Changchun, China
hxy050820@qq.com*

Abstract. Web crawler technology plays a key role in data circulation and industrial innovation. As the "gatekeepers" of cyberspace, platform service providers stand at the core of fault determination for liability division in web crawler infringement. Currently, however, fault determination is confronted with multiple problems, including vague statutory imputation principles, the absence of objective criteria for fault identification, and great difficulties in adducing evidence for subjective fault. Based on the basic principles of tort law, this paper focuses on the imputation mode, objective obligations and subjective attitude of platform service providers in web crawler infringement, and constructs a complete fault determination system for them. It establishes an imputation framework that takes fault liability as the general principle and presumption of fault liability as the exception. This paper also clarifies the boundaries of platform obligations from three dimensions: legal provisions, contractual agreements and industrial practices. In addition, it differentiates the judging standards for "actual knowledge or constructive knowledge" according to different types of information, so as to promote the balanced development between efficient utilization of data elements and protection of personal information.

Keywords: Web Crawler, Platform Service Provider, Fault Determination, Data Security, Personal Information Protection

1. Introduction

Web Crawler, also known as web spider or web robot, refers to scripts or programs that automatically collect network information in accordance with specific rules. Its core functions include traversing web pages, extracting data, as well as storing and analyzing the collected data. Essentially, it is a form of data interaction between data processors and data controllers, and also a vital technical carrier for information circulation in the data era [1]. Benefiting from the public attributes and circulation value of data in the digital economy, web crawler technology has achieved rapid development and wide application. Nevertheless, technology is a double-edged sword. The abuse of crawler technology has brought various harms. First, it infringes upon rights and interests related to personal information. Illegally collecting users' private information and identity data threatens personal information security. Second, it disrupts the competitive order of network platforms. Malicious crawling breaks platforms' data display rules and business models. Third, it

endangers network security. Frequent and large-scale crawling requests are likely to paralyze servers and disrupt the normal operation of network services.

Existing regulations do not set clear rules for determining the fault of platform service providers in crawler infringement, and unified identification criteria are still lacking in judicial practice. Some courts take "whether the robots protocol is violated" as the core basis for fault determination, while others comprehensively consider factors such as the degree of platform intervention and profit gains. The inconsistent adjudication criteria make the existing private law rules inadequate to cope with problems arising from crawler technology. Most of the current academic researches focus on regulating the behaviors of web crawler users [2]. In view of the above situation, this paper breaks the conventional thinking of merely regulating crawler users, and turns to discuss fault determination of platform service providers (data controllers) in crawler infringement. It aims to facilitate the application of imputation rules, urge platform service providers to fulfill their obligations of data security protection, and strike a balance between the protection of data rights and interests and the circulation of data.

2. Dilemmas in fault determination of platform service providers in web crawler infringement

From the perspective of platform service providers, there are three major dilemmas in current fault determination for web crawler infringement: the ambiguous application scope of statutory imputation principles, the lack of normative guidelines for objective fault identification, and great difficulties in adducing evidence for the subjective elements of "actual knowledge" and "constructive knowledge". Clarifying the above dilemmas is the prerequisite for formulating targeted and improved rules for platform fault determination.

2.1. Vague statutory imputation principles

The primary dilemma in fault determination of platform service providers in web crawler infringement lies in the ambiguity of statutory imputation principles, which leads to inconsistent judicial standards and prominent phenomenon of inconsistent judgments in similar cases. China's current legal system has few special provisions targeting web crawler infringement, and there is no unified and explicit legislative statement on the imputation principles applicable to platform service providers. In judicial practice, some courts refer to the fault liability principle for network service providers stipulated in the *Civil Code of the People's Republic of China* and require plaintiffs to adduce evidence. For example, in the country's first unfair competition case involving AI content generation services, the plaintiff proved that the "pseudo-original content" function provided by the defendant generated new articles through AI rewriting, which constituted an infringement of the plaintiff's rights [3]. In view of the concealment and complexity of web crawler technology, other courts tend to adopt the presumption of fault liability. They presume that platforms have failed to perform reasonable duty of care and require platforms to prove their innocence. A typical case is the lawsuit filed by Tencent against ByteDance for crawling articles from WeChat Official Accounts, where the court applied the presumption of fault liability. The disordered application of imputation principles deprives fault determination of a unified basis. It not only increases the plaintiff's burden of proof, but also blurs the boundary of platforms' liability, making it hard to realize fairness and rationality in tort liability determination. Meanwhile, it grants excessive discretionary power to judicial judges and hinders the unified application of laws [4].

2.2. Absence of objective criteria for fault determination

There are no clear and operable objective criteria to identify the fault of platform service providers in web crawler infringement, which makes it difficult to define what behaviors constitute fault. This is one of the core dilemmas in fault determination. Platform service providers engage in diverse behaviors, including offering technical support for crawlers, storing data acquired by crawlers, and reviewing and supervising crawling activities. Existing laws and judicial interpretations fail to distinguish fault determination criteria corresponding to different behaviors. On the one hand, for cases where platforms have actual or constructive knowledge of infringing crawling behaviors but fail to take remedial measures, the specific scope of the "reasonable duty of care" is not clarified. For instance, there are no explicit rules on the extent of review obligations platforms should perform and the standard of supervision over crawling activities. On the other hand, the boundary between platforms providing neutral technical services and actively participating in crawler infringement is blurred. Some platforms claim exemption from liability on the ground of "technological neutrality". In the case of data crawling involving Baidu Baike, the three defendants pleaded that they did not actively participate in crawler infringement [5]. In practice, it is difficult to distinguish whether a platform merely provides technical support or actively assists or abets crawler infringement. Without a unified standard for objective fault identification, judicial judges have to rely on discretionary judgment, which easily leads to biased rulings.

2.3. Difficulties in proving and ascertaining subjective "actual knowledge" and "constructive knowledge"

The determination of platforms' subjective fault, namely whether platforms have actual or constructive knowledge of infringing crawling behaviors, lacks a sound institutional framework, resulting in great difficulties in proof and identification. Subjective fault is internal and concealed, and cannot be fully reflected by external behaviors. Current laws also fail to specify the constituent elements and evidence rules for identifying platforms' subjective fault. First of all, plaintiffs struggle to prove that platforms have actual knowledge of infringing crawling activities. A platform's subjective cognition is usually reflected in internal decisions and communication records, which are inaccessible to plaintiffs. Second, there are no definite judging standards for confirming constructive knowledge. In judicial practice, judges generally make comprehensive judgments based on platforms' technical capabilities, industrial experience and the obviousness of infringing acts. However, different judges hold different judgment criteria, making the identification of constructive knowledge highly uncertain. In addition, web crawler technology updates and iterates rapidly, and platforms often lag behind in recognizing new types of crawler infringement. This further increases the difficulty of subjective fault determination. Some platforms are thus found at fault for failing to notice obvious infringements, which undermines the fairness of liability identification.

3. Statutory imputation principles for fault determination of platform service providers

In disputes over web crawler infringement, establishing scientific and reasonable imputation principles is the logical starting point for defining the liability of platform service providers. Imputation principles determine the distribution of the burden of proof and the criteria for fault identification. This paper argues that the imputation mode for platform service providers in web crawler infringement shall take fault liability as the general principle and presumption of fault liability as the exception. This institutional framework conforms to the basic legal theories of the

Tort Liability Chapter of the *Civil Code of the People's Republic of China*, and adapts to the complexity of various crawler infringement cases.

3.1. General imputation principle: adherence to fault liability

Paragraph 1 of Article 1165 of the *Civil Code of the People's Republic of China* stipulates: "A person who infringes upon another person's civil rights and interests due to fault and causes damage shall bear tort liability." For most crawler cases involving corporate data competition or the crawling of non-sensitive information, applying the general fault liability principle for torts is reasonable. First, data crawlers and data controllers (platform service providers) are generally comparable in technical capacity and management level, with no obvious information asymmetry. The application of fault liability ensures fair distribution of liability. For example, in the lawsuit between Company A and Company B, Company B adopted crawler technology to collect financial information from the news platform operated by Company A. Both parties are internet enterprises with equivalent technical strength. Second, disputes over data crawling involve complex technical and commercial judgments. The fault liability principle provides clear liability expectations and behavioral guidelines for platforms. Third, fault liability respects the freedom of conduct. It avoids imposing excessive preventive obligations on platforms, which may otherwise restrict data circulation and technological innovation.

Therefore, the application of the fault liability principle requires plaintiffs (parties whose data is crawled) to prove the existence of platforms' subjective fault and objective damages. This not only complies with the principle of fairness, but also enables both sides to conduct technical competition and market competition within reasonable boundaries. In the unfair competition dispute between Beijing Chuangrui Culture Media Co., Ltd. and Beijing Weimeng Chuangke Network Technology Co., Ltd., the Shuabao App operated by Chuangrui Company crawled a large number of short videos and comments from the Douyin platform. After trial, the court held that as a platform operator, Chuangrui Company had actual or constructive knowledge of crawling activities on its platform but failed to adopt effective prevention measures. Instead, it gained traffic and commercial benefits from such acts, which constituted obvious subjective fault and unfair competition liability [6]. This case fully demonstrates the application logic of the fault liability principle.

3.2. Special imputation principle: application of presumption of fault liability

The traditional fault liability principle can hardly cope with new types of torts emerging in the technological era, and this dilemma is particularly prominent in the field of information network infringement. Web crawler infringement is characterized by high mobility, anonymity and difficulty in investigation, so infringed parties often fail to identify the direct perpetrators. To urge network service providers to actively perform their obligations of information security protection and effectively regulate web crawler infringement, the presumption of fault liability shall be applied under the following three circumstances.

3.2.1. Active intervention by platform service providers

The presumption of fault liability shall apply when a platform not only provides passive network access or storage services, but also actively participates in or intervenes in the crawling, aggregation and recommendation of data. For example, in the lawsuit filed by Tianjin XX Technology Co., Ltd. against Guangzhou XX Computer Technology Co., Ltd., the defendant, a browser operator, was not

a purely neutral technology provider. It edited and sorted search results, aggregated sources of pirated novels, and designed structured display cards. The court ruled that as a network service provider, the defendant ought to know that the search results might link to infringing websites, yet it failed to take reasonable preventive measures and thus was at fault. In such cases, platforms engage in active editing, algorithmic recommendation or deep linking, which transforms them from "technical service providers" into de facto "content providers" or joint infringers. Under such circumstances, requiring platforms to prove their innocence, namely proving that they have fulfilled full review obligations, conforms to the legal principle that "the party who controls risks shall assume corresponding liabilities".

3.2.2. Platform gains direct economic benefits from crawling behaviors

If a platform service provider obtains direct economic benefits from third parties' use of crawler technology to collect and utilize its platform data, the platform shall bear a higher duty of care and be subject to the presumption of fault liability. In the case of Beijing Weibo Vision Technology Co., Ltd. v. Beijing Chuangrui Culture Media Co., Ltd. (Douyin v. Shuabao), the Shuabao App collected Douyin videos via technical or manual means. Although Chuangrui Company argued that it only operated a platform, the court found that the company obtained massive advertising revenue and user traffic through the Shuabao App [7]. For the original platform (Douyin), if it had actual knowledge that third parties crawled its data for commercial use but took no action, or even shared profits from such activities, it shall be presumed to be at fault. Economic gains break the neutral position of a platform, and give it both the motivation and capacity to supervise data flow.

3.2.3. Crawled information involves personality rights and causes substantial damages

This is the core scenario for the application of presumption of fault liability. Paragraph 1 of Article 69 of the *Personal Information Protection Law of the People's Republic of China* stipulates: "Where the processing of personal information infringes upon rights and interests in personal information and causes damages, the personal information processor shall be liable for compensation if it cannot prove that it is not at fault." This provision establishes the presumption of fault principle in the field of personal information processing. In web crawler infringement, platforms, as data controllers and processors, are generally in a dominant position with access to technical details and system logs, while information subjects (users) are in a disadvantaged position and cannot easily prove platforms' fault. Strict adherence to the general rule of "he who asserts shall prove" will make it impossible for users to obtain substantive remedies for infringed personal information. Moreover, once users' private or sensitive information is massively crawled and leaked in crawler infringement cases, it will cause irreparable mental damage and follow-up risks, which necessitates inclined protection for victims. For instance, in the second-instance judgment of Sina Weibo v. Maimai, the court emphasized users' right to know and right to choose concerning their personal information [8]. If a platform fails to block crawlers from collecting users' non-public information such as occupational backgrounds, educational experiences and contact lists, resulting in the exposure of users' privacy, users shall not be required to prove the platform's subjective malice. Instead, the platform shall be presumed to be at fault for failing to perform its gatekeeper obligations. This tendency of objective imputation aims to protect disadvantaged natural persons against the backdrop of unequal technological capabilities.

To sum up, the dual imputation system combining general fault liability and special presumption of fault liability can flexibly address the complex ecosystem of web crawler infringement. The fault

liability principle applies to crawler disputes arising from commercial competition between equal parties to encourage data circulation and technological innovation. The presumption of fault liability applies to cases involving platform active intervention, platform economic gains and infringement of personality rights, so as to force platforms to fulfill their obligations of data security protection.

4. Clarification of objective criteria for fault determination of platform service providers in web crawler infringement

The core of identifying the fault of platform service providers in web crawler infringement lies in judging whether they have violated specific obligations of data security protection. Relevant analysis can be carried out from three dimensions: the source of obligations, the content of obligations, and the boundaries and limits of obligations.

4.1. Classification and definition of obligations

In accordance with Article 10 of the *Personal Information Protection Law of the People's Republic of China*, Article 111 of the *Civil Code of the People's Republic of China* and Paragraph 1 of Article 32 of the *Data Security Law of the People's Republic of China*, actors including data controllers and data crawlers shall undertake negative passive obligations and positive active obligations. Combined with the scenarios of web crawling, the obligations of platform service providers can be classified into the following two categories.

4.1.1. Negative obligations: compliance with prohibitory norms

On the one hand, platform service providers shall not process information illegally. Specifically, they are prohibited from processing personal information by means of misrepresentation, fraud or coercion, and from illegally trading, providing or disclosing other people's personal information. On the other hand, platforms shall not engage in personal information processing activities that endanger national security and public interests, such as facilitating illegal crawlers or allowing crawlers to steal state secrets and core data.

4.1.2. Positive obligations: protection of data security

Drawing on the hierarchical data protection approach adopted by Ningbo Intermediate People's Court in the case of Taobao v. Meijing, the obligations of platform service providers shall vary according to the nature and authorization status of data [9].

(1) Obligation to Prevent Unauthorized or Malicious Crawling

This is the core security obligation of platforms. The identification of unauthorized or malicious crawling behaviors requires comprehensive consideration of multiple factors. First, crawling acts that violate explicit prohibitions in the robots protocol, such as crawling background data and private pages restricted by the robots protocol. Second, crawling acts that evade platforms' technical protection measures or exploit technical loopholes, such as bypassing login verification including verification codes and breaking IP access restrictions to collect non-public data. Third, large-scale crawling beyond reasonable limits, such as frequent access requests within a short period that overburden platform servers. Fourth, crawling acts targeting sensitive data, such as personal privacy information and trade secrets protected by special legal provisions. For example, in the case of Sina Weibo v. Maimai, Maimai continued to collect users' non-public information including occupational and educational backgrounds after terminating cooperation and refused to delete relevant data as

agreed, which obviously violated the platform's prevention obligation [8]. Platforms must adopt technical measures such as encryption, access control and traffic monitoring to identify and block such malicious crawlers and prevent user data leakage.

In response to the above behaviors, platforms shall establish a monitoring mechanism for crawling activities to identify and intercept abnormal access, regularly inspect crawler tools and plug-ins on their platforms, and promptly suspend services or delete data after receiving infringement notices from third-party platforms or users. In the unfair competition dispute between Hunan Yifang Software Co., Ltd. and Beijing Weimeng Chuangke Network Technology Co., Ltd., Yifang Software used crawlers to collect non-public background data of Weibo, including deleted posts and users' comments on others. Although Weibo adopted certain technical protection measures, it failed to effectively block the malicious crawling. The court ruled that Weibo failed to perform its prevention obligation and was therefore at fault [10].

(2) Obligation to Inform and Supervise Crawling of Authorized or Public Information

For authorized crawling or crawling of public information, platforms shall perform the obligation to inform users and guarantee users' right to know and right to choose. In the second-instance judgment of Sina Weibo v. Maimai, the court established the triple authorization principle: data controllers shall obtain users' consent when acquiring their personal information directly; when data crawlers acquire such information indirectly, they shall not only obtain authorization from data controllers, but also inform users of the purpose, method and scope of data use and obtain users' consent again. Violation of this principle may be deemed a breach of business ethics and constitute infringement under the *Anti-Unfair Competition Law*. This principle actually imposes an important gatekeeper obligation on platforms: they must conduct double compliance reviews in the process of data circulation to ensure that data transfer does not harm users' rights and interests [11]. In the unfair competition dispute between Beijing Taoyoutianxia Technology Co., Ltd. and Beijing Weimeng Chuangke Network Technology Co., Ltd., Taoyoutianxia collected Weibo information of non-Maimai users for contact recommendation without obtaining explicit user authorization or performing the obligation to inform users. The court held that the company violated the rules for legitimate data use and was at fault [8].

4.2. Boundaries of platform service providers' obligations: comprehensive consideration of multiple factors

The obligations of platform service providers are not unlimited. Their boundaries are determined by multiple factors, including platforms' understanding of data sources, capacity to control data usage, economic gains from data services, and the neutrality and legitimacy of applied technologies.

4.2.1. Positive correlation between control capacity and obligation level

The stronger a platform's control over data processing, the higher its duty of care. For example, large internet platforms with sophisticated technical monitoring systems and abundant management resources shall bear stricter obligations to prevent crawling. In contrast, small platforms with limited technical capacity shall be subject to appropriately relaxed obligation standards. In the unfair competition dispute between Beijing Chuangrui Culture Media Co., Ltd. and Beijing Weimeng Chuangke Network Technology Co., Ltd., the Shuabao App, as a medium-sized short video platform, possessed the technical capacity to monitor and prevent crawling activities but failed to take effective measures. The court ruled that it did not fulfill the duty of care matching its control capacity [6].

4.2.2. Positive correlation between commercial benefits and obligation level

The more commercial benefits a platform obtains, the higher its duty of care. If a platform gains huge profits directly from data crawling, such as profits from advertising and membership fees, it shall undertake stricter review and prevention obligations. If a platform only provides public welfare services or uses data for internal purposes, its obligation standards may be lowered accordingly.

4.2.3. Consideration of technological neutrality

If a platform provides neutral technical services such as general data storage and ordinary network access, and does not participate in or assist crawler infringement, its obligations shall be limited to the minimum requirements such as the notice-and-take-down rule. Some scholars hold that technological neutrality is the key to balancing platform liability and operational freedom. Platforms that merely provide technical support without fault shall not be held to excessive liability [12].

4.2.4. Consideration of platform types

Different types of platforms shall be subject to different standards of duty of care. Platforms specializing in data collection services such as public opinion monitoring and data analysis shall bear a higher duty to examine the legality of data sources than general technical service providers. They need to conduct reasonable reviews on the source, quality and legality of data. By comparison, general cloud service providers and software development platforms have relatively lower duties of care, and shall only be liable when they have actual or constructive knowledge that their services are used for infringement.

In conclusion, the determination of a platform service provider's fault in web crawler infringement shall be based on whether the platform has violated the aforesaid data protection obligations. A platform can be deemed objectively unlawful and subjectively at fault only when it fails to perform the obligations of identification, prevention, notification or supervision commensurate with its capacity.

5. Construction of subjective fault determination for platform service providers in web crawler infringement

The judgment of subjective state is the key to determining whether a platform service provider is at fault. In accordance with the spirit of the *Provisions of the Supreme People's Court on Several Issues Concerning the Application of Law in Hearing Civil Disputes over Infringement of the Right to Disseminate Information via Information Networks*, the fault of network service providers includes actual knowledge and constructive knowledge. Combined with the characteristics of web crawler infringement, this paper divides crawled data into three categories: government data, corporate data and personal data, and formulates differentiated criteria for judging "actual knowledge or constructive knowledge" for each category.

5.1. Government data: clarification of the red flag standard

Government data refers to data collected, stored and processed by government authorities, including public data and confidential data related to national security. Government data is featured with publicity and public welfare, and it underpins citizens' right to know and right of supervision. Therefore, the criteria for judging whether a platform has actual or constructive knowledge are

relatively lenient. It is generally required to prove that the platform clearly knows that crawling acts violate laws or damage public interests. To confirm a platform's actual knowledge of infringement involving government data, direct evidence is required, such as official notices issued by regulatory authorities or explicit agreements signed between the platform and crawler users regarding the crawling of government data. The judgment of constructive knowledge shall be made based on comprehensive analysis of objective circumstances, for example: government data is clearly marked as prohibited from batch crawling; crawling acts exceed the scope of reasonable use and violate the *Regulations on the Disclosure of Government Information*; the platform obtains improper benefits from crawling activities.

In short, the determination of platform fault in the crawling of government data shall prioritize citizens' right to know. When relevant data cannot be acquired from parties other than state organs either practically or economically, crawlers' right to know deserves higher priority of protection. Therefore, platforms shall not be easily deemed at fault for the crawling of government data, unless the crawling acts compromise the security of computer information systems, core data or state secrets [13].

Specific judging factors include whether the data is government information that shall be disclosed by law, whether crawling disturbs normal access of other users, and whether crawlers pay reasonable fees for data use. Given the insufficient opening of government data at present, we should give full play to the positive effect of technological empowerment. Crawling shall be permitted unless crawlers maliciously delete or alter government data, so as to promote the opening of government data and realize citizens' right to know.

5.2. Corporate data: clarification of the reasonable person standard

Corporate information mainly includes enterprises' operational data, trade secrets and non-sensitive business data. As mentioned above, the general fault liability principle applies to data crawling between enterprises, since data crawlers and data controllers are comparable in technical capacity. The reasonable person standard shall be adopted to judge whether platforms have actual or constructive knowledge of infringement. Specifically, if a reasonable network service provider with equivalent technical conditions can detect abnormal crawler traffic through conventional monitoring means — such as frequent access within a short period, access during abnormal time periods and abnormal user agent headers — and identifies that crawlers are stealing core corporate data such as price databases and inventory databases, yet the platform takes no action, the platform shall be deemed to have constructive knowledge of the infringement. For example, in the lawsuit between Company A and Company B, if Company A, as the platform operator, detects large-scale and high-frequency abnormal crawling of its financial news in the background but fails to take measures such as IP blocking, it shall be presumed to be at fault.

5.3. Personal information: establishment of objective imputation and risk prevention standard

The strictest criteria shall be adopted to judge the subjective fault of platforms in crawler acts involving natural persons' personal information. As stipulated in the *Personal Information Protection Law of the People's Republic of China*, platforms shall establish a sound hierarchical data protection system. If a platform fails to encrypt sensitive personal information or impose strict access control on API interfaces, making it easy for crawlers to steal data, it indicates major negligence in the platform's institutional design, and the platform shall be presumed to have constructive knowledge of potential risks. Claims based on personality rights generally adopt

objective imputation, which focuses on whether the act interferes with the right holder's right to privacy rather than the infringer's subjective fault. Nevertheless, the judgment of platforms' fault is still required when holding them jointly liable. The existence of actual or constructive knowledge of crawler infringement shall be verified by examining whether data crawlers have fulfilled the aforementioned positive and negative data protection obligations. Violation of these obligations will lead to a presumption of platform fault.

To sum up, the red flag standard applies to government data to safeguard public interests; the reasonable person standard applies to corporate data to balance competitive interests; the objective imputation and risk prevention standard applies to personal information to prioritize the protection of personal dignity. This layered system for subjective fault determination can accurately define the liability of platform service providers in web crawler infringement and achieve unified legal and social effects.

6. Conclusion

As a double-edged sword in the data era, web crawler technology undertakes the mission of data circulation and technological innovation, while carrying potential risks of rights infringement and order disruption. Faced with the growing number of crawler infringement disputes, simply punishing crawler users can no longer address the complex network ecosystem. It is necessary to turn attention to platform service providers that control data resources and technical capabilities. The core solution is to build a scientific system for fault determination. In terms of imputation principles, a dual structure taking fault liability as the general principle and presumption of fault liability as the exception shall be established. This framework avoids overburdening platforms and stifling digital technological innovation, and also prevents platforms from dereliction of duty due to lenient liability rules. In terms of objective fault determination, the violation of legal obligations shall be taken as the core, and platforms' prevention and supervision obligations shall be defined in accordance with the nature of data. In terms of subjective fault determination, the traditional single standard shall be abandoned. Differentiated criteria including the red flag standard, reasonable person standard, and objective imputation and risk prevention standard shall be applied to government data, corporate data and personal data respectively, so as to achieve precise balance among different legal interests.

With the improvement of data-related legislation and accumulation of judicial practice in the future, the rules for fault determination in web crawler infringement will become more explicit. We shall always adhere to the principle of attaching equal importance to technological neutrality and liability accountability, and balancing rights protection and industrial development. While curbing the abuse of crawler technology, we shall reserve space for the free flow of data elements under compliance requirements, so as to promote the sound and orderly development of the digital economy.

References

- [1] Mitchell, R. (2019). *Web Scraping of Python* (2nd ed.). People's Posts and Telecommunications Press.
- [2] Wang, H. (2023). The path to criminal liability determination for web crawling behavior: Data rights confirmation and interest balance. *Global Law Review*, 45(01), 57-73.
- [3] (2025) Zhejiang 01 Civil Final Judgment No. 3998.
- [4] (2019) Sichuan 01 Civil First Instance No. 5468.
- [5] (2024) Beijing 73 Civil Final Judgment No. 1424.
- [6] (2024) Beijing 0491 Civil First Instance No. 6175.
- [7] (2021) Beijing 73 Civil Final Judgment No. 1011.

- [8] (2016) Beijing 73 Civil Final Judgment No. 588.
- [9] (2018) Zhejiang 01 Civil Final Judgment No. 7312.
- [10] (2021) Beijing Civil Application for Retrial No. 5573.
- [11] Ni, N., & Wu, T. (2023). Research on the application of general provisions of the Anti-Unfair Competition Law to web crawling behavior. *China Price Supervision and Anti-Unfair Competition*, (03), 51-57.
- [12] Yang, N. (2024). On the foundation and connotation of the criminal law obligations of network service providers. *Journal of Ocean University of China*, (05), 94-97.
- [13] Xu, K. (2021). The legitimacy of data crawling and its boundaries. *China Legal Science*, (02), 181-183.