Science of Law Journal (2024)
Clausius Scientific Press, Canada

DOI: 10.23977/law.2024.030515
ISSN 2616-2296 Vol. 3 Num. 5

# *Typology and Judicial Limitations on the Criminalization of Crawling of Public Data*

**Sheng Shiwen**

*Beijing Normal University, No. 19, Xinjiekouwai Street, Haidian District, Beijing, China*
*202221040020@mail.bnu.edu.cn*

***Keywords:*** Web crawler, public data, crime of damaging computer information system

***Abstract:*** As internet access technology evolves, the era of big data has made web crawlers essential for data collection. However, the conflict between data collection and data protection is significant. Judicial limitations on criminalizing the use of web crawlers for public data collection must balance these interests. It is crucial to distinguish between legitimate data collection and activities that harm computer information systems. This paper summarizes the perspective of judicial cases to provide a clear definition of "public data." It uses legal hermeneutics to develop specific and actionable criteria for elements such as "violation of state regulations." This approach aims to clarify the criminalization of web crawlers' access to public data. Furthermore, it seeks to confirm the boundaries of the offense of using web crawlers to crawl public data. The paper identifies the types of data that can be collected and the technical methods that can be used. This helps avoid the indiscriminate placement of all acts of crawling public data within the scope of criminal penalties. In the current big data context, a balanced criminal law framework that recognizes the neutrality of technology and encourages innovation, while safeguarding data security and fair competition, is the basis on which to manage the complexity of web crawling activities effectively.

## 1. Elicitation of questions

In the era of big data, network information data has shown explosive growth, in which the value of the data also gradually appeared, individuals, enterprises, and so on to learn computer technology better, the use of big data analysis of the behavior of the relevant crowd, began to use the crawler program to efficiently and accurately crawl the target website data. This technology has been widely utilized in various fields of social life and has become one of the necessary means of network retrieval in the information age. However, the risk of crawling data will gradually appear, the behavior of crawling data will pose a threat to data security, constitute unfair competition, and cause damage to the computer information system. On the one hand, to enable more convenient and efficient access to data and improve the competitiveness of enterprises, the development of Internet-related technologies, including network crawler technology, is indispensable. Network crawler technology users need more stable laws to ensure the safety and legitimacy of this technology. On the other hand, considering information security and other factors, it is necessary to control network crawlers to reduce the harm they may cause to websites. The law needs to establish a clear bottom line for this kind of behavior.

## 2. Current status of the criminal law regime for the criminalization of the act of crawling public data

### 2.1. Ignoring the reality of the application of technology

This is often the case in judicial precedents, where the act of a web crawler user breaking through a website's anti-crawling measures is found to be a strong violation of the data discloser's wishes, and thus one of the bases for unauthorized crawling of data. In terms of how data exists in the real world, the data stored on a platform does not belong entirely to the platform, and in some scenarios, the platform is merely the manager of the data. This is especially evident in social network platforms, where users release information on social platforms, and the vast majority of users want to share this information, platforms cannot unilaterally restrict the collection of this information and data sharing, or else it will result in monopolization of information and freedom of speech and other issues. From the point of view of the use of technology, web crawler technology in the use of the process of presetting the website will use anti-climbing measures to restrict, then avoid anti-climbing measures as one of the basic functions of the crawler program. Due to the lack of technical understanding, technical means of web crawlers such as modifying the UA, bypassing the website access frequency control, and other measures to circumvent anti-climbing measures have mistakenly become the basis for considering the use of web crawlers to obtain data as a crime.[1] In the field of computer technology and scientific research, from the research against the crawler program can be concluded that, first, the website taking the anti-crawler measures is only a kind of maintenance of the normal order of the website program and does not mean that it completely prohibit the existence of the corresponding crawler program, and secondly, the field of computer research at the same time there is the research on the circumvention of the anti-crawler, which is not an illegal science and technology. These facts show that the confrontation between crawlers and anti-crawlers in the Internet industry is more of a technical game. If this balance is broken to identify and penalize the circumvention of anti-crawler measures as a criminal act, then numerous data-related technicians or become potential perpetrators of criminal acts. At present, there are still a large number of judicial practice to circumvent the anti-crawl measures as an objective invasive behavior as well as the subjective malignancy of the perpetrator. From the perspective of protecting social interests, it is also inappropriate to circumvent anti-crawling measures that are recognized as prohibited by criminal law. Otherwise, the technical game will transform into a legal confrontation. This could lead to platforms with open data deciding how their data can be used and accessed, which is problematic. A private entity does not have the power to prevent another private entity from collecting information that has already been disclosed to the community on the Internet. Allowing such control would result in a serious constitutional problem, where private entities could control the dissemination of citizens' speech acts in cyberspace.

### 2.2. Expansion of the scope of criminal law incrimination

Currently, criminal law has an excessive scope of criminalization for data-crawling acts, and judicial practice focuses only on whether the network crawler is unauthorized or exceeds the authorization to capture data, without making a detailed distinction between the technical characteristics of the network crawler, and the practice of not distinguishing between the objects of regulation leads to an expansion of the scope of application of the criminal law in the case of the criminal law regulating the network crawler's capture of publicly available data. From the viewpoint of the legal nature of criminal law, criminal law is the protection law of other laws, and the punishment measures of criminal law are severe. Therefore, criminal law should uphold the principle of moderation, and criminal law protection can only be applied when other sectoral laws fail to provide effective protection. At present, however, the criminal standard for data capture is too low,

and there is the problem of criminalizing data capture without distinguishing its nature and object. The expansion of criminal liability is not only detrimental to the innovative development of the digital economy but also limits the space for the application of other sectoral laws, resulting in a further blurring of civil and criminal liability for data scraping.

### 2.3. Synergies between sectoral laws

The legal regulation of data-scraping involves multiple subjects and multiple interests and involves the competing application of multiple sectoral laws. The legal nature and responsibility of different legal departments are different, and data crawling involves the comprehensive governance of multiple sectoral laws, so it is difficult to determine the specific type of responsibility in the system of judicial law, and the current status quo of the lack of legality boundaries of data crawling makes it difficult to effectively synergize between sectoral laws in the legal regulation of data crawling. In the few cases where the data captured by network crawlers is original or conforms to trade secrets, it can be regulated and remedied by intellectual property law. The anti-unfair competition law intellectual property law and other civil laws, there are adjusted by the legal relationship of the competition.[6] In the anti-unfair competition law intellectual property law and other civil laws, there is a competition. [2]In the judicial, there is no exhaustive rights protection, but directly apply the unfair competition law for unfair competition behavior, and should only be based on civil law for infringement of the case, can apply the provisions of the unfair competition law. In addition, the data crawling behavior also involves the criminal law of criminal determination, the violation of criminal law data crawling behavior will be held criminally responsible. Although the criminal illegality of the abuse of network crawlers to crawl public data for unfair competition has not been confirmed, and the boundary between the two has yet to be confirmed, from the perspective of the protection of legal interest infringement, it is considered that the Civil Law and the Anti-Unfair Competition Law are sufficient to carry out regulation, and there is no need to adopt the criminal law to carry out regulation. From the above situation, the act of crawling public data using web crawlers involves legal regulation between different sectoral laws, and effective synergy is needed in the judicial application, with criminal law as the underpinning protective law, and effective synergy with other sectoral laws in the management of the act of crawling public data.

## 3. Boundaries of criminalization of the act of crawling public data

To protect the openness of information and citizens' freedom of access to data, the use of web crawlers to collect public data should be supported by criminal law to a certain extent, excluding punishability. At the same time, when such behavior involves a violation of the crime of destroying computer information systems, the criminal law should be based on a thorough interpretation and application of the criminal act.

### 3.1. Categories of public data protected by criminal law

Data can be roughly divided into public data and non-public data according to the importance of its confidentiality and the degree of publicity. The criminal law community has not made a conceptual summary of public data. Through the search and integration of judicial cases, the author extracts the interpretation of open data and summarizes the meaning of open data in the sense of criminal law.

In the case of unfair competition for improperly capturing Sina Weibo data①, the court divided the platform's data into public data and non-public data with access privileges, and for non-public data

① Beijing Haidian District People's Court (2018) Beijing 0108 Minchu No. 28643 Civil Judgment.

with access privileges, the court held that it was not justified to crawl such data by technical means without obtaining access privileges. However, the definition of "setting access privileges" is not clear, and it needs to be further clarified whether the public data is limited to the user's public access or to the platform's authorization to use. In the case of Liang Moubing and Huifa Zhengxin Network Tort Liability Dispute[①], the court defined that the defendant reproduced the authoritative referee documents on the referee documents network as public data, and in the referee documents network, access to the public referee documents requires the subject to register, password login, so we can see that the access restrictions, in this case are excluded from the "set access privileges" field. The field of "setting access privileges". It is inferred that "setting access rights" refers to the right of access of non-arbitrary citizens. In addition, the restriction of public data includes whether general logged-in users can intuitively view and obtain it. In the unfair competition dispute case between Microcasting and Six Worlds of Companies [②], Six Worlds Company displayed user reward data and anchor live earnings on "Zucchini" on the Internet. This information is not accessible to ordinary logged-in users. Ordinary users logging into Jitterbug to watch real-time live broadcasts do not have access to information on the anchor's live broadcast earnings or the value of user rewards.The user's bounty information briefly appears only as a floating window. Therefore, the data crawled by Six Worlds Company is not public data but rather nonpublic data related to data security. In summary, we can summarize that the public data adjusted by criminal law is the data that can be obtained directly by any subject or by non-restricted means such as verification and registration. In short, there are two constituent elements for the determination of public data, one is that the data platform does not set specific access privileges; the other is that ordinary logged-in users can directly access it.

It is also worth noting that even for the data of public works, the reproduction of written works, music, computer software, and other works without the permission of the copyright holder, which is punishable by criminal penalties, may be recognized as a crime of copyright infringement. However, this does not mean that the act of crawling the data of public works belongs to the crime of illegally obtaining the computer information system, which needs to distinguish between the act of crawling the data by using web crawlers and the act of using the data after obtaining it. Only in terms of crawling this type of data, not involving the use of copying level, this type of behavior should not be subject to the regulation of the criminal law.

## 3.2. Positioning and identification of "violations of national regulations"

### 3.2.1. Positioning of "violations of national regulations"

In the case of the crime of sabotage of computer information systems, "violation of State regulations", as stipulated in the law, is a prerequisite for the establishment of the crime, but there are many different interpretations of the specific content of State regulations. The greater controversy over "violating state regulations" in this crime is centered on determining whether or not "violating state regulations" is a constituent element of this crime. In criminal law, "violating state regulations" is divided into two categories: constituent elements and non-constituent elements. This paper argues that in the case of the crime of damaging computer information systems, the description of "violating State regulations" has no substantive significance and is merely intended to indicate the harmfulness of the act.

When "violation of state regulations" as a constituent element is stipulated in the legal provisions, whether the perpetrator constitutes a crime must refer to the antecedent laws and regulations violated by the behavior, and the judicial organs must also check whether the behavior of the perpetrator

① Beijing Fourth Intermediate People's Court (2021) Beijing 04 Civil Final No. 71 Civil Judgment.
② Zhejiang Province Hangzhou Yuhang District People's Court (2021) Zhejiang 0110 Civil Final No. 2914 Civil Judgment.

violates the relevant laws and regulations, and is reflected in the reasons for the judgment. [3]In the judicial case judgments regarding the crime of destroying a computer information system, the offender's criminal behavior was determined to be illegal without indicating a violation of antecedent laws and regulations. This demonstrates that, in judicial practice, there is an implicit acknowledgment that the specific meaning of "violation of state regulations" in legal provisions does not have substantive significance. The purpose of the presence of "violation of state regulations" as a non-constitutive element in the legal provision is to emphasize that the act is not in conformity with the provisions of the law and is not legally justified. In this scenario, to affirm the unlawfulness of the act, there is no need to confirm whether the act violates the provisions of a specific law or regulation, nor is there any need to ascertain whether the act has gone through the administrative licensing procedure under administrative law; as long as the perpetrator has committed the act of destroying the computer information system, he or she has violated the provisions of the State, and there is no need for further explanation and clarification based on the antecedent law.

Based on the above analysis, the determination of "violation of State regulations" only requires consideration of whether the act falls within the three categories stipulated in Article 286 of the Criminal Law①. Web crawler crawling data is by the established procedures to simulate the access behavior of normal users, short-term large number of access or access to data and download. It does not involve deleting, modifying, or adding to the data, nor does it involve making or disseminating destructive programs. According to the above behavioral attributes of network crawlers, the use of network crawlers to crawl public data only involves "interference" with the function of computer information systems. It is further necessary to clarify two issues of factual determination in the case of the use of network crawlers to crawl public data, namely, whether the circumvention of anti-crawling measures additionally constitutes "intrusion" in the crime of invading a computer information system and the concept and extent of "interference".

### 3.2.2. Whether circumvention of anti-crawler measures constitutes "trespass" within the meaning of criminal law

From the viewpoint of the relevant legislative data, the Interpretation of the Criminal Law of the People's Republic of China that "intrusion, refers to the unauthorized or the consent of others, through technical means to enter the computer information system" [4]. Article 2 of the Interpretation stipulates that "intrusion" needs to have the nature of avoiding or breaking through the computer information security system, the most typical forms such as attacking servers, blasting, and cracking communication protocols. Anti-crawl measures include both compromising and non-compromising.

Compromise anti-crawling measures include web association conventions, robot protocols, etc. Robot protocols are public declarations by website owners that indicate the scope and type of data that web crawlers are allowed to crawl, usually by listing tags on the index of the website. In judicial practice, the violation of the Robots agreement has also become a criterion for judging whether the crawler behavior is illegal, for example, the judge wrote in the judgment of Baidu v. 360 case, "360 in the launch of the search engine at the beginning of the failure to comply with Baidu's Web site Robots agreement, its behavior is inappropriate, and should be subject to the appropriate adverse consequences." [5]The author believes that the effectiveness of the Robots agreement judgment to the principle of scenario, with or without violation of the website set up Robots agreement can reflect the subjective will of the data disclosure of the person prohibited from crawling, but it is not necessary.

① Article 286 of the Criminal Law, states that the three categories of cases refer to: "deleting, modifying, adding to, or interfering with the functions of a computer information system, causing the computer information system to be unable to operate normally", "carrying out operations on data and application programs that are stored, processed, or transmitted in a computer information system", "intentionally creating or spreading destructive programs such as computer viruses, affecting the normal operation of a computer system". operations of deletion, modification, addition", "intentional production and dissemination of computer viruses and other destructive programs, affecting the normal operation of computer systems"

For unauthorized or beyond the authorization can not only make a formal judgment but also need to be combined with the type of data being crawled to carry out substantive illegality judgment, to examine the network crawler crawling data behavior on the legal interests of the infringement or threat to the extent that the substance of the degree of punishable.[6]If the website visited by the crawler is unrestricted, meaning the data being crawled is public, then even if the website's Robots protocol indicates data that cannot be crawled, or if a termination notification letter is sent to the crawler party after the fact, or if IP blocking and other technical means are used to prohibit the crawler software from continuing to access the website. It is impossible to substantively prove that the act of obtaining data on the website was "unauthorized" data crawling.

Non-compromising anti-climbing measures refer to the use of specific IP access restrictions and other forms of crawlers to limit the data capture behavior, "anti-climbing measures", that is, to monitor, to prevent web crawlers from capturing the site's data a variety of technical measures. In practice, websites often take one or more of the above measures to prevent network crawlers from wantonly crawling, such as identifying crawlers through UA and setting restrictions on the frequency of access to specific IPs utilizing authentication code verification, and so on, in various ways at the same time. The perpetrator sets up simulated human operations in the web crawler program to complete some types of CAPTCHA tests and other breakthroughs in anti-crawling measures are recognized as circumventing anti-crawling measures. Whether crawling public data constitutes the crime of invading a computer information system mainly lies in the determination of whether the behavior of circumventing anti-crawl measures constitutes "invasion". According to the technical interpretation, the core feature of information system security protection measures is access control, i.e., verifying the identity of the user to authenticate whether the user has the authority to access data. However, the data behind the anti-crawling measures for crawling public data is generally open to the public and does not involve the issue of permissions.

In summary, it is not difficult to find that the above anti-climbing measures do not belong to the security measures, anti-climbing measures are to access the information system to limit the way, its role is to limit the rate of access, while the role of the security measures is to determine whether the user has the right to access to the computer and exclude those who do not have the right to access. In the final analysis, security protection measures always correspond to permissions, only when a particular network or system resources are in a closed state and there are access rights, there is room for security protection measures. [7] Anti-crawler measures do not involve issues of rights. Breaking through anti-crawler measures only affects the method of data access, assuming the right to access is already in place. Thus, a crawler program that spontaneously circumvents anti-crawling measures to obtain public data does not constitute an "intrusion."

### 3.2.3. Concept and definition of "interference"

Through the previous analysis of the main points of the case and the norms of the law, the focus of considering the criminal illegality of the act of crawling public data by a network crawler lies in whether or not the act "interferes" with the functioning of the computer information system and constitutes the crime of damaging the computer information system. According to the provisions of the Interpretation, "interference" refers to the destruction of the function of the computer system, so that the normal operation order of the computer information system is disrupted. The main violation is the normal operation of the computer system, and the damage is relative to the daily operation of the website. The use of network crawler crawling public data behavior generally does not have illegal but in the case of network crawler caused by the operation of the website failure, network crawler crawling public data behavior belongs to the category of "interference".

### 3.3. Definition of the subjective form of the crime

It is generally believed that the subjective guilt of the crime of intentional destruction of computer information systems is intentionality, i.e., the perpetrator's hope or indulgence in the implementation of the criminal act of destroying computer information systems and its harmful results. In the process of criminalization of network crawler behavior, from the perspective of a subjective and objective combination of the behavior of human judgment, it is necessary to examine whether the crawler subjectively has the intention to break through the network security, data security protection measures, and access, obtain relevant data. [6]From the cognitive factors, the perpetrator needs to recognize their behavior to crack or circumvent the website anti-climbing measures contrary to the data controller or data owner protection will; from the will factors, the perpetrator needs to choose to continue to crawl the data based on free will. When and only when the network crawler actors with both "awareness factors" and "will factors", the data crawling behavior can be punished for human crimes. In the scenario of crawling public data, the subjective purpose of the perpetrator can be judged in the following ways. First, to see whether the crawler program designed by the perpetrator violates the ROBOTS agreement of the website crawled, if the perpetrator sets up the program to crawl arbitrarily without regard to the content of the agreement, first of all, the perpetrator violates a certain degree of business ethics, and secondly, it can also reflect his subjective attitude, whether it is a good or bad attitude towards data security. Secondly, it is judged by whether the browsing volume of the clicked website in the crawler program set up by the perpetrator reaches one-third of the total browsing volume of the website. ①Thirdly, the purpose of obtaining the public data is analyzed by the perpetrator, if the act of obtaining data is used for study and research, commercial profit, etc., the crawling of the public data is to a certain extent for their use, and the website is not intentionally destructive.

## 4. Optimization of the Concept of Legal Regulation of Web Crawlers Crawling Public Data Behavior

### 4.1. Technology neutrality

Technology neutrality means that science and technology do not have the attributes of good and evil, and need to be utilized according to their functions. However, the specific behavior of technology use will cause different legal consequences and does not have neutrality. It can be seen that neutrality only refers to the neutrality of the function of the technology itself, and does not involve the neutrality of the use behavior, the law can not judge the good or bad of the technology itself, and the technology use behavior often does not have neutrality, and needs to be judged from the specific case to determine the legality of the case. In the case of crawling public data, the technology itself is not the object of the law, the law examines the harm caused by the technology behavior, and the development of the technology is to consider the specific function of the technology and the intention to solve the problem, the use of the technology based on the interests of the use of the consequences of the use of different. [8]Especially when the use of criminal law to regulate this type of cybercrime, should maintain the principle of criminal law modesty, should be more tolerant of the technology itself, encourage technological innovation, should not arbitrarily a technology to interpret illegal tendency. Judges need to further understand the connotation of neutrality of technology itself when adjudicating cases, accurately differentiate between the technology itself and the behavior of

① Article 16 of the Data Management Measures (Exposure Draft) published by the State Internet Information Office in May 2019, which states that "Web pollution campers taking automated means of accessing and collecting website data shall not impede the normal operation of the website; such behavior seriously affects the operation of the website, such as automated access to the collection of traffic exceeding one-third of the average daily traffic of the website, and the website requires that it be stopped Automated access collection shall cease when the website requests it to do so."

technology use, and respond positively to the technology itself, so that the establishment of technology neutrality in the legal sense can further promote the development and innovation of technology.

### 4.2. Duty of care

The individual perpetrator has the possibility of foreseeing and avoiding the processes that may cause damage initiated by the act of crawling open data by a web crawler. For open data due to obtaining the promise and authorization of the right holder, the data resource is changed from a private product to a digital public resource. [9]Even if the act of crawling open data may cause some damage to the data subject in terms of data security, the principle of victim consent should be upheld. The will of the legal subject should be respected, and the victim's promise should be uniformly characterized as a reason for blocking the conformity of the constituent elements. Out of the above principles, the network crawler technology users need to crawl the data in the full duty of care, on the one hand, should pay attention to the limits of crawling, pre-set crawling intensity, the victim's data carriers for adequate protection, on the other hand, should respect the data subject pre-set ROBOTS agreement, etc., the scope of the crawling data to limit.

### 4.3. Building a multifaceted and synergistic governance model

For the cases that do not meet the above criminal judgment standards, laws other than criminal law should be used to collaboratively regulate the act of crawling public data by network crawlers. When crawling public data behavior does not yet constitute civil law, anti-unfair competition law, and other predecessor law, criminal law should adhere to its modesty and negate the act of criminal illegality, which is to maintain the principle of unity of the legal order. Therefore, the judiciary must adhere to the principle of modesty of the criminal law, and first consider the illegal nature of the act of crawling public data in the commercial law, economic law, and other parties, rather than directly determining that the act meets the illegal circumstances stipulated in article 286 (1) of the Criminal Law, or else it does not conform to the underpinning adjusting normative nature of the criminal law.

## 5. Conclusion

Web crawlers belong to relatively new computer and network application technologies, and the use of new technologies is usually always accompanied by difficulties in legal characterization, especially in the absence of clear standards for the application of the current computer crime legislation, and the judicial authoritie’s’ will be faced with the challenge of how to accurately evaluate the criminal law of web crawlers. In the face of the risks posed by new technologies, the interpretation and application of criminal law should adhere to the bottom line rather than being overly aggressive. We should uphold the principle of technological neutrality, will crawl. the behavior of public data is limited only to the destruction of computer information system crime, so not only gives the network crawler this technology to fully develop the space but also to establish the network crawler criminal law system. At the criminal law level, three levels can be used to limit the criminalization of the act of using network crawlers to crawl public data. The first level is to determine whether the data crawled by the web crawler user belongs to the public data regulated by criminal law. The second level, judges the subjective form, if the actor for direct intent to destroy the computer system, there is no room for debate, but when the actor is in indirect form, should be combined with the actor's duty of care from the contingent level to comprehensively judge whether it is indirectly intentional. The third level, is to determine the network crawler crawling public data behavior caused by the harmful results to the degree of "serious consequences". In the criminal law system of network crawlers crawling public

data, the above process can be followed for criminalization, avoiding the complication of the penal disposition of the act, and also providing clearer guidance for the legal application of the act.

## References


*[1] Sun Yu. (2021) Does forcible crawling of public data constitute a crime. Journal of the National Prosecutors College, 29(06), 121-139.*

*[2] Zeng Fengchen. (2021) Doctrinal Development of Judicial Policy on the Relationship between the Anti-Unfair Competition Law and Intellectual Property Law. Jiao Tong University Law, 02,157-168.*

*[3] Chang Yaoxinyu. (2021) Study on "violating state regulations" in criminal law. People's Public Security University of China.*

*[4] Lang Sheng. (2015) Interpretation of the Criminal Law of the People's Republic of China (Sixth Edition). Beijing: Law press, 491.*

*[5] Liu Yanhong. (2019) Research on Criminal Regulation of Network Crawler Behavior--Taking the Perspective of the Crime of Infringing on Citizens' Personal Information. Politics and Law, 11, 16-29.*

*[6] Liu Yanhong, Yang Zhiqiong. (2020) Research on the criminalization standard and path of network crawlers. People's Procuratorate, 15, 26-31.*

*[7] Jiang Jinliang, Zhang Dandan, Xu Dongqing. (2021) The Criminal Laws and Regulations on the Excessive Use of Web Crawler Technology. People's Justice, 02, 25-28.*

*[8] Wu Taixuan, Guo Baosheng. (2020) Technology Neutrality: Evidence and Application of Defense Subject Matters in New Types of Unfair Competition Cases on the Internet. Science and Technology Management Research, 40(20), 247-254.*

*[9] Tong Yunfeng. (2022) Study on the Limits of Criminal Laws and Regulations on Web Crawler Behavior in the Era of Big Data. Journal of Dalian University of Technology (Social Science Edition), 43(02), 88-97.*