Intelligent Automation And Soft Computing, 2019

Vol. 25, no. 4, 847–853
https://doi.org/10.31209/2019.100000088

Personal Data Security and Supervision in the Age of Large Data

DU Zhen-Yuan

Nanjing University of Chinese Medicine, Nanjing, Jiangsu Province, P.R. China

ABSTRACT

Big data's fast growth and development has triggered the attention of academic circles. Most of the studies are focusing on the concept of personal data, the definition of data rights, etc. In China, the current status of laws protecting personal data is still scarce and incomplete, relying on specific regulations that are lacking an overall architecture, which has not been efficient in the age of rapid development and application of big data. In the western countries with relatively mature personal data supervision system, the supervision principle is centered on the data subject, and it is difficult to adapt to the development and change of the current large data technology. This study analyses the current situation of large data security risk, discusses the supervision principle, supervision law of data security, right relief mechanism and supervision auxiliary mode, including the regulation of self-discipline organization of the industry, the introduction of credit system and so on, in order to build an effective personal data security supervision system and standardize the operation of the using personal data in whole industry.

KEYWORDS: Personal Data Security, Large Data, Supervision.

1 INTRODUCTION

THE development of big data has given birth to new economic growth points and which is important for reforming the business ecological environment and promoting economic development. The concept of big data was not proposed in recent years, but it became a hot word in the McKinsey report in 2011. Big data's fast growth and development has triggered the attention. How should big data be defined, whether the means to be used should be regulated, and whether companies should conduct industry self-regulation?

The issue of safety supervision has been taken seriously by the academic community in recent years. Personal data is readily accessible to Big data online players in countless industries such as medicine, education, consumer services. In China, the current status of laws protecting personal data is still scarce and incomplete, relying on specific regulations that are lacking an overall architecture. Those regulations on data privacy are embedded in old frameworks that are obsolete in regards of the fast-evolving Big data conundrum.

Available research in China on the issue of supervision of personal data security focuses on the concept of personal data, the definition of personal data rights, the value of personal data, and the supervision of personal data leakage. On the concept of personal data, the scholar believes the concept of Big data should be split in two, primary personal data and secondary personal data. The first is the basic information of an individual while the second is the result of the primary personal data being processed thereby receiving added value.(Ding Dao-Qin,2017)The scholar views data and information as separate concepts needing differentiation: data would represent the basic and raw data while information would be its digital tool.(Li Hai-Ying,2016)On definition of personal data rights, the scholar considers that data resources have become an important factor in promoting social development and economic growth.(Wu Xiao-ling,2016) Data resources are also an economic asset, which requires explicit property rights. Chen Xiao-Zhen believes that the individual is the absolute owner of the data resource and enjoys the property rights of the data. Whether the data is processed or not, the ownership is owned by the individual. (Chen Xiao-Zhen, 2016) Wang Li-Min believes that personal data should be divided into the right of personality, which is a positive right and a prior right. The right holder can request the actor to stop using the personal data at any time. (Wang Li-

CONTACT DU Zhen-Yuan wsfdu@126.com

Min, 2014) As proposed above, Ding Dao-Qin thinks that the data should be thought as binary, which would imply that data rights are dual: user data should belong to the personal user him/herself, while the processed data and its added-value should be property of data mining agent.(Ding Dao-Qin ,2017)

On personal data security supervision. Wang Rong suggested that data anonymity is an effective means for both privacy protection and data utilization, and that the anonymization of data can help achieve the identity secrecy of data.(Wang Rong,2014) Personal data protection legislation is lagging in China, government control of data cannot adapt to the changing situation. The scope and types of data the government controls should be defined, while transnational lawmaking should be sped up. (Feng Wei, 2016)

Western research pays more attention to the added value of personal data and explores new ways to supervise security of personal data. Western countries' relatively mature laws on data protection have put information at the center of the issue as any Big data "collector" (firm) is required to seek approval before being granted any possible gathering and use of personal data. Yet, those data protection laws are starting to come short in the face on rapidly evolving technologies rendered possible by an ever-changing internet landscape. Big data is defined as a huge data collection that exceeds the capacity of traditional database management and analysis in the McKinsey Global Data Analysis Institute's research report Big data: the next frontier of competition, innovation and productivity.(McKinsey Global Institute ,2011) Mehmet Cudi believes that the arrival of big data's era will inject new vitality into all industries and bring enormous benefits to society, such as medical care, education, communications, transportation, etc., all of which can make use of big data technology to create more value for the industry while improving the service level. (Mehmet Cudi, 2008) Ira S. Rubinstein thinks in the era of big data, the rule of "taking data as the main subject" is too backward-looking and it is difficult to adapt to the development and changes of current big data technology. The reason is because of the rapidly changing nature of the Internet, the use of data is becoming multidimensional rather than unidimensional and new ways are emerging that may deviate from the original purpose set by the data mining agent. Yet, through the course of operations, enterprises can neither predict the way data will be used nor supply immediate notification to the data subject (the user), from whom, therefore, they will not have received consent.(Ira S. Rubinstein,2013) British scholar Victor Tuo believes that the traditional means of supervision cannot meet the development of large data. For example, in the data collection stage, in principle, the acquisition and subsequent use of data requires the informed consent of the user, but the development of large data will make this "consent" difficult to achieve. (Victor Tuo,2013)

The above literature research generally discusses the development of big data and various aspects of personal data. Most of it focus on concept of personal data focuses, ownership of rights, protection and political debate. The theory and practice in data security supervision have been the basis for research, but it only as an argument to support a certain part of the research, does not discussed from overall data supervision. The theory of personal data security supervision has not been deeply expounded and analyzed, and the construction of the regulatory framework has not been involved.

2 BIG DATA SECURITY RISK ANALYSIS

2.1 The massive accumulation of data leads to a significant increase in the risk of leakage of personal information

THE Big Data exacerbates personal information leakage is mainly manifested in the following aspects: Firstly, Hackers attack big data systems that store personal information resulting in personal information being stolen. Organizations and enterprises that store personal information fail to protect information, also lead to the personal information be disclosed or be sold. Chinese internet Users Rights Protection Survey Report (2015)" shows that 63. 4% of internet users' online activity information such as call records and online shopping records were leaked and 78. 2% of internet users' personal identity information has been leaked. In addition, all kinds of applications and websites collect users' personal information other than the information they need to provide services, and dig this information deeply for accurate marketing, resulting in more information being leaked and used.

The main reasons for the above problems are: Firstly, a large collection of data makes it possible for a hacker to get more data at once, so hackers prefer to attack big data targets directly. Secondly, there are inherent contradictions in both security and privacy, consumers benefit from massive data mining, which could bring lower prices, more suitable goods, and the improvement of life quality, but also have to face the risk of disclosure of personal purchase preference, health and financial data. Thirdly, the cost of illegal behavior is low and the profit is large, this is the root cause of the illegal trade of personal information. According to a survey by the media, the annual output value of the underground black industry that sells personal information in China can reach billions of dollars. Certainly, China is cracking down on telecom fraud in an unprecedented way. The leakage of personal information is expected to be alleviated in the future.

2.2 New Big Data Technology Brings New Security Risks

Firstly, the use of new technologies in big data brings about new security risks. new No SQL (non-relational database) used in big data, while there is no strict access control and privacy management to maintain data security tool. Semi-structured and unstructured big data brings challenge to old security technologies and require security management technology of No SQL database in the future. Secondly, big data uses distributed architecture, which brings different storage mode and query mode, and coordinate tasks between multiple network communication sessions. However, the techniques used by many security products to monitor, analyze log files, discover data, and assess vulnerabilities do not work effectively in large data environments. Thirdly, the cloud computing supported by big data has brought a lot of new risks and new challenges, many new risks such as virtual machine isolation and shared storage brought by cloud technology also affect security of big data. (Xiaofang Li, Yanbin Zhuang & Simon X. Yang, 2017) Fourthly, the use of mobile devices for data collection, storage, access transmission, etc. brings new security risks, various phishing websites in the wireless Wi-Fi environment are hard to detect. (Xiong Yan, 2018)

The main reason for these problems is: Firstly, security protection technology always lags behind industry development. Due to the maturity of big data analysis tools and storage devices, more and more companies are able to collect and store large amounts of data while the corresponding security protection technology is progressing slowly. The security field needs to change its long-term passive status. Secondly, most of the security products have not been adjusted so that they cannot meet the big data cluster field, nor can they fully understand the information they collected. To solve the security problem of big data, we need to re-implement or fully design the architecture of most security tools.

3 BIG DATA SECURITY PREVENTION SYSTEM IS NOT PERFECT

THE governments of various countries are vigorously promoting the opening of government data, but there is no standard process of big data security. Lack of openness, protection, and processes for big data, it is difficult to achieve effective protection of data. Laws and regulations of big data security around the world's countries mainly focus only on personal information protection. Such as: Germany issued the Digital Protection Act in 1977; the Electronic Communications Privacy was promulgated in 1986 at the United States. China also issued the The NPC Standing Committee's decision to strengthen Internet information protection at the end of 2012. But countries do not cover other areas of big data security, such as what data are illegal and how are they legally used. Firstly, countries have not amended relevant laws and regulations for big data security.

4 NEWEST SUPERVISION PRINCIPLE

BIG data has gradually developed into a scheme that integrates various sectors, applying to all industries. Its supervision is no longer limited to some specific situations and will set the trend in data security protection. With characteristics of diversity, processing complexity, virtual transactions, the big data industry is producing a wide range of data security risks that are complex and take various forms. The traditional regulation theory has been unable to adapt to the big data industry under the new situations, therefore, it is necessary to update the inherent regulatory principles and build a personal data security monitoring system, which is more in line with the development of the times.

4.1 The Principle of Passive Permission And Restriction

At all stages of processing personal data requires consent from the data subject. There are two commonly methods for obtaining permissions, namely positive permission and negative permission. Active consent means that the user is deemed to have granted permission of use at every stage of data processing. Negative permission means that if the user have not refused permission by clearly, this is considered agreed. In the era of big data, if it is necessary to obtain absolute consent before collecting and processing personal data, it is very difficult and cost of operation are enormous.

During the negative consent process, not only does the final user receive notification that his/her data will be processed, but also is provided with a clear path to refusal. (Ye Wen-hui, 2015) Therefore, individual user rights could be protected with less social resources spent. It is a principle that adapts to the current development of big data. Except the data that involves personal privacy or personal safety, the principle of negative permission can be universally applied.

Personal data cannot be collected without restrictions. When the data be collected in larger number in wider scope, the potential damage will be greater. Therefore, one of the important principles of personal data security supervision is to prevent the collection of personal data without restrictions. The principle of limited personal data processing is reflected in the following three points: Firstly, the time limit for data collection and processing. It hinders data processors keeping user data indefinitely. The data should be promptly deleted or destroyed after the achievement of the established objectives. Secondly, the purpose of data collection and processing is to be limited. It requires that the collection and processing

of data should have a specific purpose, and that the purpose is legal and necessary. The collection and processing of personal data should strictly comply with the scope required by this purpose. Thirdly, the restriction on data collection and processing. It requires that the processing result of personal data cannot be misused, and must follow the agreed purpose.（Guo Yu，2012）

4.2 The Principle of Multilevel Supervision And Unification Of Control

AS big data is integrated with banks, internet, medical care, education, e-commerce and other fields that blend with each other, it is difficult to identify the body of regulation. At present, Chinese country has not specified the regulatory body for the big data industry, and currently mainly relies on adding to existing regulation governing the supervision of data-related businesses in different industries. Based on the distinct universality and multiple characteristics of the big data industry, a single regulatory framework with relatively simple rules on big data for various industries can no longer be used as a cornerstone of big data industry regulation development.(Dong Qin-Tan,2015)The big data industry currently lacks guiding principles on data control, Therefore, it is very important to establish a reasonable and legal unified regulatory body.

5 SUGGESTIONS FOR BIG DATA SECURITY SUPERVISION

5.1 Improve the Big Data Regulatory System

THERE are not many laws and regulations concerning personal data security in China at present. Big data monitoring system in China, mainly consisting of regulatory documents issued by the state or the local authorities, which has played a positive role in the development of the big data industry and the coordination of interests among the main bodies, but these regulatory documents have a limited role to be a leading. There is no clear provision to the data processing of each link of the standard operation. So it can't prevent big data's security risk. (Ma Can, 2016) Only in Provisions for the Protection of personal Information of Telecommunications and Internet users, Consumer rights and interests' protection law, Management of population Health Information (Trial), Network Security Law, Administration of Internet Information Services are some of the laws and regulations mentioned. As far as the current regulatory situation is concerned, the degree of supervision has been unable to adapt to the new situation of big data's rapid development of technology.

Chinese country urgently needs to improve the legal guarantee mechanism relating to big data industry, perfect the supervision system. It will protect the legitimate rights and interests of the parties involved in the data and provide effective institutional basis for ensuring and standardizing the development of big data's industry.

Therefore, it should combine the characteristics of big data industry, amendments and add them to existing laws and regulations, in order to increase the degree of supervision and protection. legal regulation measures as follows: further clarification of the characteristics of rights in personal data legislation: whether personal data is a personal right or a property right, or both having similar attributes should lead to a more detailed definition. At the same time, there should be a clear stipulation of the right to relief for the data subject in the legal regulations, leading to the introduction of the principle of inversion of proof, reducing the burden of proof of the victim, improve the judicial relief system and enforcement procedures, all in order to guarantee and relieve the rights of data subjects.

5.2 Establish an Effective Internal Control Mechanism

To have an effective regulatory system, in addition to a reasonable and perfect external laws and regulations, but also the effective internal functions of the organization and the good supervision of internal control mechanism construction, which is the basis for big data security supervision and play its due role. The regulatory authorities should regulate their own operating mechanism, clarify the regulatory areas and regulatory responsibilities, clarify their own powers and responsibilities, and cooperate with other relevant departments, in accordance with the principles and regulations of external supervision, the big data industry is supervised and the regulatory policy should be put in place. The internal construction of supervision departments should make an objective evaluation of their own shortcomings, and improve their own quality in many aspects, and update management thinking, establish a set of effective supervision and control procedures for big data, and provide better service to the data subject.

At the same time, coordination among different regulatory authorities is also indispensable for personal data security. Based on the characteristics of the current big data supervision, and the different regulatory departments do their things in their own ways. Separate supervision will easily lead to the unclear division of the scope of supervision and the ambiguous in supervision boundary, and may lead to regulatory gaps or overlapping. Therefore, the supervision of personal data security needs to coordinate with different regulatory departments, improve their work efficiency, and ensure the effectiveness of supervision. In particular, it can set up a big data supervision and coordination group under the Ministry of Industry and Information Technology, to carry out comprehensive coordination of data security issues, and co-ordinate the work of other

regulatory departments, other regulatory departments may share their policy changes with the coordination group in a timely manner, and cooperate with the coordination group's supervision work. Through joint cooperation to improve law enforcement efficiency, protect personal data security, and provide a strong legal guarantee for the healthy and sustainable development of the big data industry.

5.3 Perfect Right Remedy Mechanism

The rights and interests of data subjects are closely related to the sustainable development of big data. The current legal system for the supervision of the personal data security is still in its starting phase, and various regulations are not yet complete. The risk undertaken by the data subjects is relatively large. In addition, there might be illegal data transactions and personal data leakage. By perfecting the right remedy mechanisms, it could ensure the healthy development of the big data industry. Since the rights protection on the violations against the personal data rights could only rely on the appeals to the courts, this greatly increases the budget cost of individual rights protection. In addition to the prolonged duration of rights protection process, the actual successful right protections situations are less likely to be seen. Hence, there is a need to establish a new form of right remedy mechanism with lower budget cost and higher efficiency of rights protection. Big data supervisory departments could attempt to lead the establishment of relevant mediation committees who will be guided by judicial authorities. Together with the cooperation provided by the industry self-regulatory organisations, theoretical research provided by the social experts, and the regulations done on data security issues, the legal rights of data subjects could be secured.

5.4 Promote Self-regulatory Organization of Supervision of Industry

The advantage of industry self-regulation lies in the large amount of supervision coverage, flexible mode of supervision, and more obvious efficiency and effectiveness. As it is the first threshold of supervision, self-regulatory organization and supervision of industry can solve many problems in the early stage of development of big data industries. Therefore, big data areas should promote the establishment of big data class self-regulatory organization. At present, China has issued some data trading rules, in order to carry out self-regulation, but the industry has not formed a national self-regulatory organization with the new authority of the state. Big data industry can learn from other industry regulatory model, trying to set up big data industry association, the rules and risks of big data in the field of service standards guiding regulations for management. The association can establish different types of special management committees according to different processing links of the data, and make self-regulation and supervision on different links of data processing: the association also it can promote the healthy and sustainable development of big data industry according to the industrial rules of management, control, protection and other categories in different aspects of data processing.

Industry self-regulation should formulate industry rules in the course of operation and establish behavioral preferences of participants. It should make a binding clause on the standardized operation of the industry, put forward the draft business rules of the industry to form a more comprehensive set of self-discipline rules. Through the introduction of certification mechanisms, industry associations identified by law as eligible to assist in the regulation of data operators. Encourage companies to join trade associations, propagandize self-discipline rules, guide companies to consciously abide by the collection, storage, trading and use rules, and maximize the role of industry associations as far as possible.

5.5 Perfect Aid Techniques Of Supervision

5.5.1 Accelerate the Technological of Supervision

In the development of big data, the processing methods have become more updated day by day, and so do the ways to process transactions and payments. It would be difficult to sustain within the fast-changing big data industries with the single way of supervision measurement. For instance, a network data monitoring platform could be established under the guidance of supervisory department, in order to have real-time monitoring on every actions taken in every phases of data processing that could reduce the risk during data transaction. On the same time, appropriate admittance thresholds could be set and strict control should be put on companies that are involved in the transaction, so that the data security issues could be identified and tackled firstly.

5.5.2 Introduction of Information Credit System

As the most significant feature of the big data industry, information asymmetry is the main reason for the existence of data security risks. Data processing parties have a large amount of data and actual control rights, and it is difficult for data subjects to control their own data continuously after providing the data. This information asymmetry can lead to various data risks easily. Therefore, the introduction of information credit system is imminent. The credit system facilitates data exchange between data subjects and data processing institutions and captures the authenticity of transactions. The supervisory department can also set up a public platform for information disclosure. Both parties could inquire about the trustworthiness of the transaction counterpart. The supervisory department could

supervise and direct the transaction process, and publish the monitoring report regularly. It is possible to introduce the data transaction blacklist which contains names of the untrustworthy companies so that their transactions will be restricted, thereby creating disciplinary and deterrent effects on criminal offenses.

6 CONCLUSION AND RECOMMENDATION

THE age of big data, the amount of data grow with geometry speed, Personal data that exists in various industries is collected and is at risk of being abused. The legislation and technical measures for data privacy security are not perfect and effective. The existing regulatory model needs to be updated. It is necessary to build an effective personal data security supervision system through the improvement of regulatory principles, the advancement of regulatory technology, and the strengthening legal regulation of personal information collection and use.

During the process of information security management, pure technology cannot replace the sanctions and constraints by laws and society morality to privacy violation, the U.S. and the European Union has promulgated the privacy act, in order to regulate the behaviors of collection, usage and spreading of individual data; regarding the ″Individual information protection regulation of telecommunication and internet users″issued by China's Ministry of Industry and Information Technology on Jun, 2013, which has provided safety and laws and regulations guarantee over the collection and usage of individual information on the internet. It is shown that in the process of big data, the government is supposed to formulate, improve and complete the relative privacy act, therefore provide a stronger individual information protective barrier from the perspective of laws and regulations. Further research should be carried out in the area of strengthen information security and privacy legislation.

7 ACKNOWLEDGMENTS

THIS research is supported by the philosophy and Social Science Founding of Jiangsu Province (Project No.16FXB002) and Nanjing university of Chinese medicine. (Project No. ZYWH2017-16)

8 REFERENCES

Bad and Spam Information of networking of Reporting and Accepting Center in China Internet Association, (2015), A Survey Report on the Protection of Netizens' Rights and Interests in China.

Chen Xiao-Zhen, (2016), The Typing Analysis of the Title of Big Data: Logical Starting Point of Big Data Industry, Law and Economic, 2016(3),11-12

Ding Dao-Qin, (2017), Secondary Division of Basic Data and Value-added Data, Financial Law. 2017(2),5-6.

Dong Qin-Tan (2015), Personal Information Security Protection in Big Data Era, Computer Knowledge and Technology. 2015(29), 32-33.3.

Feng Wei, (2016), Suggestions on Promoting the Construction of China's Data Sovereign Legal System, China Information Security. 2016(3), 113-115.

Guo Yu (2012), Research on the Protection of Personal Data, Peking: Peking University Press, 2012, 213.

Ira S. Rubinstein, (2013), Big data: The end of privacy or a new beginning, International data privacy law advance access published.2013(5),11.

Li Hai-Ying, (2016), Legal Challenges and Recommendations for Big Data, Big Data. 2016 (2), DOI: 10.11959/j.issn.2096-0271.2016023.

Ma Can, (2016), User Privacy Protection and Countermeasures in Big Data Environment, Library and Information Libraries. 2016(5), 170-171.

Mckinsey Global Institute (2011), Big Data: The next frontier for innovation, competition and productivity. 2011 (5), 35-37.

Mehmet Cudi, (2008), On ethical and aspects of data mining, Journal of Yasar University. 25(3),13-16.

Merv Adrian (2016), Big Data, Teradata Magazine.(2016)1:11.

Viktor Mayr Schönberge & Kenneth Cooke (2013), The Age of Big Data: The Great Transformation of Life, Work, and Thinking,. Zhejiang: Zhejiang People's Publishing House, 2013,89.

Wang Li-Min (2014), What is the difference between personal information rights and privacy rights,[EB/OL].http://www.cssn.cn/zm/zm_bjtj/201403/t20140324_1040048.shtml.

Wang Rong, (2014), Legal Regulation of Data Anonymization, Information and Communication Technology. 2014(4),39-40.

Wikipedia. Big data [DB/OL].[2016-9-28]. http://en.Wikipedia.org/wiki/Big_data.

Wu Xiao-lin, (2016), Whose data is up to whom, China Securities Network., [EB/OL]. http://news.cnstock.com/news/sns_bwkx/201607/3838818.htm.

Xiaofang Li, Yanbin Zhuang & Simon X. Yang, (2017), Cloud Computing for Big Data Processing, Intelligent, Automation & Soft Computing, 2017(4),545-546.

Xiong Yan, (2018), Latest Technology of Big Data and Information Security, [EB/OL].[2018_03_18] http://wenku.baidu.com

Ye Wen-Hui, (2015), Operational model and supervision measures of big data credit bureaus—taking Alibaba's Sesame Credit as an example, Wuhan Finance, 2016(2),66-68.

9 NOTES OF CONTRIBUTORS

DU Zhen-Yuan, associate professor in Nanjing university of Chinese medicine, Research Field: Information law; Public health policy.