Privacy Revolution in the Big Data Age: from Privacy Protection to Personal Data Governance

Shujie Cui
Department of Law, Xi’an Jiaotong University, Xi’an 710049, China
cuishujie1990@163.com

Abstract. The arrival of the big data era also means the emergence of the privacy crisis. Nowadays, privacy is challenged and privacy protection theories have obvious limitations. In the big data age, the focus of privacy protection should be shifted from traditional physical protection to personal data governance. Personal data has many similarities with privacy but also some noted differences. Compared with privacy’s static and exclusive features, personal data is characterized by extraversion and fluidity, and is a combination of personal interests and property interests. Based on the above conditions and China’s data development practices, it is crucial to pay attention to the transformation process of the data governance model especially from the privacy protection to the personal data governance, which is the core of achieving the balance between personal data protection and data application. To promote personal data governance, the following strategies should be applied: a) adhering to the purpose of balancing data application and personal data protection; b) attaching importance to data values and allocating data interests on data flowing process; c) improving personal data protection legislation; d) inducing technological innovation.

1. Introduction
In the big data age, the amount of data grows rapidly and the application of data is deepening day and night. When big data releases data benefits, it also causes a series of social ethical worries, among which privacy infringement is the most serious one. Problems are caused by data application, such as data collecting and data mining. Privacy is exposed in the application process, and most of the time, people don’t know when and how their personal data being collected and the including privacy comes to light. Meanwhile, it is difficult for privacy protection to cope with challenges brought by big data technology, making privacy protection in trouble. Therefore, some scholars even argue that privacy has always died in the big data age.

Nowadays the issue of personal data governance has become a hot topic. More and more people prefer to focus on data application and data value rather than just pay attention to the privacy protection. In 2015, Chinese government issued the Action Guideline for Promoting the Development of Big Data, raising big data to the strategic level for the first time. Data governance, as a new governance model, is of great significance in national governance. At the same time, personal data governance, as an important part of data governance, is highly valued. The core of personal data governance is to balance the relationship between personal data protection and data application, promoting data development as well as ensuring the safety of personal data. This study focuses on the privacy revolution in the big data era, hoping to realize the transformation from privacy protection to personal data governance. Firstly, through the analysis of privacy crisis, the challenges posed by big data to privacy is presented. Secondly, we try to demonstrate the transformation process of the data
governance model especially from the privacy protection to the personal data governance. Last but not least, some strategies are proposed for promoting personal data governance.

2. Privacy Challenges in the Big Data Age
Big data not only brings revolution in data technology, but also all-around changes in economy, society and human thoughts. People tend to be transparent digital person and the privacy is challenged under such circumstances.

2.1. Big data technology increases the possibility of privacy invasion
In the industrial information age, data mainly exists in the form of words, sounds and graphics. Data resources are scarce and difficult to obtain. Government and the news media are the controlling parties of data resources. At that time, data is usually used for record and storage, which has high security and confidentiality. Many changes have been taken place in the data area due to the application of the big data technology. Big data technology has brought a large number of data industries and changed people’s lifestyle. Daily life, thoughts, actions and demands of people are digitized. Therefore, people are no longer just passive manufacturers of data, but data contributors that actively participate in the process of data revolution. With the surge of data volume, much privacy-related data comes out frequently. Although related enterprises adopt explicit privacy terms in data collecting and mining process to protect privacy, contents of the terms are complex and numerous and they are rarely seriously read. Worse still, a large number of data collecting platforms do not inform people that their personal data have been collected, more and more data traces are recorded unconsciously, which greatly increase the possibility of privacy infringement. Through anonymous processing and in-depth data mining, derived data and derived data relationships come into being, which are then widely used in carrying out targeted advertising, precision marketing, business judgment and so on. Data application is not controlled by individuals and privacy is likely to be exposed through deep using procedure. When we use personal health data to predict the future health status and disease status, personal health data may be leaked. Moreover, personal shopping preferences can be calculated according to the shopping data, and this may influence service attitudes towards customers, which is closely related to personal privacy and personal dignity[1].

2.2. Big data economy makes it inevitable for privacy to be shared
The new data processing and analysis mechanism have posed great impacts on the property economic system. In the era of industrial economy, economics laid more emphasis on the scarcity and limitation of resources. The property system is a supporting production relationship for the protection and effective utilization of resources. Nowadays, the big data economy, which relies heavily on the big data technology, is rising. Comparatively speaking, it is a sharing economic system, in which people are willing to share goods, services, data and intelligence[2]. It is supported by big data technology, which includes three important components of idle resources, sharing platforms and sharing participants. It connects people with different nationalities, races but with the same demands. In such circumstances, people take part in the sharing procedure and become active contributors. This kind of economy mode is based on massive data and data analysis, which exerts great challenges on the traditional property regime. Under the sharing economic style, data value is fully exploited and resources are shared of all the time. It is evitable that a large amount of privacy data has also been shared and becomes basic cells of big data economy.

2.3. Big data shakes the way of thinking in privacy category
People tend to be more acceptable to the openness of privacy in the big data era. With the continuous development and changes of people’s privacy views, their traditional awareness of privacy becomes diluted in the trend of big data. Long before, people couldn’t tolerate their photos being published, but now they are willing to upload their private photos on Facebook, Wechat and other social platforms. In the last century, the United States issued a special video privacy protection law to guarantee the
privacy of personal video rental records. However, in the big data age, films and television resources can be downloaded at will on the website. People do not care whether the downloaded records are collected by big data. In addition, another important change that big data has brought to modern society is the communication mode. The sharing of data has extended the traditional acquaintance communication to stranger communication and shifted materialistic thinking to relationship thinking. The acquaintance social circle established by blood ties and geography in the industrial age has been broken down because of the connection of big data. There, people of different regions and races could obtain each other’s data through big data communication so as to realize point-to-point communication and establish digital relationship. In this process, privacy is likely to be exposed. People are not sensitive to or even pay no attention to the exposure of privacy under this relationship mode, which is a reflection of the mismatch between privacy view and personal behaviour[3].

3. Changes from Privacy Protection to Personal Data Governance

Privacy in the traditional sense refers to the rights that citizens to enjoy the peace of private life and the privacy of private information without being illegally disturbed, known, collected, used and disclosed by others. To be specific, it includes the confidentiality of personal data and personal address as well as the tranquility of personal life. According to traditional privacy protection theory, privacy belongs to negative protection category, meaning that individual can only passively seek relief when his privacy is violated. Although privacy includes data and information privacy, it is more inclined to physical privacy and can be solidified in fixed time and space. The mobility and uncontrollability of data in the big data age has broken the original physical privacy protection. People begin to realize that the protection not only refer to the physical privacy and passive protection, but also the control and utilization of personal data[4].

Personal data can be used or associated with personal identification, which is some kind of two-dimensional electronic code in the Internet. Personal data has different name in different countries. Euro-American laws tend to call it personal data while Japanese likes to use personal information. Here, we regard them as personal data, the general term of electronic data code and the information it represents. Personal data is used for personal identification that distinguishes people from others. It includes personal identity data, personal social data and personal privacy data. Personal identity data refers to direct external identification such as name, telephone number, certificate number, home address, marital status and so on. Personal social data consists of personal occupation, social reputation, which is the main basis for social evaluation and class classification of human society. Personal privacy data, including physical and mental status, intelligence level, gene, preference, behavioral habit and other potential expression, is private and difficult for outsiders to grasp.

There are similarities between privacy and personal data, but they also have great differences, which are shown as follows: Firstly, privacy is exclusive while personal data is naturally extroverted. Privacy, such as personal body privacy and residential privacy, has physical entity attribute. Therefore, privacy right belongs to absolute right category in the civil right system. However, personal data is quite different for it related to but also separated from physical forms. It exists in the form of two-dimensional electronic code in Internet system and flows freely and invisibly within the system. Meanwhile, big data technology increases the possibility and frequency that personal data been obtained. Therefore, personal data is not fixed and exclusive, but naturally extroverted[5].

Secondly, privacy is static while personal data concerns dynamic data value. Privacy interest is static which can be limited to a fixed domain along with physical entities such as residential areas, information, emails, etc. Personal data can flow at various stages. It has diversified expressions and is more difficult to control than privacy. For example, in data collecting procedure, personal data mainly includes name, age, telephone number and address, while in data mining stage, it reflects in anonymously electronic code. Personal data is transferred alternately between different data subjects in data flowing processes. Now, people tend to pay more attention to the flowing and controlling of data rather than physical boundaries.
Thirdly, privacy only has personality interests, but personal data is an interest combination of personality and property. Personal data has personal attribute to identify individuals as well as economic values. It is widely used in many fields of economy and daily life, which helps to realize the digitalization, intellectualization, customization and greening of industries and promote the development of both Internet and sharing economy.

Privacy and personal data have different connotations and denotations and they also focus on different matters. We should concern the changes of the meaning of privacy in the big data age and focus on promoting the transformation from privacy protection to personal data governance.

4. Strategies for Promoting Personal Data Governance

4.1. Adhering to the purpose of balancing data application and personal data protection

Recently, the view of data opening and applying is widely accepted. Moreover, data value is being discovered day and night. In the small data age, data resources are scare and centralized, so data owners tend to possess and use data privately. The mainly use of personal data is storage and record. Since modern times, the value of personal data has been put forward, and the application of personal data has led to great economic and social changes. Digitalization has become a social trend and big data economy has also turned to the mainstream economic model. Therefore, data technology has already penetrated and changed people’s production and lifestyle. People enjoy intelligent and data-oriented services and are willing to share personal data, so they tend to become contributors, diggers and also enjoyers of data resources. Under such circumstances, we can’t limit the use of personal data and should encourage data application. Certainly, we must attach importance to the protection of personal data, so as to solve the problems caused by the use of personal data. This is the principle and main purpose in the big data age.

4.2. Concerning about data value and allocating data interests on data flowing process

Personal data is different from privacy in that it has more diversified value behind. Personal data governance should focus on data value and data interests distribution. According to contents of interests, personal data interests include personal interests and property interests. Personal interests contain undisturbed privacy and data controlling ability, while property data interests emphasize the economic values. We must regard personal data as a collection of personal and property interests so as to coordinate the conflicts and contradictions between personal interests and property interests. According to different interests subjects, personal data interests involve personal interests, public interests and third-party interests. As provider of personal data, individuals’ main interests demands are the protection of personal privacy and the relief of personal interests infringement. Government and public departments want to hold public interests on public issues. For example, the public hospitals hope to use personal medical data for in-depth and targeted medical analysis. The public transportation departments aim to apply personal travel data for the construction of intelligent transportation and smart city. Such kind of personal data is not exclusive, but can be used for many times, for many purposes and in many fields, which helps to make profits for public interests. In addition, big data has spawned a series of data industries and enterprises. They collect and mine personal data, so as to generate in-depth value of personal data and promote the development of big data economy. We should acknowledge that such behaviours to promote the innovation of data application but at the same time, they challenge the protection of personal data to a large extent.

Personal data governance must balance the relationship between data interests and interests subjects. As far as individuals are concerned, their personal data shall be mainly protected and their intrinsic value and dignity shall be safeguarded. Therefore, it is important for us to transfer some general personal data so as to promote data application and strengthen people’s awareness of data protection. What’s more, as the leader of public interests, the government and public departments should exploit public value of personal data and use them for social governance, as well as regulate the use of personal data through issuing policies without interfering the freedom of private life. In addition,
third-party interests subjects, such as data enterprises, are main creators and promoters of data value. They shall focus on the creation of economic interests of personal data through providing technology services, and they also need to ensure the security of personal data. For example, they could build self-disciplines, and constantly innovate data application technology and privacy protection technology, which will play a fundamental role in data practice.

4.3. Improving personal data protection legislation

Personal data is different from privacy because it covers a variety of types, so we need to distinguish different personal data in legislation and give protection and application licenses of different degrees, aiming to mediate the conflicts between personal data protection and data application. Firstly, from the perspective of contents, personal data can be divided into personal identity data, personal social data and personal privacy data. As mentioned above, personal identity data is the most directive external identity of individuals, while personal social data is the main basis for social evaluation and classification. Both of them are easy to be obtained, collected and applied. However, they are closely related to individuals. When we use such kinds of data, we must adopt certain restrictions to protect them, such as privacy clauses, notification and consent principles. Personal privacy data is closely related to privacy. It involves personal dignity and needs to be well protected. The application of such kind of data should be strictly limited unless required for the public interests and anonymously exploited. Besides, its usage process must be confidential. Secondly, personal data can be divided into personal original data and personal derived data according to the forms. Original data is highly identifiable which should be protected seriously. Derived data is processed by big data technology and mainly reflected in anonymous data sets and data flows. This kind of data is generally controlled by data developers and has great application value. Therefore, it should be allowed to be circulated and used in a wide range. However, due to the reverse identification and associated speculation of big data technology, derived data also can be used to identify individuals. So we must keep on concerning and regulating. Thirdly, according to the existence status, personal data can be divided into disclosed data and undisclosed data. The former data has been known to the public through legal approaches, such as personal data related to public interests. There is no need to get the agreements of individuals when such kind of data is collected and applied, but individuals should have rights to inquire about application situation of personal data related to public parts.

Besides, we should examine our existing legal system and find out the punishments and legal remedies for personal data infringement. According to current civil law systems in China, we can resort to the tort law, the contract law, the criminal law and so on. The tort law can protect privacy right, and the contract law helps to regulate data contractual relationships, moreover, the criminal law will punish the criminal offences. Certainly, features of personal data are not quite in line with existing legal rules and rules should be taken into account practically and measured when they are used. We will not discuss it here.

4.4. relying on technological innovation

Personal data, as a binary bit stream, has technical and implemental neutrality[6]. The application and protection problems of personal data in the big data age are caused by the use of technology ultimately. Technological problems need to be solved by technological innovation to a certain extent. Technological innovation has two key points. First, we should solve existing legal regulatory problems through technology. Privacy risk assessment is a typical example. When data enterprises collect personal data, they are usually equipped with privacy protection clauses, that is, they inform users which data will be collected and applied and what privacy protection risks they will meet. Currently, the principle of inform and consent is widely applied worldwide. However, big data technology has left such regulatory tools on the shelf because privacy clauses are too long and complex to read. Moreover, the in-depth mining of personal data has been strengthened in an unprecedented breadth and depth, which makes it difficult for data enterprises to foresee all future uses at the beginning of their collecting personal data. Therefore, the informed consent rules exist in name only. One of the
technical means to solve this regulatory problem is privacy risk assessment technology. That is to say, we shall evaluate privacy risks all the time through technical means so as to protect privacy to an extreme[7]. Second, original technical leaks shall be solved through technological innovation. One of the important technical means for data protection is the “anonymization”[8]. Anonymization is used to remove data that can identify individual. However, it is not absolutely unrecognizable. Reidentification at the technical level is common, so data reidentification technology needs to be solved through technological innovation.

5. Conclusion
Privacy has been challenged in the big data era. Big data increases the possibility of privacy invasion, makes privacy sharing inevitable, as well as shakes the way of thinking in privacy category. Now, the focus of data protection is changing from privacy protection to personal data governance. Both of them are important but have big differences. As a big data country, China should attach great importance to personal data governance. Four approaches should be highlighted for supporting the governance of personal data: a) adhering to the purpose of balancing data application and personal data protection; b) concerning about data value and allocating data interests on data flowing process; c) improving personal data protection legislation; d) relying on technological innovation. Admittedly, there are many ethical and legal problems not included in this paper, such as the legal duties and remedies in data application and governance process, the laws and policies worldwide and their interaction, the changes and applications of the current legal system etc., which are important parts in personal data governance. Due to time and length limitation, these parts are not included in this paper but will be studied later. Personal data governance is a crucial theme, we should encourage the development of data technology, meanwhile paying attention to the protection of personal data. What’s more, we need to confront the challenges so as to make some changes of our ethical views and legal system.

Acknowledgments
The author is most grateful to the Law School of Xi’an Jiaotong University. The research is funded by the projects of “Promoting the Construction of the Silk Road Economic Belt” (14JZD022) and “Research on Governance Innovation in Network Society” (15ZDA047).

References
[1] Xue Fu, Chen Hongbin. Analysis of privacy ethics of big data[J]. Studies in Dialectics of Nature, 2015,31(2):45(in Chinese).
[2] Li Zhen. The rise and the challenges of the sharing economy-an analysis from the perspective of economy and law[J]. Contemporary Economic Management, 2016,38(8):7(in Chinese).
[3] Lu Xianxiang. Sharing economy: minimize transaction cost, system reform and system supply[J]. Social Science Front, 2016(9):52-55(in Chinese).
[4] Liu Yahui, Zhang Tieying, et al. Personal privacy protection in the era of big data[J]. Journal of Computer Research and Development, 2015,52(1):229-243(in Chinese).
[5] Wu Weiguang. Criticism on personal data privacy protection under big data technology[J]. Political Science and Law, 2016(7):119 (in Chinese).
[6] Mei Xiaying. The legal properties of data and the position of data in civil law[J]. Social Science in China Press, 2016(9):164-183(in Chinese).
[7] Jing Yao. The legal basis standards of de-identification of personal information[J]. Law Review, 2017(3):128-130(in Chinese).
[8] S. L. Garfinkel. De-identification of personal information. Information Technology Laboratory National Institute of Standards and Technology, (Oct. 2015), 1-27. http://dx.doi.org/10.6028/NIST.IR.8053