Challenges and Countermeasures of Data Technology from the Perspective of Protection of Personal Data

Liang Zhao

Law School, Sichuan University, Chengdu, Sichuan, China

Keywords: Autonomy of citizens' personal data; Data technology; Incorporeal property; Protection

Abstract: The innovative development of data technology makes the social operation more efficient. At the same time, we should also realize that there are some hidden dangers in the collection, storage, analysis and research of data information. Therefore, we need to fully protect the autonomy of citizens' personal data from the dual aspects of law and business management to make a win-win situation.

1. Introduction

With the development of Information Technology, Data age is coming all-round. Data Age changes the way of information acquisition, transmission and application. Our daily life is recorded by digital signals through Internet, cookies, magnetic stripe, and so many information resource systems. Whether government agencies or public service centers, even private subjects such as companies and individuals, anyone could easily get others' basic information and manipulate data [1]. Data Technology (DT) completely changes the way people are known. However, we have to realize that the development of DT relies on the collection, storage, analysis and judgment of citizens' personal information. Meanwhile, it represents and declares that there is an expansion of 'power' in the private sphere. What we want to talk about in this essay is remind readers to pay attention to data security.

2. Challenge: Conflict between DT and Users' Data Protection

The basic principle of Data Technology is to group a signal data to form a regular data set systematically, so as to provide a reliable reference for forecasting trends and decision-making. Data Technology transforms useless data information possessed by individuals into intangible asset [2]. There are serval basic steps to finish this work: collect, store, analyze, and so on. Thus, this information could help companies to understand market pain points and users' demand, and finally help them to make profits. But we have to know that personal data also belongs to individuals [3], anyone who wants to use it have to get permission from users. So, there are two main arguments between Data Technology and users: one is that there is no strictly authorization when collecting and these data will be abused; the other one is that these data may be violated when it is stored by service providers, it must have a negative influence on people's privacy.

So there lies a challenge that how to use Data Technology to provide better products or services, at the same time, fairly use data and protect it. The current Data Technology using is not satisfactory in fact.

2.1 Personal data acquisition and use without strictly authorization

According to the laws and regulations of China and international practice, when collecting users' personal data, service providers need to obtain clear authorization from users, and make comprehensive use of the collected information under the limitation of contract with conscious of rational use. Although service providers and users are equal subjects in law, in fact, service providers are leaders and promoters of trade behavior who have more power. That means customers' choices are limited by service providers which causes the inequality of substantive rights and obligations.

DOI: 10.25236/icbdem.2020.020

When service providers collect users' personal data, the greatest issue is that users are forced to permit apps collect their personal details and operation data. The specific performance is, if users did not allow service providers to collect their data, they could not use this product or service. Next issue is that the service providers always get general authorization from users to ensure get full data they need for further use [4]. For example, when people surf the Internet, it will recommend some goods or service which users are interested in. It means browser software extracts user behavior data when using the device. Besides, the scope of authorization is uncertain. Users not only authorize service providers to collect, store and use data, but also agree they could share information with uncertain third-party organizations. All of these may expand the reasonable scope of information collection, and may infringe on users' privacy and autonomy of personal data.

When the data is utilized, there are two arguments to discuss. The first one is how to ensure the compliance in data use, including the usage condition, limitation, demands of permission, information desensitization. Secondly, data processed by specific algorithms will add its value, so whether the original owners of the data have the right to share equal profit [5].

Thus, it leads to an urgent problem to be solved, that is, how to define the attribute of value-added data. Of course, it belongs to intangible property, but it is different from trademark right and patent right. Even though it may contain some features of copyright when combined with some algorithms, it cannot be simply equated with copyright. So how to protect the intangible assets of enterprises?

Undoubtedly, these problems are seriously neglected in the development of Data Technology in China.

2.2 Potential risk of personal data leakage

After service providers collect users' data, do they protect it well? Obviously not. Users' data is at risk when it is stored and used. Recent years, a large number of data leakage events happened. Some of them are caused by business cooperation in the legal limbo, such as sharing data with third-party organizations. Some of them happened because of hackers attacking, which cause damage to both users and network platforms. The data might be stolen and platforms also suffer losses.

For users, there are three levels in personal data leakage risk. Firstly, personal basic data, which includes name, birthday, ID, phone number and so on, is stolen. It may totally expose a person in broad daylight [6]. Secondly, personal data is abused. That means someone falsely use users' account to steel money or cheat others [7]. Thirdly, some people analyze users' data and then do researches or make decisions so that to get profits [8], but users, who are data providers, have no chance to share the profits.

No matter which level the risk is, once the risk events happen and expose, the platform would suffer both reputation and economic losses. But sometimes, service providers are responsible for data leakage intentionally or not intentionally. It is not only for providing higher services, but also for earning more from users.

3. Analysis: Information Resource Systems and Collection Models Causes Conflict

When people enjoy benefits from Data Technology, for customers, they want to get convenient and high-quality products and service with cheaper price, including providing fewer personal details or other consideration. For service providers, they want to get more commercial gains and save costs as much as possible. Thus, how to get more valuable users' data is a way to make the target come true. Obviously, there is a conflict between them. Try to solve this problem, what we can do is to find the rules how the information resource systems work.

3.1 Basic information resource system in China

The core mechanism and elements of Data Technology are the establishment of various information platforms. There are three basic information resource systems in China, as in Figure 1.

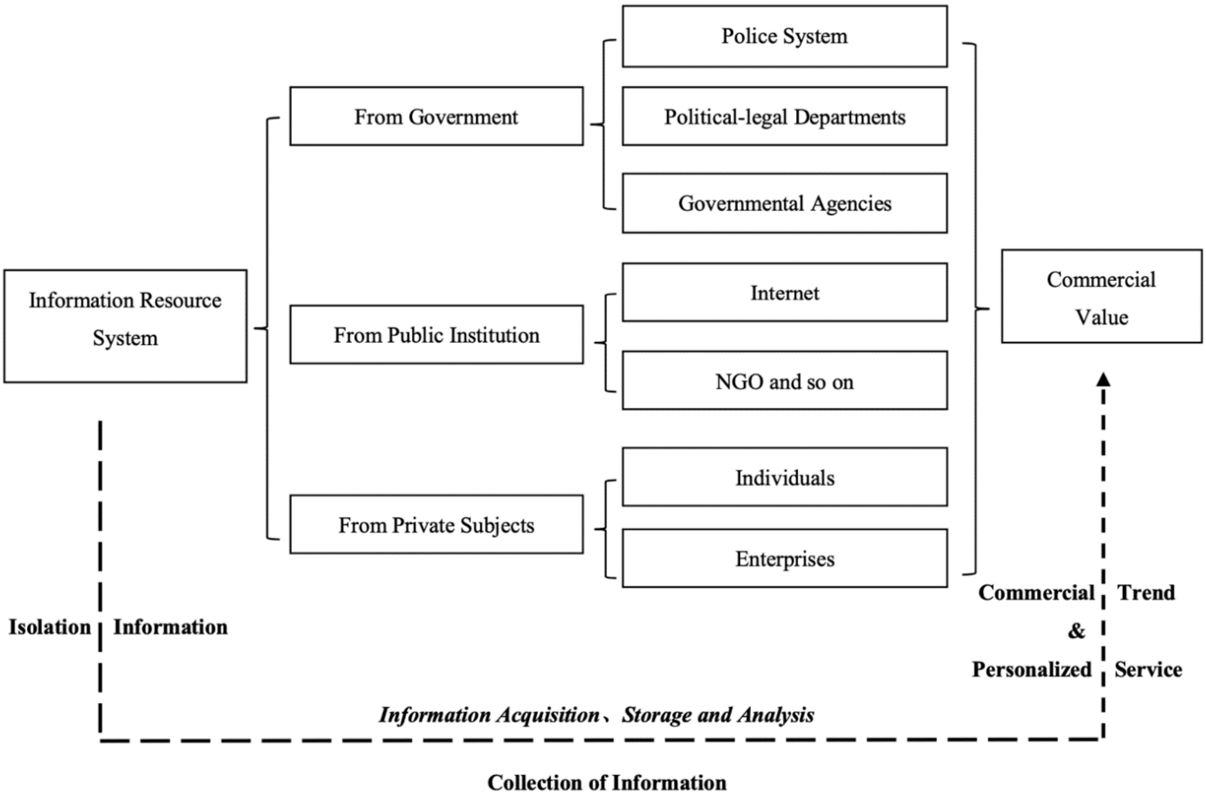

Figure 1 Information resource system

The first one is Government Information Resource System. It is established by government and functional departments. The data they collecting is used for national security or administrative management. It contains police system, political-legal departments information platforms, and other governmental information system. Because the government information resource system is controlled by state authorities, and it is very important to both nation and individuals, so it enjoys high security protection. Generally speaking, it has an isolation system with private LAN which is hard for everyone to access. And, it also has a relatively perfect information management standard.

Second one is Public Information Resource System which belongs to Non-governmental Organizations with social service attribute. The function of the system is to do academic research, pull social resource, and publicity non-secret information. This system always open to nearly all people over the world, public could find messages they need that published through Internet or some institutions. Of course, parts of core data about personal details or national security will be restricted access.

The third one is Private Subjects Information Resource System. Individuals and enterprises always establish their own data bases to collect information based on daily or commercial need, such as communication software, medical record systems, mobile payment. Every system has their own standards of data collection, storage, analysis, operation, and management and is not open to the public. But they are more likely to engage in illegal information transactions in order to gain benefits, because they have less fund and technological support.

3.2 Current modes of data collection and use

Ideally, information requester collecting, storing, and using information should obtain unambiguous and complete authorization from users, and use these data in a strict way. Contract will definite clearly about the scope of collection, ways of data storage, application of data desensitization, and limitation of data use [9]. But in practice, in order to improve the efficiency of information utilization and consider the further use of data, service providers will avoid it intentionally. So, current models of data collection include the following situations:

The first one is that information requesters steal data directly from users or third-party organizations who possess these data. It means they collect and utilize the data without any permission [10]. This phenomenon is rare happens in large enterprises with normalized operation, but we can often see it in some small and medium-sized enterprises. The most typical situation is that some commercial entities selling personal data as their main business often use this method to obtain illegal benefits.

The second one is collecting users' data in an indirect way. Information requester doesn't make use of terminal equipment or technical means, it relies on benefit exchange to buy and change users' data [10]. It contains illegal transactions, as well as strategic and cooperative information resource sharing models.

The third one is to obtain users' authorization in form but forcing them to provide personal data in reality. It is also the most common way of using various information platforms. We can find items with general authorization in the contract. If users want to get products or service, they have to agree all rules in the format contract [11]. Some provisions specifically target users to voluntarily agree to the service provider to collect and use their data. If users do not agree with any item of the contract, they are not allowed to get the product or service. Comparing with different contracts, the content is relatively long, and a large number of professional vocabularies is used in the text. Contents related to data collection and privacy protection are not separately indicated by special reminders. Meanwhile, the proportion of users reading and understanding the relevant provisions is low, because they are lack of legal risk awareness, or they just do not want to waste time. Therefore, although service providers have obtained the authorization from users, it does not represent they get real authorization intention from users.

3.3 Conflict caused by information resource systems and collection models

At present, there is a basic principle that information requesters have to get authorization if they need to collect, store and utilize users' data, except for extraordinary situation. But there are significant differences existing in different information systems and collection models. (See Table 1)

Table 1 Difference between information resource system and collection models

Information Resource System	Authorization Method	Scope of Authority	System Openness	Privacy Risks to Individual Donors
From Government	Forced Collection	Specifies Information	High Security	Low Risk
From Public Institution	Mostly No Special Authorization Required	All Information Mentioned or Provided	All Public	Medium Risk
From Private Subjects	General Authorization	The Contract Shall Prevail (But it has a wide range)	Limited Access but Unsafe	High Risk

The aim of government and public institutions getting personal data is to achieve effective governance of society and enhance prevention of social risks, so they could collect citizens' data based on law. Citizens are required to share part of their exclusive right of personal data to authorize social departments use their data. Even in some specific situations such as criminal investigation, data could be used without permission. In order to ensure data security, these data are placed in an independent system for proper storage, which is difficult for other units and individuals to access. But different from this, public information resource system is more open. When people get service from public institutions, they will provide their personal details, and the public institutions may keep it, such as the operation records when people surf the Internet. Some messages also have to be announced to the public. It means that these data are more easily to be violated.

Comparing with the last two systems, the system owed by individuals and enterprises always will be isolate. But in order to make more profit, the private subjects may share the data with others. And because of the willing to limit operating costs, they don't have enough capital to establish high-level data security system. So, the system is not safe enough as we think, though the system will not open to the public theoretically.

Overall, generally speaking, sort by system openness, information resource system established by government has the minimum openness, next is owed to private. People could get prodigious amounts of information through public information resource system, because it has to ensure the publicity and credibility of specific information. But there is a more complex situations in fact. Government and public institutions always have complete regulations to constraint their behavior, but there is no supervision mechanism for individuals and enterprises. So, the risk of information being leaked and abused is higher in the systems owed by private. Meanwhile, we find the users' data involved in this system is more comprehensive and private. Once the information is disclosed, the value of the information-based resource may cause more serious damage to people's rights.

4. Conclusion: Balance Data Use and Data Protection

We could find the risk of data collection and use from the analysis above. Firstly, although the Act of Personal Data Protection is being drafted, we do not have specific law to protect the right of autonomy of personal data until it is promulgated. The only regulation is our constitution which protect citizen's freedom and secrets of communication. It means protection for citizen's data is not comprehensive. Secondly, abuse of personal data may lead to wrong business judgment and decision-making which may cause economic losses directly. Thirdly, data collection and use without permission will increase the risk of data leakage and exposure, and encroach citizens' privacy. To cope with the conflict, we should focus on both sides of commercial benefits and protection of autonomy of citizens' personal data strategically, and make a balance between them.

4.1 Coping strategies

4.1.1 Protect the autonomy of citizens' personal data

The earliest investigation of personal data could be found in 1093 that Professor Prosser published an article named Right of Privacy. But this paper just come up with the idea that 'personal data' should be served as private affairs, it did not mention the concept of autonomy of citizens' personal data. In 20 centuries, the case of portraiture right infringement in the United States has promoted the legislation to protect citizens' personal data. The German Census Act of 1982 recognized autonomy of citizen's personal data as a basic constitutional right, and formed a series of landmark cases such as T-Grundrecht and Diary of Tagebuch. After that, countries around the world have followed suit, such as the British Data Protection Act 1998, the European General Data Protection Regulation.

What is personal data? The concept should include the following information (see Table 2):

Table 2 The Meaning of Personal Data

Information	Examples
Personal Details	Name, Date of birth, Sex, ID, Phone Number, etc.
Identity Characteristics	Biometric: face, print, iris, line, auricle shape, DNA, etc.
Social Information	Family Members, Marriage, Occupation, Education Background, Health Condition, credit information, Social Relationship, etc.
Behavior Data	Travel Record, Preference of Goods, Interests, etc.
Other Information	All about recognition of someone.

What 'autonomy of citizens' personal data' we call is a right that citizens could independently determine how to store and use their own personal information which can be used to identify their personal image and preferences. That is to say, all personal data that can be used to identify and confirm the one from others will become the object and scope of the protection of autonomy of personal data. Therefore, in a broad sense, all collection, storage and use of citizens' personal data without legal authorization will constitute a violation of autonomy of citizens' personal data.

In order to achieve autonomy of citizens' personal data, users have rights to know what their data is used and how to manage their information access. It needs the help of law and business ethics self-control.

4.1.2 Sustainable commercial benefits development

Pursuing of commercial benefits is the aim of data use. However, when develop Data Technology, leakage of users' data will make huge damage to commercial value and stock performance. For example, Facebook's stock declined about 2.59% after attacked by hackers in August, 2018 and its market value has fell down about 600 billion since by Cambridge Analytica Ltd abusing users' data events exposed in May, 2018 [12].

In some ways, users' information safety is consistent of sustainable commercial benefits [13]. If service providers cannot protect their clients' information safely, customers may choose another service provider, and the investors are also disappointed of them. Thus, commercial benefits will face a long term of recession.

In order to keep sustainable business profits, enterprises should build trust relationship with their clients. It cannot finish in one day, instead, it needs enterprises maintain good business ethics and take advantages of technology to protect users' data constantly.

4.2 Key point of coping

We should concentrate on these two aspects to make a balance between data protection and commercial profits increasing. One is to keep the end of the negotiation to protect citizens by laws, the other is to regulate service providers behave of using data to avoid legal risk.

4.2.1 Legal Principles have to be followed in data technology

There are five principles we have to insist when develop Data Technology.

No.1: Principle of Purposefulness. It demands data collection, storage and use shall conform to users' authorization of certain purposes. It is not allowed that service providers could get users general authorization and share the data with others to make new items or profits exchange. In essence, data exchanging is an indirect way to sale users' data which is beyond initial purpose of information collection, storage and use. So, it should not be encouraged.

No.2: Principle of legitimacy. It emphasizes data collectors, and savers have to conform to the law and regulations, including international practice. They cannot take illegal measures to get others' data by theft or cheat. Data savers need get relevant qualification or license. If they do not have conditions to store data, they should entrust other companies to do this work instead. The use of data also needs to be in accordance with the law and contract [14].

No.3: Principle of Proportion. This principle originates from administrative law. Its basic meaning is often explained as administrative organs should consider both the realization of administrative objectives and the protection of rights and interests of the counterpart. If there is a disadvantage effect on someone, it should be limited to the smallest scope, so that the two are in a moderate proportion. In fact, this principle is a further supplement to the principle of purposiveness. It demands service providers to choose the way of minimizing the infringement of autonomy of personal data.

No.4: Principle of Information Reduction and Saving. It derives from Article 3a of the Federal Personal Data Protection Act 2002 in Germany. The principle requires that information resources be saved as much as possible, the same information is collected only once, and establish information sharing mechanism to improve the rate of information utilization.

No.5: Principle of privacy protection. Citizens' personal data covers a large number of information, which is related to one's preference, home location, transportation ways and other comprehensive life information. So to speak, all the data involves the privacy of citizens' personal life. From the perspective of privacy protection, the data collected by service providers should be disclosure to others and be destroyed in time if it is not related to the purpose of the contract.

4.2.2 Service providers' duty

For service providers, they have to pay attention to both sides that raising compliance awareness and innovating the way of data use.

Firstly, they have to follow the international legislative trend and update the way of acquisition of authorization. There is no doubt that the owners of personal data have the right to possess, use, benefit and dispose of its data. So, service providers should get users' authorization when they collect, store and use these data. Service providers need to consider the reasonable standard terms about data collection, storage and use, and show it to users clearly. They have to avoid general authorization, and make sure the words used in announcement or contract are concise which is easy for every user to understand. Besides, the most important thing is that data collection should not be used as a condition for users to use related products or services. Providers need to consider the function of products or services completely to ensure users have the right of free choice.

Secondly, in a technological view, service providers need to innovate data encryption technology to ensure information security, such as preventing hacker intrusion.

Thirdly, following the international trend of compliance operation, enterprises need to innovate methods of data analysis to improve data value and quality, and save cost in a reasonable way.

Lastly, service providers have to concentrate on the limitation of data use in management, and establish tort compensation liability system. Once the users' privacy is improperly disclosed, service provider should bear the corresponding liability for compensation [15]. If they get information gains based on the data collected from users, service providers should also distribute the profits. The profits distribution plan can be diversified, and it is also the next question in my further study.

References

[1] Y.Q. Tu, J. Xu, L. Guo, Big Data Economy. Data Cost and Enterprise Boundary, Journal of Graduate School of Chinese Academy of Social Sciences, 2015, 5: 40-46.

[2] R. Lu. The Value and Exploitation of Data Goods – A Critical Analysis of Christian Fox's Theory of Digital Labor, Economic Review Journal, 2019, 5: 11-17.

[3] X.L. Yuan, E. Chang. Information Ethics of Native Digital Resources, Library Journal, 2011, 30(10): 2-6.

[4] L. Xu, C. Jiang, J. Wang, et al. Information Security in Big Data: Privacy and Data Mining, Ieee Access, 2014, 2: 1149-1176.

[5] C. PETRIE. The Proper Use of the Internet Digital Private Property, Ieee Internet Computing, 2016, 20: 92-94.

[6] J.H. Li. Data Leakage of 5000 Facebook Users, Chinese Platforms Should Take Warnings, CBN, 2018-10-15.

[7] X.Y. Zhang. Take Stock of the Leaks That Put Us At Risk, Big Data Time, 2018, 08: 64-73.

[8] K. Zhao. Data Leakage Reflects Privacy Protection Issues, Chinese Social Science Today, 2018-11-08.

[9] C. Wu. From Raw Materials to Assets: Challenge and Thinking of Data Assets, Bulletin of Chinese Academy of Sciences, 2018, 33(08): 791-795.

[10] Y.H. Shi. Legal regulation of Personal Data Transaction, Information Studies: Theory & Application, 2016, 39(5): 34-39.

[11] J. Chu. Who Will Protect Your Rights with Advent of Big Data Era, China WTO Tribune, 2018, 08: 45-47.

[12] J.C. Zhao. The Biggest Flaw in Facebook's History, World Affairs, 2018, 20: 77.

[13] X.C. Yang, K. Tu. Research on the Effect of Platform Support Quality on User Value Co-creation Citizenship Behavior under the Background of Sharing Economy, Business

Management Journal, 2018, 40(3): 128-144.

[14] W.G. Wu. Criticism on Protection of Personal Data Privacy Rights under Big Data Technology, Political Science and Law, 2016, 7: 116-32.

[15] X.Z. Wang. (2018) Reconstruction of Personal Information Legal Protection System in the Era of Big Data, Legal Forum, 2018, 33(6): 115-25.