Study on Data Security of the National Level under the Trend of Liberating Public Agencies Data

Biao WANG and Ding WANG

University of International Relation, Beijing, 100091, China

*Corresponding author

Keywords: Liberating Data, National Security, Information Security, Institutional Improvement.

Abstract. Liberating public agencies data has become an irresistible trend. China has also started the process of liberating public agencies data, meanwhile liberating public agencies data may have an impact on national security and society stability because the safeguards system of liberating public agencies data is yet imperfect and national data security cannot be guaranteed. Thus, this paper is trying to learn from foreign experience of data system construction and maintenance of information security, analyze risks of liberating public agencies data, and propose corresponding countermeasures that could be helpful to accurately reach the balance between open and safe in the process so as to enhance the advantage of the information security strategic in the future, prevent the liberating public agencies data from being analyzed and used by other countries to pose a threat to our national security.

Introduction

As an important strategic resource of national and social development, government data contains a huge potential value. Liberating public agencies data is an important way to protect the citizens’ right to learn the truth, enhance the quality of public service and upgrade the government image. However, it may threaten our national security if the data were analyzed and used by other countries. In a manner of speaking, liberating public agencies data may have an impact on national security and society stability. Therefore, faced with the opportunities and challenges brought about by liberating public agencies data, we should combine China’s national conditions with the security environment. Not only to grasp the balance between liberating public agencies data and national security accurately, but also to emphasis strategically on liberating public agencies data so that we can put forward suitable measures to maintenance of data security based on our national conditions, which are multi-level, multi-dimensional, multi-faceted in safeguarding national information security.

Current Situation of Liberating Public Agencies Data in China and US

Although the Chinese government has a certain effect on the construction of liberating data, but is still in its infancy. The security system needs improvement, the ability of security protection must be enhanced urgently and the maintenance of national data security cannot be guaranteed. By comparing the current situation of liberating public agencies data in China and US and learning from foreign system construction and practice, it has important implications for China’s liberating public agencies data.

Current Situation of Liberating Public Agencies Data in China

Up to the first half of 2014, Beijing, Shanghai, Wuhan and other places has launched data liberating, other local governments are actively doing related preparations. Liberating data has strategic meaning for “information benefit people” services, local government need to combine their own reality basis and treat it rationally.

Currently, 29 departments in Beijing have announced more than 400 data packets, covering tourism, education, transportation, health care and other categories. Shanghai has been clear about
190 public data resources to be key areas, a total of 209 data packets and more than 20 applications. They are including 11 fields, such as public safety, public service, transportation service, education and science, financial services, energy and environment, health and hygiene as well as cultural entertainment. Departments of Wuhan’s municipal government have been disclosed in various categories of public data more than 500, involving more than 30 areas. For instance, government, police, environmental protection, urban management, health, agriculture, transportation, home care, trade and logistics and so on. In addition, Qingdao, Guiyang, Guangzhou, Xiamen and other cities are also carrying out relevant preparations. With the transformation of the development of information services, liberating public agencies data will become a trend[1][2][3].

**Current Situation of Liberating Public Agencies Data in US**

As part of the Obama administration's Open Government Plan, Data.gov (http://www.data.gov/) was established in 2010 and opens "high-value" data sets to the public. The liberating public agencies data platform is now directly able to provide users with large amounts of raw government data, and the users are expected to find new value in the use of these data in order to deepen citizens’ cognition to government activities and more complex social affairs in the way that is impossible in the past[4][5][6].

At the same time the United States has a set of complex and evolving information policy (laws, regulations and memoranda), used to manage the life cycle of information, from production and dissemination of information to the information processing and archiving, which involves how to seek a balance among accessibility, privacy, security, digital asset management, archiving and preservation and other issues. Although this policy framework has also been adjusted many times, but is still lagged behind the advances in technology, this discrepancy makes current US policy framework unable to cope with the problems caused by big data: accessibility, quality and accuracy of data, big data age privacy, etc.[7][8]

**Main Risks after Liberating Public Agencies Data in China**

The development of big data is a double-edged sword, the gathering and analysis of big data increases the risk of privacy disclosure. Liberating public agencies data in big data environment makes information security and privacy protection face challenges. In the network environment, the public cannot live without social networking, e-mail and other tools, these personal data is often put together, these data relate to all kinds of personal information including customer information of corporate and other privacy information. Despite the wide usage of network, our country did not develop comprehensive laws to protect personal information, if these data are used improperly, there will be a risk of leakage of personal privacy[9].

In addition, with the orderly liberating public agencies data, collection and collation of information in big data technology contain precious value, which may become the target of many attacks by criminals. And the attack technology of criminals in big data also improves synchronously, which poses a great threat to liberating public agencies data. Specific risks are as follows.

**Risk of Data Quality**

The accumulated data resources of the government can be divided into two categories, one is collected via automation equipment such as sensors, and the other is furnished by artificial collection, these two types of data have different difficulty in the open process. The accuracy of data from automated collection channels can be ensure in an easier way and does not involve human factors, which makes it not prone to disputes when released, while artificial collection data is prone to error, and is extremely prone to problems when released, even causes data disputes[10].

When the mechanism of liberating public agencies data is relatively mature, after the formation of credibility, if some agencies use public data directly regardless of data disturbance, data quality and other factors, then it may cause erroneous interpretation of current areas of national economy,
policy, scientific research, leading to wrong decisions and affect the comprehensive national security[11].

Risk of Government Leaks

With social development, the public awareness of the importance of government data on their life is increasing, requiring expanding the scope of disclosure and deepening content continuously. Liberating public agencies data is no longer limited to official documents, but extends to all kinds of written materials or data generated in the daily work of government, and this caused considerable difficulties to the implement of secrecy censorship for administrative organ. Take Yancheng city as an example, in the year of 2014, all kinds of liberating public agencies data is counted up to more than 100,000. Especially with the website, micro blogging, Wechat and other new media has become the new channel for administrative organ to disclosure government data and a little carelessness will cause irreparable damage. As in 2014, Shanxi Province Environmental Monitoring Center carelessly leaked Dongfeng-41 ballistic missile-related information in daily work data on their official website.

Heavy Burden of Protecting Personal Privacy

Information security issue is a new challenge brought by liberating public agencies data, and it is even harder to protect privacy and the enterprise's business secret. For a long time, due to the lack of necessary legislation, civil servants paid great attention to government secret, but paid little attention to privacy protection, such imbalance must be reversed. In the long run, government data will be more open, the content that needs to be confidentiality will be less and less, but the responsibility to protect privacy and corporate trade secrets will become increasingly heavy, especially after data disclosed to the public, how to further protect the public's privacy and corporate trade secrets will become a new problem. A little mistake will cause great dissatisfaction from personal or business. This requires the government to develop clearer, more specific regulations to prevent infringing use of the data, and use more automatic technology to desensitize open data, shielding privacy leaks that may occur[12][13].

Risk of Society Instability

The general view is that data not being open enough is the contributing factors to social instability. For example, the PX refinery project in many cities including Xiamen, Dalian and Ningbo are protested by the public, which has a significant relationship with the opaque data. Social stability problem causing by improper disclosure of data should share the same concern. In 2011, Yuwen Road railway accidents, the improper words of the spokesman of Railways Ministry, making the investigation of the occurrence and the aftermath of the accident questioned by citizens. After 2012, Qidong incident, the reporter learned that the most participators were aware of the July 28 parade by the dissenting open letter and text message from local government. Hence, government data is open or not, how the data to be open are likely to affect social stability, which has a greater risk of unpredictable.

Suggestions on Data Security Protection at National Level under the Trend of Liberating Public Agencies Data

In the epoch of Mobile Internet, data has become a very important strategic resource for national and social development, and the security of liberating public agencies data has direct impact on national security and social stability. Therefore, the policy-making relating China’s data security strategy requires sufficient consideration on opportunities and challenges brought by liberating public agencies data, great emphasis on liberating public agencies data at a strategic level, and a multi-level, multi-dimensional and multi-aspect protection of national data security[14][15].
Integrate Present Laws and Regulations, and Conduct Specialized Legislation

In order to protect liberating public agencies data, at first, we should ensure that our top-level design to be clear, explicit and long term, and the execution of relative strategies and policies receive the guard of laws and regulations.

For example, in legislation and strategy planning, the US has characteristics and advantages in its multi-level systems, coordination between rules, and emphasis on integrity and good structure. The US has a perfect legal system, with its core laws and legislations including Freedom of Information Act, the First Amendment, Electronic Freedom of Information Act, government in the Sunshine Act, Privacy Act and the GPO Electronic Information Acquisition Promotion Law. This legal system has supporting laws, such as Paperwork Reduction Act and Electronic Government Act, and has issued relative files to guard liberating public agencies data[16][17].

However, the present laws and regulations of China has problems and weakness in its low legal hierarchy, outdated contents, delayed legislation, lack of systematic and a narrow range of liberating public agencies data. Therefore, China should promote national cooperation and ensure that law plays a leading role in liberating public agencies data, under the conditions of secured national data and citizen privacy. We should revise and modify present laws and regulations, such as Regulation of the People's Republic of China on the Disclosure of Government Information, National Secrecy Law and the Archives Law of the People's Republic of China, and timely make specialized and comprehensive laws on liberating public agencies data. What’s more, we should also clarify obligations and duties, and provide guarantee for the long-term strategy of liberating public agencies data[18].

Clarify the Border of Liberating Public Agencies Data

Liberating public agencies data is a historical trend. Faced with this, Chinese government has already broadened the range of public agencies data, and deepened the contents of public agencies data. However, the confidential examination system has thus faced great difficulties, because various kinds of paper files and data in the daily work of the government have to be revealed. Great loss will occur in case of inattention to the opening-up process, especially when New Media, including websites, micro blogging and Wechat, has gradually become important routes of liberating public agencies data. For example, information relating to the confidential Dongfeng-41 Ballistic Missile was leaked through the opening-up of daily work on official website by Environmental Monitoring Station of Shaanxi Province in 2014.

Liberating public agencies data involves many aspects. We will absolutely have a better and more comprehensive understanding in macroscopic view after collecting and analyzing all the data, however national secrets will be more probable to be leaked in this situation. For example, data about natural and social resources will perhaps reveal the significant circulation and exchange centers of China’s strategic resources, data about economics will be applied into the prediction of situations and breakthrough points of national economy, and the disclosure of agricultural data at a too excessive and detailed level will probably leak pivotal information of national grain reserves and economic situations[19][20][21].

Therefore, we should clarify the border of public agencies data, explicitly list the information that can be immune to disclosure, and decide through further examination whether a particular kind of information can be immune to disclosure. Highly sensitive information can be immune to disclosure without any appraisal, and other information will be checked according to whether the information itself is suitable for disclosure. If the disclosure of a certain kind of information has great benefits for the public, it should absolutely be disclosed. Besides, we should also establish a system of technical standards on security appraisal, which will analyze the security extent and flaw situations in data opening-up.

Coordination in liberating public agencies data at national level should also be strengthened, and a specialized administrative or coordinative institution should be established to take in charge of national opening-up of governmental data. At the initial stage of data opening-up, we should give most emphasis on the increasing scale of open data and the imitation of data opening-up method. A
list of public institution data resource, including government departments and public institutions, and a list of “Data Opening-up In Advance” should be built, and planned steps should be taken in the promotion of data opening-up.

**Ensure the Information Security of Liberating Public Agencies Data Network Platform**

Our government will publish list and guidebook of liberating public agencies data in recent years. This action will enable the access of the public and enterprises to use, reuse and develop these original data according to personal and commercial purposes, thus providing new and innovative services based on these data. A one-stop data opening-up network platform set up by the government will be required, and governmental data will be disclosed on it in a positive way[22][23].

Since the amount of data is gradually increasing, the application, reserve and conservation of big data faces various kinds of challenges, which include the inability of a single department to share and administrate data in a concentrative way. Usually, foreign countries share big data by adopting a method in which the government coordinates and administrates the disclosure and the third-party institutions undertake the particular disclosure work. For example, most data at Data.gov, the US liberating public agencies data network platform, are made publicly and reserved physically in third-party institutions.

A unified network platform increases possibility of attacks by network viruses and other malicious software. On this account, we should ensure, at infrastructure and network security technique level, that information can be immune from modification, and enhance the back-up of important data resources. If we adopt the method that information disclose by the third-party institution, before the disclosure of relative data, our government should appraise the sensitive information carefully, and urge the third-party institution to apply data perturbation technique and encryption technique. The sensitivity of the information disclosed should be dealt with on the base of the intactness of data statistics characteristic and scientific interpretation.

**Establish Third-party Data Risk Appraisal Institutions.**

When relative properties of data form different sources are combined. Recognition of these properties will become much more complicated, since relevance of these data will cause leak of pivotal information. This requires deliberate relevance appraisal and risk appraisal of these data before the disclosure. At present, the United States, the European Union, the United Kingdom, the Netherlands, Japan, Singapore, South Korea and Australia have established independent data risk appraisal and administration institutions in charge of sharing coordinative administration and risk appraisal[24].

These institutions give critical emphasis on horizontal and historical analysis, especially historical analysis. Since historical data have long-term and stable data accumulation, trend in relevant aspects will be revealed through analysis and census of historical liberating public agencies data at particular times. Historical liberating public agencies data is very complex in the process of data opening-up. Analysis of historical data will enable us to predict and test what is and what will be going on. As for the government, analysis of data with special characteristics will predict trends of future and then provide guidance and directions for our future job. Meanwhile, our government will also have a clear estimate of overseas hostile forces and make relevant preparations.

Therefore, these institutions are required for standardized risk appraisal of data ready for disclosure. Sustained attention on continuative data should be given, and once security risk of data is discovered, relative information should be reported to government institutions for further specialized processing.

**Encourage Relative Academic Researches and Engineering Practices**

The United States has nominated Data Promotion Specialists and set up Expert Communication Teams, whose main task is to conduct publicity and research of liberating public agencies data. Data Policy Office, affiliated to Department of Justice, often meets Chief Data Officers of different institutions and organized training and discussing conferences. In the US, government departments
and the Army decompose their own security risk by themselves, or decompose data with potential risk after desensitization, and then ask research institutions or social forces for help by means of network attack and defense competitions or other academic matched, thus acquiring contents in need from academic discussions. Project teams discovered from research that update frequency and sensitivity of data on Data.gov vary inversely, i.e. liberating data on the 911 incident in New York, where disclosure on time, location, geological conditions and weather became extremely cautious[25].

Compared with their foreign counterparts, relevant researches in our country are fewer. Researches in China concentrate on contents and application of open data, and pay less attention to profound probe into the guarantee mechanism of data opening-up. Especially in the trend of big data, analysis of liberating public agencies data will probably reveal national secrets and thus endanger national security. China may also adopt a similar strategy as the US, taking as the appraisal standard part of the open data which can represent our intention, or publish relative academic tasks, and then conduct independent operation and analysis after acquiring outcomes[26].

Therefore, our government institutions should positively interact with relative academic groups to learn and adopt innovative techniques, and gradually perfect data risk controlling strategy in accordance with modern technique development. We should relax approving constraints for relative academic associations and organizations, give them more legal permission, and render them specialized support to perform their intermediary role. Academic groups and activities are highly sensitive to inclined fund for their neutrality. To avoid these academic activities from turning into mere funding with the appearance of confined and superficial discussions, we suggest that our government conform to the conventions widely used in third-party funding, avoiding too many inclined constraints and receiving more advices. Meanwhile, to protect our own rights and benefits, we should apply third-party appraisals in examining security risk when adopt the go-and-see strategy.

Acknowledgments

This paper is supported by the Fundamental Research Funds for the Central Universities (No. 3262015T70).

References


