Science of Law Journal (2023)
Clausius Scientific Press, Canada

DOI: 10.23977/law.2023.020608
ISSN 2616-2296 Vol. 2 Num. 6

Definition of the Scope of Sensitive Personal Information

Yuhang Shen*

Law School, China Jiliang University, Xueyuan Street, Qiantang Town, Hangzhou City, Zhejiang, China
**Corresponding author*

Keywords: Scope of sensitive personal information, Identification, Contextual integrity theory, Personal information

Abstract: The definition of the scope of sensitive personal information is crucial for the scope of data collection, but in China, the scope of sensitive personal information does not be clearly defined. Although Article 28 of the Personal Information Protection Law defines the scope of sensitive personal information to a certain extent through the method of "listing and summarizing", it is difficult to apply it in practice. Plus, it is inappropriate to simply classify information into general personal information and sensitive personal information. To identify sensitive personal information properly, we must consider specific scenarios, adopt the scenario integration path which is based on contextual integrity theory to establish objective criteria for determining whether the information is sensitive.

1. Introduction

China's Personal Information Protection Law is a special law to protect citizens' personal information. The law distinguishes personal information into two types, namely general personal information and sensitive personal information, and sets up special sections to regulate the processing of sensitive personal information. Article 28 of the law stipulates that sensitive personal information refers to "personal information that, once leaked or illegally used, can easily lead to infringement of the personal dignity of natural persons or harm to personal and property safety." However, processing personal information is a relatively complex process. The sensitivity of personal information varies in different scenarios, so simply dividing information into general personal information and sensitive personal information cannot meet the needs of various situations. In addition, the term "easy" in the phrase "easily leading to infringement of the personal dignity of natural persons or endangerment of personal and property safety" is a highly subjective term, and judges may have different understandings of it, this can easily lead to different judgments in similar case. We should analyze the risk of personal information leakage in specific scenarios when identifying risks. This article compares the advantages and disadvantages of various theoretical viewpoints, and fully combines the current legislative situation at home and abroad to conduct an in-depth analysis of Article 28 of the Personal Information Law. It attempts to introduce scenario theory as a framework for defining sensitive personal information, and proposes specific standards for determining information sensitivity, making the determination of sensitive personal information more operational in practice.

2. Legal Definition of Sensitive Personal Information in China

2.1. Current Status of Legal Provisions

The expression of "sensitive personal information" was first adopted in China in the Personal Information Protection Law, while the Civil code adopted the expression of "private information", but the connotation and extension of them are different. [1]Abroad, the European Union and the United States have not yet formed a unified definition similar to "sensitive personal information", and this concept is used in this article to refer to it.

Article 28 of the Personal Information Protection Law mainly uses the method of "listing and summarizing" to define sensitive personal information, listing seven typical types of sensitive personal information and providing a generalized definition.

As for the definition method, from the perspective of Comparative law, the enumeration method is still the mainstream. Most relevant legislation in foreign countries adopts an enumeration method to divide the scope of sensitive personal information, and the enumeration method can clearly and clearly reveal a certain range of sensitive personal information.

Taking the European Union as an example, Article 6 of Convention No. 108 of 1981 defines the scope of "sensitive personal information" by way of enumeration. It stipulates that "personal data revealing race, political opinions, religion or other beliefs, as well as personal data relating to health, sexual life and criminal judgments, shall not be processed automatically." It can be inferred that the European Council believes that such data as race, political orientation, religion, criminal record and sexual orientation are "sensitive personal information". Similarly, Article 9 (1) of the General Data Protection Regulations (GDPR) of the European Union in 2018 stipulates: "It is not allowed to disclose information such as race or national origin, political concept, religion, philosophical belief, trade union membership, etc. in the process of personal data processing. The law prohibits the disposal of individual genome data information, biological characteristics data for the purpose of determining the identity of natural persons, and also strictly prohibits the analysis of individual Health data, sexual life, sexual orientation, etc." [2]

Taking the United States as an example, American law also lists many types of sensitive personal information. The definition of "sensitive personal information" in the United States is integrated into many separate laws. In order to meet the needs of social development, various industry norms have constantly spawned legislation on information protection, such as the Financial Privacy Act, the Consumer Privacy Bill of rights, and the Child Privacy Protection Act. In 2018, California, the United States, passed the Consumer Privacy Bill of rights (CCPA), in which article 4 stipulates that "personal information" refers to all data managed by legal subjects and generally unavailable to ordinary citizens through legal channels. The bill lists some common types of personal information, including business name, contact information, tax identification number, Biomarker, etc. On this basis, Virginia passed the Consumer Data Protection Act (CDPA) on February 19, 2021. The provisions in this bill are more refined based on the reference of the former CCPA. Compared to CCPA, CDPA separately defines sensitive data: "1. Personal data that displays race or ethnic origin, religious belief, mental or physical health diagnosis, sexual orientation or citizenship or immigration status; 2. Processing genetic or biological characteristic data for the sole identification of natural persons; 3. Personal data collected from a child; 4. Accurate geographical location data." Sensitive data is subject to separate supervision. It also stipulates the general methods for determining sensitive personal information, mainly the ability to disclose residents' names, gender, family status, physiological and psychological information in digital form, as well as other situations that can be used to identify or assist in determining the above information.[3] It can be seen that the relevant legislation in the United States is complex, and the scope of sensitive personal

information listed is also inconsistent among different norms. But this approach is more detailed and can meet the information processing needs of different industries and fields.

Although the methods listed are accurate, they are fixed judgment methods that lack flexibility and are difficult to cope with the changing information society. The international community has also foreseen the drawbacks of listing methods and taken certain measures. For example, the Security Council clarified the "sensitive" elements in the definition and protection of "special information" in the Interpretation of the 108 Act, emphasizing the direction of protecting personal interests, which is closer to the standard of general legislation. For example, although the "108 Draft" determines sensitive personal information by listing methods, it still emphasizes that protecting personal interests is the essential purpose of information legislation, and encourages contracting parties to add information categories according to their respective situations: the content of sensitive personal information listed in the draft is not all, and may vary due to differences in national legislation and social customs. Other forms of sensitive information will be added by contracting parties based on local conditions.

From this, it can be seen that the added general clauses in China can find their source in international legislation. The rationality of general clauses can be extracted from the content of the clauses, which can be divided into two defining criteria: firstly, whether information is prone to causing damage to the personal dignity of natural persons, and secondly, whether there is a risk of endangering the safety of personal life and property. Based on these two criteria, make a judgment on whether the information is sensitive, and then determine whether the information is general information or sensitive personal information based on its sensitivity. This regulation starts from the possibility of citizens' rights being vulnerable to harm, while also determining the connection between sensitive personal information and the personal dignity, personal safety, and property safety of natural persons. [4]China's summary clauses focus on risk prevention, which shares similarities with EU legislation represented by GDPR. Both pay attention to the risk of information leakage and adopt risk control as the legislative path, which is also the trend of information legislation today. However, although China has added the method of defining generalizations, it has not yet broken away from the rigid "binary model", and generalizations themselves have their own uncertainty and relativity.

2.2. The Drawbacks of the 'binary Model'

2.2.1. Lack of Operability

The so-called "binary model" refers to the classification of personal information into two types (general personal information and sensitive personal information) and two types of information are protected separately. Therefore, the method of identifying sensitive personal information is the core of this model. However, the articles of law that stipulates the method is a general clause, which can lead to a lack of operability in the identification of sensitive personal information and difficulties in applying it in practice. The clause is general for the uncertainty of relevant expressions, for example, the phrase 'easily causing harm to the personal dignity of natural persons and endangering personal and property safety' can be summarized as' easily causing risks', which is uncertain. People's perception of "easy" is subjective, and different people have different understandings. Giving judges complete judgment and excessive discretion may lead to unfair phenomena such as different judgments in the same case. This article believes that contextual integrity theory should be considered to analyze the risk of personal information leakage in specific scenarios, and distinguish whether personal information is sensitive based on the degree of risk. The risk of sensitive personal information is not a static "presence or absence" judgment, but a dynamic judgment, changing the judgment method as the scene changes. Therefore, the potential risks caused by the processing of

sensitive personal information can only be identified in specific scenarios.

2.2.2. The 'binary Mode' Is Too Absolute

The law sets the boundaries between sensitive personal information and general personal information, and sensitive personal information will be protected more strongly. However, it should be noted that information may be transformed between sensitive information and insensitive information, so an absolute protected mode is inappropriate. Firstly, 'sensitive' is a relative and uncertain term. In different specific scenarios, the standards for sensitive information and insensitive information are different. The sensitivity of the same information varies with the change of the environment, so it is necessary to evaluate and determine the possibilities faced by individuals in the dynamic environment of personal information processing. Secondly, with the development of science and technology, new types of information (such as location data of mobile phones and data generated by network activities) are constantly emerging. For example, with technological innovation, artificial intelligence may also collect personal information. [5]And this information requires continuous evaluation by legislators, so sensitive personal information and general personal information are not always clearly defined. [6]

3. Definition Method based on Contextual Integrity Theory

3.1. Contextual Integrity Theory and Its Advantages

Contextual integrity theory originates from American scholar Nissenbaum, which advocates that the determination of information properties should be analyzed in conjunction with specific scenarios and should not be limited by rigid differentiation thinking, as information sensitivity can change in different contexts. The professor's "contextual integrity theory" was originally proposed to address privacy protection issues. This theory combines privacy protection with specific scenarios, transforming how privacy is protected into information flow, how disclosure meets the rationality requirements of specific scenarios, and information flow criteria. [7] By applying this theory, a scenario integration path can be proposed to define sensitive personal information and place personal information under specific scenarios for comprehensive consideration.

Application of the theory has considerable advantages in defining sensitive personal information. Firstly, sensitive personal information is of great significant to the subject of personal information. It is easy to identify the corresponding subject through sensitive information. Once leaked or improperly used, it can affect the personal dignity of natural persons and even threaten the safety of human life and property, this is the root cause of emphasizing the protection of sensitive personal information. Secondly, as a new category of information, the scope of sensitive personal information is constantly evolving, and its legal enumeration is limited. After the formation of new sensitive personal information, its category must be determined according to the situation. The categories of sensitive personal information in the Personal Information Protection Law are specifically listed using the term "etc.", this indicates that the scope of sensitive personal information includes not only the seven types of information listed but also those not listed. As the types of sensitive personal information will continue to expand with the development of the economy and society, the scope is changing. Therefore, it is necessary to determine whether new types of information can be classified as sensitive personal information through individual cases. Thirdly, the scope may vary in different social environments. For example, disclosing someone else's home address and phone number on a small scale and directly processing this personal information on the internet are completely different situations, which have different impacts on determining whether sensitive personal information has been violated. Therefore, to determine sensitive personal information, contextual integrity theory must be adopted. [8]

At present, the academic community highly recognizes the use of scenario integration path to

determine information sensitivity. According to the scenario integration path, to figure out how the close relationship produced from information and personal dignity, personal safety, and wealth security based on specific scenarios, it is necessary to particularly examine the feasibility, reality, and potential degree of infringement.[9] However, some researchers advocate for a more fundamental judgment benchmark to better define sensitive information, and scenarios are often external factors that trigger information processing risks.[10] Based on this, some scholars have proposed a scenario extraction path, which strips out scenario factors and studies the fundamental features of sensitive personal information, and only examines whether the information content itself conforms to this feature. Some researchers suggest using scenario extraction and scenario integration pathways, while considering both scenario and content. [11]Through the analysis and comparison of the above two theories, this article believes that a simple scene integration path should still be used to apply scene theory, because there are too many factors to consider when using scene extraction and scene integration, which is too subjective and difficult to implement in practice. However, the impact of the scene cannot be completely ignored. Therefore, a compromise method should be adopted to apply scene integration path.

3.2. Elements of the Theory

According to the "contextual integrity theory", violating the flow of information norms in a certain scenario is the reason for people's concerns about personal information rights. Information norms are the sum total of norms in specific scenarios, which constitute information categories, involved subjects, and transmission principles. [12]Under this theory, the rationality judgment of information flow is influenced by specific scenarios, as the three elements of information regulation will play different roles in different scenarios. Professor Niesenbaum has determined five specific parameters to correctly express information processing norms: message subject, sender of message, receiver of message, type of message, and its dissemination principles. The first three can be uniformly summarized as the participants of information content, and the determination of the participants can determine the different rights of individuals with different identities to manage information; The type of information mainly refers to a classification of the nature of information content, including but not limited to email information content, medical information content, insurance information content, personal property information content, etc. Some information only exists in certain scenarios, while others (such as names) can be applied to multiple scenarios. The transmission principle stipulates the behavior of transmitting information from one subject to other subjects, as well as the realization of certain constraints on the free flow of information. Confidentiality is the most common feature, which restricts message recipients from sharing messages with others in the future; Secondly, there is reciprocity, which is often reflected in interpersonal relationships, allowing for two-way communication of information.

Scholars have identified the interests of data controllers, the purpose of data collection, and potential consequences as scenario elements. Article 51 of China's Personal Information Protection Law also lists several elements for information processors to consider in order to ensure the legitimate processing of personal information, such as processing methods, purposes, types of information, and potential risks.

3.3. The Central Principle of Scenario Theory

3.3.1. Principle of Appropriateness

The principle of appropriateness means that personal information must be shared differently between individuals, and it also specifies what personal information is allowed, expected, and required to be disclosed in different contexts. For example, when a doctor sees a patient, they will share information about their health status with the doctor; When chatting with banks or creditors,

people will talk about their personal financial situation; When chatting with teachers, people will talk about their grades. If you talk to a teacher about personal health information or financial status, or if you talk to a doctor about recent work situations, it will be considered a violation of the appropriateness criteria for information. The appropriateness of information ensures that people share information differently. German philosopher Ferdinand Schoeman once said, "Maintaining completely different interpersonal relationships with completely different people is crucial." [13] The above analysis can provide us with some inspiration - placing appropriate information in one scenario in another may cause infringement. [14]

3.3.2. Principle of Liquidity

Dr. Niesenbaum believes that the principle of circularity in scenario integrity theory includes freedom of choice, discretion, and confidentiality, allowing people to freely choose what information to transmit and keep others' information confidential. Other free circulation rules also involve necessity, authority and obligation. In actual medical situations, when patients share information about their bodies with physicians, it is not at the patient's personal discretion. This rule is similar to restrictive regulations, where physicians determine the information content that patients must inform them according to the requirements of the examination. Patients should also share information content with physicians that is conducive to rational examination and treatment. In the context of medical treatment, the transmission of information is not bidirectional, and medical personnel do not need to require patients to disclose their personal information, but they need to keep the patient's security information confidential.

4. Criteria for Judging Sensitive Personal Information under Scenario Theory

4.1. Using Flexible Methods to Evaluate Risks

Processing information involves various aspects such as obtaining and utilizing information. Therefore, static risk assessment is obviously not enough. It is necessary to consider various factors such as the context and purpose of information utilization, analyze the possibility of information subject being violated through dynamic data, and then evaluate the sensitivity of information content. In fact, the term "sensitive" is a vague concept, and fixed evaluation criteria will lead to difficulties in mastering too many information elements. So it is necessary to fully utilize the discretion of judges to make substantive interpretations in order to achieve fairness in individual cases, rather than relying solely on national agencies to solidify or qualitative evaluation standards through the promulgation of regulations or judicial interpretations. [15]

4.2. Blur the Boundaries between Sensitive Personal Information and General Information

General personal information can only be converted into sensitive signals in special situations. [16]For example, generally speaking, gender is not included in personal information for protection, but if there is a special situation, such as a boy or nine girls in this situation, then this gender information has an indirect identification function. Therefore, personal information cannot be simply divided into general personal information and sensitive personal information. Either inherently identifiable information or information processed by means of anonymization and pseudonymization can be reidentified under certain circumstances.

4.3. Establish a Risk Assessment Mechanism

The purpose of establishing a risk assessment mechanism is to provide differentiated protection of personal information based on different sensitivities. Sensitive personal information can be classified into different levels, and various types of sensitive personal information can be

quantitatively analyzed. The sensitivity level can be determined by examining various indicators (such as damage degree, identification difficulty, etc.) that cause risks. The sensitivity level can be set to three gradients: severe sensitivity, general sensitivity, and low sensitivity.

5. Conclusion

Personal data has been the focus of attention in recent years, and the traditional protected mode of personal information has been unable to respond effectively to such challenges. Therefore, people need to change the traditional protection thinking and integrate the scenario theory into the Information assurance system to reconstruct the information Protected mode.

Accurately defining sensitive personal information is to better ensure information security and promote the reasonable circulation of personal information. The legal definitions of the European Union and the United States each have their own characteristics. By comparing the legislation of the two countries and incorporating their reasonable parts, the definition method of sensitive personal information in China can be formed. Based on contextual integrity theory, specific suggestions for defining sensitive personal information can be proposed: using flexible methods to evacuate risks, blurring the boundaries between sensitive personal information and general personal information, and establishing a risk assessment mechanism.

References

[1] Shi Jiayou, The Private Law Dimension of Personal Information Protection. A Discussion on the Relationship between the Civil Code and the Personal Information Protection Law, Comparative Law Research, No. 5, 2021, p. 26.

[2] Tal Z. Zarsky. Incompatible. The GDPR in the Age of Big Data), 47 Seton Ha11L. Rev. 1013-1014 (2017).

[3] Wei Qianglin, Balance of dual attributes of personal Information - Innovative application based on" Scene Theory ", Kunming University Daily, No. 4, 2022.

[4] Zhang Yong, "The Integrated Protection of Sensitive Personal Information by Public and Private Law", Oriental Law, No. 1, 2022.

[5] Muige Fazlioglu, Beyond the "Nature"of Data. Obstacles to Protecting Sensitive Information in the European Union and the United States, 46 Fordham Urbar Law Journal 271, 289 (2019).

[6] Wang Yuan, "Definition of Sensitive Personal Information and judgment of its Elements - Centered on Article 28 of the Personal Information Protection Law", Global Law Review, No. 2, 2022.

[7] Helen Nissenbaum, Privacy as Contextual Integrity, 79 Washington Law Review 101, 137 (2004).

[8] Cheng Xiao. "On the Basic Principles of China's Personal Information Protection Law", Journal of the National Academy of Prosecutors, No. 5, 2021.

[9] Cheng Xiao. "On the Personal Information Processing Rules in China's Personal Information Protection Law", Tsinghua Law, No. 3, 2021.

[10] Shen Weixing, "On the Construction and Systematization of Personal Information Right", Comparative Law Research, No. 5, 2021

[11] Ning Yuan. "The Legal Basis and Scope Definition of Sensitive Personal Information", Comparative Law Research, No. 5, 2021

[12] Helen Nissenbaum, Wang Yuan, "What is a Scene? -- An Analysis of the concept of Scene in Privacy Scene Theory", Network Information Law Research, No. 1, 2021.

[13] Hong Lingxiao, "On Data Utilization and Protection of Personal Information in Epidemic Prevention and Control", Local Legislation Research, 4th issue, 2022.

[14] Wang Liming, Ding Xiaodong, "On the Highlights, Characteristics and Application of the Personal Information Protection Law", Jurist, 6, 2021.

[15] Zhou Jianyu and Zhou Chenyang, "Analysis on the Definition and protection of Sensitive Personal Information", Credit Investigation and Law, No. 11, 2021.

[16] Xie Lin, Wang Xuan, "Connotation and Extension of Sensitive Personal Information in China", Electronic Intellectual Property, No. 9, 2020.