

A Study on Privacy Issues and Solutions of Public Data in Education

Woochun Jun

Professor, Department of Computer Education, Seoul National University of Education, Korea
wocjun@snue.ac.kr

Abstract

With the development of information and communication technology, various data have appeared and are being distributed. The use of various data has contributed to the enrichment and convenience of our lives. Data in the public areas is also growing in volume and being actively used. Public data in the field of education are also used in various ways. As the distribution and use of public data has increased, advantages and disadvantages have started to emerge. Among the various disadvantages, the privacy problem is a representative one. In this study, we deal with the privacy issues of public data in education. First, we introduce the privacy issues of public data in the education field and suggest various solutions. The various solutions include the expansion of privacy education opportunities, the need for a new privacy protection model, the provision of a training opportunity for privacy protection for teachers and administrators, and the development of a real-time privacy infringement diagnosis tool.

Keywords: *Public Data, Education, Privacy, Information and Communication Technology.*

1. Introduction

Public data refers to materials or information that are created or managed by public institutions. All data or information in various forms (text, numbers, images, video, audio, etc.) generated by the agency's work is covered. In Korea, the law was enacted in 2013 to promote public data. In addition, the government's core policy is to mandate open and free access to public data for free use. Public data is available through public data portal sites[1].

Opening public data means opening and providing public data in various forms so that users can freely use public data held and managed by the government, local governments, and public institutions[2]. In Korea, the Korean government, local governments and public institutions open and provide public data in various forms so that users can freely use it. However, information that is subject to personal information protection, information including third party rights (copyrights, etc.), and information related to national security and order are excluded[2]. Anyone can use public data freely, and through public data portal, they provide various ways such as file data, open API, and visualization so that people can use public data easily and conveniently.

Open Data Portal that is a representative public data portal in Korea has the following public data use guidelines[1]. As a provider, they have the following policies.

- Principle of equality
- Principle of free use
- Guarantee of commercial use

On the other hand, Open Data Portal requires the following obligations for uses to comply.

- Guarantee of universal access rights: Anyone can use the right to use
- The principle of equality: The principle of equality in access and use
- Prohibition of use inhibition principle: Prohibition of restriction or prohibition on access and use
- Principle of guarantee of commercial use: Ensuring various use including the commercial use
- The obligation of good faith: Users are required to use public data for the good of the public interest, others' rights, laws, etc.

With the spread of digital data, the spread and utilization of public data is becoming active. In addition, various studies on the use of public data are being conducted[3-6]. Privacy is the one of issues that are becoming increasingly serious since privacy issues can cause a variety of problems, including personal information violations and hacking. Thus, in order to use public data properly, it is necessary to understand privacy issues and solutions dealing with privacy issues.

In this paper, we deal with the privacy issues of public data in education. The rest of the paper is organized as follows. First we introduce definition and characteristics of privacy. Next privacy issues in public data are presented. Also, various solutions dealing with privacy issues are presented. Finally, conclusions and further research issues are discussed.

2. Related Works

2.1 Definition of Privacy

According to [7], privacy is defined as follows. Privacy refers to personal privacy, personal affairs, or the right not to interfere with others. Privacy is sometimes translated into private life, but privacy is generally a larger category than private life.

On the other hand, the privacy of IT, that is, the computer and the communication, is defined as follows[8]: "The right of an individual or an organization to be controlled. Information about data, individuals, or organizations owned by an individual or organization should not be collected and used without permission, and personally identifiable information belonging to an organization may not be used by other individuals or organizations not related to personnel, employment, and work. "

2.2 Types of Privacy

According to [9], privacy can be classified as follows.

① Physical privacy

Physical privacy is the freedom from physical infringement of an individual, property, or space.

② Information privacy

Informational privacy refers to the expectation of privacy when personal information is collected, stored and shared in digital or other forms.

③ Organizational privacy

Organizational privacy refers to privacy in the sense that their activities or secrets are not leaked to others in government agencies, organizations and enterprises.

④ Communication privacy

Communication privacy refers to privacy in that most people think that activities in e-mail, telephone, and social networking are private and also want to be protected.

⑤ Privacy of behavior

Before the information age, it was difficult to record and share the daily life of people, where they go, what they buy, and whom they meet. However, in the modern knowledge and information society, many behaviors that people do not want may be represented and shared by digital information, and various purposes can be used unintentionally. For example, you can predict the purchase of goods by analyzing people's consumption patterns.

⑥ Privacy for individuals

Privacy for an individual means the right to maintain anonymity at a specific time or location from the perspective of an individual.

2.3 Viewpoints of Privacy

The viewpoint of privacy can be largely classified as follows in terms of the US and Europe[9].

1) US perspective

The point of privacy in the United States can be summarized as "right to be alone". There is no mention of privacy in the United States' highest constitutional constitution. However, Article 4 of the Constitution (search and arrest warrant) has the following provisions[10]: "The rights of the people who are guaranteed the safety of bodies, houses, documents and communications from unlawful search, arrest and seizure cannot be infringed. A warrant for arrest, search or seizure may be issued for a very good reason, supported by an oath or affirmation, and may not be issued unless the place to be searched, the person to be arrested or the goods to be seized is listed."

Privacy issues in the United States are mainly focused on confession and the body. In other words, a private house or body must be protected from infringing activities such as illegal search or seizure, and especially protected from public power such as government.

2) European perspective

Privacy in Europe can be summarized as 'honor and dignity'. The European Convention on Human Rights, ratified in 1953, provides the following rights to privacy: "Everyone has the right to respect for his privacy, his family life, his family, and his correspondence."

Also, according to [9], Europeans' common concept of privacy emphasizes preserving individual honor and dignity in public space. In other words, in the public space, individuals can be interpreted as having the right to be respected and the dignity of the individual. Unlike the United States, anyone has to have the right to self-control and to control when and how to disclose personal information.

3. Privacy Issues and Solutions

3.1 Privacy Issues

Like other public institutions, the privacy of educational institutions is serious[11]. The reason for this is as follows. First, the security of educational institutions is weak. This comes from budget problem. That is, limit of a budget, privacy protection issue is not taken into consideration. Second, there is no detailed policy for privacy protection in schools and educational institutions. There is a general guideline in personal information protection law, but there is a lack of specific guidance and manpower to do it. Third, there is a lack of publicity and education on privacy protection. Prior to strengthening the privacy policy in educational institutions, privacy protection education should be preceded.

Meanwhile, a recent study found that even students were not very interested in their privacy[12]. In other words, students do not know about privacy protection. In addition, the e-mail password is not changed periodically because it is annoying, and even students with personal information leakage have not reported or consulted because it is annoying, and most students do not know how to report or consult.

3.2 Solutions for Privacy Issues

In this paper, solutions for privacy issues in educational institutions and public data in education are proposed as follows.

1) Privacy Protection Education

In order to understand and prevent infringement of privacy and to deal properly with infringement incidents, it is first necessary to provide an understanding of privacy and educational opportunities in the elementary and secondary education curriculum. At present, elementary and secondary school public education is operating privacy education in the form of information ethics.

In addition to the urgent need for personal information protection education, it will be operated as a small area under public education. Considering equity in other subjects or other areas, it is expected that various difficulties will arise in restructuring the curriculum. To compensate for this, various special lectures, hands-on activities, and promotional videos can be used to recognize the importance of personal information protection and to prevent damages.

2) Development of New Privacy Protection Model

In the literature, the following protection scheme to protect privacy is proposed in [13]. The following table 1 shows privacy protection model that consists of three techniques.

Table 1. Privacy protection model

Technique	Meaning	Application Example
K-Anonymity	Review whether you can infer that you are a specific person, and make sure you are not more than a certain probability level	Make at least K records with the same value. In this case, the probability of identifying a particular individual is $1 / K$.
L-Variety	Even if a specific inference is not made, the technique of lowering the possibility of inferencing by raising the diversity of sensitive information	Each record has at least one diversity to prevent inference based on homogeneity or background knowledge.
	L-Diversity as well as techniques to lower the	The distribution of the

T-Proximity	distribution of sensitive information to further reduce the probability of inference	information of the entire data set and the distribution of the specific information are set to T or less to prevent inference
-------------	--	---

In addition to the above three privacy protection models, more privacy protection models should be developed in educational institutions. In other words, a model must be developed that compromises privacy and information disclosure. If too much emphasis is placed on privacy, information disclosure may be reduced, and privacy may be violated if information disclosure is given priority.

3) Provision of Training Opportunity for Teachers and Managers Dealing With Privacy

In order to fundamentally solve the educational institution's privacy problem, a variety of training opportunities must be offered to the teachers and the administrators of the educational institution. Regular and thorough training can prevent privacy infringement and can cope well in the event of an infringement incident. In addition, detailed and specific privacy guidelines should be established and promoted at the school and institutional level.

Recently measures or actions to ensure the safety of personal information has been emphasized for public institutions and public data[14]. The followings are main concerns.

- Establishment of internal management plan
- Access right management
- Access control
- Encryption of personal information
- Storage and inspection of access record
- Prevent malicious programs
- Safety measures for management terminal
- Physical safety measures
- Safeguards against disasters
- Destruction of personal information

The above guidelines are very practical so that they must be enforced and observed in educational institutions. Also, the guidelines must be included in training contents for teachers and administrators of privacy protection in schools.

4) Development of Real-time Privacy Infringement Diagnosis Tool

In accordance with the spread of universal design principles, the government started the Web Accessibility Research Institute (<http://www.wah.or.kr>) operated by the Korea Internet & Security Agency (KISA) to reflect the diverse information and communication technologies. Web accessibility guidelines and mobile accessibility guidelines have been developed by the institute. In addition, the Web Accessibility Research Institute distributes software tools that provide access quality certification and diagnose accessibility on websites.

However, periodic checks are not enough to respond quickly to flooding privacy violations. In other words, in order to respond quickly and completely to privacy infringement threats, tools for diagnosing and responding to infringement threats in real time are needed. Such a privacy infringement diagnosis tool should have the following functions.

- Real-time privacy diagnosis

- Real-time privacy response function
- Real-time privacy infringement reporting function
- Privacy violation prediction function

4. Conclusion

In this paper, we deal with privacy issues in the public data environment. In Section 1, the privacy issues of public data are introduced. In Section 2, privacy in general is presented. In specific, the definition, types and characteristics of privacy are introduced. In Section 3, the privacy issues of public data and educational institutions are introduced first. Also, various solutions are proposed for dealing with the privacy issues of educational institutions. As a solution, we introduce the expansion of privacy education opportunities, the need for a new privacy protection model, the provision of a training opportunity for privacy protection for teachers and administrators, and the development of a real-time privacy infringement diagnosis tool.

As public data becomes more available in the future, the number of relevant privacy infringements is expected to increase. Too much privacy can have a negative impact on activation, and the more negative the protection of privacy, the greater the risk of damage. In other words, privacy must be carefully protected like a double-edged knife and good use of two-sided policies to avoid adversely affecting the overall public data.

Acknowledgement

This work was supported by the 2018 Research Fund of Seoul National University of Education

References

- [1] Open Data Portal, Public Data, <http://www.data.go.kr>
- [2] Wikipedia, Opening Public Data, https://ko.wikipedia.org/wiki/%EA%B3%B5%EA%B3%B5_%EB%8D%B0%EC%9D%B4%ED%84%B0
- [3] E. Shin, "An Analysis on Trends and Tasks of Open Data Policy in the Digital Era", Journal of the Korean society for information management, Vol. 32, No. 3, pp. 49-68, September 2015. <http://dx.doi.org/10.3743/KOSIM.2015.32.3.049>
- [4] M. Min, "Modeling and Implementation of Public Open Data in NoSQL Database", International Journal of Internet, Broadcasting and Communication, Vol. 10, No. 3, pp. 51-58, August, 2018. <https://doi.org/10.7236/IJIBC.2018.10.3.51>
- [5] M. Min, "A Data Design for Increasing the Usability of Subway Public Data", International Journal of Internet, Broadcasting and Communication, Vol. 11, No. 4, pp. 18-25, November, 2019. <http://dx.doi.org/10.7236/IJIBC.2019.11.4.18>
- [6] S. Yoon and J. Hyun, "An Analysis of Open Data Policy in Korea: Focused on National Core Data in Open Data Portal", Korean Public Management Review, Vol. 33, No. 1, pp. 21-247, March 2019. https://www.kci.go.kr/kciportal/landing/article.kci?arti_id=ART002456940
- [7] Wikipedia, Privacy, <https://en.wikipedia.org/wiki/Privacy>,
- [8] Naver Knowledge Encyclopedia, Privacy, <https://terms.naver.com/search.nhn?query=privacy&searchType=&dicType=&subject=>
- [9] T. Craig T and M. Ludloff, Privacy and Big Data: The Players, Regulators, and Stakeholders, O'Reilly Media, 2011
- [10] Wikipedia, Constitution of the United States, https://en.wikipedia.org/wiki/Constitution_of_the_United_States
- [11] Security News, Security Problems in Educational Institutions, <https://www.boannews.com/media/view.asp?idx=66551&page=1&kind=4>
- [12] H. Kim, "A Study on the Recognition Level of Personal Information Protection for Students : Focusing on the G

University”, in Proc. of 2019 The Korean Entertainment Industry Association Conference, pp. 155-166, May, 2019.

- [13] Korea Internet & Security Agency, Guidelines for Identification of Personal Information Names, https://www.kisa.or.kr/public/laws/laws2_View.jsp?cPage=1&mode=view&p_No=282&b_No=282&d_No=3&ST=T&SV=
- [14] Personal Information Protection Portal, Guidelines for Safety of Personal Information, <https://www.privacy.go.kr/inf/gdl/selectBoardList.do>