

A Study to Improve the Public Data Management of the City of Busan

Hyung Rim Choi, Choul Hoon Kwak and MinJe Cho*

Department of Management Information Systems, Dong-A University, Republic of Korea;
mjcho@dau.ac.kr

Abstract

The government of the Republic of Korea has recently released a new administrative paradigm called “Government 3.0” to encourage the use of public data. Although this has caused a consistent increase in public data’s usage, users and managers still experience difficulties in the system because of management issues such as disordered naming, unstandardized classification, and the unprofessionalism and insufficient knowledge of on-site operators. To respond to these problems, this study focused on identifying the critical issues and points for improvement of public data management with case studies and focus group interviews with on-site persons in charge of public data belonging to Busan’s city government.

Keywords: Government Data, Information Management, Open Data, Public Data, Public Information

1. Introduction

1.1 Background and Significance

At present, many nations have been progressively encouraging the use of various kinds of data by opening their public data, and Korea is not an exception as its public data have been opened since 2001. At that time, the South Korean national government established the electronic government system to make national administrative works effective via online civil complaint services. In addition, the government started opening public data gradually on home pages of diverse authorities, including the national government itself and other public agencies. However, the government’s method of information provision at that time only observed one-direction communication. Because of this, users had to search and find materials by signing in on each home page to read the open information. The system became public-

oriented when Government 2.0 started to be observed in 2009—from one-direction communication to mutual communication. At this period, the public data portal (data.go.kr) was established with its goals of sharing, openness, participation, and interaction. To further develop the system on hand, the new government offered a new paradigm called Government 3.0 in 2013 in which the open public data are to be expanded and used more by private sectors to create new jobs and to turn national administration into a scientific system based on the data managed. With this, the opening and use of public data, which were considered as tools and methods to ensure the securement of Government 3.0, have been considered as the system’s core issues to be considered.

For the purposes above, the government established and proclaimed laws associated with opening information such as the Laws about Providing and Encouraging Use of the Public Data to designate obligation to open

* Author for correspondence

information on the maintenance and management of public data by public agencies. However, it is true that the open public data and the respective operation and management are in their very early stage. Thus, this emphasizes the importance of the elimination of public data management issues to provide well-organized services and increase the usage of public data.

This study points out the current problems of public data management and suggests the following actions to improve the system as a whole: (a) observance of standards and guidelines on public data collection, (b) setting of standards for naming and classifying public data, and (c) establishment of a public data exclusive division to train experts in the field. This study also claims that it is necessary to collect public data via Internet of Things (IOT) and Big Data, which are considered new information technologies, to improve the system's overall quality.

1.2 Methodologies

As the purpose of this study was to identify how public data operators recognize management issues and how to improve the system as a whole, it was expected that quality-based methodologies would be a more effective source of suggestions for possible plans than quantity-based ones in consideration of this study's attributes. Because of this, Focus Group Interviews (FGIs), which are commonly used for various strategic plans, were conducted in this study.

Questionnaires include the topic of issues and points for improvement of public data management, and case studies and FGIs were designed to draw out a more comprehensive view on the issues related to public data. The questionnaires were delivered prior to actual interviews to provide a better explanation of the study's targets. The interviewees were composed of 12 public data operators belonging to the local and public agencies of the Busan area and each group was made up of 6 members; Group A was composed of persons from the city hall and several Gu offices in charge of providing and registering public data, and Group B was composed of the persons in charge of collecting public data and are engaged in public agencies including the Busan Water Authority and the Busan Metropolitan City Institute of Health & Environment (IHE). The interviews were conducted 2 times, and each interview took about 2 hours to 2 hours and 30 minutes.

Table 1. Expert Groups

Group	ID	Agency	Group	ID	Agency
A (Persons in charge of provi- sion and registra- tion)	AA	Busan City Hall	B (Persons in charge of collec- tion)	BA	IHE, Busan
	AB	Busan City Hall		BB	IHE, Busan
	AC	Suyeong-gu Office		BC	IHE, Busan
	AD	Nam-gu Office		BD	IHE, Busan
	AE	Haeun- dae-gu Office		BE	IHE, Busan
	AF	Haeun- dae-gu Office		BF	Busan Water Au- thority

2. Definition of Public Data

Before the term "public data" was introduced, various terms including administrative information, information, knowledge information resources, and administrative information data sets were used in South Korean laws: "administrative information" was defined as "materials that were treated with optical or electronic methods and expressed by media such as signs, voices, sounds, or movies"¹; "information" was defined as "all kinds of materials or knowledge that were treated with optical or electronic methods and expressed by media such as signs, voices, sounds, or movies for specific purposes"²; "knowledge information resources" was defined as "materials in values of maintenance and that make use of nationally developed and digitalized materials or in need of digitalization on scholars, cultures, science and technology, and administrations"² and "administrative information data sets" was defined as "any letters, numbers, figures, images, and other information generated by administrative information systems."³

In July 2013, the Laws about Providing and Encouraging Use of the Public Data were established. Here, the laws defined public data as "materials or information, such as database and electronic files, treated with optical or electronic methods, which public agencies created or attained, and managed for purposes designated by laws." This study calls all the similar terms above "public data" according to the Laws about Providing and Encouraging Use of the Public Data.

This era allows society to live smart without any constraints in terms of time and space thanks to what the wireless Internet and smart devices have to offer⁴⁻⁵. This change on lifestyle consistently increases the need to open public data for the purposes of new businesses and job creation, aside from everyday convenience⁶. In addition, the national government strives to encourage its use of private sectors and to expand the range of the open public data to secure the people's right to know and the clear government operation and the increasing usage of public data by private sectors⁷.

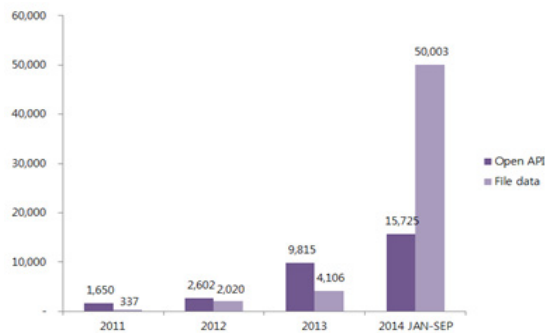


Figure 1. Number of Public Data Inspection Applications⁸.

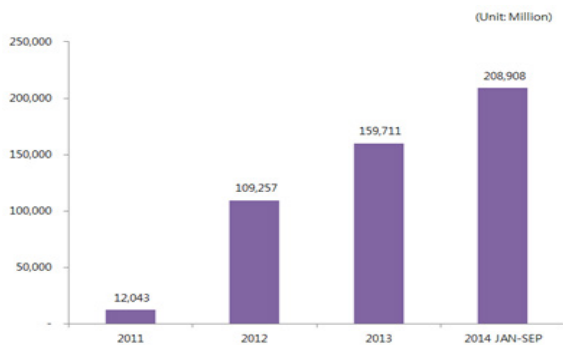


Figure 2. Number of Uses of Public Data⁸.

Figure 1 shows the number of public data applications by users for an inspection of specific Application Programming Interfaces (APIs) permitted, while Figure 2 shows how many of the permitted APIs were actually used. Seeing that the number of applications on file data rapidly increased in 2014 in Figure 1, it can be assumed that the public data demands of non-IT relations not only increased IT relations who formerly used the open APIs. With this, the national government should make public data more accessible and easy to use and more specific for the increasing variety of the growing number of users.

Today, public data are used in various fields, including transportation, tourism, health, weather, and environment, and they are also utilized as contents of the mobile applications and Web pages merged with information technologies such as open API and mashups (Web applications or pages)⁹⁻¹⁰. These services made public data more easily used and processed by private sectors and allowed the creation of new contents¹¹.

Most public data are now provided via the public data portal managed by the Ministry of Government Administration and Home Affairs, and it provides about 11,903 file data and 1,621 open APIs, and 251 data visualization services classified by 16 business reference models, including education, territory management, public administration, and finance. Files are in formats of xls, hwp, xlsx, html, and pdf, and are of large quantities, with most of the files made up of documents that provide information with a simple, reader-friendly approach¹².

The number of public data of Busan in the public data portal is about 1,008 currently in February 2015, and it is the second largest number in the South Korean cities and dos. After interviewing the operators, the number of public data provisions of the area has rapidly increased. This shows that it is necessary for the city of Busan to ensure the management of its public data. In addition, the standard deviation among the gus and guns of the city reduced from 24.31 in May 2014 to 14.11 as of 2015 showing that such local-level authorities now open public data more actively.

Table 2. Changes of the Public Data of the 16 Local-Level Authorities in Busan⁸

Division	Total	Average	Standard Deviation
'14.05	474	29.62	24.31
'14.08	591	36.93	25.78
'14.11	801	53.40	14.49
'15.02	887	55.43	14.11

3. Issues on the Public Data Management

3.1 Issues in Collection and Provision

Busan has no standards for public data collection and opening; because of this, each local office provides only the existing information on the administration status. It ensures only the provision and collection of data that

are relevant to the city and that are determined by their respective operators as necessary, despite their collection of new data.

“We used to provide only the existing materials rather than collect new public data with specific instruction or guidelines.” (AB)

“Current open data in the public data portal can also be found in our home page.” (AD)

“I don’t know what the citizens need so we investigate only what I feel is necessary. I would reflect the citizens’ requests if I can get them.” (AE)

Current public data provided are mostly made up of existing administration information, and here, there are no reflections of the needs of users¹²⁻¹³. Although users can freely request any public data that they need, it is hard to accomplish such because of issues including the data opening and collection period. In addition, there is almost no mutual communication between providers and users; users request for public data using only simple memos, causing confusion among the operators as they cannot understand the exact needs of the users.

“Although a user requests some public data, the data cannot always be provided because of diverse issues like open range differences among divisions and the large number of laws associated with this.” (AC)

“Current data on our Gu Office home page can also be found in the public data so we sometimes feel it is too ineffective. Actually, it makes us worried because we need to manage the same data twice if they are required to be modified.” (AD)

“We are adding public data via civil complaints; however, it is not easy to understand or find what users exactly need because they request only with memos that are too simple.” (BA, BD)

3.2 Standardization Issues

Busan provides public data according to the national government’s functional classification system on the public data portal. However, the system is only for the

views of the providers, making it hard to supply the exact materials needed by private sectors [9]. In addition, there is no guideline to categorize and name the data to be provided.

“We don’t have any guidelines for classification. It’s only up to me. If I get confused, I can refer to similar data opened by other areas.” (AF)

3.3 Unclear Management System

There are only a few public data standardization and management organs in local-level authorities; sometimes, there are no organs established to begin with. These organs’ on-site operators have insufficient knowledge on public data to help private sectors in the data’s usage, lowering the level of user support. Some operators are even in charge of other tasks outside public data management, making the provision of quality services to private sectors difficult.

“We had training at the beginning of public data provision. However, the training covered only what ‘public data’ means instead of how to manage such data so the training did not help us at all.” (AC)

“I work with other tasks aside from public data management so it is impossible to just focus on public data.” (AB, AD)

“I am only in charge of registration. In fact, I don’t have enough knowledge on public data, so I can’t help users that much.” (AA)

4. Proposals to Improve Public Data Management

4.1 How to Improve Public Data Collection and Provision

To improve quality of public data at a local level, it is required to offer clear standards and guidelines for collection and provision. If clear standards and guidelines are given, both the users and providers will work more effectively.

It is also necessary to improve passive attitudes of the providers. They need to be active and positive to provide

the public data instead of collecting and providing these only with requests from the users. To do so, they need to strive to estimate the user demands, collect and analyze much data as possible, and provide quality public data.

In addition, it is needed to encourage the use of public data with much more positive promotions and training provided to private sectors.

4.2 How to Improve Public Data Standardization

It is required to set standard guidelines to register public data. The public authorities register and open public data in the public data portal without specific guidelines, producing inconsistent classifications and names. By applying standard guidelines on the public data, the system can be made more convenient to users, and policy-based responses will be also possible by estimating how public data are used.

Currently, the national government works to standardize the public data expected to be frequently used. However, because of policy and priority differences among areas, standardization should be carried out in each city and do, and the national government should offer guidelines and supervise works at a local level.

4.3 Setting the Public Data Management System

For the public data of Busan, classification management is unclear and naming of items also varies. To solve these issues, it is required to establish expert organs to standardize and manage public data and to have a more organized training for operators about management and services on public data instead of the previous training, which only covers the definition of public data. Training the operators and providing incentives should be made prior to the establishment of expert organs for genuine improvement.

“I feel that public data expert divisions are necessary.” (AA)

“I also perform other tasks aside from public data management, so some compensation like incentives would really motivate me in my work.” (AC)

“Setting rules for public data management and training us in a more organized manner will be very helpful to our work.” (AE)

4.4 Efforts to Improve Quality

Because the national government and the private sectors are highly interested in public data, these are provided much more than the past years. However, the current public data are mostly composed of plain information that may be insufficient for actual use. To solve this issue, it is necessary to set expert divisions to research on public data and to support such researches in the field.

In addition, it is also helpful to improve the quality of public data by using Information and Communications Technology (ICT)¹⁴. By collecting an enormous volume of data from sensors of or IOT, expansion of the kinds of public data is made possible. The public data collected can be analyzed with the same methodology for Big Data¹². Using those advanced information technologies to collect and manage public data will improve the data's quality as well as prevent unnecessary expenditure of workforce and other resources.

“There are many cases in which public data were collected or managed through IT technologies in other countries. Seeing those cases, IT technologies are available to be applied to many fields; however, this is not easy to do in our (Korea) country because of the large number of laws associated with this.” (BC, BE)

“Introducing IT technologies like IoT or RFID may contribute to various public data collected.” (BB)

5. Conclusion

This study aims to identify issues in and suggest points for improvement of the public data management system established under the new government operation paradigm called “Government 3.0.” To do this, the analyzed the public data portal and drew out deeper issues from interviews with public data operators in Busan based on the result of the analysis. Using the data accumulated, this study suggested several ways to improve the system and make public data management more effective.

First, this study identified issues on collecting, processing, and providing data as well as how to improve the system as a whole. Currently, there are huge gaps among the numbers of the public data currently provided,

and it is difficult to reflect requests from private sectors for specific conditions. To fix these issues, it is necessary to set standards and guidelines to collect and provide the data and suggest which public data should be collected and provided at each local authority. In addition, the providers should be positive and active on the public data demanded by private sectors before they make a request.

Second, this study identified those that are considered issues on nonstandardized public data and the respective points for improvement. The most significant issue identified by the study was the fact that there is no guideline in the naming and classification of items of public data. With the current classification system, operators should determine the name and classification of the items by themselves. The national government also recognizes this issue and strives to standardize the public data; however, the national government should play the role of a supervisor for such works given the differences on local policies and conditions.

Third, this study pointed out the lack of expert divisions and human resources on public data. At present, most public data operators are also assigned with other tasks and have no professional knowledge on public data. To solve this issue, it is needed to establish expert divisions to standardize and manage public data and motivate the operators with forms of compensation like incentives along with the provision of well-organized training to equip them with the necessary knowledge.

Finally, to have better public data services, it is required to collect quality public data. Here, the introduction of IT technologies like IoT (Internet of Things) and Big Data will be significantly helpful.

This study identified issues on public data management and how to solve them by interviewing on-site operators to suggest on-site solutions. However, this study focused only on persons belonging to the city of Busan, serving as the study's limitation. It is therefore necessary to conduct further field-oriented studies on the collection and management of public data using the advanced IT technologies suggested by this study.

6. Acknowledgement

This work was supported by the Dong-A University research fund.

7. References

1. Korea Ministry of Government Legislation, Electronic Government Laws Available from: <http://www.law.go.kr>; 2014.
2. Korea Ministry of Government Legislation, Fundamental Laws of the National Informatization Available from: <http://www.law.go.kr>; 2014.
3. Korea Ministry of Government Legislation, Acts on the Public Record Management Laws Available from: <http://www.law.go.kr>; 2015.
4. Korea Internet & Security Agency, Guide for Public Information Literacy; 2011.
5. Kim Y M, Lee S M, Kim S H. A Study on plan for promoting innovation and utilization of public information sharing, KISDI; 2013.
6. Kim E D, Chin Y. Policy Measures for Enable the Open and Re-Use of Public Sector Information, Seoul: Y. Chin Parliamentary chamber, E.D. Kim Parliamentary chamber; 2012.
7. Kim H R, Choi J H. The State of the Art : Open Government Data-the Semantic Web Perspective, The e-business studies; 2012; 12(3): 371-93.
8. Korea Ministry of Government Administration and Home Affairs, Public Data Portal Available from: <http://www.data.go.kr>; 2015.
9. Kim D G, Joo W K, Kim E J, Lee Y H. A Case Study on Classification System Design for Public Sector Information Typology, Journal of Digital Convergence; 2014; 12(4): 51-68.
10. Heo P S, Park G M, Park W J, Cho G S, Ryu W. Domestic Market and Economic Impact of the Re-use of PSI(Public Sector Information), Electronics and Telecommunications Trends; 2013; 28(4): 118-31.
11. Kim Y M. A Study on plan for promoting innovation and utilization of information sharing, Journal of Digital Convergence; 2014; 12(4): 43-49.
12. Hong Y W. A study on the invigorating strategies for open government data, Journal of the Korean Data & Information Science Society; 2014; 25(4): 769-77.
13. Kim H R. Emerging Considerations for Accelerating Open Data, Architectural Research; 2013; 57(11): 46-49.
14. Lee K Y, Sim J C, Cho K S. Construction of knowledge base for the utilization of big data in public domain, Communications of the Korean Institute of Information Scientists and Engineer; 2012; 30(6): 40-46.