Knowledge Sharing Series

Data-Driven Smart Government





Knowledge Sharing Series

Data-Driven Smart Government

National Information Society Agency



APCICT ASIAN AND PACIFIC TRAINING CENTRE FOR INFORMATION AND COMMUNICATION TECHNOLOGY FOR DEVELOPMENT

Acknowledgements

This publication was prepared by the National Information Society Agency, Republic of Korea. Substantive comments were provided under the guidance of Kiyoung Ko, Director, APCICT. This publication is presented as a complementary resource material for the Academy training module on Realizing Data-Driven Governance. The views expressed herein are those of the author, and do not necessary reflect the views of the United Nations.

Cover credit: Yanni Ma

Cover photo: Peter Stuckings / Shutterstock.com

Disclaimers: This publication has been issued without formal editing, and the designations employed and material presented do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries. The opinions, figures and estimates set forth in this publication are responsibility of the author. Any errors are the responsibility of the author. Mention of firm names and commercial products does not imply the endorsement of the United Nations. The United Nations bears no responsibility for the availability or functioning of URLs. Correspondence concerning this report should be addressed to the email: apcict@un.org

Contact:

Asian and Pacific Training Centre for Information and Communication Technology for Development G-Tower, 5th Floor, 175 Art Center-daero Yeonsu-gu, Incheon City 406-840 Republic of Korea Email: <u>apcict@un.org</u>

Table of Contents

1	The	e 4th Industrial Revolution, Data, and Data Divide1		
	1.1	The 4th Industrial Revolution and Data	1	
	1.2	Data Divide	2	
2	Dat	ta-Driven Smart Government	5	
	2.1	Data and Data-Driven	5	
	2.2	The Concept of Smart Government	6	
	2.3	Conditions of a Successful Smart Government	7	
	2.4	Data-Driven Smart Government	8	
3	"Op	pen Data Release and Use" Framework	10	
	3.1	Conceptual Approach to Open Data	10	
	3.2	Data Institutional Framework	15	
	3.3	Open Data Control Framework	20	
	3.4	Open Data Technical Framework	23	
	3.5	Data Use Framework	25	
	3.6	Human Capability-Building for Data	31	
	3.7	Data Feedback Framework	37	
4	The	eory and Practice of Data-Driven Smart Government	45	
	4.1	Raise Awareness about Open Data Release among Public Institutions	45	
	4.2	Data Privacy Protection Issue	45	
	4.3	Open Data Use Fee	47	
5	Bib	oliography	50	

List of Case Studies

Case Study 1. Example of a Data-Driven Smart Government	9
Case Study 2. Digital Democracy	11
Case Study 3. Decide Madrid	13
Case Study 4. Uber	14
Case Study 5. Return to Fukushima Home support application	15
Case Study 6. Trends of Open Data Law/Policy around the World	16
Case Study 7. Open data implementation policy of Korea	17
Case Study 8. Open data implementation policy of the Philippines	19
Case Study 9. Open data implementation policy of Bangladesh	19
Case Study 10. Introduction to Overseas Policies and Guidelines Related to Data Quality	21
Case Study 11. Open Data Portal of Singapore (data.gov.sg)	24
Case Study 12. Health Care Service: Real-Time Disease Forecast Service (Korea)	25
Case Study 13. Open data used in culture and tourism	29
Case Study 14. Open data in environment and weather	30
Case Study 15. Open data used by the private sector for land management	31
Case Study 16. Case of enhancing human capabilities in UK ODI	33
Case Study 17. Enhancement of human capabilities in data in the public sector	34
Case Study 18. Support for expanding Open Data Use by ODX, Canada	35
Case Study 19. Start-up support centre for open data in Korea: "Open Square – D"	36
Case Study 20. Multi-stakeholder Forum related to open data	39
Case Study 21. Support for improving demand for data	43
Case Study 22. Legislations on the protection of personal information around the wor	ld in the
context of open data	46
Case Study 23. HIPPA privacy rules in the medical field of the USA	46
Case Study 24. Global trend in open data charge	47
Case Study 25. Case in Korea – Open Data Use Fee	48

List of Figures

Figure 1. The evolution of the industrial revolution	2
Figure 2. Data type by Joel Gurin	10
Figure 3. A Scene of Digital Democracy Service	12
Figure 4. The main page of Decide Madrid	13
Figure 5. Screen for Uber web service	14
Figure 6. Return to Fukushima Home service process	15
Figure 7. Open data strategy of Korea	18
Figure 8. Data quality control guidelines of the U.S	21
Figure 9. Before and after open data standardization	22
Figure 10. National Health Map service	26
Figure 11. Red Table service	29
Figure 12. Weather feed service by Kweather	30
Figure 13. Screenshot of Zillow.com	31
Figure 14. ODI Skill Framework	34
Figure 15. The ODX website (codx.ca)	36
Figure 16. Perspective image of Open Square-D	36
Figure 17. Startup collaboration project	37
Figure 18. Structure and roles of open data forum	40

Figure 19. Tenants of OpenGov Hub	43
Figure 20. OpenGov Hub office	44

List of Tables

Table 1. E-government vs. Smart Government	6
Table 2. Open data used for the service	26
Table 3. Cases of development of standard analysis model by sector in Korea	27
Table 4. Capabilities in data utilization	32
Table 5. Training programs for the public sector in Korea	34
Table 6. Type of participation by civil society	40
Table 7. Data-based monitoring in politics and administration	41
Table 8. Participation in politics and administration based on data or Internet technology	42

List of Boxes

Box 1.The 4th Industrial Revolution	1
Box 2. Conceptual Definition of Big Data	7
Box 3. Defining a Data Type	10
Box 4. Data-driven administration legislation around the world	25
Box 5. OGP (Open Government Partnership)	
Box 6. Governance	
Box 7. Civic hacking	42

This section aims to:

- Examine the cause and features of the 4th industrial revolution
- Explain the importance of data in the 4th industrial revolution context
- Discuss the definition and the cause of the 'data divide'.

1.1 The 4th Industrial Revolution and Data¹

While natural resources such as coal, oil, and electric power were the key resources of the previous phases of the industrial revolution; data, meanwhile, is the key resource of the current 4th industrial revolution.

Box 1.The 4th Industrial Revolution

The 4th Industrial Revolution first began with the "Industry 4.0²" implemented by Germany in 2011. Its innovation results from combining manufacturing industry with ICT. In 2016, Klaus Schwab, President of the Davos Forum, said: "Now a fourth industrial revolution is building on the third³, the digital revolution that has been occurring since the middle of the last century." Likewise, the Switzerland global banking group UBS (Union Bank Switzerland) defined the third industrial revolution as changes based on automation and connectivity, and the fourth industrial revolution as the phase of extreme automation and connectivity due to the advancement of computer and Internet technologies, artificial intelligence, and intelligent information technology.

Technological development was a key driver to each of the industrial revolutions. The first revolution was based on machine technology with a focus on steam engines, whereas the second industrial revolution was based on the development of electric power and energy technology, while the third revolution was focused on information and communication technology. These days, the development of intelligent information technology is a key driver to the 4th industrial revolution. "Intelligent information technology" refers to a technology that implements high-level information processing capability of human beings in machines by fusing artificial intelligence (AI) technology and data utilization (ICBM⁴) technology. Intelligent information technology is capable of handling complex and vast amounts of data, and it is rapidly changing society by combining diverse fields such as health care, administration and finance.

At the heart of the 4th industrial revolution is the goal of saving and controlling big data generated

¹ See Chapter I "The Data Revolution and Data Justice", section 1.1 and 1.2 of "Realizing Data-Driven Governance" for additional information on the recent data revolution.

² Industry 4.0: The German government conceived and announced the implementation method on April 2013 after 2 years of research by employing 150 experts in related industries, research centers, and academic circles. The government decided that the existing mass production system could not work anymore due to high price competition with countries with low labor cost such as China, and aimed at reducing cost at the mass production level to secure high profitability. As a result, the government implemented the "Industry 4.0," which includes the implementation of the CPS (Cyber Physical System) in which humans and machines partially cooperate in the production site and partially respond with automation.

³ 3rd industrial revolution: Following the argument of scholars represented by the economist Jeremy Rifkin, the 3rd industrial revolution occurred from the mid-20th century to the late 20st century and is also called the "information revolution."

⁴ ICBM: IoT, Cloud Computing, Big Data, Mobile

by the Internet of things (IoT), mobile technology with cloud computing, as well as having the capacity to analyze and utilize high-performance IT. In the era of the 4th industrial revolution, the information paradigm is shifting from hardware (H/W) to software (S/W) and from network to data. Data already plays a central role in creating value and momentum for economic growth as the oil of the 21st century. Therefore, competitiveness depends on how to create, save, control, and utilize more data.

Figure 1. The evolution of the industrial revolution



Source: Ministry of Science and ICT, Korea (2016)

Global leading companies that dominate the world market such as Google, Facebook, and YouTube base their business on data, and existing companies focusing on hardware, software, and network, such as IBM and Microsoft, are in transition to a data-based approach to their enterprise.

The government is also shifting its paradigm from an e-government based on systems and networks to a smart government based on data and artificial intelligence. Personalized specialized services as well as AI-based work and decision-making processes, which are the key to a smart government, are also changing towards using data as their key resource. We're living in an era where data competence defines the competitiveness of individuals, organizations, and governments.

1.2 Data Divide⁵

While the importance of data continues to grow with the 4th Industrial Revolution, readily available and high-quality data is still often neither a public or free good that can be easily obtained by anyone, anywhere and at any time. It is still more alike to a private good whose use and access differ depending on one's economic, political, and social competence.

As a result, data capacity has become a key determinant to wealth and competitiveness in the era of the 4th industrial revolution, essentially causing a divide at personal, organizational, and

⁵ See Chapter I "The Data Revolution and Data Justice", section 1.3 "Data Justice" of "Realizing Data-Driven Governance" for additional information on how data can affect differently and disproportionately certain groups in society.

international levels.

Data Divide Beyond Digital Divide

Generally, data divide refers to the difference in capabilities among individuals or groups that produce, collect, analyse, and use data.

Alleviating the digital divide⁶ was especially important in the early phases of the informatization of society, but in the era of the 4th industrial revolution, where each person's activities are converted to digital numbers and collected as data, 'Data Divide' reverts an even greater importance.⁷ This is because decision-making based on information as well as the quantity and quality of knowledge that can be practically used are affected by the gap between the use of data, data production, collection, analysis, and utilization caused by technology convergence and fusion in various areas. (Andrejevic, 2014).

Where does the Data Divide come from?

Absence of Quality Data

It is difficult for a government to root its decision on science and data evidence when government or state produced data is not properly shared, or even if shared, if the data is insufficient in terms of quality and quantity. Therefore, individuals and groups that fail to establish a framework for producing and sharing high-quality data may have difficulty with decisions-making and problems-solving.

From a gap in the development of data collection and processing technology

The data divide also relates to the technical aspect of continuously collecting, processing, and utilizing constantly generated data. Primary issues such as the establishment of a database are also important. However, developing a platform for data collection and processing is the most important issue in creating a link between data production and processing-consumption of data.

From a gap in the data utilization phase

Data becomes useless if not utilized by end users. The gap in the diversity and utilization level by end users is called "data utilization gap." The existence of an application that enables data utilization makes consumption possible, which is the last link in the data production-processingconsumption process. The data utilization gap, which is mostly related to the IT infrastructure level of a given country, can be eliminated by private enterprises. For example, if private enterprises use data, develop apps actively and provide free of charge services, the gap can be reduced.

From a gap in data literacy

'The data literacy gap' refers to the discrepancy in the active ability of producing information that fits context or promoting the distribution of information that is deemed valuable beyond simple

⁶ Traditional "digital divide" means the gap of accessibility to the digital device (including the Internet) and infrastructure or the gap recognized by social members regarding the individual's technological competence of utilizing digital (skill, the second digital divide) (Norris, 2001; Hargittai, 2001).

⁷ McCarthy(2016)

data consumption by users. The gap is closely related to digital literacy⁸ which is instrumental in understanding the overall data context and environment.

⁸ Digital literacy: Capability of someone for living, learning and working in a digital society. Source: JISC, the British institution dedicated to education informatization

This section aims to:

- Define data and data-driven
- Discuss the difference between a smart government and e-government
- Describe the requirements to become a smart government
- Define a data-driven smart government

2.1 Data and Data-Driven⁹

Changes in the Value of Data

The current rapid development of ICT has brought about high performance and increased integration of the computing environment, mobile computing, and hyper-connection. This led to a stratification in data quality and a sharp increase in quantity. While in the past, there was mainly 'passive data' that data holders or owners created while passively responding to the request of the data collecting actors, there is now a rise in 'active data,' which data holders volunteer to produce.

In addition, data is evolving into a new resource that substitutes to the role natural resources played in the industrial age and grows as a new policy tool in the public domain. Data here refers to big data that includes structured data on the properties of a person or an object, observation data from people's behaviors (e.g. clicks, Internet logs, reply, locations, etc.) and unstructured data. As big data is proven valuable for new businesses or projects both in the private sector and public policy, demand on data collection, analysis, and utilization has dramatically increased recently. Recognizing the importance of data, many governments have begun to share open data and use it as a policymaking tool.

Definition of "Data-Driven"¹⁰

The term "data-driven" in public policy means that **data becomes an important source of policy analysis and policy making** as a tool to solve issues and reach policy objectives.

Traditionally, for most governments, it was the institution with public officials in charge that inspired important government decision-making to address simple social problems. Often this led to governments that took unscientific decisions, and thus, failed to earn public consensus. However, social issues are becoming increasingly complex with problems such as youth unemployment, breakout of communicable diseases, economic slowdown, all of these with various stakeholders at play in a variety of government decision-making processes have made old approaches unviable to addressing social and policy issues.

The government should be able to use data as a policy tool and practice scientific public administration, and proactively respond to national agenda and social problems. Government

⁹ For various definitions of data including big data and small data, see Chapter 3, "A World Awash in Data" of "Realizing Data-Driven Governance".

¹⁰ See Chapter II "Data and Governance" of "Realizing Data-Driven Governance" for clear understanding of definition of "Data Driven".

policies should also be built upon data, with evidence-based policy-making pursued.

2.2 The Concept of Smart Government

Although the concept of smart government has yet to have an agreed definition among scholars and practitioners, it can generally be defined as a government that:

- Converges data from government services based on advanced IT technologies,
- Provides government services wanted by the citizens at anytime, anywhere, and on any media,
- Evolves by inviting citizens to participate in decision-making and by communicating with them.

The smart government is more alike to an intelligent government that transcend the current existing e-government.

The e-government makes use of ICT for efficiency improvement of public administration offices, which enhances productivity, transparency, and democratic processes of public administration, eventually contributing to better life of the people.

The existing e-government concentrates on business processing and the digitalization of public service online with a focus on computers, the Internet, and application processes. The smart government (also known as intelligent government) is different in that it pursues scientific and precise decision-making, automated and autonomous business processing, and customization and real-time provision of public services with the help of data, algorithms, and artificial intelligence robots. The table below provides a comparison between e-government and smart government.

Table 1. E-government vs. Smart Government	Table 1.	E-government vs	s. Smart Governmen
--	----------	-----------------	--------------------

	e-government	Smart government
Public	Citizens and public servants	Automatic recognition of rising
administration	propose agenda \rightarrow Improvement	agenda through digital brain $ ightarrow$ Self-
		Ideation \rightarrow Improvement
Policy-making	Government-led policy operation	People-led policy-making
Actual	Simple business processing	Enabled resolution of complex
administration		problems
Service	Quantitative and efficient service	Co-production of qualitative and
objective	provisioning	empathetic service
Service content	Customised to different life cycles	Close-to-life secretary service to
		different phase of life cycles
Delivery	Online, mobile	On-demand, on-offline, multichannel
channel		

Data, artificial intelligence as well as citizens and customers are at the core of the smart government.

Until now, the supplier (government) was providing services to the consumer (citizens, enterprises) unilaterally. The smart government, however, understands the consumers' needs and provides personalized services to everyone using data. Consumers can also search for the government services they need and leverage them independently using the data. Customers (private sector and individuals), not the government, are at the center of the service in the smart government, and

services are provided in a customized way that is intelligent, autonomous, and automated.¹¹ As a result, the very essence of the smart government is both data and customer; data when referring to technology and customer when referring to content of a smart government.

2.3 Conditions of a Successful Smart Government

Several conditions should be met to become a smart government, including the creation of a cloud computing environment, the utilization of big data and artificial intelligence, open data release and data security, and smart technology capabilities.

Creating a Cloud Computing Environment

Each government department should create a cloud-enabled¹² working environment.

The cloud offers an optimised environment for efficiently managing the vast amount of ICT resources in the public sector. The cloud provides a shared platform as an ICT service environment for the public sector, integrating both the hardware that is needed for application and transactions, and software altogether.

Moreover, IT resources can be available as needed and at the right time. This enables the actors to respond to ICT changes, cut costs, and enjoy flexibility, productivity, and efficiency.

In addition, creating a cloud work environment can help break down barriers between ministries and promote data sharing and cooperation. If cloud-related legislations governing the various government departments is different, guidelines for cloud development and operation as well as standardization should be facilitated.¹³

Utilization of Big Data and Artificial Intelligence

Efficient administration should be promoted using big data and artificial intelligence. If big data and artificial intelligence are used, civil service requests can be anticipated. This process is enabled by accumulated data so that customized services for each individual can be provided in real time. Expenses on the budget can be saved by automating routine services with an artificial intelligence robot, and excessive workload of civil servants can be alleviated. Public administration can develop into a preceptive and predictive system and be built into a "smart decision-making system" that can recommend optimal ideas, alternatives and policies for various areas like disasters, public safety and security, as well as healthcare.

Box 2. Conceptual Definition of Big Data

¹¹ Myeong Seunghwan (2012)

¹² Cloud is short version for cloud computing, which provides users over the Internet integrated shared computer processing resources and data into a system environment that is provided on demand to computers and other devices, or IT resources that exist in different physical locations, with virtualization technology.

¹³ See Chapter I, section 1.2 of "Realizing Data-Driven Governance" for more benefits government can enjoy from datafication.

(Technical definition) A new generation of technologies and architectures designed to extract value economically from very large volumes of a wide variety of data by enabling high-velocity capture, discovery, and/or analysis. (IDC, 2011)

(Size definition) Datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze. (McKinsey, 2011)

(Methodological definition) Initially, big data refers only to an enormous data set itself that amounts to tens-thousands of terabytes, but it gradually changed to a term that also encompasses related tools, platform, and analysis techniques. (Samsung Economic Research Institute, 2010)

Open Data and Security

As the power of a smart government comes from the distribution of freely shared data, it becomes imperative to develop data security measures and a clear privacy protection standard. That way, a favorable environment can be set up for quality data accumulation, distribution, and utilization.

Open data should be gradually increased along with strict security procedures. Data related to national defense, healthcare data with personal information, and trial data through which personal identity can be inferred should not be released as open data right away. A tiered-security regime, with various level of openness, such as classified-internal use only, is a good safeguard for these types of data. Therefore, strategies should be devised where data release and security are in balance in a way that public confidence in the government is not jeopardized.

Data Technology Capabilities

The successful implementation of smart government requires sufficient skills with data technology including but not limited to cloud, big data, artificial intelligence, open data and security.

More specifically to develop a pool of talent, the government need to develop training curricula to build data technology capabilities and to nurture promising people to help them apply this knowledge to their work. In addition, college students and other citizens should also be provided with opportunities to learn technical skills and offered more learning opportunities through field and practical experiences.

2.4 Data-Driven Smart Government

Definition and features of Data-driven Smart Government

Unlike e-government that are designed to provide better public services by making administration efficient through ICT, a data-driven smart government runs state affairs with people at the center relying on open data, and providing preemptive services by capturing people's needs in advance through analysis of both structured and/or unstructured data.

The data-driven smart government has the following characteristics:

• The government and the citizens participate in state affairs management together.

- Good quality data is provisioned in a single platform (of an app, a portal, a webpage) to proactively respond to citizens' needs.
- The government services are developed to the extent that meets citizens' satisfaction.
- The citizens themselves can customise and enjoy those services.

Below is an example of a Data-Driven Smart Government.

Case Study 1. Example of a Data-Driven Smart Government

("Before Service", the anticipated service) Social benefit services are provided in Push method after consideration of individual means for living, location, and circumstantial needs. The beneficiaries do not wait or seek for the service.

("Smart personal assistant" service) The government platform (app, portal, homepage, etc.) automatically recognizes the so-called "niche needs" and even emotions of citizens beyond the standard needs related to different phases of their lifecycle.

(Re-design the preemptive administrative process based on machine(self)-learning) Artificial intelligence acting as a personal assistant to the public servant in charge, diagnoses and improves the business process and identifies "hidden policy needs" by analyzing the existing policy data.

(Up close and personal digital service for the underprivileged) Providing an environment for the geographically/physically underprivileged with stable services (without having to travel or move).

(Implement cognitive "intelligent safety net") Preventing incidents preemptively and strengthening the disaster forecast capability by analyzing the crime rate by area, disaster (earthquake, heavy snow, etc.) occurrence statistics, and citizen's behavior pattern.

This section aims to:

- Understand the concept and characteristics of "open data"
- Discuss the impacts of open data use
- Learn about the framework for data opening and open data use.

3.1 Conceptual Approach to Open Data

According to the OECD, open data has two main aspects. First, it consists of "government data," which is created and/or authorized by the government. Second, it is truly "open data" in the sense that it is available to anyone and it can be used and reused for free, letting users the freedom to put together their own creation.

Box 3. Defining a Data Type

Open data can be used, re-used, and re-distributed free of charge. The "open" in "open data" refers to the fact that the data is machine-readable for re-use, offered free of charge, under an openlicense, and available to anyone. Types of open data include geodata, transport, environment, weather, statistics, financial, science, and culture. Moreover, the policy of releasing data held by the government is referred to as "open data policy." Open data is also technically referred to as open public data or open government data. In this context, Open data can be distinguished from big data and open government.

Joel Gurin¹⁵ described the relationship between big data, open government, and open data using the following Venn diagram.



Figure 2. Data type by Joel Gurin

¹⁴ This chapter mainly deals with the framework of open data that the governments is responsible for management. See Chapter VI. "Enabling Policies" of "Realizing Data-Driven Governance" to better understand the concept of open data, institutiionalisation of data governance and the benefits of opening data.

¹⁵ He was a senior advisor of the GovLan, the open data research center in New York University in the U.S., and a founder of the Open Data Enterprise Center.

As previously mentioned, the government and public institutions produce and manage an enormous amount of data and information. In the past, data was regarded as a simple input material, or raw material, in the information production process and thus, its independent value was hardly acknowledged. The unique value of the data, however, is now acknowledged as a resource needed for new industries and job creation. It is more than an input material, because of the influence of information and communication technologies such as artificial intelligence (AI), Internet of things (IoT), the cloud, and smart phones.

In particular, the importance of open data as a key resource is highlighted by the benefits it brings across a wide spectrum of sectors when produced continuously and stably.

If main actors such as governments, enterprises and citizens have access to and use open data, the following social and economic effects are to be expected.

Impact of Open Data Release and Utilization¹⁶

The economic impact of open data includes increased efficiency and effectiveness of public services, creation of new industries and jobs, development of new products and services, increased production index, and increased income level. Besides these economic effects, there are also other social and cultural values in open data such as boosted citizenship, information literacy, and social capital.¹⁷

Verhulst & Yong at GOVLAB¹⁸ (2016) studied the cases of 19 countries and presented the social and economic impact of open data release by grouping them into the following categories; improving government, empowering citizens, creating new economic opportunities, and helping to solve several serious public problems.

Improving government

Open data release enables governments around the world to curb corruption, make functions and responsibility more transparent (especially regarding budget), and improve government efficiency in terms of public service and resource distribution.¹⁹

Case Study 2. Digital Democracy

The policy-making information and documentations of the state assemblies of the United States are archived and offered in the form of video and scripts, enhancing legislative transparency. The archive serves as a web platform that encourages the participation of stakeholders, such as citizens, press, and interest groups, in policymaking.

All bills, public hearings, speakers, committee reviews, public depositions and testimonials are opened to the public in the form of video and texts without restriction. Anyone can search and view the information they need.

¹⁶ Stefaan Verhulst and Andrew Young, "Open Data Impact when demand and supply meet: Key findings of the Open Data Impact case studies", March 2016, pp14~15.

¹⁷ Gonzales-Zapata et al. (2015); Jetzek (2016)

¹⁸ GOVLAB: Open data research center in the New York University in the U.S.

¹⁹ Open Budget Transparency Portal in Brazil brings about responsibility and citizen supervisory for the national budgeting process. Central Registry in Slovakia is a world-famous open contract model. "Tax return data opening" submitted by a charity in Canada was the first step to the improvement of responsibility and transparency of the charity.

A keyword search engine is available so that various bills, hearings, speaking sessions in state assemblies, and meeting minutes are also easily found and downloadable.

All the content is downloadable and allowed to be stored in local computers. Social media links including Facebook, Twitter, and LinkedIn are also offered.

Introduced by the State of California in 2015, the service is being rapidly adopted by other states including New York, Florida, and Texas.



Figure 3. A Scene of Digital Democracy Service

Empowering citizens

With open data release, citizens can make smarter decisions, and exercise more control over their life through a new form of social mobilization. This is made possible with a new way of communicating and of accessing information.

Case Study 3. Decide Madrid

Decide Madrid is a participatory democracy platform to collect citizens' ideas and proposals on city ordinances and policies and ensure procedural transparency for policy development and legislation at the city council in Madrid.

It intends to leave no one behind in the policy-making process. It leverages collective intelligence from the citizens and enables the implementation of inclusive policies with a wider impact.

anguage: English 🗸	Transparency Open data Local Forums Blo
MADRID	Sign in Register
Debates Proposals Voting Processes Participatory budgeting Help	
Film library Citiz	zen of
Moratalaz Choose movies	
Moratalaz Choose movies Most active proposals	Most active debates
Moratalaz Choose movies Most active proposals Ofa-Sanchinaro heated pol Build the municipal sports center with a swimming pool that was promised many years ago. The neighborhood lacks sports facilit for all Parking Colonia de Lardín	Most active debates Abuses at the les Barajas Underground parking spaces for residents (PAR)- Control and sanction

Source: Decide Madrid

This platform comprised sections on Proposals, Participatory Budgeting, Consultation Process and Debate Process.

- **<u>Proposals</u>**: Individuals can propose new legislations and put them to vote in the platform.
- <u>Participatory Budgeting</u>: Citizens determine budget execution plans for the city's different projects.
- <u>Consultation Process</u>: Citizens make a proposal, put it on vote, and express opinions about policy-making of the city government.
- **Debate Process:** The city council hears and collects grassroot opinions before policy decision-making.

Initially it was run as a civic platform to deliver citizens' voices to the government, but it is now evolving into a platform for governance enhancement.

(Available data) budget data at project level, geological project mapping (addresses), and bills proposed to the council.

Creating new economic opportunities

Open data release creates new economic opportunities for citizens and the civil society and improves transparency in cities and provinces all over the world. Economic growth is promoted, new domains are created, and innovation is nurtured as more information is opened. In this process, open data release creates new jobs and new methods for citizens to prosper in the world. These effects can be sometimes found in the application and platform development process laced with open data.

Case Study 4. Uber

Ride with Uber
Sign up to drive
Erral Last Name
Phone
Create Password
Oty

Founded as UberCab in 2009 in California, USA, Uber was officially launched in June 2010. Uber provides a ride sharing service wherein it connects shared or Uber-owned vehicles with passengers who pay their ride through the application. Uber then earns profit from a commission on the ride fee. As of June 2014, it was in operation in over 100 cities worldwide. To deliver its services, Uber uses open data. Major data used includes information on commercial area, climate data, statistical data, and demographic / local crime data from local governments.

The corporate value of Uber in 2015 is estimated at USD 68 billion. Different versions of Uber have been launched, and it has recently expanded its business to food delivery with its service called UberEATS as well as freight delivery service dubbed Uber Freight. It is also playing a leading role in researching autonomous driving vehicles.

Source: Uber website

Helping solve several serious public problems

Recently, open data release (released public data) contributes to the resolution of serious social problems that cannot be otherwise handled easily. Citizens and policymakers can analyze social problems using new methods such as data-based evaluation and participation. This has had a significant impact on the public health crisis and other recent emergencies.²⁰

²⁰ Open data was helpful to communicate Ebola-preventive measures in Sierra Leone. The Singaporean government and

Case Study 5. Return to Fukushima Home support application

(Overview) An application service was launched for former Fukushima's residents who had lived in the disaster zone until 2016. The application provides information related to the return permit and the most-up-to date information regarding their residential areas.

The application feeds various types of information about the disaster zone and let the almost 100,000 evacuees check with through their smartphone the possibility of returning home.

In addition to utilizing open data, the evacuees and returnees of Fukushima Prefecture share local information directly through the app to establish a data distribution system that can be reused as open data.



Figure 6. Return to Fukushima Home service process

(Available data) Names of facilities, addresses, phone numbers, open hours of clinics and hospitals, capacity of welfare facilities, public housing information, goods and supplies for sale.

Resource: Resas Community

3.2 Data Institutional Framework

The establishment of a Law/System

With the value of public data becoming increasingly important, a growing international concern that governments have begun to recognize is the need to come up with a legal basis and systems to ensure the opening and stable use of open data. Accordingly, each government started to build a legal framework governing the following:

- The government's obligation to open data and protect people's right to open data;
- The implementation system;
- The open data standard, and
- The applicable procedure and methodology.

its people use the dengue fever cluster map to prevent the spread of the disease in Singapore. Open data was also used in the reconstruction efforts in Christchurch, New Zealand, after the earthquake.

This legal framework has to be put in place in order to meet the expectations of the people and the global society as well as to create an enabling environment for open data.

Case Study 6. Trends of Open Data Law/Policy around the World

The Republic of Korea

The Korean government has introduced and applied various policy means to promote open data. The Korean government enacted the "Act on Promotion of the Provision and Use of Public Data" (abbreviated name: Public Data Act)²¹ in October 2013 to prepare the legal basis for guaranteeing the people's right to open data, giving the obligation of open data provision to public organizations, and facilitating the effective use of open data by the private sector. This move was directed towards enabling open data to play crucial roles in promoting the open government and private sector.

The Korean government endeavoured to secure transparency in the government's decision-making process and earn popular confidence by enacting the "Official Information Disclosure Act" (abbreviated name: Information Disclosure Act)²² in 1996. Aside from this, the "Act on Promotion of the Provision and Use of Public Data" was separately enacted and enforced to promote the provision and utilization of open data by the private sector in order to demonstrate the government commitment to open data release.

European Union

The EU released a "Green Paper on Public Sector Information Society" dealing with "Public Sector Information: A Key Resource for Europe" in 1998, covering the concept and kind of public information, accessibility, and copyright conditions. The EU also enacted the "Directive on Re-use of Public Sector Information" in November 2003, laying legal foundation for the private service providers to use the public information commercially. This directive was designed to promote the creation of cross-border information services among EU members and stipulates the practical means and procedure for re-use of public information.

The European Commission also released the revised version of the directive, which puts emphasis on the economic effect of data re-use, and the Open Data Strategy in December 2011. As a result, the 27 member countries are required to provide all open data online and allow all citizens and service providers to use the data commercially. That is, commercial use of open data is now supported, and open data is serviced for free.

The United States

The U.S. government enacted the "Freedom of Information Act (FOIA)" in 1966 that allows anyone to access the information retained by the federal government. The copyright law enacted in 1976 stipulated that the intellectual property right of the federal government regarding the open data cannot be protected. In addition, the notification document of the Office of Management and Budget A-130 in 2001 (OMB Circular No. A-130) defined the use and re-distribution of the public information and guaranteed the commercial use of the open data.

²¹ Official name in English "Act on Promotion of the Provision and Use of Public Data"

²² Official name in English "Official Information Disclosure Act"

With the inauguration of the Obama Administration, the "principle of open data disclosure presumption" was announced in March 2009, wherein release of government data prevails when there is uncertainty over consent to release. The government also revealed the guideline to FOIA, a proof of more active government commitment. The public information provided under this principle is not only accessible by the public but can also be used by the private sector for commercial purposes, providing the private sector with capacity to open, provide and re-use public information. Based on those laws and policies, the Obama Administration opened the data.gov platform in May 2009 to provide open data to the public and induced the participation of managing government departments and affiliated organizations.

The United Kingdom

The U.K. government has taken a proactive stance on open data release and re-use. The "Freedom of Information Act" was enacted in November 2000 and the "Advisory Panel on Public Sector Information (APPSI)" began in April 2003, which facilitated the re-use of open data across the information industry. In June 2005, the "Re-use of Public Sector Information Regulation 2005" was enacted as the country internalized the EU's "Directive on Re-use of Public Sector Information" to the domestic law. In addition, the ODI (Open Data Institute) was set up to manage the open data policy throughout the country.

Indonesia

The Indonesian government enacted the "Public Information Disclosure Act No.14" in 2008 to open up the government and increase transparency. In 2011, the Indonesian government declared the disclosure of the government information by launching OGP (Open Government Partnership) to strengthen transparency, accountability, and citizen's rights²³. The "data.go.id" portal was opened in 2014. In 2015, the government adopted principles for transparency, accountability, information access, and anti-corruption using open data as a leverage.

Source: Open Data Strategy Council in Korea, http://www.odsc.go.kr/ Open data portal in the U.S., https://www.data.gov/ Open data portal in the U.K., https://data.gov.uk/ Open Government in Indonesia, OECD, 2016

Data Policy Implementation Structure

To effectively implement open data policies, the relationship among internal government actors should be well established and communication with external actors is essential.

First, a government-wide implementation system involving both the government and the private sector for overseeing and coordinating the open data policy at a higher level and promoting the actual release of public data, is needed. For government-wide implementation system to exercise and display a strong leadership in releasing open data, the highest government administration officer (president, prime minister, etc.) should be given the authority to set the open data policy as a national agenda and lead the policy implementation process.

Case Study 7. Open data implementation policy of Korea

²³ Indonesia is one of 8 countries that started the OGP (Brazil, Mexico, Norway, Philippines, South Africa, U.K., and U.S.)

The Korean government stipulated the implementation system as shown in Figure 7 and specified the roles and rights of each actor in the Public Data Act. The prime minister appointed by the President and a representative from the private sector are the co-chair of the Open Data Strategy Council. 'The Open Data Strategy Council' is composed of the overall control agency (Ministry of Public Administration and Security), specialized organization (Open Data Utilization Support Center), persons in charge in each department and local government (provision officer), and open data dispute mediation committee. In addition, the "Open Data Forum" is established for private/government cooperation, aside from the Open Data Strategy Council.



- **Open Data Strategy Council:** Deliberate and coordinate the open data policy and plan, check on progress and evaluate it.
- **Ministry of the Interior and Safety:** Announce open data release list, conduct actual operation evaluation, draw up a master plan, and lay the provision's foundation.
- Officer in Charge of Providing Open Data: Oversee the open data release and promote open data use.
- **Open Data Center:** Provide comprehensive and systematic support for open data release and use.
- **Open Data Mediation Committee:** Resolve disputes regarding suspension of data release or refusal to release open data.
- **Open Data Forum:** Proposing policy ideas for private and public cooperation and exploring projects on the capacity of data providers and users.

Source: (Korea) Open Data Strategy Council (<u>http://www.odsc.go.kr</u>)

Second, a **dedicated unit** is needed to plan for consistent implementation of open data policy and secure practical momentum for implementation at the government level. Additionally, it is also important to secure policy consistency by nominating people in charge of open data release and working-level staff in each organization. Moreover, in order to avoid undue pressure on the government to open data, a channel for civil and private sector complaints for raising objections can also be set in place. Other options that could also be considered include appointing and running specialized organizations that would support the establishment and foundation-laying for open data and facilitate open data use.

Case Study 8. Open data implementation policy of the Philippines

The system of implementing the open data policy in the Philippines is led by the presidential initiative to establish a system involving the top decision-making organization to the lowest level in the government.

The Philippine government opened the open data portal, "data.gov.ph", in 2014. By that time, however, the government organizations had already installed and operated their own open data portals and programs. The Philippine government modified the system in such ways that all open data including all public information should be consolidated into the central database, according to the Open Data Philippines Action Plan 2014-2016 and the presidential order of "Freedom of Information (FOI)" in 2016. Accordingly, the Department of Information and Communications Technology (DICT) changed the address of its portal homepage to "ogv.ph/data" (2017.6). The Open Data Philippines policy aims to integrate all released government data onto a single portal.

Open data-related policies in the Philippines are led by the Open Data Philippines Task Force, which involves the Office of the Presidential Spokesperson (OPS), Presidential Communications Development and Strategic Planning Office (PCDSPO), and Department of Budget and Management (DBM). The DICT, which runs and manages the portal, and the Advanced Science and Technology Institute (DOST-ASTI), are running the iGovPhil program in order to improve overall government processes for better services to the citizens and enterprises and promote public participation. The Open Data Philippines policy is one of the projects of this program.

Source: DICT (<u>http://www.dict.gov.ph/</u>) and DOST-ASTI (<u>http://www.asti.dost.gov.ph/</u>) homepage

Third, to achieve success, it is of utmost importance that the government open data policy and its implementation be carried by private demand. To this end, a **communication channel should be operated** enabling the exploration of data demand, identification of bottlenecks in data use, and resolve difficulties through active communication between the government and private sector. Data communication channels between the government and the private sector can take a variety of forms, including communities, forums and consultative meetings, where the government, the private sector and related experts can discuss on a regular basis.

Case Study 9. Open data implementation policy of Bangladesh

The Bangladesh open data policy is mainly promoted by the Executive Committee and Working Group. The Cabinet Division manages the overall open data initiative in the Executive Committee, and is composed of experts in the private sector, domain experts who can focus on the innovation and demand aspect, and a representative of the government agency that operates the data system and related regulatory system. The Working Group is composed of government officials and related experts in the private sector. It supports the open data initiative and develops catalogues and plans to recruit its members from the civil society, ICT industry, development partners, academic members, students, researchers, and related experts.

The Bangladesh government developed the open data portal (data.gov.bd) with the active participation of the Prime Minister's Office (PMO), Cabinet Division, Computer Committee (BCC), Bangladesh Bureau of Statistics (BBS), and Statistics and Information Department (SID). The innovation laboratory "Access to Information (A2I)" under the management of the PMO provides technical support for the portal, and the A2I, UNDESA²⁴, and BBS promoted open data release to support the SDG (Sustainable Development Goals) of the UNDP (United Nations Development Programme). The SID established OGD Strategy and the opening of the portal is the last phase of open data development.

Source: A2I homepage (http://a2i.pmo.gov.bd/)

3.3 Open Data Control Framework

Data Quality Control²⁵

Quality of data is key capital

Even if there is a large volume of data available, the inclusion of error data or the opening of data produced in different organizations in different ways will undermine the reliability of the data and pose a barrier to private sector usage. In particular, as a result of the convergence and fusion process, where various data get linked to each other, the distribution of some error data will affect the quality and corrupt the entire dataset.

The key success factor for data to realize its infinite possibilities is determined by "quality." Therefore, the government should systematically perform quality control activities throughout the data life cycle, from the data creation phase to utilization phase, as to provide high-quality data that can be reliably supplied and used, and should improve data utilization and release the data in a form that can be easily utilized by the private sector.

Understanding Data Quality Control

"Data quality" can be defined as a "level that can give useful value to the user by securing up-to-date, accurate, and interrelated data". That is, "data quality management" can be defined as a series of actions of setting the quality target, quality check, and improvement to ensure data

²⁴ UNDESA: United Nations Department of Economic and Social Affairs

²⁵ See Chapter VI. "Enabling Policies", section 4.2.10 of "Realizing Data-Driven Governance" for additional information on Data Quality Management.

quality, so that useful value can be provided to the user.

Quality Control Policy and Actual Cases

The data quality issues weaken the momentum of data policy, undermine confidence in the public sector, and impede on administrative tasks handling. Therefore, data must be effectively managed to ensure that the data held by the government and each agency is up-to-date, accurate, and inter-operable. In addition, each agency should establish a data quality control policy to provide quality data that meets the needs of the actual users and provides reliable high-quality data.

Case Study 10. Introduction to Overseas Policies and Guidelines Related to Data Quality

Data Quality Act (2001) and the government-wide OMB guidelines (2002) in the U.S.

- The federal government in the U.S. enacted the Data Quality Act (DQA) or Information Quality Act (IQA), and the OMB (Office of Management and Budget) issued the data quality control guide at the pan-government level.
- Agencies are obliged to prepare policies and instructions to improve data objectivity, utility, and integrity according to the OMB Guidelines, and to report the progress on a regular basis.
- As the OMB takes the initiative, government agencies have a strong obligation of compliance, but the guidelines do not cover the entire aspect of data quality in depth.



- The Ministry of Social Development in New Zealand established the quality index and quality improvement procedures to solve quality issues and secure data quality in the social welfare area.
- Although the guidelines are limited to the welfare sector, it is regarded as an advanced case in that the government department established and applied a concrete quality management system that is linked with the quality control target.
- The Ministry of Justice in New Zealand has also established and implemented the quality index and quality improvement procedures to secure data quality in the legal affairs area.

Canada – Data quality control guidelines of the Canadian Institute for Health Information (CIHI) (2001)

• The CIHI has established and implemented a data quality control framework that is composed of data quality activity cycle, data quality control tool, and data quality control documents to secure data quality in the health care area.

Korea – Open data management guideline (2014), Quality control manual (2012), Quality assessment (2014~)

- The MoIS defined the roles and the standards that have to be enforced at the governmentwide and the individual agency level for quality enhancement of open data held by each organization. The quality control manual details what and how the quality assurance officer in each organization can control data quality for each data lifecycle phase.
- Korea choose public data every year that commands high utility in the private sector and affects the quality out of the public data pool held by each office and has the selected public data diagnosed for its quality and improved in quality at the state level.
- In 2014, the "Open data maturity model" was developed as an objective quality control criteria and assessment system. After the pilot assessment with the model was conducted in 2015, 21 databases in 2016 and 42 databases in 2017 were taken for quality assessment with the model.

Source: NIA

Data Standardization

"Data Standardization" refers to unified and open standards. Services are implemented at the national level and data can be used easily by the private sector through standardization for the same kinds of data produced and managed at each province and organization.

For example, when each agency collects parking lot information as shown in figure 6, agency A provides 10 items (parking lot name, operating hours) in a PDF format; agency B provides 15 items (parking space names, opening time, close time) in an Excel format; and agency C collects the statistical data on the number of parking lot by zone within its area and provides it in a Microsoft Word format.

To use that information, the public needs to put up with the inconvenience of processing the data due to the difference in data presentation and definition. If the headings of parking spaces nationwide and data presentation are standardized, and data is open in specific formats such as open API, XML, JSON, or RDF, and visual analysis service like tables and charts are also provided, users may be able to save both time and energy on data reprocessing and make better use of the data.

Figure 9. Before and after open data standardization



Source: Open Data Utilization Support Center, NIA (2016)

These steps are necessary to push for data standardization:

- Development of the open data release standard

- Development of standard guideline for open data release
- Development of standard for building core dataset for open data release
- Pooling of standard terms and glossaries for open data

- Establishment of foundation for standardized open data release

- Organizing the open data release standard system
- Solidifying the support system for standardized open data release
- Developing standard management infrastructure for open data release

- Dissemination of open data release standards

- Project undertaking to support standardization of data set for open data release
- Enhancement of awareness of standardization and strengthened external cooperation

3.4 Open Data Technical Framework

Open data portal

The "Open data portal" is a representative example of open data release platform²⁶. The open data portal plays the role of a channel, through which the government data is opened to and shared with the private sector. Recently governments around the world undertake open data portal project by

²⁶ Generally, a platform provides an environment to run the application program in software development, whereas a framework is a frame that provides the source of the application running on the platform. The platform is an upper-level concept.

making use of an open source data platform (e.g., CKAN) in order to make open data available to the public.

The trend goes that the open data released by the central governments, the local governments, and public institutions is to be accessed through a single portal and is to be registered and uploaded to the portal in accordance with the **"Five Stars of Linked Data"**²⁷. Some countries present open data in API or visualized formats for data developers.

Case Study 11. Open Data Portal of Singapore (data.gov.sg)

The Government Technology Agency in Singapore opened the open data portal "data.gov.sg" in 2011 to enhance economic value and encourage research excellence and popular participation. The portal provides a one-stop access to open data. There are 9 categories in the portal (economy, education, environment, finance, health care, infrastructure, society, technology, and transport) and 70 public institutions released their data (as of 2017). As of 2017, about 900 refined quality datasets have been released and about 30,000 datasets were downloaded in 2016. CSV was the most popular format for provisioning.

In addition, Singapore opened a data portal blog (blog.data.gov.sg) in July 2015. The blog explains the trend and data usage to improve everyday life of the citizens through visualized data. Various government departments and agencies cooperate to present the data related articles and visualized data for a better understanding of social trends and government policy.

The portal for data developers "developers.data.gov.sg" was also opened in April 2016. Developers and analysts can access the open data API of government agencies in real time by visiting the portal only, without signing up with each individual agency. This developer portal also provides 13 frequently used APIs.

Open data is generally released in the forms of open API, files, or LOD (linked open data). Open data has a potential of use for the private sector when it is machine-readable²⁸, and carries economic values when it is released in source data format.

First, "**Open API (Open Application Program Interface)**" refers to an open interface, which enables data users to write and develop applications and services in addition to unilaterally receiving data through the web. Open APIs have the advantage of receiving data with short renewal cycles in real time and being able to be used directly for private app/web services.

Second, "Files" enables the user to download desired data in bulk²⁹ at once. To use this data

²⁷ Five Stars of LOD: Timothy Berners-Lee, a founder of the web and linked data, proposed the multi-linear phase deployment plan of the open data using stars. That is, one star is given if data is released on the web with an open license regardless of format, two stars are given if data is released as structured data (e.g., Excel), three stars are given if released in open formation, 4 stars are given if the object is presented using the URI, and five stars are given if data is provided in a linked state with other data (LOD) to provide the data context and background.

²⁸ Machine Readable Data refers to a state of data, where respective data content or internal structure can be checked into with software or, where modification, transformation, and extraction are enabled. In other worlds, machine-readable data is the easy-to-process data on a computer (ex. RDF, XML, JSON)

²⁹ When a user can download the whole data to his or her system with ease and efficiency, data is released in bulk. On the other hand, data is not in bulk when data access is limited. The case in point is that you as a user have to make thousands or millions queries to get the whole dataset, because one access is allowed only to a limited number of objects. Releasing data in bulk is a critical requirement for open data release. (Source: Open Knowledge International "Open Data")

smoothly, machine readability and up-to-dateness should be guaranteed.

Third, the "LOD (Linked of Data)" type of release connects all data and databases on the web so that users can find the intended information accurately. It is a data format that is open, linked, and shared on the web by building linked data using web-standard technologies such as HTTP, RDF, and Internet Identifier (URI), and linking data to each other through URI, despite being from different sources.

Open data delivery should be provided using various methods rather than just one to foster greater utilization. Consider, for example, a start-up that plans to start a business using car registration information. The start-up should first download all car registration information in bulk to be able to start its business. They should receive information using the open API when a new car is registered, or when registration information is changed. If owner information, car inspection information, and law violation information related to the car registration information can be searched using LOD, the value of the service will be enhanced further.

3.5 Data Use Framework

Data Utilization: Public Sector

A data-based smart government should pursue objective evidence, or data-based scientific administration, rather than experience and intuition-based administration. "Data-driven administration" means that public institutions collect, process, and analyse large amounts of data from the private and the public sector (structured and unstructured). These are then reflected in administrative works such as policy establishment, execution, and appraisal. For this, each government conducts various research and makes efforts to prepare laws and policies as well as infrastructures for data-driven administration.

Box 4. Data-driven administration legislation around the world

The U.S. enacted the Act of Commission on Evidence-Based Policymaking in March 2016 and through its one-year of activities, published a final report that includes the Federal government's proposals for capacity-building and the evidence system. The U.S. government plans to make a mid- to long-term investment in infrastructure development and lay the foundation to implement data-driven administration based on the report and establish the National Secure Data Service (NSDS).

The United Kingdom has established the Administrative Data Research Center (ADRC) to support government-generated administrative data to be used properly in research for policy purposes to develop evidence-based policies. The ADRC is carrying out technical support for the use of data, such as addressing issues related to data opening, functions for reviewing research suitability, and non-identification.

The Korean government enacted the "Act of promotion of data-driven administration", based on which it is trying to establish the "Open big data centre" in 2017, in order to link data among agencies and implement scientific administration through data merging and analysis.

Case Study 12. Health Care Service: Real-Time Disease Forecast Service (Korea)

Handbook")

The service is provided in the Republic of Korea with the aim of changing the direction of the nation's health care policy to be more focus on prevention rather than treatment, reducing medical costs and improving public health in face of an aging society. This service model is designed to predict citizens' future health ailments in real time by pooling medical expense invoice data, real time Drug Utilization Review (DUR) information, weather forecast information of the Meteorological Office, air pollution observation data of the Korea Environment Corporation, and population information data from the National Statistical Office.

-		
Agency in Possession	Target Information	Major Properties
Health Insurance Review	General information in	- Basic patient information in the billing
& Assessment Service	the billing statement	statement
	Treatment details in the	- Treatment details information like
	billing statement	medicine
	Illness details in the billing	- Illness details information
	statement	
	Extramural prescription	- Extramural prescription details
	details in the billing	information
	statement	
	Nursing home information	- Nursing home information
	in the billing statement	
Meteorological Office	Meteorological	- Meteorological observations
	observations information	information such as climate,
		precipitation, humidity, etc.
Korea Environment	Air pollution observation	- Air pollution observation information
Corporation	information	such as SO2, CO, and PM10
National Statistical Office	Demographic information	- Mid-year population information by
		city, district, and gu (neighborhood),
		based on resident registration

Table 2. Open data used for the service

Service plan and achievement

Providing the GIS-based National Health Map service on health care and medical big data system of the Health Insurance Review & Assessment Service (http://opendata.hira.or.kr/).



In the Republic of Korea, in addition to preparing laws and policies as well as infrastructures for datadriven administration, a standard analysis model was established for each data sector in order to prevent overlapping investments in the budget and to make effective use of the data.

The standard analysis model is developed in a versatile form so that different agencies and organizations can use the model to enable standardization of data, analysis, and operational procedure of open data analysis models by domain. Thanks to this model, Korea was able to reduce duplicated investment in similar data analysis projects and cut down on project time and costs, resulting in budget saving. In addition, public officials in charge of data analysis in the field experienced greater convenience and efficiency for their work.

Sector	Data analysis method	Visualized output
Transportation	- Analysis of traffic blind spots, bus dispatchment intervals, positions of transit stations, and priority level for the introduction of low-floor buses.	Development of standard analysis model for traffic jam, transit zones, and accessibility to transportation by the disabled
Tourism	- Analysis of floating population of nationals and foreigners, commercial viability and revenues, and festivities.	Improvement of accessibility to tourist destinations after analyzing tourist routes, commercial outlet density, and

Table 3. Cases of development of standard analysis model by sector in Korea

		parking spaces.
ссту	- Analysis of CCTV vulnerability index, crime vulnerability index, crime forecast index, and priority CCTV installation zones.	Adding more CCTVs in CCTV-vulnerable areas, and beefing up policing activities in
		crime-prone areas
Apartment	 Density analysis, correlation analysis, trending, comparative analysis. Analysis of irregularities related to administration fee, bidding, and 	Exposing administration fee fraud (Shared
	contracting and subcontracting.	facilities)

Data Utilization: Private Sector

Open data makes the government smarter, but also helps job creation, production, and profitmaking in the private sector, contributing to national economic growth.

Open data usage in the private sector increases every year and outstanding achievements have emerged in various fields. In particular, five major areas demonstrated many success cases of open data utilization, including in culture and tourism, transport and logistics, environment and weather, health and medical treatment, and land management. In this context, the government should create the right environment to encourage the private sector to use open data more actively and reflect about what policies should be implemented to facilitate while considering the following:

- Data quality control and standardization
- Identification of data with high demand from the private sector
- Discovery, production, and release of the previously identified high demand data
- **Specialized consulting services** covering legal, financial, and data analytics aspects, which may hinder businesses in the early stages.
- Building a sense of community and cultural ambience through events such as Hackathons and competitions, where privately developed services using open data are recognized and awarded.

Case Study 13. Open data used in culture and tourism



[Caption] Red Table service screenshots (Menu, how to order, and payment through Alipay screens serviced in Chinese)

The "Red Table" service analyses big data in restaurant reviews written by foreign tourists to Korea and ranks the restaurants. Content is serviced in multiple languages such as Chinese, Japanese, and English. The core service of Red Table involves analysing factors created by users such as reviews, visit frequency, and major hangout areas and providing information after reliability analysis of the reviews.

Red Table used standardized **open data** by the Korea Tourism Organization and Korean Food Promotion Institute for standard menu names and menu explanations. There are approximately 300 menus provided by the websites of the Korea Tourism Organization and Korean Food Foundation, all of which are classified in a systematic manner, focusing on Korean foods.

Red Table received the Grand Prize (Presidential Award) in "The 2nd Start-up Competition Using Open data" hosted by the Republic of Korea. In 4 years since its foundation, the company succeeded in attracting investments of KRW 1.1 billion in 2015. In 2016, it signed a business contract for food tourism goods with Dianping, one of the largest companies in China. It also signed a contract with KLOOK, the largest tourism platform in Southeast Asia in 2017.



Open data (weather information and air pollution information), Kweather news, and service output on screen

In keeping with the explosive increase of smart phones in Korea, Kweather jumped into the development of a weather information feed service quickly and accurately with an application for smart phone that came out in 2010. Kweather began its App service one year before the Korea Meteorological Administration opened its weather feed service in March 2011.

Using the **weather information of the Korea Meteorological Administration,** which is open data, Kweather's weather App offers various content including the current weather, weekly forecast, special forecast, radar images, and satellite images. Its service, which was the first private weather forecast in Korea, has been known for accuracy, and is featured with weather broadcasting and news for fresh information feed to the users. On top of local weather information, it provides themed and customised weather information for different situations such as airport, mountain, golf course, and baseball stadium. At present, Kweather generates over KRW 10 billion (USD 1 billion) in annual sales.

Case Study 15. Open data used by the private sector for land management



Zillow provides all kinds of real estate information required by homeowners, buyers, sellers, renters, brokers, lenders, landowners, and valuators when they simply enter the postal number of a residence. Following its launch in August 2006, Zillow help funnel more than USD 3 billion worth of capital into the real estate market; it is also posting information on over 110 million houses in the USA. It has essentially opened up a large-scale real estate market online.

Zillow utilizes open data such as local GIS, demographic information, school information, and employment information. Zillow has established itself as a major online real estate platform in the States as evidenced by its 2016 sales of USD 850 million.

3.6 Human Capability-Building for Data

Classification of data capability and importance of human capability

In 2013, the United Kingdom announced 'A Strategy for UK Data Capability', classifying data capability into human capital, including efforts on data-related expertise, tools and infrastructures for storing, analysing, and utilizing data, and data itself.³⁰

The UK government announced the need for the following human infrastructure to strengthen human capital.

- Citizens, who interpret data and know how to use the data,
- Business leaders, who are aware of the potential value of data analysis,
- Experts, who are capable of analysing data.

Despite the expansion of the data industry market, there is an important shortage of skilled experts in data analysis. Such dearth of skilled experts is one of the biggest impediments in in-depth data analysis activities. McKinsey (2011) predicted that by 2018, there was going to be a shortfall of between 140,000 - 190,000 people in the advanced analysis workforce and of 1.5 million data managers in the United States.

³⁰ HM Government (2013.10). Seizing the data opportunity; A strategy for UK data capability. p.5

Notwithstanding the excellence of new technologies and tools, actual outcomes depend on the capabilities of people who use them. Personnel with data processing and analytic skills capabilities are key elements of the workforce that should be acquired not only in the IT field but also by most businesses, organizations, and national agency. As such, data-related human capability is a very important factor in determining data capabilities. Nurturing and acquiring a professional workforce who can create values through data is emerging as a global challenge.

Strategies for enhancing human capabilities in data

It is difficult to give a clear, universal definition or classify tasks related to "Human capability in data utilization." It is generally defined by the data utilization and related capabilities that one should possess to become a professional and is largely classified into "Data analysis," "Data visualization," and "Data science."³¹

"Data analysis" uses various analysis tools to analyse and derives new insights from data. In terms of technical aspects, data analysis capabilities are based on an understanding of Oracle, SQL, and Java programs and include the ability to analyse data with statistics, machine learning, and text mining, with required capabilities in the utilization of statistical software such as R³² or SPSS ³³.

"Data visualization" converts data into an easy-to-understand-form and delivers the meaning and implications of the data effectively through interface design, user experience (UX) design, visual design, etc. Technical capabilities such as interface design and user experience (UX) are required.

Lastly, **"data science"** extracts data from various internal and external sources and provides solutions for a given task. Enhancing data capabilities requires capacities for utilizing various programming languages, such as Hadoop, Java, NoSQL, C++, Python, and R. Knowledge in various fields other than programming – such as statistics, analytics, and mathematics – is also necessary.

Category	Capabilities in data utilization
Data analysis	 Capability in assisting data science and utilizing various analysis tools in a given system in order to conduct actual data analysis Derives new insights from data
 Capability to visualize data and convert them t easily understandable to consumers Discovers new patterns and insights thro method of thinking 	
Data science	 Establishes a system for managing, utilizing, and analyzing data inside and outside the organization and supports

Table 4. Capabilities in data utilization

³¹ Korea Data Agency (2016). Report on the plan for cultivating professional workforce for data. p.18.

³² A language for statistical analysis and graphics and an open source software that can be used freely. First developed by Robert Gentleman and Ross Ihaka of the University of Auckland, provides various data analyses and graphs (Source: Wikipedia).

³³ (Statistical Package for the Social Sciences) Statistical analysis program that is used for statistical analysis and data mining, etc. First developed in 1968, currently sold by IBM (Source: Wikipedia).

major decision-making processes such as process
innovation, new product development, and decision on
marketing strategies

Source: Korea Data Agency (2016). Partially extracted from the Report on the plan for cultivating professional workforce for data

Tasks for enhancing human capabilities in data

Development of data specialists

To deal with the shortfall of professional data workforce, the government needs to cultivate data specialists by cooperating with universities and businesses. Development of a pool of specialists should be done with a proper training and education plan with consideration that demand for such manpower could change depending on the ICT environment and levels of technological capability, data infrastructure, and access. Moreover, such education is needed not only for the private sector, but for the public sector as well.

Case Study 16. Case of enhancing human capabilities in the UK ODI

The United Kingdom emphasizes improvement of K-12 educational processes and enhancement of analytical capabilities of public officials with its "strategy for enhancing data capabilities." In cooperation with employers, e-skills UK, Universities UK, and the Open Data Institute, the British government investigates the shortcomings of data analysis technology and search for areas of cooperation between the government and the industries. Notably, Universities UK reviews how data analysis technologies are taught in various fields and examines if more work is required to bring such technologies into various areas of knowledge.

In order to nurture a professional data workforce, the government came up with the following detailed action items: improvement of computing and science capabilities as essentials for data analysis jobs; government reforms directed to the secondary education including the development of educators; development of educational programs targeting employers on big data analysis technique; and reinforcement of data capabilities for public officials.

The ODI is running various online and offline education programs; it has developed a skill framework to schematize technologies related to open data by field – declined into introduction, publishing, management, business, analysis, and leadership – and established a system for interconnecting relevant ODI educational programs by individual technique.



Case Study 17. Enhancement of human capabilities in data in the public sector

Capability reinforcement is also needed for the staff in public offices, who should be able to use data analytics in their day-to-day business. Public sector-specific training programs should be drawn up to include data analysis tools for frequent business areas (civil complaints, tourism, and transportation), analytics, how to plan and manage big data projects, and output interpretation and implication-finding.

Training for public officials in the public sector can be broken down into the following categories: open data conceptual training, practical training, and online training. The following Korean cases in the table below can provide additional details:

Category	Training content
Conceptual training	Conduct conceptual training that focuses on the methodology related to task utilization for assisting in data-based administration, including pre-planning, planning, execution, and utilization of big data analysis
Practical training	Learn about big data analysis tools (R, Q-gis, etc.) and acquire methods for analyzing results through actual practice using sample data by specific areas
Online training	Create online training contents on the concept of open big data, domestic and foreign cases of utilization, process and method of performing open big data tasks, expected effects, etc.

Enhancement of open data capabilities for business leaders

It is advised that the government provides various educational opportunities to college students, job seekers, and start-up aspirants so that they can grow as leaders in start-ups and new businesses using open data.

In the training, the government needs to provide trainees who seek to start new businesses with **basic understanding of data, data collection and analysis, and service implementation through visualization and development of simple app**, which cover the overall data lifecycle.

The training will be helpful even for businesses or start-ups that are not necessarily or specifically related to data. Many countries have centres and organizations in place for the promotion of data utilization and better capabilities of business leaders.

Case Study 18. Support for expanding Open Data Use by ODX, Canada

In 2014, the Canadian government opened the ODX (Open Data Exchange) to promote the use of open data across the private sector. Although the word "Exchange" is part of its name, its primary role is to encourage commercial utilization of data in the private sector. The government and private partners are involved in the ODX operation³⁴, forming a network of 5 between the private sector, the government, the industries, academia, and research institutions.

In Canada, the release of open data is assigned to the portal (open.canada.ca), while the use of open data is operated with ODX, with both support and operational system. The ODX provides expert consulting for commercial utilization of data by the private sector, requests for release of open data according to private sector's demand and provides training on the analysis and

³⁴ Partners in funding, training, start-ups support, etc. include CDMN (Canadian Digital Media Network; operates start-ups and commercialization programs related to data), COMMUNITECH (Consulting), D2L (Desire2Learn, training on data analysis), University of Waterloo (engages in policy design), and OPENTEXT (business and customer management / Big data Cloud solution).



Expanding open data training to the general public

Apart from business utilization, there is a need to expand educational opportunities for general citizens so that they can directly and indirectly participate in **activities aimed at solving social problems and improving the transparency of society using data** through enhanced understanding of data.

Case Study 19. Start-up support centre for open data in Korea: "Open Square – D"

"Open Square-D" serves as a control tower for revitalizing start-ups using open data. It is a space, where people with ideas on open data gather to exchange their experiences and techniques so that they can achieve growth through commercialization and start-ups. It offers support programs for materializing ideas on open data and providing assistance in the entire life cycles of open data use from start-up to sustainable growth. At present, 4 centres are in operation, and the South Korean government plans to increase the number across the nation.

"Open Square-D" provides offices for start-up aspirants using open data, with collaboration spaces that are open to the public, where anyone can exchange and share their ideas and skills, and test labs, where testing facilities are offered for free. In addition, Open Square- D runs training and consulting programs for beginners up to professionals, for people planning their own business.



3.7 Data Feedback Framework³⁵

Definition of 'Data Feedback" and Needs

"Data feedback" refers to the participation of open data users in creating open data or establishing open data policies (ex. general citizens, civic groups, research centres, corporations), wherein data suppliers (ex. government and public institutions) reflect the opinions of data users. The advantage of "data feedback" is that it can provide optimized data to those who actually use it.

Past data policies have been established around government and suppliers. Therefore, data that was not necessarily wanted by data consumers or that was inconvenient to use has been often produced, failing to encourage active utilization. However, when open data users participate in the government's data policy decision-making process, data that users actually want, that meets their needs and that is easy-to-use can be produced, boosting the use of data.

Reflecting users' opinions in data shaping and policy making, high-quality data optimized for data utilization can be produced and provided in large quantities, increasing the industrial utilization of data and contributing to new data-based industries and job creation.

"Data feedback" environment

³⁵ See Chapter VI "Enabling Policies" of "Realizing Data-Driven Governance" for additional information on citizens or public generated data.

Co-creation of policies

"Co-creation of public policy" refers to the mechanism, wherein government and nongovernment stakeholders with complex interest in policy issues participate in the decision-making process on equal footing and come up with policies together. This is also known as multistakeholderism, "Cooperative governance", and "Civilian-government cooperation governance."

Co-creation – wherein civilians and the government decide major policies through discussions – is drawing attention in recent open government and open data policy. Co-creation in terms of open government is under active discussion in the Open Government Partnership (OGP).



Box 5. OGP (Open Government Partnership)

OGP was founded when former US President Obama proposed a multilateral cooperation project for enhancing government transparency in his speech at the UN Council. As of 2017, the OGP membership stood at 75 countries.

Each member country announces and executes their own Action Plan about open government policies every two years. In this process, the governments of member countries join together to decide on Action Plans and their execution. The process is assessed by an IRM (Independent Report Mechanism), an independent evaluator.

See http://opengovpartnership.org

Box 6. Governance

The United Nations Development Programme (UNDP), in its 1997 policy paper, defined governance as "the exercise of economic, political and administrative authority to manage a country's affairs at all levels. It comprises the mechanisms, processes and institutions through which citizens and groups articulate their interests, exercise their legal rights, meet their obligations and mediate their differences".

Components of co-creation³⁶

There are 3 components:

- Spread of information

Spread of information means that all the information involved in the policy-making process should be provided on time, with feedback of stakeholders including citizens and the civil society incorporated in the process.

- Space and platform for dialogue and co-creation

There should be an inclusive and sustainable forum for dialogue, in which the government and stakeholders participate and where their collective opinions should be reflected into policies.

- "Co-ownership and joint decision-making"

Government, the civil society, and other stakeholders must make mutual contributions by having a sense of ownership over this process.

Organizations and resources to carry out the joint creation (multistakeholder model) are referred to as "multistakeholder forums," and are operated in various forms in each country as shown below.

Case Study 20. Multi-stakeholder Forum related to open data

ODUG and ODI of the United Kingdom

Gathering stakeholders in the private sector to encourage innovation, the United Kingdom has made decisions on the priority of data to be shared, provided funding support for data-based start-up companies, and supported various events for promoting data utilization.

The British cabinet has come up with a mechanism for communicating with stakeholders in the private sector through a group called Open Data User Group. As an independent advisory group that represents data users, ODUG not only provides advice on financial support related to open data but also identifies data sets that are to be used openly; it is working as an official counselling centre for setting their priority.

Open Data Forum of Korea

³⁶ Standard on OGP participation and co-creation (OGP, 2017), Handbook on Multi-stakeholder Forum (OGP, 2017).



Source: Open Data Forum (http://dok.kr)

"Open Data Forum" is a "communication and consultation council for the use of open data through private-public cooperation" and launched to unify multi-lateral communication channels and create open data-driven social values. The Ministry of Interior and Safety and National Information Society Agency – which are responsible for managing this council – in collaboration with the private sector, benchmarked against the British ODUG model and reshuffled the existing "National Open Data Forum" to launch the "Open Data Forum" in the multi-stakeholder Forum format in July 2017.

With the steering committee and the secretariat involving participants from the civil society, businesses, media, and academia, the "Open Data Forum" aims to expand and compose working groups in order to serve as a policy advisory group for data-driven social innovation and as a channel connecting the government and user groups in the private sector for open data policies.

Data-driven Social Innovation

As previously discussed in detail, change into a data-based Smart government does not merely end with internal innovation. Big data and open data are used as mechanisms of civic participation, with citizens using data and ICT technologies on their own to solve social problems, in which case "datadriven social innovation" occurs. Outcomes that are economically measurable – such as start-ups using open data – can be included here. Beyond this, innovations in areas that are difficult to measure with economic values, such as solving data-based social problems in non-profit sectors, are included as well.

Areas and characteristics of data-driven civic participation

The tables below show major areas of civic participation driven by open data.

Table 6. Type of participation by civil society

	Political participation	Administrative participation	
Relationship of discussion & partnership	 Listen to advice or response of citizens on legislation Listen to opinions of interest groups related to legislation Reflect the opinions of citizens and interest groups in enacting legislation using civic committee, etc. 	 Listen to the advice or response of citizens on administrative decisions Listen to the opinions of groups subject to specific policy programs Reflect citizens' opinions in the decision-making process related to operating policy programs using civic committee 	
Constant monitoring	 Monitor political activities through third-party members 	- Monitor administrative activities through third-party members and have them participate in the performance evaluation of policy programs	
Consumer choice & direct participation	 Participate in elections Have citizens directly suggest and enact legislations 	 Have citizens directly plan and suggest administrative programs Introduce competition in government services that are provided so that people can choose service providers 	

Source: Civic participation in the Open Government – Category of civic hacking, Cho, Byung Woo and Yoon, Sang Oh, Korean Association for Policy Studies (2017).

Country	Case	Details
USA	Open Disclosure California	Information on campaign funds to Oakland regional election candidates <u>http://opendisclosure.io</u>
	Asheville Budget Explorer (Code for Asheville)	Budget information in Asheville http://www.avbudget.org/Overview
Korea	How are my taxes being used? (Code for Seoul)	Templates on visualizing budget execution <u>http://wheredoesmymoneygo.org</u>
	All about Korean politics (Team POPONG)	Information on the legislative activities of assemblymen from the first to the nineteenth terms

	http://pokr.kr

For categories and cases, refer to Civic participation in the Open Government – Category of civic hacking, Cho, Byung Woo and Yoon, Sang Oh, Korean Association for Policy Studies (2017).

Country	Case	Details
USA	Vote ATX (Open Austin)	Service for facilitating citizens' voting participation by showing voting locations when an address is entered <u>http://voteatx.us</u>
	DemocracyOS	Platform for providing a system of gathering of opinions, discussions, and decision-making process in a democratic manner <u>http://democracyos.org/</u>
	Countable	Platform where citizens can discuss and gather opinions on issues of interest to be legislated by Congress <u>http://countable.us</u>
Republic of Korea	National Assembly Talk (Wagl)	Platform where citizens directly suggest legislations to the National Assembly <u>http://toktok.io</u>
	Parti (Parti)	Provides a communication frame for democracy in daily lives or within organizations <u>http://parti.xyz</u>

Table 8. Participation in politics and administration based on data or Internet technology

Source: Civic participation in the Open Government – Category of civic hacking>, Cho, Byung Woo and Yoon, Sang Oh, Korean Association for Policy Studies (2017).

Box 7. Civic hacking

Civic hacking is a new type of civic participation that has emerged in the United States at the end of the 2000s. Using open data provided by the government and Web/App technologies, it aims to achieve various goals through genuine cooperation between citizens and the government, such as government innovation, improvement of public services, assured transparency, and improvement of civil rights.

Unlike previous researches on online civic participation, civic hacking, which presents the role of participants, showcased a new form of participation: developer of government services using open data.

Source: Extracted from Civic participation in the Open Government – Category of civic hacking, Cho, Byung Woo and Yoon, Sang Oh, Korean Association for Policy Studies (2017).

Tasks for expanding data-based civic participation

Civic participation based on data can be expanded in terms of quantity and quality only when data supply and demand increase as shown below. In terms of specific methods, the government can provide support using the national budget, or an official private-sector foundation that has traditionally supported civil society movements which can make contributions.

- Supply: Expanding the supply of quality open data

First, sufficient amount of high-quality data should be supplied. The success of this lies in the effort for internal innovation by government departments that manage open data policies. Political leadership for effective control on all processes of open data production and management within the government and restructuring the government to tackle such tasks is required. Efforts for establishing partnership with the private sector – the consumer – are also essential to improve the quality of open data.

- Demand: Expanding the data literacy of citizens

In order to revitalize data-based civic participation, data literacy – or the capacity of citizens to understand and utilize data – should be adequately acquired as well. No matter how much quality data are opened, social innovation may be difficult to achieve if there is a limited number of citizens – data consumers – who can use them.

To accomplish this, education on data literacy for civic society should be expanded along with support for physical foundation, such as physical space and facilities, to develop the data-related capacity of the civic society.

Case Study 21. Support for improving demand for data

Non-governmental support / Non-profit sector: OpenGov Hub of the USA (http://opengovhub.org/)

Located in Washington D.C., USA, OpenGov Hub is supported by main sponsors such as "OMIDYAR NETWORK," a co-working community and global investment group consisting of over 35 media and non-profit groups related to Open Government and open data, as well as "HEWLETT FOUNDATION," a charity that supports educational institutions.

It is being used as an office space for tenants and as networking space for education, innovation, and cooperation for issues related to the Open Government. It has branches in Kathmandu and Nepal.

Figure 19. Tenants of OpenGov Hub



List of NPOs residing in OpenGov Hub. As shown in the figure, many institutions and groups related to OpenGov work in a single space, producing synergistic effects through cooperation and joint seminar, etc.

<image>

Source: opengovhub.org

OpenGov Hub office: Tenants are working together in a single, large space without walls in the coworking space

This section aims to:

- Discuss the factors for a successful data-driven smart government.
- Learn the challenges of using open data, including privacy protection and usage fee

4.1 Raise Awareness about Open Data Release among Public Institutions

Issue of lack of "awareness about open data release"

The economic and social benefits and profitability from open data release and usage are becoming increasingly visible. However, the need for "open data release" was not equally shared among public institutions early on, mostly because the issue was not regarded as urgent when compared to other serious social issues. In addition, public institutions, which are the main actors behind data openness, are reluctant to open their data as they do not see immediate benefits from the release and risk of having an added workload. This lack of accountability will degrade the quality and accuracy of the data.

Solution

It is first necessary to be confident that public officials in charge of open data release share conviction about the need of open data. As mentioned earlier, the 4th industrial revolution and ICT development are a global phenomenon, and it is now imperative that data-based smart government should be transformed into a data-driven one.

To foster a change of mindset and perception among public officials, government-wide policies should be established. Moreover, continuous training and event-holding could serve as incentives for leading institutions in promoting open data release and use as reward and recognition (i.e. awards and scores).

4.2 Data Privacy Protection Issue

Issue of privacy protection

Open data, is in a collision course with personal and private information, which requires to be protected and not to be disclosed. Withstanding differences in legislations of different countries, generally speaking, personal information should not be collected and used without the consent of the person owning the information. Even open data, which utilization was authorized by the owner, is limited in use. Indeed, public data portals can be opened to the private sector for public use, but even open data that the use has been agreed to by the information owner, are posted to public data portals with personal information preemptively masked. As such, conflict between open data and personal information is not possible, at least theoretically.

³⁷ See Chapter VI "Enabling Policies" of "Realizing Data-Driven Governance" for additional information on policymaking that can enable data-driven governance.

However, the reality is often quite different. The definition of personal information differs from country to country, and in particular, the scope of personal information is often abstract, leaving it largely up for interpretation. Personal information in one country is not considered personal information in another country. With the government possessing much of personal information, there is very high likelihood of violating privacy unintentionally in the process of utilizing open data. Personal information issues should remain on the main agenda and be considered, as enormous amount of data is to be processed increasingly fast in the open big data environment, with the purpose of scientific policy-making by the government with the help of a variety of technologies.

Solution

Multiple solutions can be considered to protect personal information without impeding on the utilization of open data. Anonymization, de-identification, and hiring specialists, experts and specialized agencies in data public use, are some of the possible options:

- Anonymization and de-identification for public purpose use

Countries around the world regulate the issue of personal information with anonymization or non-identification in their legislations.

Case Study 22. Legislations on the protection of personal information around the world in the context of open data

EU GDPR(General Data Protection Regulation) taking effect in 2018: Processing personal information to archive for public interest, scientific or historical research, or statistical purposes (Paragraph 1(b), Article 5; Paragraph 1, Article 89).

Processing of personal information for archiving purposes for public interest, scientific or historical research, or statistical purposes is permitted without consent as long as they implement safeguards, including pseudonymization³⁸.

Personal Information Protection Act of Japan: Anonymously processed information refers to personal information obtained after being processed to prevent the identification of a personal identity in accordance with measures stipulated in each paragraph of article 2.9. Once processed, original personal information cannot be recovered (article 2.9)

In 2017, the act became effective, which covers the new concept of anonymously processed information to regulate personal information protection.

Source: Official Journal of the European Union, EU General Data Protection Regulation and National Personal Information Protection Commission, Private Information Protection Act

- Hiring specialists (or groups) related to public purpose use

Case Study 23. HIPPA privacy rules in the medical field of the USA

³⁸ Pseudonymisation: The processing of personal data in such a way that the personal data can no longer be attributed to a specific data subject without the use of additional information, No. 5, Article 4

The rule largely presents two types of non-identification plans for the conditions of sharing data, and one of them is "Expert Determination Method." Experts review data and determine appropriate means of de-identification for minimizing the risk of re-identification.

4.3 Open Data Use Fee

Issue related to Open Data fee

When freely using government (public) data, there is a need to discuss whether the government, as a data producer and administrator, should assume responsibility for all costs of collecting and processing data, or whether citizens (users) who request and use such data should be partly responsible for such costs.

Provision of open data requires hardware (HW) such as network and physical servers, and possibly additional workforce for extracting and processing data. Therefore, it is not easy to determine whether it is appropriate to have users shoulder the cost of providing open data, and if so, how much should they pay for. Furthermore, since the scope of public institutions that manage open data is wide and they are distinctive on their own, a uniform standard to apply across the spectrum should be given additional thought.

Solution

While the beneficiary principle is established for goods and services supplied by the public sector, most countries offer open data free of charge as public goods. That is because they do not want fees to limit people's right to use open data. The global trend is that minimum price is charged for open data and the user's right to open data should not be limited with an excessive price tag.

Case Study 24. Global trend in open data charge

- Given the increased interest in recycling open data and in enacting the relevant policies, countries tend to lean toward policy for the free provision of open data.
- Countries like the USA, EU, and Japan are announcing the prohibition of limiting citizens' right to access public information by charging costs of providing data.
- Notably, the US asserts that all open data created by the government are national assets that can be utilized by citizens to the hilt to maximize the economic benefits. Therefore, it does not charge any cost, including copyright, in providing data other than costs related to distribution.
- In Europe, most data-providing institutions used to regard their data as profitable products, making it mandatory to defray a little or a significant part of the operation expenses using the profit from data sales and having no principle of free provision. Recently, however, there are growing opinions that institutions should provide public information at a marginal cost unless there are special reasons to prevent doing so.

• The United Kingdom has legislated and enacted the Freedom of Information Act of 2000. It has a cost system that classifies data into core material and value-added material as well as for charging costs for value-added material. As such, the British government introduced a license policy on public information in 2001 to classify license on core material and value-added material, providing core material free of charge and charging a separate cost for value-added materials.

It is deemed desirable that the government assumes responsibility for the costs of production and management. If the government wishes to charge the cost of data utilization to users for specific reasons – for example, if the data size is large, if extra lines are required for sending data in real time, or if much effort is required for processing data – however, it would be beneficial to have users shoulder such costs.

In order to charge the minimum actual costs required for providing publicly owned data to citizens, a cost calculation standard that is acceptable to everyone must be available. In other words, a standard for determining the cost based on various standards – such as data type, method of providing data, value of data, and premium for intellectual property rights of data – should be presented.

Moreover, although it is difficult to provide a single, comprehensive standard since public institutions that collect and manage information are characterized by diverse forms of operation³⁹, it is necessary to think about rational plans for charging costs according to the characteristics of each institution. There is also a need to consider how to return and utilize the profit from providing open data created from citizens' taxes.

Case Study 25. Case in Korea – Open Data Use Fee

Korea aims to offer open data for free in principle. Pursuant to Article 35 of the Public Data Act, however, a minimum cost is allowed to be charged to the users. In the case of charged service, the public institution is banned from excessive charge, which might constrain use of open data. The charging principle for open data is to put as little a price as possible on open data. The applicable Korean law and enforcement rules specify the charging criteria to be employed by a public institution as following:

- A public institution is required to specify whether its open data is charged or not, and if charged, what criteria and unit are in use for charging when it register its open data to the open data portal (<u>www.data.go.kr</u>).
- Prior to the enactment of the Public Data Act, information such as patent information, meteorological information, and statistical information used to be provided only to businesses with specific conditions or for non-profit purposes such as academic research. Even in such cases, free or minimal costs have been charged. After the enactment of the Public Data Act, however, institutions have been trying to restructure their criteria for cost imposition in keeping with the purpose of policies for opening up open data.

³⁹ There are public institutions that operate using the profit from creating and providing open data.

Source: Adapted from "Study on cost sharing plan for promoting the use of open data" (Choi, Jin Won, et al, November 2014)

5 Bibliography

1. Introduction

Andrejevic, M. (2014). Big data, big questions | the big data divide. International Journal of Communication, 8, 1673-1689.

Hargittai, E. (2001). Second-level digital divide: Mapping differences in people's online skills. arXiv preprint cs/0109068.

McCarthy, M. T. (2016). The big data divide and its consequences. Sociology Compass, 10(12), 1131-1140.

Norris, P. (2001). Digital divide: Civic engagement, information poverty, and the Internet worldwide. Cambridge University Press.

UN (2014). A World That Counts. A report from UN Data Revolution Independent Experts Advisory Group.

2. Data-driven Smart Government

Jeong, Guk Hwan (2008). How will the government in the future society work? – Materialization of next-generation e-Government through completion of information sharing and system linkage. e-Government Focus.

KCERN (2016.10.25). Block Chain and Governance Innovation. The 30th KCERN Regular Forum

Ministry of the Interior and Safety. (2017). Presentation material on "Seminar on the Future Strategy of IT Service by the New Government." Korea Society of IT Services.

Ministry of the Interior and Safety (2016). "Basic Plan on e-Government 2020."

Myeong, Seung Hwan (2012). On Smart e-Government. Yulgokbook Publishing Co. National Informatization Strategies Committee. (2011. 11). Materialization of Smart Government using Big Data. The 8th National Informatization Strategies Forum.

National Planning Advisory Committee (2017). Five-year Planning on National Administration of the Moon Jae-In Government.

Nam, Goong Geun (2011). Policy. Beobmunsa.

NIA (2013). "Policy tasks for data-based national future strategy".

Press release by the Ministry of the Interior and Safety (2017.3.8). "Smart Government Pours Priming Water for the Fourth Industrial Revolution."

Steering Committee of Government 3.0 (2017). Government White Paper 3.0.

Thom Rubel, Smart Government: Creating More Effective Information and Services, IDC http://smarterub.com/Solution.aspx?pkid=643

3. Framework for Open Data and the Use

Cho, Byeong Woo and Yoon, Sang Oh (2017). Civic participation in the Open Government – Category of civic hacking. Korean Association for Policy Studies.

Choi, Yoo and Cha, Hyeon Sook (2016.12). Ex Post Legislative Evaluation on the Information Act. Korea Legislation Research Institute.

Gonzalez-Zapata, F., & Heeks, R. (2015). The multiple meanings of open government data: Understanding different stakeholders and their perspectives. Government Information Quarterly

Gruen, Nicholas, John Houghton, and Richard Tooth. (2014). "Open for Busines: How Open Data Can Help Achieve the G20 Growth Target". Omidyar Network.

Gurin, J. (2014). Open Data Now. Mc Graw Hill Education.

HM Government. (2013. 10.) "Seizing the data opportunity; A strategy for UK data capability"

Jeong, Guk Hwan, et al (2013). "Study on the evaluation model for measuring performance on the openness and utilization of open data." KISDI.

Jeong, Joon Hwa (2014.12). Status and tasks of services on opening open data and supporting big data utilization. National Assembly Research Service.

Jeong, Myeong Woon (2015.9). Research on structuring data-related legislations for materializing evidence-based scientific administration. Steering Committee of

Government 3.0.

Jeong, Yong Chan (2017). "Strategies for promoting data economy during the Fourth Industrial Revolution." Korea Information Society Development Institute.

Jeong, Yong Chan and Han, Eun Young. (2014.11). "Study on the promotion of big data industries." Korea Information Society Development Institute.

Jetzek, T. (2016). Managing complexity across multiple dimensions of liquid open data: The case of the Danish Basic Data Program. Government Information Quarterly

Korea Data Agency (2016). Report on the study for cultivating professional data workforce.

Korea Development Institute (2016). "Analytical study on the economic effects of opening open data."

Kulk, S., & van Loenen, B. (2012). Brave new open data world?. IJSDIR

Lee, Si Jik (2014.12). Plan for legislative and policy improvement on the provision and promotion of open data utilization. KISDI.

Manyika, James, Michael Chui, Diana Farrell, Steve Van Kulken, Peter Groves, and Elizabeth Almasi Doshi. (2013). "Open Data Unlocking innovation and Performance with Liquid Innovation". McKinsey Global Institute

McKinsey Global Institute. (2011). 'Big Data: The next frontier for innovation, competition, and productivity'

Ministry of the Interior and Safety (2013). Basic Plan on Government 3.0.

National Information Society Agency (2014.11). Study on cost sharing for the promotion of open data utilization.

National Information Society Agency (2012.9). Report on legislative review for the Act on the Promotion of Provision and Use of Public Data.

National Information Society Agency (2012.4.12). 3 factors for successful big data utilization: Resource, Technology, and Workforce. IT & Future Strategy No. 3.

National Information Society Agency (2012.12.20). Prospect on creating big-data-based jobs. IT & Future Strategy No. 15.

NIA. (2013). Function and role of the open data platform. A report from the National Information Society Agency.

OPENGOV. (2014). Open Government Platform (OGPL). Park, Sung Hyun (2017). Arrival of the data economy era – Reinvent open data. Future Korea.

Stefaan Verhulst and Andrew Young. (2016). "Open Data Impact when demand and supply meet: Key findings of the Open Data Impact case studies"

World Wide Web Foundation. (2016). Open Data Barometer 4th edition <u>www.opengovplatform.org</u>.

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/25413 6/bis-13-1250-strategy-for-uk-data-capability-v4.pdf

http://www.etnews.com/201001290062

http://www.futurekorea.co.kr/news/articleView.html?idxno=41799

The UK ODI <u>https://theodi.org/course</u> <u>https://theodi.org/open-data-skills-framework</u>

Canada ODX https://codx.ca/

www.oecd.org

https://www.opendatabarometer.org

https://theodi.org/open-data-skills-framework

https://opensquared.org

http://opengovpartnership.org

https://opengovhub.org

4. Theory and Practice of Data-driven Smart Government

Choi, Jin Won, et al (2014.11). Study on cost sharing plan for promoting the use of open data. Korea Association for Informedia Law.

Kim, Kyung Hwan (2017.7). Response to the Fourth Industrial Revolution / EU GDPR. Presentation at the Seminar on Personal Information Protection.