

Basic Principles on Open Data (Provisional Translation)

May 30, 2017

Approved by the IT Strategic Headquarters* • Strategic Committee
for the Advancement of Utilizing Public and Private Sector Data

Since the Great East Japan Earthquake on March 11, 2011, there has been growing awareness of the publication and use of data held by Japan's national government, local governments, and companies.¹

The government has been promoting open data initiatives based on the Open Government Data Strategy (approved by the IT Strategic Headquarters on July 4, 2012), which demonstrated the understanding that public data is a common property of citizen.

The documents Toward the Development of New Open Data (approved by the IT Strategic Headquarters on June 30, 2015) and Open Data 2.0 (approved by the IT Strategic Headquarters on May 20, 2016) were indicative of a change of mindset, with a shift away from initiatives centered on data publication, toward the promotion of Solution-oriented Open Data utilization.

The Basic Act on the Advancement of Utilizing Public and Private Sector Data (hereinafter “Public–Private Data Utilization Act”) was promulgated and entered into force on December 14, 2016. Its objective is to contribute to the realization of a society in which the citizens are able to have safe, secure and comfortable life. The act includes provisions concerning the easy use of public and private sector data held by the State, local governments, and companies.

This document sets out the basic principles for the State, local governments and companies to publish and promote utilization of public data, based on the concept of Open Data by Design,² in keeping with initiatives to date.

1. Impact of Open Data

The primary meaning and objectives of publishing public data in a form that permits secondary use and promoting utilization thereof are as follows.

(1) Resolving social challenges and revitalizing the economy through promoting public participation and public-private partnership.

*Abbreviation of Strategic Headquarters for the Promotion of an Advanced Information and Telecommunications Network Society

¹ Requests were made for cooperation in providing disaster-related information in a file format that would facilitate the processing of the information. In addition, information held by companies about whether or not roads were passable was published and used.

² This is the practice of planning, developing, and operating all information systems and operational processes relating to public data on the premise of open data.

As the use of public data by a wide range of entities progresses, a diverse array of creative services will be offered in a speedy and efficient manner, while public services will be delivered and improved through public-private partnerships. As well as enabling increasingly diverse needs and values, and environmental changes such as technological innovation to be addressed appropriately, this will help to solve challenges faced by our nation, such as the severe financial situation and Japan's rapid development into a super-aging society with a low birthrate.

This will also encourage the creation of diverse new services and businesses by startup companies, and promote greater efficiency in corporate activities, thereby leading to the economic revitalization of the nation as a whole.

(2) More advanced and efficient public administration

Evidence-based policy making (EBPM) — the use by national and local governments of information gained from data utilization as the basis for planning policies and measures — will help to promote more effective and efficient public administration.

(3) Increasing transparency and trust

Publishing the public data used in policy making will enable the public to fully analyze and judge policies, thereby increasing the transparency of public administration and the public's trust therein.

2. Definition of Open Data

Open data is defined as public and private sector data held by the national government, local governments, and companies, which is published in a form to which all of the following applies, so that every citizen can easily use (process, edit, redistribute, etc.) it via the internet, etc.³

- ① Data published under a rule allowing secondary use, whether or not for commercial purposes
- ② Data published in a machine-readable⁴ format
- ③ Data that can be used free of charge⁵

3. Basic Rules on Open Data

(1) Scope of data published as open data

Based on the understanding that public data is common property of citizen, data held by each ministry and agency — including data that has been the basis for policy making

³ However, measures such as requiring users to register in advance or other restrictions on methods of accessing the data may be imposed for security reasons.

⁴ “Machine-readable” means that a computer program can automatically process and edit the data.

⁵ While unable to be described as open data, there are cases in which the costs of maintaining and managing the data provision system are covered by imposing a charge on a limited range of users.

(including laws, regulations, and budgets) — must be published as open data, in principle.

If a request is made for publication of information that is not appropriate for publication (such as (i) information containing personal information; (ii) information that could impede national or public safety, or the maintenance of order; or (iii) information that risks harming the rights and interests of a corporation or individual), the reason why it cannot be published as open data must be published,⁶ in principle.⁷

Moreover, even if the information is not appropriate for publication as open data at present for reasons such as those described in (i)–(iii) above, it is recommended that the possibility of releasing it as open data in the future should be opened up by publishing information other than the problematic data items or actively utilizing the limited publication⁸ technique, in which the information is shared among a limited range of relevant parties.

(2) Rules on secondary use of published data

The data published on ministry and agency websites⁹ shall, in principle, be subject to the Government of Japan Standard Terms of Use¹⁰ and the secondary use of published data shall be actively encouraged, except data whose secondary use is not permitted on specific and reasonable grounds.

(3) Publication environment

Each ministry and agency shall publish data on its website in an easily searchable and usable form. Where there is thought to be a particularly high need for the data, mechanisms permitting bulk download or provision via APIs shall be promoted, for the convenience of users and to reduce the burden on systems and increase efficiency.

Furthermore, to encourage data utilization by facilitating searches of all government open data across organizational boundaries, metadata such as an outline of the data and its format shall be registered and published on the government's data catalog site :

⁶ If the information cannot be published as open data for legal reasons or on other reasonable grounds (including cases in which there are restrictions on secondary use), this fact will be indicated in specific terms.

⁷ If publishing the reason why the information cannot be published would have the effect of disclosing the information that cannot be published (Non-Disclosure Information), the reason may be withheld.

⁸ For example, it is envisaged that information, which risks impeding the maintenance of public safety if widely published, could be shared among a limited range of relevant parties and utilized for safety measures. In this situation, it would be desirable to publish information about what kind of information is being shared and among which parties.

⁹ Used in the broad sense, including websites, database sites, and sites for carrying out individual operations on the internet, which ministries and agencies (including organizations belonging to each ministry and agency, such as facilities and other organs, and local branch offices) have established under their own name.

¹⁰ Rules concerning the use of ministry and agency websites displayed as a uniform template in order to encourage the secondary use of data published on the websites of governmental organizations. Basic Approach to the Publication of Ministry Data to Encourage Secondary Use (Guidelines): Appendix 1, Government of Japan Standard Terms of Use (Version 2.0).

DATA.GO.JP , in accordance with the standards.

(4) Published data format, etc.

Data for publication shall have a structure and data format that is machine readable,¹¹ in principle.¹² Efforts shall be made to publish data using terminology and formats that make it easier to utilize, with reference to IMI: Infrastructure for Multi-layer Interoperability and the 5-star¹³ indicators used for evaluating open data achievement levels.

Information will continue to be published in data formats suitable for the conventional form of use — namely, being read by humans — so that information can be made available to the public, but even in this case, it will be necessary to ensure that text search and extracting can be carried out.

The National Strategy Office of IT, Cabinet Secretariat will continue to promote the standardization of data structures and formats in partnership with relevant ministries and agencies.

Data containing information on corporations shall also specify the Corporate Number.

(5) Updating of already-published data

Data that needs to be published swiftly or kept up to date shall be published as swiftly as possible and updated in a timely and appropriate manner. The data update cycle shall be disclosed and steps taken to ensure that users can ascertain the timing of updates in advance.

4. Mechanisms to Promote Publication and Utilization of Open Data

(1) Promoting Open Data by Design

Based on the concept of Open Data by Design, each ministry and agency will take the necessary measures to publish data held by administrative bodies in a form that is easy for users to utilize, starting from the planning and design stage of administrative procedures and information systems. Under the Government CIO, the National Strategy Office of IT will provide the necessary advice to each ministry and agency.

¹¹ There are a number of levels of machine readability, determined on the basis of the extent to which a computer can reuse the data without human intervention. To enable computers to reuse data automatically, it is necessary to ensure that the logical structure of the data in question can be determined (read) and that the values in the structure (such as figures and text in a table) can be processed.

¹² See Basic Approach to the Publication of Ministry Data to Encourage Secondary Use (Guidelines): “3. Approach to Expanding Publication Using Data Formats That Are Easily Machine-readable” and Appendix 2 “Important Points Concerning the Preparation of Figures (Tables), Text, and Geospatial Information as Data.”

¹³ A five-rank indicator advocated by Tim Berners-Lee, the inventor of the Web. Data made available for secondary use under an open license (in whatever file format) is one-star Web data, while that made available in the comparatively easily machine-readable csv format is three-star Web data.(<http://5stardata.info/en/>)

(2) Reflecting user needs

In promoting open data, it is crucial to ensure that user needs are reflected accurately as initiatives progress by publishing¹⁴ a list summarizing the data that they hold and its publication status. Accordingly, each ministry and agency will strive to identify user needs, including potential needs, and publish data in a form tailored to those needs.

5. Promotion Framework

(1) Establishing help desks

A general open data help desk will be established within the National Strategy Office of IT to proactively gather information and gain an understanding of user needs (requests) and opinions concerning open data. In addition, help desks will be established in each ministry and agency and shall proactively respond to individual inquiries from users concerning open data.

(2) Promotion body

Under the leadership of the Government CIO, the National Strategy Office of IT will work in partnership with each ministry and agency to formulate a government-wide open data policy and undertake general coordination. In addition, it will promote government-wide open data initiatives, such as reviews and follow-ups concerning each measure, including the publication status of data by each ministry and agency. Moreover, the National Strategy Office of IT will cooperate with relevant ministries and agencies in encouraging initiatives by local governments.

Each ministry and agency's Chief Information Officer (Ministry CIO) will promote open data initiatives within their own ministry or agency, as well as encouraging open data initiatives by independent administrative agencies and public utilities, among others. In addition, each ministry and agency's Deputy Chief Information Officer (Deputy Ministry CIO) will assist the Ministry CIO in providing direction and supervision within the ministry or agency regarding open data measures.

6. Open Data Initiatives by Local Governments, Incorporated Administrative Agencies, and Companies

(1) Local governments

Article 11 (1) of the Public-Private Data Utilization Act stipulates that local governments,

¹⁴ To enable such lists to be prepared, each ministry and agency will conduct an inventory of the data that it holds before the end of FY2017. Once the lists concerned are published, the government will organize a public-private round-table as trial (here, "private sector" includes companies, researchers, and citizens) to discuss approaches to the publication and utilization of data in fields where needs are high. At the round-table discussion, participants will discuss approaches to the publication and use of public and private sector data, including data held by local governments and companies (including incorporated administrative agencies).

just like the national government, are to implement the necessary measures to enable citizens to easily use the data that they hold.

Local governments are required to promote open data in keeping with the spirit of the Public–Private Data Utilization Act and these basic principles.

In doing so, it would be desirable for them to strive to publish data in a standard format and in compliance with standard rules, so that the data published by national and local governments can be used across organizational boundaries. One effective measure would be for multiple organizations to jointly set up an open data portal site.

The government will actively support open data initiatives by local governments, by such means as deploying open data experts; identifying the fields and data items regarding which publication is particularly desirable in the case of local governments; providing guides and anthologies of examples of advanced initiatives; and providing human resource development tools.

(2) Companies

Article 11 (2) of the Public–Private Data Utilization Act stipulates that companies (including incorporated administrative agencies) are to endeavor to implement the necessary measures to enable citizens to easily use data that contributes to enhancing the public interest.

It would be desirable for incorporated administrative agencies, universities, and the like to promote open data initiatives compliant with these basic principles, since their operation, projects, and/or research are funded by the government.¹⁵

Moreover, given the public nature of companies operating in the field of public utilities such as electricity, gas, communications, broadcasting, and transport, it would be desirable to promote open data in keeping with these basic principles and user needs.

7. Review of These Basic Principles

Flexible reviews of these basic principles will be carried out in response to changes in technological trends and the international environment, while also taking into account such matters as the results of the various measures, and the views and requests of the public and companies.

¹⁵ Promoting Open Science in Japan (March 30, 2015) states, “Among the results of publicly funded research, published results and underlying data that serve as evidence for the findings of those published results shall, by default, be publicly disclosed. It is desirable for all other research data obtained as the results of research and development activities to be made public as well, to the extent that this is possible.” (http://www8.cao.go.jp/cstp/sonota/openscience/150330_openscience_en1.pdf)