

---

# The Classification and Potential of Business Archetypes by Using Open Data

**Run Duan**

Guangdong University of  
Technology  
Guangdong, China  
carrotmiao@126.com

**Tetsuo Noda**

Shimane University  
Matsue, Japan  
nodat@soc.shimane-u.ac.jp

**Abstract**

Public data collected or possessed by administrative agencies and subsequently released as Open Data is expected to bring about positive economic effects. The purpose of this paper is to summarize the business archetypes of using Open Data to establish whether this expectation holds true, and to classify Open Data business archetypes into 7 types to predict their commercial potential.

**Author Keywords**

Open Data; Open Innovation; Business Archetypes; Classification

**ACM Classification Keywords**

Social and professional topics~Economic impact

**Introduction**

Open Data is defined as “public data, from government of other sources, that’s available for anyone to access for personal or business use” (Gurin, 2014. p.2). This definition first developed by Open Knowledge Foundation in 2006, that is “a piece of data or content is open if anyone is free to use, reuse, and redistribute it—subject only, at most, to the requirement to attribute and/ or share-alike”(ibid.,p.267).

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

OpenSym '19, August 20–22, 2019, Skövde, Sweden  
© 2019 Copyright is held by the owner/author(s).  
ACM ISBN 978-1-4503-6319-8/19/08.  
<https://doi.org/10.1145/3306446.3340830>

**Business model:** Business models were descriptions of how a firm or, more generally, a person or an organization does business. They identified five different perspectives in business model research. They are business model activities, business model logics, business model archetypes, business model elements, and business model alignment. Each of the five perspectives has a distinct way of defining business models and all of them are meaningful in their own right. business model archetypes (Ritter and Lettl,2017).

**GovLab:** GovLab is an action research center based at New York University's Tandon School of Engineering and MIT media lab of Massachusetts Institute of Technology's School of Architecture + Planning. Retrieved July 8, 2019 from <http://www.thegovlab.org/about.html>

In this paper, we follow the definition of Open Data by Gurin (2014), and state it as open accessed data from governments, companies, organization, and people, both public sector and private sector, that can be reused, redistributed, and create value. We aim to analyze the value of Open Data from the classification of business archetypes by case studies. Business archetypes defined by Ritter and Lettl (2017), which describes the generic logics of how firms do business. It is separated from the business model research.

### Literature Review

When estimating the value created by Open Data, there are several perspectives. In the term of quantitative research, the European Commission Vickery research project (2011) estimated the value of Open Data is approximately 200 billion Euro annually or 1.7% of GDP across the EU 27 countries. Omidyar Network (2014) announced that Open Data will strongly support the G20's economic growth and concluded that the contribution of Open Data contribute as much as 55% of the goal of 2% growth of G20's GDP.

In term of qualitative research, Goldstein (2013) admitted Open Data risen in public sector had sparked innovation, driven efficiency, and fueled economic development (p. IX). They used case studies to examine Open Data's impact of civic innovation.

Gurin (2014) described the business application of Open Data with cases of companies (p.16). GovLab contributed Open Data 500 Global Network which is an international network of organizations that seek to study the use and impact of Open Data.

Gurin (2014) introduced 5 business archetypes classified by consulting firm Deloitte which worked with Open Data Institution in London. They are:

- Suppliers publish their data as Open Data that can be easily used. While they don't charge for the data-if they did, it wouldn't be Open Data-they increase customer loyalty and enhance their reputations by releasing it.
- Aggregator collect Open Data, analyze it, and charge for their insights or make money from the data in other ways.
- Developers design, build, and sell web-based, tablet, or smart-phone applications" using Open Data as a free resource.
- Enrichers are typically large, established businesses that use Open Data to enhance their existing products and services, for example by using demographic data to understand their customers better.
- Enablers charge companies to make it easier for them to use Open Data (Gurin,2014, p. 13).

Promotion organization of open data&big data utilization and regional creation in Japan (VLED) classified Open Data business archetypes into 3 types. They are:

- Value-added type that companies use open data to strength their existing business and increase value.



Figure 2. The website of Socrata.  
Retrieved July 8, 2019 from  
Source: <https://socrata.com>



Figure 3. The website of Opencorporates.  
Retrieved July 8, 2019 from  
<https://socrata.com>

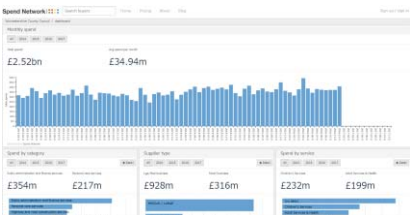


Figure 4. The website of Spend Network.  
Retrieved July 8, 2019 from  
<https://www.spendnetwork.com>

- New value creation type that transform data into a completely new value, such as future prediction by performing advanced analysis.
- Platform type characterized by collecting data from data holders and publishing it on the platform (VLED, 2016, p. 3).

Atsuma(2015) proposed 5 business archetypes. They are:

- Date opening support, such as formatting and cleaning public data before it opened.
- Construction of data portal, hosting and providing technical support.
- Visualization such as data mapping.
- New service utilizing open data.
- API development for accessing data sets.

### The Classification of Open Data Business

Data collected from various sources cannot be directly understood by the audience. It should be cleaned, classified and compiled as dataset, then become information through visualization. Information are tables and figures which convey information to the audience. Some of the companies' business stopped at transmission of information, but some stepped future to utilize information to provide services.

In this paper, we follow the form of data changing and classify representative business archetypes of Open

Data into 7 types based on previous research (Figure 1). The 7 types of business archetypes are:

1. Opening Support: Process and publish data kept by organizations as Open Data.
2. Cataloging: Build a catalog of Open Data by collecting and classifying a large amount of data.
3. Visualization: Support businesses to make Open Data more understandable and accessible through visualization.
4. Management: Manage Open Data on behalf of the people or organizations and create places to provide their data.
5. Analysis: Develop value-added services by analyzing Open Data and providing insights.
6. Forecast: Create new services by using Open Data to create predictive archetypes.
7. Marketing: Obtain know-how by the development of services around Open Data and apply this knowledge (business methods) in other markets.

In the following sections we will introduce 7 types of business archetypes and their representative companies and businesses.

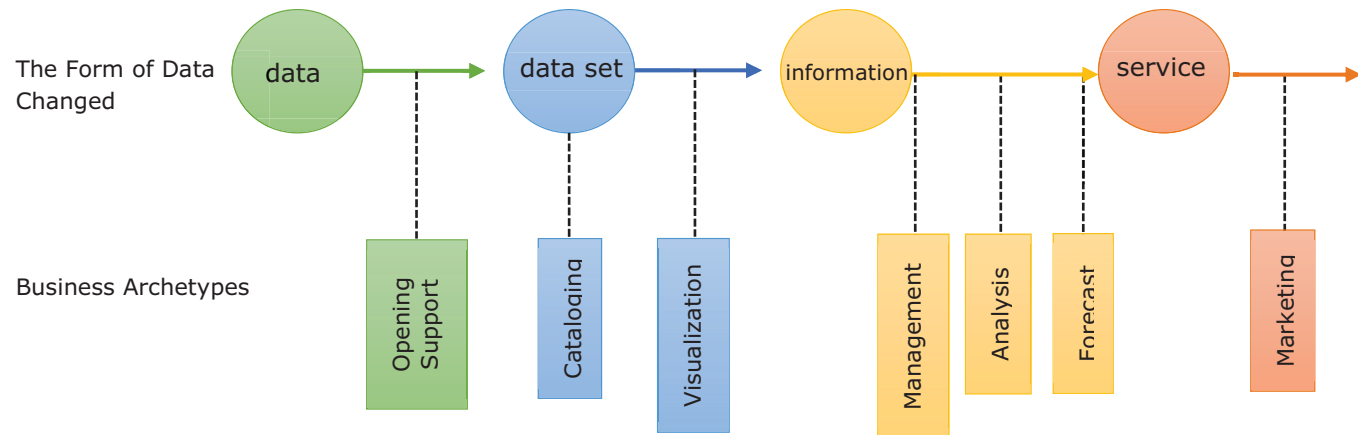


Figure 1. Seven types of Open Data business archetype

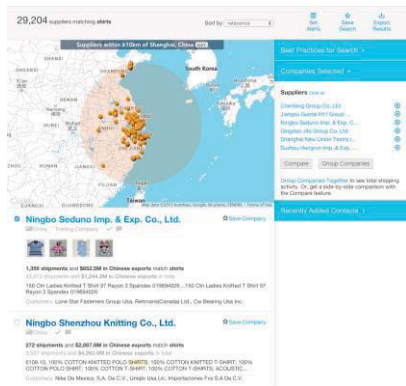


Figure 6. The website of Panjiva.  
Retrieved July 8, 2019 from  
<https://panjiva.com>

### 1. Open Data Support Business

Socrata provides cloud-based solutions allow government organizations to put their data online, make data-driven decisions, operate more efficiently, and share insights with citizens. They provide 4 kinds of solutions and 7 kinds of products(Figure 2). Their solutions include :

- Financial Insights Solution which makes financial transparency and accountability.
- Open Data&Citizen Engagement Solution helps government to make their program move up the digital maturity curve, and engage the participation of citizens, businesses, researchers and analysts.
- Performance Improvement & Accountability Solution empowers government to monitor

progress, adjust activities, evaluate the effectiveness of programs, and make decisions based on internal data metrics.

- Federal Solution includes a central, secure, scalable and APIbased data platform, visualization tools, interactive dashboards, web applications, and open data network.

The products include:

- Capital Projects Explorer which is a dashboard of agency's spending.
- Citizen Connect provides key information about public services and community facilities, alerts, and community-level analysis of incidents.

**CKAN:** CKAN is open source, free software for making Open Data websites. It uses internal arche-type to store metadata about the different records, and presents it on a web interface that al-lows users to browse and search this metadata. It also offers an API that allows third-party participation. Retrieved July 8, 2019 from <https://ckan.org/about/>

**Open Street Map:**

OpenStreetMap is a map of the world, created by people like you and free to use under an open license. Retrieved July 8, 2019 from <https://www.openstreetmap.org>

- Insights Dashboard of budgetary and expenditure data.
- Open Finance which is an applications enable citizens to understand their government’s collecting and spending tax revenues.
- Open Performance helps governments identify specific, measurable goals founded on actual data.
- Perspectives allows governments to create narrative web pages that put data in context, and enhance the data with links to other resources, including additional data, supporting web pages, images, videos, and other types of multimedia.
- Public transforms data from words and numbers into understandable images and publishes on the web.

OpenDataStack provides 4 kinds of open data services for government. They are:

- Open Data Promotion Support. They take inventory of public data from local government, and transform data into open format such as CSV, XML, and Json.
- Open Data Generation Service. They provide tools to upgrade government websites, making date generated automatic.
- Construction and Operation of Data Catalog. In addition to ordinary government website, they also introduce data catalog sites using CKAN,

promoting data opening as local government required.

- Open Data Utilization Support. They provides various of utilization support, such as organizing idea and application contests, management of social communities on facebook, etc.

*2. Open Data Catalog Business*

OpenCorporates is the largest open database of companies and company data, with in excess of 100 million companies. It provides structured data to proprietary data users for charge, but to journalists, NGOs and academics for free. Structured data can be accessed in 3 ways. Firstly, it provides OpenCorporates’ API with queries and results as JSON or XML. Secondly, as for core company data in bulk, it can provide available files such as CSV. Thirdly, it collects, collates and standardizes hundreds of specialized datasets, such as subsidiary data, licenses, and branches, for purchase(Figure 3).

Spend Network gathers and standardizes data from over 300 different public sector bodies, and provides analysis for public sector spending to help suppliers and government to forge lasting partnerships. For suppliers, it provides 5 years of detailed data on categories from key sectors, and data on the individual companies trading with government. It helps suppliers to know better about their competitors and make decisions. For buyers, it provides fully categorized spending data to UNSPSC standard and data of companies serving the public sector. It helps buyers to find the right suppliers, manage contracts and control spending(Figure 4).

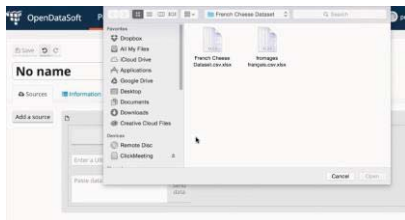


Figure 5. How to upload data onto OpenDataSoft.  
Retrieved July 8, 2019 from <https://www.opendatasoft.com>

Enigma.io links internal datasets with public datasets to provide the big picture. It provides data infrastructure which are flexible tools for data operations, metadata infrastructure tools called Enigma Assembly and Enigma Concourse. Enigma Assembly is a platform for data improvement, discovery, and application development. It makes data and metadata searchable and available for powering applications. Enigma Concourse is a data operations platform that standardizes data ingestion and enables companies to build pipelines that are maintainable and transparent. Enigma Public is a collection of public data from companies, governments, organizations, and universities. Data can be accessed through Enigma public.

### 3. Open Data Visualization Business

OpenDataSoft provides tools and a platform for organizations to launch their open data portal in a quick and cost-effective way. It provides back-end interface for users to collect and process structured data. Datasets can be turned into visualized form, such as tables, graphs, charts, maps, and pictures. It also provides API for business users to share structured data. Developers can interrogate, refine, and sort data in open formats such as CSV, JSON, Excel, GeoJSON, and Shapefile (Figure 5).

Panjiva provides platform of data visualization based on trade data from the U.S., Mexico, Brazil, 9 countries of the Central & South America, China, and India. It visualizes trade relationships, source of new products, latest trends, and the company's basic situation into graph, map, and image to provide clear, actionable insights on global trade for buyers, suppliers, logistics, government agencies, and analysts (Figure 6).

### 4. Open Data Management Business

Agri-note is a tool to manage agriculture logs. It uses data and photography from Google Maps and GPS for farmers to locate their field and create record. The record includes farm working record, growth record such as plant height and number of leaves, quality, standards and harvest record for each field, and shipping record with income and expenditure. It also uses open data from Food and Agricultural Materials Inspection Center (FAMIC) which is a dependent administrative agency, for farmers to check information necessary for agricultural work. All of the data can be outputted and printed in Excel or CSV file for further analysis and verification (Figure 7).

Zaim is an online household account book service for users to manage their daily spending. It provides both web interface and mobile application for users to scan and record their receipts. It also provides Balance Diagnosis to analyze user's spending. Recently, it launched a new service called "My Benefit" which is a service provides information about the benefits, allowances and deductions that users may receive based on users' living area, family composition, and household account book record (Figure 8).

### 5. Open Data Analysis Business

Spikes Cavell Analytic is an agile and innovative data-centric organization. It provides spending analysis and procurement analysis including transforming, improving and adding value to the raw data, and making spend visibility. It also provides dashboard, cloud-based analytical toolset, training and comprehensive support.

Development seed analyzes latest imagery from satellites, sensors, and phones using machine learning



Figure 7. The website of Agri-note.  
Retrieved July 8, 2019 from <http://www.agri-note.jp>



Figure 9. Travel times to education facilities near Vientiane, the capital of Laos. Retrieved July 8, 2019 from <https://developmentseed.org>



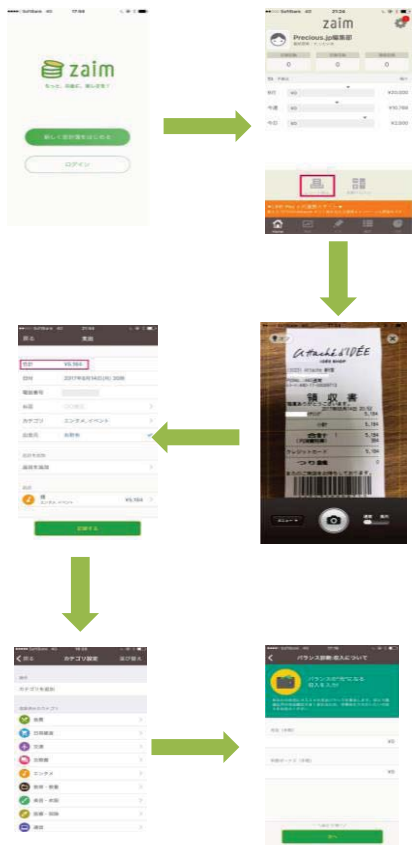


Figure 8. How to transform the data from receipts on Zaim. Retrieved July 8, 2019 from <https://precious.jp/articles/-/2269>

and open data, and works with other organizations, such as the World Bank and NASA. They work with the World Bank developing Rural Accessibility Map (RAM) based on the data from OpenStreetMap (OSM). RAM uses state-of-the-art routing algorithms to archetype the impact of road networks on access to crucial services. They work with NASA’s Earth Science Data System (ESDS) Program to build Cumulus, a cloud-based prototype to ingest, process, catalog, archive, and distribute NASA’s Earth Data streams(Figure 9).

#### 6. Open Data Forecast Business

The Climate Corporation multiplies weather data, soil data, harvest volume data, etc., creates more than 10 trillion detailed simulation points by its unique algorithm, and predicts the harvesting risk of agricultural products. This analysis algorithm has been commercialized as a revenue compensation insurance for farmers, then became an advice service for farmers, and now it is also used as a monitoring service of growing condition (VLED, 2016). It now has built the Climate FieldView platform to help farmers sustainably increase their productivity with digital tools(Figure 10).

GEEO is a service to predict real estate price. It predicts the selling price of real estate using proprietary algorithm based on Open Data such as Route Value<sup>1</sup>, Census, Residential/ Land Statistics Survey, and etc (VLED, 2016). It provides GEEO Pro which includes Account and API for real estate and financial professionals. It also provides GEEO Free which is a

<sup>1</sup> Route Value is the appraisal value per 1 square meter of the residential land facing routes. Retrieved July 8, 2019 from

free service to calculate the price of the real estate by one click on the map(Figure 11).

#### 7. Open Data Marketing Business

Thingful is a search engine for the Internet of Things, providing a geographical index of connected objects around the world. It enable users to find and use connected objects, sensors, datasets and real-time data sources across many popular Internet of Things networks. In the future, it aims to build up an interoperable Internet of Things which connects objects find and use each other's data with the active consent of their owners, and enables IoT data owners to search, organize, access and respond to realtime data, and make them to do more valuable and effective decision(Figure 12).

Transportapi is a digital platform provided a comprehensive open solution for transport data. It has 1500 developers and organizations building on the platform with open data which standardized in JavaScript Object Notation (JSON). It provides services such as processing, visualizing and analyzing telematics and sensor data for developers and organizations, and live departures, timetables, journey planning, tweet mapping, performance indicators, and fares for users.

### Conclusion

#### 1. Problems of Open Data Business Archetypes

When public sector data is proprietary then adjacent business industries must acquire data independently. Issues as small as disparate file formats storing data reduces interoperability and reuse, and the basic act of

<https://kotobank.jp/word/%E8%B7%AF%E7%B7%9A%E4%B%E%A1-154970>



Figure 10. The mobile application of Climate Corporation.  
Retrieved July 8, 2019 from <https://www.climate.com/features/data-visualization>



Figure 11. The website of GEEEO.  
Retrieved July 8, 2019 from <http://otani.co/pdf/GEEEO-geospatial-analysis-service.pdf>



Figure 12. The concept of Thingful.  
Retrieved July 8, 2019 from <https://www.thingful.net>

collecting data requires considerable cost. Conversely, when at least some public sector data is freely available, business archetypes that generate profit by organizing data such as case 1 - Open Data Support Businesses - benefit from a considerably reduced cost of data acquisition. This provides an incentive for the public sector to open further data in areas where third-party analysis can be of benefit. Since there are many jobs contracted from the public sector by this business archetype, it provides an opportunity for stable profit for a certain period. However, this type of business archetype has no fundamental difference from traditional contracted IT business archetypes, and Open Data itself does not create business values per se. It simply reduces the cost of collecting data for analysis.

The business archetypes listed in situations 2 to 6 create value by providing, processing and analyzing Open Data. In these business archetypes human resources with appropriate technical skills and management know-how are necessary to analyze large amounts of data to create such value and there is a commensurate requirement for financial resources to sustain this activity. Therefore, the conditions for monetization around these business archetypes is the technical ability to analyze data, the management ability to refine the applied profit archetype, and the financial strength to support the chosen approach. Write in a straightforward style. Try to avoid long sentences and complex sentence structures. Use semicolons carefully.

## 2. Possibilities of Open Data Business Archetypes

The business possibility of each business archetypes is plotted in Figure 13. The vertical axis shows business potential and the horizontal axis indicates the

ownership of the data (= publicness), either by public or private.

The business archetype 1, raising revenues by Open Data itself, will sustain itself but will gradually decrease in value according to the progress of Open Data conversion in the public sector. Companies with this archetype can extend their service range, such as not only providing Open Data support, but also providing management, analysis and feedback services.

In the business archetype 2, even after creating solutions for cataloging and providing data, companies involved can continue to generate revenue by undertaking maintenance and management of such services, but it is difficult to get more benefit if they only stay in warehousing service instead of extending their business scope.

Most of the business archetype 3 involve trial services or - when paid - have a tier that can be used free of charge. It is necessary to increase the quality of the provided services and convert free service users to paid users to see sustainable growth.

When talking about the business archetype 4, it becomes a competition of economics of scale. The more data Open Data companies manage, the more profit they make. Advertisement revenue is also the main source of the profits. The larger user size, the better advertising effect.

The business archetypes offered by approaches 5 and 6 depend on developing valuable service provision based on data analysis and on creating competitive advantages for others. Technical abilities, management



**UNSPSC:** UNSPSC is the abbreviation of The United Nations Standard Products and Services Code which is an open, global, multi-sector standard for efficient, accurate classification of products and services. Retrieved July 8, 2019 from <https://www.unspsc.org/>

abilities and fiscal strength are the key requirements for success. Archetype 5 targets existing businesses and archetype 6 targets new services. In both scenarios it is possible to build competitive advantages for startup industries and to drive new services. On the two archetypes, archetype 6 appears to offer the greatest potential to become a driving force for the advancement of Open Data businesses in the future.

In the archetype 7, profitability is obtained indirectly, with the businesses involved making use of data and know-how obtained by various Open Data in other business areas or services. If this approach can be linked to increasing profits as a result, it is considered to be a viable high potential business archetype regarding future use of Open Data.

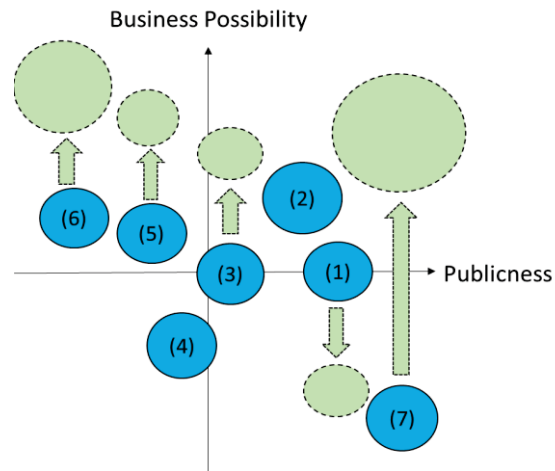


Figure 13. Market Possibilities of Open Data Business Archetypes

## References

1. Atsuma T. 2014. Case Study and Analysis of Open Data Business. Chieba, 119: 28-29.
2. Goldstein B. 2013. Beyond Transparency: Open Data and the Future of Civic Innovation. America Press, San Francisco.
3. Gurin J. 2014. Open Data Now. McGraw-Hill Education.
4. Omidyar Network. 2014. Open for Business: How Open Data Can Help Achieve the G20 Growth Target. Retrieved July 8, 2019 from [http://www.omidyar.com/sites/default/files/file\\_archive/insights/ON%20Report\\_061114\\_FNL.pdf](http://www.omidyar.com/sites/default/files/file_archive/insights/ON%20Report_061114_FNL.pdf).
5. Osterwalder A., Pigneur Y., Tucci C. L. 2005. Clarifying Business Models: Origins, Present, and Future of the Concept. Communications of the Association for Information Systems, 15: 1-15.
6. Ritter T. and Lettl C. 2017. The wider implications of business-archetype research. Long Range Planning. 51, 1: 1-8.
7. The European Commission Vickery Research Project. 2011. Review of recent studies on PSI reuse and related market developments. Retrieved July 8, 2019 from [http://ec.europa.eu/information\\_society/newsroom/cf//document.cfm?doc\\_id=1093](http://ec.europa.eu/information_society/newsroom/cf//document.cfm?doc_id=1093).
8. VLED (promotion organization of open data & big data utilization and regional creation). 2016. Business Case Collection of Open Data Utilization. Retrieved July 8, 2019 from [http://www.vled.or.jp/results/opendata\\_business\\_usecases.pdf](http://www.vled.or.jp/results/opendata_business_usecases.pdf)