The Open Data Canvas-Analyzing Value Creation from Open Data

Gao, Yingying; Janssen, Marijn

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

The Open Data Canvas–Analyzing Value Creation from Open Data

YINGYING GAO, College of Public Administration, Huazhong University of Science and Technology
MARIJN JANSSEN, Faculty of Technology, Policy & Management, Delft University of Technology

Expectations to derive value from open data are high. However, how value is created from open data is still largely unknown. Open data value is usually generated in constellation of actors in which each player has different capabilities and roles. To understand the open data value creation process, the business model canvas is introduced in this article. The typical components of the business model canvas and open data value creation are derived from the literature. By combining these two research streams, the open data value model canvas is created. The case of Coronavirus disease 2019 (COVID-19) worldwide dashboard developed by the Johns Hopkins University is used to evaluate the model's utility. Key components of the open data value model are creating an overview of various data sources from public and private organizations, having capabilities to combine heterogeneous data, and connecting data and needs. In this way, the open data canvas helps to grasp the value creation logic.

CCS Concepts: • **Applied computing → E-government**;

Additional Key Words and Phrases: Open data, value creation, business model canvas, design science research, data infomediaries, open data canvas

1 INTRODUCTION

Public agencies are the largest creators and collectors of data in many different domains such as traffic, healthcare, economics, and so on [32]. Potentially, a huge amount of value can be derived from these data by exploiting data mining techniques and analytics [19, 23]. For example, during COVID-19 epidemic, many countries' governments published the number of positive infections and deaths every day and continuously adjusted the quarantine policies based on the open data [27], resulting in public values like improving government transparency and public trust. Not only the public health agencies but also scientists can contribute to the anti-coronavirus actions by analyzing these open data. Some of them conducted research about the spread of virus [63], providing insights into the current and future situation to the citizens [58].

Open data refers to publishing data in open and reusable formats without restriction or monetary charge for their use by society [6]. These data can come from scientific researches, the public sector, and so on [9]. However, in the scope of this research we will focus on **open government data (OGD)**. The OGD movement aims to transform the government into a transparent administration and stimulate citizen participation and engagement [50]. Other types of value are also expected from open data, including innovating and creating new businesses based on it [2, 48] and improving public services [33, 44]. Despite this, understanding how value is created from open government data remains largely unknown, particularly how public value is generated [37, 75].

Value creation from open data is a complex process that is created in constellations of actors in which each player has different capabilities and roles [68, 75]. Therefore, an ecosystem metaphor can be taken to model the interactions among these different elements [57]. There are several types of ecosystems in the private sector, like business ecosystem, innovation ecosystem, and platform ecosystem [31]. Reference [1] reviewed these researches and defined ecosystem as "a co-evolutionary business system of actors, technologies, and institutions" (p. 24). Recently, the ecosystem metaphor has occurred in the open data area [16, 28]. An ecosystem is comprised of interacting components, whereas specific components vary from ecosystem to ecosystem [61].

The business model is a tool to clarify the components in the value creation ecosystem to understand the rationale of how organizations create, deliver, and capture value [54]. Each enterprise explicitly or implicitly has a business model [13, 65]. Many researchers outline those business models and their components [53, 60]. Among all the business model tools, the business model canvas is the most popular method for analyzing the value creation by companies [10], as it is easy to understand and contains the main components and their relationships. Therefore, we will use the business model canvas as a starting point in this article. However, although the business models can be used to discuss the issues of information systems in public sectors, there is still some difference in the logic of value creation between the enterprises and government.

As such, the aim of this article is to develop the open data value creation model to answer the question *How is value created from open data*? This model will be labelled the "*open data canvas*" to avoid the association with the scope of business. This work will focus not only on public values but also on the value for private organizations will also be taken into account as they are the important motivation to support the data re-use. We followed the **Design Science Research (DSR)** method to design the model. The open data canvas was built based on a systematic literature review. Then the case of Johns Hopkins University COVID-19 dashboard was used to evaluate the model's utility. The article is structured as follows. In the next section, the literature review about business models is presented, followed by the research approach. Thereafter, the open data canvas is derived from the literature and evaluated by the case. Finally, conclusions and recommendations are drawn.

2 RESEARCH METHOD

The DSR is a fundamentally problem-solving paradigm in Information Systemthat refers to created new artifacts with existing knowledge to address problems and improve the performance of technologies or organizations [25]. These artifacts could be "construct," "models," "method," and "instantiation" [49]. The DSR method has been adopted in a wide range of research settings including decision support systems design [4, 52] and business model innovation [2, 53]. This research aims to build the open data canvas based on the business model canvas to understand the value creation logic of open data. So, it is suitable to follow the design science approach to conduct the research. A common DSR framework developed by the authors of Reference [30] was adapted for our research, as shown in Figure 1.

The relevance cycle connects the environment of the research with the design research activities. Motivated by the desire to solve the research question, this work built the open data value model and evaluated it, which is the design cycle and the heart of research activities [29]. The rigor cycle connects the design activities with the knowledge base, including the scientific theories and methods that provide research foundations [29]. In this
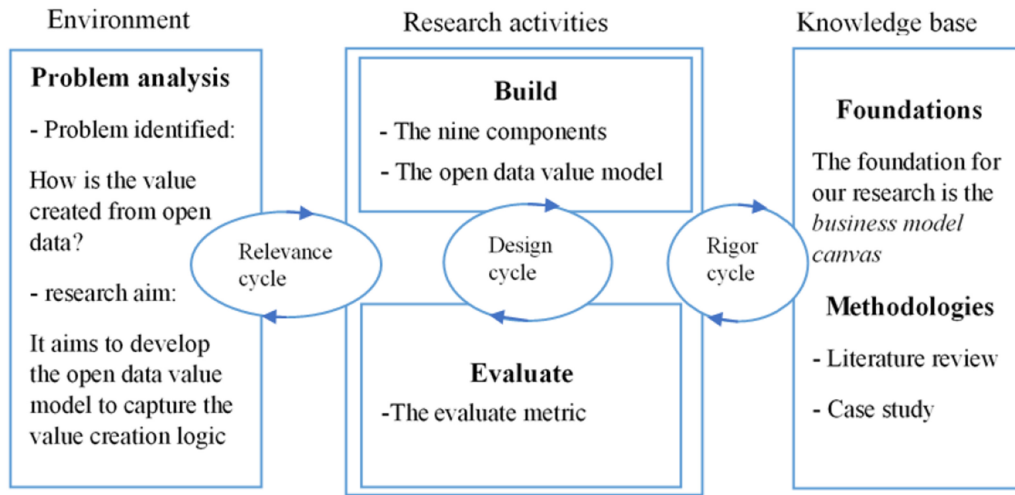
Fig. 1. The design framework adapted from Reference [30].

Table 1. Research Activities and Methods

Main DSR activities	Specific steps	Section structured
1. Problems and aims	Give an introduction to the problem and research aim	Introduction
2. Open data canvas model design	Review business model canvas literature	Literature review
	Reviewed open data value literature and looked for typical elements needed	and the open data canvas model
	Categorized the elements in an open data canvas.	
3. Demonstration and Evaluate	Demonstrate that the open data canvas model could be used to understand the case – The COVID-19 worldwide dashboard developed by Johns Hopkins University	Case study
4. Contribution	Discuss about the contribution to the knowledge and practice	Discussion

work, we used the business model canvas as the foundational framework. The literature review and the case study method are separated and applied to build and evaluate the model. Table 1 presents the specific method used in this work.

The first step was to build the model. First, a literature review was conducted into the business model canvas. Recognizing the key elements of business models provided the basis to understand the open data value model. Second, the open data value literature was systematically reviewed to identify the typical elements relevant for the open data field. Third, the elements derived from the business model canvas were adapted to the open data value creation situation. We categorized the elements in the open data model canvas and defined them.

The second step is to design the open data canvas model. After analyzing the business model canvas and its application, a systematic literature review was used to survey the typical elements of open data use. To cover the largest range of relevant publications possible, the databases ACM Digital Library, Science Direct, and Web of Science were searched. According to our question, the search terms are "open data," "value," and "benefits." We conducted the search on the title and abstract fields. Then, we performed a manual selection, retaining only those results relevant to the research question. Figure 2 presents our process for identifying high relevant research. After investigating each data source, we merged all the results and removed the duplicate literature.
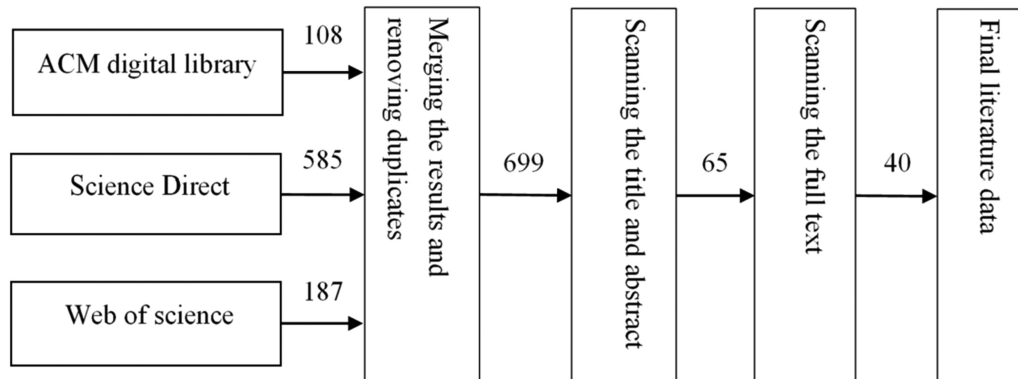
Fig. 2. Procedure for identifying the relevant literature.

This resulted in 699 publications. The following step was to scan the title and abstracts manually. In this step, most of the literature was excluded, because they were not related to the open data, the linked open data, or open government data fields and do not cover open data value creation. Finally, we went through the full text of the studies. This resulted in a total of 40 studies, and the overview can be found in Appendix. These papers discuss the value creation ecosystem from different perspectives, including stakeholders, the value creation process, and practices. They help to clarify the elements of the open data canvas. Finally, the elements were categorized, and the open data canvas model was developed.

The third step is to demonstrate and evaluate the model. The case study method is useful to test the hypotheses [30] and very appropriate to answer the "how"- and "why"-type questions [39]. Case studies can deal with complex situations having many actors [8]. Even generalization is possible based on a single case [22]. Thereafter, the case study approach was applied to evaluate the utility of the model. The criteria for selecting cases are as follows:

(1) The cases should have many data sources that can be used for more than one type of decision.
(2) Different kinds of stakeholders should be involved.
(3) Societal values should be created.

Given the relevance and availability, the COVID-19 worldwide dashboard developed by Johns Hopkins University was investigated. The fourth step is to give a discussion of contribution to the knowledge and practice.

3 LITERATURE REVIEW AND OPEN DATA CANVAS MODEL

3.1 The Business Models and Public Sector

The concept of **business models (BM)** originates from the field of e-commerce. Reference [67] was the first to classify e-commerce business models. The author defined a business model as an architecture of the information, product, and financial flows in which gives a description of the various business actors and their roles, the potential benefits for business actors, and the sources of revenues [67]. A number of researchers defined the concept [14, 73] and discussed the relationship between the strategy and business models [12, 72]. Reference [65] addressed the dynamics in business models to keep competitiveness.

The application of the business model concept has evolved over the years. Researchers focused on classifying business models for electronic markets [56, 74] and on identifying the e-business models components. For example, Reference [60] based their business model on four components, e.g., strategic choices, value network, value creation, and value capture. Reference [3] put forward a synthesized business model framework including

Table 2. The Elements Summary in the Existing Literature

Main Elements	Specific contents	Literature source
Stakeholders	Data Providers Data Infomediary Technology support The end -users and public beneficiaries	4, 5,10, 28, 12, 13, 27, 14, 33, 36, 17, 22, 15, 18
Barriers or resources	Need Organization or Policy support Finances Legal risks such as privacy protection ICT infrastructure	29, 21, 24, 30
Key activities	Data discovery and collection Data integration and cleaning Data processing and analysis Application design development and deployment Visualization or application	4, 36, 9, 12, 22, 38, 39, 40, 13, 15, 25, 35
Benefits and values	Revenue models Open data value model The Open data Hackings	1, 3, 7, 6, 15, 16, 36, 37, 26, 19, 21, 32, 33, 34

value proposition, value architecture, value finance, and value network. By consolidating, research in this area [54] **shaped business model canvas (BMC)** based on nine components.

Although the business models are often associated with revenue and profit, they can also be aimed at generating public value. As opposed to private value, public value refers to the value or benefits for the society and public created by the government through services, laws, regulations, and other actions [41]. The usefulness of the BM concept in e-government research has also been recognized in a number of studies. Reference [34] surveyed e-government websites and derived eight e-government business models. Drawing on the business models proposed by Osterwalder et al., Reference [20] revealed four typologies for transforming public administration using open innovation.

The business models were broadly applied in government to analyze the value creation logic by public services, which paved the way for its use in specific contexts—open data use. Among all the business models, the BMC is a popular visualization tool to provide an overview at a glance to make it easy to understand. It also enables us to describe the open data use ecosystems details.

3.2 Open Data and Value Creation

In total, 25 conference papers were found in the systematic literature review results, which shows that open data value creation research is still in infancy. After reading these papers, the elements referring to the existing literature are summarized in Table 2. The literature number is linked to the Appendix.

Stakeholders. There are multiple stakeholders involved in open data use and value creation. Like the open data ecosystems designed in Reference [16], the data providers and the end-users are the key stakeholders [5, 42]. The data providers can be public sectors and non-government sectors [40]. The role of data infomediaries is also important. They connect the data suppliers and the end-uses and aim to improve and enrich the open data value [51]. There are different types of infomediary, like data aggregators, data developers, and so on [66]. The infomediary are the direct data users and could provide enriched data and services to the public [46]. However, the data users also need technology support from other actors [43]. Unlike the previous business model in which the organizations mainly chase profits, these stakeholders have different value propositions, either for economic value or for the public interest [71].

Barriers and Resources. The value of open data can only be created by its use. However, there are barriers or costs associated with both data open and use [33]. These barriers or costs range from technology support [59] to policy support [69] to legal risks [59, 64]. Existing literature classified those challenges, but for the specific users, the classification is still too broad.

Key activities. Research classified key activities from different perspectives like data lifecycle [5, 6, 45], the specific projects [43], and value creation [70]. Overall, for the direct data users, before data processing they need to collect, aggregate, and clean the data for their purposes. Then, enabled by the data analytic technologies, the knowledge and enriched data or services are produced for the final users [15]. Sometimes the visualization process is adopted to make the results easy to understand.

Benefits and values. Some studies focused on the open data value creation mechanism. For instance, Reference [47] illustrated the value creation process of the Chinese open data hackathons. Reference [37] summarized two kinds of value generation mechanisms: the information-sharing mechanism and the market mechanism. The former refers to delivering useful information to society, and the latter means offering products or services to the market [11, 37]. Reference [64] analyzed three kinds of open weather government data value creation mechanisms, including decision making, new offerings, and transparency. However, these discussions did not open the value creation box, and we still do not know how to gain benefits from open data. To answer this question, Reference [21] investigated the profit-oriented actors involved in the value generation process and summarized the revenue model for each actor. This model is a revenue model or pricing model, which discussed how to make money for the companies [35]. References [43, 71] applied the business model canvas to analyze the innovation based on open data. But they did not adopt existing models to create an open data model canvas focused on the value creation of open data in ecosystems. Janssen analyzed 12 open data application cases and summarized six infomediary business models as Single-purpose apps, Interactive apps, Information aggregators, Comparison models, Open data repositories, and Service platforms. These models are taxonomies, and they do not give a further analysis of the data infomediaries specific activities. Reference [2] developed a 6-V open data business conceptual framework. But this model was not tested with any cases, making it hard to reuse it in other open data contexts.

The existing literature discussed the value creation process stakeholders, the barriers or costs, and the key activities, providing useful experience for developing an open data canvas model. Some research applied the business model thinking in the open data context. But they are broad and do not focus on the specific value creation process. Furthermore, these models do not focus on costs analysis, required resources, and other key elements. Hence, the focus of our research is to develop an open data canvas to give an overview of all the key elements of the open data creation process.

3.3 The Open Data Model Canvas

The business model canvas provides us with understanding of the open data value generation process from the direct data user's lens. Existing researches discussed some typical elements of open data value creation, such as the stakeholder types, data use type, data processing activities, profit models. To comprehensively understand the key elements of value creation, the following changes were made to the business model canvas to construct the open data value model canvas, as shown in Figure 3:

(1) The *customer relationship* component was removed. Unlike each exchange of coffee, cola, or any other items that could generate profits, there is no payment needed for the delivery of open data and information, and the delivery might often be a one-time exercise. Therefore, maintaining customer relationships is not the main task, although for policy-makers, stimulating open data usage is important [24].

(2) The *data providers* item was added. Data are the core material to support the operation of a business model. So, it is necessary to clarify where to get the data. *Data providers* provide datasets in different formats, which are collected in various manners.
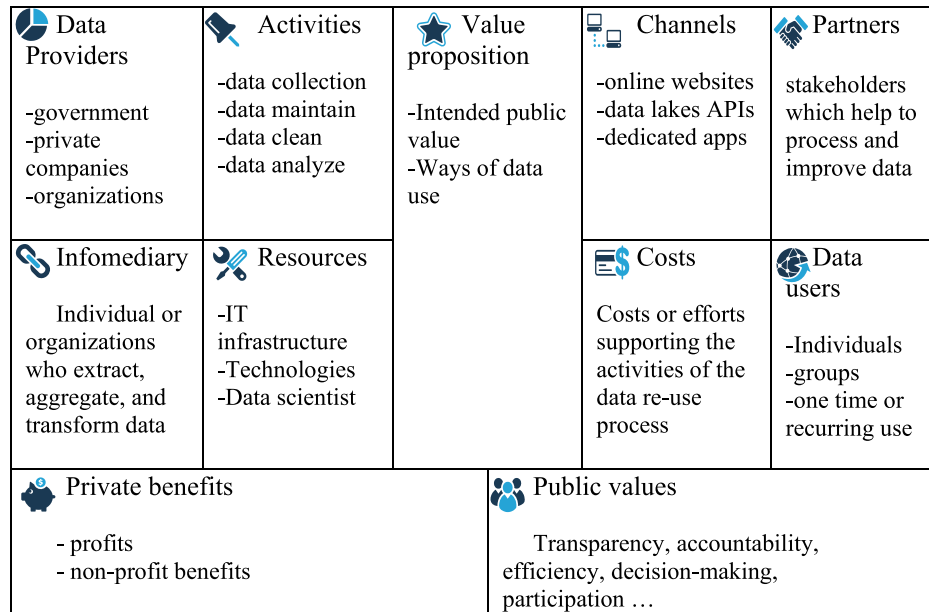
🔵 Data Providers -government -private companies -organizations	🔧 Activities -data collection -data maintain -data clean -data analyze	⭐ Value proposition -Intended public value -Ways of data use	💻 Channels -online websites -data lakes APIs -dedicated apps	🤝 Partners stakeholders which help to process and improve data
🔗 Infomediary Individual or organizations who extract, aggregate, and transform data	🛠️ Resources -IT infrastructure -Technologies -Data scientist		💲 Costs Costs or efforts supporting the activities of the data re-use process	🌐 Data users -Individuals -groups -one time or recurring use
🐷 Private benefits - profits - non-profit benefits		👥 Public values Transparency, accountability, efficiency, decision-making, participation …		

Fig. 3. The open data value canvas.

(3) The *infomediary* components were included. Infomediaries are a type of stakeholders focusing on com-mercial and social information collection and distribution [17]. While in the open data domain, the *in-fomediary* can be defined as an intermediary who extracts, aggregates, and transforms data, making it more beneficial and important than raw data [38] for the end-users. Often, they hide the complexity of end-users and develop apps to enable the creation of value from data.

(4) The *customer* component was changed to *users*. Users is the preferred term, as no products are sold to customers; instead, data are made available for free to users. There are several stakeholders in the open data situation that can be classified into three main types: data provider, infomediary, and the users [35]. So, *user*-item is used here to refer to groups of people or entities the opening of data aims to serve.

(5) The *revenue stream* was split into "*private value*" and "*public value.*" Although the aim of data opening is to obtain public value like transparency and participation, infomediaries must have sufficient benefits as motivation. The *private value* refers to the benefits that the infomediary could generate for itself. These benefits could be for-profit or non-profit advantages, like reputation, attention, and so on. In contrast, *public value* refers to the societal benefits for the public, like transparency, accountability, and improved decision-making.

(6) The *value proposition* explains how to solve the target user's needs and obtain profits. In the context of open data, the data re-users should also pay attention to users' pursuit of public value.

(7) The *channel* clarifies how to distribute the products or services to the target group. In the open data use situation, these channels are almost online, including websites, portals, dedicated apps, and so on.

(8) The *resources* refer to the assets needed to realize value, like human capacity, IT infrastructure, software, and money. To deliver useful information or other services, data scientists with knowledge and technolo-gies are important. An IT infrastructure is also necessary to support open data use, because the whole process from data input to service output is online.

(9) The *activities* clarify the important thing that a team to process data to create value. In the commercial area, the activities include marketing, production, logistics. In the open data context, activities mainly focus on data integrating, processing, enriching, and visualization.

(10) *Cost* defines what finance or efforts are needed to re-use data.
(11) *Partners* refer to the different types of stakeholders that help process and improve data, such as technology support partners, infrastructure partners, and so on.

4 EVALUATING THE OPEN DATA CANVAS: COVID-19 DASHBOARD BY JOHNS HOPKINS UNIVERSITY

After having developed the model, we evaluated the model's utility by using it in practice. From December 2019, COVID-19 became a worldwide pandemic. In the open data domain, the government was encouraged to open as much data as possible to the public. This data can be used for a variety of purposes, including governments for determining their crisis response, hospitals for understanding the effects of medicines, and citizens to understand the outreach and consequences of measures to avoid its effect. The Johns Hopkins University developed a COVID-19 dashboard that collects data from all over the world and is used by many and diverse stakeholders. The university team collected the data about infection, recovery, and death about every country and area and combined them into a map to give a clear picture of the evolvement of the pandemic.

4.1 Background of COVID-19 Dashboard

The COVID-19 dashboard was initially developed by the Lauren Gardner research team from civil and systems engineering at Johns Hopkins University. Their motivation is to *"let the public have an understanding of the situation as it unfolds, with transparent data sources"*[1] and provide a useful tool for the researchers, health departments, and the public to track the epidemic situation [18].

During January 22–31, since most infection cases were identified in China, and there were only a few cases in other countries, the research team collected and processed the data manually and updated them twice a day. But as the epidemic evolved across the world, it is unsustainable to report cases manually. Therefore, after February 1, on the one hand, the research team collected the data from aggregated websites like Dingxiangyuan (DXY) and **World Health Organization (WHO)**. On the other hand, they developed code to automatically monitor the related data releasing website [18].

4.2 The Elements of Open Data Value Generation through the COVID-19 Dashboard

In the following section, the open data canvas was used to analyze the COVID-19 dashboard developing cases to understand how public value is generated. The overall analysis of open data value generation through the COVID-19 dashboard is presented in Table 3.

4.2.1 The Value Propositions. Almost all countries published data about the number of infected people, cured patients, hospitalized patients, and so on. One can easily know the situation in their living area and country, but it takes much effort to compare among different countries to understand the trend of coronavirus and the pandemic severity of the living area. As such, the purpose of this dashboard is to let *the general public quickly understand the situation and provide transparency*. In addition, the research team also aims to provide a user-friendly tool for researchers who can *innovate* with the data underlying the dashboard for new knowledge and the public health authorities who could use the data as the *basis for policy-making*.

To generate the intended value, on the one hand, the research team provided the *information service* that is creating a real-time world map of the pandemic and interpretation of one or more data sources. On the other hand, the research team provided the *data service,* and they opened up all the data they aggregated on GitHub in a CSV format.[2]

---

[1]From the interview https://www.pbs.org/newshour/health/track-the-spread-of-novel-coronavirus-with-this-map.
[2]The open source on the GitHub: https://github.com/CSSEGISandData/COVID-19.

Table 3. An Analysis of the COVID-19 Dashboard through the Open Data Value Canvas

Data Providers	Activities	Value propositions	Channels	Partners
Worldwide national or local health departments and government	The main activities are data collection and data visualization.	– let the general public quickly understand the situation; – provide a tool for researchers to innovate with the data for new knowledge; – provide the basis for the policy-makers.	– their own website for the map – GitHub and ArcGIS platform for the downloadable database – social media for interpretation of the data	– Amazon company-Slack company – GitHub platform-the Esri Living Atlas team – Applied Physics Laboratory of the Johns Hopkins University
Infomediaries	**Resources**		**Costs**	**Data users**
– Inter-government organization – The aggregated websites – The social news	– IT infrastructure: hardware: could computing server software: Slack and GitHub – Technology support: Esri Living Atlas team, Applied Physics Laboratory in Johns Hopkins University		Financial support from JHU, National Science Foundation, National Institute of Allergy and Infection Diseases and so on to pay for the required resources	– Citizens – Researchers – The policy-makers – The media
Private benefits – Presented the strong scientific capacities – Reputation increased			**Public values** Transparency, accountability, innovation, participation	

4.2.2 Data Providers. The *national or state government and health departments* are the owners of the epidemic data and therefore are the largest data provider. The health department combined the data from death information systems, positive infection testing system, and so on to submit the final data. There are differences in measurement between countries. Some countries count asymptomatic infections as positive cases, while others do not. Moreover, due to the difference in detection capabilities, the accuracy of the reported data is also different. However, in this case, the research team did not respond to these differences but truthfully aggregated the data from various countries.

4.2.3 Infomediaries. Collecting all the countries' coronavirus epidemic data is challenging for a university team. So infomediarie are particularly important. In this case, there are three types of data sources. The first is *inter-government organizations* like the WHO. The second is non-government *aggregating websites* such as DXY, 1point3acres, and Worldometers.info, which rely on a combination of reporting data from local health departments. The last one is the news, like BNO News and The Mercury News, which is the main source to know new infection cases in the early stages of the epidemic when the government has not yet opened the data regularly. Some data sources are no longer relied on now.

4.2.4 Activities. For the research team, there are two main activities, e.g., data collection and data visualization. First, they need to collect the data from the health departments or the aggregate websites around the world as accurately and comprehensively as possible. Second, they performed visualization analysis to give a clear map of coronavirus. Ensuring the accuracy and timeliness of the data collection is a huge challenge for the data re-users.

4.2.5 Resources. The research team listed two main types of resources required to develop the coronavirus map. The first one is the IT infrastructure resources, including the hardware facilities such as cloud computing servers and the software that allows team members to communicate and work collaboratively. And the second

one is the data scientists, who can provide technology support on data collection, systems engineering, and data analysis.

4.2.6 Channels. The channels describe the way to distribute the information or the service to the users. In this case, the coronavirus resource center provided the downloadable database through the GitHub and ArcGIS platforms. In addition, they provide a worldwide map through the website developed by themselves. At the same time, the expertise from medical, data analytics, and other disciplines of Johns Hopkins University gave their professional interpretation of the data and distributed the results through social media platforms.

4.2.7 Cost. According to the information the research team published, they received financial support from Johns Hopkins University, National Science Foundation, National Institute of Allergy and Infection Diseases, and National Aeronautics and Space Administration. This financial support paid for the required IT resources and technology support.

4.2.8 Users. Although the dashboard website does not provide the access numbers, there are many discussions quoting the data from the dashboard on (social) media. Many publishers, like *The Washington Post* and NBC 15, usually reported the data and the latest policies to allow their readers to evaluate the situation. The other important group of users are researchers. Some researchers used the integrated data to make predictions [36]. Some scholars have combined these data with telecom data to verify the effectiveness of the social distancing measure [7]. For the health authorities, there is no formal evidence that this dashboard directly affects policy-making.

4.2.9 Partners. First, Amazon, Slack, and GitHub provided the research team with the required IT resources. Second, except for the research team members, data scientists from the Esri Living Atlas team and **Applied Physics Laboratory (APL)** of the Johns Hopkins University also provide technology support. The data gurus from APL created a code that periodically travels to trusted websites and scrapes for data, which greatly improved data collection efficiency.[3]

4.2.10 Private Values. The main private value for the Hopkins University is increasing their reputation and becoming well known to the public. The Hopkins University has demonstrated its strong scientific capabilities, including the ability to rapidly processing data and generate scientific knowledge. The website and the university have become well known to the public through news media and social media, which has increased the popularity of the university.

4.2.11 Public Values. The public values describe the benefits this project delivered to the whole society. Overall, the data help to create a better understanding of the pandemic and helps to deal with the pandemic better. First, the general public can clearly see trends and make an intuitive comparison of different countries, which could improve the public perception of government transparency. Second, researchers use data to construct new models predicting the trend of the outbreak, analyse the effectiveness of the measures, judge the pressure of the medical system, and so on. They try to use their expertise to participate or impact policy-making. In addition, in the media industry, the news usually cites data and reports the current measures, trying to give pressure and accountability to the government.

For the case analysis, the open data canvas was found to be useful. Describing the components increased the understanding of the open data value creation process. In comparison with the previous business models, it included the components who both the private value while emphasizing open data public value at the same time. This case analysis shows the many data sources used and that infomediaries play an important role in processing the data. Infomediaries enriched the data by integrating the numbers released by government departments, which greatly improves the data collection and processing efficiency of Hopkins University. Ensuring the accuracy and timeliness of data is a challenge for data re-use. An appropriate IT infrastructure is needed for data

---

[3]Information comes from https://edition.cnn.com/2020/07/11/health/johns-hopkins-covid-19-map-team-wellness-trnd/index.html.

re-use. Furthermore, the open data canvas shows that technical support for the data storage, calculation, processing from other partners is also necessary and needed. All components together increase the understanding of the value creation logic.

5 DISCUSSION

5.1 The Contribution to the Knowledge

Some components of the business model canvas like profit revenue, customers, and so on were not found to be fully suitable, as the open data value creation essential elements are missing. The business model canvas was extended and modified for the open data research domain. The main components of the value creation logic for open data have been identified. The role of infomediaries has been included and emphasized (e.g., References [35, 51]), and the role of data infomediary is proved to be essential to understand the value creation logic as found in the case study. The whole model focuses on public value instead of private values as these are motivation of open data for the public sector [33, 55]. The open data canvas model enables us to more easily understand the required resources and the value creation process from the user's perspective. The model focussed on the essential aspects, is easy to use, and can be used within a short time frame.

This model provides a concise framework to analyse other types of open data use cases. For instance, to analyse the data to service cases and compare the differences between them. In our recent research, we used this model to analyse the value creation logic of the Dingxiangyuan COVID-19 dashboard developed by a company. Future research can pay attention to the open data value creation's relationship to the type and abilities of the data users.

5.2 The Contribution to the Practice

The open data canvas can help the data users to understand the need for the technology and infrastructure and learn the possible challenges they may encounter. Open data providers often neglect the user side [62] and pay little attention to the value creation process of users [33]. resulting in a low usage and uptake of open data [26]. For public sector practitioners, this is a useful tool to capture and better understand the entire data utilization process, so as to know the user demand of the data type and quality. This should result in higher utilisation of open data.

Furthermore, the open data canvas draws the attention to the roles of other parties. The open data canvas emphasizes the role of infomediary. In our case, the role of infomediaries was found to be pivotal for open data use. This open data canvas helps to clarify such an information chain. The canvas can be used to identify the need for new components to improve public value creation, like integrating data targeting for the end-users needs.

5.3 Future Research

In further research, the model can be tested and refined. Insights from practices might result in the refinement of some of the components. An example of this is that the required resources or the activities might be different for other types of use of open data, such as for the downloading of data in batches or real-time use for creating visual dashboards. Finally, whereas costs are quantifiable, public values are not. Hence, further research is recommended on measuring public values and quantifying the income-output ratio.

6 CONCLUSION

Researchers and practitioners are struggling to understand how value is created from open data. In this research, the open data canvas was developed to give a clear picture of the key elements of value creation, including data providers, infomediaries, value proposition, activities, resources, channels, cost, partners, users, private value, and public value. There are different stakeholders who play a role in value creation, such as the data

providers, infomediaries, technology partners, and users. In general, the government might be providing data sources, but in our case study, the infomediaries are crucial by integrating and processing open government data from different sources. Infomediaries are a bridge to link the government and the data users. The open data canvas draws the need to analyse the activities and required resources. The main activities focus on data processing activities, including data collection, combination, processing, and so on, which require IT resources and technology support. The open data canvas suggests to analyse the cost, private benefits, and public values. Although the value proposition is to achieve public value, the infomediary also needs to gain private benefits, which can be non-monetary, such as popularity and reputation. The cost and benefits are often dissimilar for organizations. Although one organization might need to make investments, other organizations might have the gains. In our case, the journalism community, the public, and the researchers get benefits from the investment made by the Johns Hopkins University.

One limitation of this work is that we conducted only one case study to evaluate the open data canvas. Future research is needed by using the data canvas in a large number of case studies to seek general patterns of value creation by combining the elements of the open data canvas. In turn, this can be used to analyse the different types of value.

REFERENCES

[1] Leena Aarikka-Stenroos and Paavo Ritala. 2017. Network management in the era of ecosystems: Systematic review and management framework. *Industr. Market. Manage.* 67 (2017), 23–36. DOI : https://doi.org/10.1016/J.INDMARMAN.2017.08.010

[2] Fatemeh Ahmadi Zeleti, Adegboyega Ojo, and Edward Curry. 2016. Exploring the economic value of open government data. *Govern. Inf. Quart.* 33, 3 (2016), 535–551. DOI : https://doi.org/10.1016/j.giq.2016.01.008

[3] Mutaz M. Al-Debei and David Avison. 2017. Developing a unified framework of the business model concept. *Eur. J. Inform. Syst.* 19, 3 (2017), 359–376. DOI : https://doi.org/10.1057/EJIS.2010.21

[4] David Arnott. 2006. Cognitive biases and decision support systems development: A design science approach. *Inf. Syst. J.* 16, 1 (2006), 55–78. DOI : https://doi.org/10.1111/j.1365-2575.2006.00208.x

[5] Judie Attard, Fabrizio Orlandi, and Sören Auer. 2017. Exploiting the value of data through data value networks. In *Proceedings of the 10th International Conference on Theory and Practice of Electronic Governance* (2017). DOI : https://doi.org/10.1145/3047273

[6] Judie Attard, Fabrizio Orlandi, Simon Scerri, and Sören Auer. 2015. A systematic review of open government data initiatives. *Govern. Inf. Quart.* 32, 4 (2015), 399–418. DOI : https://doi.org/10.1016/j.giq.2015.07.006

[7] Hamada S. Badr, Hongru Du, Maximilian Marshall, Ensheng Dong, Marietta M. Squire, and Lauren M. Gardner. 2020. Association between mobility patterns and COVID-19 transmission in the USA: A mathematical modelling study. *Lancet Infect. Dis.* 20, 11 (2020), 1247–1254. DOI : https://doi.org/10.1016/S1473-3099(20)30553-3

[8] Pamela Baxter and Susan Jack. 2008. Qualitative case study methodology: Study design and implementation for novice researchers. *Qualit. Rep.* 12, 4 (2008), 544–559. DOI : https://doi.org/10.1039/c6dt02264b

[9] Carla M. Bonina. 2013. New business models and the value of open data : Definitions, challenges and opportunities. In *NEMODE–3K Small Grants Call.*

[10] H. Bouwman, E. Faber, T. Haaker, B. Kijl, and M. de Reuver. 2008. Conceptualizing the STOF model. In *Mobile Service Innovation and Business Models.* Springer, Berlin, 31–70. DOI : https://doi.org/10.1007/978-3-540-79238-3_2

[11] Sharon Buteau, Preethi Rao, Anshuman Kr Mehta, and Vigneshraja Kadirvell. 2018. Developing a framework to assess socioeconomic value of open data in India. In *Proceedings of the 14th International Symposium on Open Collaboration (OpenSym'18).* Association for Computing Machinery, Inc, New York, NY, USA, 1–6. DOI : https://doi.org/10.1145/3233391.3233532

[12] Ramon Casadesus-Masanell and Joan Enric Ricart. 2010. From strategy to business models and onto tactics. *Long Range Plan.* 43, 2–3 (2010), 195–215. DOI : https://doi.org/10.1016/J.LRP.2010.01.004

[13] Henry Chesbrough. 2010. Business model innovation: opportunities and barriers. *Long Range Plan.* 43, 2–3 (2010), 354–363. DOI : https://doi.org/10.1016/J.LRP.2009.07.010

[14] Henry Chesbrough and Richard S. Rosenbloom. 2002. The role of the business model in capturing value from innovation: Evidence from xerox corporation's technology spin-off companies. *Industr. Corp. Change* 11, 3 (2002), 529–555. DOI : https://doi.org/10.1093/ICC/11.3.529

[15] Tim Davies. 2010. Open Data, Democracy and Public Sector Reform: A Look at Open Government Data use from Data. gov. uk. Retrieved from http://www.opendataimpacts.net/report/

[16] Sharon S. Dawes, Lyudmila Vidiasova, and Olga Parkhimovich. 2016. Planning and designing open government data programs: An ecosystem approach. *Govern. Inf. Quart.* 33, 1 (2016), 15–27. DOI : https://doi.org/10.1016/J.GIQ.2016.01.003

[17] David L. Deephouse and Pursey P. M. A. R. Heugens. 2009. Linking social issues to organizational impact: The role of infomediaries and the infomediary process. *J. Bus. Ethics* 86, 4 (2009), 541–553. DOI : https://doi.org/10.1007/s10551-008-9864-3

[18] Ensheng Dong, Hongru Du, and Lauren Gardner. 2020. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* 20, 5 (2020), 533–534. DOI : https://doi.org/10.1016/S1473-3099(20)30120-1

[19] Brandon Fan, Weiguo Fan, Carly Smith, and Harold "Skip" Garner. 2020. Adverse drug event detection and extraction from open data: A deep learning approach. *Inf. Process. Manage.* 57, 1 (2020), 102131. DOI : https://doi.org/10.1016/j.ipm.2019.102131

[20] Joseph Feller, Patrick Finnegan, and Olof Nilsson. 2017. Open innovation and public administration: Transformational typologies and business model impacts. *Eur. J. Inform. Syst.* 20, 3 (2017), 358–374. DOI : https://doi.org/10.1057/EJIS.2010.65

[21] Enrico Ferro and Michele Osella. 2013. *Eight Business Model Archetypes for PSI Re-Use*. In *Open Data on the Web Workshop 23rd – 24th, Google Campus*. Shoreditch, London.

[22] Bent Flyvbjerg. 2006. Five misunderstandings about case-study research. *Qualit. Inq.* 12, 2 (2006), 219–245. DOI : https://doi.org/10.1177/1077800405284363

[23] Yingying Gao and Marijn Janssen. 2020. Generating value from government data using ai: An exploratory study. In *Electronic Government 2020*, G. Viale Pereira, Marijn Janssen, Habin Lee, and Ida Lindgren (eds.). Springer, Cham, 319–331. DOI : https://doi.org/10.1007/978-3-030-57599-1_24

[24] Yingying Gao, Marijn Janssen, and Congcong Zhang. 2021. Understanding the evolution of open government data research: Towards open data sustainability and smartness. *Int. Rev. Adm. Sci.* 00208523211009955 (2021). DOI : https://doi.org/10.1177/00208523211009955

[25] Shirley Gregor and David Jones. 2007. The anatomy of a design theory. *J. Assoc. Inf. Syst.* 8, 5 (2007), 312–335..

[26] Juyeon Ham, Yunmo Koo, and Jae Nam Lee. 2019. Provision and usage of open government data: Strategic transformation paths. *Industr. Manage. Data Syst.* 119, 8 (2019), 1841–1858. DOI : https://doi.org/10.1108/IMDS-04-2019-0218

[27] Teresa M. Harrison and Theresa A. Pardo. 2021. Data, politics and public health. *Dig. Govern.: Res. Pract.* 2, 1 (2021), 1–8. DOI : https://doi.org/10.1145/3428123

[28] Teresa M. Harrison, Theresa A. Pardo, and Meghan Cook. 2012. Creating open government ecosystems: A research and development agenda. *Fut. Internet* 4, 4 (2012), 900–928. DOI : https://doi.org/10.3390/fi4040900

[29] Alan R. Hevner. 2007. A three cycle view of design science research. *Scand. J. Inf. Syst.* 19, 2 (2007), 4.

[30] Alan R. Hevner, Salvatore T. March, Jinsoo Park, and Sudha Ram. 2004. Design science in information systems research. *MIS Quart.: Manage. Inf. Syst.* 28, 1 (2004), 75–105. DOI : https://doi.org/10.2307/25148625

[31] Michael G. Jacobides, Carmelo Cennamo, and Annabelle Gawer. 2018. Towards a theory of ecosystems. *Strateg. Manage. J.* 39, 8 (2018), 2255–2276. DOI : https://doi.org/10.1002/SMJ.2904

[32] Katleen Janssen. 2011. The influence of the PSI directive on open government data: An overview of recent developments. *Govern. Inf. Quart.* 28, 4 (2011), 446–456. DOI : https://doi.org/10.1016/j.giq.2011.01.004

[33] Marijn Janssen, Yannis Charalabidis, and Anneke Zuiderwijk. 2012. Benefits, adoption barriers and myths of open data and open government. *Inf. Syst. Manage.* 29, 4 (2012), 258–268. DOI : https://doi.org/10.1080/10580530.2012.716740

[34] Marijn Janssen, George Kuk, and René W. Wagenaar. 2008. A survey of Web-based business models for e-government in the netherlands. *Govern. Inf. Quart.* 25, 2 (2008), 202–220. DOI : https://doi.org/10.1016/j.giq.2007.06.005

[35] Marijn Janssen and Anneke Zuiderwijk. 2014. Infomediary business models for connecting open data providers and users. *Soc. Sci. Comput. Rev.* 32, 5 (2014), 694–711. DOI : https://doi.org/10.1177/0894439314525902

[36] Alireza M. Javid, Xinyue Liang, Arun Venkitaraman, and Saikat Chatterjee. 2020. Predictive analysis of COVID-19 time-series data from johns hopkins university. arXiv:2005.05060. Retrieved from http://arxiv.org/abs/2005.05060.

[37] Thorhildur Jetzek, Michel Avital, and Niels Bjørn-Andersen. 2019. The sustainable value of open government data. *J. Assoc. Inf. Syst.* 20, 6 (2019), 702–734. DOI : https://doi.org/10.17705/1jais.00549

[38] Peter A. Johnson and Sarah Greene. 2017. Who are government opendata infomediaries? A preliminary scan and classification of open data users and products. *URISA J.* 28, 1 (2017), 9–18..

[39] Juliet Kaarbo and Ryan K. Beasley. 1999. A practical guide to the comparative case study method in political psychology. *Polit. Psychol.* 20, 2 (1999), 369–391. DOI : https://doi.org/10.1111/0162-895X.00149

[40] Maxat Kassen. 2018. Adopting and managing open data: Stakeholder perspectives, challenges and policy recommendations. *Aslib J. Inf. Manage.* 70, 5 (2018), 518–537. DOI : https://doi.org/10.1108/AJIM-11-2017-0250

[41] Gavin Kelly, Geoff Mulgan, and Stephen Muers. 2002. Creating public value: An analytical framework for public service reform. Strategy Unit, Cabinet Office, UK Government (2002), 1–35. Retrieved from http://webarchive.nationalarchives.gov.uk/20100416132449/http://www.cabinetoffice.gov.uk/media/cabinetoffice/strategy/assets/public_value2.pdf

[42] Tomi Kinnari. 2012. Open data business models for media industry—Finnish case study. Master's thesis. Aalto university, Helsinki, Finland.

[43] Fotis Kitsios and Maria Kamariotou. 2019. Open data value network and business models: Opportunities and challenges. In *Proceedings of the 21st IEEE Conference on Business Informatics (CBI'19)*, 296–302. DOI : https://doi.org/10.1109/CBI.2019.00040

[44] Jan Kucera and Dusan Chlapek. 2014. Benefits and risks of open government data. *J. Syst. Integr.* 5, 1 (2014), 30–41. DOI : https://doi.org/10.20470/jsi.v5i1.185

[45] Atif Latif, Anwar Us Saeed, Patrick Hoeer, Alexander Stocker, and Claudia Wagner. 2009. The linked data value chain: A lightweight model for business engineers. In *Proceedings of the 7th International Conference on Semantic Systems (I-SEMANTICS'09)*, 568–575.

[46] P. Linna, T. Makinen, and K. Yrjonkoski. 2017. Open data based value networks: Finnish examples of public events and agriculture. In *Proceedings of the 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO'17)* (2017), 1448–1453. DOI : https://doi.org/10.23919/MIPRO.2017.7973649

[47] Wenzeng Lyu and Lei Zheng. 2017. Value research on open government data: Cases of Shanghai open data apps in China. *ACM Int. Conf. Proc. Ser.* F128275, (2017), 582–583. DOI : https://doi.org/10.1145/3085228.3085250

[48] Gustavo Magalhaes and Catarina Roseira. 2020. Open government data and the private sector: An empirical view on business models and value creation. *Govern. Inf. Quart.* 37, 3 (2020), 101248. DOI : https://doi.org/10.1016/j.giq.2017.08.004

[49] Salvatore T. March and Gerald F. Smith. 1995. Design and natural science research on information technology. *Decis. Supp. Syst.* 15, (1995), 251–266..

[50] Patrice McDermott. 2010. Building open government. *Govern. Inf. Quart.* 27, 4 (2010), 401–413. DOI : https://doi.org/10.1016/j.giq.2010.07.002

[51] Eunice Mercado-Lara and J. Ramon Gil-Garcia. 2014. Open government and data intermediaries: The case of aiddata. In *ACM International Conference Proceeding Series* (2014), 335–336. DOI : https://doi.org/10.1145/2612733.2612789

[52] Shah J. Miah, Don Kerr, and Liisa von Hellens. 2014. A collective artefact design of decision support systems: Design science research perspective. *Inf. Technol. People* 27, 3 (2014), 259–279. DOI : https://doi.org/10.1108/ITP-04-2012-0041

[53] Alexander Osterwalder. 2004. The Business Model Ontology: A Proposition in a Design Science Approach. Université de Lausanne. Ph.D. Dissertation. University of Lausanne, Lausanne, Switzerland.

[54] Alexander Osterwalder and Yves Pigneur. 2010. *Business Model Generation*. John Wiley & Sons, Ltd..

[55] Gabriela Viale Pereira, Marie Anne Macadar, Edimara M. Luciano, and Maurício Gregianin Testa. 2017. Delivering public value through open government data initiatives in a smart city context. *Inf. Syst. Front.* 19, 2 (2017), 213–229. DOI : https://doi.org/10.1007/s10796-016-9673-7

[56] Michael Rappa. 2001. Business Models on the Web. Retrieved November 5, 2021 from http://digitalenterprise.org/models/models.html

[57] Luigi Reggi and Sharon Dawes. 2016. Open government data ecosystems: Linking transparency for innovation with transparency for participation and accountability. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer Verlag, 74–86. DOI : https://doi.org/10.1007/978-3-319-44421-5_6

[58] Tobias Röddiger, Michael Beigl, Dominik Dörner, and Matthias Budde. 2020. Responsible, automated data gathering for timely citizen dashboard provision during a global pandemic (COVID-19). *Digit. Gov.: Res. Pract.* 2, 1 (2020). DOI : https://doi.org/10.1145/3428471

[59] Todd Sanderson, Andrew Reeson, and Paul Box. 2019. Optimizing open government: An economic perspective on data sharing. *PervasiveHealth: Pervas. Comput. Technol. Healthcare Part F148155*, (2019), 140–143. DOI : https://doi.org/10.1145/3326365.3326383

[60] Scott M. Shafer, H. Jeff Smith, and Jane C. Linder. 2005. The power of business models. *Bus. Horizons* 48, 3 (2005), 199–207. DOI : https://doi.org/10.1016/J.BUSHOR.2004.10.014

[61] Andrew Shipilov and Annabelle Gawer. 2020. Integrating research on interorganizational networks and ecosystems. 14, 1 (2020), 92–121. DOI : https://doi.org/10.5465/ANNALS.2018.0121

[62] Goran Smith, Hosea Ayaba Ofe, and Johan Sandberg. 2016. Digital service innovation from open data: Exploring the value proposition of an open data marketplace. In *Proceedings of the Annual Hawaii International Conference on System Sciences*. 1277–1286. DOI : https://doi.org/10.1109/HICSS.2016.162

[63] Barry Smyth. 2020. Lockdowns & rebounds. *Dig. Govern.: Res. Pract.* 1, 4 (2020), 1–7. DOI : https://doi.org/10.1145/3411762

[64] Xiaoling Song, Charles Shen, Lin Zhong, and Feniosky Peña-Mora. 2017. SEM-based value generation mechanism from open government data in environment/weather sector. In *Proceedings of the 11th International Conference on Management Science and Engineering Management*. 351–366. DOI : https://doi.org/10.1007/978-3-319-59280-0_29

[65] David J. Teece. 2010. Business models, business strategy and innovation. *Long Range Plan.* 43, 2–3 (2010), 172–194. DOI : https://doi.org/10.1016/J.LRP.2009.07.003

[66] Piyushimita (Vonu) Thakuriah, Lise Dirks, and Yaye Mallon Keita. 2017. Digital infomediaries and civic hacking in emerging urban data initiatives. In *Springer Geography* (2017), 189–207. DOI : https://doi.org/10.1007/978-3-319-40902-3_11

[67] Paul Timmers. 1998. Business models for electronic markets. *Electr. Markets* 8, 2 (1998), 3–8. DOI : https://doi.org/10.1080/10196789800000016

[68] Pasquale del Vecchio, Gioconda Mele, Valentina Ndou, and Giustina Secundo. 2018. Creating value from social big data: implications for smart tourism destinations. *Inf. Process. Manage.* 54, 5 (2018), 847–860. DOI : https://doi.org/10.1016/j.ipm.2017.10.006

[69] Tung-Mou Yang, Jin Lo, Hui-Ju Wang, and Jing Shiang. 2013. Open data development and value-added government information: case studies of taiwan e-government. In *Proceedings of the 7th International Conference on Theory and Practice of Electronic Governance*. DOI : https://doi.org/10.1145/2591888

[70] Fatemeh Ahmadi Zeleti and Adegboyega Ojo. 2016. Open data value capability architecture. *Inf. Syst. Front.* 19, 2 (2016), 337–360. DOI : https://doi.org/10.1007/S10796-016-9711-5

[71] Hans Dieter Zimmermann and Andreja Pucihar. 2015. Open innovation, open data and new business models. In *–Proceedings of the 23rd Interdisciplinary Information and Management Talks (IDIMT'15)*, Johannes Kepler Universitat Linz, 449–458. DOI: https://doi.org/10.2139/ssrn.2660692

[72] Christoph Zott and Raphael Amit. 2008. *The Fit between Product Market Strategy and Business Model: Implications for Firm Performance.* John Wiley & Sons, Ltd. DOI: https://doi.org/10.1002/SMJ.642

[73] Christoph Zott and Raphael Amit. 2010. Business model design: an activity system perspective. *Long Range Plan.* 43, 2–3 (2010), 216–226. DOI: https://doi.org/10.1016/J.LRP.2009.07.004

[74] Christoph Zott, Raphael Amit, and Lorenzo Massa. 2011. The business model: Recent developments and future research. *J. Manage.* 37, 4 (2011), 1019–1042. DOI: https://doi.org/10.1177/0149206311406265

[75] Anneke Zuiderwijk, Natalie Helbig, J. Ramón Gil-García, and Marijn Janssen. 2014. Special issue on innovation through open data - A review of the state-of-the-art and an emerging research agenda: Guest editors' introduction. *J. Theor. Appl. Electr. Commerce Res.* 9, 2 (2014). DOI: https://doi.org/10.4067/S0718-18762014000200001