A Conceptual Framework for Open Government Data: Drawing upon Conditions Surrounding Use

Alizée Francey
Swiss Graduate School of Public Administration, University of Lausanne, alizee.francey@unil.ch

Follow this and additional works at: https://aisel.aisnet.org/itais2021

Recommended Citation
https://aisel.aisnet.org/itais2021/2
A Conceptual Framework for Open Government Data: Drawing upon Conditions Surrounding Use

Alizée Francey

Swiss Graduate School of Public Administration, University of Lausanne, Lausanne, Switzerland
alizee.francey@unil.ch

Abstract. Public decision-makers’ interest in digital technology has fuelled the debate about the potential of open government data (OGD) as a foundation and driver for economic growth, competitiveness, innovation, job creation, and societal progress in general. These potentials create high expectations sustained by a plethora of economic and societal forecasts about OGD. However, there appears to be a lack of scholarly literature about OGD use from which value stems. This paper addresses this concern by developing a conceptual framework scoping conditions surrounding OGD use. The conceptual framework, built on the literature, gathers three major components, i.e. (1) the characteristics of the datasets, (2) the features of the OGD portal, and (3) the organizational resources and capabilities, which altogether afford use. The paper also provides some insights for researchers and policymakers. By structuring conditions surrounding OGD use in organizations, the framework allows to further distinguish between the opportunities offered by the OGD portal and respective datasets from the actualization of those opportunities by organizations throughout use.

Keywords: Open Government Data, OGD, Framework, Use, Portal, Datasets

1 Introduction

Once generated and collected, government-produced data are encouraged to be published to generate useful information for the public [1, 2]. Although encouraging the use of open government data (OGD) and acknowledging OGD technologies are necessary elements for OGD to create value, the open data debate has mainly been oriented towards OGD provision leaving aside OGD use [2, 3]. The paucity of research regarding OGD use can be attributed to several factors such as the diversity of users, the predictors influencing acceptance and use of OGD as well as the variability of uses. This complexity around OGD makes altogether OGD use difficult to examine. Even though the lack of research still makes it unclear how much value stems from OGD use, public decision-makers’ recent interest in digital technology has fuelled the debate

---

1 OGD refer to non-sensitive, digitized data produced by state bodies (e.g., census data, cartographic material, statistical data) and made available to everyone with minimal mechanisms of control like copyright, price, or repurposing.
about digital government transformation [4], particularly the potential of OGD considered as a foundation and driver of economic growth, competitiveness, innovation, job creation and societal progress [5]. From the mentioned considerations, governments are promoting a plethora of economic and societal forecasts about OGD, predicting billions of annual benefits and cost savings. To mention an illustrative example, the total direct economic value of government data is expected to increase from a baseline of €52 billion in 2018 for the 27 EU countries and the United Kingdom, to €194 billion in 2030 [6].

Given the expected value for the economy and society stemming from OGD, several countries have focused their intention on increasing OGD accessibility by devoting funds and resources to develop technical infrastructures and build OGD portals, which increase the accessibility of government-produced data [7]. However, despite the increased accessibility of OGD, previous studies have shown that the use of OGD is currently low in practice and that few organizations are using OGD due to a lack of resources, skills, and opportunities needed to make further use of the datasets [8-10]. However, OGD has limited value in itself [11, 12] yet depends on its associated purposes becoming utmost valuable when users know how to use it [13]. Hence, to know if OGD honours its economic and societal expectations, there is a need to further examine OGD use from which value stems [12].

To further examine OGD use yet account for its complexity, this paper seeks to decompose this complexity by answering the following research question: What are the conditions surrounding OGD use? To achieve that goal, a conceptual framework scoping conditions surrounding OGD use is developed based on the literature. By developing a conceptual framework, the paper aims to get a comprehensive understanding of the conditions surrounding OGD use. With this purpose in mind, the conceptual framework provides different axes of scoping conditions surrounding OGD use in a coordinated approach to clarify the mechanisms for OGD to occur. In doing so, the contributions to the OGD literature are twofold. While better understanding the use of OGD is a major concern for policymakers and practitioners, considering the current empirical and theory-building research in OGD, this research is also input to theory. In sum, identifying conditions surrounding OGD use allows to further distinguish between the opportunities offered by OGD from the actualization of those opportunities throughout OGD use.

The paper is structured as follows: I first outline the background of the paper, then present the developed conceptual framework before discussing and drawing some conclusions.

2 Background

Before introducing the conceptual framework scoping the conditions surrounding OGD use, the underlying dimensions, and respective criteria, this section aims to briefly outline the three-stage process embedding OGD as well as the actions needed for OGD use to occur.
According to Ubaldi [14], OGD use is the final phase of a process, from which I outline three stages and respective stakeholders (i.e. data production, data distribution, and data use) (also see [15]). I thus hold that OGD use is surrounded by several conditions emerging from each stage of the process and respective stakeholders (i.e., OGD producers, OGD intermediary, and OGD users). As the aim of this paper is not to detail the process, I only address the stages and respective stakeholders from the standpoint of further use to collect conditions surrounding OGD use. By conditions surrounding OGD use, I understand the main conditions emerging from each stage and which, by intertwining together, alter the possibilities of use. Indeed, while the stages and respective stakeholders are independent within the process, they are strongly interrelated as they reinforce one another [16]. Hence, this research seeks to undertake an ecosystem perspective by considering the stakeholders and respective conditions emerging from each stage of the process.

2.1 From OGD production to OGD use

Firstly, datasets are generated by OGD producers (e.g., public administrations, organizations with a state mandate…). During this stage, datasets are aggregated and processed to obtain certain characteristics, defining the quality of the datasets. I consider the characteristics of the datasets as conditions surrounding OGD use because if the datasets do not have certain characteristics (i.e. machine-readable formats, updated datasets…) with regards to their users, they may not be exploited to their full potential [17]. Secondly, data distribution ensures that government-produced data are shared through the OGD portal, which plays the role of intermediary allowing the users (e.g., private, public, non-governmental organizations) to access the datasets [18]. Given that the OGD portal is key to further use by allowing users to access the datasets, most of the national governments in Europe have developed such portals [19]. Accordingly, I acknowledge the features of the OGD portal as conditions surrounding OGD use. Thirdly, OGD use relies upon the user and is dependent on certain capacities [20]. While OGD use can be studied at any level of analysis, I chose the organizational level (i.e., private, public, or non-governmental organizations). Given that certain resources and capabilities are needed to take proper advantage of OGD, I consider the organizational resources and capabilities as conditions surrounding OGD use.

2.2 From a goal-oriented activity to value creation

While I argue that the characteristics of the datasets, the features of the OGD portal, as well as the organizational resources and capabilities, alter the possibilities of use, OGD use occurs when a goal-oriented activity triggers the user to use the datasets, from which value stems. Concerning the goal-oriented activity, I consider that each type of user has particular interests and intended use [9, 21]. While public sector organizations may use OGD to improve public service or decision and policy-making processes [9, 22, 23], private sector organizations may use OGD for innovation purposes [23-25], for developing better services to customers, or for making profits [23]. Other types of users such as
research institutes, journals, or non-governmental organizations may use OGD to better understand a thematic, to influence policy, or simply to create knowledge in a different format or to inform citizens [9, 10, 23, 26, 27]. While this diversity shows that each type of user has different goals, it also shows that each user has a defined objective when accessing the datasets through the OGD portal. Hence, this research assumes that OGD is never used just to be used but rather to achieve a goal-oriented activity. This research thus recognizes that there are as many goals triggering the use of OGD as users’ profiles and understand by goal-oriented activity whatever the user aims to achieve using OGD. As OGD has limited value in itself [11, 12] yet depends on its associated purposes when used [13], I define OGD use as the action of a user exploiting OGD in a way that helps the latter achieves a goal-oriented activity. This research thus asserts that OGD use becomes utmost valuable when the characteristics of the datasets, the features of the OGD portal, as well as the organizational resources and capabilities altogether afford the achievement of the user’s goal-oriented activity creating OGD value. Measuring the economic and societal value of OGD encounters several issues as identifying the factors contributing to the success are diverse and for now little attention has been paid to this phenomenon in the academic literature [20]. However, as I consider that OGD use is triggered by a goal-oriented activity and most importantly that any goal achievement, even when hard to assess, is never completely subjective, understanding the conditions of OGD use opens the door to further investigate value creation.

3 Conceptual framework

To develop the conceptual framework, I built on the literature. I first identified a set of papers through manual search using starting conditions and specific terms related to the characteristics of the datasets, the features of the open government data portal as well as the organizational resources and capabilities of open government data users. I used Google Scholar to avoid bias in favour of any specific publisher [28]. I restricted the language to English and decided against OGD as a keyword because it is also extensively used to abbreviate other scientific terms such as oxygen-glucose deprivation in biochemical and biophysical research, or oesophagal-gastro-duodenoscopy in the medical field. Moreover, I used backwards snowballing as a research approach and thus conducted a backward search within the references section [28]. I ended the process once no new papers were found in the iteration because reaching a saturation point [28].

From the synthesis of the literature identified (i.e., references to be found in Tables 1, 2, and 3), I developed the following conceptual framework aiming to examine the conditions surrounding OGD use. By defining how the user appraises the datasets and the OGD portal, this research aims to further identify the conditions surrounding OGD use and thus value creation. Figure 1 presents an overview of the conceptual framework gathering the three major components surrounding OGD use, i.e. (1) the characteristics of the datasets, (2) the features of the OGD portal, and (3) the organizational resources and capabilities, which altogether afford use. Each component is further detailed in the following sub-sections.
3.1 Characteristics of the datasets

This research considers the characteristics of the datasets as conditions surrounding OGD use. They refer to OGD quality without which the use of OGD may be jeopardized [29]. There are a couple of dimensions that contribute to the quality of the datasets and directly alter OGD use such as the accessibility, accuracy, comprehensibility, formats, or completeness of the datasets. To mention an illustrative example, if the datasets are available in formats that are not machine-readable such as PDF files and cannot be processed on a computer using spreadsheets software, their quality is hampered and so is their further use because the user cannot take the best advantage of the datasets. Hence, if the datasets generated by OGD producers do not have some characteristics for their users, then the data will not be fully exploited [17].

This paper relied on the literature to identify ten main dimensions and respective criteria to operationalize the characteristics of the datasets. They are presented in Table 1.

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Criteria</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accessibility</td>
<td>The access to the datasets available on the OGD portal is convenient (the ease with which the relevant dataset is discovered).</td>
<td>[17, 30]</td>
</tr>
<tr>
<td>Ease of use</td>
<td>The datasets available on the OGD portal can be used easily.</td>
<td>[17, 31]</td>
</tr>
</tbody>
</table>
Accuracy

By correctly describing the respective information, the datasets available on the OGD portal are correct and accurate. [7, 17, 29, 31-34]

Completeness

The datasets available on the OGD portal are complete. [17, 29, 30, 33-36]

Consistency

The datasets available on the OGD portal are consistent. [17, 29, 32]

Timeliness

The datasets available on the OGD portal are associated with a time and are up to date. [7, 17, 29, 30, 32, 37-39]

Formats

The formats provided on the OGD portal are machine-readable and convenient for further use. [14, 29, 30, 32, 34, 36, 38, 39]

Catalogue

The size of the dataset catalogue (the number of datasets it includes) is sufficient. [14, 34, 38, 40]

Comprehensibility

The datasets available on the OGD portal are easily understandable. [17, 29, 37]

Description

The datasets available on the OGD are provided with a description. [14, 32, 34, 39]

### 3.2 Features of the OGD portal

This research considers the features of the OGD portal as conditions surrounding OGD use because the portal serves as the intermediary between the data providers and the data users [7]. By cataloguing the data from various authorities, services, and websites, the OGD portal affords the user access to government-produced data [17]. I thus understand the OGD portal as a “one-stop-shop” that enables OGD users to search and access the published datasets [17].

The development of OGD portals has proliferated since the mid-2000s but such portals vary greatly and such variability influences OGD further use [7, 14]. For example, criticisms have been raised against OGD portals that cannot be used by non-technical users or do not offer an overall friendliness (i.e., by constraining the interaction between data providers and data users) [7, 41-43]. To afford further use and ensure a high level of use, the OGD portal should be fit for use and go beyond providing datasets by assisting the user in finding the datasets needed [7, 38]. For example, an OGD portal can contain the most fruitful datasets but if a bad functioning of the search engine does not allow the user to find the datasets, OGD use is held down.

Hence, there are a couple of dimensions such as the search engine or the extensive documentation to mention few examples that determine if the OGD portal affords the user to find and access the datasets. Relying on the literature, this research identifies five dimensions and respective criteria (i.e., Table 2) to operationalize the features of the OGD portal.

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Criteria</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usability</td>
<td>Navigating on the OGD portal is intuitive.</td>
<td>[7, 34, 44]</td>
</tr>
<tr>
<td>Search engine</td>
<td>The search engine provides strong datasets’ search and finding capacity.</td>
<td>[14, 34, 37, 38, 40]</td>
</tr>
</tbody>
</table>
Application Programming Interface (API) | The OGD portal offers a sufficiently detailed description of the API for searching and downloading datasets. | [14, 34, 38, 40]
---|---|---
Documentation | The OGD portal provides documentation to use the portal. | [7, 14, 34, 36, 38, 39]
Feedback functionality | The OGD portal provides an opportunity to provide feedback or ask questions on the datasets (from the users to providers). | [7, 12, 14, 34, 38, 39, 45]

3.3 Organizational resources and capabilities

To make the best use of OGD, there is a need for the users to exploit the portal and datasets in ways that are beneficial for them [46]. This research considers that certain organizational resources and capabilities are needed to take proper advantage of the OGD portal as the access to the digital infrastructures, to the hardware or software as well as to the financial or training resources and capabilities differ between stakeholders [46].

While the characteristics of datasets and the features of the OGD portal are equal to all users, making use of the portal and respective datasets is dependent on certain organizational resources and capabilities, which I understand as the abilities of an entity to achieve an objective [20, 47]. Hence, the use of OGD requires organizational resources, which are factors controlled by the organization as well as organizational capabilities, which refer to the organization’s ability to deploy its resources [48].

Applied to our case, it is the collective ability of an organization to use OGD for a goal-oriented activity. Without the organizational resources and capabilities enabling to access, use and make sense of the data, the goal-oriented activity, as well as the value stemming from OGD, are constrained. I thus argue that several dimensions can afford the collective ability of organizations to use OGD.

Firstly, human resources are essential to make use of the data [48]. I understand by human resources, employees using their abilities and skills to explore the OGD portal and use the respective datasets. Secondly, IT resources such as storage infrastructures or computers (hardware) as well as analytical tools or platforms (software) are needed to make use of OGD [48]. Thirdly, using OGD for further purposes (e.g., innovation) requires an average amount of financial resources [3, 23]. Fourthly, given that the lack of organizational capabilities and knowledge hinders OGD use, technical capabilities are an essential condition to make use of the data [12, 23, 26, 48]. To be able to analyse or repurpose the data for further use, employees need to have certain skills (e.g., selecting the appropriate tools to download, interpret, visualize, map, and link the data) [48]. Finally, given that some capabilities are needed to make use of the datasets, there is a need for organizations to acquire such capabilities through continuous training. Hence, offering continuing training bolsters organizations’ employees with the ability to acquire technical skills (e.g., statistics, programming…) [3, 23, 48].

Altogether, this research identifies four criteria to operationalize the organizational resources and two criteria to operationalize the organizational capabilities. The six criteria are presented in Table 3.
Table 3. Organizational resources and capabilities

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Criteria</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human resources</td>
<td>The organization possesses the human resources needed to use the datasets available on the OGD portal.</td>
<td>[48, 49]</td>
</tr>
<tr>
<td>IT resources</td>
<td>The organization possesses the hardware needed to use the datasets available on the OGD portal.</td>
<td>[46, 48, 49]</td>
</tr>
<tr>
<td></td>
<td>The organization possesses the software needed to use the datasets available on the OGD portal.</td>
<td>[46, 48, 49]</td>
</tr>
<tr>
<td>Financial resources</td>
<td>The organization supports the use of the OGD portal through financial resources.</td>
<td>[3, 23, 46]</td>
</tr>
<tr>
<td>Technical capabilities</td>
<td>The organization possesses the capabilities and knowledge required to use the datasets available on the OGD portal.</td>
<td>[9, 12, 23, 26, 48, 49]</td>
</tr>
<tr>
<td>Training offer</td>
<td>The organization supports the use of the OGD portal by providing specific training courses.</td>
<td>[3, 9, 23, 46, 48]</td>
</tr>
</tbody>
</table>

4 Discussion and conclusion

In this paper, I developed a conceptual framework scoping the conditions surrounding OGD use at an organizational level. By doing so, I aim to further distinguish between the opportunities offered by the OGD portal and respective datasets from the actualization of those opportunities by organizations throughout use. I thus expect the presented conceptual framework to have several implications.

Firstly, the conceptual framework may be a starting point for managing an OGD ecosystem. While several countries have focused their intention on building OGD portals to increase the accessibility of government-produced data [7], the current practical challenges related to OGD (i.e. low OGD use in practice [8-10]) show that there is a need to look beyond data provision. To do so, I believe that it is necessary to manage an OGD ecosystem to enable manipulation and use of the datasets [51]. The conceptual framework is a first step in managing an OGD ecosystem as it shows that the coordination and development of a whole set of conditions stemming from OGD producers (e.g., public administrations, organizations with a state mandate…), OGD intermediary (e.g., the OGD portal), and OGD users (e.g., private, public, non-governmental organizations) are needed for OGD to be used and unfold its expected value. By allowing to identify what specific conditions stemming from OGD producer, OGD intermediary, and/or OGD users are missing for OGD use to occur, the conceptual framework may be the basis for a more coordinated approach.

Secondly, the conceptual framework can be used to further examine OGD value creation. While the expectations related to OGD seem to be, like most digital transformations in the public sector, guided by hopes and dreams rather than confirmed by empirical evidence [52], I believe that after more than a decade of OGD initiatives worldwide it is time to provide better estimations around OGD. Given the low OGD use in practice [8-10], I believe that the current estimations predicting billions of annual
benefits and cost savings seem to be unrealistic. Hence, the conceptual framework may be a starting point to better estimate the value stemming from OGD use. By providing three different axes scoping the conditions surrounding OGD use, the conceptual framework may be used as a point of departure to clarify all the mechanisms for OGD use to occur and get a more grounded estimation of the value stemming from it [12]. In doing so, the conceptual framework may be useful not only for researchers but also for policymakers.

Finally, being able to consider respective contexts related to OGD use and the value stemming from it offers some opportunity to conduct research, which is not only relevant for theoretical development but also practice. While many countries in Europe have focused on the development of strong data infrastructures [52] rather than ecosystems, previous studies have shown that despite such developments, the use of OGD is low in practice [8-10] hampering in turn value creation. A good example has been Switzerland, because, like many other countries in Europe, Switzerland has focused on developing technical infrastructures and building an OGD portal – opendata.swiss – to increase the accessibility of government-produced data [53]. Hence, like other governments’ strategies, the Swiss strategy is based on the fixed idea that sharing a large amount of data is sufficient for further use. However, an official audit reported several shortfalls in governance and utility of the national OGD portal and as of yet, the Swiss economic and social benefits fall far away from the predictions made [53]. A case analysis, such as from Switzerland (but also other countries), is useful given the practical challenge related to the lack of OGD use and value creation. I thus seek to apply the conceptual framework to examine the conditions surrounding OGD use in a specific context. To do so, I will next conduct a national-scale survey targeting users of the Swiss national OGD portal through a questionnaire. By studying a given portal using the conceptual framework, I seek to determine if the characteristics of the datasets and the features of the OGD portal allow organizations to achieve a goal-oriented activity. In addition, this study will also enable to examine whether organizations hold the resources and capabilities to carry it out. I thus expect our national scale survey to provide a more holistic and generalizable picture of how the characteristics of the datasets, the OGD portal features, and organizational resources and capabilities are interlinked in OGD use.

While this research offers some implications for research and practice, it also has some limitations. The limitations are notably related to the fact that the conceptual framework is built on the literature. Hence, an obvious limitation of manual search is coverage as it cannot be sure that all studies scoping conditions surrounding OGD use have been identified. This limitation is partially offset because of the iteration process realised with snowballing as a research strategy, which enables the identification of the most relevant articles. However, another limitation related to coverage touches upon the terminology, which is too often not standardized and may hamper the identification of relevant articles through specific terms. As illustrative examples, some authors refer to ease of use, format, or consistency of the datasets, while others refer to usability, compliance, or incongruent datasets respectively. The limitation related to the terminology also pertains to the understanding of the dimensions. For example, while this research refers to accessibility as the simple fact of accessing the datasets available on
the OGD portal and to completeness for thorough datasets, other research computes completeness with availability and accessibility [29, 50]. Regarding the dimensions, some are based on several references, while others are only based on two references such as accessibility or human resources. While this shows some limitations, the discrepancy may not relate to the criticalness of these dimensions. It could be that these dimensions are so intrinsic to OGD use that there are not always mentioned. In that sense, accessibility or human resources are primary and intrinsic conditions for other conditions to be considered. Indeed, the format of the datasets becomes irrelevant if the accessibility of such datasets is itself inconvenient. Another example is the training offer, which becomes useless if there are no human resources to be trained.

References


