The Impact of Data Governance on OGD Publication – An Ethnographic Odyssey

Auriane Marmier
Swiss Graduate School of Public Administration (Idheap), University of Lausanne, Rue de la Mouline 28, CH-1022 Chavannes-près-Renens
auriane.marmier@unil.ch

ABSTRACT

Over the past decade, Open Government Data (OGD) strategies have become a continuing concern in administrative services. This is even truer than at any time. Given the current situation, data management, specifically consistent data publication, has been central to public institutions. The Covid-19 pandemic has shown that data collected by public administrations could make valuable contributions. However, in Switzerland, the pandemic has highlighted the limitations of public organizations' capability to lead the publication of their data. Based on an ethnography and a literature review, this paper explores how data governance components impact OGD publication process and presents a model of OGD governance. For this purpose, we identify key data governance components necessary to OGD publication - structural, procedural, and relational - and illustrate how OGD challenges rarely arise from the publication of OGD or the open nature of data itself, but a lack of data governance.

CCS CONCEPTS

• **Information systems** → Data management systems.

ACM Reference Format:
Auriane Marmier. 2022. The Impact of Data Governance on OGD Publication – An Ethnographic Odyssey. In *DG.O 2022: The 23rd Annual International Conference on Digital Government Research (dg.o 2022), June 15–17, 2022, Virtual Event, Republic of Korea.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3543434.3543438

1 INTRODUCTION

With the emergence of the open government policy under the Obama presidency, Open Government Data (OGD) - which are transformed government data to be openly published, shared, and reused by anyone for any purpose - became a continuing concern in worldwide administration services. The Washington Post recently reported that business groups, including the Software Alliance, the Information Technology Industry Council and the Internet Association, pressure the Biden administration to coordinate open data efforts across the government [1]. From an academic perspective, many authors such as [2] and [3] claimed that open access to government data has increased rapidly and represents a growing part of government management activities. Due to their daily activities, governments not only collect and create a plethora of data, but they also appear as the central actors in charge of these strategic assets [4]. This is even truer now than at any time. The Covid-19 pandemic provides concrete examples and has shown that public administrations' data might help better understand the situation, identify clusters, and plan adapted measures [5]. Moreover, the pandemic has also demonstrated how necessary the availability and exchange of data in crisis management has become. While collecting data has always been a structural component in public organizations occupations, the Covid pandemic recently reveals limitations of data handling in public organizations. [6] notably state that government test results may contain numerous errors, and it is unclear if these data provide valuable information to decision-making and generate value [7]. In the same vein, leading media outlets describe a chaotic data collection and information exchange situation. In Switzerland, the country where this study originates, there seems to be a lack of interoperability between the Federal Government and lower-tier administration (i.e. the 26 Cantons the Swiss's administration is divided in), skills deficit as well as a lack of communications between the actors concerned [8]. In many cases, disciplines, practices, tools, and techniques for collecting, cleaning, and organizing government data seem different. This makes data ownership, strategies as well as data publication, among others, often unclear, adding new data management and governance complications [4, 7] and leading public sectors organizations face an increasing number of OGD challenges [9].

Although there is a growing body of literature on OGD publication regarding sociological, technical, and legal challenges [7, 10-15], few papers consider data governance as the root of the problem. Based on an in-depth literature search, [16] define data governance as *"a cross-functional framework for managing data as a strategic enterprise asset. In doing so, data governance specifies decision rights and accountabilities for an organization's decision-making about its data. Furthermore, data governance formalizes data policies, standards and procedures, and monitors compliance"*. Recognize as the main challenge in (big) data value creation, data governance provides organizations insights and information necessary to run a data-driven organization [9] and appears relevant for facilitating open data use [17]. [18] highlighted that only effective data governance may allow value creation, and Soares [19] stated that to leverage data as valuable assets, people, processes, and technologies must be effectively managed through data governance programs [9]. [20] describes a data governance program as a framework providing data security, allowing data policies and standards development,

dg.o 2022, June 15–17, 2022, Virtual Event, Republic of Korea

ACM ISBN 978-1-4503-9749-0/22/06…$15.00
https://doi.org/10.1145/3543434.3543438

and thus assisting decision-makers. While data governance receives increasing attention in the private sector as an appropriate solution to resolve organizational issues with data [7], so far, little is known regarding the impact of data governance on OGD publication [21]. [22] show that only three percent of the data governance papers written between 2007 and 2017 focus on the e-government research field. Hence, we believe that the publication of government data in open access is not the only reason that makes public organizations overwhelmed by OGD publication, but that OGD barriers find their origins deeper in data governance issues. Based on an extensive review of the literature on OGD and data governance and combined with an ethnography inquiry, this paper attempts to formulate theoretical assumptions in order to develop an OGD governance model. In doing so, we seek to extend the current discourse and better understand to what extent data governance practices may impact the publication of OGD [4]. Accordingly, we aim to answer the following research question:

- To what extent do data governance practices impact the OGD publication?

For this purpose, we first discuss the data collection approach of both literature review and ethnography as well as the data analysis. Then, to make the ethnography easier to understand, we present the context in which it happened. In the fourth part, we present and discuss results by cross-checking the findings of existing literature with the analyses of ethnographical data collected. In doing so, we seek to identify the crucial data governance practices and their effects on the OGD publication process. This step allows us to define convergences between the theory and the practice and formulate eleven hypotheses clustered in three majors' components. Finally, we conclude this paper by presenting our OGD governance model.

2 METHODOLOGICAL APPROACH

2.1 Data collection

To a better knowledge of data governance practices and their potential impacts on the OGD publication, we conducted a comprehensive literature review on two types of publication (Table 1). We first examined (i) practice-oriented publications, i.e. relevant official documents such as *Swiss OGD strategies*, *cantonal strategy*, *legislative agenda* [29-31], *reports of the European Commission [32, 33]*, as well as (ii) scientific literature on OGD and data governance. From a practice-oriented publication perspective, we led research throughout the project and focused on the city and cantonal websites and the Swiss and European OGD platforms. We notably researched with different web browsers, using keywords such as *data governance strategy, OGD strategy, or data governance practices in Switzerland*. From a scientific publication perspective, we conducted a systematic literature review in September 2021 and integrated materials published between 2001 and 2021 on two platforms - *Web of science* and *AIS library*. We considered these two platforms a good starting point to study data governance state of knowledge given that the first tends to regroup public administration papers while the latter clusters literature on Information System (IS) research. On both platforms, we used the two following search strings - "*data governance" AND "public sector"*, and *"Open Government Data" OR "open government data"*. As a result, we found a total of 496 papers. After removing duplicates, editors' comments, and introductions to mini tracks, we finally obtained 464 scientific papers. To further reduce the number of relevant papers, we created an excel database including an ID, the title and the abstract of the papers, authors' name, and year of publication, in which we made a key words research. Thus, we were able to selectively focus on papers that contained the terms "open data","data governance" and "data governance practices" words in their abstract. This operation leads us to focus on 22 papers intensely.

In order to identify the daily practices of municipal departments involved in the OGD municipal sample preparation process and closely investigate their mechanisms, we conducted an ethnographic inquiry. Ethnography is described by [23] as a methodological and practice-based approach that seeks to pinpoint human interactions with other humans, objects, environment or institutions to better understand their operations. According to many authors, this qualitative research design facilitates exploration in a real-life context of departments and strongly contributes to pinpointing fractures and rifts workers [24-26]. Furthermore, through multiple sources of evidence, ethnography aims to describe both the group members' point of view and the perceptions and interpretations of researchers [27]. Hence, for a researcher, ethnography allows exploring the roles of departments members, their actions and contributes to a better comprehension of the flow of organizational activity, events, and dynamics in their daily work [26, 28]. Therefore, to observe how the OGD publishing process works and from a more global perspective to understand better how municipal departments apprehend management and data governance, we participated in six meetings, exchanged hundreds of emails, and participated in one hackathon. We collected data using direct observations and unstructured interviews realized during this period with three housing department members and three IT department members supporting digital initiatives (Table 1). In addition, as recommended by ethnography literature, we kept a logbook. [24] argued that an ethnography logbook helps limit participants' perceptions and provide researchers with objective information. In the logbook, we resumed and described interactions with department members, registered the meeting minutes, and added a copy of the email exchange. We also included the results of our observations (e.g. feelings).

2.2 Data Analysis

To analyze data collected through the review of material published, the ethnography, and thus deeper understand the data publication process, we followed a content analysis approach [34]. According to scholars, the content analysis approach is generally used to examine qualitative data such as interviews, semi-structured interviews, documents but can also be applied to various nonverbal data, such as feelings or gestures [35-37]. In doing so, this approach allows to analysis a large amount of data by revealing different categories. Furthermore, as the qualitative content analysis also focuses on the underlying meaning of words [38], this analysis is especially useful for understanding more profound a phenomenon [39]. To this end, we thus followed a well-defined step-by-step process proposed by [35]. We started by studying data governance practices mentioned in the literature. As recommended by the authors, we first familiarized ourselves with the data by reading and rereading material and

Table 1: Data gathering

Field Data	Literature sources
Meetings minutes	Scientific papers:
Direct Observations	AIS Library
Unstructured interview	Web of Science
Logbook	Practice – oriented publications:
Emails	European commission reports
Information models	Cantonal and Federal digital strategies
Database extractions	Legislative agenda

Table 2: Definitions and references of data governance components

Definition	References
Structural components refer to the roles and responsibilities of a public organization and the allocation of decision-making authority from a data governance perspective.	[19] [51] [52] [53] [54] [55] [56] [57] [58] [59] [60] [61]
Procedural components comprise data strategies, policies, standards, processes, and procedures that provide data to be appropriately recorded, maintained safely, exploited effectively, and shared adequately.	[16] [19] [51] [52] [53] [54] [55] [56] [57] [58] [59]
Relational components relate to business alignment, employee competencies and organizational culture that support decision-making.	[16] [17] [19] [59] [60] [61] [63] [64]

keeping an eye on the recurrent information. This enabled us to identify the essential practices for the success of data governance programs. However, while the literature comprises similar data governance practices, the semantic diversity of terminology hampers information understanding and makes compilation and analysis difficult. To address this issue, we followed the results of [16], which gathers the different practices observed in the literature under three major data governance components - structural, procedural and relational (Table 2).

To study the ethnographic results, we also applied the content analysis approach. Following the same logic, we first familiarized ourselves with the content of the ethnographic material. Then, from the initial information noted in the first step, we generated a code for the information that seemed relevant. Next, we organized previously defined codes according to components highlighted in the literature review. Finally, to examine to what extent practices observed in the literature and during our ethnography were convergent, we applied the triangulation technique [40]. This enabled us to cross-check the ethnography inquiry data against scientific and practice-oriented literature and thus compare theoretical perceptions of data governance practices with actual municipal practices.

While the term triangulation may lose some clarity when it refers to science methodology [41], it is a metaphor that describes a technique allowing the investigation and comprehension of a situation from several perspectives [42-45]. Based on the cross-checking of different data sources such as quantitative survey, participant observation, historical analysis, comparison and discourse analysis [42], the triangulation technique is used to increase the validity and trustworthiness of the results [43]. By comparing multiple data sources, triangulation pinpoints convergence or divergence in data collected [45, 46]. The triangulation technique allows ethnography to thus struggle with bias implied by a single source of data or method as well as improve the veracity of the study [42, 47, 48]. Thus, as many authors in organizational research [45], geographical studies [44] or computer science [46], we triangulated obtained results in our ethnographic observations, practice-oriented literature and academic studies.

Before discussing the results obtained by triangulation, we describe the context of our ethnography in the next part. While ethnography appears well suited to investigating specific social and cultural phenomena, the inquiry's contextualization is needed to better understand them.

3 CONTEXT

In Switzerland, discussions to make government data publicly available appeared for the first time in the E-Government Strategy 2008-2015 [49]. After that, the Federal Council developed a new strategy for "*an information society in Switzerland intending to optimize information resources*" [50]. In this context, several initiatives concerning data availability have been launched, followed by creating the first Swiss Open Data Strategy, known as "Strategy OGD 2014-2018" [50]. This strategy notably pursued objectives such as encouraging innovation and economic growth, fostering transparency and participation of citizens, and increasing the efficiency of public administrations on all political levels [51]. In line with prior researches [e.g. 52, 53], the Swiss Federal Council considered that the best way to achieve the strategic objectives of the OGD was to develop a national OGD platform that made public data accessible

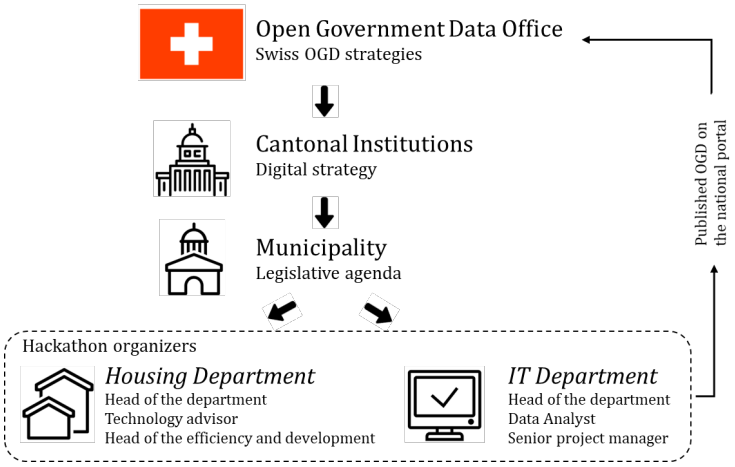

Figure 1: Context and actors of the ethnography

to a wide range of individuals. As a result, in 2016, the Swiss Federal Council ordered the development of an OGD platform [30]. However, as advanced by several scholars and confirmed by the strategy results, focusing only on technical aspects was not sufficient to support the utilization of OGD. Consequently, the "Strategy OGD 2014-2018" evaluation revealed that the strategy's objectives were only partially met [54]. Despite implementation costs approaching several million, the commitment of public service actors only enabled the development of critical infrastructure elements intended to host public administration data. Following the experience of the first strategy, the Swiss Confederation revised its objectives and encouraged joint planning and harmonization of activities related to the publication of data, the creation and use of a central register of official data. In addition, the 2019-2023 strategy for developing OGD has become more active in drawing up the legal framework, improving data quality and describing metadata [54].

When the OGD publication project discussed in this paper took place in 2018, the first national OGD strategy had just ended, and the objectives of the second strategy were under discussion. In this context, the Canton of Vaud launched its digital strategy to promote innovation and position the Canton as a *data territory*. To this end, the Canton of Vaud intends to concentrate its actions on five cross-cutting and interrelated themes, including developing an open data policy. The central city of the canton naturally decided to align its digital city strategy with the objectives of the Canton. In this sense, the city decided to organize a Hackathon in May 2019 to boost the implementation and deployment of the open data policy (Figure 1). The purpose of the hackathon was to show public service employees and citizens specific applications of the OGD. For this purpose, the IT department planned the development of an OGD sample from municipal departments data, notably through the understanding of data, the identification of those that deserve interest, their description, and the search for appropriate anonymization rules. In doing so, the IT executive considered this hackathon as a preview to better understand what the OGD publishing process involves. Following a smart city policy, the housing department participated in the project. Consequently, personnel of the IT department, with the collaboration of the housing domain experts, engaged in the OGD samples preparation so that hackathon participants could use them. Given that one objective of the city was to strengthen political commitment to partnerships between its departments and the University, we joined the project as open data experts to prepare an OGD municipal sample in September 2018.

4 FINDINGS AND ANALYSIS

The triangulation technique has rapidly pinpointed that the data governance practices promoted in the literature often appear to be OGD publication challenges in ethnography. Based on these observations, we present and build our hypotheses through the three components previously identified – structural, procedural, and relational.

4.1 Structural Components

Results of the literature review show that practices such as the role and responsibilities of stakeholders as well as decision-making authorities [55] received particular attention. Clustered under the term *Structural components* and frequently called *organizational data structure* [19, 56], *standardization* [57] or *formalization [58]*, several authors stressed that the distribution of roles and responsibilities is a crucial factor of data governance [59-64]. [63] proposed the establishment of data roles and responsibilities as one of its critical success factors. According to the author, the first step to guaranteeing a successful data-driven strategy should be to establish distinct roles and responsibilities attached to any data governance activities. While literature does not provide a common understanding of roles and responsibilities [16, 65, 66], there is a consensus that IT guys cannot handle data alone (and even less published as OGD). In this sense, the literature strongly recommends establishing a global data governance office, with at least a data governance leader and data steward from different organization domains to support data-driven activities, from acquisition, management and storage to re-utilization [67].

However, the results of our ethnographic inquiry show that the role and responsibilities of the individuals involved in the project were not well defined, and data-related activities were still de facto considered IT duties rather than as a global management discipline. Although departments discussed the project's objectives and proposed some actions, we never defined or understood each actor's roles and responsibilities. For example, as data producers, the housing domain experts were committed to making data collected by their department available to the IT department to perform anonymization tasks. Nevertheless, there were no more specific tasks than "making data available" and "finding anonymization rules" assigned to project participants. Moreover, we did not notice any data steward, data officer, or decision-maker appointed to lead this project, leading to decision-making issues. While the data producer department looked skeptical about giving data access, nobody could force them to share information relative to their data. Given this lack of decision-making responsibility, we did not access databases and metadata. Thus, we have not been able to anonymize data, which considerably affected the initial project and led members of the IT department to take part in the hackathon as participants. Consequently, in this OGD publication project, the absence of roles allocation with no authority for decision-making negatively affected the project's progress and achievements.

Thus, we believe that the enforcement of structural data governance practices such as *the roles and responsibilities of a public organization and the allocation of decision-making authority* positively affect the publication of OGD. [59] explained that these components are extensively reported in data governance frameworks to foster quality, value, and suitable data reuse. Especially, the leading roles of data governance bodies such as who is responsible for data-related activities (e.g. data producer, owner, publisher), who decide processes and policies to implement (e.g. data stewardship) and who assign data activities' duties (e.g. data governance office) [58, 63, 68]. Therefore, this leads us to present three first hypotheses:

- H1: The OGD publication is positively affected by the definition of roles
- H2: The OGD publication is positively affected by the distribution of responsibilities
- H3: The OGD publication is positively affected by a decision-making authority

4.2 Procedural Components

Our findings also show that although procedural components have been widely studied in OGD barriers literature, especially, data quality standards, metadata procedures, data platforms and infrastructures as well as described as enablers of OGD sharing [12-14, 69, 70], the existence and application of clear policies, process and standards remain largely insufficient. The small data and information collected by the IT personnel confirm this idea. They were incoherent, unclear, and poor quality (e.g. reports, links to websites). When the IT department requested information or metadata, it received an Excel file like an empty survey with no explanation. Then we received what was supposed to be a relational model (i.e. how data are stored in a database) in PDF file forms. Many variables had neither names nor inscriptions to understand the file. It was just a complex mix of numbers and strange characters (e.g. X22H6T8) in thousands of boxes. One of the researchers mentioned: *"I have tried to "understand" the structures, and it seems almost impossible to make sense of the documents obtained. There is no legend, and I cannot guess what X22H6T8 means only by screening this PDF"*. When the IT department asked for more information, they finally received some extractions (i.e. PDF version of the database) with no details but a sentence explaining that they cannot do better.

We cannot say if data quality standards or more internal organization documents regarding metadata procedures exist. However, apart from the Canton's desire to appear as a "*data territory*", mentioned in the legislative agenda, we did not find any materials to plan, organize, or conduct the process of OGD publication or even a data governance program. The recruitment of external workers by the IT department to identify personal and sensitive data and investigate de-identification solutions (i.e. anonymization et pseudonymisation) for the departments tends to show that guidelines or policies for handling data are not defined. In that sense, an employee revealed that while the legislation exists on data protection, the standards, processes, and strategies to provide accurately recorded data, store them securely, foster effective reuse, and allow appropriate shareability are still scarce.

Yet, these fundamental elements for the data governance implementation play a critical role in data treatment as they stipulate actors what to do and in which domains [55, 60]. Data procedural components are widely reported in the data governance literature and put data strategies, policies, standards, processes, and issues management at the center of data governance research [16, 55, 60]. Several authors notably mentioned data processes and procedures and interpreted them as "*guidelines and rules necessary for dealing with data*" [63]. For some authors, they reflect the desired organizational behavior in terms of data quality [19, 56, 58, 64], data access [60, 71], data collection and storage [72], metadata management [56, 58, 73], data lifecycle [19, 56, 64], and data platform and architecture [19, 56, 74]. As they concern all steps of the data lifecycle, from the recording to the sharing [60, 64, 75], some authors stated that adopting procedural components ensures data management as a strategic asset [59]. For instance, a metadata strategy provides content that makes data understandable and reusable (Khatri and Brown 2010), while quality standards facilitate interoperability [69]. Therefore, as we remarked that the IT department struggled to access and comprehend the data management of the housing department, we believe that there is no consistency in the management of departmental data and no specific rules or standards for municipal data. Accordingly, we argue that the implementation of data governance procedural practices presented above may positively impact the publication of OGD. We thus present a second set of hypotheses:

- H4: OGD publication process is positively affected by standards
- H5: OGD publication process is positively affected by policies
- H6: OGD publication process is positively affected by data strategies
- H7: OGD publication process is positively affected by data process and procedure

- H8: OGD publication process is positively affected by data architecture and platform rules

4.3 Relational Components

From the beginning to the end of the project, we have never been able to identify a shared common discourse. Although we defined key objectives during the kick-off meeting - preparing an OGD sample - we rapidly understood that the strategies of both departments differed from the main one. While the IT personnel seemed more concerned with preventing confidentiality risks (i.e. by applying an appropriate level of anonymization), the housing domain experts sought to promote their activities in developing a smart city. We noticed that while the IT department strived to access and understand housing department data, domain experts focused on hackathon sponsorships and press conferences. It is interesting to note that the domain experts never took part in strategic discussions except for meetings on the hackathon's practical organization. Nevertheless, we observed a greater involvement of the housing domain experts when the project's main objective changed (i.e. when the hackathon's goal was no longer to present the OGD sample but to find solutions to the problems of the housing department). Domain experts attended meetings and actively organized the event. This study's observations echoed the findings of [76], who found that the alignment of departments frequently fails because there are no overall objectives. [77] also noted these alignment issues. The authors did not observe a universal vision or a common perception involving common expectations to complement, cooperate and contribute to organizational effectiveness in public sector managers. Yet, according to [33], effective data governance relies primarily on an alignment with business objectives and both, academic and practice-oriented literature have found that strategic alignments were crucial challenges in OGD publication [57, 60, 75].

[78] suggest that stakeholders encounter divergent interests because public sector culture stems from complex institutional cultures with different political and administrative interests. For the authors, cultural issues may impact the alignment of objectives and the behavior and attitudes of public employees regarding OGD. This was clear during the ethnography. While the housing department was not legally considered the data owner, we noticed tensions in sharing and communicating regarding data and metadata. Employees of the housing department were always speaking about "their data" and remained deaf to our demands. Consequently, we rapidly observed a strong resistance between the two departments. Although IT personnel repeated several requests for access to the housing department's data set and metadata, it took them three months to access limited and inadequate data knowledge and four months to get a PDF version of some databases. Until the end of the project, we never accessed to information needed to anonymize data. Given the housing domain experts' behavior on data matters, it did not surprise us when they first delayed, then reprogrammed, and finally cancelled sessions organized to go forward. As the IT executive later said: "As you probably noticed, the IT department suffers from a huge lack of support from other departments, and for their data, the city's departments have little respect for collaboration". However, the attitudes and behaviors expressed by the housing domain experts and IT personnel regarding the OGD publishing process go against data governance recommendations. Furthermore, they have been identified by [78] as a factor that could substantially affect the success of OGD.

While some authors have started to stress the importance of the impact of employees' data competencies [55], and knowledge [63] as well as training [16] on data governance, our findings suggest gaps in data knowledge skills and competencies. After talking to various people working for the city, we noticed that OGD and data management remain subjects ignored by public sector employees and highly controversial when people know them. While IT department personnel had been working on the project for more than six months, we were surprised when they questioned the definition of sensitive data, the usefulness of open data, and how to benefit financially. Furthermore, when the IT department requested database and metadata access, we noticed that the housing department interlocutors were unfamiliar with data management and governance vocabularies, such as metadata or databases. Yet, more and more papers discussed *self-organizing ability* [17], *organization* [60], *capacity for coherent implementation* [33] or *facilitating connections between data producers and users* [59] as an enabler to treat data as strategic assets. [16] showed that training employees involved in data processes help them to act accordingly to data policies, processes, and procedures [63, 79]. [63] put the employee data competencies as the first critical success factor. According to the authors, they have a fundamental role in the success of a data governance program.

Thus, this study produces results in line with previous works in this field and shows that the lack of relational practices such as alignment, culture, and competencies negatively impact the OGD publication process. For this reason, we add three more hypotheses to our model:

- H9: OGD publication process is negatively affected by the lack of alignment between stakeholders
- H10: OGD publication process is negatively affected by the lack of communication and culture
- H11: OGD publication process is negatively affected by the lack of knowledge and capabilities

5 CONCLUSION

Many researchers focus on distinct challenges present in the OGD ecosystem without considering that they can derive from data governance interferences. This study goes further by establishing a link between data governance components and OGD publication and shows a convergence between the application of data governance practices and the completion of data-driven processes. Although this ethnography represents local departments struggling to address OGD publication, the paper illustrates how these challenges rarely arise from the publication of OGD or the open nature of data itself, but a lack of data governance practices. Based on the triangulation approach of [40], we cross-checked our ethnographical observations with a review of the extant literature and noticed that little or no key data governance practices recommended by the literature to solve organizational data difficulties had been implemented. Therefore, we derive accurate components, specify hypotheses, and finally build the model of OGD governance (Figure 2). The model states that three structural components (H1, H2, H3), five

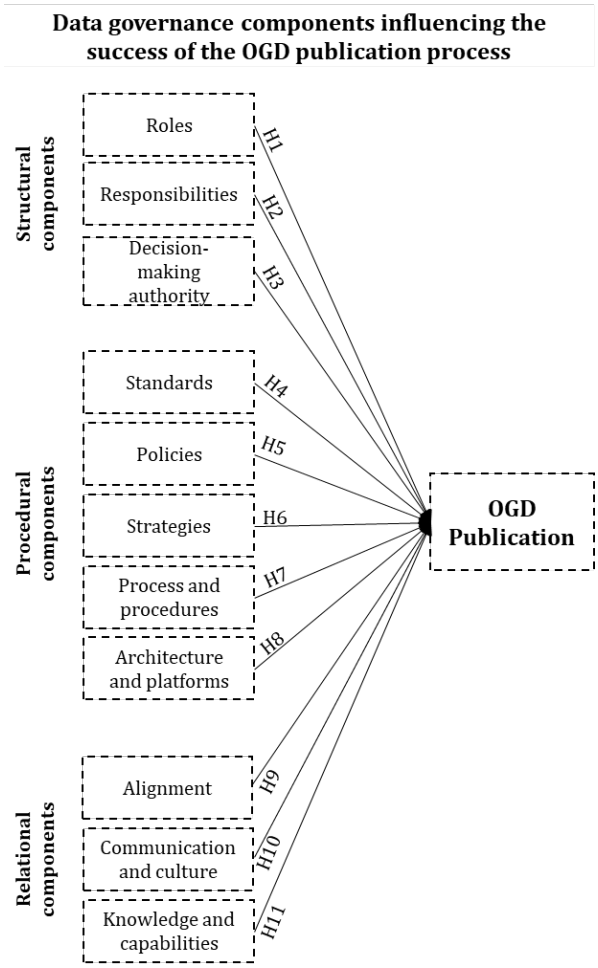

Figure 2: The model of OGD governance.

procedural components (H4, H5, H6, H7, H8) and three relational components (H9, H10, H11) seem to influence the success of the OGD publication process.

By establishing a causal relationship between data governance practices and OGD publication success, this paper extends existing OGD research and enlarges data governance literature in the public sector. It aligns with [7] research that encourages case studies to investigate how data governance is framed, which are the discourses and the consequences. In addition to providing a case to the literature on OGD publication in a context where only a few empirical studies have been started, this paper develops a first model that explains the importance of data governance on public sector data-driven activities. By cross-checking detailed observations obtained through an ethnography inquiry with a literature review, this study proposed a model that helps researchers better comprehend the OGD publication process and the dynamics that underlie data-driven activities. Furthermore, this first model appears as a starting point for OGD governance research and may form the basis for general theorization [80].

Besides extending data literature in the public sector, this study suggests ways forward by outlining avenues for practitioners. The OGD governance model illustrates and depicts the connection between data governance and OGD publication and highlights critical practices encountered by municipal departments in data-driven activities. By showing that challenges faced in OGD initiatives and, to a broader extent, data initiatives correspond to the lack of data governance, this paper seeks to help inexperienced public organizations comprehend that data-driven activities are not as simple as uploading dataset on a computer. These activities require more attention and imply a large-scale change effort at the structural, procedural, and relational levels. Furthermore, this model shows the public sector that implementing these data governance practices is of primary importance in data-driven activities and may lead to the success of, or on the contrary, its failure. As stated by the general data protection regulation (GDPR), implementing an appropriate governance strategy considerably reduces the risks of data exploitation. This is particularly important since the great majority of data owned by the public sector may be personal or sensitive. Leading media frequently speak about the infiltration of government databases to steal personal data. Yet, according to the experts, data hacking could be reduced by implementing a data governance program, including training, sufficient prominence in organizations and board support, and transparent processes and standards, among others [81]. For these reasons, we call on the departments involved in any data-related projects to establish a public-private partnership to set up a concrete data governance program and fill the lack of data capabilities.

Although our paper provides food for researchers and practitioners, additional case studies are necessary to further probe the OGD process. This study presents only one case study. Therefore, it may appear as a limitation as it does not generalize findings. We also recognize that the lack of terms homogenization may represent a limitation. The multiplicity of terms to define a similar concept led us to create clusters, which may reduce the level of detail. Finally, the highly contextualized nature of the study as well as problems of control mechanisms, do not provide the replicability of the study. Yet, this approach is common in IS literature and considered a good research practice [82]. Indeed, the case study approach may be a starting point for developing exploratory and explanatory contributions [80]. Consequently, we plan to continue expanding this paper by testing our hypothesizes through survey research conducted alongside OGD publishers. We schedule to distribute questionnaires to key employees of the most prominent Swiss and French public sector organizations involved in the publication process and then analyse the questionnaires' results through the structural equation modelling (SEM) approach. To further extend this study, we also recommend using the findings to study the complex relationships between OGD actors inside the OGD ecosystem. While previous research has pointed out the importance of OGD ecosystems, findings reveal the need to focus on symbiosis in ecosystems, i.e. "the living together of unlike organisms" [83]. Given the presented results, we call for more practice-based research regarding the data governance practices in the public sector in order to support public organizations and help them to find concrete solutions and adapt their practices to the rapid technological evolution. Also, observations that municipal departments lack data governance knowledge,

we see a need for more support using, preferably, some techniques such as formation, information, and the provision of sufficient resources (i.e. experts, communication tools).

REFERENCES

[1] Tonya Riley, T.W.P. The Cybersecurity 202: NSA Director Says Intelligence Has a Big Blind Spot: Domestic Internet Activity. 2021 [cited 2021 March 26]; Available from: https://www.washingtonpost.com/politics/2021/03/26/cybersecurity-202-nsa-director-says-intelligence-has-big-blind-spot-domestic-internet-activity/.

[2] Máchová, R., M. Hub, and M. Lnenicka, Usability Evaluation of Open Data Portals: Evaluating Data Discoverability, Accessibility, and Reusability from a Stakeholders' Perspective. Aslib Journal of Information Management, 2018. 70(3): p. 252-268.

[3] Matos, U.C. and J. Corbett. Creating Knowledge for Value Creation in Open Government Data Ecosystems. in Proceedings of the 25th Americas Conference on Information Systems. 2019. Cancún, Mexico.

[4] Vancauwenberghe, G. and B. van Loenen, Governing Open Spatial Data Infrastructures: The Case of the United Kingdom, in Governance Models for Creating Public Value in Open Data Initiatives. 2019, Springer: Cham, Switzerland. p. 33-54.

[5] Le Parlement Suisse. Réutilisation des Données Personnelles et Techniques Collectées et Gérées par les Administrations Publiques et les Institutions qui en sont Proches. 2021 [cited 2021 March 30]; Available from: https://www.parlament.ch/fr/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20204547.

[6] Yiannakoulias, N., *et al.*, Open Government Data, Uncertainty and Coronavirus: An Infodemiological Case Study. Social Science & Medicine, 2020. 265: p. 113549.

[7] Nielsen, O.B., A Comprehensive Review of Data Governance Literature. Selected Papers IRIS, 2017. 8: p. 120-133.

[8] Seydtaghia, A. and M. Farine. La Suisse, Véritable Cancre Numérique: Sept Exemples Marquants. 2021 [cited 2021 March 30]; Available from: https://www.letemps.ch/economie/suisse-veritable-cancre-numerique-sept-exemples-marquants.

[9] Trom, L. and J. Cronje. Analysis of Data Governance Implications on Big Data. in The proceedings of the Future of Information and Communication Conference. 2019. San Francisco, USA.

[10] Lv, H. and H. Ma, Performance Assessment and Major Trends in Open Government Data Research Based on Web of Science Data. Data Technologies and Applications, 2019. 53(3): p. 286-303.

[11] Afful-Dadzie, E. and A. Afful-Dadzie, Liberation of Public Data: Exploring Central Themes in Open Government Data and Freedom of Information Research. International Journal of Information Management, 2017. 37(6): p. 664-672.

[12] Benitez-Paez, F., *et al.*, Roadblocks Hindering the Reuse of Open Geodata in Colombia and Spain: A Data User's Perspective. ISPRS International Journal of Geo-Information, 2018. 7(1): p. 1-6.

[13] Barry, E. and F. Bannister, Barriers to Open Data Release: A View from the Top. Information Polity, 2014. 19(1, 2): p. 129-152.

[14] Janssen, M., Y. Charalabidis, and A. Zuiderwijk, Benefits, Adoption Barriers and Myths of Open Data and Open Government. Information Systems Management, 2012. 29(4): p. 258-268.

[15] Zuiderwijk, A. and M. Janssen, Barriers and Development Directions for The Publication and Usage of Open Data: A Socio-technical View, in Open government. 2014, Springer. p. 115-135.

[16] Abraham, R., J. Schneider, and J. Vom Brocke, Data Governance: A Conceptual Framework, Structured Review, and Research Agenda. International Journal of Information Management, 2019. 49: p. 424-438.

[17] Welle Donker, F. and B. van Loenen, How to Assess the Success of the Open Data Ecosystem? International Journal of Digital Earth, 2017. 10(3): p. 284-306.

[18] El-Emam, K., *et al.* A Policy Forum on the Use of Big Data in Health Care. 2013 [cited 2021 Septembre, 13]; Available from: https://bipartisanpolicy.org/download/?file=/wp-content/uploads/2019/03/Use-of-Big-Data-in-Health-Care.pdf.

[19] Soares, S., The IBM Data Governance Unified Process: Driving Business Value with IBM Software and Best Practices. 2010: MC Press, LLC.

[20] Ndamase, Z., The Impact of Data Governance on Corporate Performance: The Case of a Petroleum Company. 2014, University of Cape Town, South Africa.

[21] Lämmerhirt, D. Avoiding Data Use Silos. How Governments Can Simplify The Licensing Landscape. 2017.

[22] Benfeldt, O., J.S. Persson, and S. Madsen, Data Governance as a Collective Action Problem. Information Systems Frontiers, 2020. 22(2): p. 299-313.

[23] Kitchin, R. and N. Thrift, International Encyclopedia of Human Geography. 2009, London: Elsevier.

[24] Hancock, D.R. and B. Algozzine, Doing Case Study Research: A Practical Guide for Beginning Researchers. 2006, Columbia University, New York, NY: Teachers College Press

[25] Barbour, A., Exploring Some Ethical Dilemmas and Obligations of the Ethnographer. Ethnography and Education, 2010. 5(2): p. 159-173.

[26] Fusch, P.I., G.E. Fusch, and L.R. Ness, How to Conduct a Mini-ethnographic Case Study: A Guide for Novice Researchers. The Qualitative Report, 2017. 22(3): p. 923.

[27] Holloway, I. and C. Daymon, Qualitative Research Methods in Public Relations and Marketing Communications. 2010, London, UK: Routledge.

[28] Noor, K.B.M., Case Study: A Strategic Research Methodology. American Journal of Applied Sciences, 2008. 5(11): p. 1602-1604.

[29] Ville de Lausanne. Agenda et Actualité. 2020 [cited 2020 November, 17]; Available from: https://www.lausanne.ch/.

[30] Confédération Suisse. Opendata.swiss Homepage. 2020 [cited 2020 August 20, 2020]; Available from: https://opendata.swiss/en/.

[31] Etat de Vaud. Strategie Numérique du Canton de Vaud. 2018 [cited 2021 August 3,]; Available from: https://www.vd.ch/strategie-numerique-du-canton-de-vaud-digitale-strategie-des-kantons-waadt/.

[32] OECD, Data Governance in the Public Sector, in The Path to Becoming a Data-driven Public Sector. 2019: OECD Publishing, Paris. p. 23-57.

[33] OECD, Digital Government Review of Argentina. 2019, OECD Publishing, Paris: OECD Digital Government Studies.

[34] Robson, C., Real World Research: A Resource for Social Scientists and Practitioner-Researchers. 2002, Oxford, UK: Wiley-Blackwell.

[35] Braun, V. and V. Clarke, Using Thematic Analysis in Psychology. Qualitative Research in Psychology, 2006. 3(2): p. 77-101.

[36] Vaismoradi, M., H. Turunen, and T. Bondas, Content Analysis and Thematic Analysis: Implications for Conducting a Qualitative Descriptive Study. Nursing & Health Sciences, 2013. 15(3): p. 398-405.

[37] Figgou, L. and V. Pavlopoulos, Social Psychology: Research Methods. International Encyclopedia of the Social & Behavioral Sciences, 2015. 22: p. 544-552.

[38] Mayring, P., Qualitative Content Analysis. A Companion to Qualitative Research, 2004. 1(2): p. 159-176.

[39] Eisenhardt, K.M. and M.E. Graebner, Theory Building from Cases: Opportunities and Challenges. Academy of Management Journal, 2007. 50(1): p. 25-32.

[40] Flick, U., Triangulation in qualitative research. A companion to qualitative research, 2004. 3: p. 178-183.

[41] Erzberger, C., *et al.*, Handbook of Mixed Methods in Social and Behavioral Research. Handbook of Mixed Methods in Social and Behavioral Research, 2003: p. 457-90.

[42] Nguyen, L., Research Methodology, in Chandos Asian Studies Series,Guerilla Capitalism, L. Nguyen, Editor. 2009, Woodhead Publishing Limited. p. 29-39.

[43] Williamson, C., Ethnographic Research, in Research Methods: Information, Systems, and Contexts. 2013, Tilde University Press. p. 287-309.

[44] Nightingale, A.J., Triangulation, in International Encyclopedia of Human Geography (Second Edition), A. Kobayashi, Editor. 2020, Elsevier. p. 477-480.

[45] Heath, L., Triangulation: Methodology, in International Encyclopedia of the Social & Behavioral Sciences, J.D. Wright, Editor. 2015, Elsevier. p. 639-644.

[46] Barnum, C.M., Analyzing the Findings, in Usability Testing Essentials (Second Edition), C.M. Barnum, Editor. 2021, Morgan Kaufmann. p. 287-319.

[47] Strauss, A. and J. Corbin, Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory 2014: Sage publications.

[48] Yin, R.K., Discovering the Future of the Case Study. Method in Evaluation Research. Evaluation Practice, 1994. 15(3): p. 283-290.

[49] Swiss Confederation. E-governement Strategy. 2018 August 14, 2018]; Available from: https://www.egovernment.ch/index.php/download_file/force/440/3614/.

[50] Federal Office of Communications. Digital Switzerland Strategy. 2018 [cited 2018 August 14, 2018]; Available from: https://www.bakom.admin.ch/bakom/en/homepage/digital-switzerland-and-internet/strategie-digitale-schweiz/strategy.html.

[51] Federal Council. Open Government Data Strategy for Switzerland 2014 - 2018. 2014; Available from: https://www.egovernment.ch/index.php/download_file/force/761/3631/.

[52] Robinson, D.G., *et al.*, Government Data and the Invisible Hand. Yale Journal of Law & Technology, 2009. 11: p. 160-175.

[53] Kalampokis, E., E. Tambouris, and K. Tarabanis, A Classification Scheme for Open Government Data: Towards Linking Decentralised Data. International Journal of Web Engineering and Technology, 2011. 6(3): p. 266-285.

[54] Federal Statistical Office. Open Government Data strategy 2019–2023. 2021 May 21, 2021]; Available from: https://www.bfs.admin.ch/bfs/fr/home/services/ogd/strategie.html.

[55] Alhassan, I., D. Sammon, and M. Daly, Data Governance Activities: A Comparison Between Scientific and Practice-oriented Literature. Journal of Enterprise Information Management, 2018. 31(2): p. 300-316.

[56] Marchildon, P., *et al.*, Data Governance Maturity Assessment Tool: A Design Science Approach. Projectics/Proyéctica/Projectique, 2018(2): p. 155-193.

[57] Kim, H.Y. and J.-S. Cho, Data Governance Framework for Big Data Implementation with NPS Case Analysis in Korea. Journal of Business and Retail Management Research, 2018. 12(3): p. 36-46.

[58] Stanford University's Data Governance Office, Data Governance Maturity Model. 2013.

[59] Spy Pond Partners, L. and I. Iteris, Data to Support Transportation Agency Business Needs: A Self-assessment Guide. 2015: Transportation Research Board.

[60] Panian, Z., Some Practical Experiences in Data Governance. World Academy of Science, Engineering and Technology, 2010. 62(1): p. 939-946.
[61] Cheong, L.K. and V. Chang. The Need for Data Governance: A Case Study. in Proceedings of the 18th Australasian Conference on Information Systems. 2007. University of Southern Queensland, Toowoomba, Australia.
[62] Weber, K., Otto, B., Österle, H., One Size Does Not Fit All - A Contingency Approach to Data Governance. ACM Journal of Data and Information Quality, 2009. 1(1): p. 1–27.
[63] Alhassan, I., D. Sammon, and M. Daly, Critical Success Factors for Data Governance: A Theory Building Approach. Information Systems Management, 2019. 36(2): p. 98-110.
[64] Khatri, V. and C.V. Brown, Designing Data Governance. Communications of the ACM, 2010. 53(1): p. 148-152.
[65] Otto, B., Organizing Data Governance : Findings from the Telecommunications Industry and Consequences for Large Service Providers Communications of the Association for Information Systems, 2011. 29(1): p. 45–66.
[66] Dreibelbis, A., Enterprise Master data Management: An SOA Approach to Managing Core Information. 2008, India: IBM Press. 87.
[67] Tallon, P.P., Corporate Governance of Big Data: Perspectives on Value, Risk, and Cost. Computer, 2013. 46(6): p. 32-38.
[68] Yang, L., *et al.*, Towards Big Data Governance in Cybersecurity. Data-Enabled Discovery and Applications, 2019. 3(1): p. 1-12.
[69] Marmier, A. and T. Mettler, Developing an Index for Measuring OGD Publisher Compliance to Good Practice Standards: Insights from Opendata. swiss. Information Polity, 2020(Preprint): p. 1-20.
[70] Martin, S., *et al.* Open data: Barriers, Risks and Opportunities. in Proceedings of the 13th European Conference on eGovernment. 2013. Insubria Varese, Italy.
[71] Brackett, M. and P.S. Earley, The DAMA Guide to The Data Management Body of Knowledge. 2009.
[72] Khatcherian, S. and G. Jefferson, Information Governance: Challenges for Financial Services Firms Operating in the Information Age. Convergence, 2009. 5: p. 80.
[73] CMMI. The Capability Maturity Model Integration. 2021 [cited 2021 November 8,]; Available from: https://cmmiinstitute.com/dmm.
[74] Thomas, M.A., *et al.*, Data Management Maturity Assessment of Public Sector Agencies. Government Information Quarterly, 2019. 36(4): p. 101401.
[75] Borgman, C.L., Big Data, Little Data, No Data: Scholarship in the Networked World. 2015, Cambridge, MA: MIT press. 416.
[76] Vitale, M.R., B. Ives, and C.M. Beath. Linking Information Technology and Corporate Strategy: An Organizational View. in ICIS 1989 Proceedings. 1986.
[77] Marmier, A. and T. Mettler, Different Shades of Perception: How Do Public Managers Comprehend the Re-use Potential of Open Government Data?, in Proceedings of the 41th International Conference on Information System. 2020: Hyderabad, India.
[78] Rusu, L. and G.M. Jonathan, IT Alignment in Public Organizations: A Systematic Literature Review, in Information Technology Governance in Public Organizations. 2017, Springer, Cham. p. 27-57.
[79] Randhawa, T.S., Incorporating Data Governance Frameworks in the Financial Industry, in Walden University 2019: Walden Dissertations and Doctoral Studies.
[80] Recker, J., Scientific Research in Information Systems: A Beginner's Guide. 2012, Berlin: Springer Science & Business Media.
[81] Roussel, A., Notre si Précieuse Intégrité Numérique, in L'invité de La Matinale, A. Roussel, Editor. 2021.
[82] Gauld, R., Public Sector Information System Project Failures: Lessons from a New Zealand Hospital Organization. Government Information Quarterly, 2007. 24(1): p. 102-114.
[83] de Bary Heinrich Anton "Ueber Symbiose" [On Symbiosis] Tageblatt für die Versammlung Deutscher Naturforscher und Aerzte (in Cassel) [Daily Journal for the Conference of German Scientists and Physicians] (in German), 1878. 51: p. 121–126.