



UNIL | Université de Lausanne

Unicentre

CH-1015 Lausanne

<http://serval.unil.ch>

---

Year : 2023

## Open Government Data: Holy Grail or Pandora's Box ?

Marmier Auriane

Marmier Auriane, 2023, Open Government Data: Holy Grail or Pandora's Box ?

Originally published at : Thesis, University of Lausanne

Posted at the University of Lausanne Open Archive <http://serval.unil.ch>

Document URN : urn:nbn:ch:serval-BIB\_D6F08C554BB06

### **Droits d'auteur**

L'Université de Lausanne attire expressément l'attention des utilisateurs sur le fait que tous les documents publiés dans l'Archive SERVAL sont protégés par le droit d'auteur, conformément à la loi fédérale sur le droit d'auteur et les droits voisins (LDA). A ce titre, il est indispensable d'obtenir le consentement préalable de l'auteur et/ou de l'éditeur avant toute utilisation d'une oeuvre ou d'une partie d'une oeuvre ne relevant pas d'une utilisation à des fins personnelles au sens de la LDA (art. 19, al. 1 lettre a). A défaut, tout contrevenant s'expose aux sanctions prévues par cette loi. Nous déclinons toute responsabilité en la matière.

### **Copyright**

The University of Lausanne expressly draws the attention of users to the fact that all documents published in the SERVAL Archive are protected by copyright in accordance with federal law on copyright and similar rights (LDA). Accordingly it is indispensable to obtain prior consent from the author and/or publisher before any use of a work or part of a work for purposes other than personal use within the meaning of LDA (art. 19, para. 1 letter a). Failure to do so will expose offenders to the sanctions laid down by this law. We accept no liability in this respect.



UNIL | Université de Lausanne

---

FACULTÉ DE DROIT, DES SCIENCES CRIMINELLES ET  
D'ADMINISTRATION PUBLIQUE

INSTITUT DE HAUTES ÉTUDES EN ADMINISTRATION PUBLIQUE  
(IDHEAP)

Open Government Data: Holy  
Grail or Pandora's Box?

THÈSE DE DOCTORAT

présentée à la

Faculté de droit, des sciences criminelles et d'administration  
publique

de l'Université de Lausanne

pour l'obtention du grade de

Docteur en administration publique

par

Auriane Marmier

Directeur de thèse

Prof. Tobias Mettler, Université de Lausanne

Jury

Prof. Christine Legner, University of Lausanne

Prof. Gianluca Miscione, University of Dublin

Prof. Peter Rohner, University of St-Gallen

LAUSANNE

2023



UNIL | Université de Lausanne

IDHEAP  
Institut de hautes études  
en administration publique

## Imprimatur

Vu le rapport présenté par le jury de thèse, composé de

Prof. Dr. <b>Tobias Mettler</b>	Directeur de thèse
Prof. Dr. <b>Christine Legner</b>	Autre membre du corps professoral
Prof. Dr. <b>Gianluca Miscione</b>	Expert externe
Prof. Dr. <b>Peter Rohner</b>	Expert externe

le Décanat de la Faculté de droit, des sciences criminelles et d'administration publique autorise l'impression de la thèse de

**Madame Auriane Marmier**

intitulée

**Open Government Data: Holy Grail or Pandora's Box?**

Lausanne, le 16 janvier 2023

Décanat de la Faculté de droit,  
des sciences criminelles et  
d'administration publique



Professeur Nils Soguel  
Vice-doyen

*Tous les hommes pensent que le bonheur se trouve au sommet de la montagne, alors qu'il réside  
plutôt dans la façon de la gravir.*

*Confucius*

# **Open Government Data: Holy Grail or Pandora's Box?**

Auriane Marmier

## Contents

Synopsis .....	4
Publication .....	36
Dance your PhD Contest.....	36
Paper # 1 .....	37
Paper # 2 .....	57
Paper # 3 .....	74
Paper # 4 .....	87
References.....	103

## *Synopsis*

## Synopsis

Open Government Data: Holy Grail or Pandora's Box? .....	6
Introduction.....	6
Background.....	8
Open Government Data.....	8
Platform Definitions: platform vs ecosystem .....	9
Platform models.....	10
Platform operation.....	11
Platform traps .....	12
The Chicken-or-egg problem .....	12
The Penguin dilemma.....	13
Emergence.....	13
Data Governance .....	13
Research approach .....	14
Research aims .....	14
Research philosophy.....	14
Research questions .....	15
Research design .....	16
Research context.....	18
Research procedure & articles summary .....	19
Paper 1 .....	19
Paper 2 .....	21
Paper 3 .....	22
Paper 4.....	24
Discussion.....	26
Platform traps .....	27
Data and OGD platform governance issues.....	29
Implications.....	32
Practitioners.....	32
Academic research.....	33
Limitations and further research .....	34



# Open Government Data: Holy Grail or Pandora's Box?

## Introduction

With the recent increase in social media, the ever-faster information technology developments, and the race for innovation, the volume of data is continuously growing. While digital-native giants such as Facebook and Uber built their prosperity by collecting and processing data, numerous data are also generated, collected, and stored through companies' or governments' daily activities' (Ciuriak 2018). Such is the case for many public organisations that gather and store vast quantities of data to complete various tasks (Magalhaes and Roseira 2017; Pollock 2011). Over the past few years, public organisations have by far been the richest holder of data (Jetzek et al. 2019; Máchová and Lnénicka 2017; Zhao and Fan 2018), holding archives of historical and contemporary data in their hands (Sandoval-Almazán et al. 2017). Given the economic potential of data, many governments perceive them as resources of strategic relevance (Bates 2014; Munné 2016). Data are frequently employed to study patterns and trends, serve as strategic assets to improve internal processes or as a springboard for innovation (Davenport 2013; Gelhaar and Otto 2020). They are unique assets that can be leveraged simultaneously, several times and without value depreciation (MIT Technology Review Custom 2016). Traditionally to monetize data, i.e. convert intangible values of data into real value, famous companies sell them (Najjar and Kettinger 2013). Yet data valorisations may also take numerous forms, such as business analytics or personal advertisement, which may improve decisions-making (Jetzek et al. 2019; Moro Visconti et al. 2017; Piovesan 2017).

In light of this, the public sector has started to develop open government data (OGD), promising to lead stakeholders to re-use data to develop new businesses and propose innovative solutions (Magalhaes and Roseira 2017). OGD are data produced by state bodies which are freely accessible, modifiable, sharable, and usable by anyone for any purpose (Open Knowledge Foundation 2018b). Perceived as a fundamental pillar of modern democracy (Kassen 2018), it was after Obama's Administration's efforts to promote transparency and civic engagement that action took place, and open government and OGD initiatives started to be implemented (Sandoval-Almazán et al. 2017). According to Tauberer (2014), "*technologists saw the important role of government data in democracy and started the trend of civic hacking, which often involves using government data to make governments more accountable*". The development of OGD initiatives has also been perceived as a driver to help public organisations offer more transparency and improve citizen participation but also reduce their cost and help them deliver better public services (Manyika et al. 2013; Vickery 2011; Zeleti et al. 2016; Zuiderwijk and Janssen 2014b). In 2013, Switzerland estimated OGD annual economic benefits between CHF 0.9 and CHF 1.2 billion (Bürgi-Schmelz 2013). According to the European Commission (2021) the 27 EU countries and the United Kingdom expect an increase in the economic value of public data from €52 billion in 2018 to €194 billion in 2030. Thus, the forecast of billions of annual benefits and cost savings make the OGD the founder and a driver for economic growth, innovation, competitiveness and societal progress (Francey 2021).

With an expected increase in worldwide data creation from 4.4 zettabytes in 2013 to 44 zettabytes in 2020 (Abraham et al. 2019) and intense pressure from OGD advocates, several countries dedicated funds and resources to increase OGD accessibility (Nikiforova and McBride 2021). Hence, governments started to make public sector information available by building OGD platforms (Magalhaes and Roseira 2017; Nikiforova and McBride 2021). Since 2009, many governments have adopted OGD initiatives, increasing between 2 to more than 300. Until 2013, more than 280 OGD catalogues have been created, and more than one million datasets have been published on OGD platforms (Máchová et al. 2018; Matos and Corbett 2019). A recent study highlighted eight countries with an average number of datasets published superior to 20'000. At the top of the ranking, Tang and Jiang (2021) identified the USA and India with respectively an average of 276'951 and 140'698 datasets published. Accordingly, facilitating open access to public data

proliferated during the last decade and became a major government management activity (Máchová et al. 2018; Matos and Corbett 2019). Thus, *opening public sector data* became a strong motivator for public organizations to show their willingness to be transparent (Constantinides et al. 2018).

Therefore, Switzerland, similarly to several countries, launched its first OGD Strategy in 2014 (E-Government Switzerland 2014). Following a planning horizon for 2014-2018, the strategy aimed to develop a centralized platform that facilitates access to public data. The strategy was expected to foster the sharing and re-use of government-produced data by improving access to public data (E-Government Switzerland 2014). Thus, intending to cluster public data and encourage the Swiss authorities (i.e. the Confederation, cantons, communes, and other organizations), the Federal Archives built the *opendata.swiss platform* in 2016 (E-Government Switzerland 2016; Federal Statistical Office 2020a). After two years, the Federal Council reiterated the Swiss OGD strategy for 2019 to 2023 (Federal Statistical Office 2020a). Closely aligned with the previous strategy 2014-2018, the new strategy focuses on improving OGD re-use (Federal Statistical Office 2020a), encouraging government agencies to publish more, and paying particular attention to the coordination of data publication (Conseil fédéral 2018). Yet, like in other countries (Jetzek et al. 2014; Nikiforova and McBride 2021; Sandoval-Almazán et al. 2017), Switzerland remains stuck on the idea that the re-use and value creation necessitate sharing a considerable quantity of data and faces many challenges in producing valuable data and quality (Swiss Federal Audit Office 2019).

Although thousands of initiatives have been launched to achieve the goals of OGDs, previous research indicates that the re-use of OGDs is lagging behind (Attard et al. 2015; Dawes et al. 2016; Janssen et al. 2012; Krasikov et al. 2020; Safarov et al. 2017). According to Bonina et al. (2018), realizing the economic and social benefits of OGD is rarely easy to obtain. In particular, the authors blame a lack of resources and skills for leveraging datasets further (Gascó-Hernández et al. 2018; Martin 2014; Ruijter et al. 2017). It is therefore not surprising that the Swiss Federal Audit Office leading at the end of the first Swiss OGD strategy reveals several shortfalls in governance and the utility of the national OGD platform. Moreover, contrary to the expected forecast, the Audit discovered that no economic benefit had been established (Swiss Federal Audit Office 2018). According to the literature, simply creating platforms and depositing government data on them is insufficient to produce valuable and reusable OGD and even less to generate their re-use (Gascó-Hernández et al. 2018). As stated by Safarov et al. (2017), it is not as simple as *drag and drop*. The re-use of OGD necessitates several technological and social pre-conditions. The Open Data Barometer argues that there is still a long way to go for datasets published on OGD platforms to be accessible and re-usable (Yang and Wu 2016). Similarly, Hervé Bourlard (i.e. the director of the not-for-profit research institute), declared “*il faudra beaucoup de temps pour que le traitement des données récoltées en masse fournisse vraiment des informations*” (Parvex 2013).

Given the complexity of OGD, it has become a relevant and significant topic of study during the last decade. Academic interest started in 2011 with the first empirical study dedicated to OGD and significantly increased between 2014 and 2020. With the rise of dedicated conferences in 2016 and 2017, it was during this period that the number of OGD-related publications exploded (Safarov et al. 2017). Francey and Mettler (2021) and Ubaldi (2013) cluster the OGD process according to three major phases (1) data production, (2) data distribution and (3) data re-use. Accordingly, the OGD literature investigates OGD challenges and difficulties through these different lenses (Bonina and Eaton 2020; Dahbi et al. 2018; Danneels et al. 2017; Dawes et al. 2016; Máchová et al. 2018; Susha et al. 2015a; Zuiderwijk 2015). Surprisingly, although the literature portrays OGD as the output of digital platforms (i.e. models, technologies or tools facilitating interactions among potential users, intermediaries and suppliers by sharing available products or services (Cunningham Dahl-Jørgensen and Parmiggiani 2020), few authors question the concept of the platform (Bonina and Eaton 2020; Danneels et al. 2017). Except Danneels et al. (2017), who followed the platform literature to identify the different types of OGD platform ecosystems, and Bonina and Eaton (2020), who investigated how the demand and the supply side are governed to foster a platform ecosystem development, the OGD as the product or service of a digital platform had never been questioned. Similarly to Uber or Airbnb, which propose listing services' (Parker et al. 2016), the overarching aim of OGD platforms is to

foster datasets exchange by facilitating interaction between data suppliers and OGD re-users. Therefore, data are the fundamental resource that drives these kinds of platforms (Srnicek 2017), and the opening of public data becomes thus a precondition to the OGD re-use. Consequently, applying platforms' mechanisms and how the public sector apprehends them appear as a path to explore to increase the current low re-use of OGD and thus advance in the OGD goals achievement.

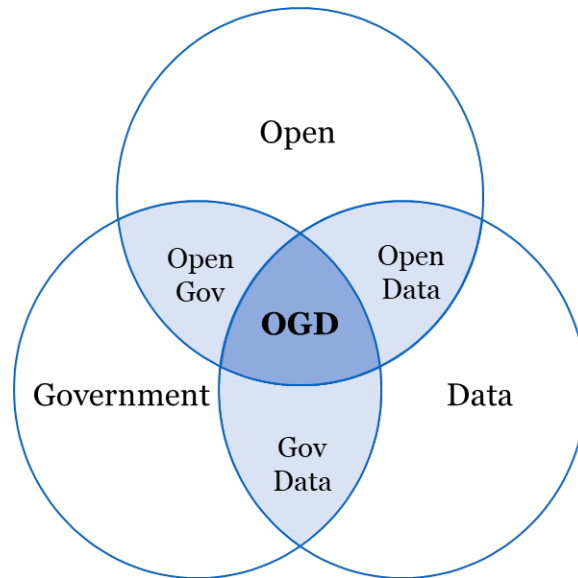
Consequently, to advance in the OGD goals achievement, this dissertation seeks to better understand the mechanisms underlying the functioning of a platform and their effect on the opening public data process. Drawing on the platform's theory and the fact that several traps may strongly impact the platform's success, this dissertation mainly seeks to observe if data providers fall into platforms' traps during the opening process of public data and, if so, the reasons why. In doing so, this dissertation seeks to better understand the reasons that prevent OGD Swiss platform data providers from opening public data. As this project is meant to be a cumulative thesis, this dissertation first presents and explains some theoretical aspects of the research and investigates major concepts relative to OGD and digital platforms. Then, the dissertation provides information regarding the research design and questions, describing the storytelling of the four articles selected for this thesis. After presenting the papers, this dissertation discusses the implications of this research for practitioners and sciences. To conclude, this thesis provides some guidelines for public sector organizations seeking to expand and improve their OGD productions.

## **Background**

Since this thesis aims to better understand the Swiss OGD platform and the barriers met by public sector organizations in the opening process of public data, we first provide a description of OGD and an overview of OGD barriers research. We then review the digital platform stream to motivate and position this thesis.

### ***Open Government Data***

Although many countries have pushed open data strategies over the past two decades, it is not new. According to Bates (2012), the concept of OGD for civil society and business communities can be traced back to the 1970s. The author examines OGD in the United Kingdom in detail and explains how the re-use of Public Sector Information (PSI) in Europe is *the precursor to OGD*. Other scholars traced the historical origins and development of OGD in the constitutional right-to-know act (Luna-Reyes et al. 2014). According to the authors, new regulations and policies from other countries could explain the evolution of the constitutional right-to-know act. For Tauberer (2014), OGD originates in open access to laws. The author defines OGD as a movement and a "*small part of the broader open government movement that includes both the classic open government movement (such as the Freedom of Information Act) and the newer areas of citizen participation and citizen experience*" (Tauberer 2014). Like others, the authors noticed that the passage of the Freedom of Information Act (FOIA) in the United States was key to the development of OGD, concluding that the rise of new information and communication technologies marked a crossroads in the open government movement (Luna-Reyes et al. 2014; Robinson et al. 2009). Then the movement hit other milestones, such as the Sunlight Foundation's creation and the Ten Principles for Open Government Data. Finally, the actual definition of OGD was finalized with the Open Government Directive, which focuses on the three main principles of Open Government-transparency, participation, and collaboration (European Commission 2020; Swiss Confederation 2018), which Gonzalez-Zapata and Heeks (2015) summarized as the combination of three simple concepts: open government, open data, and government data (Fig. 1).



**Figure 1. OGD adapted from Gonzalez-Zapata and Heeks (2015)**

In addition to Gonzalez-Zapata and Heeks (2015), who saw a bureaucratic dimension to OGD (i.e. improving information quality and public sector efficiency), other authors also found technological, political, and economic dimensions to OGD (Wirtz et al. 2022). McBride et al. (2019), Zhao and Fan (2018), Berrone (2017), and others see OGD as a source for developing economic value. McNutt et al. (2016) and Meijer (2015) agreed that OGD opens the door to technological innovation in the public sector, and Zhao and Fan (2018) and Meijer (2015) also see OGD as a way for governments to fulfil their responsibilities to citizens, particularly by providing transparent insights and encouraging citizen participation. Given these different perceptions, Wirtz et al. (2022) identified six key themes explored in the OGD literature, including (1) the overall evolution of OGD, (2) barriers and drivers, (3) adoption, use, and implementation, (4) success and performance, (5) acceptance and satisfaction, and finally (6) policy, regulation, and law. The authors found that topics (3) adoption, use, and implementation, and (4) success and performance accounted for nearly 60% of OGD studies, while studies on OGD barriers and drivers accounted for only 15% and the other topics accounted for less than 25%.

In terms of barriers preventing OGD strategies, findings from previous studies distinguish three types of barriers arising from the OGD ecosystem-institutional, technical, and societal (Barry and Bannister 2014; Janssen et al. 2012; Ruijter et al. 2017). Institutional barriers refer to difficulties arising from the lifecycle of public data (i.e. collection, storage and sharing of data) (Yang et al., 2015; Barry & Bannister, 2014), such as the complexity of government structures or the lack of public sector employees skills related to data-driven activities (Ruijter et al. 2017; Welch et al. 2016; Yang et al. 2015b; Young 2020). When considering technical barriers, studies focus either on data infrastructure characteristics such as data quality, format, or interoperability or on platform issues (i.e. lack of support, poor architecture) (Barry and Bannister 2014; Janssen et al. 2012; Smith and Sandberg 2018). In the third category, which examines users' ability to re-use OGD, the most commonly identified barriers relate to users' inability to interact with the platform or data and their lack of knowledge (Barry and Bannister 2014; Ruijter et al. 2017) or non-use of data published on the platform (Janssen et al. 2012; Smith and Sandberg 2018) (Smith & Sandberg, 2018; Janssen et al., 2012).

### ***Platform Definitions: platform vs ecosystem***

Platform (business) models are ubiquitous in today's industries and are operated by more than half of the top ten companies - Apple Inc, Microsoft Corp, Alphabet Inc (formerly Google), Amazon.com Inc, and Meta Platform Inc (formerly Facebook) are notably among them (Parker et al. 2016). The most popular

platforms notably include social media (e.g. Facebook and LinkedIn), mobile operating systems (e.g. Android and iOS), payment systems (e.g. PayPal and Apple Pay), and peer-to-peer (e.g. Uber and Airbnb) (Blaschke et al. 2019). The term platform is merely used as a label to describe government-social media interaction (Hong and Kim 2016), cloud computing (S. Paquette 2012), general open government research (Nam, 2012), and to characterise an object such as a portal or a national or local website (Bonina and Eaton 2020; S. Ganapati 2018). The platform is also understood as a digital tool that enables value-added interactions between providers and consumers (Parker et al. 2016), a set of resources that foster service innovation through increased stakeholder participation (Cunningham Dahl-Jørgensen and Parmiggiani 2020), and an intermediary between two or more markets (Thomas et al. 2014) or a technology-based business system centre (Gawer 2008; Gawer and Cusumano 2002). While some authors understand digital platforms as complex supply chains that bring together buyers and sellers of products or services (i.e. a business model) (Parker et al. 2016; Tiwana 2013), Gawer (2010) defines a platform as *"a building block that could be a product, service, or technology that serves as a foundation upon which other organisations can develop complementary products, services, or technologies"*. For Wajcman (2014) and Pasquale (2016), a platform is an organisational form and digital or platform capitalism strategy.

Last but not least, Andersson Schwarz (2017) described a platform as *"infrastructural arrangements that situate digital operability on proprietary systems that are, to some degree, programmable and/or customizable by the system users, making possible one- or multisided market exchanges"*. In doing so, the author emphasises platform's programmable and customizable nature and the concept of a platform's infrastructure and architecture. This is especially important because, contrary to what can be found in the literature, the concept of open data platforms is often confused with the concept of an open data ecosystem. Although the two concepts are very similar, they are distinct. Unlike a platform, the ecosystem concept does not rely on a customizable infrastructure. Biologically, an ecosystem is described as *"evolving systems that are dynamic, constantly reforming and responding to natural disturbances and competition between and among species"* (Immonen et al. 2016). Based on this idea, James Moore adapted the concept of ecosystems in business environments to describe new forms of economic activity that enable the production of valuable goods and services for customers (Aarikka-Stenroos and Ritala 2017). From a social science perspective, the idea has been retained, and an open data ecosystem (ODE) has been described as *"a complex multidimensional variety of interdependent actors that create dynamic multidirectional and multi-institutional exchanges through a variety of activities such as collection, management, re-use, and connection to online services"* (Corbett et al. 2018; Ding et al. 2011). Thus, in this dissertation, we understand an open data platform as part of the open data ecosystem and as an adaptable infrastructure, tool, or business model that facilitates the achievement of these goals.

### ***Platform models***

The literature shows that there are many different conceptions and visions of platforms. Asadullah et al. (2018) examined 96 articles from IS and management journals to conduct a systematic literature review on digital platforms. The study's results revealed considerable ambiguity in conceptualizing the digital platform concept. While the authors noted that previous research on this topic classified platforms according to the following dimensions - business model, interaction mode, governance mode, and ownership structure (Asadullah et al. 2018), Porter (2004) took a different view. The author ranked digital platforms according to the range of actors and organizations that operate them. (Faulkner-Gurstein and Wyatt 2021). In this vein, Kordzadeh and Warren (2013) created a typology with four types that ordered platforms according to their structural connection between consumers and providers. The typology includes platforms from professionals to professionals (P2P), from professionals to consumers (P2C), from consumers to professionals (C2P), and from consumers to consumers (C2C) (Faulkner-Gurstein and Wyatt 2021). Another interesting way of classifying digital platforms is based on the function of the platforms. Srnicek (2017) identified five different functions: Advertising (e.g. Google, which collects data to better target ads), cloud platforms (e.g. Salesforce, which provides networks of software and hardware), industry (e.g., Siemens, which creates infrastructures), manufacturing (e.g. Spotify which rents out its assets), and lean

platforms (e.g. Uber, which facilitates the relationship between producer and consumer by enabling economies of scale) (Faulkner-Gurstein and Wyatt 2021). Evans and Gawer (2016), also focusing on platform functions, highlighted four other types of digital platforms: transactional, innovation, integration, and investment. Finally, Gawer (2008) clusters platforms into four different types: (1) internal platforms, (2) supply chain platforms, (3) industry platforms, and (4) double-sided platforms. Thus, in this jungle of definitions, OGD platforms could be seen as a mix of lean platforms that seek to connect data providers with potential re-users to benefit from cost reductions and production platforms that provide their resources through a licensing system. From a typological perspective, OGD platforms are P2C platforms because they allow all types of consumers to access the data, and P2P platforms because they allow public sector organizations to re-use the data provided on the platforms. However, OGD platforms can also be considered innovative platforms, as their assets can be re-used by anyone to create new products, services, or technologies, and transactional platforms that facilitate interaction between stakeholders.

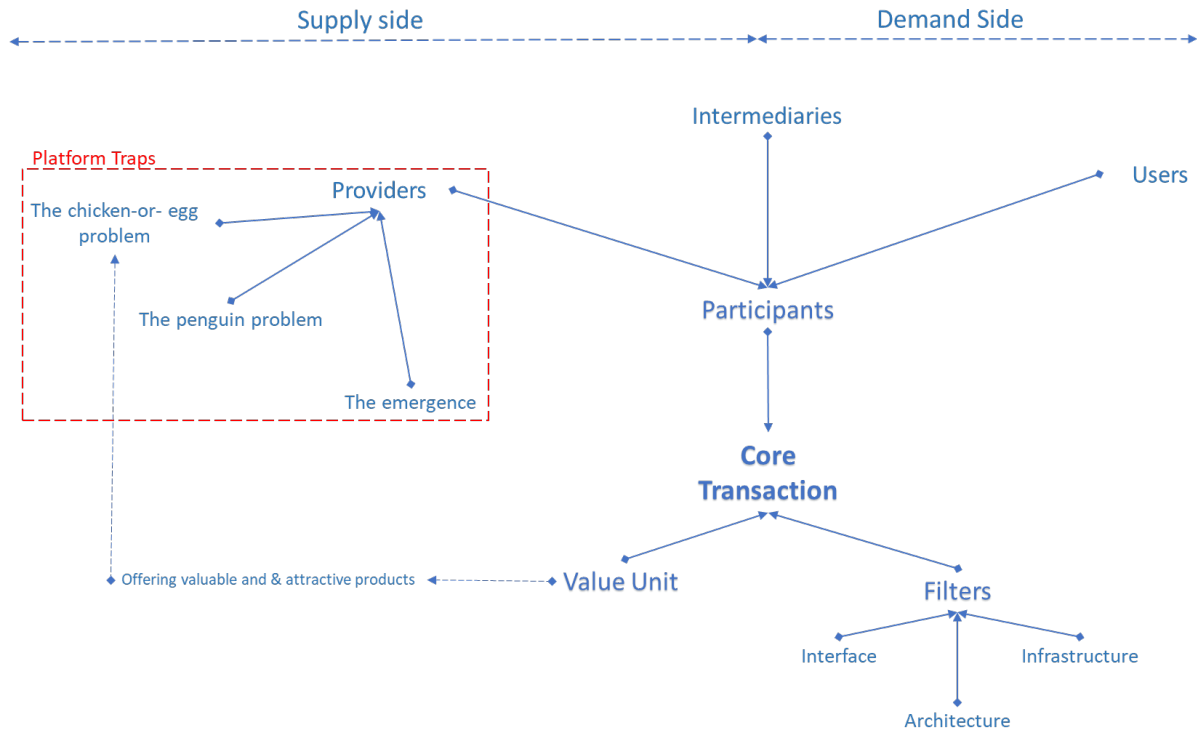
Understanding the Swiss OGD platform from a business mode perspective, it is the fourth typology presented by Gawer (2008), the double-sided markets, also known as two-sided, multi-sided, and multi-sided platforms (Walton 2017), that has been retained for this work. The double-sided markets design symbiosis and direct interaction between two or more groups of participants to create value through technologies, products, or services (Gawer 2008). Multisided platforms are building blocks that other companies use as the basis for developing other technologies, services, or products (Gawer 2008). Following Langley and Leyshon (2017) vision, the strength of a platform in a digital ecosystem depends on intermediation practices and capitalization processes. While the authors define mediation practices as "convening, facilitating, and coordinating relatively autonomous actors that enable them to function as a network," they identify capitalization processes as *"the process by which this mediated network is seen as potentially revenue-generating and thus able to channel investment"* (Langley and Leyshon 2017). Accordingly, the various steps that define a multi-sided platform, such as platform construction, operation, and exploitation, as well as the transformation of existing organizational forms, often referred to as platformization (Casilli and Posada 2019; Nieborg and Poell 2018), are of great importance to the success of multi-sided platforms. Therefore, to succeed and survive, multi-sided platforms must achieve platform building, operation, and exploitation (Faulkner-Gurstein and Wyatt 2021).

### ***Platform operation***

However, for the platform to work, multi-sided platforms must shape the creation of symbioses between the actors on the platform (Robertson & Ulrich, 1998). As described above, a platform is not just a technology, a website, or a mobile app but primarily a business model that generates network effects by connecting people, organizations, and resources through technology in an interactive ecosystem. According to Parker et al. (2016), network effects are the most important source of value creation. Therefore, to initiate network effects, the platform model must get the core transaction of the platform right (Parker et al. 2016). This means focusing on the three essential elements of a platform - filter, value unit, and participant (Fig.2) (Parker et al. 2016). Parker et al. (2016) refer to the filter as an interface that enables the exchange of value units between different actors. In other words, the filter can be seen as a marketplace where, for example, Kindle users can buy e-books, smartphone owners can download apps, and data analysts can obtain datasets. Then the authors define the value unit as the goods, services, or information exchanged. In the case of Airbnb, the exchanged product is a list of apartments. In the case of the Android Play Store, the value unit takes the form of applications. In the case of OGD platforms, the value unit corresponds to OGD or data sets. Finally, the term participant includes all actors in the platform, including providers, consumers, and intermediaries. To support interactions, a platform-based model is generally governed by a platform owner (Bakıcı et al. 2013). In addition to the platform owner, who ensures the platform's governance, various actors participate in the platform ecosystem (Baldwin and Woodard 2009). While the supply side may include all actors involved in product creation (i.e., suppliers, producers etc.), the demand side includes platform adopters such as end users, third-party developers, and complementors who develop new and innovative services or products (Bonina and Eaton 2020; Tiwana 2013). Some actors may play different



roles at different times. For example, third-party developers may develop apps and download and use them (Parker et al. 2016). Similarly, data providers involved in OGD production, such as public organizations, may also act as future re-users.



**Figure. 2 Platform Process and platform traps**

***Platform traps***

The literature and experts' experiences show that the success of platforms is not only based on building user-friendly and coherent platforms with a wide range of products. Fostering network effects and creating core transactions largely depends on the ability of providers to pay more attention to start-up and evolutionary principles of the platform, i.e. (1) offering valuable and attractive products, (2) creating interactions with potential consumers of the products, and (3) creating a self-organising order that makes the ecosystem work. In other words, platform success depends on product providers' ability to avoid platform traps - the chicken-or-egg problem, the penguin dilemma, and emergence (Fig. 2) (Parker et al. 2016; Tiwana 2013). The following sections briefly define these basic principles for successful platforms that are often perceived as traps.

**The Chicken-or-egg problem**

According to the literature, one of the biggest challenges for platform models is to simultaneously attract both the supply and demand sides to the platform (Tiwana 2013). There is no question that a platform must have both sides on board, i.e. the suppliers and the consumers, to ensure a valid value proposition. However, as long as one of the two parties does not play along, the second is unlikely to join (Caillaud and Jullien 2003). This dilemma is called the chicken-or-egg problem (Caillaud and Jullien 2003). The reason for this dilemma is that the perceived attractiveness of a platform to consumers lies in what they can derive from it. In other words, the absence of valuable and attractive products limits the value and utility of the platform and reduces the number of interested consumers. However, suppliers only have a strong interest in participating in platforms if there is a sufficient consumer base. Due to economies of scale, providers only

benefit from the platforms' business model if they reach enough customers (Tiwana 2013). To avoid this dilemma, providers should therefore offer valuable and attractive products so that consumers are interested in joining the platform.

### **The Penguin dilemma**

A successful product is not enough to generate successful platforms; it is only the first step. History is made of encouraging platforms with promising products that stopped at this point, forgetting that *"it takes two to tango"* in addition to one standalone product. The next trap identified in the literature that can affect a platform's chances of achieving its goals is getting all stakeholders on board (Tiwana 2013). Experts call this the penguin dilemma, i.e. no one moves unless everyone moves, so no one moves. The penguin metaphor, which economists also refer to as the "bandwagon effect" (Rohlf's 2003), is used to illustrate the question, *"Who will be the first to jump in?"* Even if there is a predator in the water, a hungry penguin is more likely to jump into the water to find food if the rest of the group also jumps with it. This reduces the individual risk of being thrown at the predator. The penguin dilemma is often seen in introducing new platforms with great potential for network effects, but the fear of being the only one to join the platform keeps stakeholders from taking practical action (Farrell and Saloner 1986). As discussed above, the potential for the success of a platform is an arduous undertaking that requires significant investment by stakeholders. However, the fewer providers there are, the fewer consumers there will be, and as there are fewer and fewer consumers, there will be fewer and fewer providers, and so on. This keeps a vicious cycle going that, if not reversed, can nip a platform in the bud.

### **Emergence**

The concept of emergence is the third principle of the platform model that contributes to the evolution of a platform. At the ecosystem level, emergence is defined as a self-organising order that results from the various actions of platform actors. This self-organising order is not only rooted in the plans of the platform owner but is mainly motivated by the pursuit of self-interested interests based on personal knowledge and continuous adaptation to the other ecosystem actors (Chiles et al. 2004). To this end, the system must move beyond equilibrium to achieve this self-organising order that contributes to the prosperity of the platform ecosystem (Dougherty and Dunne 2011). This will trigger changes that spontaneously generate new energy, products, knowledge, and innovations. However, the emergence and resulting innovations can only occur if platform owners effectively facilitate and shape this emergence (Tiwana 2013). Platform owners can advise platform actors to some extent by providing a guiding vision, but they cannot dictate how this should occur. However, without the emergence of a self-organising order, it is difficult to say whether the platform will be able to sustain the ecosystem interaction that is essential to platform models.

### **Data Governance**

In the 1990s, the ambiguous term governance emerged in academic, political, and professional circles. The widespread adoption of the term governance has created new opportunities for investigation across a variety of social, political, and information science disciplines (Goodwin 2009). Although the term governance is overused and overly generalized, it is always the same idea which refers to the stewardship of different kinds of assets through implementing strategies, policies, and roles (Abraham et al. 2019). Specifically, Dawson et al. (2016) described IT governance as *"the organizational capability exercised by the board [of directors], executive management, and IT management to control the formulation and implementation of IT strategy and, in this way, to ensure the fusion of business and IT"* (Dawson et al. 2016). From a public sector perspective, IT governance is defined as *"a service provided to the public through the exercise of power and authority aimed at meeting public needs and interests"* (Elpez and Fink 2006). When referring to data governance, Ladley (2010) explained that it corresponds to implementing and organising different policies, decision-making procedures, and so on to manage information assets. According to Sebastian-Coleman (2022) data governance provides the opportunity to set policies and establish accountability for



data, which ensures better organization to limit data-related problems and good data standards. In this sense, Ladley (2010) compared data governance to the area of quality assurance (as opposed to quality control) in manufacturing or regulatory compliance in financial services. While data governance focuses on the data, platform governance focuses on the platforms that need to be "controlled." In other words, platform governance is described by Tiwana (2013) as one of the two gears of the motor of platform evolution. Indeed, the author indicates that the architecture on which a platform relies must be properly managed for it to function, otherwise data or other assets on a platform won't be valuable or useful. In particular, he points to the notion of pricing, control, and decision rights (Tiwana 2013).

## **Research approach**

In this part, we discuss the research approach adopted for this dissertation. First, we describe the aims of this dissertation and explain the chosen research philosophy. Then, we present the research question and explain the research design used to achieve the goals of the dissertation. Finally, we explain the context in which this study was conducted.

### ***Research aims***

By understanding OGD as a digital platform product, this thesis draws not only on the literature on Open Government Data and digital platforms but also on the practical implementation of OGD in Switzerland. In doing so, this thesis intends to decipher and understand the platform barriers that prevent OGD Swiss platform data providers from opening public data. Based on the primary consideration that the sustainability of an OGD platform depends on the interactions between data providers and re-users and thus on the ability of providers to generate and sustain core transactions, this dissertation aims to explore how OGD providers deal with these platform traps, namely *the chicken-or-egg problem*, *the penguin problem*, and *the emergence concept* (Parker et al. 2016; Tiwana 2013). In other words, this dissertation seeks to understand whether Swiss OGD providers are caught in platform traps and explores the consequences of the public data opening process. In addition, this thesis also seeks to better comprehend the reasons OGD providers are stuck in this early stage of platform development. With this goal in mind, this dissertation focuses on data governance in public sector organizations and seeks to establish a link between data governance practices and OGD reusability.

### ***Research philosophy***

In a context where this dissertation seeks to better understand the different reasons that discourage Swiss OGD providers from opening public data and then offer possible solutions, it seems appropriate to follow the pragmatism paradigm. While some authors use the term paradigms or research philosophy (Holden and Lynch 2004; Lincoln et al. 2011; Mertens 2010), others use the term worldview to describe the "*basic set of beliefs that guide action*" (Cresswell and Plano Clark 2011). A paradigm, described as the core of research methodology (Guba and Lincoln 1982; Holden and Lynch 2004), can be viewed as a belief system that guides how research data are collected, analysed, and used (Guba and Lincoln 1982). In particular, the paradigm orientates researchers in their choice of research design and methodology but also underlies the selection of a particular research method (Holden and Lynch 2004; Luthfi 2021; Saunders et al. 2007). Cresswell and Plano Clark (2011) identified the pragmatism paradigm as one of the four most commonly used paradigms in mixed methods research. As an overarching philosophy (Erzberger et al. 2003), the pragmatism paradigm is a real-world practice oriented toward "*what works*", focusing primarily on research consequences, research questions, and diversity of data collection methods (Cresswell and Plano Clark 2011). In other words, the pragmatism paradigm is a solution-oriented worldview that mixes several data collection methods to increase knowledge about the problems under study (Cresswell and Plano Clark 2011; Kankam 2019). Therefore, the pragmatism paradigm seems appropriate to achieve the goals of this dissertation and provide avenues for further reflection.

## ***Research questions***

This dissertation follows the problematization approach to derive the research questions and achieve the abovementioned goals. Alvesson and Sandberg (2011) define problematization as a method for "*identifying and challenging the assumptions underlying existing theories*". In literature, the prevalent way of generating research questions relies on gap-spotting. However, focusing on the gap-spotting technic suggest that the literature already understands the boundaries of the field under study. This tends to limit the studies from others' perspectives and think out of the box. Even when a wealth of research has appeared on a topic, it does not systematically make an exciting contribution (Alvesson and Sandberg 2011). More is being produced, but the high-impact papers remain elusive theories. To address this problem, Alvesson and Sandberg (2011) suggest generating research questions through problematization to challenge the assumptions of others. In this way, researchers may approach research with new questions based on their perspective and other points of view. Because it opens critical debates, gives voice to unusual ideas, and allows for alternative perspectives, the problematization approach is especially recommended in research areas where researchers have colonized a particular worldview. The approach seems particularly relevant in the OGD context because although OGD research is relatively new, researchers already seem already stuck on similar viewpoints. Indeed, OGD barriers are mostly associated with technical, societal, and organizational problems (Barry and Bannister 2014; Janssen et al. 2012; Ruijter et al. 2017), and the abovementioned literature reveals that the concept of OGD as a platform product has not even been mentioned (Bonina and Eaton 2020; Danneels et al. 2017). Hence, this dissertation follows the principle of problematization to derive research questions and expand knowledge about OGD barriers. First, we identified the field of literature and key texts that constitute the OGD field. We then examined the main assumptions underlying the literature in the OGD field and found that the opening of OGD, and consequently their re-use, was often linked to technical, social, or organizational issues. After considering whether these assumptions were worth challenging, we developed an alternative assumption: OGD opening issues may also result from the model used, i.e. the platform model, as well as the upstream treatment of the data, i.e. data governance. From this, the following overarching research question was developed in this dissertation:

***RQ: How do the platform's traps and data governance practices impact the opening process of public data?***

The first trap to avoid for a platform to thrive is to offer attractive products or services (i.e. the chicken-or-egg problem). Therefore, this thesis starts with Paper#1 on the attractiveness of the products offered and first diagnoses the metadata quality of the Swiss platform. By investigating and diagnosing the core of the OGD platform, this thesis seeks to comprehend whether the Swiss platform still faces the chicken-or-egg problem and whether the published OGD are attractive enough to convince re-users to join the platform. Yet, offering an attractive product is a necessary but not sufficient condition for creating successful platforms. For a platform to be successful, two sides must be on board - the users and the providers. Experience already shows that few users seem to be interested in OGD. We thus decided to investigate the views of OGD providers and question their opinions and readiness for OGD in Paper # 2. This second study aspires to better understand the attitudes of public sector employees towards OGD and to determine whether the Penguin Dilemma affects OGD providers. Finally, the literature argues that the evolution of a platform after solving the start-up principles lies in the self-organising order resulting from the various actions of the platform actors, i.e. emergence. Therefore, in paper #3, we focus on the third platform trap and investigate the intentions and behaviours of public authorities in the production of open data in the context of data governance practice in a public service of a large city in Switzerland. Through the first three papers, i.e. Paper#1, Paper#2, and Paper#3, this dissertation seeks to understand whether platform traps influence the

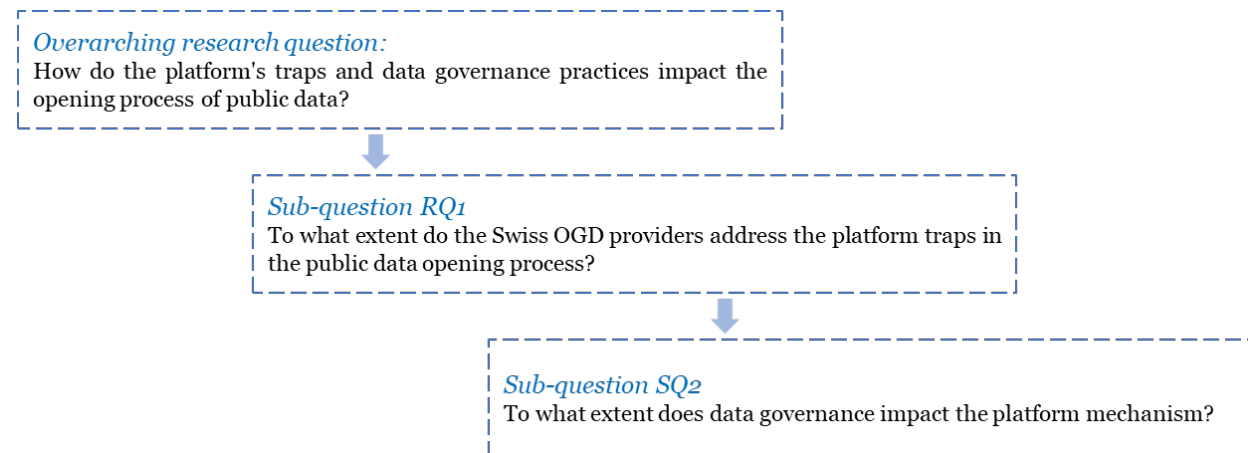
opening process of public data and play a role in the current low re-use of OGD. Thus to address this point, this dissertation seeks to answer the following *sub-question*

***RQ1 – To what extent do the Swiss OGD providers address the platform traps in the public data opening process?***

Exploring the extent to which Swiss OGD providers address platform traps in opening public data is not sufficient to provide policymakers recommendations. Therefore, this dissertation intends to go a step further and investigate the reasons that might explain why OGD providers remain blocked with platform traps. Therefore, the fourth paper of this dissertation (i.e. paper #4) focuses on the level of data governance practices in Swiss public organizations. By attempting to better comprehend the level of public data governance, this final paper aims not only to provide an overview of data and OGD governance activities in Swiss public organizations but also to show that a lack of data governance may prevent OGD providers from addressing platform challenges. Thereby, this dissertation seeks to answer *the second research sub-question*

***SQ2 - To what extent does data governance impact the platform mechanism?***

In doing so, the aim of this thesis is twofold. The first objective of this dissertation is to better understand the possible causes that undermine the opening of public data. By identifying the causes, this thesis can therefore focus on possible solutions. Accordingly, the second objective of this thesis is to address these issues by providing guidelines to public sector employees and policymakers involved in OGD publishing.



**Figure 3. Outline of the research questions**

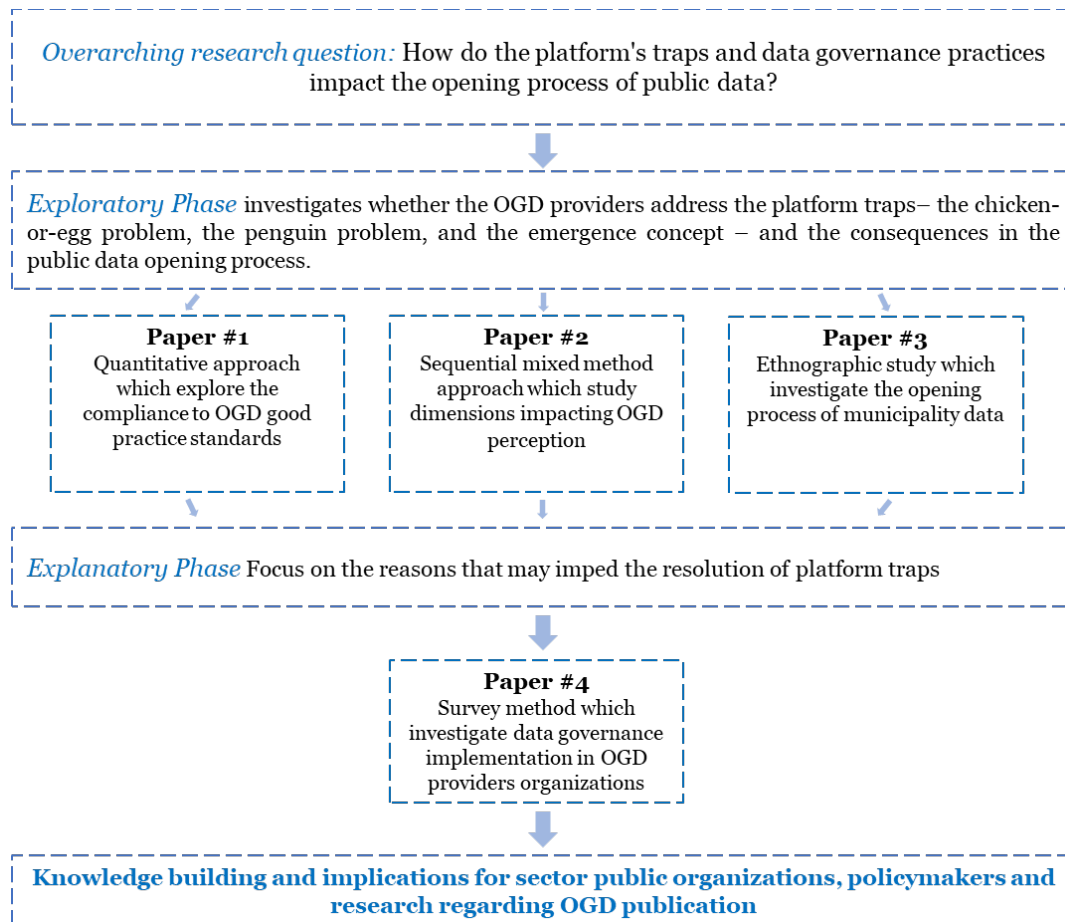
### ***Research design***

The rapid evolution of new technologies has created tremendous opportunities for public organizations to collect and utilize data but has also brought new challenges for government bodies to make it useful. In addition, the willingness to publish OGD on a national platform to promote transparency, stimulate innovation, and support business has made the opening of public data an integral part of public organizations' activities (Constantinides et al. 2018). Yet, opening public data and understanding the challenges of reusing OGD is a complex phenomenon. It involves many stakeholders, linking public policy issues, information technology, data management, and organizational relationships. In this rapidly changing environment, OGD researchers face situations where previous literature does not provide adequate explanations or offer significant insights into the issues under study. However, taking advantage of the flexibility offered by mixed methods design for a dissertation on the opening process of public data could provide robust results for OGD researchers. Indeed, integrating the features of qualitative and quantitative research approaches provides a more comprehensive and complete view of the question under study (Creswell and Clark 2011). According to Newcomer (2016), the accumulation of factual and subjective

knowledge contributes to improving the understanding of complex issues that affect project implementation. It enables in-depth analysis and the collection of appropriate evidence (Newcomer 2016).

Therefore, mixed methods should help explore the complex phenomenon of platform traps, how it affects OGD openness, and help advance OGD practices. Thus, to address the overarching research questions and shed light on the reasons that may impact the opening of public data in Switzerland, this dissertation uses a mixed methods approach to gather data from a survey, an ethnography, a focus group, and an API containing thousands of metadata. Yet, the *mixed-methods approach* does not just mean collecting both data types. It also necessitates careful consideration of the relationship between the two types of data and the ability to contrast findings (Creswell and Clark 2018; Tashakkori and Teddlie 2003). For this purpose, different mixed-methods research designs exist, notably the multiphase research design. Because the multiphase research design is a problem-oriented design that aims to “*address a set of incremental research questions that all advance one project research objective*” (Creswell and Clark 2011), a multiphase design was chosen in this dissertation to explore the importance of platform traps as well as the implementation of data governance in the OGD context. Early writings use the image of the sandwich to illustrate how multiple layers lead to a complete result (Creswell and Clark 2011). Through an iteration of interconnected qualitative and quantitative studies, multiphase research design provides triangulation of findings on the same topic and allows for the construction of a complete research project, i.e. answering the sub-questions underlying the public data opening process. Therefore, this dissertation compiles and synthesizes the results of four qualitative and quantitative studies, divided into exploratory and explanatory phases, to provide a better overview of how platform trappings and data governance affect the opening of public data. While the first three articles belong to the exploratory phase, the last article covers the explanatory phase.

This dissertation starts with the exploratory phase to investigate how Swiss OGD providers address the platform traps in the public data opening process. To this end, the first three papers of this dissertation each examine one of the three traps most commonly described in the literature - the chicken-or-egg problem, the penguin dilemma, and the implementation of the principles of platform evolution principles, i.e. emergence. To study the chicken-or-egg problem, the first paper (Paper #1) seeks to understand the value and attractiveness of the Swiss digital platform product (i.e. the OGD) using a quantitative approach. For this purpose, we extract from the *opendata.swiss* API about ten different metadata from almost 20,000 datasets. Following the literature on data quality, we then built an index that measures OGD providers' compliance with good practice standards. Subsequently, we used a sequential mixed methods approach to understand public managers' perceptions and understandings in the face of OGD implementation (Paper #2), thus exploring the Penguin Dilemma. As a first step, we collected data using the Repertory Grid technique - a sort of focus group - in an afternoon with 18 public managers. To highlight and identify the different behaviours, we conducted a quantitative data analysis, the Principal Component Analysis (PCA). The last paper of the exploratory phase ( paper # 3) concentrates on the third platform trap and seeks to analyze the activities in the opening process of Swiss public data to better understand their impact on OGD opening. To do so, the third paper has taken the form of an ethnography, collecting multiple meeting notes, hours of observations, thousands of emails, and a wealth of documents over the course of a year. The results of the exploratory phase helped us refine our understanding of the platforms' traps and highlighted that the level of data governance of OGD provider organizations could influence them. The second phase of the research design - the explanatory phase - therefore explored the extent to which data governance may influence platform traps and, thus, the entire chain of the OGD process, from data opening to OGD re-use. To investigate data governance practices and their impact, the final paper of this dissertation (Paper #4) developed a 45-question survey for OGD data providers that have already published datasets on the Swiss OGD platform. These represent about 100 organizations at the communal, cantonal, and federal levels.



**Figure 4. Outline of the research design**

### ***Research context***

The term open data first appeared in 1995 in a document from a U.S. government agency regarding the publication of geographic and environmental data. Although the idea can be traced back to 1966 with the Freedom of Information Act (FOIA) (Abdelrahman 2021), 2009 marked the start of the OGD movement when the administration of U.S. President Barack Obama launched transparency and open government initiatives. Since then, several governments have followed suit and announced new initiatives to disclose their public information (Daniel et al. 2012; Tang and Jiang 2021). The data.gov USA OGD platform, launched in May 2009, and the United Kingdom's data.gov.uk platform, launched a year later, have contributed significantly to the development and success of the OGD movement (Abdelrahman 2021). This proliferation of OGD platforms and new strategies also provided the impetus for the scientific community to make OGD a new area of research (Tang and Jiang 2021). Although research on open data began just over a decade ago, studies have been conducted on every continent and cover topics as diverse as there are authors (Tang and Jiang 2021). A quick search of the Web of Science platform (i.e. string search = "open government data" conducted on November 30, 2022) shows that at least 79 research areas have covered research on OGD. The top 15 are presented in Figure 5.



**Figure 5. Overview of OGD research topics**

This thriving field, which encompasses information systems, computer science, public administration, management, telecommunications, and many other areas, is thus an intersection of many disciplines conducted from different perspectives and with contrasting methodologies (Tang and Jiang 2021). As a result, the actual understanding of the movement and its implications are blurring, and important currents of thought are struggling to emerge. While González-Zapata and Heeks (2015) adopt a bureaucratic perspective, McNutt et al. (2016) and Meijer (2015) follow a more technological approach, Zhao and Fan (2018) embrace a political perspective, and McBride et al. (2019) opt for an economic view. Given this, theories have not yet been identified, making it difficult for young researchers to get an overview of the situation. In addition, the few literature reviews that have been conducted indicate a wide variability in the terminology used in OGD-related research, which adds to the complexity of understanding the issue (Hassan and Twinomurizi 2018; Kvamsdal 2017; Safarov et al. 2017; Wirtz et al. 2022). For example, when authors from the private perspective address data management issues, researchers from the public side refer to governance issues. Thus, it is in this fuzzy context that this dissertation was developed.

## Research procedure & articles summary

This section presents in more detail the logical and structural connections between the four papers that, taken together, achieve the stated research objectives. Therefore, to better understand the impact of platform theory on the opening process of public data, this dissertation organizes the articles according to the three main platform traps that often affect the platform mechanisms and application level of data governance practices. While Paper#1 investigates the chicken-or-egg problem, Paper#2 explores the penguin dilemma, and Paper#3 considers the concept of emergence. Finally, Paper#4 examines data governance levels in Swiss public organizations.

### *Paper 1*

A sufficient consumer base is needed for providers to join platforms and derive strong interests from them. Due to economies of scale, providers only benefit from the platforms' business model if they reach enough adopters (Tiwana 2013). To achieve a sufficient customer base, providers must address the chicken-or-egg problem. Otherwise, the other side will not join the platform, and the platform will undoubtedly fail. Thus, an OGD platform will not be able to attract subsequent re-users if it does not have a wide variety of valuable

datasets, and public organizations are unlikely to publish datasets on the platform if they are not sure that these datasets will be re-used. According to Tiwana (2013), the safest way to overcome this problem is to attract the consumer side with an inherently valuable and attractive product. Among the most successful platforms, such as iOS, Microsoft Windows, Amazon, and Google, all started with standalone products that then evolved into platforms. In the case of the iOS ecosystem, for instance, the platform's development and success began only after iPhone enthusiasts embraced the product (Tiwana 2013).

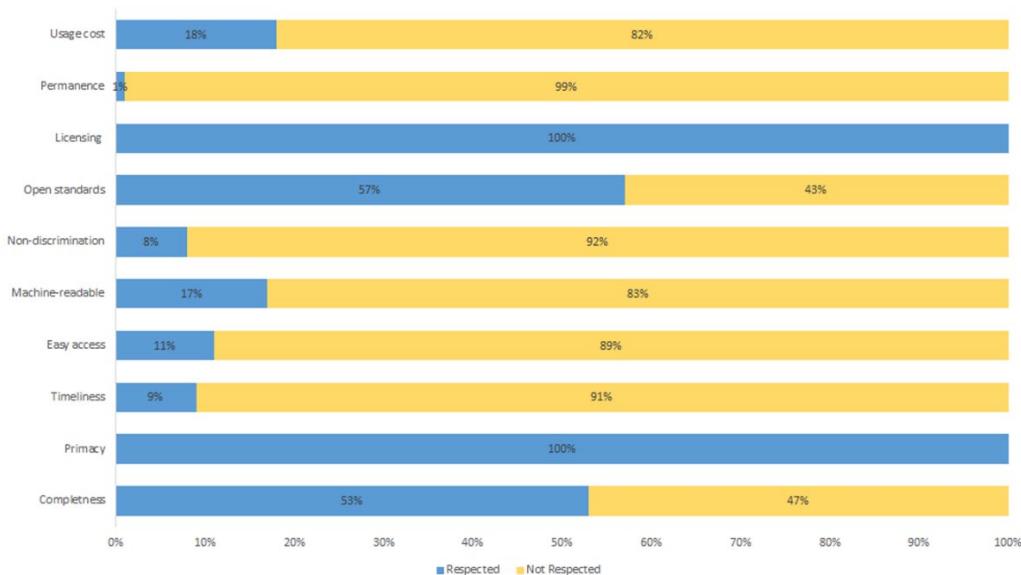
This is probably why the literature on OGD reinforces the importance of valuable and attractive data through data quality. We have found in the literature that it is essential for governments to promote the development of data quality to increase the attractiveness of platforms (van Veenstra and van den Broek 2013). Data quality is considered one of the most important factors for the attractiveness and development of a web platform (Ciancarini et al. 2016). Recent events confirm this view. For instance, the case of the U.S. Census Bureau, which faced data collection issues (Mervis 2020) or the way Switzerland tried to keep track of coronavirus cases (Swissinfo 2020) led to the publication of incorrect and incomplete data, prompting citizens to ask questions about the reliability of the data. Data quality is a major challenge for OD platforms and can affect the implementation of OD projects (Open Knowledge Foundation 2018c; Torchiano et al. 2017; Umbrich et al. 2015). However, as Kubler et al. (2017) discussed, ensuring a high and consistent data production process on OGD platforms - including data quality, metadata access, application of standards, etc. - is one of the biggest challenges today. Poor data sets or a lack of information may severely impact OGD discovery and processing; even if one type of data is available on a nationwide OGD platform, machines and humans may not be able to use or find it.

We assumed that an OGD platform would not be able to attract end users if it did not have a large variety of valuable datasets and good quality for end users. In order to verify if the Swiss OGD platform offers valuable products that allow avoiding the chicken-or-egg problem, this thesis starts with an investigation of the potential attractiveness of the proposed products. The first phase of this investigation examines and diagnoses the core product of the OGD platform ecosystem - the datasets published on the Swiss platform. While data quality has been primarily studied over the past two decades (Sadiq et al. 2011), defining data quality is a recurring challenge. Data quality may refer to *"the characteristics associated with the data and the processes used to measure or improve data quality"* (Askham et al. 2013), but it also depends on the context as well as the perspective of data consumers (Sidi et al. 2012b). In information systems research (IS), data quality is often defined as a multidimensional concept based on key dimensions such as accessibility, timeliness, completeness or data quantity, data accuracy, and consistency (Abate et al. 1998; Krasikov et al. 2020; Wang and Strong 1996). However, as Batini et al. (2009) noted, defining data quality and the dimensions are chosen to assess it is easier said than done. For this reason, in this thesis, we are guided by the general definition used in research from IS. This choice is consistent with (1) the lack of consensus in the literature regarding data quality for optimizing the use of OGD platforms, (2) a pragmatic approach, as the goal of this thesis is not to reflect on a consensus regarding data quality definitions, and (3) the fact that OGD advocates such as the Sunlight Foundation (SF) consider similar dimensions for improving OGD publishing. Indeed, the dimensions most commonly cited in the OGD literature for measuring data quality are consistent with the good practices proposed by OGD advocates. The SF has developed ten good practices-including completeness, primacy, timeliness, ease of access, machine readability, non-discrimination, open standards, licensing permanence and usage cost -to *"empower the public's re-use of public data held by governments"* (Sunlight Foundation 2018). It is essential to remember that while the SF talks about good practices, they are applied to metadata and not directly to the datasets themselves.

Accordingly, we decided to measure the compliance of public organisations to good practices proposed by the SF. Emphasizing compliance may provide an interesting overview of the OGD quality produced and insight into the reliability and potential attractiveness of the platform ecosystem. Therefore, through the underlying question - To what extent do OGD published on such platforms follow existing good practices? - the purpose is twofold: First, to determine whether the datasets offered on the platform are valuable and



attractive, and second, to investigate whether public organizations in Switzerland follow good practices when publishing in open access. To this end, we developed a compliance index and conducted a comprehensive empirical analysis based on 17,777 published data resources from the Swiss OGD platform *opendata.swiss* (Paper #1). Thanks to this first paper, we have shown that in Switzerland, in addition to not providing valuable and attractive datasets, the adherence to good practices for publishing OGD is fairly low. On average, public organizations in Switzerland complied with only 4 out of 10 principles when publishing OGD in 2018.



**Figure 6. A general overview of adherence to the ten good practices**

As presented in Figure 6, few resources were 100% compliant with the principles of primacy and licencing. Two principles, completeness, and open standards were adhered to by more than 50% of data resources, while permanence, timeliness, and non-discrimination were adhered by less than 10% of data resources. In addition, we found that communes were more likely to follow best practices than cantonal and federal organisations—they are the only group of data providers that largely followed the easy access and usage cost principles. Although some principles were less followed at the cantonal level than at the municipal level, we still found efforts to comply with 9 out of 10 principles. At the federal level, only the principles of completeness, primacy, open standards, and licencing seem to be followed. Finally, the results of paper#1 show that the OGD published on the Swiss platform does not represent attractive products of sufficient quality, suggesting that the Swiss platform is still struggling with the chicken-or-egg problem.

## ***Paper 2***

Given the results of the first paper (Paper #1), it is legitimate to wonder why organizations are not applying the standards of good practice. While the poor quality of the data might discourage future users from joining the platform, might not the cost-benefit ratio for publishing OGD limit the adoption of the platform by OGD providers? Hence, can public managers, as key actors initiating the OGD platform ecosystem, influence the decision to publish data, thereby increasing (or decreasing) the critical mass needed to attract future end users? Thus, to understand the extent to which the Penguin Dilemma may influence public managers in their intention to open public data, this dissertation develops the second paper - *Different Shades of Perception: How Public Managers Comprehend the Re-use Potential of Open Government Data?*

According to Yang et al. (2015a), uncertainties in OGD influence public organisations' attitudes toward OGD and cause them to act differently. In IS, project management approaches, organisational resistance, and expectation misalignment are often the reason for high project failure rates (Anthopoulos et al. 2016;



Coelho and Valente 2017; Holgeid and Thompson 2013; Korzaan and Erskine 2018; Sauser et al. 2009; Varajão 2018). As those responsible for project implementation in their department, organisation, or even at the national level, public managers play a key role in aligning institutions' strategic goals with national or international policies (Lee and Kwak 2012). According to Whitley and Hosein (2008), managers' attitudes and perceptions toward technology may affect not only policy but also the outcomes of policy decisions. Zhao and Fan (2018) argue that project managers need to understand the concept of OGD and its implications to have a clear picture of the value of OGD. Otherwise, they will continue to resist OGD production, which may result in government data never being opened and government agencies being more opaque than ever. In the European Strategy for Data, the European Commission (2020) goes further and emphasises that fragmentation of perceptions among stakeholders (i.e. governments, interest groups, application developers, and researchers) is a major risk in implementing a common data strategy. Yang and Wu (2016) confirmed this view, pointing out that *"the intention of government agencies to publish open data is indeed a strong predictor of government agencies' open data implementation activities"*. Studies at IS confirm this idea and show that the cognition and perception of managers matter. They are critical in determining an organisation's strategic direction and impact (Tan and Gallupe 2006). According to Carayannis et al. (2003) and Hardless et al. (2015), cognition - the human ability to perceive, interpret, and think about the environment - is closely related to people and their actions. Smelser and Baltes (2001) share the same idea and argue that action and perception are interdependent.

We, therefore, assumed that the perceptions of public managers, who directly influence the required critical mass of end users, may help to understand the reasons that prevent the publication of Swiss OGD. Thus, we assumed that if public managers have different views on data re-use, this may lead them to apply good practices differently and consequently provide either valuable and attractive or unusable datasets. Guided by the question - How do public managers comprehend the potential re-use of OGD? - we have developed a second paper (Paper #2) that addresses public managers' comprehension and perception of OGD. We intended to shed some light on why OGD does not follow publication practices and faces major obstacles or even stalemates in certain countries, so we used the Repertory Grid technique to explore and investigate the understanding and perception of the potential re-use of OGD. With this in mind, we conducted interviews that allowed for the development and analysis of the content of 18 grids of public managers working in different branches and levels of government. Based on the affordance theory and the assumption that one can perceive a potential before acting and achieving real impact, we further explored whether public managers share the same vision regarding the re-use potentials of OGD in different application domains. We attempted to understand the extent to which public managers apply and similarly comprehend the concept of OGD and which application domains and dimensions lead to variability and conflicting opinions. In doing so, we were able to shed light on public managers' current intentions to share OGD. The analysis results show that public managers seem to share different views on possible applications and comprehension of OGD. They do not seem to place the same value on free accessibility, reusability, and data sharing for everyone sharing, to anyone, for any purpose, and without legal, technological, or social constraints. On the contrary, they seem to adapt these principles depending on the scope of OGD, which means that we do not observe "OGD universality" among the 18 public managers interviewed, contrary to what is often believed by open systems advocates. Furthermore, public managers seem to share different views on how OGD should be re-used. There is no clear and shared vision about who should potentially re-use OGD, nor what level of openness should be applied. These findings make it clear that there is no team mental model. In other words: Neither a global vision nor a common purpose justifies jumping in the water. In contrast to the comparison with the penguin dilemma, the results suggest that it is not the development or survival of the project that drives OGD providers but something else. Consequently, as with the first platform trap, the penguin dilemma also seems to impact the public data opening process.

### ***Paper 3***

The results of the second paper (paper #2) make the emergence of a self-organizing order challenging to conceive. While the guiding vision is generally expressed through OGD initiatives by governments, it is up

to OGD providers and other stakeholders to create and develop an ecosystem that supports that vision. However, if data providers already hold different views, how can the various stakeholders agree on how the OGD ecosystem will work? Several press articles have reported exciting facts about the interactions and organization of open data provision during the pandemic. Le Temps' article on the Federal Statistical Office (FSO) is a good example. The news Paper noted, "*Jusqu'à présent, les médecins, les hôpitaux et les cantons étaient libres de choisir la manière dont ils souhaitaient transmettre les données, bien qu'il existe un formulaire électronique standardisé (...) l'envoi par fax était également possible*" (Temps 2020). This justifies research into what specifically happens in opening practices and whether a self-organizing order may emerge in the public data publication process. Following the recommendations of Yang and Wu (2016), it is crucial to study government agencies' intentions and behaviours in producing open data and how public entities administer OGD (De Tuya et al. 2017). The literature shows that fear of losing work autonomy and perceived quality of information or social influence impact innovation in the public sector and significantly increase project failure (Meier et al. 2013).

Therefore, we decided to investigate more concretely this self-organizing order in the opening process of public data. We had the opportunity to conduct an ethnographic study when we were an external stakeholder realizing an assignment for a municipal project. This project involved two main stakeholders: the SOI (i.e. the IT department at the municipal level) and the housing department (i.e. the data owners). With the SOI, our role was to "facilitate" the opening of the data held by the housing authority. To this end, meetings and interviews were held between November 2018 and April 2019. During this period, we developed a research paper (Marmier and Mettler 2019a) proposing an Action Design Research (ADR) to test solutions for de-identifying data. The ADR included the organization of a hackathon, so we organized a public event, the Smart City Boost Hackathon, in collaboration with SOI, the Housing department, and the University of Lausanne, which took place in Lausanne on May 3 and 4, 2019. After the hackathon, we wrote a whitepaper for public organizations (Marmier and Mettler 2019b), summarizing basic advice on de-identification rules. However, while the project aimed to explore solutions for the de-identification of government data, the frequent interactions with the different actors provided important information about the state of mind of the different participants as well as the general atmosphere that prevailed in this self-organized order. Thus, with in mind to dig deeper and better comprehend the organization on the supply side of the Swiss OGD platform ecosystem, this ethnography gave us the opportunity to examine *the extent to which data governance practices impact the OGD publication*.

For this purpose, we first identified the key components of data governance required for OGD publishing in the literature - structural, procedural, and relational - and then compared them to our observations during this immersion. This immersion in a self-organized municipal department was crucial because it gave us a complete overview of the situation and made us aware of the lack of emergence in this context. In addition to revealing that data governance practices in this context significantly impacted project implementation, the results also showed that the two services (i.e. SOI and the housing department ) maintained a good working relationship but could not collaborate to produce OGD. We found fierce resistance and a staunch refusal to cooperate on the part of the housing department and discovered that the definition and implications of OGD were poorly understood by both stakeholder actors (i.e. the data owners and the department responsible for the opening process). We also noticed the low enrolment of civil servants, and few had time to move forward with the project. These observations allowed us to develop an OGD governance model that includes the data governance components influencing the success of the OGD publishing process ( Fig. 7). By showing the complex interaction and collaboration between the different actors, which reveals the lack of self-organization, the results of this paper seem to go in the same direction as the other two. They seem to confirm that OGD providers face platform traps and that these traps affect the opening process of public data.

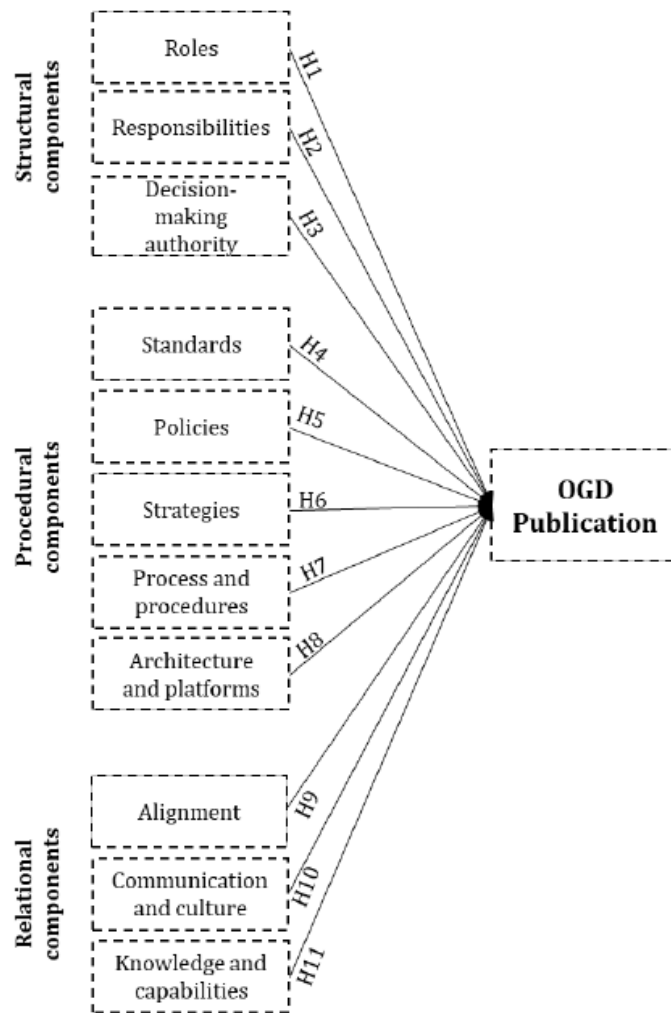


Figure 7. The OGD governance model

#### Paper 4

In addition to the findings of the previous papers, which offer exciting insights into the barriers that hinder the OGD opening process, the third paper also highlights that the absence of data governance in public organizations may explain why data providers struggle with these traps. Many studies reveal that activities related to dataset collection, transformation, and publication are resource-consuming for public organizations (Attard et al. 2015; Barry and Bannister 2014; Janssen et al. 2012; Zuiderwijk et al. 2012). According to Yang and Wu (2016), public organizations are expected to understand OGD concepts and take on the work of dozens of intermediaries. For instance, organizations involved in the process must expend significant resources to identify only their own datasets (Zuiderwijk and Janssen 2014b). Publishing OGD also requires public organizations to meet and consult with experts to avoid disclosing personal data or making it public under the law (Zuiderwijk and Janssen 2014b). A recent example shows that the FSO's new goal regarding OGD strategy is to coordinate the creation of a list of "priority OGD" before the end of 2020 (Federal Statistical Office 2020b), which requires the cooperation of many data owners and government offices. So far, no additional aid has been announced to support them. Last but not least, the news Paper 24 Heures recently reported on access to personal data, noting that "Certains offices considèrent que les demandes relatives à la loi sur l'information perturbent leur "quotidien" et tente " d'affaiblir la loi

*sur la transparence par des dérogations inscrites dans des lois spéciales" (24 heures 2020).* Thus, we may easily understand that opening public data without a data governance program already in place can be a nightmare for public organizations. Accordingly, this prompted us to learn about data governance practices in Swiss public organizations.

	<b>Level 1 Initiate</b>	<b>Level 2 Managed</b>	<b>Level 3 Defined</b>	<b>Level 4 Managed</b>	<b>Level 5 Optimized</b>
<b>Technical</b>	Identifying the different data storing systems (cloud, local server etc.)	Identifying data origin (e.g. data owner, services or department where data come from)	Cleaning and storing data according to OGD standards  Preparing a registry of publishable data	Transforming data given personal and sensible data guidelines as well as OGD standards	
<b>Procedural</b>	Defining personal and sensible data guidelines		Defining an OGD publication policy according to the data origin (e.g. business, IT, sensible, non-sensible)	Establishing rules and managing quality and metadata of OGD	
<b>Relational</b>				Enforcing training for employee engage in OGD governance Enhancing collaboration between IT and business domains	
<b>Monitoring</b>				Defining a strategy of OGD publication Monitoring OGD strategy advancement	
<b>Structural</b>				Determining OGD decision-making authorities	Establishing roles & responsibilities of the OGD governance stakeholders and completing the data governance program

**Table 1. OGD Governance Maturity Model**

Although the adoption of OGD strategies is a serious data governance challenge, few articles investigate data governance practices from the public lens (Al-Ruithe and Benkhelifa 2020; Brous et al. 2016; Tiwana et al. 2013), making public sector data an under-researched area (Al-Ruithe and Benkhelifa 2020; Tiwana et al. 2013). Yet, judging from the results of seven of the top ten market capitalizations, i.e. GAFAM, data governance practices have become increasingly important to be competitive in data-driven activities. For this reason, in this fourth paper, we examine *the extent to which OGD, and data governance practices are implemented in Swiss public organizations*. Following the perspective of the problematization logic proposed by Alvesson and Sandberg (2011), Paper#4 has two objectives. The first objective was to evaluate the maturity of OGD governance in Swiss public organizations through a survey distributed to OGD publishers. Based on this survey, it was then intended to identify which potential data governance mechanisms have been successfully implemented and which could benefit from further development. Accordingly, the second objective of Paper#4 was to develop a Maturity Model (MM) for OGD governance. Based on the list of Swiss OGD publishers available through the API, we reached out to several publishers

of 83 organizations, such as Bernmobil, SITG, ETH Library, etc. (Confédération Suisse 2020) registered on the Swiss platform. Questionnaires were emailed to respondents twice between February 2022 and April 2022. We received 252 answers from respondents, which we analyzed using the Rasch model. This gave us a more detailed overview of the data governance situation in Swiss public administrations (Confédération Suisse 2020) and allowed us to identify the "hot spots" of the OGD platform ecosystem. This gave us the opportunity to develop a tool to assess the maturity of data governance adapted to the Swiss public administration. In addition, the study found that technical governance practices are most adopted by Swiss organizations, regardless of the political level of the organization (i.e. communal, cantonal, or federal), while monitoring and structural practices face further difficulties. The results also show that implementation occurs at two speeds: data-oriented organizations (e.g. archives) versus less data-oriented organizations (e.g. teaching hospitals). These results also offer intriguing insights into the relationship between data governance and platform traps. Some data governance aspects appear essential to avoiding platform traps, particularly data cleansing, preparing data registries, and defining decision-making authority and roles and responsibilities. Yet, important practices required to avoid platform traps emerge only at the third and fifth levels of the MM.

	Title	Publication outlet	Research question	Research design	Main findings
Exploratory Phase	Paper #1	Information Polity	<i>To what extent do OGD published on such platforms follow existing good practices?</i>	Quantitative approach	In Switzerland, the adherence to good practices for publishing OGD is fairly low. The study reveals that in average in 2018, public organisations in Switzerland only respect about 4 out of 10 principles when publishing OGD.
	Paper #2	International Conference on Information System	<i>How do public managers comprehend the potential re-use of OG</i>	Sequential mixed method approach	Results highlight that there is (1) no OGD Universality, i.e. No universal vision of free accessibility and reusability of data to anyone, for any purpose and without any constraints and (2) no team mental model, i.e. Public Managers seem to share different views regarding how OGD must be re-used.
	Paper #3	DG. O Conference	<i>To what extend data governance practices impact the OGD publication?</i>	Ethnographic study	Findings display eleven data governance practices that may impact the opening process of OGD. The different practices are clustered in three components (1) structural, (2) procedural, (3) and relational.
Explanatory Phase	Paper #4	ECIS	<i>To what extent OGD and data governance practices are implemented in Swiss public organizations?</i>	Survey method	Results indicate that among the five data governance practices – technical, procedural, relational, monitoring, and structural – the thee last seems the more complicated to implement

**Table 2. Summary of the articles**

## Discussion

Based on the emerging literature on platforms, this dissertation examined in depth the opening process of public data from the perspective of platform theory. The overall goal of this dissertation was to investigate whether platform traps block Swiss OGD providers during the process of opening public data and identify whether not addressing these platform traps may prevent Swiss OGD providers from opening their data.

This dissertation pursues this goal in two main phases. First, the dissertation focuses on an exploratory phase in which the first three contributions (i.e. paper#1, paper#2, and paper#3) investigate the extent to which Swiss OGD providers address platform traps. The dissertation then conducts an explanatory phase to understand and explain the reasons that prevent public organizations from not applying golden platform rules. By examining how Swiss data providers have approached the logic of the platform model for OGD publishing, these articles give insights into the reasons that may prevent OGD providers on Swiss platforms from opening their data. Even though (Big) Data management techniques and technologies are progressing, this dissertation seems to show that there are still many problems in the OGD context that prevent OGD platform providers in Switzerland from making public sector data openly available. Contrary to what is claimed in the OGD literature, this compilation of articles tends to show that the challenges do not stem only from the technology or the open nature of OGD (Bonina and Eaton 2020; Danneels et al. 2017). Insufficient knowledge of platform models and data governance appears to play an important role and severely impacts the success of OGD strategies. The findings of the dissertation show that the public sector in Switzerland seems to be very engaged in OGD implementation and accepts the discourse of data as an asset, while OGD providers face significant platform model issues that could arise from important data governance issues. Noting that platform traps and inadequate data governance prevent OGD publication and re-use, this work can provide specific and adapted solutions to OGD providers who want to promote their participation in the OGD movement. By discussing the results of this work and relating them to the current literature, this dissertation also suggests how OGD platform providers can address the challenges of platform and data governance.

### ***Platform traps***

The set of papers highlights the chosen model's importance for sharing public data. By offering lists of datasets, the Swiss platform follows the same model as Uber or Booking, which offer lists of drivers or hotels instead. However, unlike these thriving private platforms, the *opendata.swiss* platform seems to fall into what experts call the platform trap. Indeed, the results of the four studies show that OGD providers still face platform start-up problems, i.e. the chicken-or-egg problem and the penguin dilemma, as well as platform evolution issues, i.e. problem of emergence. Although Swiss authorities developed the OGD platform years ago, the dissertation shows that the platform is still stuck in these fundamental problems. Paper#1 shows that metadata quality hinders OGD re-use and consequently affects the mechanisms determining a platform's attractiveness, addressing the platform model's limitations. Product quality has significant value in accelerating the platform model, and potential users will not join the platform if data quality is poor, but only if it is valuable and monetizable (Parker et al. 2016; Tiwana 2013). Furthermore, a platform is only valuable if actors join it. The more data providers are on the platform, the more valuable the platform is for re-users and vice versa. As the literature suggests, it takes two to tango. But as Paper#2 and Paper#3 suggest, data providers do not seem interested in "joining" the OGD platform, which keeps plunging the Swiss platform into the penguin dilemma and emergence issues. No one wants to be the first to jump into the water, and no self-organising movement around OGD initiatives can emerge. Therefore, these issues make it difficult to implement the mechanisms of the platform model.

While previous research has made valuable contributions in explaining why OGD providers struggle with OGD strategy and identifying the barriers that limit data access on an OGD platform (Barry and Bannister 2014; De Reuver et al. 2018; Janssen et al. 2012), they often attribute OGD barriers to institutional constraints (i.e. relationship, lack of knowledge), technical limitations (i.e. technical failure, the dysfunctionality of data portals), and societal barriers (i.e. civic engagement) (Wirtz et al. 2022). Yet, they omit the influence of public organisations' data governance and the relationship between OGD sharing and the platform models used. This dissertation generally supports previous research findings on OGD barriers, particularly data quality (Torchiano et al. 2017; Vetrò et al. 2016) and organizational resistance (Meier et al. 2013; Wirtz et al. 2016). However, while the literature recommends working on data properties (e.g. lack of interoperability, data formats (Barry and Bannister 2014; Smith and Sandberg 2018)), platforms (e.g. technical support, updating data platforms (Janssen et al. 2012)), users knowledge (Syed-Ikhsan and

Rowland 2004), developing an appropriate architecture to leverage the features of platforms, and transforming public sector data into actionable insights (Parker et al. 2016; Tiwana 2013), this dissertation, in line with several other studies, shows that this is only part of the work. These elements, considered essential in the OGD field, are not enough to ensure the success of a platform (Bonina et al. 2021; Danneels et al. 2017). As cited by experts, history is full of platform projects with good ideas and best intentions that have not led to their goal (Tiwana 2013). Therefore, when considering the OGD platform as a platform model, it is essential to overcome the platform trap if OGD platforms are to have a chance of success.

Considered the main challenges to successful platforms, several studies have shown that solving platform traps is the most important insight to generate network effects and enable the development of a healthy platform. Since network effects are at the core of a platform-based business, a platform that does not create network effects seems doomed to disappear. According to experts, failing to address these traps is particularly problematic because data quality prevents interactions between providers and downstream users (Parker et al. 2016; Tiwana 2013). Bonina et al. (2018) noted that the OGD platform could not provide valuable resources to OGD developers and users without high-quality datasets. Tiwana (2013) also emphasised that *"not addressing the chicken-or-egg problem is guaranteed to break a platform even before it has a chance of taking off"*. In line with Cennamo and Santaló (2015), this thesis also highlights the reluctance of the public sector to jump in the water first and openly publish public data. The authors pinpointed that while Amazon had overwhelming success with the Kindle in 2007, the company had difficulty convincing book publishers. Publishers were reluctant to release their latest bestsellers in e-book format (Cennamo and Santaló 2015), resulting in a loss of potential customers. Similarly, De Reuver et al. (2018) showed how difficult it was for a start-up platform to convince potential platform actors to join. The authors found that while this new platform based its strategy on working with known partners to secure initial meetings with providers and increase its platform implication, they were sceptical of joining a platform without a reputation.

For De Reuver et al. (2018), platform issues can be explained by the fact that public organizations not only have a very heterogeneous IT infrastructure and technical systems but are also unfamiliar with business models based on digitalization, which requires a different mindset and new perspectives. For this reason, Gawer and Cusumano (2014) argue that the platform operator should articulate and communicate a coherent vision of the platform ecosystem and its products in order to attract and convince potential stakeholders to join the platform and take the plunge (Faulkner-Gurstein and Wyatt 2021). For Ondrus et al. (2015), the best way for data providers to reach critical mass is to open the *"platform to players from the same industry"*. Salminen (2014), with the same thought in mind, shows that providers are more likely to join *"existing platforms with an established customer base"* (De Reuver et al. 2018). In other words, this overlaps with the idea of the data space, which has recently been developed in other countries. Data spaces generally have a specific sectoral focus, such as health, education, energy, and mobility, and seek to link fragmented and scattered data from these different sectors. In doing so, the data space aims to provide an *"interoperable, trusted IT environment for data processing and a set of rules of a legislative, administrative and contractual nature that determine the rights to access and use data"* (European Commission 2020). According to European Strategy for Data, *"Common European Data Spaces will ensure that more data are available for use in the economy and society, while the companies and individuals generating the data remain in control"* (European Commission 2020).

Finally, in the platform literature, authors have also pointed out that platform development is often driven by the Red Queen effect (e.g. Apple vs Microsoft, PlayStation vs Xbox). Borrowing from Lewis Carroll's book - Alice in Wonderland - the Red Queen effect illustrates the competitive paradox and the pressure to innovate (Barnett and Hansen, 1996). Although Alice and the Red Queen are constantly running faster and faster, they always stay in the same place compared to each other (Tiwana 2013). This paradox illustrates the behaviour described by Meadows (2008) called escalation in systems thinking. The author explains that in competitive cases, the innovative actions of one party are perceived by a second party as an imbalance in the system. This situation may create a sense of threat to the second party, which in turn will adjust its



behaviour to maintain the equilibrium. Consequently, in order to survive, both parties will increase their activities to adapt quickly enough to their competitors, which will lead to outbidding competition (Tiwana 2013). In the context of OGD, this means that governments should produce OGD at least as fast as their competitors in order to survive and stay in the game. Unlike Facebook or Google, however, the primary role of public organizations is not to monetize public data or develop activities to do so. Yet in the age of globalization, where public policies reflect countries' ambitions and serve to project the image of governments' power on the world stage, building OGD platforms that make governments more transparent has become an important issue for national sovereignty (Conseil fédéral 2014; Conseil fédéral 2018). This new enthusiasm for OGD platforms even led to the developing of several indexes and websites, such as the Global Open Data Index and the Open Knowledge Foundation (Jacob Arturo Rivera Perez et al. 2020; World Wide Web Foundation 2018). In the process, these indexes began to rank platforms, highlighting countries' performance but also increasing competition, leading OGD platforms to publish more and more (Constantinides et al. 2018). This race to gain a reputation as an open and transparent country could thus explain the rush to develop the platform. The "red-queen effect" creates so much pressure that governments worldwide are forced to open their data quickly. As Thomas Schulz, the head of the FSO's data department, explained at the opening of the Swiss Open Data Forum in June 2022, *"To boost the OGD platform, we put everything that public organizations owned on the platform. But then we had a lot to clean up, mainly because of the amount of PDF"*. This could also explain why those organizations that are more advanced in data-driven activities or already accustomed to transparency seem better equipped for OGD adoption than those that have never had to report on their data.

To avoid platform traps and enable platform prosperity, Parker et al. (2016) and Tiwana (2013) recommend establishing platform governance and naming a platform owner. Tiwana (2013) explains in her book that platform governance adapted to the different phases of the platform lifecycle may help overcome start-up problems (Tiwana 2013). While Faulkner-Gurstein and Wyatt (2021) showed that the platform model could advance public interests, such as developing new medical treatments, the authors also show that none of this would have been possible without platform governance. The platform owner may develop network effects through platform governance by balancing three essential functions: pulling users, facilitating collaboration and exchange, and matching the right users with appropriate providers (Parker et al. 2016). To this end, several strategies have been proposed in the literature. Parker et al. (2016) proposed a list of eight proven strategies to help platforms address the chicken or egg issue - the follow-the-rabbit strategy, the piggyback strategy, the seeding strategy, the marquee strategy, the single-side strategy, the producer-evangelism strategy, the big-bang adoption strategy, and the micro market strategy (Srinivasan and Lakshmiopathy 2017). In addition, Cennamo and Santaló (2015) highlighted that developing a platform without a strategic focus may increase the risk of failure. In particular, they mentioned the case of Groupon Inc. (i.e. the online coupon company), which ceased operations after unfocused growth (Cennamo and Santaló 2015). Another solution Cennamo and Santaló (2015) proposed is to pursue an intermediate path between the mass market and a niche (e.g. data space). Bonina and Eaton (2020) also suggest cultivating and building a chain of organizations with common goals to provide solutions to the traps of start-up platforms.

### ***Data and OGD platform governance issues***

While this dissertation focuses on traps encountered by data providers on the Swiss platform, it also highlights another element that prevents the opening of public data - the lack of data governance in Swiss public organizations. The set of studies seems to indicate that many public organizations have difficulties implementing data governance activities. In particular, Paper#3 and Paper#4 directly examine the governance aspect in public organizations. Besides showing a two-speed OGD governance implementation that seems to depend on the area of activity, this dissertation mainly shows a lack of data governance practices in Swiss public organizations. The survey of OGD publishers conducted in paper#4 shows that 38% of the surveyed public organizations do not have or are thinking about a data governance program, while 28% of the respondents have such a program underway. This means that just over a third of the



organizations surveyed have a data governance program in place, while 100% have already published OGD on the Swiss platform. These results are not surprising when the first two papers are put into perspective. Indeed, the results of paper #1, which provides a general overview of compliance with the ten good practice standards, already predicted such a conclusion. Non-adherence to OGD practice standards such as timeliness, machine readability, and durability already indicated deeper problems, such as disruptions in the management and governance of public sector data in Switzerland. The fact that only 17% of data resources respected a machine-readable format or that only 9% were published in a timely manner (i.e. no later than six months after their collection) was a harbinger of the level of data governance applied in the Swiss public sector organizations. The same observations can be made for Paper#2 and Paper#3. Paper#2 indicates that some public managers tend to adapt their publishing practices according to the effects of the OGD re-use potentials while others do not, raising the question of a data governance program. The ability to adjust their behaviour in terms of the publishing process suggests that the public administrators interviewed for Paper#2 do not appear to be subject to an OGD governance program and are questioning their digital data governance to a greater extent. Following this line of reasoning, it could be argued that depending on the importance or utility of the data and the perspective of the managers responsible for publishing data, they are likely to implement varying levels of data governance activities. However, this tends to confirm the problems and biases in data governance practices of OGD and, more generally, public sector digital data. Finally, Paper #3 is probably the most revealing document highlighting the problems faced by Swiss data providers regarding data governance issues. By underpinning that structural, procedural, and relational data governance components seem to influence the success of the OGD publication process and re-use of OGD, this paper clearly brings to light that problems of a data governance program prevent the OGD data providers of the Swiss platform from opening public data.

Although more than half of the top ten companies - Apple Inc, Microsoft Corp, Alphabet Inc (formerly Google), Amazon.com Inc and Meta Platform Inc - seem to be leaders in data governance practices, Holt et al. (2015) showed that only 55% of private organizations surveyed had a data governance program in place in 2015 (Alhassan et al. 2016). As shown in this dissertation, there are also deficiencies in data governance in public sector organizations. Thompson et al. (2015), who studied two Western Australian public information systems, also take this line. The authors reached alarming conclusions regarding data governance activities in the Police Firearms Management System and the Department of Health Information System. While in the former case, the auditors acknowledged that they had "no confidence in the accuracy of the information on the number of individuals licensed to possess firearms or the number of firearms licensed," *the Department of Health acknowledged not really knowing "which medical personnel were responsible for the clinical data entries made"*. Kim and Cho (2018) and Morabito (2015) also noted similar issues, particularly the inaccuracy and incompleteness of organizational data (Abraham et al. 2019). A recent OECD report also goes on this line, showing that public sector organizations face challenges in managing and governing data relevant to their success but different from traditional assets (OECD 2019a). From the OGD perspective, Ruijter et al. (2017) and Young (2020) have already addressed the lack of data governance as a barrier to OGD, and Wirtz et al. (2022) show that data governance seems to be of particular importance in the OGD context. Finally, Chiesa (2019) has shown that the difficulties caused by data-related action and understanding jeopardize what he calls the "*datization of a vast amount of previously intangible information*" (OECD 2019a). Yet, considered as the main challenges in Big Data value creation, several authors point out that data governance is essential for all data-driven activities. In addition to enforcing the development of data standards and policies (Ndamase 2014) data governance provides companies with the insights and information needed to run a data-driven organization (Trom and Cronje 2019) and the benefits of a data-driven strategy (Ernst & Young 2014). Moreover, the absence of an overarching data governance model opens the door to numerous excesses, such as the proliferation of data standards and technical solutions for data sharing. As a result, the lack of data governance impacts data interoperability, integration, and processes, as well as organizations and citizens. Rather than promoting the "once-only" concept (i.e. requesting information from businesses and citizens only once), the lack of data governance may lead to the public sector asking citizens to provide the same personal data multiple

times (OECD 2019a). Therefore, the question of data governance is critical as it not only addresses organizational issues, but also directly affects the value of the platform. Without a robust data governance program that specifically addresses data quality issues, the opening and sharing of Swiss OGD will remain an illusion. Thus, implementing a data governance program in public organizations not only gives the public sector with a clearer view of its data assets and data-related actions, but also supports open data initiatives by contributing to good platform practices.

Due to the multitude of disciplines, principles, and even actors involved in data governance, several hypotheses have been proposed in the literature to explain data governance issues. According to Russom (2011), the volume of data is not the only reason why traditional data management techniques are outdated when analysing and processing (big) data. Their complexity and volatility also play an important role. As data is defined by its large-scale team-based projects that bring together interdisciplinary teams, data is collected by different organisations with different institutional approaches. As a result, data ownership, privacy, and data processing are often unclear. Such complications lead organisations to face significant challenges in establishing data governance (Trom and Cronje 2019). Moreover, in the absence of a common programme, disciplines, tools, and techniques for collecting, cleaning, and organising government data vary, resulting in heterogeneous data production (Butts-Wilmsmeyer et al. 2020). For Siddiq et al. (2016), data management problems, especially with heterogeneous data, lie directly in data governance issues. Therefore, challenges such as decision-making mechanisms, operational frictions, or lack of regulation may directly impact Big Data management. In addition, the OECD report suggests that companies face difficulties in data governance because they still face organisational culture and its resistance to digital innovation (OECD 2019a). Unlike today's social media culture, where people may easily share information with friends and strangers on a daily basis, the policies, standards, and procedures that govern data governance are more difficult for many of us to understand (Thompson et al. 2015). For example, it takes many hours to digest and fully understand the difference between private and sensitive data and other important documents (Thompson et al. 2015). Many studies agree with previous observations and suggest that the activities associated with collecting, transforming, and publishing datasets are resource intensive for public organisations (Attard et al. 2015; Barry and Bannister 2014; Janssen et al. 2012; Zuiderwijk et al. 2012). According to Yang and Wu (2016), public organisations are expected to understand data and OGD concepts and take on the work of dozens of intermediaries. Thus, organisations involved in the process must devote substantial resources to identifying only the datasets they own (Zuiderwijk and Janssen 2014b). OGD production also requires public organisations to meet and consult with experts to avoid disclosing personal data or making it public under the law (Zuiderwijk and Janssen 2014b).

Switzerland's lack of a program or systematic data governance practice is not surprising. This dissertation was written the same time the European Data Governance Act (EDGA) came into being. The EDGA was just passed in May 2022, and its application will take place in September 2023. Aiming to *"ensure a better distribution of the value derived from the use of personal and non-personal data between the actors of the data economy"*, the EU proposes four broad sets of measures, in particular measures to facilitate data sharing (The European Parliament 2022). In practice, this consists of developing additional categories of actors, such as data stewards and chief data officers in organizations, but also delineating the roles and responsibilities of actors involved in the process. The data governance literature recommends that organizations further analyze the extent to which the role of the data owner is useful to a data governance program. If so, they should define it (Abraham et al. 2019). To this end, organizations should further explore how data ownership is determined. Similar to Vilminko-Heikkinen (2019), in the study of data management projects in a Finnish municipality, it is often unclear whether an organization defines data ownership based on the data user, where it is stored, or other elements. This ambiguity around data ownership can create complications for the rest of the data governance program. Organizations should also focus on which function is assigned decision-making authority. As Otto (2011a) notes, assigning decision authority to a business function has different consequences than assigning decision authority to a function IT or another function. Organizations should also be mindful of data governance mechanisms that may support organizations in inter-organizational relationships, particularly when sharing and exchanging sensitive data,

which may raise new privacy concerns (Winter 2018). Among other things, organizations should monitor compliance with good data governance practices and also measure interest in them in inter-organizational settings (Abraham et al. 2019). Finally, public organizations should develop a standardized and trusted data-sharing environment for data governance to be successful (Abraham et al. 2019; Cohn 2015; Rasouli 2016).

## Implications

### *Practitioners*

In line with this dissertation, the conclusions of a cross-sectional audit on the implementation and evaluation of the Swiss OGD Strategy 2014-2018 revealed that the public sector has only partially achieved the OGD objectives. Based on these conclusions, the Federal Council mandated the Federal Statistical Office (FSO), which already oversees the OGD strategy, to take measures to enable multiple data uses. Thus, the FSO developed a new program for National Data Management (NaDB). The main objectives of this program follow the OGD movement but emphasize the once-only principle (i.e. companies and individuals must provide certain information to the administration only once) and the interoperability of data sets. In this way, Switzerland has developed three platforms dealing with public data over the past decade. The *opendata.swiss* platform was launched in 2016, followed by the interoperability platform (I14Y) in 2019, and the trusted data space and digital self-determination (i.e. the data space) are expected in 2023. Although the term OGD is not mentioned in any of the documents related to these new projects, their goal remains to facilitate the use and valorization of public data. However, the failure of OGD can be explained not only by architectural and technical problems but also by strategic and organizational issues. By showing that some of the reasons that prevent Swiss data providers from opening public data are more organizational and strategic, this dissertation provides authorities with new elements and priorities to integrate into developing their new platforms. In particular, these findings inform agencies about the importance of platform governance practices and data governance rules that must be implemented across the various organizations that handle data. While the discussion provides some recommendations for solving platform traps, Paper#3, and Paper#4 present elements to focus on when developing data governance in public organizations.

This thesis shows policymakers and public organizations that platform models are not just business models or websites that enable interaction but also sophisticated architecture and complex organizational model that requires strategy. Thus, as an organizational model, platforms not only fulfil an economic logic but simultaneously pursue various goals (Faulkner-Gurstein and Wyatt 2021). By understanding how platforms work and proposing solutions to improve them, this dissertation facilitates the development of market-oriented initiatives for policymakers in Switzerland. The findings also suggest ways the public sector can improve its self-organization and better meet its obligations. As Faulkner-Gurstein and Wyatt (2021) note, platforms provide "*opportunities for discipline and oversight*" and an ecosystem that all share the same purpose. Therefore, this dissertation offers policymakers and public organizations involved in opening public data new elements to achieve the OGD strategy and improve the public sector's performance. While this strategy aims to make public data accessible to everyone without restrictions (i.e. the definition of open data), this dissertation clearly exposes that some organizations seem reluctant to make public data accessible to everyone. Therefore, this thesis also focuses on the question of users. Before making public data accessible to anyone and for any purpose, this dissertation offers elements for policymakers and public services to rethink the development of platforms by and for public administrators and public sector employees. Furthermore, by proposing solutions to promote the integration of public organizations into the platform (i.e. solving the Penguin dilemma), this dissertation helps policymakers improve the once-only principle pursued in the new strategy. This dissertation also contributes to and promotes the development of the data space. By showing that some sectors are more willing to share datasets but are also ready to do so at the organizational and technological level, this work shows which areas to focus on first. For the first time, therefore, agencies may choose to develop an OGD platform or data space with organizations that

have a well-implemented data management program. Focusing on the "*most advanced*" organizations would allow the Swiss platform to address the data quality issues and thus deal with the chicken-or-egg problem. Thus, this thesis also helps to narrow down the problems related to the emergence concept. The most common reasons that prevent public organizations from entering the OGD process are the fear of getting it wrong and the time and cost involved in managing the data. Working directly with the trailblazers, for whom data governance is a common practice, will likely help resolve the question of who jumps in the water first. It is also more likely that public administrators working with shared resources will join a process to facilitate their interaction. This could create the self-organizing system so important in a platform ecosystem that fosters collaboration, sharing, and network effects.

This thesis not only contributes to improving the development of open data in Switzerland but also fosters the importance of data governance and data management in all organizations that deal with data daily. In fact, this thesis is an excellent example for the public sector regarding the usefulness of data governance programs in administrations. As they form the basis for many future developments such as smart cities, artificial intelligence (AI), etc., data collected by public officials must be "treated" before it is used. Otherwise, data-driven projects could not be realized. Thus, by showing the impact of the absence of a standard data program or data governance practices, this dissertation enables public organizations to better understand the impact of their daily work on the exploitation and valorization of public data. In addition, identifying a lack of data governance in many public organizations already involved in data sharing allows authorities to become aware of the extent of the problem. This may then allow for developing new and more targeted policies to address these gaps. By demonstrating that all public administrations are unequal in data governance and openness, this dissertation, and papers #2 and #4 in particular, can help agencies better target their policies and support to meet different needs and adapt their strategies. Authorities can thus draw on the work of data governance trailblazers and leaders to better understand what works and what does not and then seek to apply that to other areas. Better data governance would solve some of the platform problems, as it would be easier for data providers to open data, which would directly benefit OGD re-users and citizens. Users would be more inclined to join the platform if the data quality is better. Researchers or entrepreneurs could rely on the proposed products and use them long-term without fear of possible complications. By ensuring that published data follow a common governance program with Swiss public organizations, users would no longer have to worry about the timeliness and individuality of OGD publication or the interoperability of data from different services. Over time, the attractiveness to users would reinforce and promote the network effect and strengthen the platform.

### ***Academic research***

First, this dissertation contributes to and expands the knowledge of OGD research. This is especially true for research that deals with the suppliers' side of OGD platforms. By taking the perspective of data providers, this dissertation brings to light the tensions and issues that public sector workers face in the emergence of OGD and throughout the data-opening process. The dissertation shows that OGD, which is often seen as the new gold of public administration, enabling innovation, transparency, and new business, is only at an early exploitation stage and does not yet enable systematic interaction between public services. This provides new directions and food for thought for OGD researchers interested in better understanding one of the core problems of OGD - how the public sector deals with digital innovation. Moreover, to date, findings and insights on OGD are mainly based on public administration and public management research (Wirtz et al. 2022) but are less often oriented towards platform logic. By focusing on the platform model approach to study OGD barriers, the results of this dissertation provide a different perspective on understanding OGD challenges. This allows for the inclusion of a IS perspective that is too often missing in OGD research (Wirtz et al. 2022).

Second, this work contributes to extending platform theory to the public sector. To date, little is known about how platforms operate in the public sector. Yet platform models have become increasingly present in public sector research, and since the advent of this model, public platforms have emerged for anything

and everything. Therefore, it seems essential to have a better knowledge of the concept of platforms, their goals, and how they work to adapt them to the public sector. It also seems essential to deepen the knowledge of public sector platforms to avoid creating more unused platforms or dead links because the chosen model was not adapted to the purpose. As this thesis and previous research have shown, applying private-sector models to the public sector is not always successful. Like De Reuver et al. (2016) and Bonina et al. (2018), we, therefore, believe that the OGD literature still has much to learn in the area of platform theory.

Finally, this dissertation enhances the collection of scientific evidence on data governance in a public organization, on which few observations have been available. Combined with the focus on platform models, the thesis offers a new approach to the OGD literature. Rather than attributing data publishing and sharing challenges to purely institutional issues, such as staff resistance to public sector digitalization, lack of knowledge, and collaboration issues, this work takes them a step further and traces them to data governance challenges. While data governance is becoming more prevalent in private-sector companies and has been discussed at the government level in recent years, it is still in its early stages in the public sector. In conclusion, we believe that a different research perspective is particularly relevant to the study of OGD because, as Wirtz et al. (2022) noted and as we demonstrated in this dissertation, OGD research does not depend solely on public sector domains. OGD research seems to relate to a variety of adjacent fields, such as data governance, data security, Big Data analytics, Open Data, interface design, etc. Therefore, this thesis offers new opportunities for researchers in each of these different areas.

## **Limitations and further research**

Although this dissertation provides food for thought for researchers and practitioners, it is not without limitations. The pragmatic approach led this dissertation to focus on publishing articles with concrete objectives, i.e., studying specific problems to propose solutions to public organizations or policymakers that could improve their situation. This approach resulted in this work quickly getting to the heart of the matter (i.e., the barriers encountered by data providers) without taking the time to conduct a literature review. Therefore, this thesis relied on existing literature reviews rather than starting with a formal literature review. For more veracity and homogeneity of the term used, it would have been necessary to conduct a literature review of the articles dealing exclusively with the supply side of OGD platforms and the barriers data providers encountered.

In addition, although this dissertation follows the pragmatism paradigm, the evaluation phase sometimes proposed for this paradigm could not be completed. In one of the articles, the implementation of solutions for a major Swiss city was to be tested, but experience showed that this step is not possible in the context of a PhD thesis. The lack of coordination among the various stakeholders, the unwillingness of the authorities, and the time allotted for the final abandonment of the original project indicated that the assessment was too risky to be carried out in the context of a PhD. Yet, according to Venkatesh et al. (2013), *“if a mixed methods approach helps a researcher find theoretically plausible answers to his or her research questions [...], he or she should undertake such research without much consideration of paradigmatic or cultural incommensurability”*.

Further, while a rigorous and transparent mixed methods approach offers several advantages to the field discipline compared to a qualitative or quantitative research design (Hadi et al. 2013), its appropriate application remains difficult to implement (Bazeley 2004). In addition to developing extensive experience with qualitative and quantitative experiments (Almeida 2018), we needed to develop knowledge and practice generalisation to assemble research methods. Otherwise, it would have been complicated to reap the benefits of mixed methods and gain a broader understanding of the question under study (Bazeley 2004). However, combining the strengths of qualitative and quantitative methods requires significant resources and is very time-consuming (Almeida 2018). Therefore, while the mixed method was a fruitful experience, the construction of this dissertation required more effort.

Moreover, this dissertation does not explore the potential impact of Swiss federalism on data governance mechanisms. Although it might be logical that the centralization or decentralization of a country can strongly influence data governance, the literature on this topic is not so clear. Some authors, such as Safarov (2019), argue that data governance's centralized or decentralized nature seems to have a large impact. In particular, the author shows that the strong and centralized support in the United Kingdom, in contrast to the Netherlands and Sweden, seems to allow for a more systematic and strategically well-aligned implementation of OGD. In contrast, Buchmann (2017) concludes that neither a centralized nor a decentralized governance approach is appropriate for information management. The author recommends polycentric governance approaches, i.e. an optimal degree of decentralization, by allowing competition between decentralized governance approaches. He argues that the application of a centralized data governance approach leads to a unified approach. Unitary governance approaches generally do not consider public sector data preferences and constraints. In the same vein, Loukiala et al. (2021) showed that decentralizing data governance can give organizations a competitive advantage and better value, contrary to monolithic data storage. Given the complexity of this issue, the implications of centralized or decentralized data governance should be explored in more detail.

Finally, this thesis approaches data governance from the unique perspective of organizations that open their data on the Swiss OGD platform. Consequently, this approach does not fully reflect the current data governance situation in the Swiss public sector. Consequently, it would be interesting to further develop the maturity model to better identify the organizations and/or sectors that need concrete support in terms of data governance. Indeed, the current results do not provide a visualization of the classification of the organizations by points. This could allow more accurate identification of the organizations and sectors that already have actionable data and require little effort to publish on the Swiss OGD platform. This could also encourage OGD users to join the platform and perhaps even keep it. Finally, working with high-maturity organizations will allow us to better understand what model works in Switzerland, why it works, and the extent to which it is applicable to organizations that struggle to govern and openly publish public data.

## ***Publication***

### ***Included in this thesis***

Marmier, A., and Mettler, T. 2020a. "Developing an Index for Measuring OGD Publisher Compliance to Good Practice Standards: Insights from Opendata. swiss," *Information Polity*:Preprint, pp. 1-20.

Marmier, A., and Mettler, T. 2020b. "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*. Hyderabad, India.

Marmier, A. 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*, pp. 235-243.

Marmier, A. 2022. "Towards a good data governance in the public sector: Myth or reality?" (under review in the 30th European Conference on Information Systems)

### ***Not included in this thesis***

Marmier, A., and Mettler, T. 2019a. "Challenging the Robustness of OGD De-identification Rules Through a Hackathon," in *Proceedings of the 18th IFIP WG 8.5 International Conference EGOV*, San Benedetto Del Tronto, Italy, pp. 81-89.

Marmier, A., and Mettler, T. 2019b. "Proposition pour la Publication des Données Ouvertes Publiques: Working Paper de l'IDHEAP," IDHEAP.

## ***Dance your PhD Contest***

The publication of Open Government Data by Auriane Marmier - Dance Your PhD 2022 - <https://www.youtube.com/watch?v=g0pK0TZ61bY>

# *Paper # 1*



# **Paper #1 - Developing An Index For Measuring Ogd Publisher Compliance To Good Practice Standards**

**Title:** Developing An Index For Measuring Ogd Publisher Compliance To Good Practice Standards: Insights From Opendata.Swiss

**Author:** Auriane Marmier and Tobias Mettler

**Published In:** Information Polity

**Abstract:** In many countries, public organisations are among the largest creators and gatherers of data. To increase economic growth, governments have therefore begun to liberate access to large parts of government data by developing open government data (OGD) initiatives. Since the emergence of OGD initiatives, many OGD portals have been launched. There is a common belief that sharing OGD throughout platforms would be sufficient to motivate companies to re-use data and improve economic growth. However, there is very little evidence about the quality of shared OGD. For companies to be able to re-use, share and create value from OGD, data publishers must meet certain good practice standards. Following a pragmatic research approach, in this paper we present an index that can be applied for the quality assessment of the published OGD on portals. On the basis of 17,777 published data resources gathered from the Swiss OGD portal (opendata.swiss), we demonstrate the logic of the index and discuss the key learnings we obtained from applying the index to this concrete case. We conclude that, in Switzerland, the adherence to good practice standards for publishing OGD is fairly low.

## Paper 1

Paper #1 - Developing An Index For Measuring Ogd Publisher Compliance To Good Practice Standards .....	38
Introduction.....	40
Background.....	41
Data quality in the literature .....	41
Data quality and open data .....	41
Data quality and standards.....	43
Methodology.....	45
Input data for building the compliance index.....	46
Treatment of the metadata .....	46
Results.....	47
Discussion and conclusion.....	50
Appendix.....	53
Appendix A: the 10 data quality principles of the sunlight foundation.....	53
Appendix B: CKAN metadata fields used to create the Compliance Index.....	55
Appendix C: Question set and chaining logic .....	56

## INTRODUCTION

In the current economic context, in which data is driving innovation, many governments have perceived their data as a resource of strategic relevance (Bates 2014; Munné 2016). Through their daily activities, governmental authorities and public organisations generate and collect vast quantities of data (Pollock 2011) and, over the past few years, governments have by far been the major creators of data (Máchová and Lnénicka 2017). So that governmental authorities may benefit from such resources, they have begun to discuss the implementation of different initiatives. With the emergence of open movements, initiatives have been formed to liberate the access to large parts of government data (Attard et al. 2015; Kalampokis et al. 2011b). These initiatives include open governmental data (OGD) platforms that allow everyone to have access to data were founded.

Torchiano et al. (2017) noted that many authors (Alexopoulos et al. 2014a; Barry and Bannister 2014; Zuiderwijk et al. 2016) recommend facilitating access to public data sets. In their view, this could result in interesting new types of re-use for various (i.e. industrial, individual, scientist, etc.) actors. According to Davies (2010), accessible governmental data may generate important revenues, not only for the public but also for the private sector (e.g. applications, smart city tools). This data could allow the development of new value-added services, commercial purposes as well as political issues (Alexopoulos et al. 2014b). Thus, the common belief that sharing OGD throughout platforms would be sufficient to motivate companies to re-use data and improve economic growth was born (Gascó-Hernández et al. 2018; Vickery 2011). OGD platforms are digital infrastructures, which allow everyone to have access to the data, download them and use them for any purposes (Danneels et al. 2017; The European Commission 2018). This belief has therefore led to major investment for the creation of many OGD platforms (Alexopoulos et al. 2014b; Jetzek et al. 2014; Ubaldi 2013).

However, there is very little evidence that OGD platforms in fact enable data re-use and foster innovation and economic growth (Martin 2014). Although several governments release a large amount of data open to the public, Danneels et al. (2017) ask about whether the OGD platforms will allow reaching the targets strategic goals and will live up to expectations. Many authors continue to express high hopes regarding the OD potential (Charalabidis et al. 2018a; publishing 2016; Zuiderwijk et al. 2018) but according to our knowledge, none of the studies has demonstrated the actual re-use of data available on OGD platforms (Danneels et al. 2017). According to van Veenstra and van den Broek (2013), it became imperative for governments to encourage the development of ways allowing practical OGD re-use as well as the reinforcement of platforms attractiveness. It is an illusion to believe that the publication of OD is automatically followed by the download and re-use of published data (Danneels et al. 2017).

Concerning non-re-use of OGD, one of the reasons explored in research relates to quality. According to Torchiano et al. (2017), public administrations (PAs) provide low-quality data. As Umbrich et al. (2015) point out, low-quality of data can seriously affect their re-utilisation. Also, Allison (2010) stressed the fact that the quality of data on OGD platforms is not always appropriate for use in applications (e.g. non-machine-readable formats, no licence on the data sets). To address these quality issues, the literature has presented different models for analysing and evaluating OGD (Charalabidis et al. 2018a; Conradie and Choenni 2012; Heimstädt et al. 2014). Researchers have analysed the many characteristics of OGD, have identified the dimensions that define their qualities, and have measured OGD quality. In the same vein, open data (OD) advocates have developed principles (e.g. use of commonly owned or open formats) encouraging organization to follow standards (e.g. type of formats such as .cvs, .txt. etc.), frequently inspired by data quality dimensions, and have developed indices in order to measure OGD platforms' quality. Although many papers and indices have measured data quality in terms of OGD platforms' usability (European Data Portal 2017), availability (Open Data Monitor 2018) or openness (Open Knowledge Foundation 2018c), few have focused on the quality of the data and none have reported on the data providers' compliance to these standards when publishing OGD.

To advance the research on the actual use of OGD platform, we propose and developed a *Compliance Index with the aim to analyse public organisations' compliance levels to existing OD standards*. In our view, the compliance to good practice standards on OGD platforms is essential to improve the quality of data publication, foster data exchange and consequently improve its re-use. We use the Swiss OGD platform *opendata.swiss* as an example and scrutinise the metadata published by the OGD platform, seeking to shed some light on this research question: *To what extent do data published by public organisations on OGD platforms comply to existing standards?*

The remainder of this paper is structured as follows. In Section 2, we briefly explore how data quality is treated in the literature and what dimensions are used most often in OD standards. Next, we describe the methodology used to develop our Compliance Index. We will then present the results of the index from its application in Switzerland, concluding with the discussion of our research's implications and limitations.

## BACKGROUND

### Data quality in the literature

In the literature, the definition of data quality is a recurring challenge. According to Sadiq et al. (2011), data quality has been studied largely over the two last decades (Wahyudi et al. 2018; Zhang et al. 2019), without finding a clear consensus on the subject (Corsar and Edwards 2017). As Sidi et al. (2012a) noted, data quality depends on the context, as well as on the perspectives of data consumers, which complicates the process of defining quality. In the Data Management Book of Knowledge (edited by DAMA UK), data quality refers to “the characteristics associated with, and to the processes used to measure or improve the quality of data” (Askham et al. 2013). According to CKAN Association Steering Group (CKAN Organization 2018), the data quality is a complex measure of data properties and utilises various dimensions. These dimensions make it possible to know whether data are appropriate for their purpose. Strong et al. (1997) and Wang and Strong (1996) defined data quality in terms of their ability to be re-used by data consumers, while Wand and Wang (1996) defined data quality as high when the data meet their objectives. In the information system (IS) research domain, data quality is often presented as a multidimensional concept based on key dimensions (Abate et al. 1998; Fox et al. 1994; Huh et al. 1990; Redman and Blanton 1997; Wand and Wang 1996). IS researchers have also presented a wide range of dimensions that are useful in assessing this multidimensional concept of quality. For instance, some authors have mentioned dimensions such as accessibility, timeliness, completeness or amount of data, while others have focussed on data accuracy, consistency or timeliness (Wand and Wang 1996). But as Batini et al. (2009) noted, defining data quality and the dimensions chosen to assess it remains hard.

In this paper, we align ourselves with the general definition of data quality used in IS research. Our choice resonates with (1) the lack of consensus in the literature regarding data quality for the optimization of OGD platform' usage and (2) a pragmatic approach since the aim of this study is not to reflect on a consensus regarding data quality definitions. As a matter of fact, the literature shows that platforms - whether they are OD or OGD - do not use necessarily the same dimensions to measure data quality (Yang et al. 2015a). Máchová and Lnénicka (2017) found a lack of harmonisation on the dimensions used and underlined the need for quality standards. Furthermore, the literature shows that most of the studies focus on the evaluation of the platform itself rather than on the quality of its resources (Vetrò et al. 2016). To better understand which dimensions are used and at which level, we describe the different data quality dimensions proposed in the literature to assess OD and OGD platforms.

### Data quality and open data

According to Yang et al. (2005), data quality is one of the most relevant factors in the evolution of a web portal (Ciancarini et al. 2016). For several authors, data quality is a serious risk for OD platforms; they argue that data quality could have impacts on OD projects' implementation (Open Knowledge Foundation 2018b; Torchiano et al. 2017; Umbrich et al. 2015). Prior research (e.g. Henninger 2013; Janssen et al.

2012) pointed to the importance that published data (such as catalogues, dataset or resources) must be comprehensible, complete, consistent, and machine-readable. Also, an unclear licensing of the data or inconsistent pricing might additionally discourage the use of OGD (Vickery 2011). With OGD initiatives, many OGD portals have been launched therefore, various efforts have been made to study the quality of open data. Scientists and data advocates have worked on many models that propose different dimensions for evaluating quality.

Umbrich et al. (2015) and Reiche and Höfig (2013) monitored and assessed the quality of platforms using quality dimensions similar to those in the IS literature. Umbrich et al. (2015) chose a total of six dimensions (data retrievability, use, completeness, accuracy, openness and contactability), while Reiche and Höfig (2013) focussed on quality dimensions such as formalism, completeness, accuracy, richness of information, accessibility and availability, among others. Both base their measurement on datasets but provide an evaluation of the quality of the platforms. Maurino et al. (2014) realised a study based on 50 OGD Italian datasets. They evaluate the quality in terms of completeness, accuracy and timeliness and also perform the quality evaluation at platforms level. Through their platforms analysis, only a few researchers propose solutions for the data providers. Reiche et al. (2014) focussed on the assessment of OGD platforms' metadata to define their quality. Among others, they used dimensions such as the format used, the format's machine-readability, the licence, the URL's accessibility and the existence of a contact e-mail address. In order to give feedbacks to local publishers, Neumaier et al. (2016) proposed an instrument to evaluate open data portals in small and medium-size cities. Chatfield and Reddick (2017) did the same by providing a longitudinal cross-sector analysis for several cities in Australia. They used for these purposes dimensions such as policy intensity, open data provision and data format variety, among others.

Indices available on the Internet also used dimensions as metrics to measure some OGD platforms' quality (European Data Portal 2017; Open Data Monitor 2018), such as the data's completeness, machine-readability and licence. The Global Open Data Index (GODI) provides a snapshot of available OGD regarding different domains, such as government spending, draft legislation, election results, or land ownership (Open Knowledge Foundation 2018c). The goal of GODI is to illustrate the range of data that is available on nation-wide OGD platforms. The Open Data Barometer aims at comparing the readiness, implementation, and impact of OGD platforms worldwide on an aggregate, generic level (World Wide Web Foundation 2018). Again, the index does not deliver immediate information about the quality of shared data but relies on responses of intermediaries or operators of national OD platforms. Lastly, the Open Data Monitor shows indicators for measuring the readiness and maturity of OGD platforms across Europe (European Data Portal 2017). Given that its goal is to compare the maturity level of European countries, it refrains from analysing data providers directly and rests on the level of national OGD platforms.

We noticed that, all of these studies and indices concentrate, first and foremost, on the *intermediaries or operators* of nation-wide OGD platforms and less on the process of data publication itself. To our knowledge, there is no studies or indices that focus on the compliance or adherence to good practice standards from the perspective of *data providers*. As discussed by Kubler et al. (2017), assuring a high and consistent data publication process (i.e. data quality, metadata access, standards applications, etc.) on OGD platforms is one of the main challenges today. Poor datasets or a lack of information may largely affect the discovery and processing of OGD; while some kind of data might be available on a nation-wide OGD platform, it might not be used or found by machines and humans. Zhang et al. (2019) underline that data users such as scientists and citizens should be able to explore, investigate and understand the quality of datasets to foster their re-use. We assume that even with a high degree of data quality, a low level of metadata requirements will negatively impact the use or reutilization of OGD. According to the W3C Egov Interest Group (W3C 2017), without a proper documentation of the nature and content of the data it is hardly reusable. To guarantee a more effective use of OGD platforms, we strongly believe that good descriptions of metadata are crucial: they are the keys that allow data consumers to explore and understand the signification of data and evaluate if the provided quality is suitable for their purposes. We presume that

ensuring data quality begins by ensuring that good practices standards are respected and applied by data providers, but also ensuring an access to a comprehensive documentation.

#### **Data quality and standards**

Data are described as the most valuable asset of the century. Nevertheless, many authors are raising questions about why the promised data-driven innovation still not happening. One of the hypotheses is that data quality dimension do not follow standards and that could impact data exchange (W3C 2017). The Web Best Practices Working Group studied 26 OD use cases in order to understand how the lack of standards could retard the development of the data-driven economy (W3C 2017). Umbrich et al. (2015) compared 82 OGD platforms' and showed the utilisation of a wide range of formats. He explains this diversity by a lack of standards defining the resources formats. Many authors further notice claim the necessity to standardize the dimensions used to assess data quality (Ciancarini et al. 2016; Máchová and Lnénicka 2017; Vettrò et al. 2016; Yang et al. 2015a). Thus, in our view, governments' application of and compliance to good practice standards on OGD platforms are essential to improving the platforms' quality as well as to attaining OGD objectives (e.g. re-use of their content).

As seen above, although there is still no consensus on how to define data quality for OGD portals, and even less on the dimensions used to measure data quality, we saw that several data quality dimensions have emerged from the OGD literature. Given that there is an unstructured development of data quality literature, Paré et al. (2015) recommend conducting a comprehensive narrative literature review to manually gathered dimension that occurred the most. Therefore, we listed the different dimensions of data quality used by researchers to assess OGD platforms' characteristics in Table 1. Among them, several form part of the good practices published by advocates of OD and the most frequently used dimensions to measure data quality in the OGD context are: licence, easy access, machine-readable format, timeliness and completeness.

Data quality dimensions	(Vetrò et al. 2016)	(Máchová and László 2016)	(Reiche and Höfig 2013)	(Umbrich et al. 2015)	(Reiche et al. 2014)	(Sunlight Foundation 2016)	(Open Knowledge Foundation)	(Open Data Monitor 2016)	(European Data Portal 2017)	(World Wide Web Foundation)
Accessible			X							X
Accuracy	X		X	X						
Availability		X	X					X	X	
Completeness	X		X	X		X				X
Compliance	X									
Consistency	X									
Dataset url		X			X					
Easy access				X	X	X	X		X	X
Free usage cost						X	X		X	X
License				X	X	X	X	X	X	X
Machine readable format					X	X	X	X	X	X
Metadata availability			X							
Metadata completeness			X					X		
Non-discrimination						X				
Open format						X	X		X	X
Openness				X						
Permanence						X				
Primary/ primacy						X				
Re-usability of data			X	X					X	
Timeliness/ up to date	X	X				X	X		X	X
Use of data				X					X	

**Table 1: Most frequently used dimensions to measure data quality in OGD**

By listing dimensions that occur the most in the literature, we rapidly realized that most of the authors use different dimensions to describe the same idea. For instance, we constantly encounter the dimensions *availability*, *accessibility* or *re-usability of the data*, which without clear definition may lead to the same interpretation. This is maybe the reason why the ODI (Open Data Index 2016) recommended that researchers, developers and policymakers adhere to common data standards. In light of the foregoing considerations, we chose to follow the good practice standards of the Sunlight Foundation (SF), testing whether they were respected by Swiss OGD platform publishers. Its principles have been developed in order to "empower the public's re-use of public data held by governments". They re-group most of the important open source publishing rules' (Sunlight Foundation 2018) and cover a great variety of dimensions such as data accessibility, availability, technicality and legality. Even if the vocabulary used is not always similar, we also chose to follow the OGD good practice standards propagated by the SF (Sunlight

Foundation 2018) because its principles correspond to the quality dimensions most often cited in the literature and most often recommended by the most advanced OGD platforms (e.g. *data.go.uk*, *data.gouv.fr*, *data.gov*, etc.). Contrarily to other recommendations, the meaning of each 10 principles is deeply and comprehensively defined. While some principles meaning could be discussed (e.g. why speaking about usage cost when we speak about open data, or commonly owned formats such as Excel while it is subject to licence requirements), these 10 principles provide a very good representation of the challenges of publishing and using OGD, from the perspectives of both data providers and consumers. In our view, they remain essentials to ensure feasibility and continuity in the OD re-use. Finally, the SF is one of the first foundations to take an active interest on the special case of government data. We summarise the 10 principles in Table 2 (for a full-length description of the principles, see Appendix A).

#1 Completeness	Resources published on OD platforms should contain all raw information and metadata defining and explaining their content.
#2 Primacy	Resources published on OD platforms should also include the original information released by the government.
#3 Timeliness	Resources should be available to the public in a timely manner.
#4 Easy access	Resources published on OD platforms should be easy to find and download.
#5 Machine-readable format	Resources should be stored in a machine-readable format (i.e. should be processable by a computer)
#6 Non-discrimination	Resources published on OD platforms should be accessible without having to identify oneself (e.g. via needing to log in) or having to provide a justificatory reason.
#7 Open format	Resources should be usable without proprietary software.
#8 Open licencing	Resources published on OD platforms should use an open licencing model.
#9 Permanence	Resources published on OD platforms should be accessible by machines and humans over time.
#10 Usage cost	Resources should be available for free.

**Table 2: Summary of the 10 Sunlight Foundation’s principles for data quality**

Usually, only open data portals on the national level are evaluated and very little is known from portals at the regional, local or organizational level. The approach we present, which we named the Compliance Index, analyses whether Swiss public organisations are complying with the 10 principles recommended by the SF. We proposed that index in order to go further in data quality measurement and also to fill the gap in data quality evaluation at the organizational level. We chose to apply the Compliance Index to Switzerland, partly because the Swiss OGD platform *opendata.swiss* have been launched in 2016 and only two years after its opening, authorities already observe trouble with its use. Furthermore, the first Swiss OGD strategy (from 2014 to 2018) is coming to an end, and that is, therefore, an appropriate time to evaluate it but also to propose an appropriate recommendation to the authorities in charge of the platform and data providers. Due to its federal structure, Switzerland offers to our study different analysis levels allowing us to observe the OGD publication through three types of governance (i.e. federal, cantonal and municipal).

## METHODOLOGY

In developing an index that can be used for the quality assessment of data publication, we chose to follow a pragmatic epistemology (Goldkuhl 2012). The essence of pragmatic research lies in the interplay between actions and change: To alter certain aspects of reality (in our case data quality of published OGD), actions are required. Empirical evidence is essential to initiate change and possibly alter the current situation to a more desired state. Actions and their impact can also contribute to further cognitive clarification and



development. This way of researching contrasts with, for example, purely descriptive research that primarily seeks to explain reality by using models (or a structure of relations) and which uses methods that emphasize the discovery of new knowledge and verify existing (structural) knowledge without deliberately distorting reality. In this sense, pragmatic research is problem-driven research by necessity since complex problems do not respect philosophical, historical, or disciplinary boundaries of science. That said, we followed the common steps of pragmatic research, which can be summarized as problem analysis, artefact construction, artefact application, and interpretation and learning from the results. After having outlined the current issues regarding data quality and OGD in the previous sections, we now turn to the description of the construction of our compliance index and show its application based on the Swiss context.

### 1.1 Input data for building the compliance index

*Data collection.* To see whether governmental authorities and Swiss public administrations are following the aforementioned OGD good practice standards, we analysed the available metadata that is published alongside a data resource and that usually can be downloaded from an OGD platform. Like many OGD platforms worldwide (Kirstein et al. 2019), *opendata.swiss* uses the Comprehensive Knowledge Archive Network (CKAN), a web-based open-source management system to store and distribute OGD (Open Knowledge Foundation 2018a). We harvested the available CKAN metadata, such as resources titles, resources authors name or resources formats, by creating specific requests so as to obtain a valid catalogue endpoint for extracting all resources-related metadata. All open data catalogues generated with CKAN are organised by group, datasets and resources. One group contains many datasets and one dataset may contain at least one or more resources. We first used the application programming interface (API) offered by CKAN to create requests and then used the excel tool “power pivot” to download the metadata of all published resources. The data downloading was done on 26 May 2018, when the portal contained entries from 44 public organisations – a total of 17,777 data resources.

*Data manipulation and imputation of missing metadata values.* In phase 2, we extracted and analysed all relevant resources by means of Excel using the power pivot functionality. To remove inconsistencies in the obtained data, we cleaned and pre-selected specific CKAN metadata fields in Excel on the basis of their accessibility, interpretability and coherence in terms of the 10 principles. The final CKAN metadata fields we used to create the Compliance Index appear in Appendix B. In case that metadata had missing values for some resources, we chose the single imputation method, including the hot deck imputation solution. This approach considers missing values as part of the analysis by imputing them another value. Thus, in case of a missing value, we imputed a negative answer (*no*), as we will now explain.

### 1.2 Treatment of the metadata

*Questions set.* The coding of the metadata was realised via a set of questions we developed on the basis of the 10 principles (for more details, see Appendix C). For instance, to assess the compliance to Principle 5 (machine-readability), we asked whether the resources were available in a format such as RTF, CSV, XML or JSON, i.e. whether the value of the CKAN metadata field *resources.format* corresponded to one of the pre-defined machine-readable formats. Such questions examined whether the metadata (and metadata value) adhered to the good practice standards of publishing OGD. In this sense, each of the 10 principles is linked to a specific question. A question can engage one or many CKAN metadata fields if it is not mutually exclusive. For instance, the CKAN metadata field *resources.rights* was used in multiple questions to answer the adherence to Principles 8 (open licencing) and 10 (usage costs).

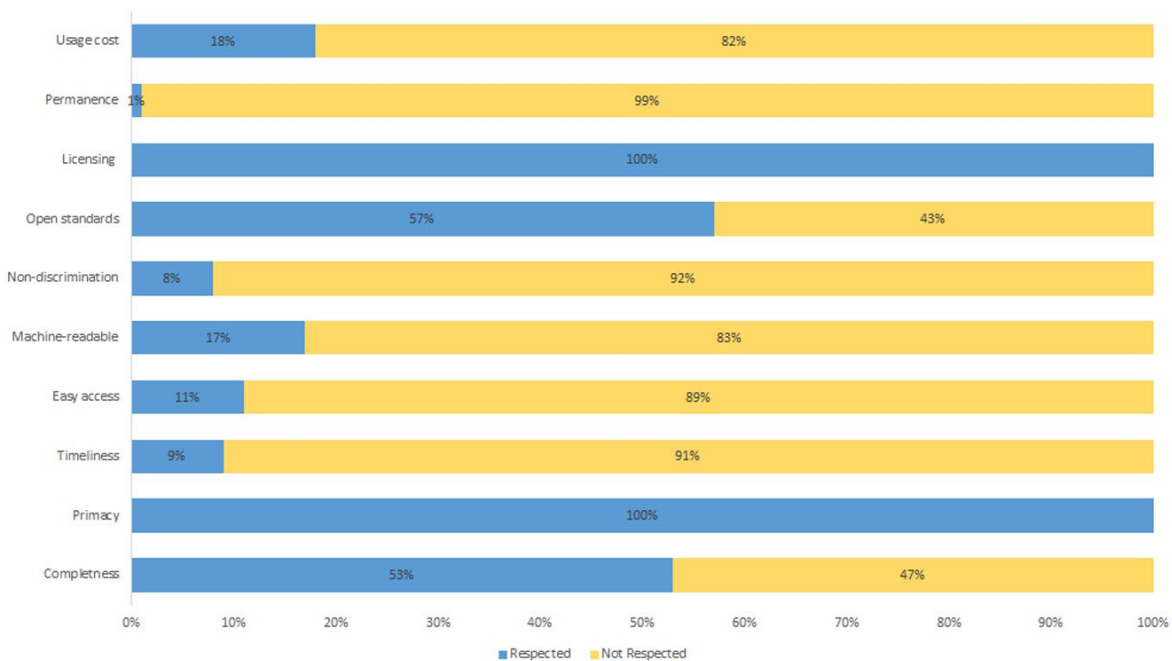
*Allocations of points.* As noted, we transformed the answers into binary numbers. A positive answer (*yes*) concerning the application of a principle to a resource = one point, while a negative answer (*no*) = zero points. By using binary numbering, nominal answers easily translated in ordinal answers. This ensured a computation of the scores obtained for each resource (index per resources) and the creation of an index for organisations (Compliance Index). To compute the index per resources and the Compliance Index, we opted for the equal weighting method. This method implies an equal status for all the principles. In this sense, the

index per resources and the final Compliance Index considered all the principles as equally important concerning publishing OGD. Thus, these two indices do not favour certain principles over others.

*Scoring of the Compliance Index.* For each published resource by an organisation, every principle grants a number of points; the sum of the points constitutes the score of the index per resources. On this basis, we computed the Compliance Index per organisation, which corresponds to the average index scores per resources. Considering our study purpose, a linear aggregation method allowed for the computation of an index per resource, but also provides information on principles applied to a given resource. For instance, a score of seven on the index per resources means that a given resource applied seven of the 10 good practice standards. This also indicates that the organisation that publishes these resources must modify some of its practices if it is to be considered a fully compliant OGD provider. In sum, the Compliance Index provides a general overview of how organisations publish OGD, and the index per resources allows for a deeper examination of how each organisation manages its resources.

## RESULTS

We will now present the results of applying our method to data providers publishing their data on the Swiss OGD platform *opendata.swiss*. Figure 1 provides a general overview of the current compliance levels regarding all 10 principles over all 17,777 data resources. Segments in blue represent the percentage of resources that respected a specific principle, while segments in yellow constitute the percentage of resources that did not. As the illustration suggests, not all good practice standards were followed equally. On *opendata.swiss*, data resources only respected the principles of primacy and licencing to 100%. Only two other principles, completeness and open standards, were respected by more than 50% of data resources. Permanence, timeliness and non-discrimination were given in only less than 10% of the data resources.

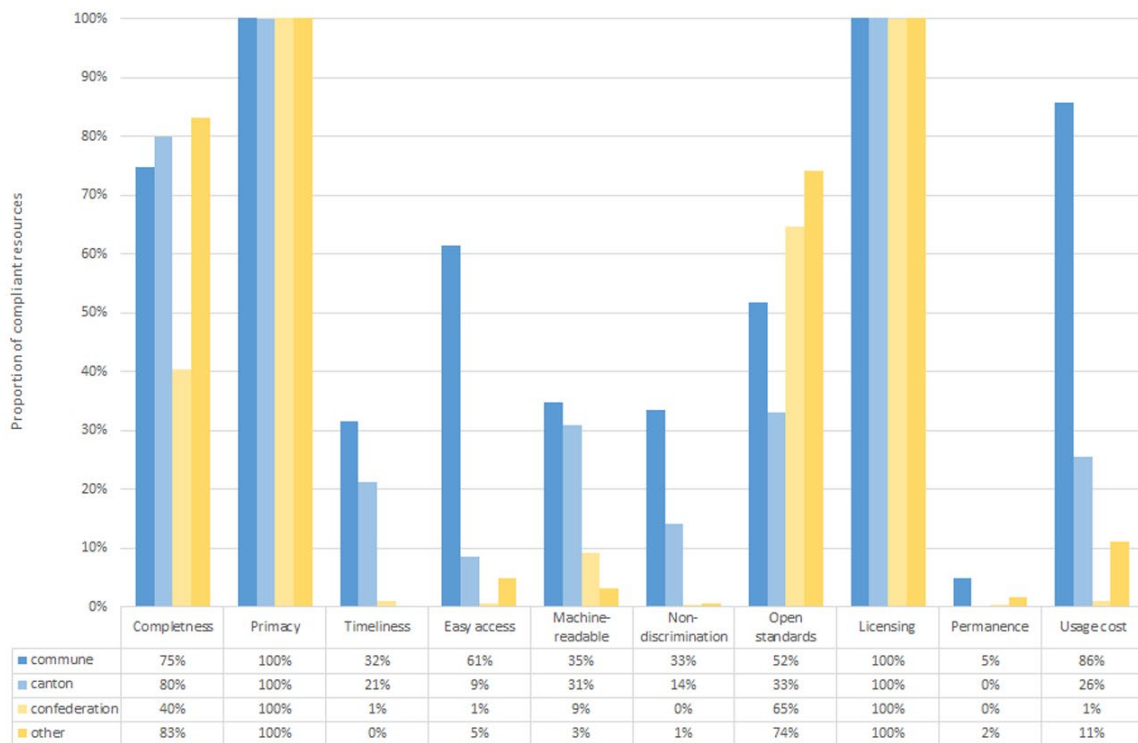


**Figure 1: General overview of adherence to the 10 good practice standards**

Figure 2 shows the percentage of data resources that complied to the 10 principles by the data provider's political level (communal, cantonal, federal and other). As we can see, communes tended to follow good practice standards more than the three others – it was the only group of data providers that largely applied

the easy access and usage cost principles. At the canton level, although some principles were not respected as much as on the communal level, we could still determine efforts to comply with 9 out of 10 principles. At the federal level, only the principles of completeness, primacy, open standards and licencing were followed.

Completeness, primacy, machine-readability, open standards and licencing are the principles that were most followed by data providers across most political levels. For instance, completeness was largely followed by communal (75%), cantonal (80%) and other public actors (83%), while these percentages only reached 35%, 31% and 3% in the case of machine-readability.



**Figure 2: Compliance of data resources by political level**

Table 2 shows the percentages of data resources that were respected for each principle based on OGD categories (e.g. data for tourism, trade, health-care or agriculture). We did this analysis in order to understand possible differences in publishing OGD across different public sector domains, assuming that data providers face different complexity levels (e.g. different regulations concerning privacy and data protection).

Despite not finding a clear pattern from this analysis, some OGD categories stood out. For instance, a small share of administrative data resources seemed to respect all good practice standards and may, overall, be more compliant than other categories. To a lesser extent than administrative data resources, OGD categories such as geography, legislation and territory also adhered to most of the principles, while educational, energy and industrial data resources appeared to have more difficulties applying the principles. Surprisingly, a principle may be broadly respected in one domain, yet may be completely neglected in others. For instance, the principle of completeness was applied in 92% of data resources published in the legislation category, while only 17% of the data resources under the trade category provided indications about their completeness. Another example is the principle relating to data resources' usage costs. While 67% of data

resources published under the trade category respected this principle, the percentage only reached 2% in case of data resources in the industrial category.

Principles:	Completeness	Primacy	Timeliness	Easy access	Machine-readable	Non-discrimination	Open standards	Licensing	Permanence	Usage cost
OGD Categories:										
<b>administration</b>	83%	100%	4%	34%	36%	23%	53%	100%	13%	50%
agriculture	50%	100%	3%	5%	11%	4%	53%	100%	1%	6%
construction	44%	100%	8%	14%	12%	5%	47%	100%	5%	18%
crime	59%	100%	0%	0%	17%	0%	68%	100%	1%	3%
culture	64%	100%	1%	5%	14%	3%	64%	100%	4%	11%
education	27%	100%	1%	3%	6%	2%	64%	100%	1%	6%
energy	60%	99%	0%	3%	8%	2%	21%	100%	0%	8%
finances	83%	100%	54%	6%	4%	2%	20%	100%	7%	13%
<b>geography</b>	61%	100%	2%	22%	33%	27%	43%	100%	0%	49%
health	54%	100%	6%	4%	6%	2%	80%	100%	0%	6%
industry	21%	100%	1%	0%	5%	0%	55%	100%	1%	2%
<b>legislation</b>	92%	100%	0%	8%	32%	0%	56%	100%	5%	15%
mobility	68%	100%	3%	25%	10%	8%	50%	100%	0%	30%
national economy	57%	100%	1%	14%	13%	3%	66%	100%	2%	16%
politics	84%	100%	4%	5%	39%	2%	50%	100%	3%	8%
population	34%	100%	3%	6%	9%	1%	60%	100%	0%	7%
prices	48%	100%	2%	3%	10%	2%	79%	100%	0%	9%
public order	91%	100%	0%	1%	34%	1%	54%	100%	3%	8%
social security	48%	100%	0%	10%	14%	7%	62%	100%	0%	16%
statistical basis	41%	100%	1%	1%	2%	1%	91%	100%	0%	2%
<b>territory</b>	71%	100%	1%	17%	28%	12%	48%	100%	0%	24%
tourism	31%	100%	0%	3%	6%	3%	73%	100%	0%	10%
trade	17%	100%	0%	0%	0%	0%	58%	100%	17%	67%
work	55%	100%	3%	0%	17%	0%	68%	100%	0%	2%

**Table 2: Compliance to data resources by OGD categories**

Table 3 provides an overview of the final scores of the Compliance Index per public organisation and political levels. For anonymisation reasons, organisation names are not displayed and have been replaced by an identification number. To recap, the Compliance Index is an average of the scores obtained by every resource (published by an organisation) for each principle. For the analysed timeframe, we found that the maximum number of principles respected by organisations was 5.9 and that the minimum number was 2. Only 16% of public organisations complied with five or more good practice standards. Our results also indicate that the maximum principles respected by a data provider at the canton level was almost 6 and that the minimum was 2. Concerning the other political levels, a minimum of 2 principles were respected, but the maximum was only 4.7 principles, with a mean index score of 3.5. On average, data providers at the communal level were more compliant with the 10 principles (4.2). Data providers at the canton and federal levels respected on average 3.8 principles. A lower mean index score was only reached by public organisations not attributable to one of the aforementioned political levels (3.5). Indicatively, the average score of all resources published on the *opendata.swiss* platform, independently of their political level, was 3.75 out of 10. In concrete terms, this means that public organisations in Switzerland only respected about 4 out of 10 principles when publishing OGD.

ID	Federal organisations
1	2.5
2	2.8
3	2.9
4	2.9
5	3.1
6	3.1
7	3.2
8	3.2
9	3.2
10	3.3
11	3.3
12	3.4
13	3.5
14	3.9
15	4.0
16	4.5
17	4.6
18	5.0
19	5.6
20	5.6
21	5.6
Average	
	3.8

ID	Cantonal organisations
22	2.0
23	2.5
24	3.1
25	3.5
26	3.5
27	4.0
28	4.0
29	4.0
30	5.8
31	5.9
Average	
	3.8

ID	Communal organisations
32	2.5
33	4.0
34	4.0
35	4.0
36	4.2
37	4.8
38	5.9
Average	
	4.2

ID	Other organisations
39	2.0
40	3.0
41	3.3
42	3.8
43	4.2
44	4.7
Average	
	3.5

**Table 3: The Compliance Index scores for every organisation**

## DISCUSSION AND CONCLUSION

Our analysis results revealed that, in Switzerland, the adherence to good practice standards for publishing OGD is fairly low. Our Compliance Index also demonstrated that not all principles were equally well implemented by data providers. For instance, while we found that all resources published on *opendata.swiss* followed the primacy principle as well as indicated a licencing model, the principle of permanence saw less respect.

A possible interpretation for this disparity could be the fact that data providers allocate different importance to each principle. For instance, disclosing the correct licencing model is essential in order to inform data consumers about their data re-use options or could also be motivated by self-interest to protect current or future copyright interests. Conversely, respecting the permanence principle means keeping data resources available online for data consumers over an indefinite period. This is complex and costly for data providers, and requires a good understanding of existing data storage and network structures.

Further, our coding procedure could have led to this impression. Not all principles were always equally measured. For some principles, we needed to develop approximations based on multiple metadata fields, since no single field explained the principle in its entirety. For others (e.g. open format) only one metadata field was sufficient. Sometimes, the range of allowed values in determining compliance was unambiguous. For instance, the principle of timeliness can be interpreted in different ways (Emran 2015). Vetrò et al. (2016) defined timeliness as the presence or absence of the updated version of a data set, while Atz (2014) measured it via the percentage of a data sets' up-to-dateness. One other case, machine-readable formats, is considered to be common knowledge, yet there is no standard coding list we could fall back on.

But there could also be a much simpler explanation. It could just be that the platform intermediary (i.e. the Swiss Federal Archives) marked certain metadata fields as compulsory, while making others optional for data providers. Accordingly, it would be impossible for data providers to publish OGD without for instance determining a licencing model or indicating the resource's originator. Since the existing technical and organisational guidelines (e.g. DCAT-AP for Switzerland) or end-user handbooks are not always written in ways that can be understood by a non-technical expert in a public administration, it could be that the technical complexity of publishing OGD, paired with ignorance, may have influenced data providers to only publish metadata they could fully understand.

When looking at the differences at the distinct political levels, we see an interesting result. On average, data resources published by communal organisations are more compliant, particularly in terms of ease of access, non-discrimination and usage costs. A plausible explanation for this could be that these actors are closer to citizens and are therefore more attentive to their needs, or that federal agencies must deal with a greater diversity and complexity of data resources, making publishing OGD more difficult and error-prone. However, it could also be that certain actors (e.g. Federal Office of Statistics, Federal Office of Topography) are not only participating in sharing OGD because they want to (e.g. as communal or cantonal organisations), but because they have to. Enforcing top-down diffusion of OGD could strengthen the belief that OGD is a burden to public administrations, rather than a service to citizens (Janssen et al. 2012). Nonetheless, our results are surprising, given that prior research has found that federal-level governmental agencies often have a higher maturity and readiness to diffuse OGD than communal or regional (cantonal) public administrations (Zuiderwijk et al. 2018).

We have developed a Compliance Index and have conducted a comprehensive empirical analysis on the basis of metadata published on *opendata.swiss*. The purpose was to see whether public organisations in Switzerland followed good practice standards when they publish their data as open access (and not to determine what defines data quality in the OGD context). We developed the Compliance Index mostly because we observed that OGD re-use might increase only if data providers improve the quality of the data, and consequently applied good practice standards. We argue that data providers such as public organizations, cities and communes, among others, not only need quality platforms analysis but also clear and details feedbacks on their resources. For this purpose, organizations need to know which resources in which dataset does not have an adequate format or is not up-to-date. This is one objective of the Compliance Index. In our view, our Compliance Index is a promising first step towards better understanding data quality dimensions and sharing practices of governmental authorities and public organisations on OGD platforms. It offers the possibility for data providers to clearly and rapidly identify their deficiencies or potential *OGD champions* who have a better understanding of the OGD publishing process. Further, combining with a platform traffic analysis the Compliance Index can also be used as a first approximation to understand what quality dimensions are the most useful to allow data to be downloadable and potentially re-used and, thus, may be of particular interest to data consumers and platform operators. Although several authors explored the question on the data quality, most of the evaluations investigated on open data platform's quality at national level (Máchová and Lnénicka 2017; Open Knowledge Foundation 2018c; Ubaldi 2013) and only a few studies examined metadata resources quality's at the organization level (Berners-Lee 2006; Maurino et al. 2014). Our Compliance Index draws on the studies of Vetrò et al. (2016) and Maurino et al. (2014) but propose to go deeper by analysing automatically the OGD platform content, from data providers perspective. It directly evaluated the published CKAN metadata of data resources to then performed evaluation at dataset and resources levels. An analysis on this level of granularity not only supports distinct data providers to improve certain aspects of their OGD publishing procedures, but also, intermediaries and operators of OGD platforms by receiving key performance indicators that helps them to monitor and understand their platform's current state of OGD quality. From the perspective of public managers, such an index could support planification and process organization. As a matter of fact, public managers could gain better knowledge on the evolution of the "opening process" their services and be able to identify the next steps for further implementing open practices. The compliance index allows a dual action approach composed of identifying problems regarding data publication and resolution. The compliance index could

also reinforce trust regarding governments and politicians' actions. For instance, thanks to this tool, citizen could see the evolution of the open government strategy and the corresponding level of realization in practice. By making the compliance index available to OGD organizations, public managers could better collaborate with each other and therefore, strengthen the most effective techniques that promote data publication (e.g. anonymization rules, recurring problems, etc.). Furthermore, CKAN is currently the most used web application in the public sector, such as used in the United Kingdom, United States, and many other countries (W3C 2019), for building open data catalogues (Kirstein et al. 2019). Thus, developing an index based on CKAN allows to not only improve OGD re-use in Switzerland but also other for other countries, with the aim to emphasize the quality assessment of data publication.

Since developing the Compliance Index required compromises and approximations, our study has limitations. First, we narrowed the analysis to the CKAN metadata available on *opendata.swiss*, because this was easier to extract and scrutinise than analysing the published data sets, which would have required multiple different approaches, including text mining, image mining and others, given that a wide variety of administrative, geographical and statistical data sets are made available by different public organisations. Also, determining content's quality is often subjective and highly context dependent. In this sense, our Compliance Index remains a proxy to measure whether or not the dimensions that define data quality are well respected by governmental authorities and public organisations but is nonetheless based on the analysis of actually published data and not on survey data about attitudes, perceptions or 'wishful thinking' of interested or involved (and thus not always unbiased) respondents. Second, we based our understanding of good practice wholly on the 10 principles defined by the Sunlight Foundation. We acknowledge that there are different standards and guidelines, suggesting fewer or more circumstantial principles for directing OGD initiatives. However, the SF regroups 30 open government advocates coming from fields such as research, public administrations and internet. We are convinced that they appear a good starting point. Third, there were also issues concerning operationalisation of measurements and the coding procedure. Some of the principles, such the ease of physical and electronic access, could not be examined in much depth, given that user feedback is not collected systematically on *opendata.swiss* (we simply assumed that it is easier and more practical for end-users to download data directly from the platform than to be redirected to another download website). Fourth, we listed the most used OD quality dimensions by realising a comprehensive narrative literature review. However, a systematic literature review would be more appropriate and could provide a more accurate idea of the different dimensions used. Then, we used binary operators when deciding whether or not a principle was respected. In certain situations, this may be a bit simplistic, since there are different levels of adherence or non-adherence (e.g. free re-use of data after registration is still better than paid use of OGD only). Finally, the use of the same metadata fields to approximate principles could cause multicollinearity problems. Using the same fields to define different principles may provoke a chain reaction, if one organization score low in one aspect. For example, organizations with empty fields in a certain category (e.g. *resources.rights*) will be penalized as often as this field is used.

A more nuanced view, particularly in this regard, would be helpful. We trust that our paper will motivate other researchers to delve into the complex dynamics and controversial nature of OGD quality and of ways to measure it. It is only if we have a proper approximation to measure OGD quality that it is possible to make reasonable projections on OGD's value and to corroborate the currently uncontested relationships between OGD, innovation and economic growth.



## Appendix

### Appendix A: the 10 data quality principles of the sunlight foundation

	<b>Definition</b>
Completeness	Data sets released by the government should be as complete as possible, reflecting the entirety of what is recorded about a particular subject. All raw information from a data set should be released to the public, except to the extent necessary to comply with federal law regarding the release of personally identifiable information. Metadata that defines and explains the raw data should also be included, along with formulas and explanations for how derived data was calculated. Doing so will permit users to understand the scope of the available information and to examine each data item at the greatest possible level of detail.
Primacy	Data sets released by the government should be primary source data. This includes the original information collected by the government, details on how the data was collected, and the original source documents recording the collection of the data. Public dissemination will allow users to verify that information was properly collected and accurately recorded.
Timeliness	Data sets released by the government should be available to the public in a timely way. Whenever feasible, information collected by the government should be released as quickly as it is gathered and collected. Priority should be given to data whose utility is time-sensitive. Real-time information updates would maximise the utility the public can obtain from this information.
Ease of physical and electronic access	Data sets released by the government should as accessible as possible, with accessibility defined as the ease with which information can be obtained, whether through physical or electronic means. Barriers to physical access include requirements to visit a particular office in person or requirements to comply with particular procedures (such as completing forms or submitting FOIA requests). Barriers to automated electronic access include making data accessible only via submitted forms or systems that require browser-oriented technologies (e.g. Flash, JavaScript, cookies or Java applets). By contrast, providing an interface for users to download all of the information stored in a database at once (known as bulk access) and the means to make specific calls for data through an application programming interface (API) make data much more readily accessible. (An aspect of this is <i>findability</i> , which is the ability to easily locate and download content.)
Machine-readability	Machines can handle certain kinds of inputs much better than others. For instance, handwritten notes on paper are very difficult for machines to process. Scanning text via optical character recognition (OCR) results in many matching and formatting errors. Information shared in the widely used PDF format, for instance, is very difficult for machines to parse. Thus, information should be stored in widely used file formats that easily lend themselves to machine processing. (When other factors necessitate the use of difficult-to-parse formats, data should also be available in machine-friendly formats.) These files should be accompanied by documentation related to the format and how to use it in relation to the data.
Non-discrimination	<i>Non-discrimination</i> refers to who can access data and how they must do so. Barriers to data use can include registration or membership requirements. Another barrier is the uses of a <i>walled garden</i> , which is when only some applications are allowed access to data. At its broadest, non-discriminatory access to data means that any person can access the data at any time without having to identify themselves or provide any justification for doing so.
Commonly owned or open standards	Commonly owned or open standards refer to who owns the format in which data is stored. For instance, if only one company manufactures the programme that can read a file where data is stored, access to that information is dependent on use of the company's processing programme. Sometimes that programme is unavailable to the public at any cost, or is available but for a fee. For instance, Excel is a fairly



	commonly used spreadsheet programme that costs money to use. Freely available alternative formats often exist via which stored data can be accessed without the need for a software licence. Removing this cost makes the data available to a wider pool of potential users.
Licencing	The imposition of terms of service, attribution requirements, restrictions on dissemination and so on are barriers to public use of data. Maximal openness includes clearly labelling public information as a work of the government and available without restrictions on use as part of the public domain.
Permanence	The capability of finding information over time is referred to as permanence. Information released by the government online should be <i>sticky</i> : it should be available online in archives in perpetuity. Information is often updated, changed or removed without any indication that an alteration has been made. Or, it is made available as a stream of data, but is not archived anywhere. For best use by the public, information made available online should remain online, with appropriate version-tracking and archiving over time.
Usage cost	One of the greatest barriers to access to ostensibly publicly-available information is the cost imposed on the public for access, even when the cost is small. Governments use a number of bases for charging the public for access to their own documents: the costs of creating the information; a cost-recovery basis (cost to produce the information divided by the expected number of purchasers); the cost to retrieve information; a per page or per inquiry cost; processing cost; the cost of duplication, etc.

**Appendix B: CKAN metadata fields used to create the Compliance Index**

	<b>Metadata fields</b>
Completeness	<b>Original metadata</b> metadata_modified contact_points.e-mail resources.issued resources.download_url resources.rights organisation.name <b>Created metadata</b> Metadata existed, RawInformationExist
Primacy	<b>Original metadata</b> contact_points.e-mail
Timeliness	<b>Original metadata</b> Accrual_periodicity Modified <b>Created metadata</b> annual, semi-annual, quarterly, monthly, monthly, weekly, biweekly, daily, continuous, sum, modifiedTransformed, time until today, tau
Easy access	<b>Original metadata</b> resources.rights resources.download_url
Machine-readable format	<b>Original metadata</b> resources.format
Non-discrimination	<b>Original metadata</b> resources.rights resources.download_url resources.url
Commonly owned or open standards	<b>Original metadata</b> resources.format
Open licencing	<b>Original metadata</b> resources.rights
Permanence	<b>Original metadata</b> resources.download_url resources.url
Usage cost	<b>Original metadata</b> resources.rights

### *Appendix C: Question set and chaining logic*

	<b>Questions</b>	<b>Chaining logic</b>
Completeness	Q1: Is the metadata complete?	If the raw information and the metadata of this resource exist = 1, else 0
Primacy	Q2: Is there an e-mail address for a contact point/support contact?	If an e-mail address to contact the originator exists = 1, else 0
Timeliness	Q3: Is the resource up to date?	If the tau of data > 0.5 = 1, else 0
Easy access	Q4: Is the data available in bulk?	If resources.download_url exists and resources.right is NonCommercialAllowed-CommercialAllowed-ReferenceNotRequired = 1, else 0
Machine-readable format	Q5: Is the resource available in machine-readable format?	If the format used is machine-readable = 1, else 0
Non-discrimination	Q6: Does people have a limited access to the resource?	If a downloadable link exists, the licence to use data is fully open and the data machine-readable = 1, else 0
Commonly owned or open standards	Q7: Is the resource in an open file format?	If the variable resources.format is filled with an open format = 1, else 0
Open licencing	Q8: Is the resource openly licenced?	If licencing information is available = 1, else 0
Permanence	Q9: Is the published resource available over time?	If a direct downloadable link exists and if it is different from the URL link = 1, else 0
Usage cost	Q10: Is the resource freely available?	If resource the resource uses an open format and open licence = 1, else 0

## *Paper # 2*

## **Paper#2 - Different shades of perception**

**Title:** Different shades of perception: How do public managers comprehend the re-use potential of open government data?

**Author:** Auriane Marmier and Tobias Mettler

**Published in:** Proceedings of the 41th International Conference on Information System

**Abstract:** Open government data (OGD) initiatives have become an important part of digital transformation strategies and a means for supporting digital entrepreneurship. However, several studies have shown that OGD is experiencing major barriers or even a standstill in certain countries. Little is known about what public managers think of OGD and what uses they ascribe to it. Accordingly, our paper tries to shed some light on the cognitive structures of public managers. Using the RepGrid technique as empirical method and affordance theory as theoretical grounding, our study explores how public managers working in different branches and levels of government interpret and make sense of OGD. Our findings allow for a better understanding of how managers comprehend OGD as concept and to what extent they share the same vision regarding how to re-use OGD in different application domains.

## Paper 2

Paper#2 - Different shades of perception.....	58
Introduction.....	60
Background.....	61
On open government, open data, and government data.....	61
On affordance theory and its application to open government data .....	62
Research Method .....	62
Repertory Grid Technique .....	62
Data Collection.....	63
Step 1 - Definition of elements .....	63
Step 2 - Elicitation of constructs .....	64
Step 3 - The scoring process .....	64
Data Analysis.....	64
Results.....	66
Structuring and limiting the re-use potential of OGD .....	66
Different perceptions on OGD re-use potential.....	66
Dimension 1: Group vs. Individual behavior .....	66
Dimension 2: Openness vs. Constraints.....	68
Dimension 3 – Utilitarian vs. non-utilitarian effects.....	69
Dimension 4 – Active vs. Passive actualization.....	70
Conclusion .....	71
Future research and limitations.....	73

## Introduction

Democracy is based on the premise that elected governments provide transparent information about their intended and actual implementation of laws and public policies as well as their use of public resources (Ruijter et al. 2017). Nevertheless, in many countries government data has only been accessible for a limited group of consignees, such as think tanks or lobbyists, while most of the citizens and businesses have to undergo a lengthy application process or are completely cut off from accessing information online (Peled 2011). With the Open Internet movement gaining a certain level of media attention (Financial Times 2020), along with an emerging rise of anti-government sentiments, many democratic countries have started to implement Open Government Data (OGD) initiatives in the last 15 years with the objective to increase transparency and trust between citizens and their governments (Lv and Ma 2019). However, the global economic crisis had a fair share in changing the direction of OGD, from being an instrument for promoting good governance, counteract corruption (Bertot et al. 2010), or empower citizen participation (Lee and Kwak 2012), to being an enabler for data-intensive business ventures or as a source for reducing government spending. For many governments, OGD initiatives have therefore become an important part of their digital transformation strategies and efforts for establishing digital entrepreneurship in the respective country (Ubaldi 2013). However, to date, very few initiatives pushed by governments demonstrate promising results (Danneels et al. 2017; Martin 2014). To address this mismatch between objectives and results, several hypotheses have been considered. Many authors argued that governments, seeing their OGD environment evolving (i.e. including more resources, systems and actors), were face to unprecedented technical and organizational challenges (Bakıcı et al. 2013; Corbett et al. 2018; Cranefield et al. 2014). A plethora of studies have therefore focused on mechanisms and actors that explain barriers and challenges of OGD initiatives (Harrison et al. 2012; Heimstädt et al. 2014; Parycek et al. 2014; Van Schalkwyk et al. 2016; Zuiderwijk et al. 2014). Among the assumptions presented in these studies, Wang and Shepherd (2020) emphasized accessibility as most prominent issue and noticed that only a small minority of government data appears to be really open. They described the current OGD as "*a swamp of non-granular, unstructured, aged and frequently inaccessible*". Sandoval-Almazán et al. (2017) presumed that government transformation through digital government builds on organizational efforts and accentuates individuals' issues. Many authors also suggested that OGD initiatives, in order to be successful, should center more on ecosystems and relations developed between OGD intermediaries and their environment (Chan 2013; Corbett et al. 2018; Gonzalez-Zapata and Heeks 2015; Sandoval-Almazán et al. 2017; Young and Yan 2017).

However, to our knowledge, few authors concentrate on the people at grassroot of data opening - public managers. As responsible for OGD implementation in their department, organization, or even nationally, public managers have a key role in aligning the strategic goals of institutions with national or international OGD policies (Lee and Kwak 2012). According to Whitley and Hosein (2008) managers' attitudes and perceptions towards technologies may not only impact policies but also results of policy decisions. Attard et al. (2015) and Yang et al. (2015a) share the same idea and noticed that cognition and perception of public managers towards OGD implantation crucially determines the strategic direction and impact of an organization. Indeed, cognition – as the human capacity to perceive, interpret, and reason about environmental - is connected to humans actions (Carayannis et al. 2003; Hardless et al. 2015). However, with the increase of OGD initiatives, government agencies (as main publishers of data) and their managers are under pressure, which may affect their perceptions and reactions. In order that OGD is re-used, it has to be released in a timely and machine-readable manner but also follow strict de-identification rules to avoid data disclosure risks, which asks extra-effort and increased work-load from public managers. Furthermore, data publication is often unrewarding, costly, and not providing any credit or acknowledgments to the ones who have made an effort (Fane et al. 2019). In the light of the above, we argue that it is important to investigate public managers because their perception of OGD concept and its potential re-use, may influence public managers behavior (European Commission 2020) and lead them to take decision on the basis of personal gain and non-objective selection criteria. Public managers' perception - as they may

directly (or unconsciously) influence the realization of an OGD initiative - could not only affect public institutions' transparency and governance but also individuals' privacy, data protection and contribute to increase what Taylor (2017) call dataveillance (Paspatis et al. 2017; Taylor 2017).

Comprehending perceptions as possible accelerators (or decelerators) of actions, we hence seek to investigate the following research question: *how do public managers perceive the re-use potential of OGD?* To answer this research question, we take an affordance theory perspective and use the repertory grid technique for systematically exploring the perceptions of public managers.

Our findings allow for a better understanding of how managers comprehend the OGD concept and to what extent managers share the same view regarding characteristics needed to actualize the potential of OGD re-use. Our results show how the *Repertory Grid* technique can be used for capturing the underlying mechanisms of perception. We conclude by offering some suggestions as how governments may use this technique in order to develop their OGD initiatives or align their strategies. Our study is structured as follows: We start with a brief introduction to OGD and affordance theory. We then explain our research design and data collection and analysis, before we explain the most prominent thought patterns (i.e. constructs) of public managers. We finish by a discussion of our results and offer suggestions for future work.

## **Background**

### ***On open government, open data, and government data***

The idea of OGD finds its origin in the concept of data-driven innovation. Almost all governments worldwide produce and collect data in their daily activities given the strategic relevance of data as resource (Bates 2014; Munné 2016). In order to get benefits from data as a resource, open strategies emerged and among them the idea to share access to government data (Attard et al. 2015; Kalampokis et al. 2011b). For Gonzalez-Zapata and Heeks (2015) three concepts need to be disentangled: open government, open data, and government data. The concept of *open government* is based on the idea that governments should increase their transparency by providing more information to citizen and businesses about public activities. In doing so, a government is supposed to strengthen citizen engagement and participation, but also enhance the collaborative efforts between all the actors (i.e. across government's agencies, profit and non-profit organizations) (Wirtz and Birkmeyer 2015). *Open data* is based on the principles of free accessibility, reusability and data sharing, by anyone, for any purposes, and without any legal, technological, or social constraints (Open Knowledge Foundation 2018b). Open data as concept aims to facilitate information diffusion, innovation and economic growth (Braunschweig et al. 2012). Different from open data which origin goes back to the digital era (Conradie and Choenni 2012), *government data* is a fairly old concept and exists since (national) governments have been founded. It simply refers to any data that is produced or gathered by public organizations (Ubaldi 2013). OGD is the interrelationship of those three concepts. It is data collected and held by a government and its agencies, shared and made freely available to anyone in order to pursue open government objectives. The value of OGD rests therefore on *universal access and universal participation – without strings attached*.

To guarantee such “universal conditions”, publishers are facing important challenges. Crusoe and Melin (2018) identified technical, legal and organizational barriers as the main issues to the publication of OGD. The quality and the format of data seems to cause the main technical troubles. Given the numerous activities of a government, data may originate from different branches of government and be of geospatial, statistical, financial, or political nature (among others) and consequently stored in various (proprietary and non-proprietary) formats. However, to be considered as truly “open”, these data sets (and corresponding meta data) should be made available in a machine-readable, non-proprietary way, which frequently requires an additional complex and time-consuming transformation. The complexity of this task might demotivate publishers to make OGD accessible (Albano and Reinhard 2014). Furthermore, Crusoe and Melin (2018), argued that this apparent complexity might be reinforced by a lack of knowledge regarding OGD



publication. Defining general rules or guidelines to motivate data sharing are often not helpful as the context and aggregation level of data may vastly differ. Yet, from a legal perspective, a public manager needs to comprehend the consequences (and subtleties) of publishing organization-level data, like department expenditure, train timetables, topographic information, or individual-level data, such as the financial situation, health status of a citizen. Not being able to judge the boundaries of privacy and (state) secrecy vs. transparency reinforces the mentioned complexity and is one of the major inhibitors of OGD publication today (Huang et al. 2017). According to Susha et al. (2015b) convincing data owners to actually publish data is a key challenge. Wirtz et al. (2016) share this observation and put forward that a risk-based attitude of public servants paired with (technical) complexity of publishing OGD are major threats to open government endeavors.

### ***On affordance theory and its application to open government data***

Although legal and technical factors already limit the success of OGD initiatives (Dulong de Rosnay and Janssen 2014), further research showed that cognition also shapes the implementation process and data publication (Safarov 2019). Many authors focused their attention on the impacts of psychological and observable characteristics of managers, such as mental models or attitudes towards organizational projects (Hambrick 2007; Hodgkinson and Sparrow 2002). Jelinek and Litterer (1994) and Schwenk (1988), among others, indicated that cognition of top managers crucially determines the strategic direction of organizations and consequently the realization and implementation of initiatives. Carpenter et al. (2004) and Hambrick (2007), asserted that top managers' interpretation of contextual information influences their organizational strategic choices. These observations are in line with Gibson's *affordance theory*. According to Gibson (1966), it is the perception about the environment which leads to some *course of action*. While the original theory defines affordance as "action possibilities" that emerge from the relation between an animal and its environment (Gibson 1986), the concept of *affordance* has been used in the IS field to identify and analyze mechanisms that underlie the relationship between an IT artefact and organizational systems (Majchrzak and Markus 2012; Volkoff and Strong 2013). The theory has been popularized because it stipulates that the actualization of affordances does not only depend on artifact properties (i.e. in our case OGD properties) but also on the properties of actors and their environment (i.e. properties of OGD publishers). Therefore, the affordance effects are jointly determined by the OGD characteristics as well as by the capabilities of the OGD publishers, i.e. their faculty to perceive the OGD re-use potential, their ability to reflect upon it and to act accordingly (i.e. actualize an affordance). To highlight the importance of the role of perception on the affordance actualization, Pozzi et al. (2014) developed a theoretical framework based on four steps (cognition, perception, actualization, effects). First, they argued that the idea of affordance existence (i.e. the idea of a potential action) is based on a cognitive process, which means that to exist, an actor and an environment must be aware of the affordance may occur. Second, an affordance needs to be perceived by actors, i.e. an actor must be aware of the affordance opportunities. Third, an actor needs to adopt his behavior according to the perceived opportunities in order to finally take actions which produce effects of the perception.

We share the same view and see affordances to be relative and reliant to human perception. This, however, means that while public managers may share similar public values or act in a similar fashion, they must not see the world in the same way. The Repertory Grid Technique (RGT or RepGrid) provides information on how people perceive the world. Consequently, to propose improvements in the OGD implementation and propose ways for solving publications issues, we used the Repertory Grid Technique (or RepGrid).

## **Research Method**

### ***Repertory Grid Technique***

Originally from Kelly's Personal Construct Theory (i.e. the psychological study of personality), RepGrid is a technique for studying cognition structures and for showing how individuals perceive a given topic,

how they construe it and represent it (Kelly 2003; Oppenheim et al. 2003; Wagner et al. 2015). Even though not very prominent in the IS research field, it has been applied in some studies in the past years (Almusharraf et al. 2015). For example, it has been used for studying IS professionals' perception on important characteristics of good team members (Siau et al. 2010) or explorations on systems analysts qualities (Hunter 1997), to mention just a few examples.

The technique involves the construction of a personal grid (i.e. representation of a mental model), composed of *elements* and *constructs*. Elements represent the objects of attention for an investigated topic. They may be people, things or events (Jankowicz 2005; Siau et al. 2010). They are abstracted from the context of the investigation and can be interpreted as representations of the topic studied (Fransella et al. 2004). For instance, Siau et al. (2010) used team members as elements. Although elements are crucial, constructs are the central notion of RepGrid. Constructs are elicited during a semi-structured interview and define how individuals perceive the contents or characteristics which they attribute to the elements (Fransella et al. 2004). A construct seeks to express a contrast and is by nature always composed of two poles (Tan and Gallupe 2006). For Oppenheim et al. (2003), the only way to understand what means "good" for individuals is to understand what means "bad". In that case, "good" represents the emergent pole, "bad" represents the implicit pole of the construct. Following the example given by Siau et al. (2010), poles of a construct are the characteristics perceived as important to be a good team member (e.g. leader, honest, good communication skills). "*Leader versus follower*", "*positive general attitude versus negative general attitude*" or "*honest versus dishonest*" are examples of 3 constructs built by Siau et al. (2010) during semi-structured interviews.

As pointed out by Pozzi in the affordance theory, the cognition (i.e. affordance existence) and recognition process (i.e. affordance perception) of "potentials actions" are essential so that the actualization produces effects (outcomes). In our study, we have opted for RepGrid because we wanted to understand better the cognition and recognition process of public managers regarding the potential re-use of OGD. By focusing on how individuals perceive, interpret and reason about their environment, RepGrid is a powerful technique to better understand how (and if) they see potentials of OGD re-use.

### ***Data Collection***

The sample for our RepGrid study comprises 18 public managers, stemming from different branches (and levels) of Swiss government, including IT, Culture, Education, Justice department or Social affairs. A purposive sampling strategy was chosen to recruit the respondents that were all following a certificate of advance studies in Digital Government. The target audience consist of managers in public organizations from federal, cantonal and communal level as well as not-for-profit enterprises who are responsible for or are engaging in public digitalization projects. The subsequent data collection followed the guidelines proposed by Fransella et al. (2004), Jankowicz (2005) and Tan and Hunter (2002) and consisted of three steps.

#### **Step 1 - Definition of elements**

We started with a focus group discussion in order to define the RepGrid elements. The moderation of the focus group discussion was led by one researcher, while another researcher was taking observational notes. As suggested by Tan and Hunter (2002), we started the focus group by first explaining the RepGrid technique and our overall goal of this study. The discussion then went straight to the heart of the matter when we asked public managers to think about the question: *what could be application domains of open government data?* To encourage the exchange of reflection and avoid redundancy, we created three sub-groups of four persons, respectively two sub-groups of three persons. After a short time for reflection, we then asked public managers to write on cards five or six application domains of OGD per group. All ideas proposed by the groups were posted on a whiteboard. After an extensive discussion, we selected the most representative applications by first eliminating the applications which were off-topic. Then, we grouped the redundant ones, and finally, the participants voted for the nine most important application domains

according to their personal experience. This consensus process yielded a list of the following OGD domains: *traffic control, natural disasters management, job creation, public protection, standards and certificates, academic research promotion, fiscal optimization, heritage sites conservation and infrastructure resources management.*

### **Step 2 - Elicitation of constructs**

After having identified nine elements, the next step consisted of eliciting the *personal constructs* of each participant. There are different ways how to elicit constructs. We followed the suggestions by Tan and Gallupe (2006) and used the triadic procedure (i.e. triples of elements). The idea through this procedure is to highlight a similitude between two of the three elements by opposing them to the third. In doing so, managers express the emergent pole and the implicit pole of the construct. To achieve the triadic procedure, public managers were separated into two groups, one group of nine managers played the role of interviewers, the other nine managers played the role of interviewees. First, each interviewer wrote nine elements on cards, one per card. Then, the interviewers randomly picked three cards and propose them to the interviewees. At that moment, the interviewers asked the interviewees to describe how two of the three elements were similar, yet different from the third. Once the interviewees abstracted one construct from the proposed triad, the interviewers shuffled and exchanged cards for the following iteration. The interviewers continued the triadic procedure until the interviewees could not find any new constructs. The total constructs of an interviewee represent a personal Repertory Grid. For that study, we asked each interviewee to get between eight and twelve constructs each. Then, after each of the nine interviewees had built his or her personal RepGrid, we asked managers to reverse role-playing. Thus, for the second round of constructs elicitation, interviewers became interviewees and vice versa. To avoid the same answers between the interviewer-interviewee group, we also asked them to change their partner during the second round of constructs elicitation. One construct elicitation round usually lasted about one hour. During that part of the experiment, we moved from group to group, to follow discussions, observe reactions and take notes.

### **Step 3 - The scoring process**

After managers had played the two roles – interviewee and interviewer – each of the 18 public managers had his or her own grid, ready for scoring the potentials of OGD re-use. At this point, each participant's grid was composed of the nine OGD domains which was defined during the focus group, and a set of at least eight personal constructs. To finalize the repertory grid, we asked managers to score each OGD re-use potential on each construct. The scoring process aims to position elements along with each construct to compare elements between each other. We used a five-point scale with 1 meaning that public managers perceive the potential of OGD re-use according to the characteristic of *the emergent pole*. On the contrary, a score of 5 means that public managers perceive the potential of OGD re-use according to the characteristic of *the implicit pole*. For example, if a manager addresses a score of 2 to the OGD re-use potential *job creation* for a construct composed of *individual* (as the emergent pole) and *group* (as the implicit pole), that means the managers perceive the affordance *job creation* according to *individual* characteristics. On the contrary, if the manager addresses a score of 5, that means the managers perceive *job creation* according to *group* characteristics. Once scored, the entirety of personal grids can be analyzed in order to explore how managers perceive the potentials of OGD re-use and to better understand characteristics that they conceive to actualize OGD re-use potentials.

### **Data Analysis**

We used a combination of different data analyses techniques associated with RepGrid to investigate public managers' views regarding the potentials of OGD re-use. A first analysis was to identify common themes across managers' constructs. For that purpose, we applied a content analysis (CA) procedure, as described by Jankowicz (2005), for the purpose of pooling and categorizing participants' constructs regarding their similarities. We realized the CA in three steps. First, we cleaned constructs and removed those that were

unusable (i.e. illegible handwriting, elicitation guidelines not respected, etc.). In the second stage, we first focused on identical constructs (i.e. constructs with the same emergent and implicit pole) such as *individual-group* or *public-private* and grouped them under the same dimension (e.g. group vs. individual). We then concentrated on constructs that were not identical but that had very close poles (e.g. *individual-group* vs. *individual-collective*) and allocated them to existing dimension. We continued this process until all of the constructs have been categorized, creating sometimes new dimensions. As recommended by Jankowicz (2005), we grouped the unclassifiable constructs under the "miscellaneous" dimension. Finally, to avoid subjectivity in the choice of dimensions, Jankowicz (2005) recommend a *reliability check*. It requires that two researchers perform the CA independently. Accordingly, we realized the CA (i.e. the moderator and the observer) in parallel, without any dialogue. After some deliberations, we compared our results and agreed on the choice of dimensions. Dimensions are combinations of managers' view of the OGD re-use potentials. They highlight public managers' reflections on the actualization of the OGD re-use potentials. They can be understood as the necessary *characteristics* to actualize the potentials of OGD re-use.

**To facilitate the qualitative interpretation of the results, we used the Principal Component Analysis (PCA).** The objective of the PCA is to observe relations between and among elements and constructs. To go further and identify *whether public managers share the same view regarding how the potentials of OGD re-use should be actualized*, we created new grids (i.e. dimension grids), one for each dimension. The *dimension grids* were composed of elements elicited during the focus group as well as constructs that composed the dimension (e.g. for the combined grid *group vs. individual*, we included the construct *individual-group*, *individual-collective*). *These grids regroup managers' reflections on the actualization of the OGD re-use potential and facilitate the visualization of common viewpoints.* According to Jankowicz (2005), only grids that possess at least six elements and six constructs should be analyzed with PCA. We present the results of four grids in the next section. We use for that purpose the package "openregrid" available on RStudio. We used the `bertinCluster` function for extracting different clusters and subsequently conducted the PCA using the `biplot` function.

A particular interest was placed on the result regarding (1) the distribution of the constructs, (2) the distribution of the elements and (3) on the length of constructs. The distribution of the constructs shows the correlation between constructs, i.e. *to what extent managers share the same view regarding characteristics needed to actualize the potential of OGD re-use.* Following Jankowicz (2005), the proximity of the constructs (i.e. the presence groups) could be interpreted differently according to analysts (i.e. psychologist, researcher, or manager). However, in this study, a group of constructs means that public managers scored similarly the potentials of OGD re-use. This implies that they share similar views regarding characteristics needed to actualize the potential of OGD re-use. Distinct groups may have different points of view concerning characteristics to apply, which may imply different ideas regarding the rules for publishing OGD. We also focused on the distribution of elements to explore which OGD re-use potentials present the highest variability (i.e. where public managers have different views regarding how to actualize the potential of OGD re-use). The closer the elements are to the horizontal axis, the less variability there is in the rates assigned to the elements and the more managers share the same view regarding the actualization. Finally, by studying the length of constructs, we explored *to what extent the characteristics to actualize the potential of OGD re-use may vary according to the OGD re-use potential.* In doing so, we were able to provide information on how managers comprehend the OGD concept. We sought to understand to what extent public managers may adapt OGD publication rules according to the OGD re-use potential. The greater the length of the construct, the more variable the scores assigned to the OGD re-use potentials and the more the characteristics to actualize the potentials of OGD re-use depends on the OGD re-use potential. This implies that each of the OGD re-use potentials is envisaged with its own characteristics of actualization and with its own publication rules. On the other hand, the shorter the construct, the less variability and the less characteristics to actualize the potentials of OGD re-use depend on the OGD re-use potential itself (i.e. independently of the OGD re-use potentials, publication rules tend to always be the same).

## Results

### *Structuring and limiting the re-use potential of OGD*

We applied CA on the 18 personal grids. With, on average nine constructs per grid, public managers elicited a total of 120 constructs. Most of these constructs were clear, comprehensive and consistent with our instructions. We only removed 11 constructs from the analysis when their poles did not express contrast (i.e. good vs. bad). A significant part of the constructs was similar, meaning that the emergent implicit poles were identic. This was the case for constructs such as *proactive-reactive*, *individual-group*, *tangible-intangible*, *past-future*, among others. In some cases, the words that managers used for expressing a pole were not exactly identic but belonged to the same family or were semantically similar (e.g. *constraint-free* or *constraint-freedom*). Given the similarities in the proposed constructs, we agreed on four dimensions (i.e. characteristics perceived by public managers to actualize the potentials of OGD re-use) – *group vs. individual behavior*, *affordance vs. constraint*, *utilitarian vs. non-utilitarian effect* and *active vs. passive actualization*. Definitions of dimensions and examples of constructs used are presented in Table 1.

<b>Dimension</b>	<b>Definition</b>	<b>Exemplary constructs</b>
Group vs. individual behavior	This dimension expresses the extent to which the actualization of OGD re-use potentials should be taken by an individual person or a group.	Individual – group, individual – collectivism
Openness vs. constraints	This dimension expresses the extent to which the actualization of OGD re-use potentials depends on a balance between constrains and complete openness.	regulated –liberalized, protection – amelioration, accessible – controlled
Utilitarian vs. non-utilitarian effects	This dimension expresses the extent to which the effects of actualizing the OGD re-use potential has rather a utilitarian or non-utilitarian effect.	Economical – environment, economic – social,
Active vs. passive actualization	This dimension expresses the extent to which the actualization of OGD re-use potential should be embrace by government agencies.	Proactive – reactive, anticipation – correction, innovation - conservation

**Table 1. Definition of dimensions**

### *Different perceptions on OGD re-use potential*

As mentioned above, the results from the PCA allow us to draw various conclusions. Using the previous dimensions, next we discuss how public managers perceive the re-use potential of OGD. In the interest of clarity, we would like to underline that the figures presented below highlight three important shades of perception: to what extent managers share the same view regarding OGD re-use, where their views differ (i.e. in which application domains) and to what extent personal point of view regarding the actualization of the potential of OGD re-use may vary according to potential applications.

#### **Dimension 1: Group vs. Individual behavior**

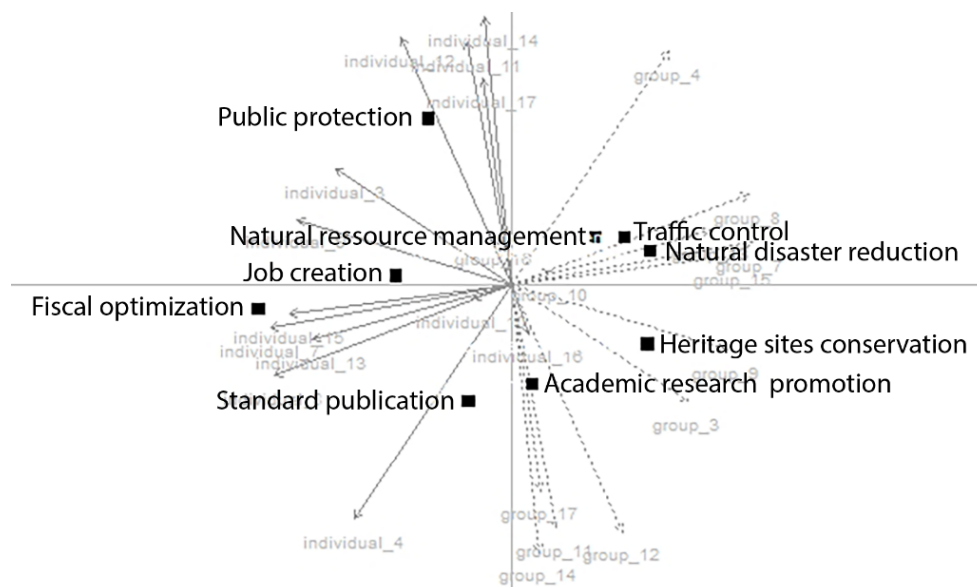
The first dimension is relative to the users and their behaviors to actualize the potential of OGD re-use. It expresses the extent to which the actualization of OGD re-use potentials should be taken by an individual user (with non-profit intention) or a group (with profit intention). Owing to the numerous repetitions of the constructs – individuals vs. groups -, public managers seem to assign importance to the user profile (i.e.

individual vs. group) and its behavior (i.e. commercial vs. non-commercial purposes) regarding the actualization of the potential re-use of OGD.

In Figure 1, the distribution of the constructs provides information to what extent public managers have similar perceptions regarding who should actualize the OGD re-use potentials. We differentiate four groups of public managers. The first cluster is composed of managers 11, 12, 14, 16 and 17 and the second includes managers number 7, 8, 10, 13 and 15. The third and fourth clusters only group few managers (i.e. two managers for cluster three, whereas manager number 4 shapes the last group). Clusters of constructs show that there are few variabilities in the rating among managers of the same cluster, which implies that managers of the same cluster assign same scores to the OGD re-use potentials. Consequently, they tend to share the same view regarding who should actualize the potentials of OGD re-use. It is interesting to mention that although managers may be subject to a different cluster, apart for the manager 16, they tend to all believe that a group with profit intentions should actualize applications such as *traffic control*, *natural disasters management* and *infrastructure resources management*. Managers also seem to agree on the fact that individuals with non-profit intentions should actualize *fiscal optimization* and *job creation*. Surprisingly, only manager number 16 seems to have a distinct view on this.

All managers seem to share a similar perception who (individuals or a group) should be concerned with harnessing the potentials of *job creation*, *fiscal optimization*, *traffic control*, *natural disasters reduction* and *natural resources management*. However, a large proportion of managers does not seem to agree on which users will profit from the actualization of the OGD re-use potentials. The potentials of OGD re-use are closer to the vertical ends, which represents an increase in the variability of the scores. This means that managers share different views regarding which users should actualize *public protection*, *standards and certificates* as well as *academic research promotion* and *heritage site conservation*.

Apparent from Figure 1 is that the length of constructs appears very similar and relatively long, expect for managers 10 and 16. The length of a construct reflects the personal perception of a manager regarding the actualization of the nine potentials OGD re-use. The greater the length of the construct, the more variable the score assigned to the OGD re-use potentials and the more the manager's perception will be influenced by the user who will actualize the potentials of OGD re-use. Consequently, according to the potential OGD re-use, managers will not have the same expectation regarding the actualization, which may influence them on how to publish OGD.



## Figure 1. Group vs. individual behavior

### Dimension 2: Openness vs. Constraints

During the focus group discussion, public managers asked many questions about legal issues, such as data ownership, responsibility, and risks associated with the publication of data. We argue that constructs used by public managers such as *regulated – liberalized* or *accessible – controlled*, reflects those legal concerns. Thus, we define the openness vs. constraints dimension as follows: *it expresses the extent to which the actualization of OGD re-use potentials should be constrained or completely open*. This dimension hence describes the openness level that should be attached to the actualization.

The shape of the constructs distribution of Figure 2 is similar to Figure 1. We identified three clusters of managers. This means managers from each cluster seems to share a similar vision regarding the level of openness that should be attached to actualize the potential of OGD re-use. By being closer from each other, the three managers' clusters tend to share the same vision regarding the level of openness to applying for *standards and certificates* and *job creation*. The first group is composed of managers 10, 11 and 16, group two is composed of managers 12 and 15, and managers 4 and 17 form the third group. Although most of the managers agree on the constraints that should be implemented to actualize *standards and certificates* and *job creation*, the distribution of the remaining affordances shows that managers tend to share different views regarding the level of openness that should be attached to the OGD re-use potentials. The score variability is more important for *infrastructure resource management*, *heritage site conservation*, and *traffic control*. This means that some managers think that the actualization of the OGD re-use potentials should be controlled, while others think the opposite. For instance, managers 1 perceives *fiscal optimization*, *heritage sites conservation* and *infrastructure resources management* as OGD re-use potentials that require less restrictions, while manager 2 thinks they need to be more controlled. Despite this, we notice that public managers tend to favor control over unrestricted re-use of OGD, which is in line with previous research we discussed in the background section.

Similar to our observations of the previous dimension, we also see in Figure 2 that constructs tend to be relatively long. Public managers do not apprehend the actualization and publication of OGD with the same openness level. While one manager may expect a low level of constraint for *job creation*, the same manager prefers a higher degree of openness for *natural disaster management*. Figure 2 shows that the degree of constraint or openness perceived by public managers seems depends on the potential of OGD re-use. For instance, for OGD re-use potentials generating privacy risks, public managers might apply binding licenses, such as non-commercial data use, which may impede the universal participation concept of OGD.

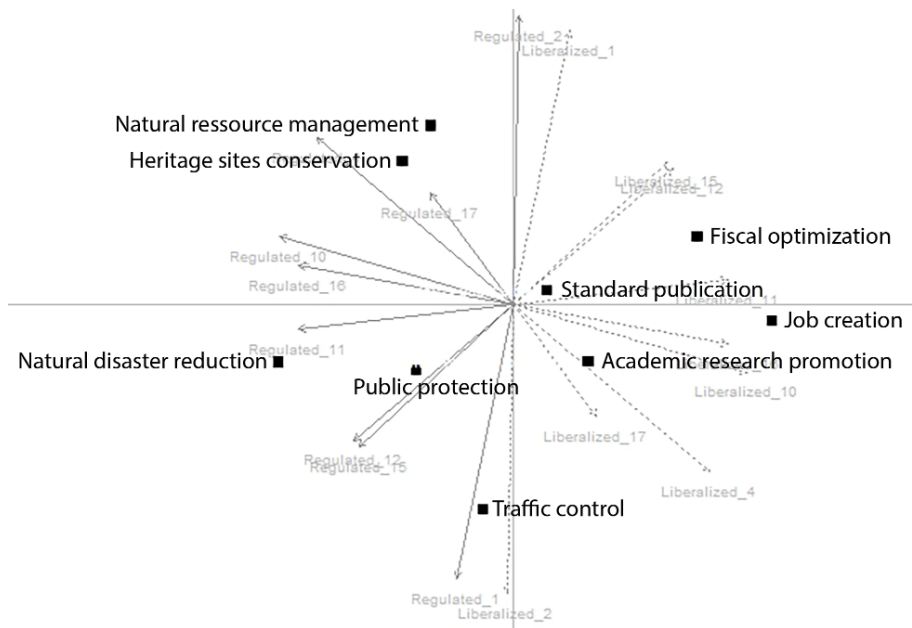


Figure 2. Openness vs. constraints

### Dimension 3 – Utilitarian vs. non-utilitarian effects

During the focus group discussions public managers expressed some concerns regarding the nature of effects when OGD re-use potentials are actualized. In the personal grids these concerns were manifested by developing constructs, such as social-economical, environmental-economical, and cultural-economical. We thus defined a third dimension that expresses the extent to which the effects of actualizing the OGD re-use potential has rather a utilitarian or non-utilitarian effect. We interpret utilitarian effects as economic effects, while non-utilitarian may be understood as social, cultural, or even environmental effects.

Figure 3 shows a lower degree of variability in the scoring of elements. It appears that managers' point of view regarding the effects of actualizing the OGD re-use potentials look more homogenous than previous categories. Contrary to the second dimension, most of the elements are close to the horizontal axis and do not deviate much from it. Only *citizen protection* and *job creation* tend to be perceived differently by managers. For the other OGD re-use potentials, managers tend to agree on the effects of the actualization (i.e. social versus economical), although it is more difficult to identify distinct groups. In Figure 3, we can further see that the distribution of constructs appears to be most concentrated (i.e. 2 groups on a bit more than 90%). This confirms the common view regarding the effects of actualizing the OGD re-use potentials. We observe that *standards and certificates* and *fiscal optimization* should provide economic outcomes, while *natural disaster* and *heritage sites conservation* should be associated with social and cultural objectives.

Contrary to the two previous dimensions, when we look at the length of the constructs we observe major differences. Some constructs are much shorter than others. These shorter constructs show that some managers always perceive the potentials of OGD re-use in the same ways, independently of the actualization effects (i.e. utilitarian vs. non-utilitarian). This is the case for managers 3, 9 and 13. For those managers, whether economic, social, cultural or environmental effects actualization will produce, this will have little impact on how they perceive the OGD re-use potentials. These three public managers will not adapt their publication rules according to the effects of the OGD re-use potentials. However, for the other managers, the OGD re-use potentials depend on the effects of the actualization. We observe it distinctly for the managers 6 and 15. Figure 3 suggests that the nature of effects (utilitarian vs. non-utilitarian), in some



cases, affect the perception on OGD re-use potentials (i.e. longer constructs). Therefore, managers may adapt their OGD publication practices according to the effects of the actualization.

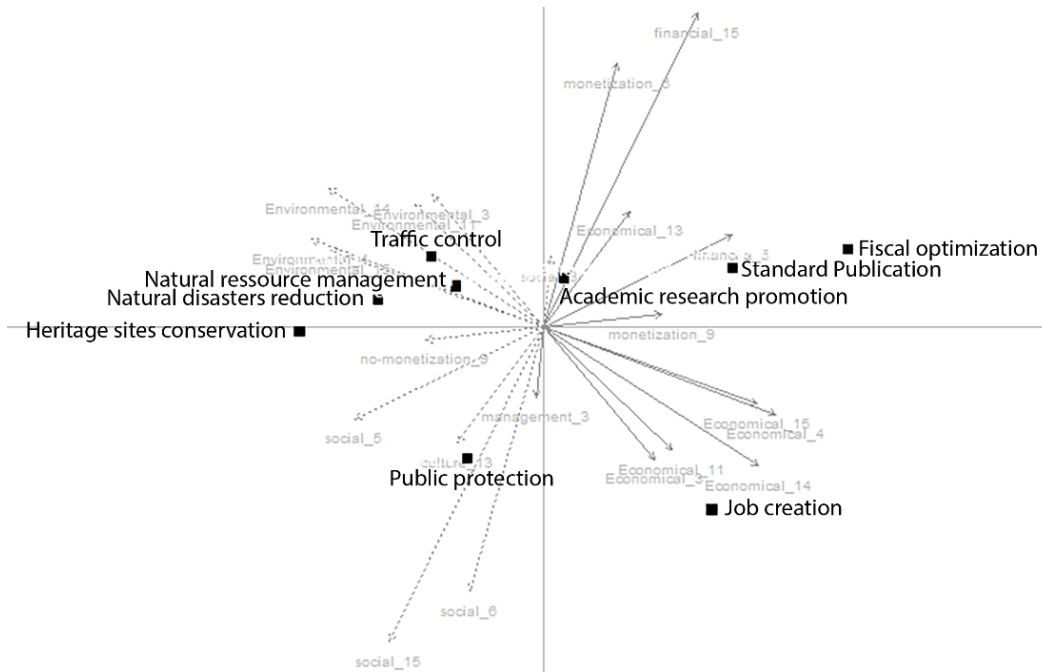


Figure 3. Utilitarian vs. non-utilitarian effects

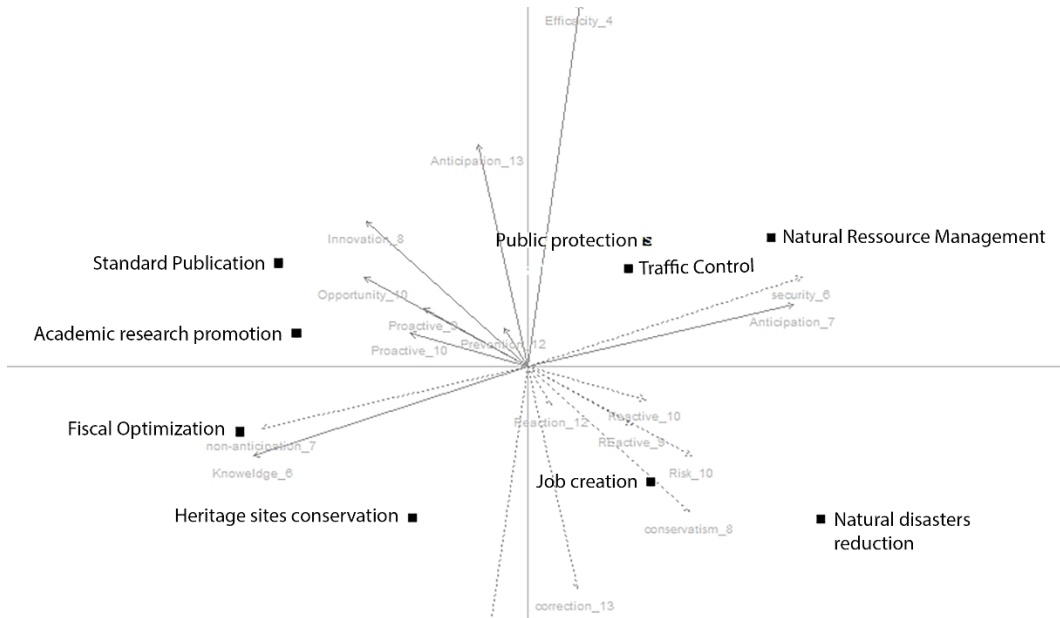
#### Dimension 4 – Active vs. Passive actualization

A large number of participants often expressed the lack of government engagement in contributing to the active development of an OGD ecosystem. With constructs, such as *proactive – reactive* or *anticipation – correction*, managers articulated some concerns related to the degree of engagement that a government should exercise for establishing an OGD initiative. The active vs. passive actualization dimension expresses, therefore, the extent to which the actualization of OGD re-use potential should be embraced by government agencies.

Two observations can be based on Figure 4. First, there is a high degree of variability of the scorings of elements. We observed OGD re-use potentials spread over the whole figure and rather far from the center. This indicates that public managers tend to disagree on the government’s role to push OGD. This means that there is no consensus on the extent to which the actualization of OGD re-use potential should really be a priority for government agencies. This is particularly observable in the case of *natural disaster management* or *infrastructure resource management*. While manager 7 argues in favor of an active implication in using OGD in the context of *natural disasters management*, manager 13 tends to advocate a passive stance by governments. The shape of constructs distribution confirms this observation: it looks like a wheel where dashed constructs lines intermingle with solid line constructs. This form of distribution does not allow the identification of a clear trend in the managers’ point of view. These various opinions on how government should embrace the actualization of the OGD re-use potential may influence the publication process of OGD. For instance, different believes regarding the role of government in OGD re-use may encourage managers to adopt different behaviors, being themselves rather active or passive.

As mentioned previously, the length of a construct gives information on the high variability of the scores that a manager addressed to the potentials of OGD re-use. While the government’s engagement degree on

the actualization of OGD re-use potentials seems to have no impact on the scoring for the manager 12, it largely affected the other managers (particularly managers 4, 6 and 7).



**Figure 4. Active vs. passive actualization**

## Conclusion

With the goal to shed some light on the reasons why OGD is experiencing major barriers or even a standstill in certain countries, our study tried to comprehend how public managers perceive the actualization of OGD re-use potential. We analyzed the content of 18 grids from public managers working in different branches and levels of government. We identified four relevant dimensions based on a CA and categorized common themes which were expressed by public managers. Drawing on affordance theory and the assumption that being able to perceive a potential before one can act and create real-world effects, we further explored if public managers share the same vision with respect the re-use potentials of OGD in different application domains. We attempted to understand to what extent public managers apply and comprehend the concept of OGD in a similar way and what application domains and dimensions lead to variability and/or conflicting opinions. In doing so, we analyzed four dimensions in more detail: (1) who should actualize the re-use potential of OGD (i.e. group vs. individual), (2) the openness level that should be considered to actualize the OGD re-use potential (i.e. openness vs. constraints), (3) the effects that possibly emerge from actualizing OGD (i.e. utilitarian vs. non-utilitarian effects) and, lastly (4) the degree of government engagement in promoting the re-use of OGD (i.e. active vs. passive actualization).

Our analyses show that, expect for the perceived affordance effects (i.e. dimension utilitarian vs. non-utilitarian), public managers seem to share different views regarding how OGD must be actualized to potentially be re-used. There is no clear and shared vision regarding who should actualize the re-use potentials of OGD, neither which openness degree should be applied. We could not find either a common view on the role of the government in the actualization of OGD. Although we extracted the dimensions from the same constructs (i.e. individual-group, regulated-liberalized and active-passive), we did not identify what Hardless et al. (2015) called a “team mental model”. According to them, it is possible to find “organized mental representations, which group members share about key elements in their environment”. A high level of similarity between individuals' perception implies common expectations regarding task

completion, interpersonal cooperation and contribute to organizational effectiveness. This was not the case with our sample and may be an indicator why OGD initiatives are developing slowly or facing resistance.

As other RepGrid studies have shown, sharing different points of view or having a strong personal point of view, may affect task completion and the success of a project. Raman et al. (2013) showed that top managers' interpretation of a situation had a major influence on their organizational strategic choices. Indeed, as demonstrated by Zhao and Fan (2018), if public managers place great emphasis in OGD, they will deploy extra resources and their subordinates will be more predisposed to participate to OGD implementation. The European Commission (2020), confirmed this idea and remarked that fragmentation of actors' perception is a major risk in the current realization of a common data strategy and may, therefore, impact the continued realization and implementation of OGD initiatives. Our results are consistent with these observations and give a more nuanced view on where perceptions on OGD differ. By identifying these distinct viewpoints, our study may help to explain current mismatches between OGD strategies and actual implementation. The fact that public managers do not stick to a "team mental model" and see OGD quite differently reveals that governments need to intensify communication and persuasion work so that OGD becomes a reality for all branches of government and not just for those areas where there is a history of publishing government data (e.g. Census bureau, public transportation). Our results reveal the need for a focus and resources toward organizational processes and not only to technical issues, which is traditionally the case (Kitchin 2014).

Moreover, we also remarked that perceptions often are constructed upon judgments who will actualize the potential, opinions about the role of government, and personal preferences for or against open government. Our results indicate that public managers do not seem to put equal emphasis on free accessibility, reusability and data sharing, to anyone, for any purposes, and without any legal, technological or social constraints. On the contrary, the different lengths of the constructs show that public managers seem to adapt these principles according to the domain of application, which is consistent with findings on affordance niche adaption (Mettler et al. 2017). There is no such thing as "universal accessibility" of OGD for the participants of our study. In practice, this means that discrimination against beneficiaries or user profiles (e.g. individuals, private group, education or citizen), type of utilization (e.g. cultural, environmental, commercial, non-commercial, or social) may happen. We argue that managers may be inclined to use different approaches to publishing OGD, because universal accessibility is an illusionary concept in practice and imposes many organizational, technical, and legal problems. According to Fane et al. (2019), the biggest challenge to publish OGD concerns the uncertainty of data ownership and the misuse of data. By making nuances in the publication of OGD, public managers may calibrate some of the risks. This becomes particularly apparent with respect to the second and fourth dimension. According to El Emam et al. (2015), sharing data creates strong pressure on data publishers and may expose them to the unintentional posting of sensitive information. This is tricky, given that Article 24 of the GDPR makes data publishers responsible for complying with data protection. Countries like India and China but also United Kingdom, already created many smart cities allowing governments gather information on every activity of every citizens (Salmasi and Gillam 2010; Taylor 2017). Improper opening of such data may engender several risks such as breach of confidentiality, mass surveillance or personal intrusion that affect individuals' privacy. According to Taylor (2017), these risks should not be underestimated given that data processing technologies advance faster than data justice. These aspects may strengthen managers' fears and demotivate them to take an active stance in publishing OGD of their department data. It may also explain why sometimes the actualization of some OGD re-use potentials seem to be clear for a large majority of managers, while there is a great disagreement in other application domains. Depending on these risks, public managers seem to adapt their OGD perception. If risks are limited, such as for datasets to be re-used in a job creation application, then openness and universal accessibility of datasets is favored. If they see threats to privacy, as could be the case with the example of public protection, a more nuanced way of liberating data is preferred. In this sense, we argue that the comprehension of OGD as concept is relative, and far away from a universal vision promoted by some OGD advocates. If governments really wanted to make OGD universally accessible, they crucially need to deploy supportive measures (i.e. formations,

budget, etc.) to educate people in charge of OGD publication, reduce risks and thus limit individual choice based on distorted perceptions.

### ***Future research and limitations***

To date, most studies concerned with open government have focused on OGD policy initiatives, technologies, and barriers of using or publishing OGD (Crusoe and Melin 2018; Dulong de Rosnay and Janssen 2014; Susha et al. 2015b). Although having a better understanding of what public managers think about OGD has been found to be essential for properly controlling and steering OGD initiatives, not much evidence exists about the way how they form decisions, what deliberations and preferences guide this process, and how they (if at all) act up on it. By following the affordance theory and using the RepGrid technique, our study tried to provide some further insights on the constructs and dimensions preceding the actualization process of OGD re-use potentials. The mechanisms arising from perception are usually not detectable at first sight. To be identified, studies on human cognition are needed. The RepGrid technique combined with the affordance theory help to capture the underlying structures of perception and shed some light on reasons that may explain different stances in favor or against publishing OGD in a certain application domain.

However, although our paper provides insights for researchers and practitioners, it is subject to certain limitations. RepGrid, as a cognitive mapping technique, does not allow predicting public managers' resistance against OGD publication. It does also not provide evidence regarding cause-and-effect relationships between the publishing of OGD and certain profiles of public managers (e.g. variation in behavior between public managers on different levels of government or in different departments). RepGrid is an analytic and descriptive tool that may be used as a starting point for developing explanatory contributions.

To go further and extend this study, we suggest to use our results in order to develop a national or international survey of public organizations that are engaged in the process of data openness in order to see if aspects such as background, personal values and organizational culture may explain these different perceptions of openness. Also, our observation that public managers do not share a common vision on the implementation of OGD requires further corroboration using, preferably, some explanatory techniques. Lastly, in view of the presented results, we see a need for more research clarifying the terminology in an "user-friendly" manner and which develops more nuanced frameworks that can be used by governments to adapt their open data strategies or specific OGD initiatives. We also call for more practice-based research that helps to develop such frameworks and which supports public managers in systematically assessing risks - organizational, technical, and legal ones – so that government data is shared and made publicly available from all branches of government (Berghmans et al. 2017).

## *Paper # 3*

# Paper #3 The Impact of Data Governance on OGD Publication

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

**Author:** Auriane Marmier

**Published in:** DG. O 2022: The 23rd Annual International Conference on Digital Government Research

**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.

# Paper 3

- Paper #3 The Impact of Data Governance on OGD Publication ..... 75
- Introduction..... 77
- Methodological approach..... 78
  - Data collection ..... 78
  - Data Analysis..... 79
- Contexte ..... 80
- Findings and analysis..... 81
  - Structural Components ..... 81
  - Procedural Components..... 82
  - Relational Components..... 83
- Conclusion ..... 85

## 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 re-used 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 (Tonya Riley 2021). From an academic perspective, many authors such as (Máchová et al. 2018) and (Matos and Corbett 2019) 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, they also appear as the central actors in charge of these strategic assets (Vancauwenberghe and van Loenen 2019). 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 (Le Parlement Suisse 2021). 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. (Yiannakoulis et al. 2020) 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 (Nielsen 2017). 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 (Seydtaghia and Farine 2021). 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 (Nielsen 2017; Vancauwenberghe and van Loenen 2019) and leading public sectors organizations face an increasing number of OGD challenges (Trom and Cronje 2019).

Although there is a growing body of literature on OGD publication regarding sociological, technical, and legal challenges (Afful-Dadzie and Afful-Dadzie 2017; Barry and Bannister 2014; Benitez-Paez et al. 2018; Janssen et al. 2012; Lv and Ma 2019; Nielsen 2017; Zuiderwijk and Janssen 2014a), few papers consider data governance as the root of the problem. Based on an in-depth literature search, (Abraham et al. 2019) 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 (Trom and Cronje 2019) and appears relevant for facilitating open data use (Welle Donker and van Loenen 2017). (El-Emam et al. 2013) highlighted that only effective data governance may allow value creation, and Soares (2010) stated that to leverage data as valuable assets, people, processes, and technologies must be effectively managed through data governance programs (Trom and Cronje 2019). (Ndamase 2014) describes a data governance program as a framework providing data security, allowing data policies and standards development, 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 (Nielsen 2017), so far, little is known regarding the impact of data governance on OGD publication (Lämmerhirt 2017). (Benfeldt et al. 2020) 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 (Vancauwenberghe and van Loenen 2019). 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.

## METHODOLOGICAL APPROACH

### 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* (Confédération Suisse 2020; Etat de Vaud 2018; Ville de Lausanne 2020), *reports of the European Commission (OECD 2019a; OECD 2019b)*, 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.

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

**Table 1: Data gathering**

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 (Kitchin and Thrift 2009) 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 (Barbour 2010; Fusch et al. 2017; Hancock and Algozzine 2006). Furthermore, through multiple sources of evidence, ethnography aims to describe both the group members' point of view and the perceptions and interpretations of researchers (Holloway and Daymon 2010). 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 (Fusch et al. 2017; Noor 2008). 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. (Hancock and Algozzine 2006) 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).

### **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 (Robson 2002). 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 (Braun and Clarke 2006; Figgou and Pavlopoulos 2015; Vaismoradi et al. 2013). 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 (Mayring 2004), this analysis is especially useful for understanding more profound a phenomenon (Eisenhardt and Graebner 2007). To this end, we thus followed a well-defined step-by-step process proposed by (Braun and Clarke 2006). 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 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 (Abraham et al. 2019), 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 literature and during our ethnography were convergent, we applied the triangulation technique (Flick 2004). 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 (Erzberger et al. 2003), it is a metaphor that describes a technique allowing the investigation and comprehension of a situation from several perspectives (Heath 2015; Nguyen 2009; Nightingale 2020; Williamson 2013). Based on the cross-checking of different data sources such as quantitative survey, participant observation, historical analysis, comparison and discourse analysis (Nguyen 2009), the triangulation technique is used to increase the validity and trustworthiness of the results (Williamson 2013). By comparing multiple data

sources, triangulation pinpoints convergence or divergence in data collected (Barnum 2021; Heath 2015). 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 (Nguyen 2009; Strauss and Corbin 2014; Yin 1994). Thus, as many authors in organizational research (Heath 2015), geographical studies (Nightingale 2020) or computer science (Barnum 2021), 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.

Definition	References
<i>Structural components</i> 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]
<i>Procedural components</i> 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]
<i>Relational components</i> relate to business alignment, employee competencies and organizational culture that support decision-making.	[16] [17] [19] [59] [60] [61] [63] [64]

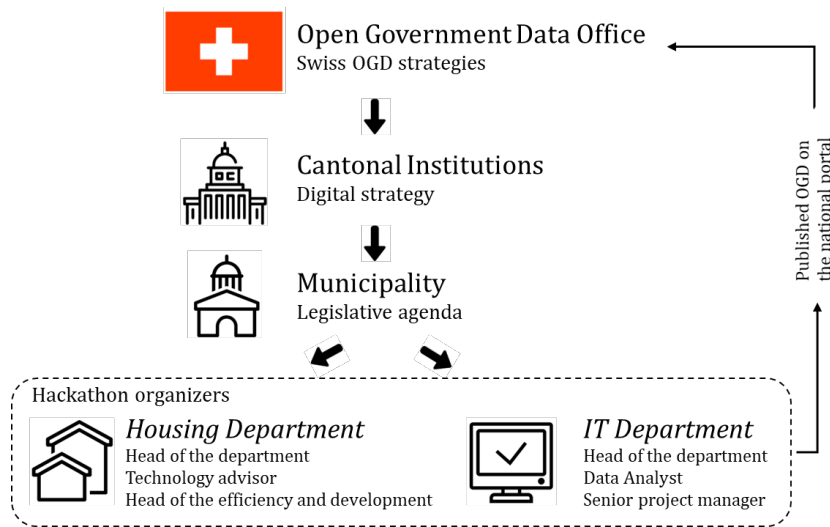
**Table 2: Definitions and references of data governance components**

## CONTEXTE

In Switzerland, discussions to make government data publicly available appeared for the first time in the E-Government Strategy 2008-2015 (Swiss Confederation 2018). After that, the Federal Council developed a new strategy for “*an information society in Switzerland intending to optimize information resources*” (Federal Office of Communications 2018). 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” (Federal Office of Communications 2018). 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 (Federal Council 2014). In line with prior researches (e.g. Kalampokis et al. 2011a; Robinson et al. 2009), 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 to a wide range of individuals. As a result, in 2016, the Swiss Federal Council ordered the development of an OGD platform (Confédération Suisse 2020). 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 (Federal Statistical Office 2021). 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 (Federal Statistical Office 2021).

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.



**Figure 1: Context and actors of the ethnography**

## FINDINGS AND ANALYSIS

The triangulation technique has rapidly pinpointed that the data governance practices promoted in 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.

### Structural Components

Results of the literature review show that practices such as the role and responsibilities of stakeholders as well as decision-making authorities (Alhassan et al. 2018) received particular attention. Clustered under the term *Structural components* and frequently called *organizational data structure* (Marchildon et al. 2018; Soares 2010), *standardization* (Kim and Cho 2018) or *formalization* (Stanford University's Data Governance Office 2013), several authors stressed that the distribution of roles and responsibilities is a crucial factor of data governance (Alhassan et al. 2019; Cheong and Chang 2007; Khatri and Brown 2010; Panian 2010; Spy Pond Partners and Iteris 2015; Weber 2009). (Alhassan et al. 2019) 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 (Abraham et al. 2019; Dreibelbis 2008; Otto 2011b), 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 (Tallon 2013).

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. (Spy Pond Partners and Iteris 2015) explained that these components are extensively reported in data governance frameworks to foster quality, value, and suitable data re-use. 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) (Alhassan et al. 2019; Stanford University's Data Governance Office 2013; Yang et al. 2019). 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*

### **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 (Barry and Bannister 2014; Benitez-Paez et al. 2018; Janssen et al. 2012; Marmier and Mettler 2020a; Martin et al. 2013), 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 re-use, 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 (Alhassan et al. 2018; Panian 2010). 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 (Abraham et al. 2019; Alhassan et al. 2018; Panian 2010). Several authors notably mentioned data processes and procedures and interpreted them as "*guidelines and rules necessary for dealing with data*" (Alhassan et al. 2019). For some authors, they reflect the desired organizational behavior in terms of data quality (Khatri and Brown 2010; Marchildon et al. 2018; Soares 2010; Stanford University's Data Governance Office 2013), data access (Brackett and Earley 2009; Panian 2010), data collection and storage (Khatcherian and Jefferson 2009), metadata management (CMMInstitute 2021; Marchildon et al. 2018; Stanford University's Data Governance Office 2013), data lifecycle (Khatri and Brown 2010; Marchildon et al. 2018; Soares 2010), and data platform and architecture (Marchildon et al. 2018; Soares 2010; Thomas et al. 2019). As they concern all steps of the data lifecycle, from the recording to the sharing (Borgman 2015; Khatri and Brown 2010; Panian 2010), some authors stated that adopting procedural components ensures data management as a strategic asset (Spy Pond Partners and Iteris 2015). For instance, a metadata strategy provides content that makes data understandable and reusable (Khatri and Brown 2010), while quality standards facilitate interoperability (Marmier and Mettler 2020a). 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*

### **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 (Vitale et al. 1986), who found that the alignment of departments frequently fails because there are no overall objectives. (Marmier and Mettler 2020b) 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 (OECD 2019b), 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].

(Rusu and Jonathan 2017) 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 (Rusu and Jonathan 2017) as a factor that could substantially affect the success of OGD.

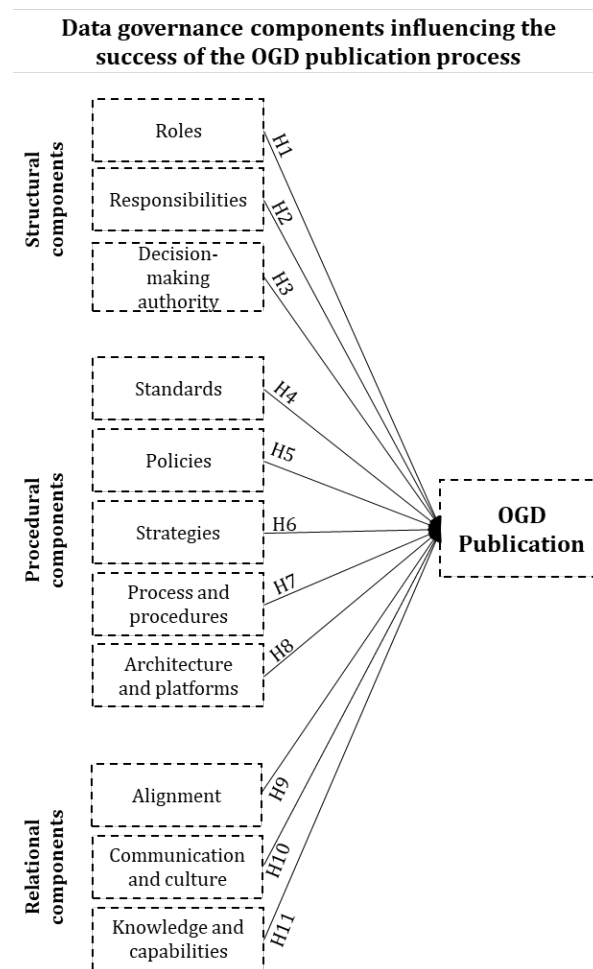
While some authors have started to stress the importance of the impact of employees' data competencies (Alhassan et al. 2018), and knowledge (Alhassan et al. 2019) as well as training (Abraham et al. 2019) on data governance, our findings suggest gaps in data knowledge skills and competencies. After 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* (Welle Donker and van Loenen 2017), *organization* (Panian 2010), *capacity for coherent implementation* (OECD 2019b) or *facilitating connections between data producers and users* (Spy Pond Partners and Iteris 2015) as an enabler to treat data as strategic assets. (Abraham et al. 2019) showed that training employees involved in data processes help them to act accordingly to data policies, processes, and procedures (Alhassan et al. 2019; Randhawa 2019). (Alhassan et al. 2019) 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*

## 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 (Flick 2004), 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 procedural components (H4, H5, H6, H7, H8) and three relational components (H9, H10, H11) seem to influence the success of the OGD publication process.



**Figure 2: The model of OGD governance.**

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 (Nielsen 2017) 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 (Recker 2012).

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 (Roussel 2021). 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 (Gauld 2007). Indeed, the case study approach may be a starting point for developing exploratory and explanatory contributions (Recker 2012). Consequently, we plan to continue expanding this paper by testing our hypotheses 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” (de Bary Heinrich Anton 1878). 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).

## *Paper # 4*

# Paper #4 Towards good data governance in the public sector

**Title:** Towards good data governance in the public sector: Myth or reality? An OGD Governance Maturity Model

**Author:** Auriane Marmier

**Under review**

**Abstract:** As a substantial actor in data creation, public sector organizations identified data as an ambitious and promising prospect to improve transparency, the economy, and government service delivery. Thus, several governments instrumented Open Government Data (OGD) strategies. However, introducing such strategies also represents considerable data governance challenges. While data governance receives increasing attention in the private sector, only some authors consider data governance as a precondition to OGD. Therefore, this paper explores to what extent OGD governance practices are implemented in Swiss public organizations and develops an OGD governance maturity model (MM). Based on a survey method, we collected 252 respondents' answers. We use the Rasch model analysis and show that technical governance practices are the most applied in Swiss organizations, independently of the political level of the organization. In contrast, monitoring and structural practices encounter further difficulties. We also reveal two-speed implementation issues between data-oriented organizations and the fewer ones.

Paper 4

Paper #4 Towards good data governance in the public sector ..... 88

Introduction..... 90

Background..... 91

    Data governance ..... 91

    Data governance maturity models ..... 92

Methodology..... 92

    From the questionnaire design to the data collection ..... 92

    From the Rasch Analysis to the Maturity Model..... 93

Findings ..... 94

    Sample Characteristics ..... 94

    A general description of OGD governance practices ..... 94

Discussion..... 97

Conclusion ..... 99

Appendix A..... 101

## Introduction

In 2019, 2.5 quintillion bytes of data were produced every day. According to the expert, this number is expected to increase by ten times by 2025 (Dhillon and Singh 2019). This continuing creation of data put into perspective their possibilities and made worldwide organizations aware of the data's strategic interest (Fadler et al. 2021). Due to this explosion of possibilities, companies have rethought (or developed) their activities to integrate data valorization. Presently, seven of the ten largest market capitalizations are data-driven companies. Consequently, as a substantial actor in data creation, public sector organizations also identified data as an ambitious and promising prospect (Jetzek et al. 2019). Thus, several governments instrumented Open Government Data (OGD) strategies. OGD strategies aim to facilitate access to those hidden data and thus increase transparency and accountability (Bertot et al. 2010; Harrison and Sayogo 2014; Parsons et al. 2011; Talukder et al. 2019; Zuiderwijk et al. 2015), empower citizens' participation (Concilio et al. 2017; Talukder et al. 2019), foster innovation (Janssen et al. 2012; Talukder et al. 2019; van Veenstra and van den Broek 2013; Zuiderwijk et al. 2015), promote economic development (Talukder et al. 2019) and increase participation and collaboration with different stakeholders in government activities (Abu-Shanab 2015; Conradie and Choenni 2014; Harrison and Sayogo 2014; McDermott 2010; Talukder et al. 2019; Zuiderwijk et al. 2015). Yet, while Governments produce large amounts of data (Ubaldi 2013) and promote their re-use through OGD platforms, the success of OGD initiatives is lagging (Attard et al. 2015; Dawes et al. 2016; Janssen et al. 2012; Krasikov et al. 2020; Safarov et al. 2017).

Actual research on OGD barriers' *"only covers parts of the open data ecosystem"*, forgetting *"open data supply, open data governance, and open data user characteristics"* (Charalabidis et al. 2014). Yet, private-sector research recognizes data governance as the main challenge in value creation (Trom and Cronje 2019). Authors pinpointed that in addition to addressing data-driven issues (Nielsen 2017; Thompson et al. 2015), data governance practices facilitate open data use (Welle Donker and van Loenen 2017). However, there is little information on data governance practices in public sector organizations (Brous et al. 2016) and even less on OGD providers. Several authors tag data governance in the public sector as an under-researched area and call for more research on the subject (Al-Ruithe and Benkhelifa 2020; Tiwana et al. 2013). Nielsen (2017) confirmed these overviews by revealing that only 3% of the paper published on data governance concerned the E-government discipline. Thus, it is not surprising that OGD barriers are not considered from the data governance perspective. Thus, following further authors, we believe that addressing OGD objectives necessitates managing and governance the vast quantity of public data (Abraham et al. 2019; Dhillon and Singh 2019). Therefore, following the perspective of the problematization logic proposed by Alvesson and Sandberg (2011), the purpose of this paper is to open the way for new reflections on reasons that limit the OGD strategies, particularly the Swiss initiative, to reach its goals. Thus, to fill the gap of data governance research in the public sector as well as investigate to what extent data governance practices may impact the Swiss OGD strategy success, this paper seeks to explore the maturity level of data governance practices in Swiss public organizations providing OGD.

For this purpose, we first seek to evaluate which OGD governance practices are set up in Swiss organizations through a survey distributed to OGD publisher organizations. Based on this overview, we will be able to identify which potential data governance mechanisms are successfully implemented and which could benefit from further development. Accordingly, the second objective of this study is to develop an empirically grounded and methodologically sound Maturity Model (MM) for OGD governance. As MM allows organizations to assess their as-is situation and determine the desired to-be situation on pre-defined items (Marx et al. 2012), we intend this MM to assess their capability and ability to implement data governance in the public sector. Thus, the proposed MM considers the five following levels - Initial, Managed, Defined, Measured and Optimized (Marchildon et al. 2018; Permana and Suroso 2018; Rivera et al. 2017; Spruit and Pietzka 2015). The five levels of the MM emphasize OGD governance practices a public organization needs to achieve an OGD strategy. Each level of the OGD governance MM is composed of a set of key data governance practices that, once achieved, advance the organizations' capabilities to govern OGD. This study results found that technical governance practices are the most applied by Swiss

organizations, independently of the political level of the organization (i.e. communal, cantonal, or federal), while monitoring and structural practices encounter further difficulties. Moreover, the results reveal two-speed implementation issues between data-oriented organizations (e.g. archives) and fewer ones (e.g. teaching hospitals). This paper is constructed as follows. We first present the background of the study, including the review of data governance and maturity models' previous research. The following section describes the methodology used. We then describe our survey findings in section four, while section five presents the MM and discusses it. We finally close this paper with the limitation as well as implications for practitioners and researchers.

## **Background**

### ***Data governance***

Although more attention has been paid to data governance in the literature in recent years, understanding data governance and what this process may entail is still a challenging task (Abraham et al. 2019; Otto 2011a). While academic and practitioner literature compiled seven literature reviews related to data governance (Alhassan et al. 2016; Alhassan et al. 2018; Nielsen 2017), we do not notice a universally accepted definition of data governance. However, since most data governance papers have been published in the Information Systems and computer science discipline, most scholars define data governance based on Weill and Ross (2004) IT governance definition (Nielsen 2017). Based on the idea of a framework encouraging data utilization, the authors define it as a framework for decision rights and responsibilities to promote desirable behavior in using data (Brous et al. 2016; Otto 2011b; Weill and Ross 2004). Following several definitions, Pierce et al. (2008) define data governance as "*the collective set of decision-making processes for the use and value-maximization of an organization's data assets*", putting forward the notion that data is an enterprise asset. For Otto (2011b), data governance is "*a companywide framework for assigning decision-related rights and duties in order to be able to handle data as a company asset adequately*". Ndamase (2014) also describes data governance as a framework providing data security, allowing data policies and standards development to assist decision-makers. To the data management association (DAMA), data governance is described as exercising authority and control over data management (DAMA 2009). This association describes data governance as exercising authority and control over data management to valorize organizational data assets (DAMA 2009) whilst keeping the focus on data-related risks management (Abraham et al. 2019; Morabito 2015).

Despite definitions and epistemological flaws of data governance, it remains one of the main challenges in (big) data value creation. El-Emam et al. (2013) underlined that solely data governance should foster value creation. Ladley (2019) presents data governance as a new business capability that may help organizations use data as a competitive asset. For some authors, data governance provides organizations with an expansive corporate agenda to derive value from data (Abraham et al. 2019; Ladley 2019; Mosley et al. 2010). In implementing data governance, organizations contribute significantly to maximizing the value of data assets (Goel et al. 2021). Since data governance provides, among other efficiency, limits expenses, reduces risk and increases management performance (Brous et al. 2020), many researchers stipulate that implementation and adoption of data governance are crucial (Brous et al. 2020). By concentrating on data quality (Otto 2011b; Weber 2009), metadata, data principle, awareness and safety and security, data governance practices provide insights and information for leading data-driven organizations (IBM 2007; Khatri and Brown 2010; Marchildon et al. 2018; Trom and Cronje 2019), and increasing data processing speed. Thus, data governance simplifies large datasets exploitation (Welle Donker and van Loenen 2017). While some research exists in the public sector, they mainly concentrate on defining the data governance meaning (Brous et al. 2016) or mapping the data governance practices (Alhassan et al. 2016). Few focus on the level of technical and structural practices, such as organizational role identification, responsibilities assignment, and data guidelines development complying with organizational strategies (Mao et al. 2021).

## ***Data governance maturity models***

In the field of data governance, we notice that to assess or evaluate progress, numerous authors used MM. Since its first mention in the 1970s (Raber et al. 2013), the MM development by academia and practitioners has evolved continuously. Developed to define an organization's capabilities and critical success factors, MM is often used to improve future organizations' development as a road map and milestones (Iversen et al. 1999). Characterized by different maturity levels, capability areas, and assessment criteria (Steenbergen et al. 2010), MM helps organizations to underline critical success factors at each stage of their development (Thomas et al. 2019). Accordingly, between 2007 and today, around twenty papers treat data governance MM (CMMInstitute 2021; Marchildon et al. 2018). While some replace the term MM with *frameworks* (Khatri and Brown 2010; OECD 2019a; Soares 2010), *models* (Alhassan et al. 2018; IBM 2007) as well as *programs* (Ladley 2019) or *maturity assessment* (Marchildon et al. 2018), all follow the same idea – *assessing the progression of data-driven initiatives in organizations*. Moreover, most of them have been designed to describe data governance practices (Alhassan et al. 2016; Brous et al. 2016), investigate influencing factors (Weber 2009) and inspect data governance applications (Begg and Cairra 2012).

In the context of OGD research, a recent paper on the metrics used to benchmark open data progress (Zuiderwijk et al. 2021) has identified 15 models to measure and rank governments' progress in OGD initiatives. Three are classified as assessment models (Hjalmarsson et al. 2015; Osorio-Sanabria et al. 2020; Srimuang et al. 2017), three as evaluation models (Charalampos et al. 2013; Dahbi et al. 2018; Reggi 2011), four as frameworks (Máchová et al. 2018; Máchová and Lnénicka 2017; Sayogo et al. 2014; Welle Donker and van Loenen 2017), one as stage model (Kalampokis et al. 2011b), and only three are officially registered as open data (OD) or OGD MM (Charalabidis et al. 2018b; Ham et al. 2015; Solar et al. 2012). While most of them concentrate on OGD platform assessment (Charalabidis et al. 2018a; Dahbi et al. 2018; Ham et al. 2015; Máchová et al. 2018; Máchová and Lnénicka 2017; Reggi 2011; Sayogo et al. 2014; Welle Donker and van Loenen 2017) and the state of e-Government openness (Srimuang et al. 2017; Veljković et al. 2014) few focus on organizational capabilities (Kalampokis et al. 2011b) and data governance (Osorio-Sanabria et al. 2020). Even though OGD MM exists for platforms or to monitor data opening progress, there is no model to evaluate the data governance capacity of public organizations. Thus, the following part presents the methodology used to construct an OGD governance MM and address the shortcomings of existing models.

## **Methodology**

Literature put forward that two major approaches exist to developing a MM. Becker et al. (2009) adopted a top-down approach which consists of first fixing a certain number of maturity levels and then validating them with items supporting the hypothesis regarding maturity evolution. The second approach that Lahrman et al. (2011) proposed follows a bottom-up perspective. First, distinct dimensions or assessment items are identified. Then, to have a better overview of the maturity evolution steps, the dimensions are clustered according to their maturity levels. Nevertheless, according to the literature, most MMs originate from arbitrary design methods or lack a solid theoretical foundation (Biberoglu and Haddad 2002; Lahrman et al. 2011), limiting their management research reputation (Marx et al. 2012). Thus, to prevent these methodological issues, this paper follows a bottom-up approach in which we first design a questionnaire by investigating literature on data governance practices and then use the Rasch model to examine data collected through the questionnaire and thus derive maturity levels.

### ***From the questionnaire design to the data collection***

In the literature, data governance is generally conceptualized with three data governance dimensions, namely structural, procedural and relational (Ababneh and Aga 2019; Mikalef and Krogstie 2018; Tallon 2013; Tallon et al. 2013), which regroup data governance practices such as role assignment or the establishment of policies and guidelines. Furthermore, the OGD literature also add technical and monitoring

dimensions (Kalampokis et al. 2011b; Osorio-Sanabria et al. 2020). Thus, we constructed the questionnaire according to these five data governance dimensions. To facilitate the respondents' task, we split the questionnaire into three distinct parts, i.e. (1) *general information on the respondents*, (2) *the current state of data and OGD governance strategy* and (3) *data lifecycle activities*. While the first part comprises ten questions, the two others consist of 30 questions, requesting the implementation level of the key data governance practices. The response options are attached to a six-point based Likert scale where 1 may vary from "strongly disagree", "terrible", or "never" to 5 = "strongly agree", "excellent", or "always". The sixth answer corresponds to "I do not know". After the first draft of the questionnaire, we conducted a preliminary investigation by submitting it to experts and civil servants. We invited them to fill out and review the questionnaire. Based on their comments, we have been able to develop the final version of the questionnaire.

We sent the invitations to complete the questionnaire to government agencies participating in OGD initiatives in Switzerland and, most notably, to employees who had already published data on the Swiss OGD platform. To identify these future respondents, we made a two-step collection. First, we extracted the metadata (i.e. emails and names of contact publishing datasets) from the Swiss OGD platform *opendata.swiss*. Like many others (Kirstein et al., 2019), the Swiss OGD platform uses the Comprehensive Knowledge Archive Network (CKAN) to store and distribute OGD. Therefore, using the application programming interface (API) offered by CKAN, we created specific requests to get the email of each person who published datasets. Then, we finally downloaded the result via the power pivot functionality on Excel on *March 1, 2022*. We distribute the survey through the online tool Qualtrics, in French, English and German. We first sent the questionnaire on *March 15, 2022*, and a reminder a month later, in *April 2022*. We only considered respondents with a good overview of their organizations' OGD initiatives to maintain an appropriate expertise level on OGD governance practices.

### ***From the Rasch Analysis to the Maturity Model***

We used the Rasch model to analyze respondents' answers and derive the MM. The Rasch model is a psychometric technique (i.e. "*the study of measuring psychological constructs and processes*" (Boone et al. 2013)) developed by *Georg Rasch* in 1960 to improve the validity and reliability of survey instruments. Initially developed for measuring education scoring, the Rasch model is a statistical technique that allows the measurement of items such as human constructs (e.g. ability, learning, rehabilitation following an injury, and attitude toward products (Boone et al. 2013; Marx et al. 2012)). Based on the analysis of pooled items, the Rasch model evaluates questionnaires by describing respondents' performance (i.e. score obtained from the questionnaire) and the items' difficulty (i.e. score of the questions). In addition to deriving maturity levels by using respondent answers (Berghaus and Back 2016), this quantitative approach allows the construction of robust instruments (e.g. tests, attitude instruments, maturity models) (Boone et al. 2013) based on current respondents' performance. Thus, the Rasch model presupposes that questionnaires' answers depend on the difficulty of the questions and the respondents' capability to answer (Bond et al. 2015).

To apply the Rasch measurement mathematic model, we used the *Jamovi* software. Following standard practices when applying the Rasch model, we generate three tools – (a) *a Wright map (Figure 1)*, (b) *the scores of organizations' abilities (Figure 2)* and (c) *an output items statistics table (Table 2)*. The Wright map visually represents practices' difficulty and organizations' ability scores. Both scores are represented on the same *logit* scale (i.e. maximum-likelihood estimates). In addition to providing a visual score of organizations' ability and practices' difficulty, the *Jamovi* software also makes available output statistics related to practices' difficulty measures. Accordingly, to obtain accurate measurements, we also generate a statistics table. Output statistics provide the measure of practices' difficulty (i.e. score of the questions in logit), standards errors, and a set of standardized fit statistics, including infit and outfit values of each item. We then arranged the score of practices' difficulty in descending order. The output statistics merged with the Wright map highlight the different levels of assessing OGD governance in public organizations and thus derive the necessary levels from managing, measuring, and controlling all aspects of OGD governance



in public organizations. We finally used the Data governance MM semantics found in the literature to label the maturity levels of our model (Table 3).

## Findings

### *Sample Characteristics*

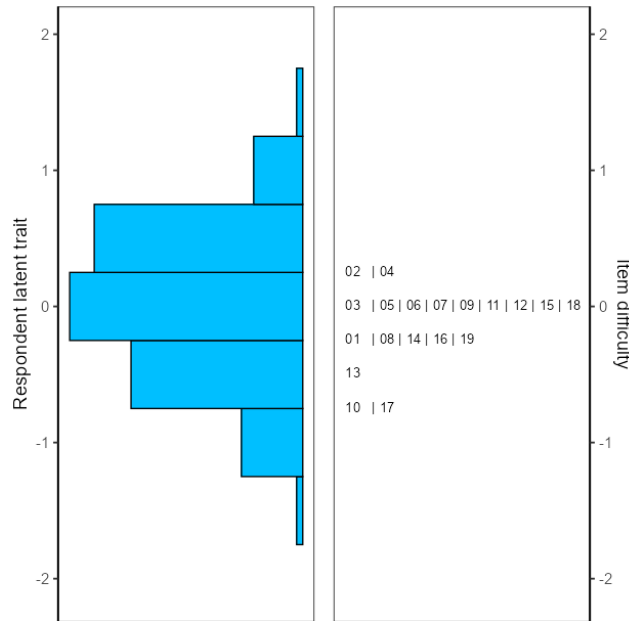
We built a sample of 252 Swiss respondents through online data collection. Given the type and length of the survey (i.e. more than 10 minutes), it appears as an acceptable response rate. Among the respondents, the majority belong to the public administration, defence, and compulsory social security sectors, while 8% come from professional, scientific, and technical activities sectors (Table 1). The other respondents originate from diverse sectors such as agriculture, education, human health, or water supply. While over half of the respondents work for the German part of Switzerland, almost 75% are men. In addition, the results show that the questionnaire mostly reached cantonal (43%) and federal (35%) workers rather than communal ones (12%).

<b>Industry Sector</b>	<b>Nb</b>	<b>%</b>	<b>Gender</b>	<b>Nb</b>	<b>%</b>
Public administration, defence and compulsory social security	146	58%	Male	181	72%
Professional, scientific and technical activities	24	10%	Female	63	25%
n/a	27	11%	Prefer not to say	8	3%
Administrative and support service activities	11	4%	<b>Total</b>	<b>252</b>	<b>100%</b>
Agriculture, forestry and fishing	9	4%	<b>Political Level</b>	<b>Nb</b>	<b>%</b>
Information and communication	7	3%	Cantonal	108	43%
Education	6	2%	Federal	87	35%
Human health and social work activities	6	2%	Communal	31	12%
Water supply, sewerage, waste management and remediation	5	2%	Other	26	10%
Transportation and storage	4	2%	<b>Total</b>	<b>252</b>	<b>100%</b>
Arts, entertainment and recreation	3	1%	<b>Language</b>	<b>Nb</b>	<b>%</b>
Electricity, gas, steam and air-conditioning supply	2	1%	German	140	56%
Other service activities	2	1%	English	12	5%
<b>Total</b>	<b>252</b>	<b>100%</b>	French	100	40%
			<b>Total</b>	<b>252</b>	<b>100%</b>

**Table 1. Sample Characteristics**

### *A general description of OGD governance practices*

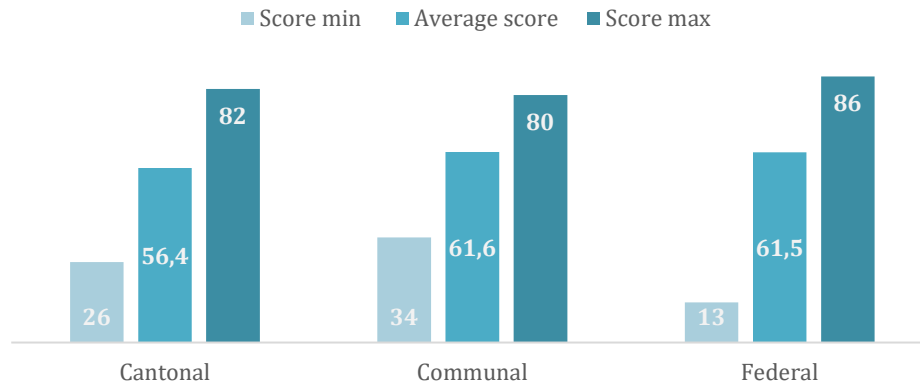
The Wright map (Figure 1) shows the difficulty of data governance practices and organizations' ability scores. While the left side of the map ranked organizations' ability to implement OGD governance practices, the right side ranked governance practices through their difficulty level. Concretely, the organizations in the left upper part of Figure 1 appear most suitable and more advanced in applying OGD governance practices, while in the left lower part, the organizations seem to encounter more difficulties.



**Figure 2. The Wright Map**

This also indicates that governance practices in the top right area (e.g. 2, 4) appear more challenging to implement for public organizations than governance practices 10 and 17 located in the downright part of the Wright map. Furthermore, Figure 1 shows that organizations with a logit superior to zero (i.e. located on the upper left side of Figure 1) have a greater than 50% probability of succeeding on practices 1, 8, 14, as well as items 13, 10 or 17 and less than a 50% probability on practices more challenging such as 3, 5, and 2 or 4. Conversely, organizations with a logit inferior to zero (i.e. located on the lower left side of Figure 1) have a lesser than 50% probability of succeeding on practices above zero logits. This suggests that competent organizations that already have implemented easy practices are more likely to implement complex practices (i.e. answer complex questions) (Bond et al. 2015), contrary to less skilled ones (Dekleva and Drehmer 1997). With a person reliability measure greater than 0.7, the Wright map provides highly reliable indicators of organizations' ability measurement.

Figure 2 summarizes the minimum, average and maximum scores of communal, cantonal, and federal organizations' abilities. We remark that the average scores for federal, cantonal, and communal organizations are broadly similar (Figure 2). However, results indicate significant variability in the within-level scores. For instance, the score minimum obtained by a municipal organization is 34, and the maximum is 80, while for a federal organization, the minimum score is 13 and the maximum is 86. This means that organization's competencies to apply data governance practices do not depend on the political level of the organization nor the compulsory nature of opening government data.



**Figure 3. The score of Organizations**

Table 2 summarizes output statistics related to scoring practices' difficulty. While the scores practices' difficulty measure may vary between -1 and 1 logit, they also allow deriving the different maturity levels (i.e. from 1 to 5). This suggests that the closer the practice's difficulty score is to 1, the more complex the dimension is to implement and the higher the level of MM, and vice versa. Therefore, results highlight that the most accessible practices to apply (i.e. MM level=1) are primarily technical practices, such as identifying the different storing systems (-0.77) or possessing a data collection system (-0.44). On the contrary, Swiss organizations' most challenging data governance practices seem to belong to the structural practices (i.e. MM level=5). Notably, practices related to implementing a concrete data governance program (0.25) and their corresponding roles (0.30) seem more complex to adopt. Results also show that monitoring and relational practices, such as the implementation of a monitoring system (0.12) or the training (-0.01) and collaboration (-0.10) of public actors, seem faced challenges in their implementation. The difficulties encountered by the implementation of procedural practices are less clear. While the management of metadata (0.04) and data quality (0.034) seem complicated to implement for public organizations, activities such as the establishment of a data registry (-0.19) or a data protection guideline (-0.65) seem relatively easy to apply. Finally, with a fit value mean of 1.02 (i.e. infit and outfit values of 0.5 to 1.5), our statistics values satisfy the quality criterion (Dekleva and Drehmer 1997) and are considered as productive for measurement (Linacre 2002).

Dimensions	Data governance practices	Practices' Difficulty Score (PDS)	Infit	Outfit	MM Level
Structural	Roles & responsibilities	0.30124	0.687	0.682	5
Structural	Data governance program	0.25073	0.950	0.944	5
Monitoring	Implementation monitoring	0.12077	1.007	1.007	4
Procedural	Metadata management	0.04054	1.089	1.088	4
Procedural	Quality management	0.03476	1.018	1.015	4
Monitoring	Publication frequency	0.03476	1.301	1.301	4
Technical	Data transformation	0.01158	0.990	0.993	4
Structural	Decision-making authorities	-0.0058	1.336	1.335	4
Monitoring	Strategy advancement	-0.01172	0.851	0.852	4
Relational	Training	-0.01756	1.189	1.198	4
Relational	Collaboration	-0.10031	0.672	0.669	4
Procedural	Publication policy	-0.11829	1.058	1.058	3
Technical	Data cleansing	-0.18508	0.794	0.792	3

Procedural	Data registry	-0.19739	1.063	1.055	3
Technical	Data storing	-0.19739	0.656	0.658	3
Technical	OGD standards comply	-0.23462	1.108	1.113	3
Technical	Own system of data collection	-0.44303	1.754	1.747	2
Procedural	Data protection guidelines	-0.65110	1.213	1.214	1
Technical	Storing system	-0.77097	0.597	0.588	1

**Table 2. Results of the Rasch Model**

## Discussion

Based on the Rasch Model output statistics (Table 2), we present in Table 3 the data governance MM of Swiss OGD providers. This MM comprises five levels - *Initial*, *Managed*, *Defined*, *Measured* and *Optimized*. The *Initial* level is characterized by the absence of relational, structural, and monitoring practices of OGD. This first level only gathers data governance practices related to basic technical activities such as identifying the different data storing systems used (i.e. cloud, local server etc.) as well as procedural practice, including the conception of personal and sensible data guidelines. The *Initial* level represents the starting point for public organizations which operate their digital transformation. According to the results, the second level, namely *Managed*, seems still focus on technical orientation, such as identifying data origin. At this level, OGD providers question the origin of the different data owners, establish where data come from, and whether organizations own their proper data collection system or the IT departments handle it. While establishing whether organizations possess their data collection system is not limited to OGD, it represents a crucial step in the governance of OGD. In doing so, organizations can clarify data ownership which will help publish OGD. Thus, by making possible the collaboration and identification of decision-maker authorities, this level tend to facilitate the step towards a higher level.

After the system clarifications have been done and the security concepts defined, Level 3 represents the first stage for any organization wishing to disclose its data openly. It comprises activities related to what specialists call the "data lifecycle", i.e. making the data ready to use, such as cleaning and registering data according to OGD standards. While there is still a strong focus on technical practices, this level marks the emergence of procedural practices. For example, besides storing, cleaning, and registering data in compliance with OGD standards, the *Defined* level also requires establishing OGD publication policies. Then, level 4 represents the final steps of technical practices. Namely, *Measure*, this level notably requests a data transformation step, i.e., applying anonymization rules to sensible and personal data. It is also at the level that organizations should process to implement metadata and quality management activities. Metadata and data quality management have been identified in the literature as key factors in the re-utilization of OGD. For this purpose, the organization should implement relational activities such as training employees engaged in OGD governance and fostering collaboration between IT and business. Literature shows that in public organizations, the IT department often leads data-oriented activities (Marmier 2022). It is also at this level that organizations set up the monitoring of OGD strategy advancement and OGD publication. Finally, it seems that only organizations that established structural practices reach the highest level of the MM, namely *Optimized*. To do so, organizations must build a robust data governance model and determine the roles & responsibilities of the OGD governance stakeholders.

	<b>Initiate</b>	<b>Managed</b>	<b>Defined</b>	<b>Measure</b>	<b>Optimized</b>
<b>Technical</b>	Identifying data storing system	Identifying data origin	Following OGD standards of cleaning and storing data Preparing a publishable data registry	Transforming data given sensible data guidelines and OGD standards	

<b>Procedural</b>	Defining personal and sensible data guideline	Defining OGD publication policies	Establishing rules and managing OGD quality and metadata
<b>Relational</b>			Enforcing OGD stakeholders' training Enhancing collaboration between IT and business domain
<b>Monitoring</b>			Defining OGD publication strategy Monitoring OGD strategy advancement
<b>Structural</b>		Determining OGD decision-making authorities	Establishing OGD stakeholders' roles and responsibilities Completing the data governance program

**Table 3. OGD Governance Maturity Model**

Currently, and at all times since the beginning of OGD research, technical issues have been studied and rapidly perceived as a critical aspect of governance and the opening of public data. Conradie and Choenni (2014) already pointed out that the way an agency collects, stores, and manages data has a crucial effect on the open data of organizations (Zhao & Fan, 2018). Consequently, several authors addressed further recommendations to solve technical issues (Caro et al. 2008; Marmier and Mettler 2020a; Wahyudi et al. 2018). Accordingly, it is not surprising to see technical practices appear on the MM's first level and throughout the whole process. By showing that the second and third levels of the MM cluster organizations encourage data publication coordination, foster data and data description's quality as well as improve data registry (i.e. technical and procedural practices), our results also put forward that OGD providers seem to implement recommendations set up by the Swiss OGD Strategy for the 2019-2023 period. While the main aims for the 2014-2018 strategy were to - release official data, coordinate the publication and provision of official data, and establish an open-data culture – the last strategy has recommended that organizations who publish OGD follow OGD standards and concentrate on legal frameworks (Federal Statistical Office 2021).

However, implementing relational and monitoring practices at the fourth level draws our attention. While all organizations surveyed already published data on an OGD platform, we expected that activities such as defining an OGD publication strategy or enforcing training for employees engaged in OGD governance would appear at the earliest levels of MM. However, the results indicate that these practices only occur at the fourth level. These observations go against Ham et al. (2015), which place practices relative to employees' capability in the first stage of the model, or Solar et al. (2012) and Welle Donker and van Loenen (2017) who estimate strategy establishment as the first step in their MM. Furthermore, we are also surprised to see that structural governance practices such as role attribution and the building of a data governance program appear at the last levels and as the practices the less implemented. While our results follow Lis and Otto (2020) by showing that the lack of formalization and the unclarity of ecosystem roles and data providers are part of the most common challenges for public organizations, many authors define the

distribution of roles and responsibilities as a critical factor of data governance (Alhassan et al. 2019; Cheong and Chang 2007; Khatri and Brown 2010; Panian 2010; Weber 2009). According to Alhassan et al. (2019), establishing distinct roles and responsibilities attached to any data governance activities is the first step to guaranteeing a successful data-driven strategy. Following the same idea, Stanford University's Data Governance Office (2013) underlined that the role structure is essential to performing data governance. Similarly, Begg and Cairra (2012), Khatri and Brown (2010) recommended defining stakeholders' decision-related roles and responsibilities to stimulate a specific behavior regarding data proceedings. In summary, the MM derive from Swiss data, and OGD governance practices follow Nielsen's observations, which suggest that data governance is still attached to IT activities and has yet to become a general management discipline (Nielsen 2017).

This apparent disorder in data governance practices implementation may be explained by what Lewis Carroll calls the *Red Queen effect* in her book *Alice in wonderland*. Often observed in platform mechanism, this metaphor highlights that despite running ever faster, Alice and the Red Queen feel like they are still at the same level (Tiwana 2013). So to win the race, Alice needs to adapt faster than the Red Queen, leading her to make mistakes (Barnett and Hansen, 1996). Put differently, this paradox illustrates the effects of competition and the pressures it may induce on the competitors (Barnett and Hansen, 1996). Accordingly, in a world where several OGD ranking platforms exist (Jacob Arturo Rivera Perez et al. 2020; World Wide Web Foundation 2018) and where being a transparent government has become of primary importance for national sovereignty (Conseil fédéral 2014; Conseil fédéral 2018), governments publish faster as they can to win the OGD race. This effect was confirmed by the head of the Federal Statistic Office, that declared: *"to boost the OGD platform, we put everything public organizations owned on it. But next we had much cleaning to do, notably due to the quantity of PDF"*, but also by several authors that showed that dragging and dropping a plethora of public documents on OGD platforms was not sufficient to create valuable data (Safarov et al. 2017).

Besides explaining the data governance surprising order, the Red Queen effect also clarifies the results described in Figure 2, i.e. the score of organizations. Results show that organizations with better scores do not necessarily belong to the federal level. They indicate that the capability of the organizations to implement OGD governance practices seems independent from their political level. However, while both Swiss OGD strategies have supported and implemented open government data publication at all federal levels, only Swiss federal-level organizations must publicly open government data (Federal Office of Communications 2018; Federal Statistical Office 2021). The information in Appendix A indicates that the data governance capability seems rather dependent on the organization's function. While a higher score seems to be attained by organizations with some experience in data-driven activities, a lower score seems to be obtained by fewer experience organizations active in the health and educational area. It brings to light a two-speed implementation of OGD governance practices, where data-driven organizations are more likely *to win the race*. These results follow previous research. While the Swiss media outlets reported major issues in forwarding morbidity and mortality data from hospitals (Seydtaghia and Farine 2021), an OGD activities audit (Swiss Federal Audit Office 2019) revealed that statistical offices and organizations dealing with meteorological data, geodata or aid funds, among others, feel more comfortable with OGD activities than other "inexperienced" organizations. Additionally, Marmier and Mettler (2020a) came to the same conclusion by pointing out that organizations, mainly where data processing is part of their principal activity, seem to perform better OGD governance practices.

## Conclusion

Intending to shed some light on public organizations' OGD situation, this paper provides an overview of OGD governance practices and designs a maturity model to measure their level of implementation. We analyzed the results of 252 questionnaires distributed to OGD publisher organizations in Switzerland. Drawing on the Rasch model analysis, we first examined the public organizations' ability to perform OGD governance practices and then focused on governance practices' difficulties. We have been able to compare

respondents and determine each respondent's overall attitude. We show that public organizations seem involved in OGD governance practices from all political levels. We also put forward that the capability of the organizations to implement OGD governance practices seems independent from their political levels but depends on their work nature, which highlights two-speed implementation issues between "experienced" and "inexperienced" public organizations in Switzerland. Then we studied the data governance practices' difficulty. We pinpointed that technical data governance practices seem more manageable to set up for public organizations than structural and monitoring practices. Based on the difficulty levels of the different governance practices provided by the Rasch model, we finally derived an OGD governance maturity model. The maturity model illustrates and depicts practices to set up published OGD correctly through five levels: *Initial, Managed, Define, Measure and Optimized*. Each level of the OGD governance MM is composed of a set of key practices (i.e. items) that, once achieved, provide an advancement of the organizations' capabilities to govern OGD.

Through the proposition of a governance maturity model, this paper allows practitioners to improve their understanding of OGD governance practices and helps them comply with the objectives of the Swiss OGD strategy. The proposed model allows practitioners to identify challenges in the OGD governance process. It notably allows public organizations to evaluate their current situation of OGD governance awareness and effectiveness and thus assess where they want to be and the actions they should take to improve. In addition to providing guidelines on how to carry out a good OGD governance, the OGD governance maturity model gives information and documented assessment of the maturity of the Swiss public organizations. It represents an excellent way of presenting results to facilitate stakeholders' decision-making. It allows decision-makers to identify and highlight the strengths and weaknesses of inexperienced and experienced public organizations, fostering knowledge sharing. By gaining insight and knowledge of existing abilities and inabilities, decision-makers can also pinpoint critical gaps in OGD strategy. Thus, they may define the stages of advancement urgently needed to make the OGD strategy most reliable. Finally, showing that the implementation level of OGD governance practices is independent of political levels suggests the willingness of the Swiss public sector organizations to open and share government data. Our paper also provides theoretical contributions. According to Nielsen (2017), while data governance researchers mainly focus on theoretical and case studies, few engage in practice-oriented methods. Thus, this paper tends to fill this gap by following survey methods. Furthermore, the literature shows that only some studies are attracted by data governance in public sectors. Therefore, this paper goes further by providing a public sector perspective on data governance research. In addition, this study also extends existing OGD research by proposing a governance maturity model. This model provides OGD researchers with an overview of critical dimensions (e.g. structural, monitoring) that need more attention and must be more intensively studied. This allows researchers to design a search prioritization strategy, where the most complicated dimensions to implement would require more research than those already implemented. Finally, it provides a MM to the OGD research field, fostering and improving studies on OGD publication and re-utilization.

By only focusing on public organizations that already have published data on open data platforms, the sample does not entirely reflect the situation of OGD governance in Switzerland. It only provides an overview of organizations already interested in adopting OGD practices. While the Rasch model analysis does not suffer from missing data issues, the sample size may limit the development of MM. According to the experts, we need a relatively large sample to identify items' difficulties and thus derive maturity levels. Moreover, as the Rasch model does not directly produce maturity levels, it requires additional analysis. This may cause more significant efforts than other MM construction methods and discourage some researchers. Finally, one of the Rasch analysis's main advantages is providing the score of the items (i.e. the OGD governance practices' difficulty) and the score of survey respondents (i.e. the organizations' ability to implement OGD governance practices). Yet, we noticed that the *Jamovi* software could neither compute nor offer visualizations of the organization's score classification. As part of more global research, the purpose of this study was also to guide organizations that responded to this questionnaire in the data governance implementation practices. Consequently, we would like to redo the analysis with another software and thus offer personalized feedback to each interested institution. In doing so, we should be able

to provide specific guidelines and recommendations adapted to each situation. To go further, we also would like to extend this research by analysing questions that are not directly related to data governance practices. Therefore, we would like to examine the survey results with a method and approach different. Given the presented results, we see a need for more research clarifying the data governance semantics. Lastly, we also call for more practice-based research that helps develop such models and supports OGD stakeholders in assessing steps they need to take so that OGD strategy becomes a reality for all branches of government.

## Appendix A

### Overview of organizations' score

Language	Industry Sector	Political Level	Organization Name	Score
FR	Public administration and defence, compulsory social security	Federal	Federal Statistical Office (FSO)	82
DE	Public administration and defence, compulsory social security	Cantonal	Statistical Office of the Canton of Zurich	81
DE	Public administration and defence, compulsory social security	Communal	City of Uster	80
DE	Water supply, sewerage, waste management and remediation	Cantonal	AWEL Canton of Zurich	78
EN	Professional, scientific and technical activities	Federal	WSL Institute for Snow and Avalanche Research SLF, Davos	78
FR	Information and communication	Federal	Swiss Federal Archives	77
FR	Public administration and defence, compulsory social security	Federal	Federal Office of Customs and Border Protection	76
DE	Public administration and defence, compulsory social security	Federal	OFSP: radiological protection division/environmental radioactivity section	74
DE	Public administration and defence, compulsory social security	Federal	Federal Office of Civil Aviation FOCA	74
DE	Agriculture, forestry and fishing	Federal	WSL	73
DE	Public administration and defence, compulsory social security	Communal	Open Data Zurich	73
DE	Public administration and defence, compulsory social security	Cantonal	Statistical Office of the Canton of Zurich	71
DE	Professional, scientific and technical activities	Federal	WSL Institute for Snow and Avalanche Research SLF	70
FR	Professional, scientific and technical activities	Federal	Federal Office of Topography swisstopo	70
	[...]	[...]	[...]	
DE	Public administration and defence, compulsory social security	Cantonal	E-Government Office, Canton of Graubünden	34
DE	Public administration and defence, compulsory social security	Cantonal	Agricultural Office Zug	32
FR	Administrative and support service activities	Cantonal	Directorate General of Post-compulsory Education (DGEP)	32



DE	Agriculture, forestry and fishing	Federal	Federal Office for Agriculture BLW	29
FR	Public administration and defence, compulsory social security	Cantonal	State of Geneva	28
FR	Human health and social work activities	Cantonal	Children and Youth Services	26
DE	Public administration and defence, compulsory social security	Federal	Federal Office for the Environment FOEN	13

---

## References

- 24 heures. 2020. "Les Documents de L'administration Fédérale Davantage Demandés." Retrieved August 23, 2020, from <https://www.24heures.ch/les-documents-de-ladministration-federale-davantage-demandes-733388545015>
- Aarikka-Stenroos, L., and Ritala, P. 2017. "Network Management in the Era of Ecosystems: Systematic Review and Management Framework," *Industrial Marketing Management* (67), pp. 23-36.
- Ababneh, T.A.M., and Aga, M. 2019. "The Impact of Sustainable Financial Data Governance, Political Connections, and Creative Accounting Practices on Organizational Outcomes," *Sustainability* (11:20), p. 5676.
- Abate, M., Diegert, K., and Allen, H. 1998. "A Hierarchical Approach to Improving Data Quality," *Data Quality* (4:1), pp. 365-369.
- Abdelrahman, O.H. 2021. "Open Government Data: Development, Practice, and Challenges," in *Open Data*. IntechOpen.
- Abraham, R., Schneider, J., and Vom Brocke, J. 2019. "Data Governance: A Conceptual Framework, Structured Review, and Research Agenda," *International Journal of Information Management* (49), pp. 424-438.
- Abu-Shanab, E.A. 2015. "Reengineering the Open Government Concept: An Empirical Support for a Proposed Model," *Government Information Quarterly* (32:4), pp. 453-463.
- Afful-Dadzie, E., and Afful-Dadzie, A. 2017. "Liberation of Public Data: Exploring Central Themes in Open Government Data and Freedom of Information Research," *International Journal of Information Management* (37:6), pp. 664-672.
- Al-Ruithe, M., and Benkhelifa, E. 2020. "Determining the Enabling Factors for Implementing Cloud Data Governance in the Saudi Public Sector by Structural Equation Modelling," *Future Generation Computer Systems* (107), pp. 1061-1076.
- Albano, C.S., and Reinhard, N. 2014. "Open Government Data: Facilitating and Motivating Factors for Coping with Potential Barriers in the Brazilian Context," in *Proceedings of the International Conference on Electronic Government*, Dublin, Ireland, pp. 181-193.
- Alexopoulos, C., Loukis, E., and Charalabidis, Y. 2014a. "A Platform for Closing the Open Data Feedback Loop Based on Web2. o Functionality," *JeDEM-eJournal of eDemocracy and Open Government* (6:1), pp. 62-68.
- Alexopoulos, C., Zuiderwijk, A., Charapabidis, Y., Loukis, E., and Janssen, M. 2014b. "Designing a Second Generation of Open Data Platforms: Integrating Open Data and Social Media," in *International Conference on Electronic Government*, Berlin, Heidelberg: Springer, pp. 230-241.
- Alhassan, I., Sammon, D., and Daly, M. 2016. "Data Governance Activities: An Analysis of the Literature," *Journal of Decision Systems* (25:sup1), pp. 64-75.
- Alhassan, I., Sammon, D., and Daly, M. 2018. "Data Governance Activities: A Comparison Between Scientific and Practice-oriented Literature," *Journal of Enterprise Information Management* (31:2), pp. 300-316.
- Alhassan, I., Sammon, D., and Daly, M. 2019. "Critical Success Factors for Data Governance: A Theory Building Approach," *Information Systems Management* (36:2), pp. 98-110.
- Allison, B. 2010. "My data can't tell you that," *Open Government-Collaboration, Transparency, and Participation in Practice*, pp. 257-265.
- Almeida, F. 2018. "Strategies to Perform a Mixed Methods Study," *European Journal of Education Studies*.
- Almusharraf, A., Dhillon, G., and Samonas, S. 2015. "Mismatched Understanding of Is Security Policy: A Repgrid Analysis," in *Proceedings of the 21st Americas Conference on Information Systems*, Puerto Rico, pp. 1-12.
- Alvesson, M., and Sandberg, J. 2011. "Generating Research Questions Through Problematization," *Academy of Management Review* (36:2), pp. 247-271.
- Andersson Schwarz, J. 2017. "Platform Logic: An Interdisciplinary Approach to the Platform-based Economy," *Policy & Internet* (9:4), pp. 374-394.
- Anthopoulos, L., Reddick, C.G., Giannakidou, I., and Mavridis, N. 2016. "Why E-government Projects Fail? An Analysis of the Healthcare.gov Website," *Government Information Quarterly* (33:1), pp. 161-173.

- Asadullah, A., Faik, I., and Kankanhalli, A. 2018. "Digital Platforms: A Review and Future Directions," in: *The 22nd Pacific Asia Conference on Information Systems*. Yokohama, Japan: p. 248.
- Askham, N., Cook, D., Doyle, M., Fereday, H., Gibson, M., Landbeck, U., Lee, R., Maynard, C., Palmer, G., and Schwarzenbach, J. 2013. "The Six Primary Dimensions for Data Quality Assessment," pp. 432-435.
- Attard, J., Orlandi, F., Scerri, S., and Auer, S. 2015. "A systematic review of open government data initiatives," *Government Information Quarterly* (32:4), pp. 399-418.
- Atz, U. 2014. "The tau of data: A new metric to assess the timeliness of data in catalogues," in *Conference for E-Democracy and Open Government*, Krems, Austria, pp. 257-268.
- Bakıcı, T., Almirall, E., and Wareham, J. 2013. "A Smart City Initiative: The Case of Barcelona," *Journal of the Knowledge Economy* (4:2), pp. 135-148.
- Baldwin, C.Y., and Woodard, C.J. 2009. "The Architecture of Platforms: A Unified View," *Platforms, Markets and Innovation* (32), pp. 19-44.
- Barbour, A. 2010. "Exploring Some Ethical Dilemmas and Obligations of the Ethnographer," *Ethnography and Education* (5:2), pp. 159-173.
- Barnum, C.M. 2021. "Analyzing the Findings," in *Usability Testing Essentials (Second Edition)*, C.M. Barnum (ed.). Morgan Kaufmann, pp. 287-319.
- Barry, E., and Bannister, F. 2014. "Barriers to Open Data Release: A View from the Top," *Information Polity* (19:1, 2), pp. 129-152.
- Bates, J. 2012. "This is What Modern Deregulation Looks Like": Co-optation and Contestation in the Shaping of the UK's Open Government Data Initiative," *The Journal of Community Informatics* (8:2), pp. 1-20.
- Bates, J. 2014. "The Strategic Importance of Information Policy for the Contemporary Neoliberal State: The Sase of Open Government Data in the United Kingdom," *Government Information Quarterly* (31:3), pp. 388-395.
- Batini, C., Cappiello, C., Francalanci, C., and Maurino, A. 2009. "Methodologies for Data Quality Assessment and Improvement," *ACM Computing Surveys* (41:3), pp. 1-16.
- Bazeley, P. 2004. "Issues in Mixing Qualitative and Quantitative Approaches to Research," *Applying Qualitative Methods to Marketing Management Research* (141), p. 156.
- Becker, J., Knackstedt, R., Pöppelbuß, J.J.B., and Engineering, I.S. 2009. "Developing maturity models for IT management," (1:3), pp. 213-222.
- Begg, C., and Caira, T. 2012. "Exploring the SME Quandary: Data Governance in Practise in the Small to Medium-Sized Enterprise Sector," *Electronic Journal of Information Systems Evaluation* (15:1), p. 3-13.
- Benfeldt, O., Persson, J.S., and Madsen, S. 2020. "Data Governance as a Collective Action Problem," *Information Systems Frontiers* (22:2), pp. 299-313.
- Benitez-Paez, F., Degbelo, A., Trilles, S., and Huerta, J. 2018. "Roadblocks Hindering the Reuse of Open Geodata in Colombia and Spain: A Data User's Perspective," *ISPRS International Journal of Geo-Information* (7:1), pp. 1-6.
- Berghaus, S., and Back, A. 2016. "Stages in Digital Business Transformation: Results of an Eempirical Maturity Study," in *Mediterranean Conference of Information Systems*, Cyprus, p. 22.
- Berghmans, S., Cousijn, H., Deakin, G., Meijer, I., Mulligan, A., Plume, A., De, R.S., Rushforth, A., Tatum, C., and Van, L.T. 2017. "Open Data: The Researcher Perspective," Leiden University Centre for Science and Technology Studies, and Elsevier, Leiden, The Netherlands, pp. 1-49.
- Berners-Lee, T. 2006. "Linked Data-design Issues".
- Berrone, P., Ricart, J., & Carrasco-Farré, C. . 2017. "The Open Kimono: Toward a General Framework for Open Data Initiatives in Cities," *California Management Review* (59:1), pp. 39-70.
- Bertot, J.C., Jaeger, P.T., and Grimes, J.M. 2010. "Using ICTs to Create a Culture of Transparency: E-government and Social Media as Openness and Anti-corruption Tools for Societies," *Government Information Quarterly* (27:3), pp. 264-271.
- Biberoglu, E., and Haddad, H. 2002. "A Survey of Industrial Experiences with CMM and the Teaching of CMM Practices," *Journal of Computing Sciences in Colleges* (18:2), pp. 143-152.
- Blaschke, M., Haki, K., Aier, S., and Winter, R. 2019. "Taxonomy of Digital Platforms: A Platform Architecture Perspective," in *Proceedings of the 14th International Conference on Wirtschaftsinformatik*, Siegen, Germany.

- Bond, T.G., Yan, Z., Stone, G., and Beltyukova, S. 2015. "Making Measures, Setting Standards, and Rasch Regression," in *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Routledge, pp. 187-225.
- Bonina, C., and Eaton, B. 2020. "Cultivating Open Government Data Platform Ecosystems through Governance: Lessons from Buenos Aires, Mexico City and Montevideo," *Government Information Quarterly* (37:3), p. 101479.
- Bonina, C., Eaton, B., and Henningsson, S. 2018. "Governing Open Data Platforms to Cultivate Innovation Ecosystems: The Case of the Government of Buenos Aires," in *39th International Conference on Information Systems, ICIS 2018*, San Francisco, California, USA: Association for Information Systems. AIS Electronic Library (AISeL).
- Bonina, C., Koskinen, K., Eaton, B., and Gawer, A. 2021. "Digital Platforms for Development: Foundations and Research Agenda," *Information Systems Journal* (31:6), pp. 869-902.
- Boone, W.J., Staver, J.R., and Yale, M.S. 2013. *Rasch Analysis in the Human Sciences*. Springer.
- Borgman, C.L. 2015. *Big Data, Little Data, No Data: Scholarship in the Networked World*. Cambridge, MA: MIT press.
- Brackett, M., and Earley, P.S. 2009. "The DAMA Guide to The Data Management Body of Knowledge."
- Braun, V., and Clarke, V. 2006. "Using Thematic Analysis in Psychology," *Qualitative Research in Psychology* (3:2), pp. 77-101.
- Braunschweig, K., Eberius, J., Thiele, M., and Lehner, W. 2012. "The State of Open Data Limits of Current Open Data Platforms," in *Proceedings of the International World Wide Web Conference*, Lyon, France.
- Brous, P., Janssen, M., and Krans, R. 2020. "Data Governance as Success Factor for Data Science," in *Proceedings of the Conference on e-Business, e-Services and e-Society*, Skukuza, South Africa: Springer, pp. 431-442.
- Brous, P., Janssen, M., and Vilminko-Heikkinen, R. 2016. "Coordinating Decision-making in Data Management Activities: A Systematic Review of Data Governance Principles," in *International Conference on Electronic Government and the Information Systems Perspective*, Porto, Portugal: Springer, pp. 115-125.
- Buchmann, M. 2017. "The Need for Competition Between Decentralized Governance Approaches for Data Exchange in Smart Electricity Grids : Fiscal federalism vs. Polycentric governance," *Journal of Economic Behavior & Organization* (139), pp. 106-117.
- Bürgi-Schmelz, A. 2013. "Wirtschaftliche Auswirkungen von Open Government Data." pp. 3-124.
- Butts-Wilmsmeyer, C.J., Rapp, S., and Guthrie, B. 2020. "The Technological Advancements that Enabled the Age of Big Data in the Environmental Sciences: A History and Future Directions," *Current Opinion in Environmental Science & Health* (18), pp. 63-69.
- Caillaud, B., and Jullien, B. 2003. "Chicken & Egg: Competition Among Intermediation Service Providers," *RAND Journal of Economics*, pp. 309-328.
- Carayannis, E.G., Gonzalez, E., and Wetter, J. 2003. "The Nature and Dynamics of Discontinuous and Disruptive Innovations from a Learning and Knowledge Management Perspective," in *The International Handbook on Innovation, Part II*. Kidlington, UK: Elsevier, pp. 115-138.
- Caro, A., Calero, C., Caballero, I., and Piattini, M. 2008. "A proposal for a set of attributes relevant for Web portal data quality," *Software Quality Journal* (16:4), pp. 513-542.
- Carpenter, M.A., Geletkanycz, M.A., and Sanders, W.G. 2004. "Upper Echelons Research Revisited: Antecedents, Elements, and Consequences of Top Management Team Composition," *Journal of Management Information Systems* (30:6), pp. 749-778.
- Casilli, A., and Posada, J. 2019. "The Platformization of Labor and Society," *Society and the Internet: How Networks of Information and Communication are Changing our Lives*, pp. 293-306.
- Cennamo, C., and Santaló, J. 2015. "How to Avoid Platform Traps," in: *MIT Sloan Management Review*. p. 12.
- Chan, C.M. 2013. "From Open Data to Open Innovation Strategies: Creating E-services Using Open Government Data," in *Proceedings of the 46th Hawaii International Conference on System Sciences*, Wailea, Hawaii, USA, pp. 1890-1899.
- Charalabidis, Y., Loukis, E., and Alexopoulos, C. 2014. "Evaluating Second Generation Open Government Data Infrastructures Using Value Models," in *The 47th Hawaii International Conference on System Sciences (HICSS)*, Waikoloa, HI, USA: IEEE, pp. 2114-2126.

- Charalabidis, Y., Zuiderwijk, A., Alexopoulos, C., Janssen, M., Lampoltshammer, T., and Ferro, E. 2018a. "Open Data Evaluation Models: Theory and Practice," in *The World of Open Data*. Springer International Publishing, pp. 137-172.
- Charalabidis, Y., Zuiderwijk, A., Alexopoulos, C., Janssen, M., Lampoltshammer, T., and Ferro, E. 2018b. "Open Data Infrastructures," in *The World of Open Data*. Springer International Publishing, pp. 95-113.
- Charalampos, A., Loukis, E., Charalabidis, Y., and Zuiderwijk, A. 2013. "An Evaluation Framework for Traditional and Advanced Open Public Data e-Infrastructures," in *Proceedings of the 13th European conference on E-government*, Como, Italy, pp. 102-111.
- Chatfield, A.T., and Reddick, C.G. 2017. "A longitudinal cross-sector analysis of open data portal service capability: The case of Australian local governments," *Government Information Quarterly* (34:2), pp. 231-243.
- Cheong, L.K., and Chang, V. 2007. "The Need for Data Governance: A Case Study," in *Proceedings of the 18th Australasian Conference on Information Systems*, University of Southern Queensland, Toowoomba, Australia, p. 100.
- Chiles, T.H., Meyer, A.D., and Hench, T.J. 2004. "Organizational Emergence: The origin and Transformation of Branson, Missouri's Musical Theaters," *Organization Science* (15:5), pp. 499-519.
- Ciancarini, P., Poggi, F., and Russo, D. 2016. "Big Data Quality: A Roadmap for Open Data," in *IEEE Second International Conference on Big Data Computing Service and Applications (BigDataService)*, Oxford, United Kingdom, pp. 210-215.
- Ciuriak, D. 2018. "The Economics of Data: Implications for the Data-driven economy." *Data Governance in the Digital Age*. Retrieved September 1, 2020, from <https://www.cigionline.org/articles/economics-data-implications-data-driven-economy>
- CKAN Organization. 2018. "Data quality, what is it?" Retrieved 6 November, 2018, from <https://ckan.org/2011/01/20/data-quality-what-is-it/>
- CMMInstitute. 2021. "The Capability Maturity Model Integration." Retrieved November 8, 2021, from <https://cmmiinstitute.com/dmm>
- Coelho, J., and Valente, M.T. 2017. "Why Modern Open Source Projects Fail," in *Proceedings of the 11th Joint Meeting on Foundations of Software Engineering*, Paderborn, Germany pp. 186-196.
- Cohn, B.L. 2015. "Data Governance: A Quality Imperative in the Era of Big Data, Open Data, and Beyond," *A Journal of Law And Policy for the Information Society* (10:3), pp. 811-826.
- Concilio, G., Molinari, F., and Morelli, N. 2017. "Empowering Citizens with Open Data by Urban Hackathons," in *E-Democracy and Open Government*, Krems an der Donau, Austria: IEEE.
- Confédération Suisse. 2020. "Opendata.swiss Homepage." Retrieved August 20, 2020, from <https://opendata.swiss/en/>
- Conradie, P., and Choenni, S. 2012. "Exploring Process Barriers to Release Public Sector Information in Local Government," in *Proceedings of the 6th International Conference on Theory and Practice of Electronic Governance*, Albany, USA, pp. 5-13.
- Conradie, P., and Choenni, S. 2014. "On the Barriers for Local Government Releasing Open Data," *Government Information Quarterly* (31), pp. S10-S17.
- Conseil fédéral. 2014. "Strategie en Matiere de Libre Accès aux Données Publiques en Suisse pour les Années 2014 à 2018."
- Conseil fédéral. 2018. "Stratégie en Matière de Libre Accès aux Données Publiques en Suisse pour les Années 2019 à 2023 (Stratégie Open Government Data)."
- Constantinides, P., Henfridsson, O., and Parker, G.G. 2018. "Introduction: Platforms and Infrastructures in the Digital Age," *Information Systems Research* (29:2), pp. 381 - 400.
- Corbett, J., Templier, M., and Takeda, H. 2018. "The Making of a Top'Open Data City: A Case Study of Edmonton's Open Data Initiative," in *Proceedings of the 51st Hawaii International Conference on System Sciences*, Waikoloa Village, Hawaii, USA, pp. 2443-2452.
- Corsar, D., and Edwards, P. 2017. "Challenges of open data quality: more than just license, format, and customer support," *ACM Journal of Data Information Quality* (9), pp. 1-3.
- Cranefield, J., Robertson, O., and Oliver, G. 2014. "Value in the Mash: Exploring the Benefits, Barriers and Enablers of Open Data Apps," in *Proceedings of the 22nd European Conference on Information Systems* Tel Aviv, Israel, pp. 1-15.
- Cresswell, J., and Plano Clark, V. 2011. *Designing and Conducting Mixed Method Research*. . Sage.

- Creswell, J.W., and Clark, V.L.P. 2018. *Designing and Conducting Mixed Methods Research*. Sage Publications.
- Creswell, J.W., and Clark, V.P. 2011. *Designing and Conducting Mixed Research Methods*. London: Thousand Oaks.
- Crusoe, J., and Melin, U. 2018. "Investigating Open Government Data Barriers," in *Proceedings of the 17th International Conference on Electronic Government*, Krems, Austria, pp. 169-183.
- Cunningham Dahl-Jørgensen, T., and Parmiggiani, E. 2020. "Platformization of the Public Sector: Assessing the Space of Possibility for Participation," in *Proceedings of the 16th Participatory Design Conference 2020*, New York, United States, pp. 35-39.
- Dahbi, K.Y., Lamharhar, H., and Chiadmi, D. 2018. "Toward an Evaluation Model for Open Government Data Portals," in *International Conference Europe Middle East & North Africa Information Systems and Technologies to Support Learning*, Marrakech, Morocco: Springer, pp. 502-511.
- DAMA, I. 2009. *The DAMA Guide to the Data Management Body of Knowledge*. New Jersey: Technics Publications.
- Daniel, Jonathan, Tim M, Antti P, Rufus P, Julian T, and Ton Z. 2012. "Open Data Handbook." Retrieved November, 30, 2022, from <https://opendatahandbook.org/guide/en/>
- Danneels, L., Viaene, S., and Van den Bergh, J. 2017. "Open Data Platforms: Discussing Alternative Knowledge Epistemologies," *Government Information Quarterly* (34:3), pp. 365-378.
- Davenport, T.H. 2013. "Analytics 3.0," *Harvard Business Review* (91:12), pp. 64-72.
- Davies, T. 2010. *Open data, democracy and public sector reform: a look at open government data use from data.gov.uk*. Oxford, UK: University of Oxford.
- Dawes, S.S., Vidiasova, L., and Parkhimovich, O. 2016. "Planning and Designing Open Government Data Programs: An Ecosystem Approach," *Government Information Quarterly* (33:1), pp. 15-27.
- Dawson, G.S., Denford, J.S., Williams, C.K., Preston, D., and Desouza, K.C. 2016. "An Examination of Effective IT Governance in The Public Sector Using The Legal View of Agency Theory," *Journal of Management Information Systems* (33:4), pp. 1180-1208.
- de Bary Heinrich Anton 1878. ""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)* (51), pp. 121-126.
- De Reuver, M., Nederstigt, B., and Janssen, M. 2018. "Launch Strategies for Multi-sided Data Analytics Platforms," in *26th European Conference on Information Systems, ECIS 2018*, Portsmouth, UK: Association for Information Systems.
- De Reuver, M., Nikou, S., and Bouwman, H. 2016. "Domestication of Smartphones and Mobile Applications: A Quantitative Mixed-method Study," *Mobile Media & Communication* (4:3), pp. 347-370.
- De Tuya, M., Cook, M., Sutherland, M.K., and Luna-Reyes, L.F. 2017. "Information Requirements to Create Public Value: Sharing and Opening Data to Address Urban Blight," *Transforming Government: People, Process and Policy* (11:1), pp. 79-98.
- Dekleva, S., and Drehmer, D. 1997. "Measuring Software Engineering Evolution: A Rasch Calibration," *Information Systems Research* (8:1), pp. 95-104.
- Dhillon, A., and Singh, A. 2019. "Machine Learning in Healthcare Data Analysis: A Survey," *Journal of Biology and Today's World* (8:6), pp. 1-10.
- Ding, L., Lebo, T., Erickson, J.S., DiFranzo, D., Williams, G.T., Li, X., Michaelis, J., Graves, A., Zheng, J.G., and Shangguan, Z. 2011. "TWC LOGD: A Portal for Linked Open Government Data Ecosystems," *Journal of Web Semantics: Science, Services and Agents on the World Wide Web* (9:3), pp. 325-333.
- Dougherty, D., and Dunne, D.D. 2011. "Organizing Ecologies of Complex Innovation," *Organization Science* (22:5), pp. 1214-1223.
- Dreibelbis, A. 2008. *Enterprise Master data Management: An SOA Approach to Managing Core Information*. India: IBM Press.
- Dulong de Rosnay, M., and Janssen, K. 2014. "Legal and Institutional Challenges for Opening Data Across Public Sectors: Towards Common Policy Solutions," *Journal of Theoretical Applied Electronic Commerce Research* (9:3), pp. 1-14.
- E-Government Switzerland. 2014. "Open Government Data Strategy for Switzerland 2014-2018." Berne: Program Office E-Government Switzerland,
- Federal IT Steering Unit, pp. 2-15.



- E-Government Switzerland. 2016. "The Project "Open Government Data 2014-2018"." Retrieved September, 8, 2020, from <https://www.egovernment.ch/en/umsetzung/e-government-schweiz-2008-2015/open-government-data-schweiz/>
- Eisenhardt, K.M., and Graebner, M.E. 2007. "Theory Building from Cases: Opportunities and Challenges," *Academy of Management Journal* (50:1), pp. 25-32.
- El-Emam, K., Gray, J., Greal, M., Hogle, M., Lichtenfeld, L., McGraw, D., Overhage, J.M., Shah, N., Marchibroda, J., and Wyden, R. 2013. "A Policy Forum on the Use of Big Data in Health Care." Retrieved September, 13, 2021, from <https://bipartisanpolicy.org/download/?file=/wp-content/uploads/2019/03/Use-of-Big-Data-in-Health-Care.pdf>
- El Emam, K., Rodgers, S., and Malin, B. 2015. "Anonymising and Sharing Individual Patient Data," *bmj* (350), p. h1139.
- Elpez, I., and Fink, D. 2006. "Information Systems Success in the Public Sector: Stakeholders' Perspectives and Emerging Alignment Model," *Informing Science: International Journal of an Emerging Transdiscipline* (3), pp. 219-231.
- Emran, N.A. 2015. "Data Completeness Measures," in *Pattern Analysis, Intelligent Security and the Internet of Things*. Springer, pp. 117-130.
- Ernst & Young. 2014. "Big data Changing the way businesses compete and operate Insights on governance, risk and compliance."
- Erzberger, C., Kelle, U., Tashakkori, A., and Teddlie, C. 2003. "Handbook of Mixed Methods in Social and Behavioral Research," in: *Handbook of Mixed Methods in Social and Behavioral Research*. pp. 457-490.
- Etat de Vaud. 2018. "Strategie Numérique du Canton de Vaud." Retrieved August 3, , 2021, from <https://www.vd.ch/strategie-numerique-du-canton-de-vaud-digitale-strategie-des-kantons-waadt/>
- European Commission. 2020. "A European Strategy for Data." Retrieved April 28, 2020, from <https://ec.europa.eu/digital-single-market/en/policies/building-european-data-economy>
- European Commission. 2021. "Open Data." Retrieved 19. April, 2021, from <https://digital-strategy.ec.europa.eu/en/policies/open-data-o>
- European Data Portal. 2017. "Open Data in Europe." Retrieved August 14, 2018, from <https://www.europeandataportal.eu/en/dashboard#2017>
- Evans, P.C., and Gawer, A. 2016. *The Rise of the Platform Enterprise: A Global Survey*.
- Fadler, M., Lefebvre, H., and Legner, C. 2021. "Data Governance: From Master Data Quality to Data Monetization," in: *the 29th European Conference on Information System*. Marrakech, Morocco.
- Fane, B., Ayris, P., Hahnel, M., Hrynaszkiewicz, I., Baynes, G., and al., e. 2019. "The State of Open Data Report 2019." Retrieved April 28, 2020, from <https://doi.org/10.6084/m9.figshare.9980783.v2>
- Farrell, J., and Saloner, G. 1986. "Installed Base and Compatibility: Innovation, Product Preannouncements, and Predation," *The American Economic Review*, pp. 940-955.
- Faulkner-Gurstein, R., and Wyatt, D. 2021. "Platform NHS: Reconfiguring a Public Service in the Age of Digital Capitalism," *Science, Technology, & Human Values*, pp. 1-21.
- Federal Council. 2014. "Open Government Data Strategy for Switzerland 2014 - 2018." from [https://www.egovernment.ch/index.php/download\\_file/force/761/3631/](https://www.egovernment.ch/index.php/download_file/force/761/3631/)
- Federal Office of Communications. 2018. "Digital Switzerland Strategy." Retrieved August 14, 2018, from <https://www.bakom.admin.ch/bakom/en/homepage/digital-switzerland-and-internet/strategie-digitale-schweiz/strategy.html>
- Federal Statistical Office. 2020a. "The opendata.swiss Portal." Retrieved September, 8, 2020, from <https://www.bfs.admin.ch/bfs/en/home/services/ogd/portal.html>
- Federal Statistical Office. 2020b. "Switzerland's "Priority Open Government Data"." Retrieved August 27, 2020, from <https://www.bfs.admin.ch/bfs/en/home/services/ogd/priority.html>
- Federal Statistical Office. 2021. "Open Government Data strategy 2019–2023." Retrieved May 21, from <https://www.bfs.admin.ch/bfs/fr/home/services/ogd/strategie.html>
- Figgou, L., and Pavlopoulos, V. 2015. "Social Psychology: Research Methods," *International Encyclopedia of the Social & Behavioral Sciences* (22), pp. 544-552.
- Financial Times. 2020. "Microsoft Throws Weight Behind Open Data Movement." Retrieved April 28, 2020, from <https://www.ft.com/content/661b16ff-f86c-4dad-a557-2e231501bf58>
- Flick, U. 2004. "Triangulation in Qualitative Research," *A Companion to Qualitative Research* (3), pp. 178-183.

- Fox, C., Levitin, A., and Redman, T. 1994. "The Notion of Data and its Quality Dimensions," *Information Processing & Management* (30:1), pp. 9-19.
- Francey, A. 2021. "A Conceptual Framework for Open Government Data: Drawing upon Conditions Surrounding Use," in *Annual Conference of The Italian Chapter of AIS (ItAIS)*, Trento, Italy.
- Francey, A., and Mettler, T. 2021. "The Effects of Open Government Data: Some Stylised Facts," *Information Polity* (Pre-press:Pre-press), pp. 1-16.
- Fransella, F., Bell, R., and Bannister, D. 2004. *A Manual for Repertory Grid Technique*, (2 ed.). Chichester, UK: John Wiley & Sons.
- Fusch, P.I., Fusch, G.E., and Ness, L.R. 2017. "How to Conduct a Mini-ethnographic Case Study: A Guide for Novice Researchers," *The Qualitative Report* (22:3), p. 923.
- Gascó-Hernández, M., Martín, E.G., Reggi, L., Pyo, S., and Luna-Reyes, L.F. 2018. "Promoting the Use of Open Government Data: Cases of Training and Engagement," *Government Information Quarterly* (35:2), pp. 233-242.
- Gauld, R. 2007. "Public Sector Information System Project Failures: Lessons from a New Zealand Hospital Organization," *Government Information Quarterly* (24:1), pp. 102-114.
- Gawer, A. 2010. "Towards a General Theory of Technological Platforms," in: *The 'Opening-up Innovation: Strategy, Organisation and Technology' Conference*. Imperial College, London, UK.
- Gawer, A., and Cusumano, M. A. 2008. "How Companies Become Platform Leaders," in: *MIT Sloan Management Review*. pp. 28-35.
- Gawer, A., and Cusumano, M.A. 2002. *Platform Leadership: How Intel, Microsoft, and Cisco Drive Industry Innovation*. Harvard Business School Press Boston.
- Gawer, A., and Cusumano, M.A. 2014. "Industry Platforms and Ecosystem Innovation," *Journal of Product Innovation Management* (31:3), pp. 417-433.
- Gelhaar, J., and Otto, B. 2020. "Challenges in the Emergence of Data Ecosystems," in *Pacific Asian Conference on Information Systems*, Dubai, United Arab Emirates, pp. 175-192.
- Gibson, J.J. 1966. *The Senses Considered as Perceptual Systems*. Boston, USA: Houghton Mifflin.
- Goel, K., Emamjome, F., and ter Hofstede, A. 2021. "Data Governance for Managing Data Quality in Process Mining," in *Proceedings of the 42nd International Conference on Information Systems*, Austin, Texas, USA.
- Goldkuhl, G. 2012. "Pragmatism vs Interpretivism in Qualitative Information Systems Research," *European Journal of Information Systems* (21:2), pp. 135-146.
- Gonzalez-Zapata, F., and Heeks, R. 2015. "The Multiple Meanings of Open Government Data: Understanding Different Stakeholders and their Perspectives," *Government Information Quarterly* (32:4), pp. 441-452.
- González-Zapata, F., and Heeks, R. 2015. "Understanding Multiple Roles of Intermediaries in Open Government Data," in *13th International Conference on Social Implications of Computers in Developing Countries*, pp. 1-10.
- Goodwin, M. 2009. *Governance*. University of Exeter, Exeter, UK: Elsevier Ltd.
- Guba, E.G., and Lincoln, Y.S. 1982. "Epistemological and Methodological Bases of Naturalistic Inquiry," *Ectj* (30:4), pp. 233-252.
- Hadi, M.A., Alldred, D.P., Closs, S.J., Marczewski, K., and Briggs, M. 2013. "A Mixed Methods Evaluation of a Nurse-pharmacist-managed Pain Clinic: Design, Rationale and Limitations," *Canadian Pharmacists Journal* (146:4), pp. 197-201.
- Ham, J., Lee, J.-N., Kim, D., and Choi, B. 2015. "Open Innovation Maturity Model for the Government: An Open System Perspective," in: *International Conference on Information Systems*,. Fort Worth, Texas, USA.
- Hambrick, D.C. 2007. "Upper Echelons Theory: An Update," *The Academy of Management Review* (32:2), pp. 334-343.
- Hancock, D.R., and Algozzine, B. 2006. *Doing Case Study Research: A Practical Guide for Beginning Researchers*. Columbia University, New York, NY: Teachers College Press
- Hardless, G., Meilinger, T., and Mallot, H.A. 2015. "Virtual Reality and Spatial Cognition," in *International Encyclopedia of the Social Behavioral Sciences*, J.Wright (ed.). Amsterdam: Elsevier, pp. 133-137.
- Harrison, T.M., Pardo, T.A., and Cook, M. 2012. "Creating Open Government Ecosystems: A Research and Development Agenda," *Future Internet* (4:4), pp. 900-928.
- Harrison, T.M., and Sayogo, D.S. 2014. "Transparency, Participation, and Accountability Practices in Open Government: A Comparative Study," *Government Information Quarterly* (31:4), pp. 513-525.



- Hassan, M.I.A., and Twinomurizi, H. 2018. "A Systematic Literature Review of Open Government Data Research: Challenges, Opportunities and Gaps," in *2018 Open Innovations Conference (OI)*: IEEE, pp. 299-304.
- Heath, L. 2015. "Triangulation: Methodology," in *International Encyclopedia of the Social & Behavioral Sciences*, J.D. Wright (ed.). Elsevier, pp. 639-644.
- Heimstädt, M., Saunderson, F., and Heath, T. 2014. "Conceptualizing Open Data Ecosystems: A Timeline Analysis of Open Data Development in the UK," in *CeDEM14: Proceedings of the Conference for E-Democracy and Open Government*, Krems, Austria, p. 245.
- Henninger, M. 2013. "The value and challenges of public sector information," *Cosmopolitan Civil Societies: An Interdisciplinary Journal* (5:3), pp. 75-95.
- Hjalmarsson, A., Johansson, N., and Rudmark, D. 2015. "Mind the Gap: Exploring Stakeholders' Value with Open Data Assessment," in *The 48th Hawaii International Conference on System Sciences*, Waikoloa, HI, USA: IEEE, pp. 1314-1323.
- Hodgkinson, G.P., and Sparrow, P. 2002. *The Competent Organization: A Psychological Analysis of the Strategic Management Process*. Buckingham, UK: Open University Press
- Holden, M.T., and Lynch, P. 2004. "Choosing the Appropriate Methodology: Understanding Research Philosophy," *The Marketing Review* (4:4), pp. 397-409.
- Holgeid, K., and Thompson, M. 2013. "A Reflection on Why Large Public Projects Fail," in *The Governance of Large-Scale Projects: Linking Citizens and the State*, H.S. A. Roemmele (ed.). Nomos Verlagsgesellschaft, pp. 219-244.
- Holloway, I., and Daymon, C. 2010. *Qualitative Research Methods in Public Relations and Marketing Communications*. London, UK: Routledge.
- Holt, V., Ramage, M., Kear, K., and Heap, N. 2015. "The Usage of Best Practices and Procedures in the Database Community," *Information Systems* (49), pp. 163-181.
- Hong, S., and Kim, S.H. 2016. "Political Polarization on Twitter: Implications for the Use of Social Media in Digital Governments," *Government Information Quarterly* (33:4), pp. 777-782.
- Huang, R., Lai, T., and Zhou, L. 2017. "Proposing a Framework of Barriers to Opening Government Data in China," *Library Hi Tech* (35:3), pp. 421-438.
- Huh, Y., Keller, F., Redman, T.C., and Watkins, A. 1990. "Data Quality," *Information and Software Technology* (32:8), pp. 559-565.
- Hunter, M.G. 1997. "The Use of RepGrids to Gather Interview Data about Information Systems Analysts," *Information Systems Journal* (7:1), pp. 67-81.
- IBM. 2007. *The IBM Data Governance Council Maturity Model: Building a roadmap for effective data governance*. Somers, NY, USA: IBM Corporation.
- Immonen, A., Ovaska, E., Kalaoja, J., and Pakkala, D. 2016. "A Service Requirements Engineering Method for a Digital Service Ecosystem," *Service Oriented Computing and Applications* (10:2), pp. 151-172.
- Iversen, J., Nielsen, P.A., and Norbjerg, J. 1999. "Situated Assessment of Problems in Software Development," *ACM SIGMIS Database: the Database for Advances in Information Systems* (30:2), pp. 66-81.
- Jacob Arturo Rivera Perez, Cecilia Emilsson, and Ubaldi, B. 2020. "Open, Useful and Re-usable data (OURdata) Index: 2019," OECD Policy Papers on Public Governance No. 1.
- Jankowicz, D. 2005. *The Easy Guide to Repertory Grids*. Chichester, UK: John Wiley & Sons.
- Janssen, M., Charalabidis, Y., and Zuiderwijk, A. 2012. "Benefits, Adoption Barriers and Myths of Open Data and Open Government," *Information Systems Management* (29:4), pp. 258-268.
- Jelinek, M., and Litterer, J. 1994. "Toward a Cognitive Theory of Organizations," *Advances in Managerial Cognition Organizational Information Processing* (5), pp. 3-41.
- Jetzek, T., Avital, M., and Bjorn-Andersen, N. 2014. "Data-driven innovation through open government data," *Journal of Theoretical and Applied Electronic Commerce Research* (9:2), pp. 100-120.
- Jetzek, T., Avital, M., and Bjorn-Andersen, N. 2019. "The Sustainable Value of Open Government Data," *Journal of the Association for Information Systems* (20:6), p. 6.
- Kalampokis, E., Tambouris, E., and Tarabanis, K. 2011a. "A Classification Scheme for Open Government Data: Towards Linking Decentralised Data," *International Journal of Web Engineering and Technology* (6:3), 2011/01/01, pp. 266-285.
- Kalampokis, E., Tambouris, E., and Tarabanis, K. 2011b. "Open Government Data: A Stage Model," in *Proceedings of International Conference on Electronic Government*, Delft, The Netherlands, pp. 235-246.

- Kankam, P.K. 2019. "The Use of Paradigms in Information Research," *Library & Information Science Research* (41:2), pp. 85-92.
- Kassen, M. 2018. "Adopting and Managing Open Data," *Aslib Journal of Information Management* (70:5), pp. 518-537.
- Kelly, G. 2003. *The Psychology of Personal Constructs: Volume Two: Clinical Diagnosis and Psychotherapy*. London: Routledge.
- Khatcherian, S., and Jefferson, G. 2009. "Information Governance: Challenges for Financial Services Firms Operating in the Information Age," *Convergence* (5), p. 80.
- Khatri, V., and Brown, C.V. 2010. "Designing Data Governance," *Communications of the ACM* (53:1), pp. 148-152.
- Kim, H.Y., and Cho, J.-S. 2018. "Data Governance Framework for Big Data Implementation with NPS Case Analysis in Korea," *Journal of Business and Retail Management Research* (12:3), pp. 36-46.
- Kirstein, F., Dittwald, B., Dutkowski, S., Glikman, Y., Schimmler, S., and Hauswirth, M. 2019. "Linked Data in the European Data Portal: A Comprehensive Platform for Applying DCAT-AP," in *International Conference on Electronic Government*, San Benedetto Del Tronto, Italy: Springer, pp. 192-204.
- Kitchin, R. 2014. "Ethical, Political, Social and Legal Concerns," in *The Data Revolution: Big Data, Open Data, Data Infrastructures and their Consequences*. London, UK: Sage, pp. 165-183.
- Kitchin, R., and Thrift, N. 2009. *International Encyclopedia of Human Geography*. London: Elsevier.
- Kordzadeh, N., and Warren, J. 2013. "Toward a Typology of Health 2.0 Collaboration Platforms and Websites," *Health and Technology* (3:1), pp. 37-50.
- Korzaan, M., and Erskine, M. 2018. "The Influence of Goal Orientation, Goal Commitment, and Implementation Mindset on Project Success," in *The 24th Americas Conference on Information Systems*, New Orleans, Louisiana, USA.
- Krasikov, P., Obrecht, T., Legner, C., and Eurich, M. 2020. "Is Open Data Ready for Use by Enterprises? Learnings from Corporate Registers," in *the 9th International Conference on Data Science, Technology and Application*, Honolulu, HI, USA, pp. 109-120.
- Kubler, S., Robert, J., Neumaier, S., Umbrich, J., and Le Traon, Y. 2017. "Comparison of Metadata Quality in Open Data Portals Using the Analytic Hierarchy Process," *Government Information Quarterly* (35:1), pp. 13-29.
- Kvamsdal, P. 2017. "Open Government Data: A Literature Review and a Research Agenda." Retrieved September 1, 2020, from [https://www.researchgate.net/publication/326160800\\_Open\\_Government\\_Data\\_-\\_A\\_Literature\\_Review\\_and\\_a\\_Research\\_Agenda](https://www.researchgate.net/publication/326160800_Open_Government_Data_-_A_Literature_Review_and_a_Research_Agenda)
- Ladley, J. 2010. *Making Enterprise Information Management (EIM) Work for Business: A Guide to Understanding Information as an Asset*. Morgan Kaufmann.
- Ladley, J. 2019. *Data Governance: How to Design, Deploy, and Sustain an Effective Data Governance Program*. Academic Press.
- Lahrman, G., Marx, F., Mettler, T., Winter, R., and Wortmann, F. 2011. "Inductive Design of Maturity Models: Applying the Rasch Algorithm for Design Science Research," in *The 6th International conference on design science research in information systems*, Milwaukee, Wisconsin, USA: Springer, pp. 176-191.
- Lämmerhirt, D. 2017. "Avoiding Data Use Silos. How Governments Can Simplify The Licensing Landscape." Available at SSRN: <https://ssrn.com/abstract=3320472> or <http://dx.doi.org/10.2139/ssrn.3320472>.
- Langley, P., and Leyshon, A. 2017. "Platform Capitalism: The Intermediation and Capitalization of Digital Economic Circulation," *Finance and Society* (3:1), pp. 11-31.
- Le Parlement Suisse. 2021. "Réutilisation des Données Personnelles et Techniques Collectées et Gérées par les Administrations Publiques et les Institutions qui en sont Proches." Retrieved March 30, 2021, from <https://www.parlament.ch/fr/ratsbetrieb/suche-curia-vista/geschaefte?AffairId=20204547>
- Lee, G., and Kwak, Y.H. 2012. "An Open Government Maturity Model for Social Media-based Public Engagement," *Government Information Quarterly* (29:4), pp. 492-503.
- Linacre, B., Beck, M. 2002. "What do Infit and Outfit, Mean-square and Standardized mean?" Retrieved March 1, 2022, from <https://www.rasch.org/rmt/rmt162f.htm>
- Lincoln, Y.S., Lynham, S.A., and Guba, E.G. 2011. "Paradigmatic Controversies, Contradictions, and Emerging Confluences, Revisited," *The Sage Handbook of Qualitative Research* (4:2), pp. 97-128.

- Lis, D., and Otto, B. 2020. "Data Governance in Data Ecosystems—Insights from Organizations," in *The Proceedings of the 26th Americas Conference on Information Systems*, Salt Lake City, USA, pp. 1-10.
- Luna-Reyes, L.F., Bertot, J.C., and Mellouli, S. 2014. "Open Government, Open Data and Digital Government," *Government Information Quarterly* (1:31), pp. 4-5.
- Luthfi, A. 2021. "Decision-making Support for Opening Government Data.." TU Delft University.
- Lv, H., and Ma, H. 2019. "Performance Assessment and Major Trends in Open Government Data Research Based on Web of Science Data," *Data Technologies and Applications* (53:3), pp. 286-303.
- Máchová, R., Hub, M., and Lnenicka, M. 2018. "Usability Evaluation of Open Data Portals: Evaluating Data Discoverability, Accessibility, and Reusability from a Stakeholders' Perspective," *Aslib Journal of Information Management* (70:3), pp. 252-268.
- Máchová, R., and Lnenicka, M. 2017. "Evaluating the Quality of Open Data Portals on the National Level," *Journal of Theoretical and Applied Electronic Commerce Research* (12:1), pp. 21-41.
- Magalhaes, G., and Roseira, C. 2017. "Open Government Data and the Private Sector: An Empirical View on Business Models and Value Creation," *Government Information Quarterly*, p. 101248.
- Majchrzak, A., and Markus, M.L. 2012. "Technology Affordances and Constraints in Management Information Systems," in *Encyclopedia of Management Theory* E. Kessler (ed.). Los Angeles: Sage Publications.
- Manyika, J., Chui, M., Groves, P., Farrell, D., Van Kuiken, S., and Doshi, E.A. 2013. *Open Data: Unlocking Innovation and Performance with Liquid Information*. San Francisco: McKinsey Global Institute.
- Mao, Z., Wu, J., Qiao, Y., and Yao, H. 2021. "Government Data Governance Framework Based on a Data Middle Platform," *Aslib Journal of Information Management* (74:2), pp. 289-310.
- Marchildon, P., Bourdeau, S., Hadaya, P., and Labissière, A. 2018. "Data Governance Maturity Assessment Tool: A Design Science Approach," *Projectics/Proyèctica/Projectique*:2), pp. 155-193.
- Marmier, A. 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*, pp. 235-243.
- Marmier, A., and Mettler, T. 2019a. "Challenging the Robustness of OGD De-identification Rules Through a Hackathon," in *the 18th IFIP WG 8.5 International Conference EGOV*, San Benedetto Del Tronto, Italy, pp. 81-89.
- Marmier, A., and Mettler, T. 2019b. "Proposition pour la Publication des Données Ouvertes Publiques: Working Paper de l'IDHEAP," IDHEAP.
- Marmier, A., and Mettler, T. 2020a. "Developing an Index for Measuring OGD Publisher Compliance to Good Practice Standards: Insights from Opendata. swiss," *Information Polity*:Preprint), pp. 1-20.
- Marmier, A., and Mettler, T. 2020b. "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*. Hyderabad, India.
- Martin, C. 2014. "Barriers to the Open Government Data Agenda: Taking a Multi-level Perspective," *Policy & Internet* (6:3), pp. 217-240.
- Martin, S., Foulonneau, M., Turki, S., and Ihadjadene, M. 2013. "Open data: Barriers, Risks and Opportunities," in *Proceedings of the 13th European Conference on eGovernment*, Insubria Varese, Italy, pp. 301-309.
- Marx, F., Wortmann, F., and Mayer, J.H. 2012. "A Maturity Model for Management Control Systems," *Business & Information Systems Engineering* (4:4), pp. 193-207.
- Matos, U.C., and Corbett, J. 2019. "Creating Knowledge for Value Creation in Open Government Data Ecosystems," in *the 25th Americas Conference on Information Systems*, Cancún, Mexico, pp. 1-5.
- Maurino, A., Spahiu, B., Batini, C., and Viscusi, G. 2014. "Compliance with open government data policies: an empirical evaluation of italian local public administrations," in *Twenty Second European Conference on Information Systems*, Tel Aviv, Israel.
- Mayring, P. 2004. "Qualitative Content Analysis," *A Companion to Qualitative Research* (1:2), pp. 159-176.
- McBride, K., Aavik, G., Toots, M., Kalvet, T., and Krimmer, R. 2019. "How Does Open Government Data Driven Co-creation Occur? Six Factors and a 'Perfect Storm'; Insights from Chicago's Food Inspection Forecasting Model," *Government Information Quarterly* (36:1), pp. 88-97.
- McDermott, P. 2010. "Building Open Government," *Government Information Quarterly* (27:4), pp. 401-413.

- McNutt, J.G., Justice, J.B., Melitski, J.M., Ahn, M.J., Siddiqui, S.R., Carter, D.T., and Kline, A.D. 2016. "The Diffusion of Civic Technology and Open Government in the United States," *Information Polity* (21:2), pp. 153-170.
- Meadows, D.H. 2008. *Thinking in Systems: A Primer*. Chelsea Green Publishing.
- Meier, R., Ben, E.R., and Schuppan, T. 2013. "ICT-enabled Public Sector Organisational Transformation: Factors Constituting Resistance to Change," *Information Polity* (18:4), pp. 315-329.
- Meijer, A. 2015. "Government Transparency in Historical Perspective: From the Ancient Regime to Open Data in the Netherlands," *International Journal of Public Administration* (38:3), pp. 189-199.
- Mertens, D.M. 2010. "Philosophy in Mixed Methods Teaching: The Transformative Paradigm as Illustration," *International Journal of Multiple Research Approaches* (4:1), pp. 9-18.
- Mervis, J. 2020. "Can the Census Bureau Actually Meet Trump's Demand to Identify Noncitizens?" Retrieved August 23, 2020, from <https://www.sciencemag.org/news/2020/07/can-census-bureau-actually-meet-trump-s-demand-count-noncitizens>
- Mettler, T., Sprenger, M., and Winter, R. 2017. "Service Robots in Hospitals: New Perspectives on Niche Evolution and Technology Affordances," *European Journal of Information Systems* (26:5), pp. 451-468.
- Mikalef, P., and Krogstie, J. 2018. "Big Data Governance and Dynamic Capabilities: The Moderating Effect of Environmental Uncertainty," in *Proceedings of the Twenty-Second Pacific Asia Conference on Information Systems*, Yokohama, Japan, p. 206.
- MIT Technology Review Custom. 2016. "The Rise of Data Capital," in: *MIT Technology Review Custom*.
- Morabito, V. 2015. "Big Data and Analytics: Strategic and Organisational Impacts," *Springer International Publishing*.
- Moro Visconti, R., Larocca, A., and Marconi, M. 2017. "Big Data-driven Value Chains and Digital Platforms: From Value Co-creation to Monetization," in *Bigdata Analytics: Tools, Technology for Effective Planning*, G. Somani A.k. And Deka (ed.). Boca Raton, USA: CRC Press - Taylor & Francis Group, pp. 355- 371.
- Mosley, M., Brackett, M.H., Earley, S., and Henderson, D. 2010. *DAMA Guide to the Data Management Body of Knowledge*. Technics Publications.
- Munné, R. 2016. "Big Data in the Public Sector," in *New Horizons for a Data-driven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe*, J. Cavanillas, E. Curry and W. Wahlster (eds.). Cham, Switzerland: Springer, pp. 195-208.
- Najjar, M.S., and Kettinger, W.J. 2013. "Data Monetization: Lessons from a Retailer's Journey," *MIS Quarterly Executive* (12:4), pp. 212-245.
- Ndamase, Z. 2014. "The Impact of Data Governance on Corporate Performance: The Case of a Petroleum Company." University of Cape Town, South Africa.
- Neumaier, S., Umbrich, J., and Polleres, A. 2016. "Automated quality assessment of metadata across open data portals," *Journal of Data Information Quality* (8:1), p. 2.
- Newcomer, K. 2016. "Evaluation Research: Methods for Studying Programs and Policies," in *The Oxford Handbook of Classics in Public Policy and Administration*. OUP Oxford, pp. 326-341.
- Nguyen, L. 2009. "Research Methodology," in *Chandos Asian Studies Series, Guerilla Capitalism*, L. Nguyen (ed.). Woodhead Publishing Limited, pp. 29-39.
- Nieborg, D.B., and Poell, T. 2018. "The Platformization of Cultural Production: Theorizing the Contingent Cultural Commodity," *New Media & Society* (20:11), pp. 4275-4292.
- Nielsen, O.B. 2017. "A Comprehensive Review of Data Governance Literature," *Selected Papers of the IRIS* (8), pp. 120-133.
- Nightingale, A.J. 2020. "Triangulation," in *International Encyclopedia of Human Geography (Second Edition)*, A. Kobayashi (ed.). Elsevier, pp. 477-480.
- Nikiforova, A., and McBride, K. 2021. "Open Government Data Portal Usability: A User-Centred Usability Analysis of 41 Open Government Data Portals," *Telematics and Informatics* (58:101539), pp. 1-13.
- Noor, K.B.M. 2008. "Case Study: A Strategic Research Methodology," *American Journal of Applied Sciences* (5:11), pp. 1602-1604.
- OECD. 2019a. "Data Governance in the Public Sector," in *The Path to Becoming a Data-driven Public Sector*. OECD Publishing, Paris: pp. 23-57.
- OECD. 2019b. *Digital Government Review of Argentina*. OECD Publishing, Paris: OECD Digital Government Studies.

- Ondrus, J., Gannamaneni, A., and Lyytinen, K. 2015. "The Impact of Openness on the Market Potential of Multi-sided Platforms: A Case Study of Mobile Payment Platforms," *Journal of Information Technology* (30:3), pp. 260-275.
- Open Data Index. 2016. "Benchmarking Data Automatically." Retrieved November 6, 2018, from <http://oldsite.theodi.org/guides/benchmarking-data-automatically>
- Open Data Monitor. 2018. "The Open Data Monitor." Retrieved November, 6, 2018, from <https://opendatamonitor.eu/frontend/web/index.php?r=dashboard>
- Open Knowledge Foundation. 2018a. "Comprehensive Knowledge Archive Network." Retrieved August 14, 2018, from <https://github.com/ckan/ckan>
- Open Knowledge Foundation. 2018b. "The Open Definition." Retrieved November, 6, 2018, from <http://opendefinition.org/>
- Open Knowledge Foundation. 2018c. "Tracking the State of Open Government Data." Retrieved August 14,, 2018, from <https://index.okfn.org/place/>
- Oppenheim, C., Stenson, J., and Wilson, R.M. 2003. "Studies on Information as an Asset II: Repertory Grid," *Journal of Information Science* (29:5), pp. 419-432.
- Osorio-Sanabria, M.A., Amaya-Fernández, F., and González-Zabala, M.P. 2020. "Developing a Model to Readiness Assessment of Open Government Data in Public Institutions in Colombia," in *Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance*, Athens, Greece, pp. 334-340.
- Otto, B. 2011a. "A Morphology of the Organisation of Data Governance," in *the 19th European Conference on Information Systems*, Helsinki, Finland.
- Otto, B. 2011b. "Organizing Data Governance : Findings from the Telecommunications Industry and Consequences for Large Service Providers " *Communications of the Association for Information Systems* (29:1), pp. 45-66.
- Panian, Z. 2010. "Some Practical Experiences in Data Governance," *World Academy of Science, Engineering and Technology* (62:1), pp. 939-946.
- Paré, G., Trudel, M.-C., Jaana, M., and Kitsiou, S. 2015. "Synthesizing Information Systems Knowledge: A Typology of Literature Reviews," *Information Management* (52:2), pp. 183-199.
- Parker, G.G., Van Alstyne, M.W., and Choudary, S.P. 2016. *Platform Revolution: How Networked Markets Are Transforming the Economy? and How to Make Them Work for You*. New York, NY: Norton and Co.
- Parsons, M.A., Godøy, Ø., LeDrew, E., de Bruin, T.F., Danis, B., Tomlinson, S., and Carlson, D. 2011. "A Conceptual Framework for Managing Very Diverse Data for Complex, Interdisciplinary Science," *Journal of Information Science* (37:6), pp. 555-569.
- Parvex, M. 2013. "«Big Brother, ce n'est pas pour demain»." Retrieved August 23, 2020, from <https://www.letemps.ch/suisse/big-brother-nest-demain>
- Parycek, P., Hochtl, J., and Ginner, M. 2014. "Open Government Data Implementation Evaluation," *Journal of Theoretical and Applied Electronic Commerce Research* (9:2), pp. 80-99.
- Paspatis, I., Tsohou, A., and Kokolakis, S. 2017. "Mobile Application Privacy Risks: Viber Users' De-Anonymization Using Public Data," in *Proceeding of Mediterranean Conference on Information Systems*, Genoa, Italy, pp. 1-14.
- Pasquale, F. 2016. "Two Narratives of Platform Capitalism," *Yale L. & Pol'y Rev* (35), p. 309.
- Peled, A. 2011. "When Transparency and Collaboration Collide: The USA Open Data Program," *Journal of the American Society for Information Science and Technology* (62:11), pp. 2085-2094.
- Permana, R.I., and Suroso, J.S. 2018. "Data Governance Maturity Assessment at PT. XYZ. Case Study: Data Management Division," in *The 2018 International Conference on Information Management and Technology*, Jakarta, Indonesia: IEEE, pp. 15-20.
- Piovesan, F. 2017. "Beyond Standards and Regulations: Obstacles to Local Open Government Data Initiatives in Italy and France," in *The Social Dynamics of Open Data*, S.G.a.Y. Verhulst, A (ed.). Cape Town, South Africa: African Minds, pp. 35-62.
- Pollock, R. 2011. "Building the (Open) Data Ecosystem." Retrieved August 14, 2018, from <https://blog.okfn.org/2011/03/31/building-the-open-data-ecosystem/>
- Pozzi, G., Pigni, F., and Vitari, C. 2014. "Affordance Theory in the IS Discipline: A Review and Synthesis of the Literature," in *Proceedings of the 20th Americas Conference on Information Systems*, Savannah, Georgia, USA, pp. 1-14.



- Raber, D., Wortmann, F., and Winter, R. 2013. "Situational Business Intelligence Maturity Models: An Exploratory Analysis," in *2013 46th Hawaii International Conference on System Sciences*, Wailea, Maui, HI USA: IEEE, pp. 3797-3806.
- Raman, R., Chadee, D., Roxas, B., and Michailova, S. 2013. "Effects of Partnership Quality, Talent Management, and Global Mindset on Performance of Offshore IT Service Providers in India," *Journal of International Management* (19:4), pp. 333-346.
- Randhawa, T.S. 2019. "Incorporating Data Governance Frameworks in the Financial Industry," in: *Walden University Walden Dissertations and Doctoral Studies*.
- Rasouli, M.R., Trienekens, J. J., Kusters, R. J., & Grefen, P. W. 2016. "Information Governance Requirements in Dynamic Business Networking," *Industrial Management & Data Systems* (116:7), pp. 1356-1379.
- Recker, J. 2012. *Scientific Research in Information Systems: A Beginner's Guide*. Berlin: Springer Science & Business Media.
- Redman, T.C., and Blanton, A. 1997. *Data Quality for the Information Age*. Artech House, Inc.
- Reggi, L. 2011. "Benchmarking Open Data Availability Across Europe: The Case of EU Structural Funds," *European Journal of ePractice* (12), pp. 17-31.
- Reiche, K.J., and Höfig, E. 2013. "Implementation of Metadata Quality Metrics and Application on Public Government Data," in *Computer Software and Applications Conference Workshops (COMPSACW), 2013 IEEE 37th Annual*, Kyoto, Japan: IEEE, pp. 236-241.
- Reiche, K.J., Höfig, E., and Schieferdecker, I. 2014. "Assessment and Visualization of Metadata Quality for Open Government Data," in *Conference for E-Democracy and Open Governement*, Krems, Austria, p. 335.
- Rivera, S., Loarte, N., Raymundo, C., and Domínguez-Mateos, F. 2017. "Data Governance Maturity Model for Micro Financial Organizations in Peru," in *Proceedings of the 19th International Conference on Enterprise Information System (ICEIS)*, Prague, Czech Republic, pp. 203-214.
- Robinson, D.G., Yu, H., Zeller, W.P., and Felten, E.W. 2009. "Government Data and the Invisible Hand," *Yale Journal of Law & Technology* (11), pp. 160-175.
- Robson, C. 2002. *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*. Oxford, UK: Wiley-Blackwell.
- Rohlf, J.H. 2003. *Bandwagon Effects in High-technology Industries*. MIT press.
- Roussel, A. 2021. "Notre si Précieuse Intégrité Numérique," in: *L'invité de La Matinale*, A. Roussel (ed.).
- Ruijter, E., Grimmelikhuijsen, S., and Meijer, A. 2017. "Open Data for Democracy: Developing a Theoretical Framework for Open Data Use," *Government Information Quarterly* (34:1), pp. 45-52.
- Russom, P. 2011. "Big Data Analytics," *TDWI Best Practices Report, fourth quarter* (19:4), pp. 1-34.
- Rusu, L., and Jonathan, G.M. 2017. "IT Alignment in Public Organizations: A Systematic Literature Review," in *Information Technology Governance in Public Organizations*. Springer, Cham, pp. 27-57.
- S. Ganapati, C.G.R. 2018. "Prospects and Challenges of Sharing Economy for the Public Sector," *Government Information Quarterly* (35:1), pp. 77-87.
- S. Paquette, P.T.J., S.C. Wilson. 2012. "Identifying the Security Risks Associated with Governmental Use of Cloud Computing," *Government Information Quarterly* (27:3), pp. 245-253.
- Sadiq, S., Yeganeh, N.K., and Indulska, M. 2011. "20 Years of Data Quality Research: Themes, Trends and Synergies," in *the Twenty-Second Australian Database Conference* Perth, Australia, pp. 153-162.
- Safarov, I. 2019. "Institutional Dimensions of Open Government Data Implementation: Evidence from the Netherlands, Sweden, and the UK," *Public Performance Management Review* (42:2), pp. 305-328.
- Safarov, I., Meijer, A., and Grimmelikhuijsen, S. 2017. "Utilization of Open Government Data: A Systematic Literature Review of Types, Conditions, Effects and Users," *Information Polity* (22:1), pp. 1-24.
- Salmasi, A.V., and Gillam, L. 2010. "DNA Dataveillance: Protecting the Innocent?," *Journal of Information, Communication and Ethics in Society* (8:3), p. 270.
- Salminen, J. 2014. "Startup Dilemmas-Strategic Problems of Early-stage Platforms on the Internet," in: *Turku School of Economics*. Turku: Turku.
- Sandoval-Almazán, R., Luna-Reyes, L., Dolores, E., Luna-Reyes, D., Gil-Garcia, J., and Puron-Cid, G. 2017. *Building Digital Government Strategies*. San Antonio, USA.: Springer International Publishing
- Saunders, M., Lewis, P., and Thornhill, A. 2007. *Research Methods for Business Students*. England: Pearson Education Limited.

- Sauser, B.J., Reilly, R.R., and Shenhar, A.J. 2009. "Why Projects Fail? How Contingency Theory Can Provide New Insights: A Comparative Analysis of NASA's Mars Climate Orbiter Loss," *International Journal of Project Management* (27:7), pp. 665-679.
- Sayogo, D.S., Pardo, T.A., and Cook, M. 2014. "A Framework for Benchmarking Open Government Data Efforts," in *2014 47th Hawaii International Conference on System Sciences*, Wailea, Maui, HI USA: IEEE, pp. 1896-1905.
- Schwenk, C.R. 1988. "The Cognitive Perspective on Strategic Decision Making," *Journal of Management Studies* (25:1), pp. 41-55.
- Sebastian-Coleman, L. 2022. *Meeting the Challenges of Data Quality Management*. Elsevier.
- Seydtaghia, A., and Farine, M. 2021. "La Suisse, Véritable Cancre Numérique: Sept Exemples Marquants." Retrieved March 30, 2021, from <https://www.letemps.ch/economie/suisse-veritable-cancre-numerique-sept-exemples-marquants>
- Siau, K., Tan, X., and Sheng, H. 2010. "Important Characteristics of Software Development Team Members: An Empirical Investigation Using Repertory Grid," *Information Systems Journal* (20:6), pp. 563-580.
- Siddiq, A., Hashem, I.A.T., Yaqoob, I., Marjani, M., Shamshirband, S., Gani, A., and Nasaruddin, F. 2016. "A Survey of Big Data Management: Taxonomy and State-of-the-art," *Journal of Network and Computer Applications* (71), pp. 151-166.
- Sidi, F., Panahy, P.H.S., Affendey, L.S., Jabar, M.A., Ibrahim, H., and Mustapha, A. 2012a. "Data quality: A survey of data quality dimensions," in *Information Retrieval & Knowledge Management (CAMP), 2012 International Conference: IEEE*, pp. 300-304.
- Sidi, F., Panahy, P.H.S., Affendey, L.S., Jabar, M.A., Ibrahim, H., and Mustapha, A. 2012b. "Data Quality: A Survey of Data Quality Dimensions," in *The International Conference on Information Retrieval and Knowledge Management, CAMP'12*, Kuala Lumpur, Malaysia: IEEE, pp. 300-304.
- Smelser, N.J., and Baltes, P.B. 2001. *International Encyclopedia of the Social & Behavioral Sciences*. Amsterdam, The Netherland: Elsevier
- Smith, G., and Sandberg, J. 2018. "Barriers to Innovating with Open Government Data: Exploring Experiences Across Service Phases and User Types," *Information Polity* (23:3), pp. 249-265.
- Soares, S. 2010. *The IBM Data Governance Unified Process: Driving Business Value with IBM Software and Best Practices*. MC Press, LLC.
- Solar, M., Concha, G., and Meijueiro, L. 2012. "A Model to Assess Open Government Data in Public Agencies," in *International Conference on Electronic Government*, Granada, Spain: Springer, pp. 210-221.
- Spruit, M., and Pietzka, K. 2015. "MD3M: The Master Data Management Maturity Model," *Computers in Human Behavior* (51), pp. 1068-1076.
- Spy Pond Partners, L., and Iteris, I. 2015. *Data to Support Transportation Agency Business Needs: A Self-assessment Guide*. Transportation Research Board.
- Srimuang, C., Cooharajanone, N., Tanlamai, U., and Chandrachai, A. 2017. "Open Government Data Assessment Model: An Indicator Development in Thailand," in *The 19th International Conference on Advanced Communication Technology (ICACT)*, PyeongChang, Korea: IEEE, pp. 341-347.
- Srinivasan, R., and Lakshmipathy, S. 2017. "Where Do My First Users Come From? Network Mobilization Strategies for Multi-sided Platforms."
- Srnicek, N. 2017. *Platform Capitalism*. John Wiley & Sons.
- Stanford University's Data Governance Office. 2013. "Data Governance Maturity Model."
- Steenbergen, M.v., Bos, R., Brinkkemper, S., Weerd, I.v.d., and Bekkers, W. 2010. "The Design of Focus Area Maturity Models," in *International Conference on Design Science Research in Information Systems*, St.Gallen, Switzerland: Springer, pp. 317-332.
- Strauss, A., and Corbin, J. 2014. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* Sage publications.
- Strong, D.M., Lee, Y.W., and Wang, R.Y. 1997. "Data Quality in Context," *Communications of the ACM* (40:5), pp. 103-110.
- Sunlight Foundation. 2018. "Ten Principles for Opening Up Government Information." Retrieved August 14, 2018, from <https://sunlightfoundation.com/policy/documents/ten-open-data-principles/>
- Susha, I., Grönlund, Å., and Janssen, M. 2015a. "Driving Factors of Service Innovation Using Open Government Data: An Exploratory Study of Entrepreneurs in Two Countries," *Information Polity* (20:1), pp. 19-34.

- Susha, I., Grönlund, Å., and Janssen, M. 2015b. "Organizational Measures to Stimulate User Engagement with Open Data," *Transforming Government: People, Process and Policy* (9), pp. 181-206.
- Swiss Confederation. 2018. "E-gouvernement Strategy." Retrieved August 14, 2018, from [https://www.egovernment.ch/index.php/download\\_file/force/440/3614/](https://www.egovernment.ch/index.php/download_file/force/440/3614/)
- Swiss Federal Audit Office. 2018. "Implementation of the Open Government Data Strategy for Switzerland at the Confederation." Berne: pp. 2-33.
- Swiss Federal Audit Office. 2019. "Implementation of the Open Government Data Strategy for Switzerland at the Confederation: A Cross-section Audit."
- Swissinfo. 2020. "Why Switzerland Struggles to Keep Track of Coronavirus Cases." Retrieved August 23, 2020, from [https://www.swissinfo.ch/eng/crunching-the-numbers\\_why-switzerland-struggles-to-keep-track-of-coronavirus-cases/45628604](https://www.swissinfo.ch/eng/crunching-the-numbers_why-switzerland-struggles-to-keep-track-of-coronavirus-cases/45628604)
- Syed-Ikhsan, S.O.S., and Rowland, F. 2004. "Knowledge Management in a Public Organization: A Study on the Relationship Between Organizational Elements and the Performance of Knowledge Transfer," *Journal of Knowledge Management* (8:2), pp. 95-111.
- Tallon, P.P. 2013. "Corporate Governance of Big Data: Perspectives on Value, Risk, and Cost," *Computer* (46:6), pp. 32-38.
- Tallon, P.P., Ramirez, R.V., and Short, J.E. 2013. "The Information Artifact in IT Governance: Toward a Theory of Information Governance," *Journal of Management Information Systems* (30:3), pp. 141-178.
- Talukder, M.S., Shen, L., Hossain Talukder, M.F., and Bao, Y. 2019. "Determinants of User Acceptance and Use of Open Government Data (OGD): An Empirical Investigation in Bangladesh," *Technology in Society* (56), pp. 147-156.
- Tan, F.B., and Gallupe, R.B. 2006. "Aligning Business and Information Systems Thinking: A Cognitive Approach," *IEEE Transactions on Engineering Management* (53:2), pp. 223-237.
- Tan, F.B., and Hunter, M.G. 2002. "The Repertory Grid Technique: A Method for the Study of Cognition in Information Systems," *MIS quarterly* (26:1), pp. 39-57.
- Tang, R., and Jiang, J. 2021. "Characteristics of Open Government Data (OGD) Around the World: A Country-Sased Comparative Meta-Analysis," *Data and Information Management* (5:1), pp. 11-26.
- Tashakkori, A., and Teddlie, C. 2003. "Major Issues and Controversies in the Use of Mixed Methods in the Social and Behavioral Sciences," *Handbook of Mixed Methods in Social and Behavioral Research* (1:1), pp. 13-50.
- Tauberer, J. 2014. *Open Government Data: The book*. Washington, USA: Joshua Tauberer.
- Taylor, L. 2017. "What is Data Justice? The Case for Connecting Digital Rights and Freedoms Globally," *Big Data & Society* (4:2), pp. 1-14.
- Temps, L. 2020. "L'OFSP va Intensifier la Saisie Numérique des Données Liées au Coronavirus." Retrieved August 23, 2020, from <https://www.letemps.ch/suisse/lofsp-va-intensifier-saisie-numerique-donnees-liees-coronavirus>
- The European Commission. 2018. "Open Data portals." Retrieved June, 12 2019, 2019, from <https://ec.europa.eu/digital-single-market/en/open-data-portals>
- The European Parliament, T.C.O.T.E.U. 2022. "Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending." Retrieved October 1, 2022, from <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R0868>
- Thomas, L.D., Autio, E., and Gann, D.M. 2014. "Architectural Leverage: Putting Platforms in Context," *Academy of Management Perspectives* (28:2), pp. 198-219.
- Thomas, M.A., Cipolla, J., Lambert, B., and Carter, L. 2019. "Data Management Maturity Assessment of Public Sector Agencies," *Government Information Quarterly* (36:4), p. 101401.
- Thompson, N., Ravindran, R., and Nicosia, S. 2015. "Government Data does Not Mean Data Governance: Lessons Learned From A Public Sector Application Audit," *Government Information Quarterly* (32:3), pp. 316-322.
- Tiwana, A. 2013. *Platform Ecosystems: Aligning Architecture, Governance, and Strategy*. Newnes.
- Tiwana, A., Konsynski, B., and Venkatraman, N. 2013. "Information Technology and Organizational Governance: The IT Governance Cube," *Journal of Management Information Systems* (30:3), pp. 7-12.
- Tonya Riley, T.W.P. 2021. "The Cybersecurity 202: NSA Director Says Intelligence Has a Big Blind Spot: Domestic Internet Activity." Retrieved March 26, 2021, from <https://www.washingtonpost.com/politics/2021/03/26/cybersecurity-202-nsa-director-says-intelligence-has-big-blind-spot-domestic-internet-activity/>



- Torchiano, M., Vetrò, A., and Iuliano, F. 2017. "Preserving the Benefits of Open Government Data by Measuring and Improving their Quality: an Empirical Study," in *The 41st Annual Computer Software and Applications Conference*, Turin, Italy: IEEE, pp. 144-153.
- Trom, L., and Cronje, J. 2019. "Analysis of Data Governance Implications on Big Data," in *the Future of Information and Communication Conference*, San Francisco, USA, pp. 645-654.
- Ubaldi, B. 2013. "Open government data: Towards empirical analysis of open government data initiatives," *OECD Working Papers on Public Governance*:22), pp. 0-1.
- Umbrich, J., Neumaier, S., and Polleres, A. 2015. "Quality Assessment and Evolution of Open Data Portals," in *the 3rd International Conference on Future Internet of Things and Cloud*, Rome, Italy: IEEE, pp. 404-411.
- Vaismoradi, M., Turunen, H., and Bondas, T. 2013. "Content Analysis and Thematic Analysis: Implications for Conducting a Qualitative Descriptive Study," *Nursing & Health Sciences* (15:3), pp. 398-405.
- Van Schalkwyk, F., Willmers, M., and McNaughton, M. 2016. "Viscous Open Data: The Roles of Intermediaries in an Open Data Ecosystem," *Information Technology for Development* (22:sup1), pp. 68-83.
- van Veenstra, A.F., and van den Broek, T.A. 2013. "Opening Moves. Drivers, Enablers and Barriers of Open Data in Semi-Public Organization," in: *Electronic Government Conference*. Koblenz, Germany.
- Vancauwenberghe, G., and van Loenen, B. 2019. "Governing Open Spatial Data Infrastructures: The Case of the United Kingdom," in *Governance Models for Creating Public Value in Open Data Initiatives*. Cham, Switzerland: Springer, pp. 33-54.
- Varajão, J. 2018. "The Many Facets of Information Systems (+ Projects) Success," *International Journal of Information Systems and Project Management* (6:4), pp. 5-13.
- Veljković, N., Bogdanović-Dinić, S., and Stoimenov, L. 2014. "Benchmarking Open Government: An Open Data Perspective," *Government Information Quarterly* (31:2), pp. 278-290.
- Venkatesh, V., Brown, S.A., and Bala, H. 2013. "Bridging the Qualitative-Quantitative Divide: Guidelines for Conducting Mixed Methods Research in Information Systems," *MIS quarterly*, pp. 21-54.
- Vetrò, A., Canova, L., Torchiano, M., Minotas, C.O., Iemma, R., and Morando, F. 2016. "Open Data Quality Measurement Framework: Definition and Application to Open Government Data," *Government Information Quarterly* (33:2), pp. 325-337.
- Vickery, G. 2011. *Review of recent studies on PSI re-use and related market developments*. Paris: Information Economics.
- Ville de Lausanne. 2020. "Agenda et Actualité." Retrieved November, 17, 2020, from <https://www.lausanne.ch/>
- Vilminko-Heikkinen, R., & Pekkola, S. 2019. "Changes in Roles, Responsibilities and Ownership in Organizing Master Data Management," *International Journal of Information Management* (47), pp. 76-87.
- Vitale, M.R., Ives, B., and Beath, C.M. 1986. "Linking Information Technology and Corporate Strategy: An Organizational View," in *ICIS 1989 Proceedings*, p. 30.
- Volkoff, O., and Strong, D.M. 2013. "Critical Realism and Affordances: Theorizing IT-associated Organizational Change Processes," *MIS Quarterly* (37:3), pp. 819-834.
- W3C. 2017. "Data on the Web Best Practices." Retrieved November, 6 2018, from <https://www.w3.org/TR/dwbp/>
- W3C. 2019. "Comparing DKAN and CKAN." Retrieved October 7, 2019, from <https://docs.getdkan.com/en/latest/introduction/dkan-ckan.html>
- Wagner, H., Pankratz, O., Basten, D., and Mellis, W. 2015. "Effort of EAI Projects: A Repertory Grid Investigation of Influencing Factors," *Project Management Journal* (46:5), pp. 62-80.
- Wahyudi, A., Kuk, G., and Janssen, M. 2018. "A process pattern model for tackling and improving big data quality," *Information Systems Frontiers* (20:3), pp. 457-469.
- Wajcman, J. 2014. *Pressed for Time*. University of Chicago Press.
- Walton, N. 2017. "Ecosystems Thinking and Modern Platform-Based Ecosystem Theory," in *The Internet as a Technology-Based Ecosystem*. Springer, pp. 85-117.
- Wand, Y., and Wang, R.Y. 1996. "Anchoring data quality dimensions in ontological foundations," *Communications of the ACM* (39:11), pp. 86-95.
- Wang, R.Y., and Strong, D.M. 1996. "Beyond Accuracy: What Data Quality Means to Data Consumers," *Journal of Management Information Systems* (12:4), pp. 5-33.
- Wang, V., and Shepherd, D. 2020. "Exploring the Extent of Openness of Open Government Data: A Critique of Open Government Datasets in the UK," *Government Information Quarterly* (37:1), p. 101405.

- Weber, K., Otto, B., Österle, H. 2009. "One Size Does Not Fit All - A Contingency Approach to Data Governance," *ACM Journal of Data and Information Quality* (1:1), pp. 1–27.
- Weill, P., and Ross, J.W. 2004. *IT Governance: How Top Performers Manage IT Decision Rights for Superior Results*. Harvard Business Press.
- Welch, E.W., Feeney, M.K., and Park, C.H. 2016. "Determinants of Data Sharing in US City Governments," *Government Information Quarterly* (33:3), pp. 393-403.
- Welle Donker, F., and van Loenen, B. 2017. "How to Assess the Success of the Open Data Ecosystem?," *International Journal of Digital Earth* (10:3), pp. 284-306.
- Whitley, E.A., and Hosein, I.R. 2008. "Doing the Politics of Technological Decision Making: Due Process and the Debate about Identity Cards in the UK," *European Journal of Information Systems* (17:6), pp. 668-677.
- Williamson, C. 2013. "Ethnographic Research," in *Research Methods: Information, Systems, and Contexts*. Tilde University Press, pp. 287-309.
- Winter, J.S., & Davidson, E. 2018. "Big Data Governance of Personal Health Information and Challenges to Contextual Integrity," *The Information Society*, pp. 1–16.
- Wirtz, B.W., and Birkmeyer, S. 2015. "Open Government: Origin, Development, and Conceptual Perspectives," *International Journal of Public Administration* (38:5), pp. 381-396.
- Wirtz, B.W., Piehler, R., Thomas, M.-J., and Daiser, P. 2016. "Resistance of Public Personnel to Open Government: A Cognitive Theory View of Implementation Barriers Towards Open Government Data," *Public Management Review* (18:9), pp. 1335-1364.
- Wirtz, B.W., Weyerer, J.C., Becker, M., and Müller, W.M. 2022. "Open Government Data: A Systematic Literature Review of Empirical Research," *Electronic Markets*, 2022/09/20.
- World Wide Web Foundation. 2018. "Open Data Barometer." Retrieved August 14, 2018, from [https://opendatabarometer.org/?\\_year=2016&indicator=ODB](https://opendatabarometer.org/?_year=2016&indicator=ODB)
- Yang, H.C., Lin, C.S., and Yu, P.H. 2015a. "Toward Automatic Assessment of the Categorization Structure of Open Data Portals," in *Proceeding of the 2nd International Conference on Multidisciplinary Social Networks Research*, Matsuyama, Japan: Springer, pp. 372-380.
- Yang, L., Li, J., Elisa, N., Prickett, T., and Chao, F. 2019. "Towards Big Data Governance in Cybersecurity," *Data-Enabled Discovery and Applications* (3:1), pp. 1-12.
- Yang, T.-M., Lo, J., and Shiang, J. 2015b. "To Open or Not to Open? Determinants of Open Government Data," *Journal of Information Science* (41:5), pp. 596-612.
- Yang, T.-M., and Wu, Y.-J. 2016. "Examining the Socio-technical Determinants Influencing Government Agencies' Open Data Publication: A Study in Taiwan," *Government Information Quarterly* (33:3), pp. 378-392.
- Yang, Z., Cai, S., Zhou, Z., and Zhou, N. 2005. "Development and validation of an instrument to measure user perceived service quality of information presenting web portals," *Information & Management* (42:4), pp. 575-589.
- Yiannakoulias, N., Slavik, C.E., Sturrock, S.L., and Darlington, J.C. 2020. "Open Government Data, Uncertainty and Coronavirus: An Infodemiological Case Study," *Social Science & Medicine* (265), p. 113549.
- Yin, R.K. 1994. "Discovering the Future of the Case Study. Method in Evaluation Research," *Evaluation Practice* (15:3), pp. 283-290.
- Young, M., and Yan, A. 2017. "Civic Hackers' User Experiences and Expectations of Seattle's Open Municipal Data Program," in *Proceedings of the 50th Hawaii International Conference on System Sciences*, Hilton Waikoloa Village, Hawaii, USA, pp. 2681-2690.
- Young, M.M. 2020. "Implementation of Digital-era Governance: The Case of Open Data in US Cities," *Public Administration Review* (80:2), pp. 305-315.
- Zeleti, F.A., Ojo, A., and Curry, E. 2016. "Exploring the Economic Value of Open Government Data," *Government Information Quarterly* (33:3), pp. 535-551.
- Zhang, R., Indulska, M., and Sadiq, S. 2019. "Discovering Data Quality Problems," *J Business Information Systems Engineering* (61:5), pp. 575-593.
- Zhao, Y., and Fan, B. 2018. "Exploring Open Government Data Capacity of Government Agency: Based on the Resource-based Theory," *Government Information Quarterly* (35:1), pp. 1-12.
- Zuiderwijk, A. 2015. "Open Data Infrastructures: The Design of an Infrastructure to Enhance the Coordination of Open Data Use." Delft University of Technology.
- Zuiderwijk, A., and Janssen, M. 2014a. "Barriers and Development Directions for The Publication and Usage of Open Data: A Socio-technical View," in *Open government*. Springer, pp. 115-135.

- Zuiderwijk, A., and Janssen, M. 2014b. "Open Data Policies, their Implementation and Impact: A Framework for Comparison," *Government Information Quarterly* (31:1), pp. 17-29.
- Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., and Alibaks, R.S. 2012. "Socio-technical Impediments of Open Data," *Electronic Journal of e-Government* (10:2), pp. 156-172.
- Zuiderwijk, A., Janssen, M., and Davis, C. 2014. "Innovation with Open Data: Essential Elements of Open Data Ecosystems," *Information Polity* (19:1, 2), pp. 17-33.
- Zuiderwijk, A., Janssen, M., and Dwivedi, Y.K. 2015. "Acceptance and Use Predictors of Open Data Technologies: Drawing upon the Unified Theory of Acceptance and Use of Technology," *Government Information Quarterly* (32:4), pp. 429-440.
- Zuiderwijk, A., Janssen, M., van de Kaa, G., and Poulis, K. 2016. "The Wicked Problem of Commercial Value Creation in Open Data Ecosystems: Policy Guidelines for Governments," *Information Polity* (21:3), pp. 223-236.
- Zuiderwijk, A., Pirannejad, A., and Sussha, I. 2021. "Comparing Open Data Benchmarks: Which Metrics and Methodologies Determine Countries' Positions in the Ranking Lists?," *Telematics and Informatics* (62), p. 101634.
- Zuiderwijk, A., Shinde, R., and Janssen, M. 2018. "Investigating the Attainment of Open Government Data Objectives: Is There a Mismatch Between Objectives and Results?," *International Review of Administrative Sciences*).