Balancing Control, Usability and Visibility of Linked Open Government Data to Create Public Value

Abstract. Linked data is a technical standard to structure complex information and relate independent sets of data. Recently, governments have started to use this technology for bridging separated data 'silos' by launching linked open government data (LOGD) portals. The purpose of this conceptual paper is to explore the role of LOGD as a smart technology and strategy to create public value. This is achieved by enhancing the usability and visibility of open data provided by public organizations. In this study, three different LOGD governance modes are deduced: public agencies could release linked data via a dedicated triple store, via a shared triple store or via an open knowledge base. Each of these modes has different effects on usability and visibility of open data. Selected case studies illustrate the actual use of these three governance modes. According to this study, LOGD governance modes present a trade-off between retaining control over governmental data and potentially gaining public value by the increased use of open data by citizens. This study provides recommendations for public sector organizations for the development of their data publishing strategy to balance control, usability and visibility considering also the growing popularity of open knowledge bases such as Wikidata.

Keywords: linked data, open data, linked open government data, data governance, public value, Wikidata.

1   Open Data and Public Value

In recent years, a growing adoption of open government data (OGD) policies can be observed (Wang and Lo, 2016), and many governments including those of the U.S., EU and UK have launched their own OGD portals online. OGD is part of the larger open data movement, which promotes the idea that non-personal and not security related data should be made available to everyone free of charge and restrictions[1]. Partly, this movement is the result of increasing pressure on governments to become more transparent (Welch and Wong, 2001). As government is one of the principal producers and collectors of data in a vast array of domains, government data is often viewed as a valuable resource offering great opportunities to stakeholders if it is openly available. Most notably, several scholars have argued that the creation of public value is one of these opportunities, if not the most important one (Attard et al., 2016; Hui and Hayllar, 2010; Lee and Kwak, 2012; Zuiderwijk and Janssen, 2014) and many practitioners mention

---

[1] Open Data Handbook: http://opendatahandbook.org/guide/en/what-is-open-data

the creation of public value as one of the key goals of open data initiatives (Janssen et al., 2012).

Generally, the promise of OGD is that it has a positive impact on society in four pivotal areas: 1) government transparency and accountability, 2) citizen inclusion and empowerment, 3) government efficiency and effectiveness, and 4) economic growth (Davies, 2013). Although, these effects of OGD have not been tested systematically and empirically (Safarov et al., 2017), there is evidence of OGD impact on the public sphere (de Kool and Bekkers, 2016; Lourenço, 2016). As such, the promises of OGD agree with many definitions of public value. Bryson et al. (2014), for instance, define public value as "producing what is either valued by the public, is good for the public [...], or both, as assessed against various public value criteria". Benington and Moore (2010) emphasize that public value means, "first, what the public values", focusing on "individual interests [...] of current users" and, "second, what adds value to the public sphere", focusing "on the longer term public good, including the needs of generations to come". Note that both of these definitions ascribe considerable interpretative power to the public when it comes to determining the public value of a phenomenon, meaning that how the public perceives the impact of government activities is key to whether they create public value or not. Additionally, to allow OGD to create public value, the data needs to be reliable and valid and it should enable citizens to create something they deem valuable (Harrison et al., 2012).

Importantly, the impact of OGD in the four pivotal areas named above is not an immediate result of making government data public, because OGD does not have any value in itself (Janssen et al., 2012). Instead, making an impact requires external actors to invest time and resources into working with OGD, for instance by data interlinking, visualization, analysis, or interpretation (Attard et al., 2016) and by eventually producing innovation (Lakomaa and Kallberg, 2013) and other forms of impact (Stuermer and Dapp, 2016). Jetzek (2016) mentions sharing and market mechanisms as the two principal ways of generating value by means of open data. To this end, OGD platforms and data formats need to be designed in a way enabling external actors to easily access, combine and use the data provided, even if it originates from multiple different public (and private) organizations.

In addition, concerns about privacy and 'big brother' surveillance threats are relevant in the context of OGD. Janssen and van den Hoven (2015) point out that both transparency and privacy are important values of governments. In their terminology big and open linked data (abbreviated as BOLD) presents new challenges to administrative organizations providing transparency within governmental activities while ensuring the privacy of its citizens.

Multiple analyses of existing OGD portals have illustrated that the use of OGD is often hampered by the multitude of different data formats and the lack of machine readability of the data (Neumaier et al., 2016; Smith and Sandberg, 2018; Umbrich et al., 2015). Additionally, Zuiderwijk, Janssen et al. (2012) have identified ten key impediments of public value generation, focusing on the perspective of the user. According to these authors, availability and access, findability, usability, understandability, quality, linking and combining data, comparability and compatibility, metadata, interaction with the data provider, and opening and uploading are key issues hampering public

value generation by users. So, merely opening to the public the data 'silos', which many public organizations maintain is, in many cases, insufficient to foster the creation of public value. While opening data to the public certainly enhances visibility of the data and thereby government transparency, the data within individual open data 'silos' are often difficult to browse, query and connect (Shadbolt et al., 2012) due to a lack of common unique identifiers of data records and commonly used vocabularies and ontologies. Thus, the use of such open data remains limited, hampering the possibilities of external actors to interact with the data and leverage their public value potential (Heath, 2008). In our view, this last aspect and particularly its connection to linked data, has thus far been widely overlooked in the literature on OGD, constituting a substantial gap in research.

In this study, alternatives are explored regarding how public organizations can improve the way they provide open data to the public, enhancing both visibility and usability and thereby maximizing their potential to create public value. To this end, a user's perspective is adopted, heeding the call by Janssen et al. (2012). The focus lies on how OGD 'silos' can be bridged by the use of linked open government data (LOGD) (Shadbolt et al., 2012) as an overlap of government data, open data and linked data (see Figure 1).
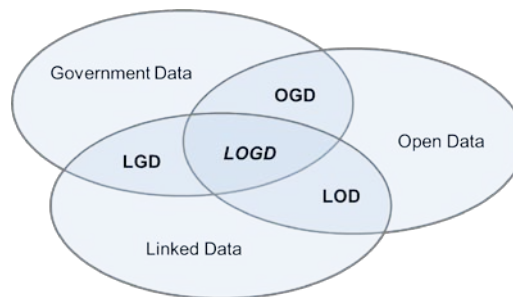

Figure 1. Linked open government data (LOGD) as overlap of government data, open data and linked data.

LOGD is already in use for several years in the public sector around the world (Ding et al., 2012). Starting in the UK the government has integrated LOGD into its OGD portal Data.gov.uk (Kalampokis et al., 2011). The US government is using linked data with their open government platform Data.gov (Bizer, 2009). Within the EU several research projects have resulted in recommendations for ontological frameworks and implementation models (Höchtl and Reichstädter, 2011). Villazón-Terrazas et al. (2011) have developed and validated methodical guidelines for publishing LOGD. Zuiderwijk, Jeffery et al. (2012) emphasize the potential of metadata in a linked data context and its value for users and publishers alike.

Consequently, the research question of this study is: *what options do public organizations have to publish their LOGD in such a way that data control, data visibility and data usability are balanced according to the organizations' requirements?*

We attempt to answer this question by deducing three different LOGD governance modes and their respective impact on control, visibility and usability, and we illustrate their use in selected case studies as suggested by Siggelkow (2007). As such, this study supports public sector organizations in assessing the existing potential, opportunities as well as risks derived from the implementation of LOGD and corresponding strategies as a smart technology to generate public value. Furthermore, this study intends to advise public organizations in handling the challenge of citizen participation in the form of data publishing via open knowledge bases like Wikidata.

This study follows an interdisciplinary approach, drawing on theory and literature from two disciplines: public management and information systems. In section 2, linked data will be introduced and its role as a solution to enhance the usability and visibility of OGD is explored. In section 3, three LOGD governance modes will be derived following a deductive approach along different levels of control, which public organizations are able to exert on their data during the publishing process. These LOGD governance modes will then be substantiated by case studies featuring these modes in section 4. Section 5 concludes the findings of this study.

2 Linked Data

Linked data is a collection of best practices to publish and connect structured data on the Web (Bizer, Heath, et al., 2009). On a technical level, linked data means using Web technologies to create links between data from different sources in a way that the data is machine-readable and has an explicitly defined meaning (Bizer, Heath, et al., 2009). The necessary mechanisms are specified in the Resource Description Framework (RDF),[2] which defines a language for representing information as linked data. Such linked data consists of nodes and directed arcs linking pairs of nodes. A subject node, predicate and object node form a so-called RDF triple. A unique identifier called Uniform Resource Identifier (URI) represents the individual nodes. Triples can be interlinked by using an object as a subject for another triple or by using new objects for existing subjects (see Figure 2). The aggregate of many interlinked triples then constitutes a so-called graph. A database suitable for storing linked data triples is called a triple store. (Bizer, Heath, et al., 2009).

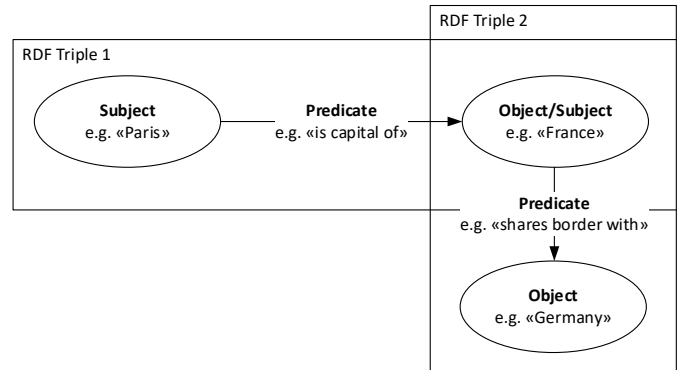[2] RDF Syntax: https://www.w3.org/TR/REC-rdf-syntax/

Figure 1. Two interlinked RDF triples.

Adherence to linked data principles can foster the use of OGD because it makes the aggregation and integration of heterogeneous data from different sources much easier (Heath, 2008; Neumaier et al., 2016). In this vein, Berners-Lee (2006) has presented a five star rating scheme for data encouraging government data owners in particular to use linked data as a possibility to bridge data 'silos'. The criteria for such five star data are: 1) availability on the web, 2) availability as machine-readable structured data, 3) availability in a non-proprietary data format, 4) availability by the use of RDF to identify things and finally 5) based on the technical possibilities resulting from 1) to 4): the inclusion of links to other related data (see Figure 3).
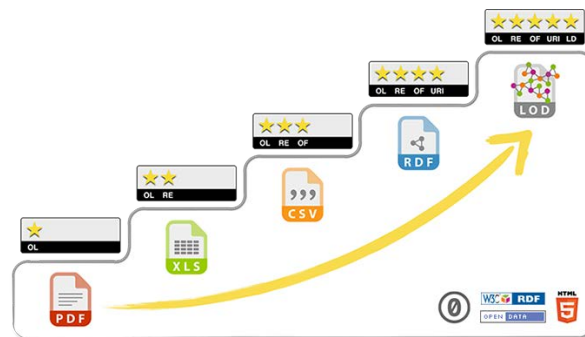

Figure 2. Five star open data rating scheme.[3]

Linked data allows for a complete separation of the data and any specific view upon it thus enabling forms of use that the creators may not have anticipated in advance (Heath, 2008). Moreover, using the established linked data principles, there is no need for the data consumer to learn a multitude of different data access techniques for different data sources (Bizer, Heath, et al., 2009; Hausenblas, 2009). Ultimately, the use of linked

---

[3] Source: http://5stardata.info

data is envisioned to enable everyone to use the Web like a single global database (Bizer, Heath, et al., 2009), called the linked open data cloud (LOD cloud). Figure 4 depicts the current state of this LOD cloud representing more than a thousand datasets from all over the Web.
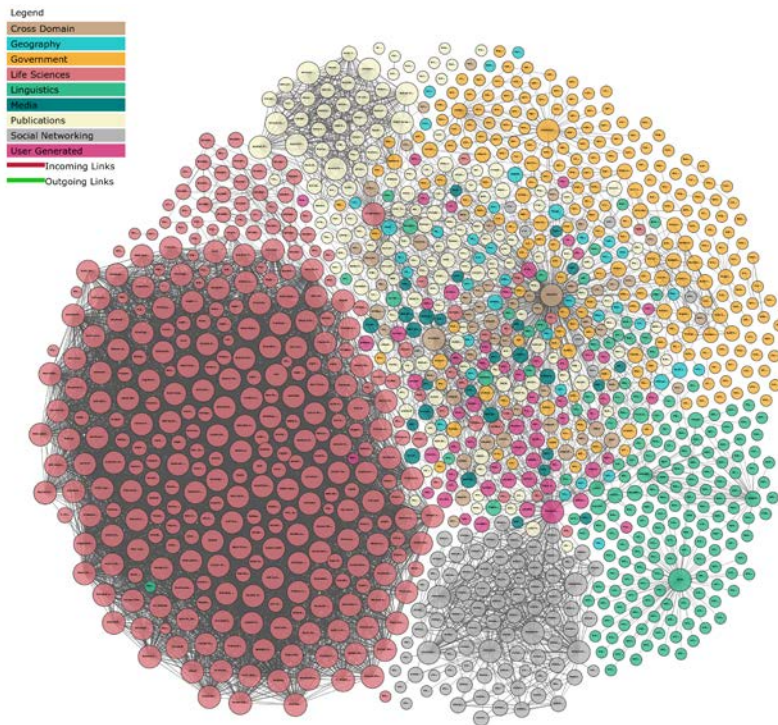

Figure 3: Current state of the LOD cloud (Abele et al., 2017) representing linked datasets.

Besides the mentioned advantages geared mainly towards data users, publishing LOGD also has some advantages for the data publishers: 1) the integration of additional data is cost effective (Bizer, Heath, et al., 2009), 2) there is no need to adhere to a particular scheme (Neumaier et al., 2016), and 3) releasing information as LOGD is cheaper than producing reports (Hendler et al., 2012).

2.1 Data Interlinking for Enhanced Usability and Visibility

The data interlinking functionality of linked data is essential to achieve the goal of five star data (Berners-Lee, 2006). Linking the data is done by the RDF mechanism with the option to choose from two linking directions: outgoing and incoming links. Figure 5 shows the effects of these two types of linking. Outgoing links enhance the *usability* by connecting the data to other sources of information within in the LOD cloud. Incoming

links improve the *visibility* by making the data easier to find for users browsing the LOD cloud.
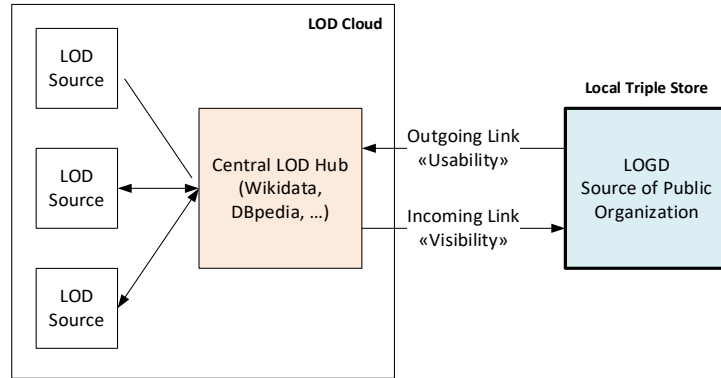

Figure 4. Outgoing and incoming links from the perspective of a local triple store.

Since an outgoing link is set up within the local triple store, creating such links requires no additional write access outside the local triple store. Incoming links however, need to be created in the LOD cloud and accordingly require write access to a triple store in the LOD cloud (see Figure 6). This emphasizes the importance of open knowledge bases acting as central LOD hubs (see Figure 5) in the LOD cloud where incoming links to a local triple store can be set up. These LOD hubs are a central element of the web of data (Bizer, Lehmann, et al., 2009).
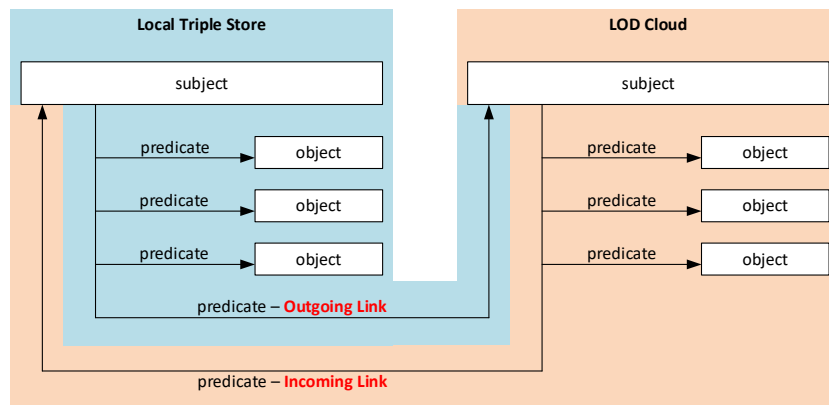

Figure 5. Outgoing and incoming links and their required write access.

3 Three LOGD Governance Modes Based on their Level of Control

Public organizations may face severe consequences if their data turns out to be inaccurate, outdated or incomplete, because they are accountable both legally and politically for the actions they take (Gilmour and Jensen, 1998). Correspondingly, they often prefer to be in full control over what and how the data they release is published (Janssen et al., 2012). As West (2003) shows, keeping control can, however, diminish some of the advantages of open knowledge. Keeping full control can even form some serious barriers to usability and visibility of the published data. Therefore, we believe public organizations should carefully choose the best option to publish their LOGD in view of the trade-off between data visibility, usability and control. In line with Stuermer et al. (2009), Boudreau (2010) and West and Bogers (2014), we consider the level of control to be an important discriminator. Hence, three main options are deduced regarding the degree of keeping control of the LOGD publishing process: a public organization can 1) keep full control over the entire publication process by using a *dedicated triple store*, 2) share it with some other public organizations in the form of a *shared triple store* or 3) share the control completely with the public by using an *open knowledge base* to publish their LOGD. We call these three options LOGD governance modes (see Table 1). Whereas option one and three present the poles of a keeping control scale, option two is a compromise. Different implementations of option two have in common the use of a single collective triple store shared by different public organizations.

Sharing the control with other public organizations has consequences by sharing information between the participating organizations, which is seen as a key strategic activity for public organizations (Yang and Maxwell, 2011). Sharing infrastructure can also lead to improved collaboration (Zuiderwijk and Janssen, 2014). However, public organizations often act independently without considering activities by other public organizations (Gil-García and Pardo, 2005). Nugroho et al. (2015) recommend assigning a designated organization for implementing open data policies and providing a shared infrastructure.

As will be shown, these modes have direct and indirect consequences on the usability and visibility of the data. In our view, these consequences should be viewed as inherent potentials as well as risks of the individual LOGD governance modes. In a particular implementation of a mode, such potentials may be realised or missed, while risks could be mitigated by the application of appropriate strategies.

3.1 Direct Consequences of the LOGD Governance Modes

The higher the level of control the higher are usually the design freedoms for the LOGD publishing process. Such design freedoms include the technical implementation, the representation of the data and the applied open data license.

Freedom to choose the technical implementation of the triple store and the query interface can be of importance for special purpose data, which profits from certain non-standard functionalities, such as sophisticated geospatial data operations. Freedom to

choose the data representation eliminates restrictions concerning the vocabulary used for predicates to describe the data. In such a case, using a self-defined as well as an established standard vocabulary is possible without additional expenditure. In a shared environment or with an open knowledge base, there are usually requirements concerning the vocabulary and accordingly, changes like establishing new predicates are complex or even impossible. Freedom to choose the open data license allows the use of a tailored licence contrary to open knowledge bases, which may define certain restrictions concerning the allowed open data licenses.

The visibility of the LOGD depends crucially on the number of users and the incoming links to the data (see Figure 5). Open knowledge bases acting as data hubs in the LOD cloud have an advantage in this area due to the large contributor and user base as well as the high amount of data and its pre-existing links. Consequently, dedicated triple stores of small and specialised public organizations will probably face difficulties in creating enough attention of users for their data.

Open knowledge bases are geared towards attracting as much data as possible. Therefore they provide tools, instructions and examples for transforming the data into linked data triples and for the associated ontological process (Vrandečić and Krötzsch, 2014). This is not the case for a dedicated or shared triple store. Operating a dedicated triple store comes with a lower cost-effectiveness because of the missing possibility to share the cost among different data providers. In contrast, open knowledge bases usually offer their services at no cost for the data provider. Open knowledge bases are built around the idea of a community supporting the project and taking responsibility for the data published (Vrandečić and Krötzsch, 2014). This is not the case for a dedicated or shared triple store operated by a public organization. Furthermore, Lee and Kwak (2012) show that choosing an open knowledge base for publishing LOGD demonstrates a higher level of engagement with the public, leading to the realization of greater public value.

LOGD governance mode	Dedicated triple store	Shared triple store	Open knowledge base
Control			
Level of control	high	medium	low
Sharing of control	low	medium	high
Design choices			
Freedom to choose technical implementation	high	medium	low
Freedom to choose data representation	high	medium	low
Freedom to choose licence	high	medium	low
Visibility			
Visibility of triple store	low	medium	high
Further characteristics			
Support for data publishing	low	medium	high
Cost effectiveness	low	medium	high
Development of community	low	low	high
Engagement with the public	low	low	high

Table 1. The three different LOGD governance modes and their potential direct consequences.

3.2   LOGD Governance Modes and Data Quality

Because of their influence on data quality, the different LOGD governance modes have some more indirect consequences on data usability and somewhat on data visibility. Wang and Strong (1996) define data quality as data that is "fit for use" by the data consumer thereby establishing a link between data quality and usability. Only if the data is fit for use, it will actually be used creating the precondition for generating public value. Wang and Strong (1996) present different data quality categories and dimensions to evaluate the various aspects of quality. Strong et al. (1997) further emphasize the close relationship between data quality and usability. Although Zaveri et al. (2016) adapt and extend the abovementioned categories and dimensions explicitly to a linked data context, we have adopted the original classification from Wang and Strong (1996) as a more general concept of data quality for our research. We will cover only the dimensions influenced by the different LOGD governance modes. Wang and Strong (1996) have introduced four data quality categories and their associated dimensions: 1)

intrinsic data quality, which is the quality of the data itself including accuracy, objectivity and believability, 2) contextual data quality representing dimensions such as added-value, timeliness and completeness, 3) representational data quality including interpretability, ease of understanding and consistency and finally, 4) accessibility data quality. The influence of the different governance modes on data quality are discussed in the following paragraph and summarized in Table 2.

1) At first sight, intrinsic data quality does not depend directly on the LOGD governance mode. As the name suggests, it seems to be an exclusively intrinsic feature of the data itself. However, if the published LOGD is mixed with data from other sources, either in a shared environment or in an open knowledge base, the collective intrinsic data quality may suffer from bad quality data from other sources, which may not be distinguishable from the original source without additional effort.

2) Regarding contextual data quality, the added value of using open knowledge bases may be their sophisticated tools to analyse and visualize the data. The data timeliness in a dedicated triple store operated by a public organization may be lower than in an open knowledge base with its many contributors, which are furthermore not tied to strict publishing procedures. The completeness of the data, however, is a challenge for open knowledge bases because an individual contributor may possess only incomplete information. As such, completeness in open knowledge bases is a subject of in-depth considerations (Ballatore et al., 2013; Färber et al., 2017). For a dedicated triple store operated by a public organization, completeness of the available data is more easily ensured based on the systematic method of operation of such organizations.

3) Representational data quality includes interpretability and ease of understanding. Due to the restrictions of the vocabulary that may be used in a shared or open environment, these governance modes have an advantage in this regard. Representational consistency however, is positively influenced by a smaller number of contributors, favouring a dedicated solution. Furthermore, this simplifies the creation of a very concise representation of the data because of the possibility to create well-tailored, new predicates instead of having to choose from a limited vocabulary.

4) Accessibility data quality is influenced by the technical realization of the triple store. Open knowledge bases can have an advantage in this regard because they offer their service to a wide variety of users, and, therefore are strongly interested in being as accessible as possible.

LOGD governance mode	Dedicated triple store	Shared triple store	Open knowledge base
Intrinsic data quality			
Accuracy, objectivity, believability	high	high	low
Contextual data quality			
Value-added	low	low	high
Timeliness	low	low	high
Completeness	high	medium	low
Representational data quality			
Interpretability	low	low	high
Ease of understanding	low	low	high
Representational consistency	high	high	low
Concise representation	high	medium	low
Accessibility data quality			
Accessibility	low	low	high

Table 2. The three different LOGD governance modes and their potential effects on data quality and subsequent data usability.

4 Case Studies

The following section provides case studies observed in Switzerland for each of the different LOGD governance modes.

4.1 Method and Case Selection

In accordance with Siggelkow (2007), the following case studies are used to illustrate and make plausible the conceptual considerations leading to the LOGD governance modes described above. Furthermore, the case studies contribute to additional insights to the LOGD phenomenon (Walsham, 1995). The case selection was done not by random selection to achieve representativeness but by choosing important cases illustrating the LOGD governance modes following King et al. (1994). As such, the selected cases represent good practice examples. Regarding the selection of the case of an open knowledge base, Wikidata was chosen as an archetype of such open knowledge bases. This Wikidata case is further supplemented by considerations on how public organizations could react to such open knowledge bases.

4.2 Case Study Dedicated Triple Store

The Swiss Federal Office of Topography operates a federal spatial data infrastructure including a dedicated triple store, which provides datasets as linked data.[4] The available datasets contain mainly spatial information about the country, its cantons, districts and municipalities. This information is version-controlled to enable the reconstruction of the temporal evolution. Furthermore, the data is enriched with some current statistical data such as population numbers. The municipalities in the triple store have additional outgoing links to Wikidata.

Furthermore, Wikidata also links back to the municipalities stored in the dedicated triple store thus creating incoming links in the triple store of the office of topography. This is possible easily because of the open nature of Wikidata where the general public is allowed to create new links. This linking between Wikidata and the dedicated triple store enables the creation of mashups combining data from both sources without effort. Figure 7 demonstrates such a mashup combining geographical boundary data for a Swiss municipality with additional information about the mayor of the municipality and its political party. The latter information originates from Wikidata. This combination of data was achieved by a single SPARQL query[5] to the dedicated triple store and did not involve any data transformation, which would be necessary if the data were not linked.

---

[4] www.ld.geo.admin.ch

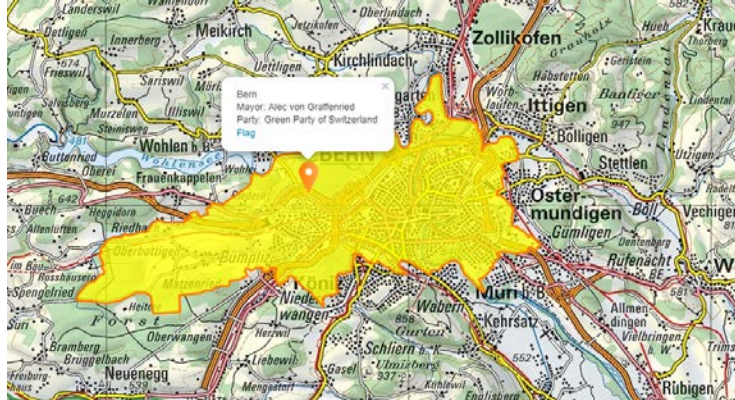[5] http://yasgui.org/short/rJAIpzpw7

Figure 6. Data mashup by using geographical boundary data from a dedicated triple store and additional information from an open knowledge base (Wikidata) by execution of a single query.

4.3 Case Study Shared Triple Store

The Linked Data Service (LINDAS)[6] operated by the Swiss Federal Archives is an example of a shared triple store. It can be used by Swiss public organizations to make their data available as LOGD. So far, it is a prototype to examine the potential of LOGD. The setting up of LINDAS included the creation of a nation-wide consistent URI naming scheme. Today, the data available on LINDAS consists of twenty-one different data sets. Many of them are often only fragments of larger data sets to enable specific use cases to demonstrate the possibilities of LOGD. The total count of LOGD triples in LINDAS is about forty million. An excerpt from the available datasets and its providers is shown in Table 3.

Dataset description	Dataset provider
Air quality in Switzerland	Federal Office for the Environment
Heavy metal legacies in Switzerland	Federal Office for the Environment
Historicised municipality inventory	Swiss Federal Statistical Office
Public transport stops	Federal Office of Transport
Animal transports data	Federal Food Safety and Veterinary Office
Governmental agency directory	State Secretariat for Economic Affairs
Meteorological weather data	Federal Office of Meteorology and Climatology

Table 3: Excerpt of available datasets and their providers in the shared triple store LINDAS.

[6] https://lindas-data.ch

Implemented application scenarios include an animal disease outbreak analysis, an energy management algorithm for railway coaches, a historicised official municipality register, and historical federal budgets.[7] Today, LINDAS does not include any links between the different LINDAS datasets and no outgoing links to Wikidata or other open knowledge bases.

4.4 Case Study Open Knowledge Base

Wikidata[8] is a crowdsourced open knowledge base. It is part of the Wikimedia Foundation and as such a member of a well-established culture of contributions by the general public. It contains more than 41 million items and currently has nearly 7900 active users with more than five edits over a one-month period.[9] Wikidata explicitly does not store facts but so-called statements since sometimes there is no global agreement on the "true" data (Vrandečić and Krötzsch, 2014). These statements are represented as linked data. Since Wikidata is not a primary source of information, these statements are supposed to contain a reference to the corresponding source. Wikidata can also deal with contradictory statements (Vrandecic, 2013). The content of Wikidata can be accessed through the Wikidata website, via the linked data query language SPARQL or by using an Application Programming Interface (API). Wikidata is designed to be used by both humans and machines.

Public organizations are affected by open knowledge bases like Wikidata for two reasons. First, Wikidata challenges such organizations by allowing everyone to publish LOGD from the manifold of available sources of government data. Even if public organizations choose not to publish any linked data at all, there would be LOGD available to the public affecting their sphere of interest. This would result in the public organizations losing the data publishing sovereignty in their own realm. Second, public organizations may use Wikidata for their own benefit and profit from the elaborated and free platform and visibility that it offers to improve the impact of LOGD thus generating potential to create public value.

Publishing data on Wikidata requires the use of a certain licence that guarantees a public domain dedication. Due to the openness of Wikidata, there is no guarantee that data put on Wikidata by public organizations will not be edited in a fraudulent or faulty way by third parties. However, Wikidata as a socio-technical system is prepared for such scenarios and provides tools enabling its community to roll back fraudulent or faulty edits. The large number of users and editors also increases the probability of detecting such edits promptly and reliably.

In Wikidata, there are no rules excluding information from certain sources nor are there any restrictions in regard to data published by public organizations. Wikidata requires published data to be in accordance with Wikidata's notability criteria.[10] These criteria include referring to a conceptual or material entity, which can be described using publicly available resources.

---

[7] https://www.egovernment.ch/en/umsetzung/e-government-schweiz-2008-2015/lindas/

[8] https://www.wikidata.org

[9] Source: http://wikidata.wikiscan.org (accessed: 18th of February 2018)

[10] https://www.wikidata.org/wiki/Wikidata:Notability

Our case study example of LOGD on Wikidata are the more than 2000 Swiss municipalities that have an entry on Wikidata. The range of available data for these municipalities varies from only a few basic statements including for example the number of inhabitants to very detailed information such as on the heads of government of the particular municipality during the 16th century.

Open knowledge bases will be of increasing importance because more and more applications are using the publicly available data stored there. An example case for such an application could include information concerning the tax rate (which in Switzerland differs from municipality to municipality) and the available building area reserve within a certain municipality. This information could be used in a housing search engine as decision guidance for someone contemplating to move to a certain place or not.

5   Conclusions

The main contribution of this study is the distinction of three different LOGD governance modes for data publishing and their resulting characteristics regarding the degree of control, data usability and data visibility. Furthermore, different possibilities for linking LOGD from different sources and the resulting effects on data usability, data visibility, and, finally the creation of public value are illustrated. These effects should be seen as inherent chances and risks, not definite outcomes. On the one hand, valuable chances to support public value creation could be missed. On the other hand, risks inherent to a certain LOGD mode may be mitigated by responding appropriately. Using the presented LOGD governance modes and linking the data will help not only to open up the data 'silos' but also effectively tear down any 'silo' boundaries. Furthermore, using open data in more effective ways by drawing on help from the outside may also help public organizations to cope with what has been described as the "big data revolution" in public affairs (Mergel et al., 2016).

As Table 1 and Table 2 show, there is no single LOGD governance mode combining all the advantages thus, we conclude that public organizations should be aware of the different forms of data governance and their strengths and weaknesses. Depending on the scenario, governments may choose to select a dedicated triple store, a shared triple store or an open knowledge base such as Wikidata. In this process, public sector IT departments should use their knowledge about open standards to play a key role (Fishenden and Thompson, 2013).

If a public organization needs to retain full control over the data during its whole life cycle, they may choose to operate a dedicated triple store. In this case, the organization should consider creating outgoing links to the LOD cloud to enhance usability and setting incoming links to the local store in open knowledge bases like Wikidata to enhance visibility. If, on the other hand, saving resources is of prime importance and some loss of freedom and flexibility is considered acceptable, a public organization may choose to publish via a shared triple store. This will enable the organization to share some of the expenditures. Finally, regarding data sets containing information usable in a wide variety of ways and possibly of interest to the general public like public transport, geography, climate and weather, a public organization should consider publication of this

data in an open knowledge base like Wikidata. Citizens using such data to create valuable applications could compensate the loss of control of the public organization. Such a use of LOGD would be promising for the creation of public value (Attard et al., 2016; Hui and Hayllar, 2010; Lee and Kwak, 2012; Zuiderwijk and Janssen, 2014).

Our case studies show, that the three LOGD governance modes are indeed used for publishing LOGD. The first case demonstrates a dedicated triple store showing the possibility of a technical realisation supporting specific data (in this case geospatial data) and the possibility to visualize such data right from the query interface. Examples and instructions how to use the data show the intention of the data publisher to help potential users benefit from the LOGD in order to create public value. The second case study involves a shared triple store (LINDAS) showing the cooperation of different public organizations to publish their data. Unfortunately, the datasets on this shared triple store are not interlinked to each other and there are no outgoing links to open knowledge bases. This shared triple store has some characteristics of a showcase targeted more at publishing organizations than at data users. Correspondingly, there is less support for potential data users like instructions and examples. However, recent actions of the Swiss Federal Archives point to a future of LINDAS as a more productive shared triple store environment. About 26 million Swiss Francs[11] are being invested for this purpose. The third study looks at the open knowledge base Wikidata. It shows one of the strengths of such an open knowledge base, taking as an example the vast amount of very detailed and specific information about individual municipalities. However, such very profound information is available only for a small subset of the entirety of Swiss municipalities demonstrating a lack of data completeness.

Using the dedicated triple store of the office of topography as an example, we have briefly illustrated some of the potential of the use of LOGD. Thanks to the linking of the underlying data, no more than the execution of a single query was necessary to combine data from the dedicated triple store and from Wikidata.

Concluding, we believe that the general concept of LOGD may help to lower the impediments to the use of open data mentioned by Zuiderwijk, Janssen et al. (2012). Linked data allows for easier understandability of OGD, linking and combining different OGD sources, and the addition of metadata to OGD. Additionally, the choice of an appropriate LOGD governance mode may help to decrease further, specific impediments. Using an open knowledge base for example, allows for easier access to OGD and stimulates the interaction with the data provider.

The main limitation of our approach is its conceptual nature. Future research should include a more in-depth, empirical analysis with the goal of formulating best practices for LOGD publishing involving open knowledge bases. Furthermore, some mixed approaches should be considered, like publishing only parts of the data to an open knowledge base while reserving the complete data for a dedicated or shared triple store. Such approaches would presumably be able to combine the advantages of both modes.

---

[11] https://www.inside-it.ch/articles/51757

References

Abele, A., McCrae, J.P., Jentzsch, A. and Cyganiak, R. (2017), "The Linking Open Data cloud diagram", available at: http://lod-cloud.net/ (accessed 28 February 2018).

Attard, J., Orlandi, F. and Auer, S. (2016), "Value Creation on Open Government Data", *49th Hawaii International Conference on System Sciences (HICSS)*, pp. 2605–2614.

Ballatore, A., Wilson, D.C. and Bertolotto, M. (2013), "A Survey of Volunteered Open Geo-Knowledge Bases in the Semantic Web", in Pasi, G., Bordogna, G. and Jain, L.C. (Eds.), *Quality Issues in the Management of Web Information*, Vol. 50, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 93–120.

Benington, J. and Moore, M.H. (2010), *Public Value: Theory and Practice*, Palgrave Macmillan, London.

Berners-Lee, T. (2006), "Linked Data - Design Issues", available at: https://www.w3.org/DesignIssues/LinkedData.html (accessed 21 September 2016).

Bizer, C. (2009), "The Emerging Web of Linked Data", *IEEE Intelligent Systems*, Vol. 24 No. 5, pp. 87–92.

Bizer, C., Heath, T. and Berners-Lee, T. (2009), "Linked Data - the story so far", *International Journal on Semantic Web and Information Systems*, Vol. 5 No. 3, pp. 1–22.

Bizer, C., Lehmann, J., Kobilarov, G., Auer, S., Becker, C., Cyganiak, R. and Hellmann, S. (2009), "DBpedia-A crystallization point for the Web of Data", *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 7 No. 3, pp. 154–165.

Boudreau, K. (2010), "Open platform strategies and innovation: Granting access vs. devolving control", *Management Science*, Vol. 56 No. 10, pp. 1849–1872.

Bryson, J., Crosby, B.C. and Bloomberg, L. (2014), "Public Value Governance: Moving Beyond Traditional Public Administration and the New Public Management", *Public Administration Review*, Vol. 74 No. 4, pp. 445–456.

Davies, T. (2013), *Open Data Barometer - 2013 Global Report*, available at: http://www.cocoaconnect.org/sites/default/files/publication/Open-Data-Barometer-2013-Global-Report.pdf (accessed 8 November 2016).

Ding, L., Peristeras, V. and Hausenblas, M. (2012), "Linked Open Government Data [Guest editors' introduction]", *IEEE Intelligent Systems*, Vol. 27 No. 3, pp. 11–15.

Färber, M., Bartscherer, F., Menne, C. and Rettinger, A. (2017), "Linked Data Quality of DBpedia, Freebase, OpenCyc, Wikidata, and YAGO", *Semantic Web*, Vol. 9 No. 1, pp. 77–129.

Fishenden, J. and Thompson, M. (2013), "Digital Government, Open Architecture, and Innovation: Why Public Sector IT Will Never Be the Same Again", *Journal of Public Administration Research and Theory*, Vol. 23 No. 4, pp. 977–1004.

Gil-García, J.R. and Pardo, T.A. (2005), "E-government success factors: Mapping practical tools to theoretical foundations", *Government Information Quarterly*, Vol. 22 No. 2, pp. 187–216.

Gilmour, R.S. and Jensen, L.S. (1998), "Reinventing Government Accountability: Public Functions, Privatization, and the Meaning of 'State Action'", *Public Administration Review*, Vol. 58 No. 3, pp. 247–258.

Harrison, T.M., Guerrero, S., Burke, G.B., Cook, M., Cresswell, A., Helbig, N., Hrdinova, J., et al. (2012), "Open government and e-government: Democratic challenges from a public value perspective", *Information Polity*, Vol. 17 No. 2, pp. 83–97.

Hausenblas, M. (2009), "Exploiting linked data to build web applications", *IEEE Internet Computing*, Vol. 13 No. 4, pp. 68–73.

Heath, T. (2008), "How will we interact with the web of data?", *IEEE Internet Computing*, Vol. 12 No. 5, pp. 88–91.

Hendler, J., Holm, J., Musialek, C. and Thomas, G. (2012), "US Government Linked Open Data: Semantic.data.gov", *IEEE Intelligent Systems*, Vol. 27 No. 3, pp. 25–31.

Höchtl, J. and Reichstädter, P. (2011), "Linked Open Data - A Means for Public Sector Information Management", in Andersen, K.N., Francesconi, E., Grönlund, Å. and van Engers, T.M. (Eds.), *Electronic Government and the Information Systems Perspective*, Springer Berlin Heidelberg, pp. 330–343.

Hui, G. and Hayllar, M.R. (2010), "Creating Public Value in E-Government: A Public-Private-Citizen Collaboration Framework in Web 2.0", *Australian Journal of Public Administration*, Vol. 69, pp. S120–S131.

Janssen, M., Charalabidis, Y. and Zuiderwijk, A. (2012), "Benefits, Adoption Barriers and Myths of Open Data and Open Government", *Information Systems Management*, Vol. 29 No. 4, pp. 258–268.

Janssen, M. and van den Hoven, J. (2015), "Big and Open Linked Data (BOLD) in government: A challenge to transparency and privacy?", *Government Information Quarterly*, Vol. 32 No. 4, pp. 363–368.

Jetzek, T. (2016), "Managing complexity across multiple dimensions of liquid open data: The case of the Danish Basic Data Program", *Government Information Quarterly*, Vol. 33 No. 1, pp. 89–104.

Kalampokis, E., Tambouris, E. and Tarabanis, K. (2011), "A classification scheme for open government data: towards linking decentralised data", *International Journal of Web Engineering and Technology*, Vol. 6 No. 3, pp. 266–285.

King, G., Keohane, R.O. and Verba, S. (1994), *Designing Social Inquiry: Scientific Inference in Qualitative Research*, Princeton University Press, Princeton, New Jersey.

de Kool, D. and Bekkers, V. (2016), "The perceived value-relevance of open data in the parents' choice of Dutch primary schools", *International Journal of Public Sector Management*, Vol. 29 No. 3, pp. 271–287.

Lakomaa, E. and Kallberg, J. (2013), "Open Data as a Foundation for Innovation: The Enabling Effect of Free Public Sector Information for Entrepreneurs", *IEEE Access*, Vol. 1, pp. 558–563.

Lee, G. and Kwak, Y.H. (2012), "An Open Government Maturity Model for social media-based public engagement", *Government Information Quarterly*, Vol. 29 No. 4, pp. 492–503.

Lourenço, R.P. (2016), "Evidence of an Open Government Data Portal Impact on the Public Sphere", *International Journal of Electronic Government Research (IJEGR)*, Vol. 12 No. 3, pp. 21–36.

Mergel, I., Rethemeyer, R.K. and Isett, K. (2016), "Big Data in Public Affairs", *Public Administration Review*, Vol. 76 No. 6, pp. 928–937.

Neumaier, S., Umbrich, J. and Polleres, A. (2016), "Automated Quality Assessment of Metadata across Open Data Portals", *Journal of Data and Information Quality*, Vol. 8 No. 1, pp. 1–29.

Nugroho, R.P., Zuiderwijk, A., Janssen, M. and de Jong, M. (2015), "A comparison of national open data policies: lessons learned", *Transforming Government: People, Process and Policy*, Vol. 9 No. 3, pp. 286–308.

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*, Vol. 22 No. 1, pp. 1–24.

Shadbolt, N., O'Hara, K., Berners-Lee, T., Gibbins, N., Glaser, H., Hall, W. and schraefel, m. c. (2012), "Linked open government data: Lessons from data. gov. uk", *IEEE Intelligent Systems*, Vol. 27 No. 3, pp. 16–24.

Siggelkow, N. (2007), "Persuasion with Case Studies", *Academy of Management Journal*, Vol. 50 No. 1, pp. 20–24.

Smith, G. and Sandberg, J. (2018), "Barriers to innovating with open government data: Exploring experiences across service phases and user types", *Information Polity*, Vol. 23 No. 3, pp. 249–265.

Strong, D.M., Lee, Y.W. and Wang, R.Y. (1997), "Data Quality in Context", *Commun. ACM*, Vol. 40 No. 5, pp. 103–110.

Stuermer, M. and Dapp, M.M. (2016), "Measuring the Promise of Open Data: Development of the Impact Monitoring Framework", *Conference for E-Democracy and Open Government (CeDEM)*, pp. 197–203.

Stuermer, M., Spaeth, S. and Krogh, G.V. (2009), "Extending private-collective innovation: a case study", *R&D Management*, Vol. 39 No. 2, pp. 170–191.

Umbrich, J., Neumaier, S. and Polleres, A. (2015), "Quality Assessment and Evolution of Open Data Portals", *3rd International Conference on Future Internet of Things and Cloud*, pp. 404–411.

Villazón-Terrazas, B., Vilches-Blázquez, Luis.M., Corcho, O. and Gómez-Pérez, A. (2011), "Methodological Guidelines for Publishing Government Linked Data", in Wood, D. (Ed.), *Linking Government Data*, Springer New York, New York, NY, pp. 27–49.

Vrandecic, D. (2013), "The Rise of Wikidata", *IEEE Intelligent Systems*, Vol. 28 No. 4, pp. 90–95.

Vrandečić, D. and Krötzsch, M. (2014), "Wikidata: A Free Collaborative Knowledgebase", *Commun. ACM*, Vol. 57 No. 10, pp. 78–85.

Walsham, G. (1995), "Interpretive case studies in IS research: nature and method", *European Journal of Information Systems*, Vol. 4 No. 2, pp. 74–81.

Wang, H.-J. and Lo, J. (2016), "Adoption of open government data among government agencies", *Government Information Quarterly*, Vol. 33 No. 1, pp. 80–88.

Wang, R.Y. and Strong, D.M. (1996), "Beyond Accuracy: What Data Quality Means to Data Consumers", *Journal of Management Information Systems*, Vol. 12 No. 4, pp. 5–33.

Welch, E.W. and Wong, W. (2001), "Global Information Technology Pressure and Government Accountability: The Mediating Effect of Domestic Context on Website Openness", *Journal of Public Administration Research and Theory*, Vol. 11 No. 4, pp. 509–538.

West, J. (2003), "How open is open enough?: Melding proprietary and open source platform strategies", *Research Policy*, Vol. 32 No. 7, pp. 1259–1285.

West, J. and Bogers, M. (2014), "Leveraging External Sources of Innovation: A Review of Research on Open Innovation", *Journal of Product Innovation Management*, Vol. 31 No. 4, pp. 814–831.

Yang, T.-M. and Maxwell, T.A. (2011), "Information-sharing in public organizations: A literature review of interpersonal, intra-organizational and inter-organizational success factors", *Government Information Quarterly*, Vol. 28 No. 2, pp. 164–175.

Zaveri, A., Rula, A., Maurino, A., Pietrobon, R., Lehmann, J. and Auer, S. (2016), "Quality assessment for Linked Data: A Survey", *Semantic Web*, Vol. 7 No. 1, pp. 63–93.

Zuiderwijk, A. and Janssen, M. (2014), "Open data policies, their implementation and impact: A framework for comparison", *Government Information Quarterly*, Vol. 31 No. 1, pp. 17–29.

Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., Alibaks, R.S. and Sheikh_Alibaks, R. (2012), "Socio-technical impediments of open data", *Electronic Journal of EGovernment*, Vol. 10 No. 2, pp. 156–172.

Zuiderwijk, A., Jeffery, K. and Janssen, M. (2012), "The Potential of Metadata for Linked Open Data and its Value for Users and Publishers", *JeDEM - EJournal of EDemocracy and Open Government*, Vol. 4 No. 2, pp. 222–244.