Open Government Data Systems
Learning from a Public Utility Perspective

Crusoe, Jonathan; Zuiderwijk, Anneke; Melin, Ulf

DOI
10.1007/978-3-030-57599-1_21

Publication date
2020

Document Version
Final published version

Published in
Electronic Government - 19th IFIP WG 8.5 International Conference, EGOV 2020, Proceedings

Citation (APA)
Crusoe, J., Zuiderwijk, A., & Melin, U. (2020). Open Government Data Systems: Learning from a Public Utility Perspective. In G. Viale Pereira, M. Janssen, H. Lee, I. Lindgren, M. P. Rodríguez Bolívar, H. J. Scholl, & A. Zuiderwijk (Eds.), *Electronic Government - 19th IFIP WG 8.5 International Conference, EGOV 2020, Proceedings* (pp. 275-289). (Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Vol. 12219 LNCS). Springer. https://doi.org/10.1007/978-3-030-57599-1_21

Open Government Data Systems: Learning from a Public Utility Perspective

Jonathan Crusoe[1(✉)], Anneke Zuiderwijk[2], and Ulf Melin[1]

[1] Division of Information Systems and Digitalization, Linköping University, 581 83 Linköping, Sweden
{jonathan.crusoe,ulf.melin}@liu.se
[2] Faculty of Technology, Policy and Management, Delft University of Technology, 2628 BX Delft, The Netherlands
A.M.G.Zuiderwijk-vanEijk@tudelft.nl

Abstract. Previous research on Open Government Data (OGD) struggles with synthesising a holistic perspective of OGD systems. A perspective that has dealt with vast, complex systems is public utility. Public utilities are, for example, water supply networks and electric power grids. This study explores what we can learn from a public utility perspective when perceiving and organising OGD systems. We used a hermeneutic literature review combined with a snowballing approach, resulting in a selection of 39 studies. We compare public utilities and OGD systems to derive five lessons: (1) an OGD system can be perceived from a node-flow view, (2) the foundational data flow of an OGD system starts at data collection and ends at data used by the public in an everyday context, (3) the organisation of OGD systems needs to consider the combinability, interpretability, and boundless reusability of data, (4) OGD systems need governance organisations that cover the whole system, and (5) OGD systems could replace existing data provision systems and be made a public utility if certain characteristic problems are overcome.

Keywords: Open data · Open Government Data · Public utility · System

1 Introduction

Public organisations openly share data over the Internet. Ideally, this data can be reused by anyone without restrictions. Data providers and data users organise as an Open Government Data (OGD) system by following the OGD principles (e.g., non-discriminatory data access and machine-readable data) [4,23,45]. The actors (people and organisations) work together for a particular purpose or conduct similar activities [44]. They believe that the OGD system will lead to benefits, such as increased transparency and improved government services [33].

Published by Springer Nature Switzerland AG 2020
G. Viale Pereira et al. (Eds.): EGOV 2020, LNCS 12219, pp. 275–289, 2020.
https://doi.org/10.1007/978-3-030-57599-1_21

An OGD system consists of interacting subsystems, such as an ecosystem of data providers or an ecosystem of data users [24,47]. These subsystems have their own rules, behaviours, and interactions [24,40]. Events or changes in one subsystem can impact other subsystems [24,25]. The OGD system is dependent on infrastructure [17]. The infrastructure help actors discover data through OGD portals [13,30] and enables the transfer of data between actors [29]. The OGD system is also limited and enabled by laws, regulations, and licenses [42].

While previous OGD research contains attempts to study OGD systems from a holistic perspective, such as ecosystem [e.g., 18], lifecycle [e.g., 21], and value network perspectives [e.g., 3], there are still three major issues related to the synthesis of a holistic perspective in OGD research. First, the OGD system is often treated as one system without subsystems and their interactions [e.g., 21,36,39], which may cause the rules and behaviours of the subsystem to be lost or merged with other subsystems. Second, the extensive infrastructure is often abated (to make something less serious) by not being fully integrated into holistic perspectives of OGD systems [e.g., 15,18,48]. Third, the legal system tends not to be incorporated into holistic perspectives of OGD systems [e.g., 1,3,15]. The legal system is one source of rules and impediments for the subsystems.

A holistic perspective that has dealt with vast, complex systems is the public utility perspective. Examples of public utilities include water supply and power supply [9]. A public utility is a large group of public service organisations that produce and manage a variety of products and services vital to modern life [38]. Their system is based on an extensive infrastructure that helps them to create products or services at one location and distribute them to several customers [22,38]. The government can be involved, oversee, or regulate this system that has to serve all customers without discrimination [38].

We argue that knowledge about public utilities could help OGD researchers synthesise a holistic perspective of OGD systems and be informative about how to organise OGD systems. The objective of this study is to explore what we can learn from a public utility perspective when perceiving and organising OGD systems.

This paper contains the following sections. First, the research approach used to identify literature about OGD systems and public utilities is described. Second, a synthesised description of public utilities with two examples is presented. Third, OGD systems and public utilities are compared. Then, five lessons are presented. Fourth, a conclusion with limitations and future work is given.

2 Research Approach

This paper is based on an exploration of the literature using a hermeneutic literature review method with citation tracking and citation analysis (snowballing) [8]. The review aimed to synthesise two bodies of literature and develop an understanding of OGD systems, drawing on literature concerning public utilities. The result also includes generic concepts (comparative categories) that were used as common points of reference for the comparison between OGD systems and public utilities [41]. Next, the two phases of the review are described.

The **first phase** focused on exploring literature about public utilities. We used the keyword "public utility" with Google Scholar and Google to identify literature about public utilities' organisation and development from a holistic perspective. Google Scholar was used since the search engine has a good coverage and recall speed with a precision above-average for simple keyword searches [49]. Google was used to identify sources that could contain relevant literature, such as online book shops or ResearchGate or synonyms to keywords. The keywords developed over time as new literature was identified. The keyword "public utility" was also combined with other keywords, such as "infrastructure" and "public administration". We identified literature by skimming their titles and abstracts, then, if needed, a quick internal search of the main text was done. In the case of books, if needed, reviews were read. In total, four books and two book chapters about public utilities were identified as relevant. Then, we mapped key ideas and insights about the topic by studying paragraphs and chapters to synthesise them, which resulted in a textual overview. Finally, we used the result to write this paper and as an input for the next phase.

The **second phase** focused on OGD systems, the comparative categories, and the comparison between public utilities and OGD systems. We used the keywords "open data" or "open government data" combined with "ecosystem", "lifecycle", and "value network". We skimmed the titles and abstracts to identify their relevance. If needed, their content was skimmed or quickly searched through to identify key ideas and insights about the topic. We considered studies to be relevant if they contributed to the understanding of OGD systems or the comparison between OGD systems and public utilities. In total, one book, one thesis, three reports, five empirical texts (e.g., OGD principles), four conference papers, and 19 journal articles about OGD were identified as relevant. We studied the relevant literature for key ideas and insights. We then compared the literature about OGD systems and public utilities to identify possible areas of comparison and create a list of the comparative categories. From this list, we selected the comparative categories "the system", "infrastructure", and "governance", as they relate to the problem described in the introduction. Also, the comparative category "the public" was included since the public (citizens) are important beneficiaries of both public utilities and OGD systems, while the comparative category "resource" was included because of its stressed and varied nature in the public utility literature. We then filled out each of the selected comparative categories for OGD systems and public utilities, which became the comparison. Lessons were drawn from each comparative category by identifying and analysing the comparison.

3 Public Utilities

This section presents the synthesis of the identified public utility literature, where two public utilities are described in detail. We define public utilities as *"[...] the large group of public service organisations that exist to locate, produce or collect, transmit, distribute and/or process and store, a variety of products and services*

that are vital to modern life. These products and services are ubiquitous." [38, p. 5]. Public utilities are based on an extensive network structure that acts as a distribution system and is built from components with strong physical linkages. Organisations create products or services at one location and then distribute them over the network to several customers. The base infrastructure is often partly owned by the government, and, typically, there is a substantial sunk cost since the infrastructure is extensive [22,38]. The infrastructure is often shared with other actors, and if not managed properly, could lead to congestion [22].

The activities of public utilities can be divided into production, transmission, and distribution components (e.g., electricity: generating plants, high-voltage lines, and local power lines) [14,22]. The governance of public utilities is complex. Governments are involved, oversee, or regulate public utilities, and the utilities have to serve all customers without discrimination within their market area [38]. Regulations can help to enact standards to ensure reasonable service [46]. On the other hand, rigid standards can handicap utilities, such as when gas utilities went from illumination to cooking and heating [46]. The enforcement of standards is an important task for public organisations and can occupy a large portion of their time [46]. The income of public utilities can come from payments, stocks, bonds, and taxes [38]. In this paper, the public refers to the end-users who use and benefit from the services and products in an everyday context.

Public utilities provide essential services necessary for modern life. Without them, there can be no cities and nations [38]. They are often commonly found wherever people are living in villages, towns, and cities [46]. They are needed to support large populations, overcome the social consequences from the concentration of people, and allow cities to grow beyond their countryside [9,46]. In the past, when towns were generally small, public utilities were often simple, such as rivers and wells [9]. Over time, city development, urbanisation, industrialisation, and digitalisation have had a significant impact on public utilities' growth and development [38,46]. Public utilities can form monopolies, which occur when a single firm can supply a market at a cost and price far lower than would be possible if several firms served the market [38] or when there are natural limitations in supply or conditions of service [46].

Two common public utilities in any modern society are the water supply networks and electric power grids. We will present each public utility in detail to gain an in-depth understanding of their characteristics. **Water Supply Networks** are the most basic of public utilities. No town can survive long without a water supply for drinking, washing, and various industrial processes [9]. If the supply becomes polluted, disease and death can follow [38]. A water supply network consists of the water source, collection, transportation, raw water storage, treatment, finished water storage, distribution, and process monitoring components [38]. The monitoring could include identifying and fixing leaks [2]. To give an example of the public utility: aquifer pumps [9] or watersheds provide water to reservoirs. Aqueducts, culverts, or large pipes then move the water in large quantities to populated areas. The water is treated or filtered to ensure consistent quality (e.g., remove dirt or kill bacteria) [2,9]. The water is then distributed

to households through local water supply networks where people can access the water with, for example, faucets, showers, or toilets [2]. Water supply utilities are also regulated. For example, the American Safe Drinking Water Act sought to protect public health by regulating the nation's public drinking water supply. The act requires the American Environmental Protection Agency (EPA) to set national health standards, which they do in a three-step process. The EPA also proposes treatment techniques and monitor the water supply networks [38].

The instant availability of electricity is assumed in the modern industry and domestic life. **Electric power grids** have become the largest public utility in modern societies. One of its greatest achievements is enabling lamps almost everywhere. Among the first power grids were pioneered as early as 1879 [9]. Electric power grids consist of four functional components: (1) power generation, (2) transmission of high voltage power, (3) distribution of low voltage power, and (4) marketing functions [38]. To give an example of the public utility: electricity is generated by power plants [2]. The electricity is sent to a substation where transformers increase its voltage, which allows it to travel long distances. The electricity travels over the high-voltage transmission grid to an area substation where it is stepped down to a lower voltage to be sent over the distribution grid. It is then distributed to homes and businesses [2]. People access the electric grid through plugging an electric device's cable into an electrical outlet. The device could be a computer or a dishwasher [2]. One major challenge with electricity is its behaviour (loop flow) to take all available routes to get from one point to another. If one line shuts down, the load is dispersed to all other lines, which can cause an overload that shuts down the entire grid [38]. Another major challenge is that electricity cannot be economically stored. It needs to be used as it is produced. Consequently, the grid has to be monitored to ensure that the grid is always met demand and does not overload [38]. Electrical power grids are also regulated. For example, Order 889 issued by the American Federal Energy Reliability Council (FERC) required all investor-owned utilities to participate in an open access same-time information system (OASIS) to make pricing and accessing processes more transparent [38].

4 Comparison Between OGD Systems and Public Utilities

This section first presents a summary of the comparison between public utilities and OGD systems (see Table 1). Then, each of the comparative categories is, in turn, presented with previous OGD research and their comparison and analysis to draw lessons learned. This section ends with an intermediate conclusion based on the comparison. The categories are the system, the public, infrastructure, resource, and governance.

4.1 The System

The system is concerned with how the system is holistically perceived as parts (e.g., elements and subsystems) and interactions. An OGD system has at least

Table 1. Comparison between OGD system and public utilities

Comparative categories	Public utilities	OGD system
The system	Objects or components with distribution and transmission [14,22,38]	Actors' interactions or elements with data flows [1,15,37,52]
The public	Served, costly products or services endowed by the public and limited access to the infrastructure [2,38]	Shared, free data, the public's interest in reusing the data is a myth, and open access to the data source [16,26,32]
Infrastructure	Distribution networks of social and technical parts over large areas [22,38]	Technical, functional artifacts (possibly on the Internet) for use by different actors [10,17,29,52]
Resource	A small set of homogeneous, standardised, public, physical, ubiquitous resources that are vital to modern life [9,38]	A large set of heterogeneous, standardized, public, intangible, recorded, interoperable data that needs interpretation [4,5,12,16]
Governance	Resource governance, regulated by law, governance organisations, and heavy involvement in the operations of actors and their interactions [38,46]	Suggested resource governance, regulated by law, international principles, more self-regulation, and unrestricted users [7,11,18,23, 32,36,42,45,47,48,51,52]

three ecosystems: data providers, (intermediate) data users, and end-users [47]. Data providers are public organisations who share their data [4,18]. They often have natural monopolies on their data because of high fixed costs, low variable costs, and a small number of potential data users [25]. Data users can be developers, journalists, and researchers [43]. They can take many different roles (e.g., data analyser or user experience provider [37]) and can use data to either provide new enhanced data back into the OGD system or products and services to the public [16,37]. End-users (the public) are any actor that uses the data directly by extracting facts or indirectly through using products and services provided by data users [16]. The data is often holistically perceived as a flow from data providers to data users and sometimes end-users [e.g., 16,37]. OGD systems can also be perceived as a collection of elements, such as data audit, data access, and data use [18,36,52].

OGD systems and public utilities follow the same basic pattern: (1) produce, (2) transmit and distribute, and (3) use products or services [e.g., 2,25]. Both holistic perspectives attempt to identify and order key activities in the system following the logic of a chain or flow. OGD systems use the terminology elements, while public utilities use components. We will continue to use 'component' for both. However, OGD research focuses on actors (e.g., data providers and data users) with data flows [e.g., 1,15,37,47,52], while public utilities focus on objects and their systematic purpose (e.g., generating plants or high-voltage lines [14,22,38]). Previous OGD research tends to focus on the interaction between

data providers and users [e.g., 24,52], while public utilities focus on a chain of objects and activities [14,22,38].

The idea of key activities, actors, and objects seem to be a promising way to holistically perceive vast, complex systems where they start at the origin of the product (e.g., data, water, or electricity) and then continue to its distribution and transmission. The activities, actors, and objects as components can roughly be divided into nodes (e.g., weather data collection, water treatment, or power generation) and infrastructure (e.g., the Internet, aqueducts, or high-voltage grids). However, if we stay at nodes and infrastructure as a chain of components, behaviours, such as monopolies or infrastructure congestion, or needs to monitor the infrastructure can be lost. We combine the idea of components with the insight that OGD systems are known to consist of subsystems [24,47,52] to get functional systems. This approach follows in McNabb's [38] components of the water supply networks and functional components of electric power systems. A functional system is a collection of nodes with a similar purpose and similar key activities in the OGD system. The system is connected to other functional systems to organise as a data flow. The nodes can have relationships and interactions that will impact functional systems later down the chain.

4.2 The Public

The public is concerned with the availability and access of products, services, or data (resources) of the systems to the public. In OGD systems, data is provided under an open license that allows anyone to acquire it for free and use and distribute it without restriction [11,23,32,45]. OGD systems and public utilities both provide their resources to the public without discrimination (anyone has the right to access and use the resource). However, they differ on at least four points. First, OGD systems share data to anyone with an Internet connection [16], while public utilities serve the public in their market area [38]. Second, the public interest in reusing OGD has been declared a myth [26], while public utilities are endowed by public interest [38]. Third, the public has direct access to the data in OGD systems [32], while the public has limited access to the infrastructure of public utilities, such as high-voltage grids and aqueducts [2]. Fourth, in OGD systems, the data is free [32], while in public utilities, the products or services can need payment (e.g., via tax or direct fees) [38].

Some of the differences are likely a consequence of the rather immature age of OGD [5] in comparison to the ancient history of public utilities [9]. While the data might not be interesting, it is arguable that the products built with it are attractive [20]. The use of data requires certain technical skills and expertise [43], which limits the development of these products to specialists. It is perceivable that OGD systems could follow in the steps of water supply networks or electric power grids. For example, data users could sell products that the public uses by plugging them into the data infrastructure. This situation is similar to how the public buys electrical devices and plug them into an electrical outlet or the water supply network (e.g., dishwashers) [2]. However, in the case of OGD systems, the resource that powers the device would be free. This possibility would continue

to build on the idea of data users as intermediaries [e.g., 25,28], but stresses that the raw data might only be for specialists and that the products and services should be sought after by the public. Consequently, the foundational data flow of an OGD system can be perceived to start at data collection and end at the use of data-based products or services by the public in, for example, their homes and businesses.

4.3 Infrastructure

Infrastructure focuses on the use of infrastructure in the systems, but also how it is holistically perceived. OGD systems are dependent on an OGD infrastructure, where the Internet is one part [16,17,23]. However, there is no agreed definition of infrastructure for OGD systems [e.g., 10,43], but it seems to be either specific artefacts and the Internet [e.g., 17,52] or something that facilitates data exchange between actors or the use of the data [e.g., 29].

OGD systems and public utilities have in common that they holistically perceive infrastructure as a tool with a specific purpose or use. In OGD systems, infrastructure is heavy technical with a focus on its use by different actors [10,17,52], while in public utilities, infrastructure is a mixture of social and technical parts over a large area [2,22,38]. Previous OGD research perceives infrastructure as functional artefacts that can exist on the Internet [17,29,52], while public utilities perceive it as a distribution network [22,38].

For the system category, we divided components into nodes and infrastructure. The idea to holistically perceive infrastructure as a distribution network seems to be promising for OGD systems, as the Internet can similarly be used to request and send data packages between locations [35]. However, public utilities differ from OGD systems in how resources are supplied. Electricity and water, once in their infrastructure, are accessible to anyone who can use the infrastructure. On the other hand, for OGD systems, the data is first published on the Internet, then must be found [13,30], and, finally, it can be downloaded. Consequently, while the Internet is a clear psychical infrastructure (e.g., servers, routers, and fibre cables [35]), OGD systems have two complementary infrastructure networks: data distribution and data discovery. The networks help to connect the nodes and functional systems and ensure that data can be found, transmitted, and distributed between them.

4.4 Resource

Resource is concerned with the properties of the central products, services, or data provided by or in the systems. In OGD systems, data is a resource (e.g., weather data [5]) that is shared to be used [36], ideally, without direct cost. The data is a selection of facts people have paid attention to and collected and becomes information when people attribute meaning to the data [12]. Data has boundless uses, as it is not used-up in its process of being used. The data of OGD systems needs to be machine-readable, complete, primary, timely, accessible, non-proprietary, and interoperable [23,32,45]. At the same time, the data, while

anonymous, could be combined with other data to re-identify individuals, which then threatens peoples' right to information privacy [34].

The resources of OGD systems and public utilities are placed under standards and provided to be used by the public. However, data of OGD systems are large sets of heterogeneous resources (e.g., weather reports, crime reports, or postal codes [5]) that are intangible, recorded, and supposed to be interoperable with other data (e.g., blog posts and interactive maps [16]). The data requires interpretation by users to be used or understood [12] and has boundless reuses, where visualisations and mash-ups can help in the process [4]. The data can be used alone or in combination, which leads to a large set of possible uses based on its heterogeneity. The possible uses of a single dataset (e.g., weather) have limited applications, but together the application is vast. On the other hand, resources of public utilities are small sets of homogeneous resources (e.g., water, electricity, gas, transportation, communication [9]) that are physical, consumable, ubiquitous, and vital to modern life [38]. These resources have a wide range of applications [2,9].

The properties of the resources in public utilities impact the organisation of the utilities. For example, water supply networks have to identify and repair leaks [2], while electric power grids have to monitor the grid to ensure that supply meets demand, and it does not overload [38]. Similarly, anonymous data can be combined to re-identify individuals [34], while there is also a fear of misinterpretation and fraud [6]. OGD systems need to be organised with consideration of the properties of data as a resource.

4.5 Governance

Governance covers both formal and informal processes and institutions that guide and restrain the collective activities of systems [31]. It is the internal and external exercise of direction, control, management, and policy-shaping of the components of the systems [38]. In OGD systems, data is a common, shared resource that should be governed by data providers to ensure it has sufficient quality for use [29]. For example, [19] suggest institutional cooperation, model frameworks, international agreement, and social certification to govern data ownership. The foundational flow of data from collection to use to function properly and realise the benefits of OGD needs to be coordinated [51], and so does the infrastructure [17]. For example, published data need to be usable and discoverable by users, responsibilities between data providers and users need to be clear, and the process should follow standards [51]. This type of governance involves feedback and discussions between data providers and users [18,51,52], licenses and principles [11,23,32,45], and legal and regulatory frameworks [47]. Laws, policies, standards, and agreements have a bearing on OGD systems [48], which can limit and enable actors [42]. Data protection laws must be followed by the actors, and the data cannot be linked back to individuals [36]. Important legislation for OGD can relate to freedom of information, public sector information, data protection, data sharing, and statistics [36].

OGD systems and public utilities are governing resources and regulated by laws. However, they differ on at least three points. First, OGD systems follow international principles and open licenses [7,11,23,32,45], while public utilities follow national regulations and standards from many levels (e.g., federal, state, and local in the USA) [38]. Second, OGD systems are more self-regulated with a focus on the interaction between data providers and data users (e.g., discussion and feedback) [18,51,52], while public utilities have governance frameworks or organisations that can monitor, regulate, or enforce standards to ensure reasonable service (e.g., EPA for water and FERC for electricity) [38,46]. Third, in OGD systems, data users are under no restrictions on their use or distribution [11,23,32,45], while in public utilities, the actors' operations and interactions can experience heavy involvement from governance organisations [38,46].

While there are similarities between OGD systems and public utilities, they are different in how they approach governance. OGD systems have a focus on governance of production and the interaction between a data provider and a data user [18,29,50,52], while public utilities cover production to consumption where every actor can be directly or indirectly governed [38,46]. The OGD system approach most likely leads to an increased perception of risks and need to build trust. However, when a system becomes vast and complex, reoccurring interactions between parties become difficult, which could be a reason for the governance organisational approach of public utilities. The use of governance organisations could help OGD systems be beneficial for the public, data providers, and data users, but can collide with OGD principles, such as unrestricted use and use by anyone. On the other hand, there are already organisations that seem to try to govern OGD systems, such as, the meeting that resulted in the eight OGD principles [45] and the Open Knowledge Foundation [32].

4.6 Are OGD Systems a Type of Public Utility?

OGD systems and public utilities are similar in their basic pattern, infrastructure, and relationship to the public. However, OGD systems provide a free, heterogeneous resource, ideally, to anyone, while not being endowed by the public. The OGD principles (e.g., unrestricted use and access to anyone), the properties of data (e.g., ability to combine and need for interpretation), and lack of interest are three characteristic problems that can be interpreted against viewing OGD systems as a type of public utility. On the other hand, public utilities vary a lot (e.g., airlines, trucking, telecoms, electricity, natural gas, and railroads [14,22]) and are still considered to belong to the same type. Public utilities are vital to modern life [38], while the benefits of OGD is more proclaimed than empirically tested [43]. On the other hand, not all public utilities started as vital for society. For example, the electric power grids were introduced in the 19th century and partly or fully replaced heating and lighting based on the gas supply networks [9]. Today, data is arguably everywhere in modern life and vital for a modern society. It can be found at many locations (e.g., parliament minutes and governmental budgets [4]). OGD systems could be a replacement of existing data

provision systems. We argue that if OGD systems overcome the three characteristic problems, OGD systems could become a public utility. We conclude that public utilities can teach OGD researchers about large-scale organisation and governance, and how vast, complex systems can be perceived.

5 The Lessons Learned

In this section, we present five lessons learned from the comparison and their implications for OGD research and OGD practice. The two initial lessons came from synthesising the holistic perspectives of public utilities and OGD systems to supplement the latter, the following two lessons are what public utilities taught us about organising OGD systems, and the final lesson gave us a possible future for OGD systems based on public utilities.

The first lesson is that *an OGD system can be perceived as a collection of nodes (functional systems) connected by complementary infrastructure networks that together organise a foundational data flow.* This perspective shifts the actor-interactions view in previous OGD research [e.g., 24,48,52] to a node-flow view similar to [37]'s actor-flow view, but the focus is on infrastructure connecting nodes of activities, objects, and actors to enable the OGD system. This shift was enabled by the public utility perspective.

The second lesson is that *the foundational data flow of an OGD system starts at data collections and ends where the public uses the data in an everyday context by plugging it into the data distribution network.* This lesson adds to [e.g., 16,18,37] by empathizing data collection, the needs for the data to have utility for the public, and the role of data users as specialists (similar to the intermediary view [e.g., 25,28]). It also shifts the focus from the interaction between data providers and data users [e.g., 4,52] to how OGD systems work to provide benefits to the public and subsystems developing from and connecting to the foundational flow.

The third lesson is that *OGD systems need to be organised with consideration to data's properties: combinability, interpretability, and boundless reusability.* Water can leak from pipes, and electricity can overload grids [2], while data could be misunderstood or abused [6,34]. This lesson is based on the need to interpret data [12] and the data's heterogeneous nature and ability to be combined.

The fourth lesson is that *OGD systems need governance organisations that govern the whole system and not only the publishers and their interactions with users towards being beneficial for the public, data providers, and data users.* OGD systems are based on principles [e.g., 11,23,32,45], licenses [e.g., 32], laws [e.g., 36,42,47,48], and feedback loops [e.g., 18,52] that are currently dispersed over different actors. Governance organisations can centralise the responsibility to enforce principles, licenses, laws, and feedback and oversee actors too, for example, help with trust and mitigate risks.

The fifth lesson is that *OGD systems could be a replacement of existing data provision systems and become a public utility if it overcomes problems from the OGD principles, the properties of data, and the lack of public interest.* Previous OGD research has viewed OGD systems from many perspectives, such as

ecosystem [e.g., 18], lifecycle [e.g., 21], and value network perspectives [e.g., 3]. If OGD systems are made into a public utility, it can include all of these perspectives. Ecosystem for the nodes and their interactions. Lifecycle for a node or two nodes and the interactions. Value network for the node-flow view. The public utility perspective also includes the notion of the public and empathise that the OGD systems exist to benefit the public. The public utility perspective is one approach that can help to synthesise a holistic perspective of OGD systems.

6 Conclusion

The objective of this study was to explore what we can learn from a public utility perspective when perceiving and organising OGD systems. The study used a hermeneutic literature review with a snowballing approach to identify 39 studies about the two topics. The topics were compared based on five comparative categories, which resulted in five lessons (see Sects. 4 and 5). The review covers key literature but recognises that there is more research about OGD systems, data properties, and public utilities to be included. Not every identified comparative category was used in the comparison, such as economics (income and costs) and system purpose. The economic side of public utilities [38,46] is a promising avenue of future research that could help solve issues with economic sustainability for OGD systems [27]. At the same time, governance was identified in the public utility literature, but sparse in the OGD literature.

This paper contributes by showing that OGD research can learn for public utilities and has three implications for OGD research and OGD practice. First, OGD systems can be perceived from a node-flow view. This view contributes to the development of a holistic perspective of OGD systems that can help OGD researchers with model building, analysis, and data collection. OGD practitioners should work to identify and organise nodes (e.g., data collection or data processing) and match the OGD system with the needs of the public.

Second, governance and the properties of data needs consideration in the organisation of OGD systems. OGD researchers can study how to govern OGD systems and the implications of the properties of data on the organisation of the OGD systems. OGD practitioners should consider the properties of data in their work and the establishment of national governance organisations.

Finally, OGD systems could replace existing data provision systems and be made into a public utility. OGD researchers and OGD practitioners should discuss the role of OGD in everyday life of the public, which may challenge the OGD principles and underlying assumptions. The governance of OGD systems is a promising avenue of future research that could help practitioners deliver value, mitigate risks, and work together. The comparison between OGD systems and public utilities has also opened for the discussion and research into if OGD systems should be made into or is a type of public utility.

References

1. Abella, A., Ortiz-de Urbina-Criado, M., De-Pablos-Heredero, C.: The process of open data publication and reuse. J. Assoc. Inf. Sci. Technol. **70**(3), 296–300 (2019)
2. Ascher, K., Marech, W.: The Works. The Penguin Group, New York (2005)
3. Attard, J., Orlandi, F., Auer, S.: Data value networks: enabling a new data ecosystem. In: 2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI), pp. 453–456. IEEE (2016)
4. Attard, J., Orlandi, F., Scerri, S., Auer, S.: A systematic review of open government data initiatives. Gov. Inf. Q. **32**(4), 399–418 (2015)
5. Ayre, L.B., Craner, J.: Open data: what it is and why you should care. Public Libr. Q. **36**(2), 173–184 (2017)
6. Barry, E., Bannister, F.: Barriers to open data release: a view from the top. Inf. Polity **19**(1–2), 129–152 (2014)
7. Berners-Lee, T.: Star open data. 5 Star Data (5). https://5stardata.info/en/
8. Boell, S.K., Cecez-Kecmanovic, D.: A hermeneutic approach for conducting literature reviews and literature searches. CAIS **34**, 12 (2014)
9. Buchanan, R.A.: Public utilities. In: McNeil, I. (ed.) An encyclopedia of the History of Technology, Chapter 20, pp. 949–966. Routledge, London (1990)
10. Charalabidis, Y., Zuiderwijk, A., Alexopoulos, C., Janssen, M., Höchtl, J., Ferro, E.: The World of Open Data. Springer, Heidelberg (2018). https://doi.org/10.1007/978-3-319-90850-2
11. The Open Data Charter: International open data charter - principles (2015). https://opendatacharter.net/principles/
12. Checkland, P., Holwell, S.: Information, Systems and Information Systems: Making Sense of the Field, vol. 1. Wiley, Chichester (1998)
13. Colpaert, P., Joye, S., Mechant, P., Mannens, E., Van de Walle, R.: The 5 stars of open data portals. In: 7th International Conference on Methodologies, Technologies and Tools enabling e-Government (MeTTeG) (2013)
14. Crandall, R.W., Ellig, J.: Economic deregulation and customer choice: lessons for the electric industry. George Mason University, Center for Market Processes (4084 University Dr.) (1997)
15. Crusoe, J.: Why is it so challenging to cultivate open government data? - Understanding impediments from an ecosystem perspective. Ph.D. thesis, Linköping University, LiU-Tryck, Linköping, Sweden (April 2019). compilation licentiate thesis
16. Davies, T.: Open data, democracy and public sector reform. A look at open government data use from data. gov.uk (2010)
17. Davies, T.: Open data: infrastructures and ecosystems. Open Data Research, pp. 1–6 (2011)
18. Dawes, S.S., Vidiasova, L., Parkhimovich, O.: Planning and designing open government data programs: an ecosystem approach. Gov. Inf. Q. **33**(1), 15–27 (2016)
19. De Beer, J.: Ownership of Open Data: Governance Options for Agriculture and Nutrition. Global Open Data for Agriculture and Nutrition, Wallingford (2016)
20. Serra, L.E.C.: The mapping, selecting and opening of data. Rec. Manag. J. **24**(2), 87–98 (2014)
21. Folmer, E., Reuvers, M., Quak, W., van den Broek, T., van Veenstra, A.F.: Walking the extra byte: a lifecycle model for linked open data. In: Folmer, E., Reuvers, M., Quak, W. (eds.) Linked Open Data-Pilot Linked Open Data Nederland, pp. 95–111, Remwerk (2011)
22. Geddes, R.: Public utilities. The Encyclopedia of Law and Economics (1998)

23. Handbook, O.D.: What is open data? (2015). http://opendatahandbook.org/guide/en/what-is-open-data/
24. Harrison, T.M., Pardo, T.A., Cook, M.: Creating open government ecosystems: a research and development agenda. Future Internet **4**(4), 900–928 (2012)
25. Heimstädt, M., Saunderson, F., Heath, T.: Conceptualizing open data ecosystems: a timeline analysis of open data development in the UK. In: CeDEM14: Conference for E-Democracy an Open Government, p. 245. MV-Verlag (2014)
26. Hellberg, A.S., Hedström, K.: The story of the sixth myth of open data and open government. Transform. Gov.: People Process Policy **9**(1), 35–51 (2015)
27. Hossain, M.A., Dwivedi, Y.K., Rana, N.P.: State-of-the-art in open data research: insights from existing literature and a research agenda. J. Organ. Comput. Electron. Commer. **26**(1–2), 14–40 (2016)
28. Janssen, M., Zuiderwijk, A.: Infomediary business models for connecting open data providers and users. Soc. Sci. Comput. Rev. **32**(5), 694–711 (2014)
29. Jetzek, T., Avital, M., Bjorn-Andersen, N.: Data-driven innovation through open government data. J. Theor. Appl. Electron. Commer. Res. **9**(2), 100–120 (2014)
30. Kalampokis, E., Tambouris, E., Tarabanis, K.: A classification scheme for open government data: towards linking decentralised data. Int. J. Web Eng. Technol. **6**(3), 266 (2011)
31. Keohane, R.O., Nye, J.S., Donahue, J.D.: Governance in a globalizing world. Power and Governance in a Partially Globalized World, pp. 193–218 (2002)
32. Knowledge, O.: The open definition (2015). https://opendefinition.org/
33. Kucera, J., Chlapek, D.: Benefits and risks of open government data. J. Syst. Integr. **5**(1), 30–41 (2014)
34. Kulk, S., Van Loenen, B.: Brave new open data world? Int. J. Spat. Data Infrastruct. Res. **7**, 196–206 (2012)
35. Kurose, J.F., Ross, K.W.: Computer Networking: A Top-Down Approach. Pearson International Edition, 5th edn. Addison Wesley, Boston (2009)
36. Lee, D.: Building an open data ecosystem: an Irish experience. In: Proceedings of the 8th International Conference on Theory and Practice of Electronic Governance, pp. 351–360 (2014)
37. Lindman, J., Kinnari, T., Rossi, M.: Business roles in the emerging open-data ecosystem. IEEE Softw. **33**(5), 54–59 (2015)
38. McNabb, D.E.: Public Utilities: Old Problems, New Challenges. Edward Elgar Publishing, Cheltenham (2016)
39. Oliveira, M.I.S., Lóscio, B.F.: What is a data ecosystem? In: Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age, pp. 1–9 (2018)
40. Pickett, S.T., Cadenasso, M.L.: The ecosystem as a multidimensional concept: meaning, model, and metaphor. Ecosystems **5**(1), 1–10 (2002)
41. Rose, R., Mackenzie, W.J.M.: Comparing forms of comparative analysis. Polit. Stud. **39**(3), 446–462 (1991)
42. de Rosnay, M.D., Janssen, K.: Legal and institutional challenges for opening data across public sectors: towards common policy solutions. J. Theor. Appl. Electron. Commer. Res. **9**(3), 1–14 (2014)
43. Safarov, I., Meijer, A., Grimmelikhuijsen, S.: Utilization of open government data: a systematic literature review of types, conditions, effects and users. Inf. Polity **22**(1), 1–24 (2017)
44. System: The Cambridge Dictionary. Cambridge University Press (2020). https://dictionary.cambridge.org/dictionary/english/system. Accessed 26 Feb 2019

45. Tauberer, J.: The 8 principles of open government data (2007). https://opengovdata.org/
46. Thompson, C., Smith, W.R.: Public Utility Economics. McGraw-Hill Book Company, New York (1941)
47. Ubaldi, B.: Open government data - towards empirical analysis of open government data initiatives. Tech. rep. 22, OECD Working Papers on Public Governance (2013)
48. Van Schalkwyk, F., Willmers, M., McNaughton, M.: Viscous open data: the roles of intermediaries in an open data ecosystem. J. Inf. Technol. Dev. **1102**(Special Issue), 68–83 (2015)
49. Walters, W.H.: Comparative recall and precision of simple and expert searches in google scholar and eight other databases. Portal: Libr. Acad. **11**(4), 971–1006 (2011)
50. Welle Donker, F., van Loenen, B.: How to assess the success of the open data ecosystem? Int. J. Digit. Earth **10**(3), 284–306 (2017)
51. Zuiderwijk, A., Janssen, M.: A coordination theory perspective to improve the use of open data in policy-making. In: Wimmer, M.A., Janssen, M., Scholl, H.J. (eds.) EGOV 2013. LNCS, vol. 8074, pp. 38–49. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40358-3_4
52. Zuiderwijk, A., Janssen, M., Davis, C.: Innovation with open data: essential elements of open data ecosystems. Inf. Polity **19**(1–2), 17–33 (2014)