

This is the accepted version of the following article

Martin Lnenicka, Anastasija Nikiforova (2021). Transparency-by-design: What is the role of open data portals?. *Telematics and Informatics*. DOI: 10.1016/j.tele.2021.101605

Publisher's version is available from:

<https://www.sciencedirect.com/science/article/pii/S0736585321000447>



This version is licenced under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Transparency-by-design: what is the role of open data portals?

Martin Lnenicka and Anastasija Nikiforova

Abstract: *Transparency in the public sector is one of the most important topics of the current debates on accountable, participatory, and responsive governance. An open government addresses these major topics and aims to encourage the relationships and flows of information between involved stakeholders. This article explores the role of open data portals in supporting these efforts and provides findings regarding the features in the design of these data infrastructures. On the basis of evidence from the concept of transparency-by-design, we argue that transparency is facilitated by open data portals and their features enabling to work with datasets. We therefore propose the list of the categories and corresponding features of open data portals that should constitute the checklist of the portal aiming to achieve the highest level of transparency. The mapping of existing features found in literature to the phases of the transparency cycle demonstrates that open data portals meet the transparency requirements.*

Keywords: *transparency; open government data; open data portal; design; features; transparency cycle*

1 Introduction

Transparency in the public sector is defined as the availability of information about an organization or actor that allows external actors to monitor the internal performance of that organization (Grimmelikhuisen and Meijer, 2012). According to Wirtz and Birkmeyer (2015), open government can be defined as “*a multilateral, political, and social process, which includes in particular transparent, collaborative, and participatory actions by government and administration.*” In view of this, Janssen et al. (2017) proposed the concept of “transparency-by-design”, according to which transparency should be considered as the main requirement for the development of the system aiming to disclose government data to the public. As reported by Meijer (2009), transparency that can be achieved through computerized systems that are available to almost everyone today should improve governments around the world by reducing corruption and at the same time improving accountability to citizens. By taking these concepts into account, modern Information and Communication Technologies (ICT) provide the resources to fulfill these goals.

The impact of ICT on public values, contributing to enhancing transparency, has been recognized since computers first appeared in public agencies and institutions. This was further boosted by the rapid growth of Internet access for most people in developed countries, making the supply of information more competitive, thereby making it significantly easier to disseminate information on corruption (Goel et al., 2012). Nowadays, global technological developments heavily rely on data, and the use of ICT in the public sector can benefit in various ways (Lourenço et al., 2017). From 2010 onward, open data have been related to wider open government reforms (Noveck, 2017), which in turn have tied open data to public management reforms (Clarke and Margetts, 2014). Open Government Data (OGD) became one of the key concepts shaping government efforts to be more transparent and accountable about the use of public resources (Attard et al., 2015; Kimball, 2011; Noveck, 2017; Safarov et al., 2017; Lourenço, 2015). Besides that, opening up government data can also ensure higher returns on public investment, while facilitating policy makers to address difficult challenges, improve public policies and the efficiency and quality of public services (Bogdanović-Dinić et al.,

2014), e.g. healthcare services (Martin and Begany, 2017), including facilitating the movement towards a smart city (Pereira et al., 2017).

Open data portals today play the role of an interface that creates transparency (Janssen et al., 2017). In the past few years, various authors have reported that the release of OGD on open data portals affects processes that have contribution to transparency in areas where corruption, wastage and inefficiency take place the most (Hogan et al., 2017; Lourenço, 2015; Murillo, 2015). For instance, Van Schalkwyk and Cañares (2019) provided an overview of the anti-corruption projects taking place in Malaysia, Moldova, Kosovo, Indonesia, Africa etc. highlighting the efficiency of open data.

However, to provide these opportunities, open data portals should offer users a wide range of mechanisms to enable them to effectively discover, extract and use data (Klein et al., 2018). Therefore, the perspective of transparency achieved through data portals was also analyzed by Lourenço (2015), Klein et al. (2018) and Thorsby et al. (2017). However, according to Janssen et al. (2017), greater attention is needed to be paid to what is the transparency promoted by ICT and how it can be achieved. For this purpose, they introduced a transparency cycle consisting of 6 phases and respective steps, which should be considered when designing ICT systems to ensure that data are disclosed to the public for creating transparency. In this regard, considering the lack of research studies focusing on setting transparency requirements and meeting them when designing open data portals, the aim of our study is to explore the role of open data portals to identify research priorities and best practices in their design. We define 8 categories which contains a list of features found in literature, which are then mapped onto phases of the transparency cycle to demonstrate that open data portals comply with the requirements of transparency.

The article is structured as follows: the following section describes the research methodology as well as the research questions. This is followed by exploring the effects of OGD and the role of open data portals in increasing transparency. The emerging research priorities dealing with features in the design of open data portals are then defined together with their mapping onto phases of the transparency cycle. Finally, conclusions summarizing the lessons learned are provided.

2 Research methodology

Under the above aim, the following research questions (RQ) are proposed:

RQ1: What are the relations between OGD, open data portals, and transparency found in the literature?

RQ2: How do the relevant articles address the role of open data portals in transparency?

RQ3: What are the research priorities and best practices in the design of open data portals?

We chose to use a methodology that supposes a research based on the review of the relevant literature. The literature review process aims to critically assess and determine priority areas and best practices in terms of transparency-by-design of open data portals. Considering the typology of literature reviews developed by Paré et al. (2015), our review is a critical review, since our goal is to draw attention constructively to weaknesses of prior studies and strengthen the development of knowledge and the direction towards studies for further improvement.

In order to answer the RQ1 a systematic search of digital libraries such as ACM Digital Library, ScienceDirect, Web of Science, IEEE Explore and Springer has been carried out. These libraries seem to be leading and should provide an overview on the popularity of the topics addressed over the years. The search was conducted by querying on these databases keywords and key phrases such as (1) “open government data”, (2) “open government data” AND “portal”, (3) “open government data”

AND “portal” AND “transparency”. These search terms were applied to both paper titles, abstracts, and bodies, which means that each paper containing the relevant key phrases was found and included in the following analysis. Only papers written in English were addressed. In terms of the scope, both journal articles, conference papers, and chapters were studied.

As regards the period covered by these searches, we covered the period 2009-2021 to get an insight of the tendencies and growth of the popularity of these topics over the years and to select the most up-to-date studies on these topics to be further analyzed. The choice of the starting date allows to focus on these topics and their popularity from the beginning of the concept of OGD, when only a few studies on these concepts have been carried out. The results of the analysis of 5 digital libraries for 2009-2021 are summarized and provided in Table 1. It can be seen that in 2009 all three topics were new and that only a few studies were carried out on them. However, from 2010 the topic of “open government data” is becoming more popular and, as a result, both related topics attracted researchers from 2011 to today. In terms of overall results for 5 libraries, the popularity of all three topics increased steadily, with a slight decrease in 2015 and 2017 in “open government data” AND “portal” subject, while the subject of transparency increased in the number of papers devoted, or at least mentioning it during this period without any exceptions, which make sure once again the importance of this topic.

Table 1: The results of the comparative analysis of 5 digital libraries

Digital Library	Key phrase	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
ACM Digital Library	<i>Open government data</i>	1	13	32	45	72	76	68	86	70	86	51	61	1
	<i>open government data + portal</i>	1	1	8	14	32	37	35	43	39	35	26	23	0
	<i>open government data + portal + transparency</i>	0	0	1	8	11	14	11	18	19	25	17	18	0
Science Direct	<i>Open government data</i>	1	2	5	9	5	25	31	28	37	43	77	63	31
	<i>open government data + portal</i>	1	0	4	3	3	15	11	14	14	18	32	23	11
	<i>open government data + portal + transparency</i>	1	0	2	3	2	10	11	11	12	12	23	16	9
Web of Science	<i>Open government data</i>	0	1	13	11	19	40	37	75	78	101	95	55	1
	<i>open government data + portal</i>	0	0	3	2	4	4	1	8	10	21	8	3	0
	<i>open government data + portal + transparency</i>	0	0	1	1	1	0	0	4	3	16	2	1	0
IEEE Xplore	<i>Open government data</i>	0	2	4	6	8	11	7	17	15	14	18	10	0
	<i>open government data + portal</i>	0	0	1	1	3	3	1	10	6	3	7	3	0
	<i>open government data + portal + transparency</i>	0	0	0	0	0	2	1	6	0	1	3	1	0
Springer	<i>Open government data</i>	0	7	38	36	64	60	70	89	94	143	164	164	28
	<i>open government data + portal</i>	0	3	21	19	26	32	41	49	50	85	90	87	18
	<i>open government data + portal + transparency</i>	0	3	17	9	17	14	29	32	38	64	56	55	11

Figure 1, however, provides (1) the results of the comparative analysis for the 2009-2019 period, presenting the state of the art regarding the number of studies for each subject and (2) the trend line, demonstrating not only how this situation is changing during these years, but also how it will change over the next 6 years. In addition, this trend line demonstrates that transparency will be the topic which popularity increase less rapidly compared to OGD and data portals. The exclusion of 2020 and 2021 was needed because this analysis was carried out in mid-January 2021, when 2019 appears to be the most recent year for which the total number of articles is already available, as not all libraries have published all results for 2020. This means that the inclusion of 2020 and 2021 would not allow for an accurate trend line for the coming years. Therefore, 2019 was chosen as the most recent year to be analyzed to figure out the popularity trends. This means an analysis of the 11 years is conducted showing how research trends have been or have not changed over time.

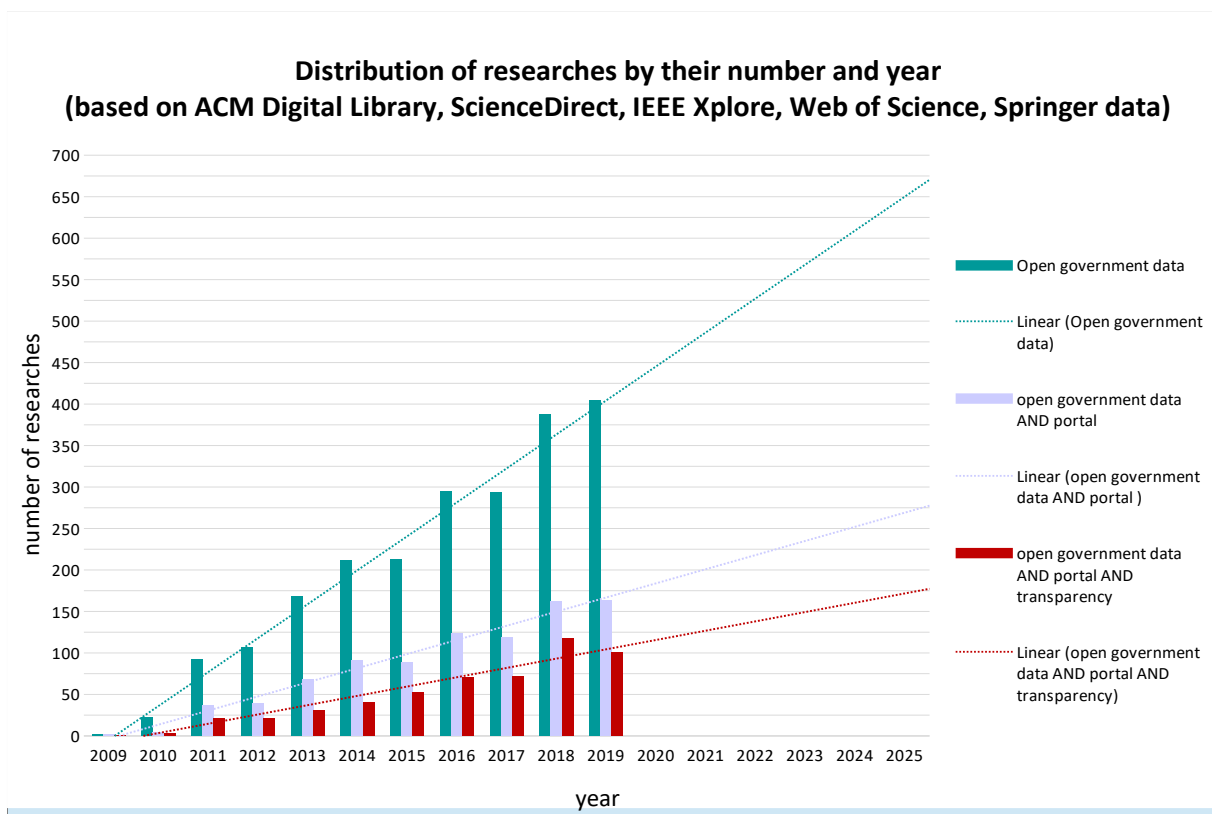


Figure 1: Results of the systematic search between 2009 and 2019 [2025]

Comparing the popularity of these subjects, our analysis shows that the number of studies focused on the transparency of OGD in portals is an average 21.4% of all studies focused on OGD. This result varies from 6% for Web of Science to 36% for Springer with the highest number of papers addressing transparency. Moreover, the number of studies focusing on transparency compared to the study on OGD portals exceeds 54.3%, with 75% for Science Direct and 67% for Springer, which means that transparency and OGD portals are closely related concepts. The same analysis conducted on Google Scholar demonstrates that these ratios in it are even higher.

Although transparency is mentioned relatively often in the literature, i.e., in 21.4% studies devoted to OGD, and the term “transparency” is closely related to the OGD portals, our analysis leads to the conclusion that this trend is growing less rapidly compared to other topics, despite being one of the most crucial aspects affecting both success of the OGD initiatives and OGD portals. One prevailing explanation of this negative trend could be the lack of comprehensive studies on transparency and its relationship with OGD portals. The mention of the term of transparency in relevant studies often

reduces to its recognition as one of the benefits of OGD and OGD portals without understanding how to achieve it or improve it. This study deals with this issue.

Taking these findings together, we can conclude on the importance of exploring the relations between OGD, open data portals, and transparency, and consider RQ1 to be answered.

3 The role of open data portals in transparency

In order to answer the second RQ2 – “*How do the relevant articles address the role of open data portals in transparency?*”, the results obtained during the systematic literature review have been adapted by selecting studies for their further detailed analysis. In addition, several articles were included, taking into account the recommendations of experts on various conferences. Therefore, this section provides a brief overview of the most relevant studies covering (1) basic concepts of public sector information and OGD, (2) ways of achieving transparency through public data, and (3) open data portals and their importance. The main points identified will be the RQ3 input data.

3.1 The core concepts behind public sector information and open government data

The topic of public sector information disclosure through a variety of information and communication channels has generated a lot of interest among practitioners in the public sector (Andrades et al., 2019; Geiger and von Lucke, 2012; Glassey, 2017; Janssen et al., 2012; Sieber and Johnson, 2015; Zuiderwijk and Janssen, 2015; Zhang et al., 2017). As access to information is the first step for citizens to understand what the government is doing, many countries around the world, more precisely 119 countries, have implemented some form of **Freedom of Information (FOI)** legislation allowing their citizens to request public sector information (Machado et al., 2018; Noveck, 2017). This is considered a mechanism contributing to anti-corruption efforts by enhancing stakeholders’ engagement (Bannister and Connolly, 2014; Noveck, 2017; Mendoza and Cruz, 2020). As reported by Janssen et al. (2012) and Sieber and Johnson (2015), FOI and e-government have been partially reviewed as an open government and OGD, with governments publishing online publicly available datasets.

According to Noveck (2017), although both the FOI and open data are related to the sharing with information, the nature of open data is based on **participation**. By promoting public engagement, open data transform monitoring relationships between governments and citizens to collaborative one, enabling the transformation of raw data into value and knowledge for the society. In 2003, the European Union (EU) adopted legislation to foster the reuse of public sector information in Member States via the Directive 2003/98/EC. It was amended by Directive 2013/37/EU in 2013 to make it more aligned with the concepts of the OGD and to encourage the development of a healthy open data ecosystem in the EU (Carrara et al., 2015).

Public sector agencies and institutions produce huge amounts of data in different formats (Carrara et al., 2015; Jetzek et al., 2014a; Kucera and Chlapek, 2014; Máchová and Lněnička, 2017). If these data can be freely used, modified, and shared by anyone for any purpose, they may be referred to as OGD. Geiger and von Lucke (2012) offers a more specific definition, according to which the OGD are all public-sector data, which the government makes available and accessible to the public without any restrictions on their use and distribution. This view is also shared by Sayogo et al. (2014), conceptualizing the definition of the OGD, pointing to the public publication of data collected and stored by national authorities in agreement with the legislation to make them available to the public **on a single data portal**.

These data should be accessible in both human- and machine-readable formats, allowing the combination and use of data in different ways (Attard et al., 2015; Geiger and von Lucke, 2012; Janssen et

al., 2012; Zuiderwijk et al., 2015). However, not all data may be published in the same way as they were stored and processed with the data owner, because confidential and classified data are not suitable for their publication (Attard et al., 2015; Jetzek et al., 2014a). These data should be excluded, while they may only remain in the datasets provided in the case if one or more data anonymizing techniques were used for the dataset. Other data should not be opened without any doubt due to security, privacy, or other reasons. The above-mentioned requirements are also in line with a set of **principles of OGD** introduced in 2007. As reported by Zuiderwijk and Janssen (2015), views and decision on which data should be opened or closed may change over time. This depends to a large extent on what kind of OGD will encourage civil engagement and government transparency (Gebre and Morales, 2020).

Open government and open data initiatives and policies are then based on OGD principles aimed at ensuring the long-term transparency of public information (Bertot et al., 2010; Kimball, 2011). Attard et al. (2016) identified the dimensions affecting or likely to be affected by value creation from the OGD, more precisely access to information, transparency, accountability, and democratic governance. Therefore, an open government movement can be understood as an extension or complement to FOI enabled by modern ICT (Eckartz et al., 2014). The main contribution of this movement to FOI is that the response to a request for information can be satisfied by disclosing this information on an open data portal where it is available for everyone (Máchová and Lněnička, 2017). As a result, Nilssen's (2010) observation showed that digital data have affected two important data features: 1) if data are readily accessible to more than one person at the same time, they shall acquire a non-competitive feature and; 2) costs arising from re-production and dissemination are substantially reduced, thereby becoming economically feasible for the reuse of data. However, the OGD initiatives enhance transparency and prevent corruption by helping stakeholders to be better informed about how their country is governed (Attard et al., 2016). Enhancing the involvement of citizen in the OGD initiatives would further encourage professionals to reuse datasets under different circumstances (Saxena and Muhammad, 2018).

On the other hand, the boost of modern ICT and data volumes poses **challenges** in terms of data infrastructure and management (Zuiderwijk et al., 2013). As public-sector systems are fragmented and data are not correctly organized and arranged, it is hard to find data and understand their value (Janssen et al., 2017). The number of intermediaries with different aims and characteristics in the domain of governmental transparency are also increasing (Glasse, 2017). Similarly, Ingrams (2017) reported that information disclosure is also accompanied by budgetary pressures, legal constraints etc., thereby creating another layer of complexity for transparency initiatives. According to Lourenço et al. (2017), focusing on the technical aspects of OGD, rather than on the goal of their release, may cause a loss of their accountability related meaning.

3.2 How to achieve transparency and other effects through open government data

Data transparency refers to the opening of public sector information and enabling the public to access data held by the government in a uniform way (Veljković et al., 2014). The disclosure of OGD, however, is a complex activity. It is increasingly recognized that it can bring numerous benefits as well as encounter risks (Charalabidis et al., 2016; Janssen et al., 2012; Kucera and Chlapek, 2014; Sieber and Johnson, 2015; Zuiderwijk and Janssen, 2015; Zuiderwijk et al., 2015; Ruijter et al., 2020). Most of these authors conducted various interviews, group sessions, and workshops with experts or even public officials to confirm or disprove these arguments. As a result, there are various lists of benefits, risks, barriers, and challenges related to the topic of OGD. In general, the benefits may be financial,

or they may be non-financial, for instance by complying with legislative and other transparency and accountability requirements.

3.2.1 Definitions and contexts

Since there are many definitions of transparency, it is difficult to achieve a common understanding of what transparency is in practice. According to Janssen et al. (2017), the main reason is the relative nature of transparency and its dependence on the views of stakeholders and the context. What one interested stakeholder may consider being transparency may not be in the view of another. Regarding the context of OGD, transparency concerns the absence of asymmetric information while it is seen as a public value (Bannister and Connolly, 2014). The transparency mechanism is also based on the network as it depends on the different cases of information exchange and receipt between stakeholders (Jetzek et al., 2014a; Jetzek et al., 2014b). Tkacz (2012) claims that modern openness and transparency are evidence of the networked and computational, even cybernetic, nature of governance. In this regard, the availability of information about “what governments do and why?” is recognized as an important prerequisite for governmental transparency and accountability (Janssen et al., 2012). On the other hand, it should be noted that while OGD relates to government transparency, the transparency of government data cannot be seen as an unambiguous factor in openness, since the data that were open may not be transparent (Bogdanović-Dinić et al., 2014).

3.2.2 Enabling processes and mechanisms

The private sector plays a critical role in this respect, as they are the biggest users of ODG (Attard et al., 2016; Carrara et al., 2015; Jetzek et al., 2014b). According to Janssen et al. (2012), economic growth and investment in societal values driven by transparency and accountability are the main incentive arguments for OGD. Rajshree and Srivastava (2012) argued that OGD can help combat corruption through new applications promoting transparency in public services. The interests of the private sector promote the use of open data portal datasets, where stakeholder can not only study the OGD available to them, but also create value from them by providing relevant feedback and cooperate more actively with the government (Thorsby et al., 2017; Attard et al., 2015; Attard et al., 2016).

Given the importance of cooperation, including collaboration between organizations from different sectors, Susha et al. (2017) defined data collaboratives as a cross-sectoral, i.e., public and private, cooperation initiative aimed at collecting, sharing and processing data to address the societal challenge. In this regard, Hivon and Titah (2017) reported an example from Montréal. Instead of simply complaining about corruption, citizens invited programmers, journalists, civil society groups and others, developing creative solutions to combat it. A similar solution was proposed in Arauz et al. (2017), i.e., a digital platform aimed at combating corruption in the urban development process, enabling citizens to use open data, thereby also eliminating citizen’s discretion in these processes.

The study proposed by Klein et al. (2018) carried out an identification of 18 mechanisms that can increase transparency of Brazil’s open data portals. In similar lines, Thorsby et al. (2017) explored the content and features of American urban data portals, as well as analysis of some common models, providing open data. Focusing on promoting transparency and engagement processes between stakeholders through the OGD, Hogan et al. (2017) described the approach to setting up the platform for this purpose. Ruijter et al. (2017b), however, introduced a variety of user requirements that should be met by platform design to promote these processes.

Other impact of OGD is cost savings by reducing administrative overhead in the dissemination of data to those already requesting these data and the ability of governments to cut costs through the

development of an outsourcing application for other stakeholders (Sieber and Johnson, 2015). On the other hand, Carrara et al. (2015) claims that a number of studies and case examples have shown that cost-based price models do not generate cost savings for public authorities in the long term, while free or marginal cost models are more favorable. Sieber and Johnson (2015) argued that providing the data could lead to greater participation in government functioning and support for decisions. Thus, the dissemination of information will also help stakeholders to take informed data-based decision (Attard et al., 2016; Safarov et al., 2017). In this regard, Zuiderwijk and Janssen (2015) provided a decision-making model that helps to compromise between OGD countermeasures. An approach supporting proactive transparency based on the provision of information as OGD was proposed by Glassey (2017). The impacts of OGD enabling factors and generating mechanisms were evaluated by Máchová and Lněnička (2016) and suggested that the greatest impact could be achieved in the educational and social development.

3.2.3 Challenges and limitations

However, while it is a well-known view that the OGD is a useful tool for combating corruption, Žuffová (2020) argues that OGD alone are not enough to deal with such a complex problem, where media and internet freedom play a greater role in changing the level of corruption as the most crucial part of transparency policy. The author also points out that countries with higher levels of media and internet freedom are characterized by a higher level of transparency which is linked to the decline in corruption in this respect. This opinion is supported by Esteban (2019), who highlights the role of social media as sources of information, pointing to the fact that it is undervalued, given that today, the largest social-media platforms unify billions of monthly active users.

There are also authors who are skeptical of the benefits of disclosing OGD and their role in increasing transparency. While data opening gives rise to the assumption that more transparency is being created, often essential details are hidden or not opened (Janssen et al., 2017). A set of interviews conducted by Barry and Bannister (2014) revealed that transparency is not always a priority of opening data. Some authors also pointed out that there is a pressure on governments to release much of the data they hold. This pressure goes beyond the public's right to know or FOI and is determined by the views on the economic, social, administrative, and political benefits of accessing such data (Barry and Bannister, 2014). Similarly, Lourenço et al. (2017) reported that the theory of public accountability for the open government has become overshadowed by the prospects of social and economic value. Janssen et al. (2012) revealed that the barriers involved both data providers that do not want to make their data public and data users who tend to not be able to easily use the data. Problems related to access, skills, and data literacy as well as the gap between technology and administrative practices were identified by Gebre and Morales (2020) and Glassey (2017). Murillo (2015) explored the formal online release of governmental data and their contribution to the government economic and institutional transparency. They concluded that modern ICT did not meet expectations regarding their contribution to government transparency and lessen knowledge divides in the areas assessed. Recent studies have also shown that transparency depends not only on how visible information is, but also on how understandable it is for users (Attard et al., 2015; Bittencourt et al., 2019; Cordasco et al., 2018; Degbelo et al., 2019; Klein et al., 2018; Ojo et al., 2016).

3.3 Open data portals and their importance

The role of ICT and the delivery of information to the public through the Internet has evolved since the beginning of this century. At first, various government websites provided basic information on citizens' rights and responsibilities, office hours, public services, or public officials. Various forms

and instructions could then be downloaded, and online public services became more accessible (Attard et al., 2015). Currently, various data portals are being launched around the world (Hogan et al., 2017; Máchová and Lněnička, 2017; Nikiforova et al., 2021; Ruijer et al., 2017b). The advantage of those data infrastructures is that they make it possible to combine data from many different organizations, on different administrative levels as well as between different types of data (Klievink et al., 2017; Zuiderwijk et al., 2013).

3.3.1 Requirements and features

The concept of **open data portal** is defined as a software platform that includes a software ecosystem that supports the interactions of different users with open data, including searching and detecting datasets, publishing datasets, analyzing and visualizing them, and sharing and development datasets stories (Ojo et al., 2016). Open data portals contain information about datasets in the form of metadata and allow datasets to be accessed, downloaded, and uploaded.

Since all datasets are organized into a searchable catalogue, searching for the relevant datasets or browsing of the datasets by categories, tags, organizations, or formats are **key features** of each portal (Alexopoulos et al., 2017; Chatfield and Reddick, 2017; Lourenço, 2015; Máchová and Lněnička, 2017). Other features include various statistics about datasets, such as a total number of downloads or views (visits) or lists of applications created with a certain dataset. As a recent trend, features allowing stakeholders to provide feedback through ratings or comments and to request open data from data providers are also sometimes available on selected portals (Máchová et al., 2018). However, while most of these features seems to be a classical set of requirements that must be met by every open data portal, recent studies demonstrated that most of them are challenging for some portals, see Máchová et al. (2018) and Nikiforova et al. (2021). On the other hand, almost every portal has room for improvements in terms of these features.

However, there are also positive examples demonstrating additional supportive facilities. For instance, in order to help stakeholders to investigate and link datasets, some portals provide built-in tools to extract, transform, visualize, and mash-up data (Attard et al., 2015; Charalabidis et al., 2014; Máchová and Lněnička, 2017). More precisely, Arauz et al. (2017) stated that the key benefit of these platforms is the provision of tools and features enabling analytical processes to be carried out using open data to improve decisions. According to Klievink et al. (2017), these infrastructures should support the provision of data in appropriate data formats and the consumption of data via easily accessible channels to become a pillar for value creation and transparency efforts. Cataloguing of OGD enables public sector agencies and institutions to better understand what data they have and how to better manage them (Kucera and Chlapek, 2014). Finally, opening and publishing these data in open data portals with open licenses saves costs because it is cheaper than converting them into reports and applications (Janssen et al., 2012).

3.3.2 Enabling processes and mechanisms

Here we would like to draw attention to the Ruijer et al. (2017a) study, which analyzed the role of open data portal in supporting the complexity of such democratic processes as monitoring, consulting, and participatory democracy. Their study used a scenario-based design methodology to explore specific contexts of open data usage. Different user requirements in efforts to develop open data portals that meet the needs of citizens were identified. Their study also found that a different approach to the design of the open data portal should be chosen for each type of democratic process. In their follow-up article, Ruijer et al. (2017b) focused on determining of information and social interaction that would solve a specific societal issue. They argued that users are attracted not only to provide

functional features, but they also require values and sense-making qualities such as data analysis and reporting tools. Lourenço (2015) established a set of requirements to analyze whether the current structure and organization of the selected open data portals are in line with the support transparency of accountability. The results showed that most of these portals do not contain crucial structural and organizational elements necessary to fully support citizens involved in transparency efforts. In this regard, open data portals should support both data and organizational transparency (Ojo et al., 2016). Here, as an example, Saxena and Muhammad (2018) examined the impact of OGD on accountability and transparency in terms of open data portals. Similarly, Gebre and Morales (2020) recognized the distinction between the use of data by internal experts and external users. They emphasized the use of interactive approaches and features and the need for increasing data literacy among stakeholders. A list of potential coordination mechanisms to be applied on open data portals to serve for transparency purposes was identified by Susha et al. (2017). These are: (a) determining potential data providers and data users, (b) maintaining control over data and their unexpected uses, (c) matching specific problem with data attributes, (d) ensuring the usability and usefulness of the data for the user, and (e) matching data provider incentives with user targets. Gebre and Morales (2020) provided insight into how to organize the open dataset to make resources available to the whole community and how consider users' perspectives for open data providers, including providing continuous support. However, Machado et al. (2018) proposed an instrument for assessing open data portals that gathers features of platforms supporting the creation of open data portals and recommendations from FOI access law. In the light of the above, it can be concluded that there are many studies dealing with the relationship between transparency and OGD, and how to strengthen them. But there is a lack of studies to develop a common understanding of their relationship in terms of OGD portals and features that contribute most to transparency. This point is addressed in the following sections.

4 The state-of-the-art and emerging research priorities dealing with features in the design of open data portals

Let us look in more detail at the studies that have attracted us most and the aspects that are emphasized in them as features for increasing transparency. In this regard, this study defines a feature as “*functionality or activity provided by the open data portal through which various stakeholders can be involved in working with datasets for their intended purpose.*” In order to answer the RQ3, the results obtained during the review of systematic literature have been adapted by selecting the most relevant studies for their further detailed research. The first requirement for the selection of the relevant articles was a limitation on the date of their release to address only current studies. This was done by selecting 2017 as the starting date, as the subject of OGD, and, in particular, open data portals are developing and changing so quickly that older articles can be considered as outdated/ obsolete, with the exception of significant contributions, focusing mainly on the frequency of their citation, that enriched the original list.

The open data portals are powerful tools that allow governments to achieve transparency, confidence and trust in government (Machado et al., 2018; Hogan et al., 2017). In addition, according to Hogan et al. (2017), this is achieved both (a) for those who are focused on monitoring government policy and its consequences, discussing them and shaping the policy decision-making process, and (b) for those directly involved in policy development and the creation of public values, using open data available on platforms. Janssen et al. (2017) stresses that the requirements applicable to transparency are met, taking into account the nature of the project, when the results of the design process meet these requirements. Therefore, the number and extent of these features constitute a unique ecosystem

that can be a useful tool that serves as an interface establishing systems that ensures effective use of OGD, allowing for real of transparency (Charalabidis et al., 2014; Lourenço, 2015; Machado et al., 2018; Saxena and Muhammad, 2018; Styryn et al., 2017).

To obtain the expected benefits of open data, they should be available and **findable** first. The data must be then **accessible** and **understandable**, as well as of **high quality**. However, even this is not enough because, even if the data are available, accessible and sufficiently qualitative, there are two more key factors that are capable to distract users from being interested and able to use published data, namely data **usefulness** and their **understandability**. In addition, statistical data, such as the total number of available datasets, categories, data publishers etc., should also be provided in order to let users understand that the portal is alive, and to provide an insight into its current state. And, of course, no less important is the presence of the **service** and its quality, establishing a two-way channel between users and the portal and ensuring citizen participation. On the one hand, this list could be considered to be defined on the basis of an intuitive approach, but in recent years a number of studies have been carried out in this area, which came to the conclusion that these factors are crucial and should constitute the checklist of the portal aiming to achieve the highest level of transparency. This list will serve as the starting point in defining the list of features to be ensured in terms of transparency-by-design.

Let us briefly address these categories and motivate our choice, mapping the defined categories to the “transparency cycle” defined by Janssen et al. (2017). According to the authors, the creation of transparency can be expressed through the following steps: (1) finding data, (2) opening-up all relevant data, (3) ensuring privacy, (4) equal access, (5) accessible formats, (6) right level of abstraction, (7) understanding data, (8) validation of conclusions, (9) social media use, (10) institutional arrangements, (11) follow-up actions. Given the fact that we focus on the open data portals, the sequence of categories examined will follow the sequence of actions performed by the users of the open data portals when using the portal. More accurate mapping of identified features in this cycle will follow in the next section. However, the portals are not launched without support in open data policies, strategies, and initiatives. Increased transparency is mostly their main objective and the open data portal is a tool to accomplish it. In addition, open data management platforms such as CKAN, DKAN, or Socrata together with relevant extensions enable to find relevant data across public sector registers and databases.

Data **findability** is one of the most popular but rarely noticeable barriers of open data portals. Ruijer et al. (2017a) revealed that, together with a user-friendly format, it is a second data access-related barrier by popularity. They found that users often admit that even when data are published, they can sometimes not be found. This was also the case in Ruijer et al. (2020) when living-lab participants experienced difficulties in finding and selecting relevant data out of wide range of the proposed datasets. While data findability is closely related to data accessibility, in terms of its nature we have separated this feature, addressing it independently as is also the case for (Barry and Bannister, 2014).

Data **accessibility** is considered as the premise of data utilization (Wang et al., 2018) reflecting the extent to which data are available, approachable, and convenient to reach. It is one of the primary open (government) data principles. It also proved to be one of the most important factors for open data portals (Wang et al., 2018; Ruijer et al., 2017b). As in Wang et al. (2018), considering its primacy, we separate the accessibility of data from other types of data quality. Accessibility refers to steps 4 and 5 in the transparency cycle (Janssen et al., 2017)

Open data must be **understandable** to citizens and other stakeholders (Lourenço, 2015). Moreover, they must be supplied with additional descriptive information, whose understandability is just as

important, thereby allowing the data to be reused, with a full comprehension of the entire dataset and each of its attribute. Here, the metadata and their quality come as well. The inclusion of metadata facilitates and accelerates the use of open data, enhancing user experience, which should contribute to the understanding and interpretation of datasets to be used by the public to create transparency (Janssen et al., 2017). Unfortunately, it appears that this category is also an obstacle (Gil-Garcia et al., 2020; Janssen et al., 2017), which open data portals have not overcome at present, thereby reducing the value of the open data published and their possible reuse. Data understandability refers to step 7 in the transparency cycle (Janssen et al., 2017). These three categories are also found to be key factors in different scenarios, such as collective accountability and collective participation (Ruijter et al., 2017b), creating a critical resource that promotes the use of open data portal and its data (Wirtz et al., 2018).

The **usefulness** of data is no less important (Purwanto et al., 2020; Gil-Garcia et al., 2020). As noted in Ruijter et al. (2017b), providing random datasets does not automatically make them useful or valuable for society. They argue that there may also be no match between released data and data in which their users are interested. This is also in line with the first stage of the transparency cycle proposed in Janssen et al. (2017). This means that it is very important to identify the most useful datasets and mechanisms, by which we are extending this category, which promotes and facilitates their usefulness and allowing to get an insight on *how they were useful for other users?* This extension is also in line with Gil-Garcia et al. (2020), where it is underlined, that it is important to ensure that not only the data by themselves but also their supplementary information is useful. According to our own observations, the datasets of higher importance for the society may vary significantly depending on the country and even the city. It also matches literature including Utamachant and Anutariya (2018), Folmer et al. (2019), Zheng and Gao (2016), Chakravarty (2018), Slobodova (2020), Lee and Kwak (2012).

One of such tools appearing more and more frequently, are discussions on the importance of the identification of **high-value datasets**, the reuse of which entails significant benefits for society, the environment and the economy (Utamachant and Anutariya, 2018; Folmer et al., 2019; Zheng and Gao, 2016; Chakravarty, 2018; Slobodova, 2020). In order to emphasize the importance of this point, it is worth to mention it refers to step 1 in the transparency cycle (Janssen et al., 2017). The concept of “high-value dataset” was specified by the Open Data Charter as one of collective actions aimed to boost innovations and increase government transparency. However, even this concept is complex enough because the nature of high-value datasets or more precisely the value of datasets depends on the list of different factors. This issue is crucial since, as recent studies have shown that only a small part of the available open data are of high value for the public. For instance, Folmer et al. (2019) has noticed that the various techniques for evaluating the reuse of data conclude that the number of “high value datasets” on the portal can sometimes be 82 out of 12000 datasets. This proves once again that an attempt to publish as many datasets as possible is not the right way because their value is more important.

In addition, even among those datasets that can be considered popular or valuable, there is a question of aspects such as **data quality** and the **frequency of open data updates**. This observation is no surprise – the practice demonstrates that this is one of the most common barriers for almost all data sectors, see Nikiforova (2020a), which may lead to the loose of users’ trust in data and interest to reuse them, and as a result this is also a risk to transparency (Klein et al., 2018; Koussouris et al., 2015). Torchiano et al. (2017) is one more study pointing on the data quality issues in OGD including data completeness, accuracy, and consistency.

Data quality, which is one of the key factors influencing users' willingness to reuse open data, has a list of quite different definitions, ranging from the most trivial as accuracy to more complex but yet sufficiently relative – fitness for use by data consumers (Attard et al., 2015; Wang et al., 2018). It also is characterized by a list of characteristics, including timeliness, precision, reliability, currency, completeness, and relevancy – see more definitions and data quality dimensions in Nikiforova (2020c). It also proved to be one of the most important factors for open data portals (Wang et al., 2018; Gil-Garcia et al., 2020; Janssen et al., 2017), because, according to Janssen et al. (2012) opening data that do not have adequate quality of information can lead to discussions, confusions, less transparency and even less trust in the government.

Here, in addition to data format, machine-readability, the correspondence of open datasets to the 5-star classification, completeness and accuracy, such more specific characteristics or data quality dimensions as **timeliness** and **frequency of updates** should be mentioned, affecting transparency and users' intention to reuse data, reducing their value, which is sometimes related to a lack of policies that ensure the constant availability of updated data (Craveiro et al., 2016; Hellberg and Hedström, 2015; Janssen et al., 2012; Nikiforova, 2020b; Nugroho et al., 2015; Saxena, 2017; Smith and Sandberg, 2018; Solar et al., 2014; Vetrò et al., 2016; Zuiderwijk and Janssen, 2014). While open data portals and data publishers in many cases satisfy other data quality-related aspects, it appears that the timeliness and frequency of updates remain an unovercome challenge. This category partly refers to steps 6 and 7 in the transparency cycle (Janssen et al., 2017). Taking a step back to the usefulness, as its level is defined by data user as the end user, this category (Ruijter et al., 2017b) is closely related to the last, but definitely not the least, **public engagement, collaboration and participation**.

Although it supposes a very broad set of functionality and activities (Gil-Garcia et al., 2020), we do not sub-categorize them. Here we are bringing together different forms of collaboration having nearly the same aim – to establish citizens' involvement and cooperation between actors promoting and facilitating transparency through communication and active participation. This category therefore refers to steps 8-11 in the transparency cycle (Janssen et al., 2017), including both, social media use, institutional arrangements, follow-up actions and feedback provision.

Perhaps one of the most important but not so often addressed feature is the provision of **open data-based applications / showcases / use-cases**. They are belonging to both data usefulness and public engagement, collaboration, and participation. Despite listing and providing users with information on the number of applications and preferably their nature, based on open data that have been reused is a “good practice” (Meijer et al., 2014; Máchová and Lněnička, 2017), these data are rarely provided. According to Zuiderwijk, et al. (2015) publishing the information on use-cases of already ready services are one of the most crucial success factors for open data, and success stories can attract more people to open data portal. According to Nikiforova (2020a), most countries provide this feature (70.7% of the 41 analyzed open data portals) and Blank (2019) states, that 20 out of EU28 countries provide such an option. In addition, in some cases, such as France, Luxembourg or Portugal, the national portal has the “**use case upload**” feature. Moreover, 18 portals offer **mapping between the use cases and the datasets** they are based on. One such example is Ireland. Since data reuse usually results in the development of services that address issues relevant to a particular country or region and trying to resolve sensitive problems, this option shows the potential of open data. Therefore, it is strongly recommended that this information is published on the portal to demonstrate the positive effects of the use of open data, thereby increasing participation and collaboration of stakeholders.

Consequently, we come to the **service and its quality**, because an active and positive experience of data use and collaboration is less likely if participants are not supported. This support is achieved

through a service that, along with their quality, is even more crucial driver for trust in OGD, which is in line with Purwanto et al. (2020). This is the case even despite the fact the data quality was always considered to be the most important driver. This conclusion was reached based on a questionnaire involving 200 participants of 27 nationalities. This study shows that human services provided on the OGD portal, including designated civil servants offering support or help to data users, in addition to the provision of OGD portal features, can improve the overall level of user satisfaction with the OGD portal and increase trust in the government. According to the authors, the service quality is essential for both first-time users, providing support to them and for experienced users, where the feedback mechanism, as an aspect of the OGD service, can encourage the correction and improvement of low-quality data.

Although it can be considered that the information on the amount of data is not as important as, for instance, ensuring cooperation between citizens and data portals, this information plays a key role in enabling users to understand the current state of the portal and its content and to make sure that the portal is alive (Wang et al., 2018). The amount and nature of information displayed to users differ from the number of datasets and publishers that are considered to be “must have” to the list of newly added and most popular datasets, that allow to track the state of the portal on a daily manner. It is also a way of overcoming the barrier identified in Ruijter et al. (2017b) as the most “popular” and most widely encountered citizen structural barrier, and the second among 38 barriers identified. As in Wang et al. (2018) we are referring this point to **data quantity**. Together with the general features of the portal this category aims to summarize the features that provide basic information about the portal and its structure.

It could also be noticed that our choice is close to Barry and Bannister (2014), where the review of literature allowed authors to classify OGD barriers into data access, data use and data deposit, further classifying them into availability and access, findability, usability, understandability, quality, linking and combining data, comparability and compatibility, and metadata. However, we go deeper, supplying each of the category, which is the first step for us, with the list of features (Table 2).

Table 2 provides a summary of the literature review of all relevant studies. Eight categories and 36 corresponding features promoting transparency-by-design were identified. Although some of these categories are not related to transparency in a very straightforward way, focusing more on participation and collaboration, which are sometimes separated from transparency, they have also been addressed as their implementation promotes transparency as a secondary object. Each category is supplied with a brief description, emphasizing its key point and a list of studies supporting this category. Although in the past open data portal was almost entirely independent of users and citizens who mainly provided data, making them open and accessible, where the open data portals and publishers were the only actors, nowadays, citizens are also considered to be co-actors and co-producers of public services (Moon, 2020). It is therefore essential to ensure that they are involved, that is beneficial for both, data users and portals. Thus, we divide open data portal actors in data portal holders (all units), data publishers, and data users. Therefore, considering that the process of publishing OGD and their further reuse supposes the involvement of both, open data holders, open data publishers and open data users, a role for each aspect was defined as well, thus referring to the responsible person or group of persons who provide or contribute to a specific category. This is also in line with Saxena (2017).

Table 2: A list of categories and corresponding features supporting the transparency-by-design concept

Category / feature	Description	Reference	Responsible role
Data quantity, structure, and general features of the portal			
Multilingualism	the content of the portal should be provided in several languages (including English) to ensure that the data can be re-used by as wide audience as possible	(Petychakis et al., 2014), (Nikiforova, 2020a), (Alexopoulos et al., 2013), (Zuiderwijk et al., 2015), (Nikiforova et al., 2021)	Data portal
Number of datasets and categories	a brief overview of the current number of datasets and categories demonstrates that data are organized and available, as well as proves that the portal is active	(Ruijter et al., 2017b), (Wang et al., 2018), (Saxena, 2017)	Data portal
Dashboards	provides key performance data on the portal, letting users to understand the portal is active; facilitates interaction	(Chatfield and Reddick, 2017), (Bui, 2015), (Lee and Kwak, 2012)	Data portal
Use of vocabulary, dictionary	ensures homogeneity and uniformity in the terms used, thereby promoting better and clearer comprehension of concepts, as well as allowing their automated processing	(Lourenço, 2015), (Attard et al., 2015), (Kim, 2019)	Data portal
Data linkage	allows users to easily find related datasets that could complement each other, thereby creating more context with the data used and facilitating their further reuse, allowing them to be combined	(Saxena, 2017), (Vetrò et al., 2016), (Attard et al., 2015)	Data portal
Data versioning	allows user to track changes (in terms of the update or even the content changed) that should follow a unified predefined standardized representation or using named graphs for versioning	(Attard et al., 2015), (Vetrò et al., 2016)	Data portal
Data quality			
Data format, machine-readability	allows data to be combined and used in various ways without the need of additional software and data transformation for their further reuse	(Saxena, 2017), (Máchová and Lněnička, 2017), (Kalampokis et al., 2011), (Chatfield and Reddick, 2017), (Corrêa et al., 2017), (Alexopoulos et al., 2013), (Moon, 2020), (Ruijter et al., 2017b) (Kassen, 2013), (Thorsby et al., 2017), (Matheus and Janssen, 2020), (Attard et al., 2015), (Saxena, 2017), (Moon, 2020), (Nikiforova et al., 2021)	Data publisher, data portal
Open data 5-star scheme	indication of dataset's format conformance with the 5-star scheme allows users to gain insight into the data format without further inspection of the format of dataset or their collection as well as educates them	(Thorsby et al., 2017), (Petychakis et al., 2014), (Böhm et al., 2012), (Lourenço, 2015), (Attard et al., 2015), (Vetrò et al., 2016)	Data publisher, data portal
Dataset description, attribute description, metadata	ensures the insight into the content of the data, their source, the collection purpose etc., determination of the purpose of each attribute and understanding the meaning of the data values provided, including units used, the list of permitted values, the explanation of abbreviations etc. and any	(Bittencourt et al., 2019), (Degbelo et al., 2019), (Cordasco et al., 2018), (Klein et al., 2018), (Attard et al., 2015), (Ojo et al., 2016), (Lourenço et al., 2017), (Saxena, 2017), (Máchová and Lněnička, 2017), (Wirtz et al., 2018), (Petychakis et al.,	Data publisher, data portal

	manipulations performed to obtain these data (if any), together with complete metadata allowing the data to be further reused	2014), (Chatfield and Reddick, 2017), (Janssen et al., 2012), (Janssen et al., 2017), (Alexopoulos et al., 2013), (Matheus and Janssen, 2020), (Ruijter and Meijer, 2020), (Nikiforova et al., 2021)	
Data timeliness and frequency of updates	allows data user to identify when the data came to the portal, how timely in relation with their accruing they came to the portal, how frequently are they updated, when was the last update made etc.	(Nugroho et al., 2015), (Solar et al., 2014), (Vetrò et al., 2016), (Janssen et al., 2012), (Saxena, 2017), (Hellberg and Hedström, 2015), (Zuiderwijk and Janssen, 2014), (Smith and Sandberg, 2018), (Craveiro et al., 2016), (Magalhaes and Roseira, 2017), (Folmer et al., 2019), (Nikiforova, 2020a, 2020b), (Nikiforova et al., 2021), (Máková and Lněnička, 2017), (Lourenço, 2015), (Corrêa et al., 2017), (Attard et al., 2015), (Janssen et al., 2012), (Lee and Kwak, 2012), (Saxena and Muhammad, 2018), (Matheus and Janssen, 2020), (Ruijter and Meijer, 2020), (Ruijter et al., 2017b)	Data publisher, data portal
Data quality (accuracy syntactic and contextual, completeness, consistency etc.)	data quality should be ensured, preferably by informing the user of the pre-made data quality checks and their results	(Folmer et al., 2019), (Torchiano et al., 2017), (Nikiforova, 2020c), (Klein et al., 2018), (Koussouris et al., 2015), (Kalampokis et al., 2011), (Böhm et al., 2012), (Lourenço, 2015), (Chatfield and Reddick, 2017), (Corrêa et al., 2017), (Attard et al., 2015), (Janssen et al., 2012), (Lee and Kwak, 2012), (Magalhaes and Roseira, 2017), (Saxena, 2017), (Moon, 2020), (Ruijter et al., 2017b), (Vetrò et al., 2016)	Data publisher
Data accessibility			
Open access, open license	data are available and retrievable freely by anyone as license-free, it is desirable that all datasets have a single common license including information on the license used	(Petychakis et al., 2014), (Chatfield and Reddick, 2017), (Attard et al., 2015), (Magalhaes and Roseira, 2017), (Janssen et al., 2017), (Alexopoulos et al., 2013), (Matheus and Janssen, 2020), (Wang et al., 2018), (Nikiforova et al., 2021)	Data portal
Download in a bulk	data must be easily downloadable in bulk for further processing without additional actions, thereby facilitating the use of a greater amount of data	(Wang et al., 2018), (O'Connor, 2015)	
API and SPARQL endpoints	developers should be provided with an entry point to have a direct access to data catalogues and their content, and to update data using external systems, ensuring their smooth use for creating applications, widgets, websites, and other tools	(Thorsby et al., 2017), (Petychakis et al., 2014), (Attard et al., 2015), (Magalhaes and Roseira, 2017), (Alexopoulos et al., 2013), (Moon, 2020), (Nikiforova et al., 2021)	Data publisher, data portal
Preview and discoverability	the data preview and discoverability options must be provided allowing the user to gain insight into the content of the datasets without downloading them	(Petychakis et al., 2014), (Attard et al., 2015), (Nikiforova et al., 2021)	Data portal

Data visualization and analytics tools	the data visualization and analytics options must be provided allowing the user to gain insight into the content of the datasets in a more user-friendly way as well as to gain initial results without their processing e.g. in charts or visualizations in maps, it also makes it considerably easier to use data for non-IT specialists and allows for results on potential trends without additional actions with data	(Saxena, 2017), (Máchová and Lněnička, 2017), (Chatfield and Reddick, 2017), (Thorsby et al., 2017), (Petychakis et al., 2014), (Kassen, 2013), (Böhm et al., 2012), (Chatfield and Reddick, 2017), (Pereira et al., 2017), (Attard et al., 2015), (Lee and Kwak, 2012), (Ruijter et al., 2017a), (Saxena and Muhammad, 2018), (Alexopoulos et al., 2013), (Ruijter et al., 2017b), (Nikiforova et al., 2021)	Data portal
Data findability			
Data search	provides a search box on a home page, allows to find data corresponding to the user's interests using keywords and by browsing topics and thematic categories	(Saxena, 2017), (Wirtz et al., 2018), (Thorsby et al., 2017), (Petychakis et al., 2014), (Böhm et al., 2012), (Lourenço, 2015), (Corrêa et al., 2017), (Ruijter et al., 2017a), (Janssen et al., 2017), (Saxena and Muhammad, 2018), (Alexopoulos et al., 2013), (Matheus and Janssen, 2020), (Attard et al., 2015)	Data portal
Datasets filter and sort	allows datasets to be filtered and sorted by a certain aspect, such as popularity, newly added, recent changes, data format, publisher etc., thereby allowing to limit the search result	(Petychakis et al., 2014), (Böhm et al., 2012), (Lourenço, 2015), (Alexopoulos et al., 2013), (Attard et al., 2015)	Data portal
Datasets categories and cataloguing	allows selecting a subset of datasets by the category, i.e., a topic of interest, thereby making it easier to find a dataset of interest to the user	(Thorsby et al., 2017), (Petychakis et al., 2014), (Kalampokis et al., 2011), (Chatfield and Reddick, 2017), (Attard et al., 2015), (Saxena, 2017), (Moon, 2020), (Ruijter et al., 2017b), (Nikiforova et al., 2021)	Data portal
Data understandability			
Dataset description, attribute description, metadata	provides sufficient information for each dataset – a description of the dataset, a description of the attributes and metadata, allowing for full insight into the data to be used and to allow further processing, thereby enhancing the user experience	(Bittencourt et al., 2019), (Degbelo et al., 2019), (Cordasco et al., 2018), (Klein et al., 2018), (Attard et al., 2015), (Ojo et al., 2016), (Lourenço et al., 2017), (Saxena, 2017), (Máchová and Lněnička, 2017), (Wirtz et al., 2018), (Petychakis et al., 2014), (Chatfield and Reddick, 2017), (Janssen et al., 2012), (Janssen et al., 2017), (Saxena and Muhammad, 2018), (Alexopoulos et al., 2013), (Matheus and Janssen, 2020), (Ruijter and Meijier, 2020), (Vetrò et al., 2016), (Nikiforova et al., 2021)	Data publisher, data portal
Preview and discoverability	allows the user to gain insight into the content of the datasets without downloading them, dataset extent (map view) spatial coverage of a dataset using OpenStreetMap or other service	(Petychakis et al., 2014), (Attard et al., 2015), (Nikiforova et al., 2021)	Data portal
Data visualization and statistics	allows the user to gain insight into the content of the datasets in a more user-friendly way without downloading them, as well as to gain initial results without	(Saxena, 2017), (Máchová and Lněnička, 2017), (Chatfield and Reddick, 2017), (Thorsby et al., 2017), (Petychakis et al., 2014), (Kassen,	Data portal

	their processing, it also makes it considerably easier to use data for non-IT specialists and allows for results on potential trends without additional actions with data, open a dataset with external applications (sites), such as plotly	2013), (Böhm et al., 2012), (Chatfield and Reddick, 2017), (Pereira et al., 2017), (Attard et al., 2015), (Lee and Kwak, 2012), (Ruijter et al., 2017a), (Saxena and Muhammad, 2018), (Alexopoulos et al., 2013), (Ruijter et al., 2017b), (Nikiforova et al., 2021)	
Data usefulness			
Definition of high value datasets	allows to provide only relevant data, i.e. data that are most beneficial to society	(Folmer et al., 2019), (Utamachant and Anutariya, 2018), (Zheng and Gao, 2016), (Chakravarty, 2018), (Slobodova, 2020), (Kassen, 2013), (Lourenço, 2015), (Dawes, 2012), (Lee and Kwak, 2012), (Wilson et al., 2021)	Data portal, data user
Views and downloads	allows identifying the overall data interest by other users	(Lourenço, 2015), (Nikiforova, 2020b), (Lee and Kwak, 2012)	Data portal
Data rating (of usefulness and quality)	allows to assess the usefulness of data in terms of both, their nature and the content and quality, provides other users with an insight on the datasets' quality and usefulness, provides data publishers with an overall assessment of the data they are publishing	(Lourenço, 2015), (Attard et al., 2015), (Alexopoulos et al., 2013), (Nikiforova, 2020a), (Máchová et al., 2018), (Nikiforova et al., 2021, (Purwanto et al., 2020)	Data portal
Reuses: use-cases, use-case upload, mapping between use-case and open datasets	allows gaining insight on how open data can be reused	(Meijer et al., 2014), (Zuiderwijk et al., 2015), (Nikiforova, 2020a,b), (Blank, 2019), (Thorsby et al., 2017), (Goldstein and Dyson, 2013), (Lourenço, 2015), (Chatfield and Reddick, 2017), (Attard et al., 2015), (Lee and Kwak, 2012), (Vetrò et al., 2016), (Moon, 2020)	Data portal, data user
Request for new dataset	allows to request for new dataset of interest and importance for their further reuse, closely related to identification of relevant and high value datasets	(Lourenço, 2015), (Nikiforova, 2020a,b), (Nikiforova et al., 2021), (Máchová et al., 2018), (Wilson et al., 2021)	
Public engagement, collaboration and participation			
Reuses: Use-cases, use-case upload, mapping between use-case and open datasets	allows sharing the audience with an example on how the data can be reused by producing valuable solutions, e.g., modified datasets (a more comprehensive dataset than the original dataset or a new resource extending it, or reuse (use case), i.e., datasets reused to publish an article, a computer graphics, or an application	(Meijer et al., 2014), (Zuiderwijk et al., 2015), (Nikiforova, 2020a,b), (Blank, 2019), (Thorsby et al., 2017), (Goldstein and Dyson, 2013), (Lourenço, 2015), (Chatfield and Reddick, 2017), (Attard et al., 2015), (Lee and Kwak, 2012), (Moon, 2020)	Data portal, data user
Comments, forums, and active feedback	allows interaction between the user and the data portal holder and / or the data publisher aiming to improve the quality of the datasets and service and all the components, facilitates users' interest and trust	(Purwanto et al., 2020), (Máchová and Lněnička, 2017), (Pereira et al., 2017), (Janssen et al., 2012), (Lee and Kwak, 2012), (Ruijter et al., 2017a), (Janssen et al., 2017), (Saxena, 2017), (Purwanto et al., 2020), (Ruijter et al., 2017b), (Nikiforova et al., 2021), (Wilson et al., 2021)	Data portal
Suggestions, forms, report on	allows interaction between the user and the data portal holder by complaints,	(Moon, 2020), (Máchová and Lněnička, 2017), (Máchová et al.,	Data portal

	observations, errors, and suggestions aiming to improve the quality of the service and all the components, facilitates users' interest and trust, a list of activities for each publisher facilitating transparent interaction with users	2018), (Nikiforova et al., 2021), (Wilson et al., 2021)	
Request for new dataset, list of requests, involvement and tracking the state of the request	allows to request for new dataset of interest and importance for their further reuse, other users are able to see the list of requested datasets, voting for them and commenting, as well as tracking the state of the request, closely related to identification of relevant and high value datasets	(Lourenço, 2015), (Nikiforova, 2020a,b), (Nikiforova et al., 2021, (Máchová et al., 2018), (Wilson et al., 2021)	Data portal, data user
Social media	establishes a channel for distribution of open data (sharing a link to a dataset) and sharing feedback	(Saxena, 2017), (Purwanto et al., 2020), (Máchová and Lněnička, 2017), (Lee and Kwak, 2012), (Janssen et al., 2017), (Nikiforova et al., 2021), (Saxena and Muhammad, 2018), (Purwanto et al., 2020), (Ruijter et al., 2017b), (Attard et al., 2015), (Wilson et al., 2021)	Data portal
Hackathons	helps to increase the involvement of citizens and get them interested in the creation of the value from the OGD, provides an opportunity to gain insight and experience on the reuse of data to develop society-centered open data-based solutions	(Goldstein and Dyson, 2013), (Kassen, 2013), (Chatfield and Reddick, 2017), (Saxena, 2017), (Lee and Kwak, 2012), (Attard et al., 2015), (Magalhaes and Roseira, 2017), (Gama, 2017), (Purwanto et al., 2019), (Wilson et al., 2021)	Data portal, data user, data enthusiast
Service quality			
Contact and technical support	civil servants provide a sufficiently timely guidance on the technical aspects, including how to download, process and interpret the data, and features provided in the portal	(Attard et al., 2015), (Alexopoulos et al., 2013), (Vetrò et al., 2016, (Matheus and Janssen, 2020), (Moon, 2020), (Purwanto et al., 2020), (Wilson et al., 2021)	Data portal
Guidelines, tutorials, manuals, FAQ	documents, which provide with guidelines both data users, that facilitate and promote the use of data and portals, and data publishers by defining guidelines and rules for data publishing to be followed to meet the high standard of transparency	(Magalhaes and Roseira, 2017), (Matheus and Janssen, 2020), (Thorsby et al., 2017), (Chatfield and Reddick, 2017), (Nikiforova et al., 2021), (Purwanto et al., 2020)	Data portal
Monitor, track	responsive personnel who receive feedback from users, not only makes sufficiently timely follow-ups, but also allows user to track and monitor this through the changes of status of the request or question	(Lourenço, 2015), (Purwanto et al., 2020)	Data portal

5 Discussion

Considering the categories and corresponding features listed in the Table 2 we can describe how they support the steps of the transparency cycle (Table 3). Open data portals at national and local levels are a key component of the national data infrastructure. The open data portal is supported by a data management system that comprises tools to find and harvest data from public sector registers and databases. This allows data publishers to find and open-up all relevant data. Corresponding features of open data portals can be also used by data users to specify the high value datasets they need and to

request for datasets. Open data policies, strategies, and initiatives clarify which data can be opened up for public usage.

During the data collecting phase open data portals provide features that enable to ensure privacy and equal access to datasets. To protect privacy, all datasets must be anonymized and some of them can only be accessed by authorized users. In addition, portals provide private sharing with individuals or groups, so authorized users can access all datasets in one place. Equal access is supported by means of data accessibility features together with multilingualism of the portal as well as guidelines and tutorials for stakeholders with different levels of skills (data literacy). Next, datasets are published and shared with the public. It is important to ensure that datasets are machine-readable and under open licenses. Data users can download them for reuse in different data formats in a bulk or through API and SPARQL endpoints. To meet the requirements of the right level of abstraction open data portals enable users to search, filter, and sort datasets.

The use of data requires their understanding and interpretation to create new datasets, reuses, and services. Data publishers should ensure that data quality is properly managed and meet the requirements. To validate the conclusions data users must use more than one feature to compare existing options available and influence the government or to take other actions. Nowadays, social media is an essential component in influencing policy formulation and decision-making processes in the public sector. Other features of open data portals to enhance public engagement, collaboration and participation include comments, ratings, forums, lists of requests, questions, hackathons etc. Feedback is essential to consider different preferences between data publishers and data users. It can also be used to review objectives, releasing new datasets, when the entire cycle is resumed from the beginning.

Table 3: Phases of the transparency cycle and corresponding features in the design of open data portals

Phase	Activity / step	Features in the design
Objectives and requirements	1. Finding data	Open data platforms such as CKAN, DKAN, or Socrata together with relevant extensions enable to find relevant data across public sector registers and databases. Data users can suggest and request high value datasets, as well as comment and influence suggestions made by other users.
	2. Opening-up all relevant data	Open data policies, strategies, and initiatives provide a strategic vision and plan for opening-up all relevant data. Data users can access policies, strategies, and initiatives on open data portals.
Collecting data	3. Ensuring privacy	Data users can register themselves and create a personal account that allows them to personalize their usage of the portal.
	4. Equal access	Mostly is ensured by data accessibility category features. Multilingualism is also important, as well as guidelines and tutorials for stakeholders with different levels of skills.
Publishing and sharing	5. Accessible formats	Mainly related to machine-readability and data formats that make data extensible for sharing. The published datasets must meet technical standards and open licenses.
	6. Right level of abstraction	Is ensured by searching, filtering, and sorting of datasets using categories, tags, and keywords. Data versioning allows data users to track changes in datasets.

Using data	7. Understanding data	Data understandability – metadata, dataset description, data vocabulary and dictionary, preview of datasets etc. Forums, guidelines, tutorials, and best practices are also important to learn new skills.
	8. Validation of conclusions	Data quality is a crucial aspect. Other features include number of downloads, list of applications using datasets (new services) and other reuses that support the validation process.
Influencing policy	9. Social media use	The key channel to online promote and share datasets on social media or news media. Portals provides a feature to create a distribution channel and share a link to a dataset on social sites.
	10. Institutional arrangements	Data users can leave comments, feedback, interact with other users, view datasets that were modified based on previous actions etc. Lists of requests, questions, and their status are provided by portals for this purpose.
	11. Follow-up actions	Suggestions on the improvement of processes, requesting for new datasets, list of requests, track the state of request. Revisited, updated, or newly opened datasets can result from these actions.
Account for	12. Provide feedback	Data users can rate datasets or reuses, leave comments, contribute to forums, share experience, e.g. through tutorials, blog, webinars or participate in hackathons.

Similarly, the open data portals can be also used to support the efforts of a transparency action cycle by Kosack and Fung (2014). The cycle aims to improve public services and consists of four elements. The first one is dealing with the accessibility to data users. They must then reuse the data to change their decisions and actions. These new actions then affect service providers in way that seems important and consequential to them. Finally, service providers react constructively and appropriately when interacting with data users. All these steps can take place on open data portals. There is also a similar approach, but with different driver in the research community that studies scientific data management. The four foundational principles are Findability, Accessibility, Interoperability, and Reusability (FAIR) to ensure transparency, reproducibility, and reusability. The FAIR approaches bring requirements of data management in broad sense instead of data quality requirements to the data portals. However, it should be noted that the releasing of the OGD to the public through portals does not guarantee the success of promoting transparency and facilitating accountability of governments (Attard et al., 2015).

6 Conclusion

The lack of research studies focusing on setting transparency requirements and meeting them when designing open data portals constitutes our study. After we analyzed the relations between OGD, open data portals, and transparency found in the literature (RQ1), we addressed the role of open data portals in transparency (RQ2) to provide research priorities and best practices in design of open data portals (RQ3).

Open data portals are an essential component of any data infrastructure, since they connect data providers with data users in terms of delivering data and enabling to establish the relationships needed to increase transparency. The concept of transparency-by-design emphasizes the phases of the

publication and reuse of OGD, and this study aimed to provide relevant findings on the categories and corresponding features of open data portals supporting these activities. Considering the identified features in the design of open data portals and their mapping onto the transparency cycle, we can provide further empirical evidence that supports the importance of these data infrastructures in increasing transparency.

Other benefits include greater levels of accountability and trust in public officials, engagement of stakeholders in policy and decision-making processes or wider public control over resource allocation. By providing an online channel for requesting and processing public sector information, open data portals help reduce costs for FOI and information exchanging. In addition, the number of requests for the same information is also decreasing since the portal provides appropriate datasets that can be accessed and reused by everyone, thereby potentially providing valuable solutions for society by transforming raw data into value and knowledge.

References

- Alexopoulos, C., Diamantopoulou, V., & Charalabidis, Y. (2017). Tracking the evolution of OGD portals: a maturity model. In *International Conference on Electronic Government (EGOV 2017)* (pp. 287–300). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-64677-0_24
- Alexopoulos, C., Spiliotopoulou, L., & Charalabidis, Y. (2013). Open data movement in Greece: a case study on open government data sources. In *PCI '13: Proceedings of the 17th Panhellenic Conference on Informatics* (pp. 279–286). <https://doi.org/10.1145/2491845.2491876>
- Andrades, J., Martinez-Martinez, D., Larran, M., & Herrera, J. (2019). Online information disclosure in Spanish municipal-owned enterprises. *Online Information Review*, 43(5), 922–944. <https://doi.org/10.1108/OIR-02-2018-0063>
- Arauz, M. R., Moreno, Y., Nancalres, R., Pérez, C. V., & Larios, V. M. (2017). Tackling corruption in urban development through open data and citizen empowerment: the case of “visor urbano” in guadalajara. In *2017 International Smart Cities Conference (ISC2)* (pp. 1–4). IEEE. <https://doi.org/10.1109/ISC2.2017.8090840>
- Attard, J., Orlandi, F., & Auer, S. (2016). Data driven governments: creating value through open government data. In A. Hameurlain et al. (Eds.), *Transactions on Large-Scale Data- and Knowledge-Centered Systems XXVII. Lecture Notes in Computer Science, 9860*. (pp. 84–110). Berlin Heidelberg: Springer. https://doi.org/10.1007/978-3-662-53416-8_6
- Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly*, 32(4), 399–418. <https://doi.org/10.1016/j.giq.2015.07.006>
- Bannister, F., & Connolly, R. (2014). ICT, public values and transformative government: a framework and programme for research. *Government Information Quarterly*, 31(1), 119–128. <https://doi.org/10.1016/j.giq.2013.06.002>
- Barry, E., & Bannister, F. (2014). Barriers to open data release: a view from the top. *Information Polity*, 19(1,2), 129–152. <https://doi.org/10.3233/IP-140327>
- Bertot, J. C., Jaeger, P. T., & Grimes, J. M. (2010). Using ICTs to create a culture of transparency: e-government and social media as openness and anti-corruption tools for societies. *Government Information Quarterly*, 27(3), 264–271. <https://doi.org/10.1016/j.giq.2010.03.001>
- Bittencourt, C., Estima, J., & Pestana, G. (2019). Open data initiatives in Brazil: an analysis of the public entities portals. In *14th Iberian Conference on Information Systems and Technologies (CISTI 2019)* (pp. 21–24). IEEE. <https://doi.org/10.23919/CISTI.2019.8760592>
- Blank, M. (2019). *Open Data Maturity Report 2019*, European Data Portal, Retrieved October 10, 2020, from https://www.europeandataportal.eu/sites/default/files/open_data_maturity_report_2019.pdf

- Bogdanović-Dinić, S., Veljković, N., & Stoimenov, L. (2014). How open are public government data? an assessment of seven open data portals. In M.P. Rodríguez-Bolívar (Ed.), *Measuring E-government Efficiency, Public Administration and Information Technology* (pp. 25–44). New York: Springer.
https://doi.org/10.1007/978-1-4614-9982-4_3
- Böhm, C., Freitag, M., Heise, A., Lehmann, C., Mascher, A., Naumann, F., ... & Schmidt, M. (2012). GovWILD: integrating open government data for transparency. In *Proceedings of the 21st International Conference on World Wide Web* (pp. 321–324). <https://doi.org/10.1145/2187980.2188039>
- Bui, L. (2015). Breathing smarter: a critical look at representations of air quality sensing data across platforms and publics. In *2015 IEEE First International Smart Cities Conference (ISC2)* (pp. 1–5). IEEE.
<https://doi.org/10.1109/ISC2.2015.7366185>
- Carrara, W., Chan, W. S., Fische, S., & Steenbergen, E. V. (2015). *Creating value through open data: study on the impact of re-use of public data resources*. Luxembourg: Publications Office of the European Union.
<https://doi.org/10.2759/328101>
- Chakravarty, R. (2018). Open Government Data (OGD) initiative in India: an empirical analysis. In *Proliferation of Open Government Initiatives and Systems* (pp. 68–90). IGI Global. <https://doi.org/10.4018/978-1-5225-4987-1.ch004>
- Charalabidis, Y., Alexopoulos, C., & Loukis, E. (2016). A taxonomy of open government data research areas and topics. *Journal of Organizational Computing and Electronic Commerce*, 26(1-2), 41–63.
<https://doi.org/10.1080/10919392.2015.1124720>
- Charalabidis, Y., Loukis, E., & Alexopoulos, C. (2014). Evaluating second generation open government data infrastructures using value models. In *Proceedings of the 47th Hawaii International Conference on System Sciences* (pp. 2114–2126). IEEE. <https://doi.org/10.1109/HICSS.2014.267>
- Chatfield, A. T., & Reddick, C. G. (2017). A longitudinal cross-sector analysis of open data portal service capability: the case of Australian local governments. *Government Information Quarterly*, 34(2), 231–243.
<https://doi.org/10.1016/j.giq.2017.02.004>
- Clarke, A., & Margetts, H. (2014). Governments and citizens getting to know each other? Open, closed, and big data in public management reform. *Policy & Internet*, 6(4), 393–417. <https://doi.org/10.1002/1944-2866.POI377>
- Cordasco, G., Malandrino, D., Pirozzi, D., Scarano, V., & Spagnuolo, C. (2018). A layered architecture for open data: design, implementation and experiences. In *Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance* (pp. 371–381). New York: ACM.
<https://doi.org/10.1145/3209415.3209466>
- Corrêa, A. S., de Paula, E. C., Corrêa, P. L. P., & da Silva, F. S. C. (2017). Transparency and open government data: a wide national assessment of data openness in Brazilian local governments. *Transforming Government: People, Process and Policy*, 11(1), 58–78. <https://doi.org/10.1108/TG-12-2015-0052>
- Craveiro, G. S., Machado, J. A., & Machado, J. S. (2016). The use of open government data to citizen empowerment. In *Proceedings of the 9th International Conference on Theory and Practice of Electronic Governance* (pp. 398–399). <https://doi.org/10.1145/2910019.2910076>
- Dawes, S. S. (2012). A realistic look at open data. *Center for Technology in Government, University at Albany/SUNY* Available at https://www.w3.org/2012/06/pmod/pmod2012_submission_38.pdf.
- Degbelo, A., Granell, C., Trilles, S., Bhattacharya, D., & Wissing, J. (2019). Tell me how my Open Data is re-used: Increasing transparency through the Open City Toolkit. *Open Cities Open Data: Collaborative cities in the information era*. Melbourne: Palgrave Macmillan.
- Eckartz, S. M., Hofman, W. J., & Van Veenstra, A. F. (2014). A decision model for data sharing. In Janssen, M., et al. (Eds): *Electronic Government* (pp. 253–264). Berlin Heidelberg: Springer.
https://doi.org/10.1007/978-3-662-44426-9_21

- Esteban, O.-O. (2019). *The rise of social media*. Retrieved October 10, 2020, from <https://our-worldindata.org/rise-of-social-media>.
- Folmer, E., Beek, W., Rietveld, L., Ronzhin, S., Geerling, R., & den Haan, D. (2019). Enhancing the usefulness of open governmental data with linked data viewing techniques. In *Proceedings of the 52nd Hawaii International Conference on System Sciences* (pp. 2912–2921). <https://doi.org/10.24251/HICSS.2019.352>
- Gama, K. (2017). Crowdsourced Software Development in Civic Apps - Motivations of Civic Hackathons Participants. In *Proceedings of the 19th International Conference on Enterprise Information Systems - Volume 2: ICEIS*, pp. 550-555. DOI: 10.5220/0006377005500555
- Gebre, E. H., & Morales, E. (2020). How “accessible” is open data? Analysis of context-related information and users’ comments in open datasets. *Information and Learning Sciences*, 121(1/2), 19–36. <https://doi.org/10.1108/ILS-08-2019-0086>
- Geiger, C. P., & von Lucke, J. (2012). Open government and (linked) (open) (government) (data). *JeDEM - Journal of eDemocracy and Open Government*, 4(2), 265–278.
- Gil-Garcia, J. R., Gasco-Hernandez, M., & Pardo, T. A. (2020). Beyond transparency, participation, and collaboration? A reflection on the dimensions of open government. *Public Performance & Management Review*, 43(3), 483–502. <https://doi.org/10.1080/15309576.2020.1734726>
- Glasse, O. (2017). Proactive transparency and open data: a tentative analysis. In *International Conference on Electronic Government (EGOV 2017)* (pp. 314–323). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-64677-0_26
- Goel, R. K., Nelson, M. A., & Naretta, M. A. (2012). The internet as an indicator of corruption awareness. *European Journal of Political Economy*, 28(1), 64–75. <https://doi.org/10.1016/j.ejpoleco.2011.08.003>
- Goldstein, B., & Dyson, L. (2013). *Beyond transparency: open data and the future of civic innovation*. San Francisco: Code for American Press.
- Grimmelikhuijsen, S. G., & Meijer, A. J. (2014). Effects of transparency on the perceived trustworthiness of a government organization: evidence from an online experiment. *Journal of Public Administration Research and Theory*, 24(1), 137–157. <https://doi.org/10.1093/jopart/mus048>
- Hellberg, A. S., & Hedström, K. (2015). The story of the sixth myth of open data and open government. *Transforming Government: People, Process and Policy*, 9(1), 35–51. <https://doi.org/10.1108/TG-04-2014-0013>
- Hivon, J., & Titah, R. (2017). Conceptualizing citizen participation in open data use at the city level. *Transforming Government: People, Process and Policy*, 11(1), 99–118. <https://doi.org/10.1108/TG-12-2015-0053>
- Hogan, M., Ojo, A., Harney, O., Ruijter, E., Meijer, A., Andriessen, J., ... & Groff, J. (2017). Governance, transparency and the collaborative design of open data collaboration platforms: understanding barriers, options, and needs. In *Government 3.0—Next Generation Government Technology Infrastructure and Services* (pp. 299–332). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-63743-3_12
- Ingrams, A. (2017). Managing governance complexity and knowledge networks in transparency initiatives: the case of police open data. *Local Government Studies*, 43(3), 364–387. <https://doi.org/10.1080/03003930.2017.1294070>
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management*, 29(4), 258–268. <https://doi.org/10.1080/10580530.2012.716740>
- Janssen, M., Matheus, R., Longo, J., & Weerakkody, V. (2017). Transparency-by-design as a foundation for open government. *Transforming Government: People, Process and Policy*, 11(1), 2–8. <https://doi.org/10.1108/TG-02-2017-0015>

- Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2014a). Data-driven innovation through open government data. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(2), 100–120. <https://doi.org/10.4067/S0718-18762014000200008>
- Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2014b). Generating sustainable value from open data in a sharing society. In *Proceedings of IFIP WG 8.6 International Conference on Transfer and Diffusion of IT: TDIT 2014* (pp. 62–82). Berlin Heidelberg: Springer. https://doi.org/10.1007/978-3-662-43459-8_5
- Kalampokis, E., Tambouris, E., & Tarabanis, K. (2011). Open government data: a stage model. In *Proceedings of the 10th IFIP WG 8.5 International Conference (EGOV 2011)* (pp. 235–246). Berlin Heidelberg: Springer. https://doi.org/10.1007/978-3-642-22878-0_20
- Kassen, M. (2013). A promising phenomenon of open data: a case study of the Chicago open data project. *Government Information Quarterly*, 30(4), 508–513. <https://doi.org/10.1016/j.giq.2013.05.012>
- Kim, H. (2019). Analysis of standard vocabulary use of the open government data: the case of the public data portal of Korea. *Quality & Quantity*, 53(3), 1611–1622. <https://doi.org/10.1007/s11135-018-0829-z>
- Kimball, M. B. (2011). Mandated state-level open government training programs. *Government Information Quarterly*, 28(4), 474–483. <https://doi.org/10.1016/j.giq.2011.04.003>
- Klein, R. H., Klein, D. C. B., & Luciano, E. M. (2018). Identification of mechanisms for the increase of transparency in open data portals: An analysis in the Brazilian context. *Cadernos EBAPE.BR*, 16(4), 692–715. <https://doi.org/10.1590/1679-395173241>
- Klievink, B., Neuron, A., Fraefel, M., & Zuiderwijk, A. (2017). Digital strategies in action: A comparative analysis of national data infrastructure development. In *Proceedings of the 18th Annual International Conference on Digital Government Research* (pp. 129–138). New York: ACM. <https://doi.org/10.1145/3085228.3085270>
- Kosack, S., & Fung, A. (2014). Does transparency improve governance?. *Annual Review of Political Science*, 17, 65–87. <https://doi.org/10.1146/annurev-polisci-032210-144356>
- Koussouris, S., Lampathaki, F., Kokkinakos, P., Askounis, D., & Misuraca, G. (2015). Accelerating policy making 2.0: innovation directions and research perspectives as distilled from four standout cases. *Government Information Quarterly*, 32(2), 142–153. <https://doi.org/10.1016/j.giq.2015.03.001>
- Kucera, J., & Chlapek, D. (2014). Benefits and risks of open government data. *Journal of Systems Integration*, 5(1), 30–41.
- Lee, G., & Kwak, Y. H. (2012). An open government maturity model for social media-based public engagement. *Government Information Quarterly*, 29(4), 492–503. <https://doi.org/10.1016/j.giq.2012.06.001>
- Lourenço, R. P. (2015). An analysis of open government portals: A perspective of transparency for accountability. *Government Information Quarterly*, 32(3), 323–332. <https://doi.org/10.1016/j.giq.2015.05.006>
- Lourenço, R. P., Piotrowski, S., & Ingrams, A. (2017). Open data driven public accountability. *Transforming Government: People, Process and Policy*, 11(1), 42–57. <https://doi.org/10.1108/TG-12-2015-0050>
- Machado, V., Mantini, G., Viterbo, J., Bernardini, F., & Barcellos, R. (2018). An instrument for evaluating open data portals: a case study in Brazilian cities. In *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age* (p. 19). New York: ACM. <https://doi.org/10.1145/3209281.3209370>
- Magalhaes, G., & Roseira, C. (2017). Open government data and the private sector: an empirical view on business models and value creation. *Government Information Quarterly*, 37(3), 101248. <https://doi.org/10.1016/j.giq.2017.08.004>
- Matheus, R., & Janssen, M. (2020). A systematic literature study to unravel transparency enabled by open government data: the window theory. *Public Performance & Management Review*, 43(3), 503–534. <https://doi.org/10.1080/15309576.2019.1691025>

- Máchová, R., Hub, M., & Lněnička, M. (2018). Usability evaluation of open data portals: evaluating data discoverability, accessibility, and reusability from a stakeholders' perspective. *Aslib Journal of Information Management*, 70(3), 252–268. <https://doi.org/10.1108/ajim-02-2018-0026>
- Máchová, R., & Lněnička, M. (2016). Exploring the emerging impacts of open data in the public sector. In Špalková, D., Matějová, L. (Eds): *Proceedings of the 20th International Conference Current Trends in Public Sector Research* (pp. 36–44). Brno: Masaryk University.
- Máchová, R., & Lněnička, M. (2017). Evaluating the quality of open data portals on the national level. *Journal of Theoretical and Applied Electronic Commerce Research*, 12(1), 21–41. <https://doi.org/10.4067/S0718-18762017000100003>
- Martin, E. G., & Begany, G. M. (2017). Opening government health data to the public: benefits, challenges, and lessons learned from early innovators. *Journal of the American Medical Informatics Association*, 24(2), 345–351. <https://doi.org/10.1093/jamia/ocw076>
- Meijer, A. (2009). Understanding modern transparency. *International Review of Administrative Sciences*, 75(2), 255–269. <https://doi.org/10.1177/0020852309104175>
- Meijer, R., Conradie, P., & Choenni, S. (2014). Reconciling contradictions of open data regarding transparency, privacy, security and trust. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(3), 32–44. <http://doi.org/10.4067/S0718-18762014000300004>
- Mendoza, R. U., & Cruz, J. P. (2020). Governing the “golden age of infrastructure”: assessing transparency innovations in Philippine infrastructure development. *Asian Politics & Policy*, 12(2), 175–204. <https://doi.org/10.1111/aspp.12527>
- Moon, M. J. (2020). Shifting from old open government to new open government: four critical dimensions and case illustrations. *Public Performance & Management Review*, 43(3), 535–559. <https://doi.org/10.1080/15309576.2019.1691024>
- Murillo, M. J. (2015). Evaluating the role of online data availability: the case of economic and institutional transparency in sixteen Latin American nations. *International Political Science Review*, 36(1), 42–59. <https://doi.org/10.1177/0192512114541163>
- Nikiforova, A. (2020a). Comparative analysis of national open data portals or whether your portal is ready to bring benefits from open data. In *IADIS International Conference on ICT, Society and Human Beings 2020* (pp. 1–8).
- Nikiforova, A. (2020b). Timeliness of open data in open government data portals through pandemic-related data: a long data way from the publisher to the user. In *2020 Fourth International Conference on Multimedia Computing, Networking and Applications (MCNA)* (pp. 131–138). IEEE, <https://doi.org/10.1109/MCNA50957.2020.9264298>
- Nikiforova, A. (2020c). Definition and evaluation of data quality: User-oriented data object-driven approach to data quality assessment. *Baltic Journal of Modern Computing*, 8(3), 391–432. <https://doi.org/10.22364/bjmc.2020.8.3.02>
- Nikiforova, A., & McBride, K. (2021). Open government data portal usability: A user-centred usability analysis of 41 open government data portals. *Telematics and Informatics*, 58, 101539. <https://doi.org/10.1016/j.tele.2020.101539>
- Nilsen, K. (2010). Economic theory as it applies to public sector information. *Annual Review of Information Science and Technology*, 44(1), 419–489. <https://doi.org/10.1002/aris.2010.1440440117>
- Noveck, B. S. (2017). Rights-based and tech-driven: Open Data, freedom of information, and the future of government transparency. *Yale Human Rights and Development Law Journal*, 19(1), Article 1.
- Nugroho, R. P., Zuiderwijk, A., Janssen, M., & de Jong, M. (2015). A comparison of national open data policies: lessons learned. *Transforming Government: People, Process and Policy*, 9(3), 286–308. <https://doi.org/10.1108/TG-03-2014-0008>

- O'Connor, J. (2015). *Towards a Profile of Open Government Data Users*. <https://doi.org/10.17615/25b1-bt94>
- Ojo, A., Porwol, L., Waqar, M., Stasiewicz, A., Osagie, E., Hogan, M., ... & Zeleti, F. A. (2016). Realizing the innovation potentials from open data: stakeholders' perspectives on the desired affordances of open data environment. In *Working Conference on Virtual Enterprises* (pp. 48–59). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-45390-3_5
- Paré, G., Trudel, M. C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: a typology of literature reviews. *Information & Management*, 52(2), 183–199. <https://doi.org/10.1016/j.im.2014.08.008>
- Pereira, G. V., Macadar, M. A., Luciano, E. M., & Testa, M. G. (2017). Delivering public value through open government data initiatives in a Smart City context. *Information Systems Frontiers*, 19(2), 213–229. <https://doi.org/10.1007/s10796-016-9673-7>
- Petychakis, M., Vasileiou, O., Georgis, C., Mouzakitis, S., & Psarras, J. (2014). A state-of-the-art analysis of the current public data landscape from a functional, semantic and technical perspective. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(2), 34–47. <http://doi.org/10.4067/S0718-18762014000200004>
- Purwanto, A., Zuiderwijk, A., & Janssen, M. (2020). Citizens' trust in open government data: a quantitative study about the effects of data quality, system quality and service quality. In *The 21st Annual International Conference on Digital Government Research* (pp. 310–318). New York: ACM. <https://doi.org/10.1145/3396956.3396958>
- Purwanto, A., Zuiderwijk, A., & Janssen, M. (2019). Citizens' motivations for engaging in open data hackathons. In *International Conference on Electronic Participation* (pp. 130-141). Springer, Cham.
- Rajshree, N., & Srivastava, B. (2012). Open government data for tackling corruption – a perspective. In *Workshops at the Twenty-Sixth AAAI Conference on Artificial Intelligence* (pp. 21–24). Palo Alto: AAAI Press.
- Ruijter, E., Détienne, F., Baker, M., Groff, J., & Meijer, A. J. (2020). The politics of open government data: Understanding organizational responses to pressure for more transparency. *The American Review of Public Administration*, 50(3), 260–274. <https://doi.org/10.1177/0275074019888065>
- Ruijter, E., Grimmelikhuijsen, S., & Meijer, A. (2017a). Open data for democracy: developing a theoretical framework for open data use. *Government Information Quarterly*, 34(1), 45–52. <https://doi.org/10.1016/j.giq.2017.01.001>
- Ruijter, E., Grimmelikhuijsen, S., Hogan, M., Enzerink, S., Ojo, A., & Meijer, A. (2017b). Connecting societal issues, users and data. Scenario-based design of open data platforms. *Government Information Quarterly*, 34(3), 470–480. <https://doi.org/10.1016/j.giq.2017.06.003>
- Ruijter, E., & Meijer, A. (2020). Open government data as an innovation process: lessons from a living lab experiment. *Public Performance & Management Review*, 43(3), 613–635. <https://doi.org/10.1080/15309576.2019.1568884>
- Safarov, I., Meijer, A., & Grimmelikhuijsen, S. (2017). Utilization of open government data: a systematic literature review of types, conditions, effects and users. *Information Polity*, 22(1), 1–24. <https://doi.org/10.3233/IP-160012>
- Saxena, S. (2017). “Usage by stakeholders” as the objective of “transparency-by-design” in open government data: case study of Sri Lanka’s open data initiative”. *Information and Learning Sciences*, 118(7/8), 420–432. <https://doi.org/10.1108/ILS-05-2017-0034>
- Saxena, S., & Muhammad, I. (2018). The impact of open government data on accountability and transparency. *Journal of Economic and Administrative Sciences*, 34(3), 204–216. <https://doi.org/10.1108/JEAS-05-2017-0044>

- Sayogo, D. S., Pardo, T. A., & Cook, M. (2014). A framework for benchmarking open government data efforts. In *Proceedings of the 47th Hawaii International Conference on System Sciences* (pp. 1896–1905). IEEE. <https://doi.org/10.1109/HICSS.2014.240>
- Sieber, R. E., & Johnson, P. A. (2015). Civic open data at a crossroads: dominant models and current challenges. *Government Information Quarterly*, 32(3), 308–315. <https://doi.org/10.1016/j.giq.2015.05.003>
- Slobodova, O. (2020). Motivations of civil society actors to reuse open data. Case study of discrete actors in Lyon. *Bitácora Urbano Territorial*, 30(3), 55–70. <https://doi.org/10.15446/bitacora.v30n3.80740>
- Smith, G., & Sandberg, J. (2018). Barriers to innovating with open government data: exploring experiences across service phases and user types. *Information Polity*, 23(3), 249–265. <https://doi.org/10.3233/IP-170045>
- Solar, M., Daniels, F., López, R., & Meijueiro, L. (2014). A model to guide the open government data implementation in public agencies. *J. UCS*, 20(11), 1564–1582. <https://doi.org/10.3217/jucs-020-11-1564>
- Styrin, E., Luna-Reyes, L. F., & Harrison, T. M. (2017). Open data ecosystems: an international comparison. *Transforming Government: People, Process and Policy*, 11(1), 132–156. <https://doi.org/10.1108/TG-01-2017-0006>
- Susha, I., Janssen, M., & Verhulst, S. (2017). Data collaboratives as “bazaars”? A review of coordination problems and mechanisms to match demand for data with supply. *Transforming Government: People, Process and Policy*, 11(1), 157–172. <https://doi.org/10.1108/TG-01-2017-0007>
- Thorsby, J., Stowers, G. N., Wolslegel, K., & Tumbuan, E. (2017). Understanding the content and features of open data portals in American cities. *Government Information Quarterly*, 34(1), 53–61. <https://doi.org/10.1016/j.giq.2016.07.001>
- Tkacz, N. (2012). From open source to open government: a critique of open politics. *ephemera*, 12(4), 386–405.
- Torchiano, M., Vetrò, A., & Iuliano, F. (2017). Preserving the benefits of Open Government Data by measuring and improving their quality: an empirical study. In *2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC)* (pp. 144–153). IEEE. <https://doi.org/10.1109/COMPSAC.2017.192>
- Utamachant, P., & Anutariya, C. (2018). An analysis of high-value datasets: a case study of Thailand’s open government data. In *2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE)* (pp. 1–6). IEEE. <https://doi.org/10.1109/JCSSE.2018.8457350>
- Van Schalkwyk, F., & Cañares, M. (2019). Open government data for inclusive development, In M. L. Smith & R. K. Seward (Eds.), *Making open development inclusive* (pp. 251–290). Cambridge, MA: MIT Press.
- Veljković, N., Bogdanović-Dinić, S., & Stoimenov, L. (2014). Benchmarking open government: an open data perspective. *Government Information Quarterly*, 31(2), 278–290. <https://doi.org/10.1016/j.giq.2013.10.011>
- Vetrò, A., Canova, L., Torchiano, M., Minotas, C. O., Iemma, R., & Morando, F. (2016). Open data quality measurement framework: definition and application to Open Government Data. *Government Information Quarterly*, 33(2), 325–337. <https://doi.org/10.1016/j.giq.2016.02.001>
- Wang, D., Chen, C., & Richards, D. (2018). A prioritization-based analysis of local open government data portals: a case study of Chinese province-level governments. *Government Information Quarterly*, 35(4), 644–656. <https://doi.org/10.1016/j.giq.2018.10.006>
- Wilson, B., & Cong, C. (2021). Beyond the supply side: Use and impact of municipal open data in the US. *Telematics and Informatics*, 58, 101526, <https://doi.org/10.1016/j.tele.2020.101526>
- Wirtz, B. W., & Birkmeyer, S. (2015). Open government: origin, development, and conceptual perspectives. *International Journal of Public Administration*, 38(5), 381–396. <https://doi.org/10.1080/01900692.2014.942735>

- Wirtz, B. W., Weyerer, J. C., & Rösch, M. (2018). Citizen and open government: an empirical analysis of antecedents of open government data. *International Journal of Public Administration*, 41(4), 308–320. <https://doi.org/10.1080/01900692.2016.1263659>
- Zhang, N., Zhao, X., Zhang, Z., Meng, Q., & Tan, H. (2017). What factors drive open innovation in China's public sector? A case study of official document exchange via microblogging (ODEM) in Haining. *Government Information Quarterly*, 34(1), 126–133. <https://doi.org/10.1016/j.giq.2016.11.002>
- Zheng, L., & Gao, F. (2016). Assessment on China's open government data platforms: framework, status and problems. In *Proceedings of the 17th International Digital Government Research Conference on Digital Government Research* (pp. 408–414). <https://doi.org/10.1145/2912160.2912213>
- Zuiderwijk, A., & Janssen, M. (2015). Towards decision support for disclosing data: closed or open data?. *Information Polity*, 20(2, 3), 103–117. <https://doi.org/10.3233/IP-150358>
- Zuiderwijk, A., Janssen, M., & Parnia, A. (2013). The complementarity of open data infrastructures: an analysis of functionalities. In *Proceedings of the 14th Annual International Conference on Digital Government Research* (pp. 166–171). New York: ACM. <https://doi.org/10.1145/2479724.2479749>
- Zuiderwijk, A., Susha, I., Charalabidis, Y., Parycek, P., & Janssen, M. (2015). Open data disclosure and use: critical factors from a case study. In *Proceedings of the International Conference for E-Democracy and Open Government (CeDEM15)* (pp. 197–208). Krems: Universität Krems.
- Zuiderwijk, A., & Janssen, M. (2014). The negative effects of open government data - investigating the dark side of open data. In *Proceedings of the 15th Annual International Conference on Digital Government Research* (pp. 147–152). New York: ACM. <https://doi.org/10.1145/2612733.2612761>
- Žuffová, M. (2020). Do FOI laws and open government data deliver as anti-corruption policies? Evidence from a cross-country study. *Government Information Quarterly*, 37(3), 101480. <https://doi.org/10.1016/j.giq.2020.101480>