Special Issue on Transparency and Open Data Policies: Guest Editors’ Introduction

Anneke Zuiderwijk1, Mila Gascó2, Peter Parycek3, Marijn Janssen4

Guest Editors
1 Delft University of Technology, Faculty of Technology, Policy and Management, The Netherlands, a.m.g.zuiderwijk-vaneijk@tudelft.nl
2 ESADE Institute of Public Governance and Management, Spain, mila.gasco@esade.edu
3 Danube University Krems, Center for E-Governance, Austria, peter.parycek@donau-uni.ac.at
4 Delft University of Technology, Faculty of Technology, Policy and Management, The Netherlands, m.f.w.h.a.janssen@tudelft.nl

September, 2014

Literature e.g., [27], [33], [46], [56], [70] and policy reports e.g., [18], [19]-[21], [49], [50], [62] suggest that open data can be used as a tool to enhance transparency. However, several researchers have challenged the idea that the release of government data will result in increased transparency [3], [34] and the idea that transparency automatically leads to more trust in the government [28]. Moreover, transparency may have undesired effects [60]. Limited research has been conducted on these sometimes conflicting findings and challenges, and on how open data policies should deal with this. It is not clear what encompasses an effective open data policy and how this influences transparency. The papers in this special issue contribute to this area. In the introduction of this special issue we present the state-of-the-art with regard to definitions, developments, research, challenges and barriers related to open data transparency and open data policies. Finally, a summary of the papers included in this special issue is provided.

This special issue forms a series of two special issues on open data. In the previous special issue Innovation Through Open Data: A Review of the State-of-the-Art and an Emerging Research Agenda the state-of-the-art with respect to understanding the context of open data innovation, developments, challenges and barriers was discussed. Moreover, the previous special issue gave an overview of open data research and outlined emerging research directions [73]. The first issue emphasized research on the relationship between innovation and open data [73], while this second issue focuses on transparency and open data policies.

Open Data Transparency

In this section we first discuss various definitions of transparency. Subsequently, developments, research, barriers and challenges related to open data transparency are described.

Defining Transparency

Transparency is an intuitively appealing term, yet it is hard to operationalize in practice. The concept of transparency is complex and multidimensional [31] and what can be transparent for one person is not necessarily transparent for someone else. A condition for transparency in the context of open data is that data can be accessed, processed and presented easily. Yet, whether something is easy to access, process or present partly depends on the capabilities of open data users. As such, transparency should always address the interplay between those who provide open data, the functionalities of the system that enable access to the data and the open data user.

Transparency has been a major characteristic of information technology already since the 1980s [43]. Meijer makes a distinction between pre-modern and modern transparency, since transparency nowadays is much influenced by the effects of new technologies. He refers to modern transparency as computer-mediated transparency. He therefore views transparency as a cultural manifestation that should be positioned within a wider perspective of broad trends in society [43]. Birkinshaw [8] pp. 189-191 and Grimmelikhuijsen [29] p. 175 also take a broad view and define transparency as the conduct of public affairs in the open or otherwise subject to public scrutiny. This implies that citizens watch agencies from the outside [29]. In line with this, Lindstedt and Naurin [39] define transparency as the release of information about institutions that is relevant for evaluating those institutions p. 301.

Meijer et al. [44] point at the importance to not just see open government as openness in informational terms, but, also, to see openness in interactive terms. Governmental information should not just be put on the internet, but participation and interaction are important to create opportunities for engagement. The authors state that, based on [15], “citizens need information to see what is going on inside government and participation to voice their opinions about this” [44], p. 11. Participation can be viewed as a way to create transparency, as interactions with each other can result in increased knowledge and understanding.
In the context of open data, transparency in itself should not be seen as an end-goal, but as a means to realize other goals, such as innovation in the public sector, building trust, enabling democratic engagement, providing access to information, including enhancing the findability, usability and intelligibility of open data and providing the possibility to trace governmental processes and policy and decision-making. Related to this, Ball [2] identifies three metaphors that are used for transparency, namely 1) transparency as a public value embraced by society to counter corruption, 2) transparency synonymous with open decision-making by governments and nonprofits, and 3) transparency as a complex tool of good governance in programs, policies, organizations, and nations p. 293. These goals and types of transparency might strengthen democracy and open data can be used as an essential lever to achieve these goals.

The internet is a medium which allows for releasing government information to citizens [14], [29], and in this way the internet can be used to increase government transparency [14]. Yet, it is a myth that transparency will automatically strengthen effectiveness and legitimacy of governments [idem]. Transparency may even weaken effectiveness and legitimacy in certain cases [idem]. Meijer [43] states that opponents argue […] that unidirectional, structured and decontextualized forms of transparency will result in a loss of societal trust p. 255, while postmodernists focus on the esthetics of transparency and argue for varied and diverse forms of computer-mediated transparency p. 255.

In this special issue, we take a broad view on transparency. Transparency can be a mean in itself, or a means to accomplish other goals. Transparency can be an outcome of a process of information finding, processing and discussing. Nevertheless, not every aspect can be and will be transparent and it is not likely that something is completely transparent to anyone. As such we view transparency is a continuous rather than a dichotomous construct. The ability to create transparency is dependent on the data availability, data format, data quality, systems for processes, resources and capabilities of users. There can be different degrees of transparency.

Developments, Research, Challenges and Barriers

Various definitions of open data have emerged in the last decade, which mainly differ with regard to whom provides the data (the public and/or the private sector), the way that users can get access to data (for free, against marginal costs, by paying for them, by obtaining permission), the way that data are defined (data, information, content, Public Sector Information [PSI]), the format of the data (e.g. raw data, including or excluding PDFs and linked data, sometimes using the five-star levels of Berners-Lee [4]) and the type of data (e.g. without any restrictions on usage, data of public interest, with or without metadata). In this paper we focus on those open data provided by the public sector which are freely available in all sorts of formats. We use this broad definition, since the papers in this special issue view open data from different perspectives and fit within this broad definition.

In general, a distinction can be made between the three following important aims or desired policy impacts of open government data release.

1) To foster social innovation or economic growth and that might lead to the development of apps and create the possibility to develop new companies, products and services,

2) To improve government services and efficiency and to fight corruption in the administration, and

3) To strengthen democracy and increase transparency (e.g. the publication of politically sensitive budget data).

Multiple aims can be combined. These aims also have to be seen in light of the relationship between open data and political systems. Non-democratic systems are more likely to focus on economic and efficiency goals rather than on transparency or democracy. Yet even high-censorship countries like China or Saudi Arabia are experiencing a growing public demand for open data. China, for example, has in recent years allowed for a greater flow of information and government criticism, and focuses its censorship on the restriction of collective action and the hampering of social mobilization [35]. This did not hinder, however, the increasing number of people calling for open data policies similar to those in the West [ibid]. In Saudi Arabia openness with information is strongly affected by religious and cultural considerations as well as national security concerns, but the Saudi Arabian authorities officially pursue an e-government concept that aims at improving their national economy [57].

The relationship between open data and the political system is not an entirely new phenomenon: Sweden’s Freedom of the Press Act of 1766 created a mandatory framework for the release of government data – triggering a development in most Scandinavian countries that created a close bond between democratic citizenship and a transparent government, for example [61]. It seems reasonable to assume that a long democratic tradition and a positive attitude towards this form of government fostered transparency in these countries, but even fairly young democracies increasingly embrace open data approaches. Slovenia, for example [65], ranks second in some international rankings when it comes to providing access to public information. Nevertheless there is often strong resistance from government agencies to make all the information available, but there is growing public and legislative pressure to ensure and broaden the access to information. It becomes increasingly clear that different political systems have different designs for open data policies that often serves the particular goals of a political regime.
These goals influence the focus on the release of certain types of data. For instance, the aim to foster social innovation and economic growth could lead to not publishing certain politically sensitive data. Much literature e.g., [27], [33], [46], [56], [70] and many policy reports e.g., [18], [19]-[21], [49], [50], [62] are based on the assumption that open data is a tool to enhance transparency. In addition, it is often argued that transparency could lead to better accountability of the government [41], [49]. However, several researchers have also challenged the idea that opening data will result in transparency [3], [34] and the idea that transparency automatically leads to more trust in the government [28]. Research has shown that open data initiatives are not per se perceived as more transparency [5], [6], [13], [25], [54]. The assumption that open data automatically results in transparency is too simple. There are at least four factors which we believe influence open data transparency:

1) The type of data opened,
2) What one can do with the opened data and how they are displayed,
3) The undesired effects of opened data and 4) the costs of open data transparency apart from the systems, resources, capabilities and other means to make sense out of data.

First, as far as the type of data opened is concerned, the data that are currently made available to the public are not typically those data which are expected to increase transparency. For instance, opened data may concern the environment, traffic and transportation. Although these data are important, they usually do not lead to decision making. Data that lead to decision making, such as the details of the budget execution, are usually not released.

Second, what one can do with opened government data and how they are displayed is important. Transparency may require, for instance, finding the open data that one searches for, processing these open data, understanding them and correctly interpreting them. This is complex because much data is available in many different places [12], [42], the required tools for processing data are often not easy to use for average citizens [28] and are often not interoperable [9], [67], [69] and it is difficult to understand datasets especially when limited metadata are available as is the case for many open datasets [40]. Moreover, it is difficult to correctly interpret data, especially in relation to the many other datasets that are released by governments.

Third, transparency may have certain undesired effects [60]. For example, more open data transparency could lead to the identification of persons in datasets through de-anonymization and thus to privacy breaches. Kalidien, Choenni and Meijer [20] describe an example of how a mash-up of social data about the mean age of female sex offenders per city per year could easily expose who this female is, and this data combined with other data could expose the full identity of a person and violate data protection laws. Kulk and Van Loenen also write that “open data policies may […] be in conflict with the individual’s right to information privacy as protected by the Data Protection Directive” [37], p. 196. Another undesired effect is that more available information can lead to less understanding and less trust [60], as illustrated by a quote of Tsoukas [66], p. 835, who states that the ideal of transparency […] undermines the trust that is necessary for an expert system to function effectively. Transparency can, under certain conditions, lead to less trust in the government [28]. Other negative effects could be that more data can lead to misinterpretations and misleading caused by variety in the quality of open data, there can be an overflow of irrelevant data and data can be misused for propaganda and political statements. Open data may not be neutral and objective in nature [36], and may empower mainly those who are already empowered [30]. Whereas Bannister and Connolly [3] state that the literature strongly supports the view that a perception of openness and honesty influences the formation of trusting beliefs p. 20, it is a myth that merely publishing more and more raw data on a website will result in a more open government and in rational decisions [34].

Fourth, open data transparency comes at certain costs for both the opening and use of data. Since it is usually argued that open data should be available for free and there are no sustainably funded initiatives [36], the government needs to highly subsidize open data transparency including open data use. Sound policies are necessary for dealing with this.

Open Data Policies

In this section we first discuss definitions of policies. Furthermore, development and research with regard to open data policies is presented, and an overview is given of challenges and barriers surrounding open data policies.

Defining Policies

The literature provides various definitions of what constitutes a policy. Dye defines a policy as “whatever government chooses to do or not to do” [17]. Policies have also been defined as the product of the activity of an authority invested with public power and governmental legitimacy [45] and as programmes of action specific to one or more public or governmental authorities within a sector of society or public area [64]. Lemieux [38] writes that policies are the product of activities aimed at the resolution of public problems in the environment by political actors whose relationships are structured. Anderson [1] p. 5 defines a policy as a purposive course of action followed by an actor
or set of actors in dealing with a problem or matter of concern. Thus, there are different perspectives on what constitutes a policy in general. Equally there are differences on what constitutes an open data policy.

Several authors have stated that the so-called policy cycle comprises four stages including problem identification, policy formulation, policy implementation and policy evaluation [45]. The last stage feeds the first one. Stewart Jr, Hedge and Lester [59] explicitly refer to the policy cycle with two additional stages, namely policy change and policy termination [59]. As far as open data policies are concerned, some of these stages have been addressed more extensively than others. For instance, there has been attention for agenda setting and policy formulation in various reports of the European Commission e.g., [18], [19], [20], [22], [24], the United States of America e.g., [48], [49]-[51], various federal governments e.g., [10], [68] and the literature [32], [33]. In contrast, less attention has been paid to policy implementations (although there are some exceptions, such as [53]), policy evaluations and policy changes or determinations in the field of open data. Despite the challenges, it is important to pay attention to the less addressed policy development stages regarding open data, since not much effect can be expected from formulated open data policies if they have not been implemented effectively. Furthermore, the effects of open data policies need to be evaluated, since this provides insights in the advantages and disadvantages of open data and the public value they may create. As a consequence it becomes possible to assess whether open data policies should be changed or even terminated.

Developments, Research, Challenges and Barriers

With regard to the first two stages of the policy cycle - problem identification and policy formulation – the following events should be mentioned. An important event in this context was the release of the European Union (EU) Public Sector Information (PSI) directive in 2003, in which a common legislative framework was presented, which regulates making data of public sector bodies available for re-use [18]. In 2009, the Obama Administration stated that its primary goal was the establishment of an unprecedented level of openness of the Government [48] and published an Open Government Directive some months afterwards [49]. In 2010 Prime Minister David Cameron of the United Kingdom sent a letter to Government departments on plans to open up Government data [10]. Building on former policies, the European Commission has presented an Open Data Strategy for Europe, in which more evident rules on making the best use of government-held information are presented [19]. Subsequently, the Obama Administration published the Digital Government Strategy in 2012, which aims to enable the American people to access high-quality digital government information and services anywhere, anytime, on any device; seize the opportunity to procure and manage devices, applications, and data in smart, secure and affordable ways; and unlock the power of government data to spur innovation and improve the quality of services for the American people [50]. Then in May 2013 the Obama Administration published another open data related memorandum [51] and released the Historic Open Data Rules to Enhance Government Efficiency and Fuel Economic Growth [62]. Thereafter the European Parliament formally adopted the updated EU open data policy in June 2013, [22]. Moreover, in June 2013 the G8 released the Open Data Charter [24]. The Charter emphasizes the significance of open data and endorses the principle of open by default, open to all and usable by machines as well as humans.

In addition to these European and American open data policies, various countries have formulated open data policies on a federal level for instance, [11], [58], [68] and/or on lower levels of bureaucracy for example, [16], such as ministries, municipalities and other public organizations [71]. There is limited research on differences between open data policies [72] and it is argued that more knowledge on open data policies is required [32]. Additionally, not all countries world-wide have adopted open data policies. In the following we describe some research that has been conducted on open data policies.

Regarding policy implementation and evaluation, Huijboom and Van den Broek [32] and Rothenberg [55] investigated countries that already have a well-defined open data policy. From a comparison of five open data policies at a national level Huijboom and Broek [32] concluded that sound evidence of the precise effects [of open data policies] is lacking (e.g. economic, social and democratic effects) [32], p. 1. Moreover, these authors write that “the acquisition of more knowledge could strengthen a well-informed debate, remove governments’ reluctance to invest in open data strategies and help them to develop an effective policy” [32], p. 1. Research of Rothenberg [55] on open data policies in four countries revealed that an important discussion about the development of open data policies relates to raw data now versus misuse of data. On the one hand based on user feedback it is advocated to release raw or imperfect data first and worry about metadata, data quality and other aspects later. On the other hand, there are arguments that issues related to privacy, quality, and semantics must be addressed first to avoid the misuse or abuse of open data, for instance, by providing sufficient documentation and metadata. This constitutes an important policy choice; where to start and under what conditions.

Research has also been conducted on implementation and evaluation of open data policies that are at an early development stage. For instance, Nugroho [47] compared five national open data policies in various development stages and derived lessons for an Indonesian open data policy. She concluded that Indonesia can learn from other countries by developing a more robust legal framework, creating an ecosystem between data publishers and data users, developing stronger Information Technology and organizational support for open data, and launching initiatives that use open data at the district government levels. Nugroho [47], p. ii also writes that the focus of the policies for countries in the developing stages are more related to the release of data from the publishers and less on
the technical processes that are involved with opening the data and identifies several forces and counter forces that affect the differences in open data policies between the countries.

Not only the level on which open data policies are formulated may differ, other aspects can also vary. Previous research has shown that open data policies are mainly different with regard to policy objectives, policy instruments, principles for opening data, the amounts and types of publicized and unpublicized data, the metadata that is provided with the open data, support for the use of the open data, the types of usage of publicized data, and expectations surrounding the positive effects and the risks [71]. This research also showed that “some organizations are truly motivated to become more open by creating an open data policy, whereas others seem to view the creation of an open data policy more as an obligation and are wary of its risks” [71], p. 22. Reasons for this might be found in an organization’s estimation of the risks of becoming truly open, such as risks on wrongful use of their data, the type of data that they produce and the risks on violating data protection laws. Gaining more insight into the differences of open data policies may be useful to help formulating open data policies.

As far as the implementation phase of policies is concerned, challenges concern both the provision and use of open data. For instance, how can departments be convinced or forced to publish government data? And how can the use of open data be stimulated? If the published data are not used, departments may stop to continue publishing data or departments which did not provide data will never start releasing them. Open data publication and use are in this way highly related, yet most open data policies seem mainly to address open data provision rather than open data use. This may mean that open data policies are more focused on transparency than on creating economic value.

Furthermore, the implementation of open data policies is difficult, because many open data policies are formulated on a high-level so that they cover all sorts of different organizations. This may sometimes be necessary because this allows for the interpretation of the policies by different organizations, but on the other hand this also makes it difficult to implement open data policies. Policies are also formulated on different levels. There is a multiplicity of open data policies at various levels of government [71]. There is no harmonized legal framework across countries [72]. Moreover, implementing an open data policy can be time-consuming and requires sufficient resources, and open data policies may shift over time.

Regarding open data policy evaluations, Bertot, McDermott and Smith [7] write that “the Open Government Directive did not offer agencies guidance on how to assess the attainment of open and transparent government through the implementation of their plans” [7], p. 2491. Furthermore, in [7] it was shown that there is a lack of evaluative metrics for open government. Moreover, open data policies, including their objectives and mission, can be very different among organizations and countries and cover different aspects, which brings certain challenges [7], such as which licenses apply when one wants to combine data that are released as part of different open data policies. Finally, sound evidence of the precise effects of open data policies is lacking [32].

Several organizations have also recently started to evaluate and to investigate maturity stages of open government data. For example, a number of indexes for measuring and evaluating the progress of open data adoption in different countries have emerged recently, such as the Open Data Readiness Assessment (ODRA) tool of the World Bank’s Open Government Data Working Group [63], the Open Data Barometer (ODB) 2013 of the Open Data Institute and the World Wide Web Foundation [52] and the Open Data Index 2013 developed by the Open Knowledge Foundation (OKFN) (Site 1). Although these indexes do not necessarily measure all aspects of open data policies, they investigate certain aspects, such as the availability of ten types of high value datasets (the Open Data Index) and various aspects of the readiness, implementation and impact of open data policies around the world (Open Data Barometer).

**Paper Overview**

This special issue covers a diversity of subjects related to open data transparency and open data policies. Six papers have been selected for this special issue. All of these papers went through a rigorous blind review process and were reviewed by three reviewers. The papers can be classified in two areas, namely 1) transparency, legislation, policies and institutional aspects and 2) designs and applications.

**Transparency, Legislation, Policies, and Institutional Aspects**

Three papers in this special issue are related to legislation, policies and institutional aspects.

- **Legal and Institutional Challenges for Opening Data across Public Sectors: Towards Common Policy Solutions** - Melanie Dulong de Rosnay and Katleen Janssens. This manuscript reports on issues and challenges related to the accessibility and reusability of public sector information. The authors first describe the current state of initiatives regarding geographic and environment data, scientific data and cultural data, and then derive common issues and assets from this description. The issues are categorised and possible solutions are identified, including the establishment of a right to reuse focusing on availability and the harmonization of formal requirements. The authors conclude that common solutions can be designed for barriers and challenges related to the accessibility and reusability of public sector information.
Diffusion of Open Data and Crowdsourcing among Heritage Institutions: Results of a pilot survey from Switzerland - Beat Estermann. This article examines the adoption of open data strategies and engagement in crowdsourcing practices for Swiss heritage organizations. Moreover, it investigates the perceived risks, opportunities, the expected benefits and the hindering factors of these strategies. The innovation diffusion approach is used as a theoretical lens. An online questionnaire was developed which was completed by 72 respondents from 65 different heritage institutions of national significance in the German-speaking part of Switzerland. It was found that very few organizations have adopted an open data policy, and that although more institutions engage in crowdsourcing practices, more organizations appear to be enthusiastic about open data. This study also provides detailed insights regarding the results in the light of the innovation diffusion model.

Reconciling Contradictions of Open Data Regarding Transparency, Privacy, Security and Trust - Ronald Meijer, Peter Conradie, Sunil Choenni. This paper addresses the gap of research on public value creation through open data and focuses on contradictions between public values of open data. From a literature review four types of public value are identified, namely transparency, trust, privacy and security. The authors state that these public values may conflict with releasing government data. The authors also identify precommitment as a policy instrument to reconcile conflicting values. Subsequently, they describe two case studies conducted at Dutch governmental organizations (a governmental ministerial research centre and a large municipality) to show how this policy instrument can be implemented to reconcile conflicting public values.

Designs and Applications

The following three papers have been accepted in the category designs and applications.

A Technical Framework for Data Sharing – Wout Hofman and Madan Rajagopal. The fifth paper in this special issue provides an overview of the required open data functionalities derived from a literature review including data source management, data preparation and other functionalities. A technical framework has been derived from the required functionalities, which was used to analyze platforms which function as intermediaries between a data source and data consumer. The analysis shows that most platforms do not meet the requirements that were found in the literature and that the market for this type of platforms is not yet mature. Additional investments are required to enable large scale implementations of open government data.

Similarities of Open Data and Open Source: Impacts on Business - Juho Lindman. The fifth paper in this special issue compares open data and open source to see what can be learned from open source to mature the field of open data. The paper draws on fieldwork and interviews with software entrepreneurs and managers who work for open data businesses and open source businesses. Similarities of open data and open source companies were found for the business model elements revenue, offering, resources and relationships. For instance, it was found that in Finland open data companies and open source companies compete in the same environment and that both companies build services on top of public goods. The manuscript also provides an overview of research and management implications with guidelines for maturing the field of open data with regard to pricing, usability, availability, maintenance, reputation, marketing of the service and legal issues and other topics.

Open Traffic Data for Future Service Innovation – Addressing the Privacy Challenges of Driving Data - Anna Rohunen, Jouni Markkula, Mariak Heikkilä, and Jukka Heikkilä. This paper addresses the gap of studies on privacy behavior from the view point of creating open data to generate new service innovations. The authors provide an extensive literature review regarding the effects of privacy concerns on individuals' intentions to use various Internet and e-commerce services. Moreover, stakeholder interviews, user interviews, and a user survey were used to investigate person's privacy concerns and their consent for disclosing personal driving data. It was found that many users are concerned about privacy issues, yet they are willing to disclose their driving data for benefits. The user's willingness to release their data was found to be influenced by informing of personal data processing, trust in organizations of the service ecosystem collecting and using the data, and users' control over their own data. The findings of this research can be used in the design of personal data based services.

This special issue is aimed at contributing to research on open data transparency and open data policies. In this introductory article we discussed definitions, developments, research, challenges and barriers related to open data transparency and open data policies. Relating to research on transparency, we took a broad view on this transparency and viewed it as positioned between broad trends in society. We concluded that research and policy reports have the general assumption that open data can be used as a tool to enhance transparency, while this idea has also been challenged by several researchers. We showed that some research has already been conducted on this topic, yet limited attention has been given to these contrasting findings. We also showed that open data policies are being developed increasingly, but these policies are not very detailed and they barely guide the publication and
use of open data. More balanced policies need to be developed taking into account the opening of data and the risks of open data. The six papers of this special issue contribute to filling these research gaps.

Enjoy reading!

Website List

Site 1: Open Data Index
https://index.okfn.org/country

References


