

Special Issue on Innovation through Open Data - A Review of the State-of-the-Art and an Emerging Research Agenda: Guest Editors' Introduction

Anneke Zuiderwijk¹, Natalie Helbig², J. Ramón Gil-García³, Marijn Janssen⁴
Guest Editors

¹Delft University of Technology, Faculty of Technology, Policy and Management, the Netherlands,
a.m.g.zuiderwijk-vaneijk@tudelft.nl

²Center for Technology in Government / SUNY Albany, USA,
nhelbig@ctg.albany.edu

³Centro de Investigacion y Docencia Economicas (CIDE), Mexico,
joseramon.gil@cide.edu

⁴Delft University of Technology, Faculty of Technology, Policy and Management, the Netherlands,
m.f.w.h.a.janssen@tudelft.nl
May, 2014

For decades good governance scholarship has focused attention on the importance of government openness [26], [34]. Since the 1960s, Freedom of Information (FOI) legislation has formed the backbone of institutional support for opening information and documents [76] and participatory processes [4], [79]. However, FOI represents a passive approach to releasing information. Persons or organizations must still request the information they want, referred to casually, as freedom of information requests. Since the 1990s, publishing documents on websites or using communication technologies to engage citizens in participation processes has signaled a more proactive approach to releasing government information and political engagement. Since 2003, governments have re-envisioned their passive *and* proactive approaches to include an *open data* agenda, [18], [62], where publishing documents and data in open formats, [35], [63] is the preferred way. Collectively, these developments have forged the basis for what has been commonly referred to as the *open government and data* movements [30].

Open data practices and policies are praised for their potential to generate public value, particularly through innovation, economic growth, and transparency, [5], [9], [18], [21], [81]. The nature and character of open data has been hailed for its innovative capacity and transformative power [19], [35], [40], [45], [80]. Various studies have confirmed that proactively releasing public and private data in open formats creates considerable benefits for citizens, researchers, companies and other stakeholders, such as business creation or having the ability to understand public or private problems in new ways through advanced data analytics, [5], [9], [18], [21], [81]. Only a handful of articles examine both the unintended consequences and negative side effects of opening data, [33] and the underlying causal mechanisms that actually lead to the desired open data benefits [5]. Open data research is still in its infancy, and as a result, the extant literature uses limited application and development of theory toward understanding the open data phenomenon. While scholars acknowledge diverse perspectives, it is not clear which theories are most relevant, nor whether a single or integrated theory is needed.

This special issue is part of a series of two special issues about open data. This issue focuses on the relationship between innovation and open data, while the second special issue emphasizes research on open data related to transparency and open data policies. To realize the practical benefits of this transformative practice and to develop theory, more research needs to focus on understanding how innovation occurs through open data activities. The papers in this special issue begin to address this gap. The introductory article discusses the state-of-the-art with respect to understanding the context of open data innovation, developments, challenges and barriers, presents an overview of open data research and outlines emerging research directions.

The Context of Open Data Innovation

Open data not only creates possibilities for private sector innovation, but also for public sector innovation. A substantial body of research concentrates on innovation in the private sector [77], while the public sector is often considered less amenable to innovation [7]. The literature characterizes innovation as a purposive act or set of acts aiming to do something better, to meet a new need, or respond to new circumstances [77]. Innovation can be defined as "a complex, diversified activity with many interacting components, and sources of data need to reflect this" [64]. p. 7. Schumpeter [69] lists several types of innovations, including the introduction of a new product or a change in an existing product; a process innovation that is new to an industry, opening a new market, development of new sources of supply for raw materials or other inputs, or any change in an industrial organization.

Each innovation type is potentially relevant for research involving open government data. Public sector innovation concentrates on using new ideas that create value for society [2] through the creation of new ventures, business models and innovative services. Provisioning data in open formats has the potential to enable different types of innovation, such as new processes in collecting, processing, and using the data to enable changes in transparency, accountability, and collaborative governance. Transparency, efficiency and accountability are mechanisms that drive

the creation of public value from opening data through innovations that create new products, services and models of service delivery, or engagements with citizens. Public sector innovation is not new for example, [32], [55], however, a more conscious and systematic approach to foster innovation has been emphasized recently to address societal challenges [2]. Research demonstrates that innovative activities must be supported by a public sector innovation ecosystem [ibid].

Perspectives on Open Data

Seven different perspectives are currently reflected in the literature, although most articles use a single perspective to study open data. Combining perspectives may be more effective in dealing with the issues related to open data and stimulating innovation. Figure 1 provides a picture of the various perspectives – legislative, political, social, economical, institutional, operational, and technical [37]. A *legislative* perspective emphasizes how open data legislation, including freedom of information acts, open data policies, open government directives, memorandums and declarations, are important. A *political* perspective emphasizes the importance of political developments and political differences between countries. A *social* perspective brings into focus the importance of cultural differences between countries and differences in agendas related to the social benefits of opening data, such as transparency and accountability. An *economical* perspective points at the financial benefits and gain which can be created with open data. It includes studies estimating the impact of reuse of open data, often estimated in billions of euros annually [20]. Innovation is seen as an important driver for stimulating economic growth. An *institutional* perspective examines the ways that institutions enable and constrain the publication and adoption of open data, such as the barriers of use from a data provider's point of view [37]. Institutional analysis reveals the importance of examining how the publication of data could become an integral part of the data collection or creation process, rather than a separate activity which is not integrated in the daily procedures and routines [82], [85]. An *operational* perspective focuses attention on the use of open data and the requirements for being able to use open data. Standards are cited as important for open data operations [78]. Finally, a *technical* perspective emphasizes the importance of open data technologies, open data platforms and open data infrastructures. It highlights the features of open data infrastructures as important, such as visualisation and analysis features [29] and the metadata that are required to enable the reuse of open data [13].

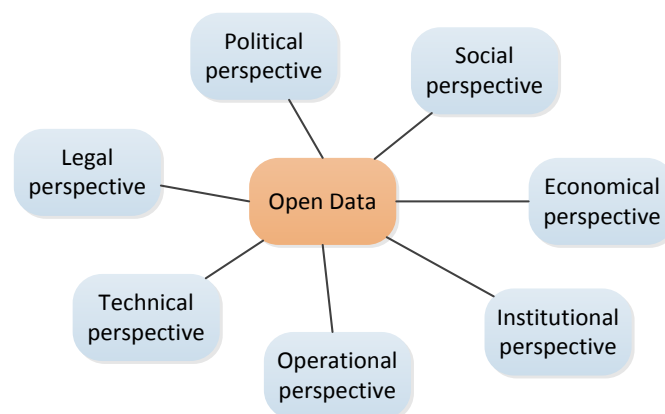


Figure 1: Perspectives on open data

Complexities of Open Data Innovations

Stimulating innovation in general is not easy [43]. Likewise, stimulating innovation through the open data process [86], the efforts required to create, open, find, and use open data, are equally complex. The complexities emerge from several factors, including the large number of actors involved in the process, the variety of social and technical contexts, uncertainty surrounding how open data will be used, and the difficulty valuing intangible impacts generated through open data innovation.

First, many stakeholders are involved in open data processes [12], [33], including open data providers, open data legislators, open data facilitators and many different types of open data users, such as citizens, researchers, journalists, developers, entrepreneurs, archivists or librarians [31]. These stakeholders have various interests and these interests may conflict [33]. For instance, a journalist may want access to and to use sensitive data which he or she would report on in an article through the news media, whereas a government open data provider will typically not publish sensitive data because of various legal and personal privacy constraints. This example demonstrates that the stakeholders are often loosely connected, and in some respects, dependent on each other's activities. The ability for journalists to use open data depends on whether open data providers publish the data. However, each stakeholder remains focused on their own activities, often performing them in non-standardized ways. The non-standardization reduces the interconnectivity between each other's activities and the opportunity to innovate within the open data process is likely lessened [82]. Learning from how users actually use their data is another potential opportunity for

innovation in the process. Currently, governments find it difficult to monitor use and few tools exist to monitor how and in what ways open data is used.

Second, there are a variety of contexts which play a part in the open data process [14], including legal [44] and cultural [59], and the variety of data content and types. Different types of data, with different content, may need a different legal, cultural, or technical treatment. Each context has its own set of characteristics which influences the way that open data are collected, disseminated, used and interpreted. For example, using data on crimes committed in a city will likely need different levels of privacy protection and interpretation of meaning than would data about playground locations within cities or tree types geocoded by city streets.

Third, the publication and use processes of open data are complex [33], [83], [88] and it is not easy to predict how users will use open data, when they will use it, and how it will be used in the future [51]. For example, government agencies may avoid publishing open data as a risk aversion strategy. The risk aversion is driven by their uncertainty in how the data will ultimately be used or combined with other data.

Fourth, understanding how value is created from open data innovation is not straightforward, particularly how public value is generated. Public value is not a new term, but a growing literature examines its definition and characteristics for instance, [42], [53]-[54], [56], [74]. Public value is seen as the product of governmentally produced benefits where public value is derived from the direct usefulness, fairness, and equitability of such benefits to a variety of stakeholders [30]. p. 5. There are many levels of public value observation – individual, group, institutional, and societal [10] and often value created from transparency, accountability, and collaboration are intangible impacts such as trust, well-being, or being more informed. These are important impacts, but they are not easily measured. The uncertainty surrounding the value of innovation through open data makes it a risky investment.

The complex nature of the open data process, and innovation processes, complicate the supply and use of open data and these complexities are often not taken into account by stakeholders. As a result, the potential of the open data process is not completely exploited [82].

Research on Open Data

In the last decade open data research has gained more attention due to the European Public Sector Information directive of 2003 [18], Obama's Memorandum for the Heads of executive Departments and Agencies [62] and other open data policy documents at the international [72] and municipal levels for example, [28], [46]. Figure 2 shows the number of publications on open data from 1974 to 2012 in the databases Scopus, Science Direct, IEEE Xplore and ACM Digital Library. The number of publications refers to those publications that used the combination of the terms *open data* in either the title, abstract, or keywords of the paper. A limitation is that this search does not include searching through the full text of the articles. Scopus reveals articles with the terms open data in the title, keyword or abstract written from 1974, IEEE Xplore from 1983, ACM Digital Library from 1986 and Science Direct from 1990. Most open data publications were found in Scopus. The graph clearly shows an increasing number of open data publications particularly after 2000, with a sharp increase in 2009.

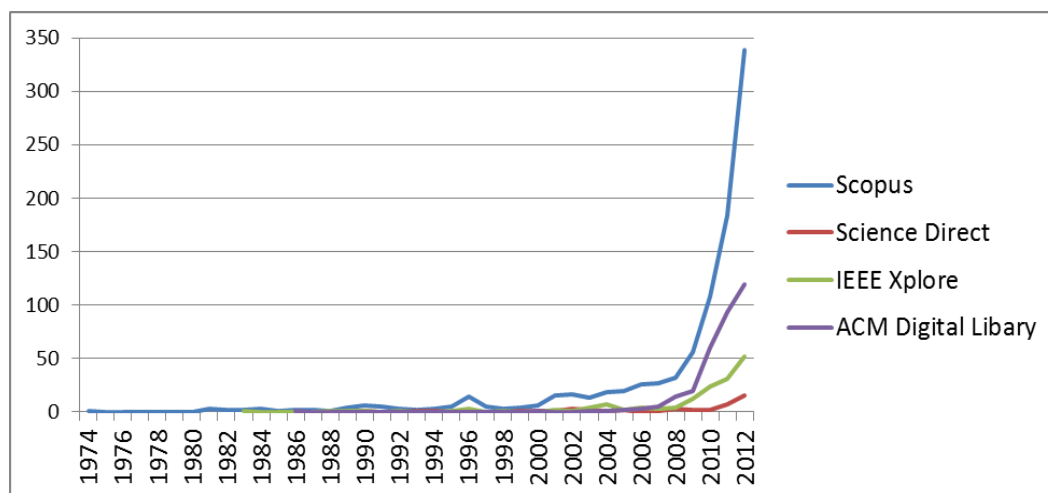


Figure 2: Publications with the term *open data* in the title, abstract, or keywords (measured in October 2013)

Our review of 143 papers on open data revealed that publications were mainly conceptual papers, descriptions of the empirical uses of open data or described the design of technology and systems [37]. The use of theory to explain changes or patterns received considerably less attention. The papers for review were collected by searching for *open data*, *open link data* and *open link government data* in various data bases, including the data bases mentioned above. The Scopus data base was used most frequently because our previous analysis showed that out of the four

examined data bases, Scopus contained the most articles about open data. To examine whether and what type of theory was used in a paper, we searched for the term *theory* in each paper and scanned the abstract and conclusion sections. Out of the 143 papers, only 19 papers explicitly used, stated, or developed theories.

Table 1 summarizes which theories are mentioned, used and developed and the related topics examined in the open data field in the 19 papers. We operationalized *the mentioning of theory* as any mention of that theory at least once. We operationalized *the use or extension of theory* as any paper that refers to the use of a theory by applying it to a certain open data topic.

We found that of the 19 articles, many were written by the same authors and that only a very small group of researchers are involved in theory development in the field to-date. Out of the 24 authors that were involved in writing these 19 papers, 8 authors were involved in writing two or more of the 19 papers. From this group, however, they have explored many different types of theories, and only rarely was the same theory used more than twice. Institutional and organizational theories and democratic theories were used several times. Institutional and organizational theories were used to investigate policy development, changing systems, and changing organizational cultures and structures. Democratic theories were used to investigate transparency and trust and participation. In addition, collaboration between different actors involved in publishing and using open data was a topic that had been investigated often, as compared to the other topics.

Table 1: Summary of open data topics and theories mentioned, used or extended in 19 of 143 articles reviewed

Open Data Topics	Theory Mentioned	Theory Used / Extended
Access to data	Mass media theories [5] Theory of justice [48]	
Accountability		Theories of public accountability [50] Stakeholder theory [50] Deutsch' [16] theory of the nerves of government [50]
Collaboration between public and private organizations to support a smart disclosure policy		Institutional theory [68]
Context and expectations	Theory of contextual integrity [60]	
Coordination mechanisms for open data activities		Coordination theory [82]
Data sharing, information sharing and knowledge sharing	Motivation theory [67]	General theory of interagency information sharing [11] Theory of planned behavior [67]
Economics of information	Theory of natural monopolies [58] Microeconomic theory [58] Competitive theory [58] Utility theory [58]	Neoclassical economic theory [58] Social welfare theory [58] Theory of public goods [58]
Information storage and communication	Information theory [47] Theory of digital objects [47]	
Innovation	Actor Network Theory [8] Theory of diffusion [8] Theory of innovation [8]	
Institutional contexts and rational choices		Structuration theory [52]
Moving from closed to open systems, organizational change	Organizational theory [1]	Institutional theory [37]
Network and power relationships between open data owners	Actor-network theory [5]	
Open data adoption	Adoption theories (e.g. Theory of Planned Behavior, Technology Adoption Model, United Theory of Acceptance and Use of Technology) [66]	Diffusion of Innovation theory [15]
Organizational cultures, structures and dynamics in relation to market characteristics	Organizational theory [5]	
Participation	Democratic theory [48]	
Policy		Theory of institutional / organizational isomorphism [84]
Power relations and power structures	Postmodernist work of Foucault [5]	
Reinforcing and transforming institutional structures		Systems theory [37]
Stakeholders		Stakeholder theory [81]
Transparency and trust	Theory of transparency [60]	Social capital theories [61] Rational choice theories [61] Deliberative democracy theories [61]

Upon further investigation of the use of theory in these 19 articles, we found that only one reference could be found in the literature and in almost half of the publications, the theory is mentioned only to justify a statement. For several topics in the area of open data, we could not find papers having a theoretical foundation, such as standard development, efficiency benefits, or economic growth. For instance, our selection of 143 papers did not include any papers with theoretical foundations for the design of open data infrastructures.

There are various potential explanations for why we see these patterns. First, different theories are used in open data research to incorporate the variety of perspectives – legal, social and technical. Different theories are needed for different perspectives on open data. For example, one theory may explain the adoption of open data, but not the process in which open data are published or the way that policies are created. For this reason, only a multi-theory approach can take into account a diversity of perspectives. Second, a diversity of theories for open data is likely to exist because it is an emerging phenomenon and it is not clear which theories are relevant to the field. For example, we found only one paper examining theoretical foundations for the topic “power relations and power structures” [5]. More research is needed to investigate this theory and to show in which context this theory can be used in the field of open data. This special issue contributes to filling this gap, as four out of the seven papers of this special issue describe the use of one or more theories in open data research.

Developments, Challenges and Barriers

Public and private organizations increasingly publish their data on the internet [27], [49], [52]. Open data is the result of publishing data in so-called open data infrastructures. Currently, open data infrastructures are being adopted at a rapid pace by different levels of government [89]. For instance, open data infrastructures have been developed by local, national, and international public and private organizations (e.g. municipalities, ministries, federal governments, the European Commission and non-country specific organizations) [ibid]. Various open data policies have been developed as well. Many organizations want to *jump on the bandwagon* and join the open data community [84]. Yet, the publication and use of open data is still accompanied with many barriers [37], [82]-[83], [86]. In [86], 118 barriers for open data were identified. The barriers were divided into ten categories, namely 1) availability and access, 2) find-ability, 3) usability, 4) understand-ability, 5) quality, 6) linking and combining data, 7) comparability and compatibility, 8) metadata, 9) interaction with the data provider, and 10) opening and uploading. Most barriers concerned the use of open data. This research revealed that little attention is paid to the user perspective, which is likely to inhibit innovation. The users need to generate value from open data [84], [86].

Figure 3 presents examples of barriers for open data processes from the eight perspectives highlighted in Figure 1. The barrier overview is based on the literature review of barriers and interviews and workshops about barriers described in [37], [86] and on other literature about barriers for example, [82]-[83]. There are many diverse barriers in each perspective, suggesting the need for research from each of these perspectives.

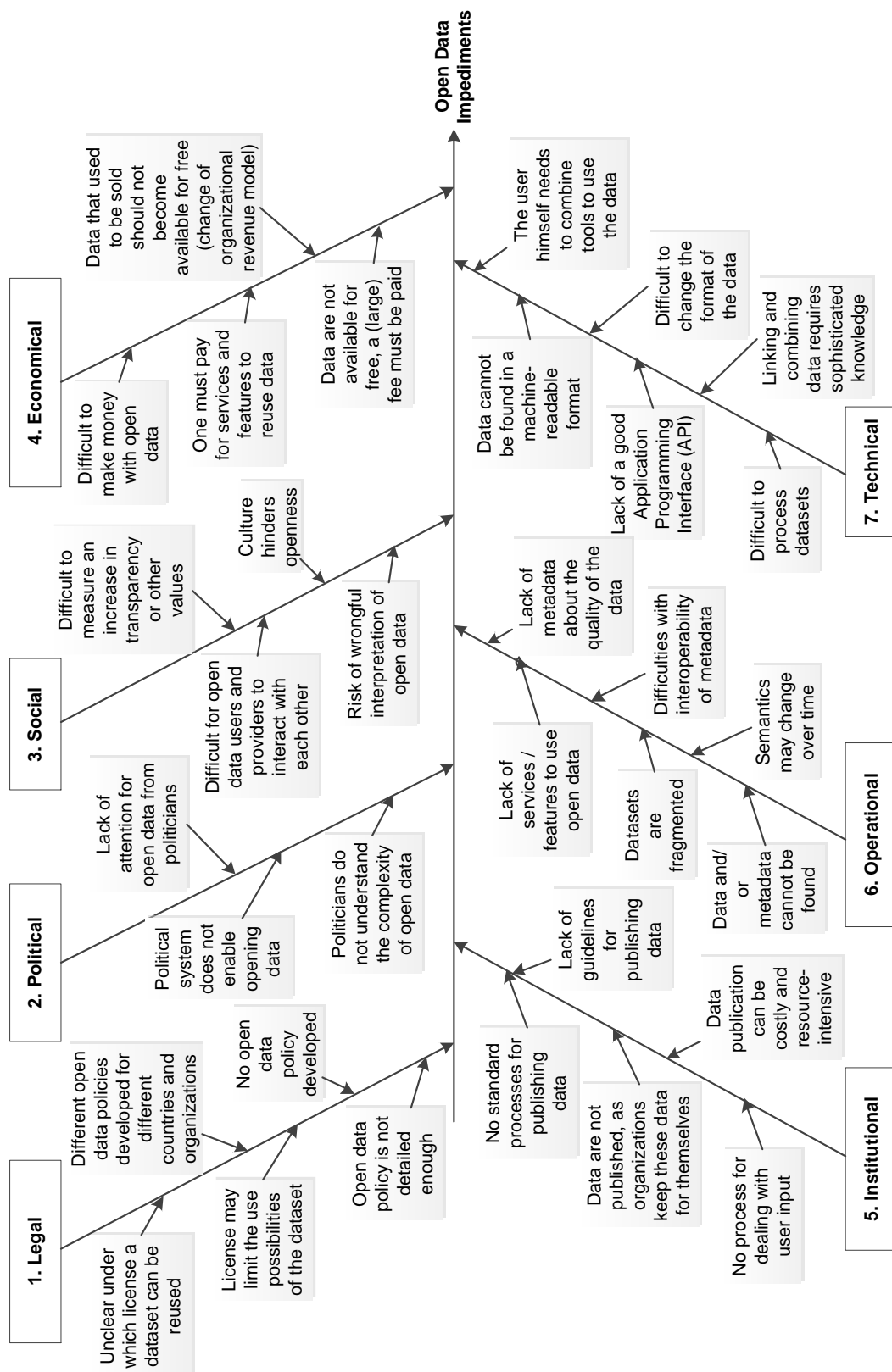


Figure 3: Examples of barriers for open data from various perspectives, adapted from [37], [82]-[83], [86]

Research Directions for Open Data

The previous sections of this introduction highlighted what has been accomplished, but also revealed that significant research remains to be conducted. In this section we identify a number of topics that lead to an emerging open data research agenda categorized by three main topics – 1) open data theory and development, 2) open data policies, use, and innovation, and 3) open data infrastructures and technologies (See Figure 4).

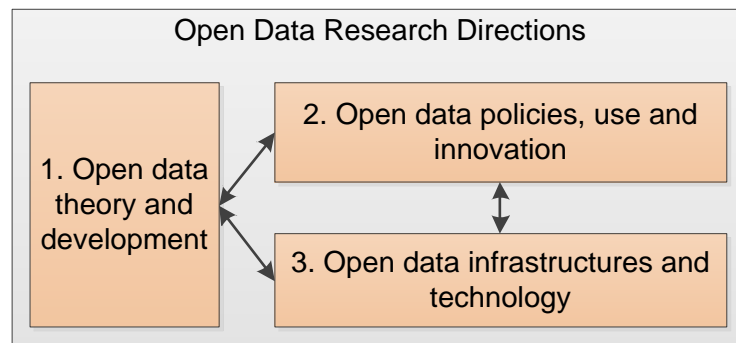


Figure 4: Proposal for open data research directions

Open Data Theory and Development

Table 1 highlighted the few articles and researchers examining the theoretical foundations for open data research. At the same time, the summary also demonstrated the diversity of theories used. Which theories are relevant in which context and to understand which phenomena is not clear yet. Future research should investigate the appropriateness of using the presented theories for open data, the benefit of taking these theoretical views, and to find out whether, and within which context, the theories can be used to understand open data. Additionally, it is important to investigate whether theories can be unified and whether an overarching meta theory can be developed. Furthermore, other theories that have not been investigated to-date may be more appropriate for application in the field of open data, some examples include: the unified theory of acceptance and use of technology, information system (IS) success model theories, organizational information processing theory, the theory of planned behavior and administrative behavior theory. Moreover, design theory of services and infrastructures is a topic that is relevant for creating value from open data innovation and is not addressed to-date.

In sum, only few papers to-date contribute to theory building and there is not a dominating theory. The diversity in the use of theories suggests that there are many opportunities for theory development in the field of open data that have not been explored sufficiently. The field is in its beginning stage and theoretical contributions are still scarce.

Open Data Policies, Use and Innovation

Freedom of information legislation and specific directives represent the legal and political landscape of open data policies to-date. The following is an elaboration of where specific research areas regarding policies, process management, innovation and stimulation of use that we believe will push the field forward.

- **Policies.** Open data policies are in place or under development in many different countries world-wide. From a practical point of view, countries may be able to learn from one another [84] or countries with no policies can learn from those who are leading the way and progressing quickly [59]. More research on open data policies could provide insights in which aspects open data policies should contain in different cultural and legal contexts and how open data policies could be implemented [59].
- **Process management.** Open data are published in different ways and it is unclear which ways are effective, efficient, and innovative. Although some guidelines have been published, further research is necessary to investigate how the processes related to open data publication and use could be managed and which coordination mechanisms are appropriate in the context of open data publication and use.
- **Innovation.** Very little research has been performed on how exactly open data can result in innovation. An example of this research comes from Ferro and Osella [24], who describe eight open data business models which are currently employed by private organizations. In another paper, Ferro and Osella [23] discuss mechanisms which allow for profit-oriented value creation based on open governmental data. Although some research has been performed on which business models could stimulate innovation [22]-[24], [38], the understanding of the open data community on this topic is still limited.

- Stimulating use. Much of the current effort in the field of open data is focused on publishing data on the internet by creating open data portals [49], [52], whereas insufficient attention is given to the problem of how data can be published and fueled in open data portals. Open data policies should focus more on stimulating the use of open data, as this is an important factor in creating the intended effects. The use of open data can be stimulated, for example, by disseminating and sharing open data usage practices among stakeholders and by organizing so-called *hackatons* and other types of competitions [84].

Open Data Infrastructures and Technology

In addition to the social aspects, technical aspects are also very important for open data research. In the following we discuss what kind of research is needed concerning infrastructures and technology.

- Infrastructures and functionalities. There is a lack of knowledge about the functionalities of open data infrastructures that are required to stimulate the innovative use of open data. Research on open data infrastructures and their functionalities can be used for comparisons, but it can also be used by developers of open data infrastructures. Users have different needs and different types of data need to be published and processed in different ways. Some initial research has been performed [87], [89]. For example, Zuiderwijk, Janssen and Jeffery [87] present an open data e-infrastructure which supports data provision, data retrieval and use, data linking, user rating and user cooperation. Moreover, Zuiderwijk, Janssen and Parnia [89] conducted a systematic analysis of three open data infrastructures and identified 35 functionalities of open data infrastructures. Furthermore, they concluded that open data infrastructures can complement each other.
- Interoperability. Janssen, Charalabidis, Kuk and Cresswell [36] concluded that a lack of interoperability results in limited progress in the field of e-government. This also counts for open data interoperability, as there are many problems with interoperability in the field of open data as well, such as problems with the interoperability of metadata standards, licenses, processes and data formats (also see the third paper of this special issue by Petychakis, Vasileiou, Georgis, et al. entitled *A state-of-the-art analysis of current public data landscape from a functional, semantic and technical perspective*). If the tools and applications which support the publication and use of open data are interoperable, then open data providers and users can select and utilize the best functionalities available. If the various open data instruments and tools complement each other and can be integrated, this will enable interoperability. By integrating the elements it should become easier to use open data and this is likely to lead to more users engaging with open data. The use of various tools and applications is a crucial element that stimulates innovation in smart ecosystems for open data.
- Metadata. Metadata are important to stimulate interoperability of open data infrastructures. Metadata may yield considerable benefits [12], [90]. For instance, metadata improve the accessibility of open data for others beyond the primary data providers by describing, locating and retrieving the data efficiently [3], [17], [41], [65], [71], [73]. Metadata improve the ability to find open data, the chances to be found by describing content and becoming searchable [6], [39], [57], [70] and the chances of a correct interpretation of open data by distilling knowledge from them [25], [39], [70], [75]. Currently mainly flat metadata, i.e. metadata for the discovery of datasets, are used and described in open data research, including Dublin Core (DC), CKAN (Comprehensive Knowledge Archive Network), eGMS (e-Government Metadata Standard) and DCAT (Data Catalog). The disadvantage of these flat metadata is that they barely provide contextual information, which is required to interpret open data. More research is needed on the use contextual metadata models, such as CERIF (Common European research Information Format).
- Data quality. Providers and users of open data often do not know what the quality of the data is. This may be a formidable barrier for using open data in meaningful ways. More research is needed on which mechanisms can help in improving data quality and dealing with low data quality, such as a quality rating system for data and so-called *seals of approval* by organizations who published the data. These mechanisms can help potential open data users in assessing the usefulness of the data for their purposes. The fifth paper of this special issue, written by Behkamal, Kahani, Bagheri and Jeremic, contributes to filling the data quality research gap (*A metrics-driven approach for quality assessment of Linked Open Data*).

Paper Overview

This special issue covers a diversity of subjects related broadly to innovation through open data processes. It consists of a collection of papers written by experts in this area. In total seven papers were selected for this special issue. All papers have undergone a rigorous blind review process and were reviewed by three reviewers. The papers can be classified in the following three areas: 1) the open data landscape, stakeholders, challenges and motivations, 2) semantic interoperability, ontologies and data quality, and 3) value creation: evaluation of innovation through open data.

The Open Data Landscape, Stakeholders, Challenges and Motivations

Three papers in this special issue are related to the open data landscape, stakeholders, challenges and motivations.

- Going beyond open data: challenges and motivations for smart disclosure in ethical consumption - Djoko Sigit Sayogo, Jing Zhang, Theresa A. Pardo, Giri K. Tayi, Jana Hrdinova, David F. Andersen, and Luis Felipe Luna-Reyes. This manuscript reports on the findings of a case study on the sustainable coffee supply chain. The case is studied to obtain a better understanding of challenges faced by data producers as well as motivating factors for greater openness. Information was gathered through a workshop and semi-structured interviews with key stakeholders to explore the challenges, motivations and perceived value of data disclosure. The identified challenges and motivating factors were grouped into four general areas influencing smart disclosure, namely: market dynamics, information policies, technological capabilities and data challenges.
- Shaping local open data initiatives: politics and implications - Josefin Lassinanti, Birgitta Bergvall-Kåreborn and Anna Ståhlbröst. This article addresses the question of how local open data initiatives can stimulate innovation through open data. Two in-depth case studies were conducted in the two Swedish municipalities Stockholm and Skellefteå, and a rich narrative is provided about their open data initiatives. The theory of Social Construction of Technology (SCOT) is used to analyze the results of the case studies. The analysis of the cases revealed that their perspectives on open data initiatives were different from each other. Two different drivers for opening up data were identified, namely a techno-economic growth and a co-created societal growth.
- A state-of-the-art analysis of current public data landscape from a functional, semantic and technical perspective – Michael Petychakis, Olga Vasileiou, Charilaos Georgis, Spiros Mouzakis and John Psarras. This paper provides a broad overview of open government data sources from all the countries of the European Union. The authors analyzed the type of data source, the language, the coverage, the license and the metadata standard that are used by each of these open government data sources. Additionally, for the countries United Kingdom, France and Greece an in-depth analysis of the open data sources is performed. The analysis shows that the quality of the investigated government data sources varies significantly depending on the country and the data provider. Moreover, many datasets are not completely open, as they have been published under restricted or non-specified licenses. Despite these limitations of existing open government data portals, the authors conclude that the quality of open government infrastructures is steadily improving.

Semantic Interoperability, Ontologies and Data Quality

The following two papers were accepted in the category semantic interoperability, ontologies and data quality.

- Using a method and tool for hybrid ontology engineering: an evaluation in the Flemish Research Information Space - Christophe Debruyne and Pieter de Leenheer. This paper describes a method to create ontologies in which the stakeholder community becomes an integral part of the ontology and ontology-engineering process, as well as the natural language definitions of concepts. These ontologies are referred to as hybrid ontologies. An experiment was conducted in which the participants used the method to build ontologies to establish semantic interoperability between various research information systems and to annotate the data of an existing system provided by a public administration.
- A metrics-driven approach for quality assessment of linked open data - Behshid Behkamal, Mohsen Kahani, Ebrahim Bagheri and Zoran Jeremic. The fifth paper in this special issue proposes a set of metrics for evaluating the inherent quality characteristics of open data. The metrics can be used to assess datasets before they are released to the Linked Open Data Cloud. Measurement theory and software measurement techniques are used to assess the quality of datasets. Various quality characteristics of datasets can be assessed, for instance, to help data publishing agencies in evaluating their data. Additionally, users of open data can use the quality metrics to assess the quality of the data that they may want to use and to filter out poor quality data.

Value Creation: Evaluation of Innovation through Open Data

Finally, two papers concerning value creation were included in this special issue. They can be summarized as follows.

- Open government data implementation evaluation - Peter Parycek, Johann Höchtel and Michael Ginner. This paper addresses the evaluation of open data initiatives. It evaluates one initiative in

particular, namely the Open Data strategy by the City of Vienna. After this city implemented its open data strategy, the authors conducted qualitative interviews and surveys. The results of this research show that various benefits can be obtained for Viennese open data and a number of success factors were identified. Five recommendations for future open data strategies of other organizations were developed, including the development of an appropriate legal framework, harmonization of data sets, more stakeholder specific target group management, education and training programs and a CC-BY or a comparatively permissive license.

- Data-driven innovation through open government data - Thorhildur Jetzek, Michel Avital, and Niels Bjorn-Andersen. The last paper of this special issue aims to explain how the use of open government data can stimulate the generation of value. A framework with four generative mechanisms is used to explain the complex relationship between openness, data and value. In addition, a critical realist approach is used as a foundation for the in-depth study of open government data in the Opower case. Enabling factors, innovation mechanisms and impacts are identified and the study has resulted in a conceptual model for the data driven innovation mechanism.

This special issue is aimed at contributing to the relationship between innovation and open data. The introductory article discussed the state-of-the-art with respect to understanding the context of open data innovation, developments, challenges and barriers, and presented an overview of open data research and outlined emerging research directions. We showed that the number of publications in the area of innovation and open data is increasing, showing the timeliness and relevance of this topic. Our analysis of open data publications showed that most papers are mainly conceptual papers, descriptions of the empirical uses of open data or descriptions of the design of technology and systems. Limited attention has been given to theory development so far and there is not a dominating theory. There are many barriers blocking innovation and the use of open data. This resulted in a research agenda consisting of three main research themes, namely 1) open data theory and development, 2) open data policies, use, and innovation, and 3) open data infrastructures and technologies.

Enjoy reading!

References

- [1] E. Barry and F. Bannister, Barriers to open data release: A view from the top, in Proceedings 2013 EGPA Annual Conference, Edinburgh, Scotland, United Kingdom, 2013.
- [2] C. Bason, Leading Public Sector Innovation. Co-Creating for a Better Society. Bristol, United Kingdom: The Policy Press, 2010.
- [3] J. C. Bertot, P. T. Jaeger, J. A. Shuler, S. N. Simmons, and J. M. Grimes, Reconciling government documents and e-government: Government information in policy, librarianship, and education, *Government Information Quarterly*, vol. 26, no. 3, pp. 433-436, 2009.
- [4] P. Birkinshaw, *Freedom of Information*. Cambridge: Cambridge: University Press, 2010.
- [5] M. Blakemore and M. Craglia, Access to public-sector information in Europe: Policy, rights and obligations, *The Information Society*, vol. 22, no. 1, pp. 13-24, 2006.
- [6] C. L. Borgman, *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. Cambridge: MIT Press, 2007.
- [7] S. Borins, Encouraging innovation in the public sector, *Journal of Intellectual Capital*, vol. 2, no. 3, pp. 310 - 319, 2001.
- [8] M. T. Borzacchiello and M. Craglia, The impact on innovation of open access to spatial environmental information: A research strategy, *International Journal of Technology Management*, vol. 60, no. 1-2, pp. 114-129, 2012.
- [9] Y. Charalabidis, E. Ntanos and F. Lampathaki, An architectural framework for open governmental data for researchers and citizens, in *Electronic government and electronic participation joint proceedings of ongoing research and projects of IFIP EGOV and ePart 2011*, Delft, the Netherlands, 2011, pp. 77-85.
- [10] A. M. Cresswell and D. S. Sayogo, Developing public value metrics for returns to government ICT investments, *The Research Foundation of State University of New York (SUNY)*, New York, 61268, 2012.
- [11] S. Dawes, Interagency information sharing: Expected benefits, manageable risks, *Journal of Policy Analysis and Management*, vol. 15, no. 3, pp. 377-394, 1996.
- [12] S. Dawes and N. Helbig, Information strategies for open government: Challenges and prospects for deriving public value from government transparency, in *Proceedings 9th International Conference on E-government (EGOV)*, Lausanne, Switzerland, 2010, pp. 50-60.
- [13] S. Dawes, T. Pardo and A. Cresswell, Designing electronic government information access programs: a holistic approach, *Government Information Quarterly*, vol. 21, no.1, pp. 3-23, 2004.
- [14] S. S. Dawes. (2012, June) A realistic look at open data. W3 org. [Online]. Available: http://www.w3.org/2012/06/pmod/pmod2012_submission_38.pdf
- [15] B. Detlor, M. E. Hupfer, U. Ruhi, and L. Zhao, Information quality and community municipal portal use, *Government Information Quarterly*, vol. 30, no. 1, pp. 23-32, 2013.
- [16] K. W. Deutsch, *The Nerves of Government. Model of Political Communication and Control*. New York: The Free Press, 1966.

- [17] E. Duval, W. Hodgins, S. Sutton, and S. L. Weibel, Metadata principles and practicalities, D-lib magazine, vol. 8, no. 4, pp. unknown, 2002.
- [18] European Commission. (2012, December) Directive 2003/98/EC of the European parliament and of the council of 17 November 2003 on the re-use of public sector information. European Commission. [Online]. Available: http://ec.europa.eu/information_society/policy/psi/rules/eu/index_en.htm.
- [19] European Commission, Communication from the commission to the European parliament, the council, the European economic and social committee and the committee of the regions. Open data. An engine for innovation, growth and transparent governance, European Commission, Brussels, COM(2011) 882 final, 2011.
- [20] European Commission. (2013, March) Digital agenda: Commission's open data strategy, questions & answers. [Online]. Available: <http://europa.eu/rapid/pressReleasesAction.do?reference=MEMO/11/891&format=HTML&aged=1&language=EN&guiLanguage=en>.
- [21] European Commission, Digital agenda: Turning government data into gold, European Commission, Brussels, P/11/1524, 2011.
- [22] E. Ferro. (2013, June) Emerging business models in PSI reuse. Enrico Ferro Blogspot. [Online]. Available: <http://enricoferro.blogspot.nl/2012/04/emerging-business-models-in-psi-reuse.html>.
- [23] E. Ferro, and M. Osella. (2012, June) Business models for PSI Re-Use: A multidimensional framework. W3 org. [Online]. Available: http://www.w3.org/2012/06/pmod/pmod2012_submission_16.pdf.
- [24] E. Ferro and M. Osella. (2013, April) Eight business model archetypes for PSI re-use. W3 org [Online]. Available: <http://www.w3.org/2013/04/odw/papers>.
- [25] M. Foulonneau and T. W. Cole. (2012, December) Strategies for reprocessing aggregated metadata. IMLS DCC. [Online]. Available: <http://imlsdcc.grainger.uiuc.edu/docs/metadataprocessing.pdf>.
- [26] S. Ganapati and C. G. Reddick, Open e-government in U.S. state governments: Survey evidence from Chief Information Officers, Government Information Quarterly, vol. 29, no. 2, pp. 115-122, 2012.
- [27] C. P. Geiger, and J. v. Lucke, Open government and (linked) (open) (government) (data), Journal of e-Democracy and Open Government, vol. 4, no. 2, pp. 265-278, 2012.
- [28] Gemeente Amsterdam. (2013, December) Amsterdam Open Data. [Online]. Available: <http://www.amsterdamopendata.nl/home>
- [29] Y. Glickman, M. Flügge, and V. Pterneas. Deliverable D7.7.6. Data provision methods and components, http://www.engage-project.eu/wp/?page_id=212.
- [30] T. M. Harrison, S. Guerrero, G. B. Burke, M. Cook, A. Cresswell, N. Helbig, J. Hrdinova, and T. Pardo, Open government and e-government: Democratic challenges from a public value perspective, Information Polity, vol. 17, no. 2, pp. 83-97, 2012.
- [31] T. M. Harrison, T. A. Pardo, and M. Cook, Creating open government ecosystems: A research and development agenda, Future Internet, vol. 4, no. 4, pp. 900-928, 2012.
- [32] J. Hartley, Innovation in governance and public services: Past and present, Public money and management, vol. 25, no. 1, pp. 27-34, 2005.
- [33] N. Helbig, A. M. Cresswell, G. B. Burke, and L. Luna-Reyes, The dynamics of opening government data. A white paper, Center for Technology in Government, University at Albany, State University of New York, New York, USA, 2012.
- [34] M. Hunt and R. A. Chapman, Open government and freedom of information, in Open government in a theoretical and practical context (R. A. Chapman and M. Hunt, Eds.). Hampshire: Ashgate Publishing Limited, 2006, pp. 1-10.
- [35] K. Janssen, The influence of the PSI directive on open government data: An overview of recent developments, Government Information Quarterly, vol. 28, no. 4, pp. 446-456, 2011.
- [36] M. Janssen, Y. Charalabidis, G. Kuk, and T. Cresswell, Guest Editors' Introduction: E-government Interoperability, Infrastructure and Architecture: State-of-the-art and Challenges, Journal of Theoretical and Applied Electronic Commerce Research, vol. 6, no. 1, pp. I-VIII, 2011.
- [37] M. Janssen, Y. Charalabidis, and A. Zuiderwijk, Benefits, adoption barriers and myths of open data and open government, Information Systems Management, vol. 29, no. 4, pp. 258-268, 2012.
- [38] M. Janssen and A. Zuiderwijk, Infomediary Business Models for Connecting Open Data Providers and Users, Social Science Computer Review, vol. 32, no. 5, pp. unknown, 2014.
- [39] K. G. Jeffery, Metadata: The future of information systems, in Information Systems Engineering: State of the Art and Research Themes (J. Brinkkemper, E. Lindencrona and A. Sølvberg, Eds.). London: Springer Verlag, 2000.
- [40] T. Jetzek, M. Avital and N. Bjørn-Andersen, The value of open government data: A strategic analysis framework. 2012 Pre-ICIS Workshop, in Proceedings International Conference on Information Systems (ICIS 2012), Orlando, Florida, United States, 2012.
- [41] A. Joorabchi and A. E. Mahdi, An unsupervised approach to automatic classification of scientific literature utilizing bibliographic metadata, Journal of Information Science, vol. 37, no. 5, pp. 499-514, 2011.
- [42] T. B. Jørgensen and B. Bozeman, Public values: An inventory, Administration & Society, vol. 39, no. 3, pp. 354-381, 2007.
- [43] S. J. Kline and N. Rosenberg, An overview of innovation, in The Positive Sum Strategy: Harnessing Technology for Economic Growth (R. Landau and N. Rosenberg, Eds.). United States of America: National Academy of Sciences, 1986, pp. 275-305.
- [44] S. Kulk and B. v. Loenen, Brave new open data world?, International Journal of Spatial Data Infrastructures Research, vol. 7, no. 1, pp. 196-206, 2012.
- [45] V. Kundra, Digital Fuel of the 21st Century: Innovation through Open Data and the Network Effect, Joan Shorenstein Center on the Press, Politics and Public Policy, 2012.

- [46] La Ville de Paris. (2013, December) Paris data. Open Data. [Online]. Available: <http://opendata.paris.fr/page/home/>.
- [47] J. Lindman, M. Rossi and V. K. Tuunainen, Open data services: Research agenda, in Proceedings 46th Annual Hawaii International Conference on System Sciences (HICSS), Wailea, Maui, Hawaii, 2013, pp. 1239-1246.
- [48] U. Maier-Rabler and S. Huber, Open: the changing relations between citizens, public administration and political authority, eJournal of eDemocracy & Open Government (JeDEM), vol. 3, no. 2, pp. 48-58, 2011.
- [49] P. McDermott, Building open government, Government Information Quarterly, vol. 27, no. 4, pp. 401-413, 2010.
- [50] A. Meijer, Publishing public performance results on the Internet. Do stakeholders use the Internet to hold dutch public service organizations to account?, Government Information Quarterly, vol. 24, no. 1, pp. 165-185, 2007.
- [51] A. Meijer, J. de Hoog, M. van Twist, M. van der Steen, and J. Scherpenisse, Understanding the dynamics of open data. From sweeping statements to complex contextual interactions, in Open Government: Opportunities and Challenges for Public Governance (M. Gasco, Ed.). New York: Springer, 2014, pp. 101-114.
- [52] A. Meijer and M. Thaens, Public information strategies: Making government information available to citizens, Information Polity, vol. 14, no. 1-2, pp. 31-45, 2009.
- [53] T. Meynhardt, Public value inside: What is public value creation?, International Journal of Public Administration, vol. 32, no. 3-4, pp. 192-219, 2009.
- [54] T. Meynhardt and S. Bartholomes, (De)composing public value: In search of basic dimensions and common ground, International Public Management Journal, vol. 14, no. 3, pp. 284-308, 2011.
- [55] L. B. Mohr, Determinants of innovation in organizations, The American Political Science Review, vol. 63, no. 1, pp. 111-126, 1969.
- [56] M. H. Moore, Creating Public Value: Strategic Management in Government, USA: Harvard University Press, 1995.
- [57] National Information Standards Organization, Understanding metadata, Bethesda: National Information Standards Organization Press, 2004.
- [58] K. Nilsen, Economic theory as it applies to public sector information, Annual Review of Information Science and Technology, vol. 44, no.1, pp. 419-489, 2010.
- [59] R. P. Nugroho, A comparison of open data policies in different countries. Lessons learned for an open data policy in Indonesia, Master thesis, Delft University of Technology, The Netherlands, 2013.
- [60] K. O'Hara, Transparent government, not transparent citizens: A report on privacy and transparency for the cabinet office, Gov. UK, London, 272769, 2011.
- [61] K. O'Hara, Transparency, open data and trust in government: shaping the infosphere, in ACM Web Science 2012, 2012.
- [62] B. Obama. (2013, February) Memorandum for the heads of executive departments and agencies: Transparency and open government. The White House. [Online]. Available: http://www.whitehouse.gov/the_press_office/Transparency_and_Open_Government
- [63] B. Obama. (2012, September) Open government directive. The White House. [Online]. Available: http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m10-06.pdf.
- [64] Organisation for Economic Co-operation and Development (OECD), The Measurement of Scientific and Technological Activities Proposed Guidelines for Collecting and Interpreting Technological Innovation Data, Oslo Manual, 2013.
- [65] S. L. Pallickara, S. Pallickara and M. Zupanski, Towards efficient data search and subsetting of large-scale atmospheric datasets, Future Generation Computer Systems, vol. 28, no.1, pp. 112-118, 2012.
- [66] D. S. Sayogo and T. Pardo, Exploring the determinants of scientific data sharing: Understanding the motivation to publish research data, Government Information Quarterly, vol. 30, no. 1, pp. S19-S31, 2013.
- [67] D. S. Sayogo and T. A. Pardo, Exploring the motive for data publication in open data initiative: Linking intention to action, in Proceedings 45th Hawaii International Conference on System Sciences, Hawaii 2012, pp. 2623-2632.
- [68] D. S. Sayogo and T. A. Pardo, Understanding smart data disclosure policy success: The case of green button, in Proceedings 14th Annual International Conference on Digital Government Research, Quebec City, Canada, 2013, pp. 72-81.
- [69] J. Schumpeter, The Theory of Economic Development. Cambridge, Massachusetts: Harvard University Press, 1934.
- [70] N. Schuurman, A. Deshpande and D. Allen, Data integration across borders: a case study of the Abbotsford-Sumas aquifer (British Columbia/Washington State), JAWRA Journal of the American Water Resources Association, vol. 44, no. 4, pp. 921-934, 2008.
- [71] C. Tenopir, S. Allard, K. Douglass, A. U. Aydinoglu, L. Wu, E. Read, M. Manoff, and M. Frame, Data sharing by scientists: Practices and perceptions, PLoS ONE, vol. 6, no. 6, pp. e21101, 2011.
- [72] The World Bank. (2013, December) Access to Information. The World Bank. [Online]. Available: <http://web.worldbank.org/WBSITE/EXTERNAL/PROJECTANDOPERATIONS/EXTINFODISCLOSURE/0,,menuPK:64864911~pagePK:4749265~piPK:4749256~theSitePK:5033734,00.html>.
- [73] United Nations Statistical Commission and Economic Commission for Europe. (2011, December) Guidelines for statistical metadata on the internet. UNECE. [Online]. Available: <http://www.uncece.org/fileadmin/DAM/stats/publications/metadata.pdf>.
- [74] A. F. E. van Veenstra, IT-induced Public Sector Transformation. 's-Hertogenbosch: Uitgeverij BOXPress, 2012.
- [75] M. Vardaki, H. Papageorgiou and F. Pentaris, A statistical metadata model for clinical trials' data management, Computer Methods and Programs in Biomedicine, vol. 9, no. 5, pp. 129-145, 2009.

- [76] A. Whitmore, Extracting knowledge from U.S. department of defense freedom of information act requests with social media, *Government Information Quarterly*, vol. 29 no. 2, pp. 151-157, 2012.
- [77] P. Windrum and P. Koch, *Innovation in Public Sector Services*. Entrepreneurship, creativity and management. Cheltenham, UK: Edward Elgar, 2008.
- [78] World Wide Web Consortium.(2013, July) Standards. W3C. [Online]. Available: <http://www.w3.org/standards/>.
- [79] B. Worthy, More open but not more trusted? The effect of the Freedom of Information Act 2000 on the United Kingdom central government, *Governance*, vol. 23, no. 4, pp. 561-582, 2010.
- [80] Z. Yang and A. Kankanhalli, Innovation in government services: The case of open data, in *Proceedings IFIP WG 8.6 International Working Conference on Transfer and Diffusion of IT, TDIT 2013* Bangalore, India, 2013, pp. 644-651.
- [81] J. Zhang, S. Dawes, and J. Sarkis, Exploring stakeholders' expectations of the benefits and barriers of e-government knowledge sharing, *Journal of Enterprise Information Management*, vol. 18 no. 5, pp. 548-567, 2005.
- [82] A. Zuiderwijk and M. Janssen, A coordination theory perspective to improve the use of open data in policy-making, in *Proceeding 12th Conference on Electronic Government (EGOV)*, Koblenz, Germany, 2013, pp. 38-49.
- [83] A. Zuiderwijk and M. Janssen, Barriers and development directions for the publication and usage of open data: A socio-technical view, in *Open Government, Opportunities and Challenges for Public Governance* (M. Gascó-Hernández, Ed.). New York: Springer, 2014, pp. 115-136.
- [84] A. Zuiderwijk and M. Janssen, Open data policies, their implementation and impact: A comparison framework, *Government Information Quarterly*, vol. 31, no. 1, pp. 17-29, 2014.
- [85] A. Zuiderwijk, M. Janssen, S. Choenni, and R. Meijer, Design principles for improving the process of publishing open data, *Transforming Government: People, Process and Policy*, vol. 8, no. 2, 2014.
- [86] A. Zuiderwijk, M. Janssen, S. Choenni, R. Meijer, and R. Sheikh Alibaks, Socio-technical impediments of open data, *Electronic Journal of eGovernment*, vol. 10, no. 2, pp. 156 - 172, 2012.
- [87] A. Zuiderwijk, M. Janssen and K. Jeffery, Towards an e-infrastructure to support the provision and use of open data, in *Proceedings Conference for e-Democracy and Open Government (CeDEM13)*, Krems an der Donau, Austria, 2013, pp. 275-291.
- [88] A. Zuiderwijk, M. Janssen, R. Meijer, S. Choenni, Y. Charalabidis, and K. Jeffery, Issues and guiding principles for opening governmental judicial research data, in *Proceedings 11th Conference on Electronic Government*, Kristiansand, Norway, 2012, pp. 90-101.
- [89] A. Zuiderwijk, M. Janssen and A. Parnia, The complementarity of open data infrastructures: An analysis of functionalities, in *Proceedings 14th Annual International Conference on Digital Government Research (dg.o 2013)*, Quebec, Canada, 2013, pp. 166-171.
- [90] A. Zuiderwijk, K. Jeffery, and M. Janssen, The potential of metadata for linked open data and its value for users and publishers, *Journal of e-Democracy and Open Government*, vol. 4, no. 2, pp. 222-244, 2012.