Understanding the evolution of open government data research: towards open data sustainability and smartness

Yingying Gao
Huazhong University of Science and Technology, China, and Delft University of Technology, The Netherlands

Marijn Janssen
Delft University of Technology, The Netherlands

Congcong Zhang
Huazhong University of Science and Technology, China

Abstract
The past decade has witnessed a rapid development of open government data practices and academic research. However, there is no systematic survey of existing research to understand the evolution of open government data. Such research can facilitate knowledge transfer within and across domains, and foster learning for countries in the early stages of open government data development. This study quantitatively extracted the evolution trajectory of open government data based on the main path analysis method and then analysed the underlying motivations. The results show that open government data research went through four main phases and that the open government data movement has spread towards developing countries and smart cities. Different challenges and issues faced by the researchers in each phase drove the evolution of open government data research. Finally, we discuss future directions of open government data research.

Corresponding author:
Marijn Janssen, Delft University of Technology, Faculty of Technology, Policy & Management, Jaffalaan 5, Delft 2628BX, The Netherlands.
Email: M.F.W.H.A.Janssen@tudelft.nl
data research based on our findings and recent development. There is a tendency to create sustainable open government data and smartness by employing artificial intelligence and creating data marketplaces.

Points for practitioners

Open government data efforts have evolved over the years into a global phenomenon. Countries have learned from each other and more and more efforts are focused on innovating with open government data by stimulating co-creation and using other incentives. The way that data are opened should focus on achieving goals like innovation, participation, transparency and accountability. There is a tendency to create sustainable open government data and smartness by employing artificial intelligence and creating data marketplaces.

Keywords
bibliometric analysis, open data smartness and sustainability, open government, systematic literature research

Introduction

Open government data (OGD) initiatives have spread rapidly in recent years (Johnson and Robinson, 2014). OGD refers to publishing public sector information in open and reusable formats, without restriction or monetary charge for use by society (Kalampokis et al., 2011). This movement’s main purpose is to ensure transparent administration and stimulate citizen participation and engagement. Besides this, OGD can help to generate public value through innovation (Janssen et al., 2012; Zuiderwijk and Janssen, 2014a). Under these motivations, the European Union (EU) and the US took the lead in launching OGD activities. There has also been a recent spurt of OGD initiatives in other parts of the globe, where both the number of OGD websites and data sets have increased steadily (Saxena, 2017).

Similarly, academic research about OGD has also developed at a fast pace. There is currently an abundance of studies on OGD. For example, some publications have examined the factors that triggered the adoption of OGD initiatives across government agencies (Coglianese, 2009; Hessain and Chan, 2015; Zhao and Fan, 2018), while some studies have explored the innovation activities driven by OGD (Mergel et al., 2018), and some have evaluated the quality and effects of OGD initiatives (Sayogo et al., 2014; Zuiderwijk et al., 2019).

There have been several literature reviews on OGD. To provide guidance for future research development, Hessain et al. (2016) assessed the status of OGD research from three levels and proposed future directions, whereas Attard et al. (2015) investigated existing OGD tools and approaches, and extracted the
challenges and issues that hinder initiatives from achieving their full potential. Finally, Safarov et al. (2017) attempted to frame the utilization of open government data to suggest future research directions. These reviews are beneficial to research development but no attempt has been made to understand research evolution.

Understanding the evolution of OGD research can help researchers comprehend the key developments, why certain developments happened and how OGD will develop in the future, facilitating knowledge transfer within and across domains (Barbieri et al., 2016; Liang et al., 2016). OGD initiatives were launched at different times in different countries, and countries might address different challenges. Research development derived from the evolution of studies in Western countries may help OGD action in developing countries (Saxena, 2018), despite the different drivers and situations.

This study carries out main path analysis to trace the OGD evolution trajectory from 2000 to 2019. The results of this are used to scrutinize the research topics, theoretical foundations and research motivations of previous research. Main path analysis, combined with motivation analysis, was used to identify emerging research frontiers. The following research questions are specifically addressed:

- How has OGD research evolved over time?
- What are the motivations underlying the evolution trajectory?
- What further development and research needs are anticipated?

The second section of this article introduces the research method and the data-retrieval process. The third section presents the results of the main path analysis and motivation analysis. The fourth section contains a discussion of the findings and a future research agenda. Conclusions, implications for research and practice, and limitations are presented in the fifth section.

**Research method and data**

**Research method**

To achieve the aforementioned research aims, we employed two research methods: bibliometric and document analysis. Figure 1 shows how the two methods work in conjunction with each other. The bibliometric analysis is used to obtain the main path of the research domain, while the document analysis is used to analyse the topics of the core articles on the main path and the research motivations. Finally, based on the analysis results and the latest research progress, future research progress is proposed.

Bibliometric analysis is a method that can be used to analyse the scientific development processes and structural relationships of a research domain based on citation, co-citation and authors (Chen, 2004). The citations in publications contain rich information on how knowledge disseminates or flows
Main path analysis is a common citation-based technique that aims to extract evolutionary trajectories from citation networks.

The first step is the construction of an article citation network. Thereafter, weights of the links in the network are calculated. One measure suggested by Hummon and Dereian (1989) is the Search Path Link Count (SPLC). The SPLC of a directed link counts all possible paths in the network traversed through the link (Batagelj, 2003). The SPLC is considered as the best measure for traversal weight since a higher SPLC value indicates a higher number of search paths (Batagelj, 2003; Hummon et al., 1990). Finally, the main path was extracted based on the traversal weight of links. Existing studies have introduced several search algorithms for selecting the most important links, namely, the local approach, global main path and key-route search methods (Liu and Lu, 2012). The local approach chooses the links with the largest traversal count from the current start point, while the global main path is the path that has the largest overall traversal count (Liu and Lu, 2012). In comparison with the local main path, the global main path emphasizes the overall importance of knowledge flow. The key-route search enables the inclusion of the top links and selects them as seeds to search backward and forward. The key-route search can be either local or global (Liu et al., 2019). To obtain the main evolution path, we constructed the **global main path by the key-route search**, as implemented in the software package PAJEK (Batagelj and Mrvar, 1998). The software enables us to select the number of important links to be 10, 20 or 30. As the number increases, the main path is gradually blurred by less significant citation links (Chen et al., 2019). Therefore, we extracted a relatively clear trajectory based on the top 10 links.

The analysis of the research topics of the core nodes in the main path can shed light on the evolution of research activities in the area (Bindu et al., 2019). One of the commonly used methods of topic modelling is based on the latent Dirichlet allocation (LDA) model, which is designed to process large archives of documents based on latent topics characterized by a distribution over words (Blei and Jordan, 2003). However, it is theoretically impossible to identify topics from a small number of documents (Tang et al., 2014). Moreover, there is no best way to decide the right number of topics (Arun et al., 2010). We selected document analysis to capture a diversity of topics and to conduct the motivation analysis.

**Document analysis** is a systematic procedure for reviewing or evaluating documents to develop an understanding of a research question. It provides a way of tracking change and development because researchers can identify subtle changes by comparing them (Bowen, 2009). This method was also adopted to conduct **motivation analysis**, which aims to identify what kind of issue or challenge drives the researcher to conduct the research. The motivations of articles with similar topics were carefully reviewed. Finally, motivations of each phase were summarized.
Data collection

The literature analysis starts from selecting appropriate data sources and metadata. The Web of Science (WoS) is widely considered the most comprehensive and high-quality database for scholarly work, as it indexes thousands of prominent journals (Dahlander and Gann, 2010). Therefore, we adopted the WoS Core Collection as our data source.

Establishing search keywords is particularly crucial to obtaining an accurate literature dataset (Lu and Liu, 2014). First, we searched the articles in the WoS and obtained a small number of studies. Then, the title and abstract of these articles were analysed to obtain a list of candidate terms. After consultation with experts and discussion among the authors, we finally determined ‘open govern*’, ‘open government data’ and ‘open data’ as final search terms. Considering that the EU published the Reuse of Public Sector Information Directive (PSI Directive) in 2003 and research articles come earlier than practice, we retrieved the literature published from 2000 to 2019. To reduce the risk of bias, we used the following inclusion and exclusion criteria to select articles for use in our review:

- we only considered articles and conference papers that were written in English; and
- we only included government data publication studies, so studies on other open data domains such as research data were excluded.

We then performed a manual study selection by reading the titles and abstracts. This resulted in a total of 1008 publications obtained for analysis.

Results

We retrieved the literature data about OGD from the WoS between 2000 and 2019. As shown in Figure 2, the period from 2000 to 2008 did not yield much literature; however, the previous research explored open data and the open government issue (e.g. Brooks, 2008; Horsley, 2006). The number of articles increased significantly in subsequent years, which can be linked to the Open Government Directive at the end of 2009. Thereafter, OGD research received widespread attention.

Tracing the main evolution path

Our first question concerns the evolution of OGD over time. We obtained the global main path shown in Figure 3. After that, we analysed the topic evolution of articles in the main path using the document analysis method. Table 1 lists the detail of milestone articles on the main path.

According to the difference in research topics, the evolution’s main path can be classified in four phases, starting with the launching of OGD initiatives and turning to evaluation and learning due to a lack of value. Motivated by the lagging behind of the actual use of OGD, researchers started investigating the adoption and actual
usage of data. Finally, the focus turned to implementation and comparison among countries to foster learning among countries.

**Phase 1: OGD launch.** The research emphasis in the first phase of OGD evolution was on the progress of the open data movement in the EU followed by the US and transforming conventional administrative systems into more open systems. The research was focused on different issues. The first branch was started by Coglianese (2009), who mainly discussed what level of transparency was needed in US OGD initiatives. This was followed by Peled (2011), who evaluated the effect of OGD in the US and found the challenge of transparency and cooperation between institutions. The other branch by K. Janssen (2011) gave an overview of the policies and regulation of open data in EU member states such as Sweden, the Netherlands and others. The last branch reviewed the laws and regulations supporting open government building from a legal perspective, as presented by McDermott (2010).

**Phase 2: Evaluation and learning.** Following the development of the OGD movement, the next issue was the evaluation of OGD efforts. The major task in the second phase was examining the anticipated benefits, the implementation barriers and performance. The research theme of Janssen et al. (2012) went beyond focusing on individual projects in the first phase and analysed the benefits of and barriers to open data systematically by synthesizing people’s experiences. This study was extended to take a multi-level perspective exploring the nature of barriers by Martin (2014). To identify the factors affecting variation in policies, Zuiderwijk and Janssen (2014b) developed a context–input–output–performance framework for comparing open data policies.

The second branch was continued by Janssen and Zuiderwijk (2014), who pointed out the need for new business models to generate benefits or value from open data. The authors derived a set of business models from case studies. The next branch of research was initiated by Harrison et al. (2012). The focus of this study was to develop the open government ecosystem in order to systematically clarify the research agenda of open government.

**Phase 3: OGD adoption and use.** The third phase’s research trend is open data adoption, which follows on from earlier mainstream research. This stage was triggered by the work of Susha et al. (2015), which compares the focus, theories and method of assessment benchmarks of OGD to foster its development. This article paved the way for open data use research by pointing out that there is less attention regarding end-user practices and individual adoption. Following this line, Zuiderwijk, Janssen and Dwivedi (2015) and, later, Saxena and Janssen (2017) conducted a user adoption study using Unified Theory of Acceptance and Use of Technology (UTAUT) to examine the factors influencing the acceptance and use of open data technologies or open data. The other line was from a provider-
centric perspective, which focused on the motivations promoting the adoption of open data initiatives by government agencies.

**Phase 4: Implementation and comparison among countries.** OGD initiatives originated from Western countries but recent studies have increasingly focused on developing countries’ OGD initiatives. Saxena (2017) employed open data maturity models to examine the nature and scope of open data in Israel, Georgia and Uzbekistan. Following her previous work, Saxena (2018) compared the development of open data in Japan, the Netherlands and Saudi Arabia to investigate the role of culture in order to show the wider range of countries involved in OGD, resulting in a global scope.

Table 1 shows the shifts to research conducted in other countries and the change of use of theory models by researchers. Most of the researchers who focused on OGD in the first phase mainly came from the US, while in the second phase, European researchers played an important role. After that, researchers from developing countries, such as China and India, joined this domain. Regarding research theories, the initial research was largely empirical and focused on the US and EU practice, whereas in the next phases, researchers started to translate existing theories to the OGD domain. For example, the technology acceptance model (TAM) and its extended model were used to analyse users’ adoption behaviour. In addition, Zuiderwijk and Janssen (2014b) developed the context–input–output model to evaluate open data policy implementation.

**Motivation of evolution**

One of the essential research motivations is to face challenges or solve current problems (Kothari, 2004), which is also the motivation for evolution. Such an analysis can help to identify which issues have not fully matured yet. The results are summarized in Table 2.

**Phase 1: Creating transparency, participation, innovation and economic value.** OGD initiatives were initially launched by the EU and US. The aim of European open data strategies was to develop the information market, enhance collaboration between public sectors and third parties, and drive innovation by disclosng data. In the US, President Obama (2007) promised to ‘create a transparent and connected democracy’ through OGD and to change the sphere of mistrust.

These initiatives have the potential to reach higher levels of transparency and participation, and to facilitate the innovative use of data, but there are many challenges faced by both practitioners and scholars. For instance, there is a need to find out how much and what type of transparency is suitable (Coglianese, 2009). It is also necessary to know what the challenges of creating citizen participation through open data action are (Evans and Campos, 2013) and how to promote innovation through the use of data (Lakomaa and Kallberg, 2013).
Phase 2: Going beyond a technical focus and the national level. The concept of linked open government data was widely spread in the information system research domain. Many discussions on collecting, publishing and using data effectively are based on a technical perspective (Shadbolt et al., 2012; Stadler et al., 2012). However, other factors besides technology play a role in the OGD process, such as legal contexts (Kulk and van Loenen, 2012) and cultural contexts (Nugroho et al., 2015). Hence, some researchers have gone beyond the technical perspective (e.g. Janssen et al., 2012; Zuiderwijk and Janssen, 2014b).

With the OGD movement’s progress, all kinds of public organizations have come under pressure to release their data since most of the data are collected at local levels (Conradie and Choenni, 2014). It is not yet clear how the potential of the OGD initiatives can be realized at the local level (Kassen, 2013); researchers have turned their focus from national-level open data activities, like the Open Government Directive and PSI Directive, towards the local level (Ganapati and Reddick, 2012).

Despite the narratives in the first phase emphasizing the transformative potential of OGD, there is limited evaluation of the actual benefits and impact of OGD initiatives and evidence of their value (Martin, 2014). The efforts of summarizing, analysing and comparing open data business models from different countries, in particular, are very limited (Ahmadi Zeleti et al., 2016). This issue is already and will continue to be well discussed in the OGD research field (e.g. Jetzek et al., 2012; Jetzek et al., 2014).

Phase 3: Usage lags behind and different open data policies. To realize the benefits of opening data, OGD need to be used by various stakeholders, particularly in private organizations. There are very few OGD usage practices by private organizations, and even fewer that are successfully building profitable products or services based solely on OGD (Kaasenbrood et al., 2015). Hence, researchers have investigated the factors influencing the acceptance of open data and primarily developed models for users’ adoption of OGD (Kaasenbrood et al., 2015; Zuiderwijk et al., 2015). Furthermore, some researchers have also analysed open data use activities by hackers (Briscoe and Mulligan, 2014).

As the OGD movement spread, researchers started tackling open data policies and implementation. Although OGD policies in different agencies and countries have similarities, their main objectives and targeted outcomes are often not the same (Cañares and Shekhar, 2016). Limited systematic and structured research has been conducted on the differences in open data policies. A comprehensive overview of existing open data policies could help countries with less developed open data policies to gain insight from the comparison and improve their own policies (Nugroho et al., 2015).

Phase 4: Less focus on open data programmes in developing countries. Academic research on OGD was more concentrated in the West (e.g. Huijboom and Van den Broek, 2011; Ruijer et al., 2017), and fewer investigations took place in developing countries
In this phase, there is an increase of open data initiatives in developing countries; however, the contexts, open data, technical platforms and standards, and government settings vary widely among different countries (Cañares and Shekhar, 2016), which makes it necessary to learn from a comparison of open data initiatives.

**Future research agenda**

Based on the preceding analysis and recent research advancements, a framework for a future research agenda was developed, as shown in Figure 4. It includes investigating the goal, process, results and direction of OGD. An integral approach towards OGD is needed that integrates the isolated efforts into a coherent whole.

**Goal: purpose and benefits of OGD**

As we mentioned in the motivation analysis, there are different objectives for OGD emphasized across countries and by different government levels. For example, in Europe, supporters of OGD often stress its economic value, whereas reasons given by developing countries are often strongly focused on transparency, accountability and citizen participation (Schwegmann, 2012). Prior research also found that the reason driving openness at the subnational level is the national or local legal framework (Cañares and Shekhar, 2016), or political leadership (Hossain et al., 2016), rather than the perceived benefits from open data (Yang and Wu, 2016).

Furthermore, the interpretation of OGD objectives such as transparency and accountability differs. People are likely to give different meanings to what constitutes transparency (Matheus and Janssen, 2019). In some countries, the publication of data is viewed as a form of transparency, whereas other countries consider that there is only transparency if the public are also able to interpret the data (van Zyl, 2014). Therefore, there is a need to clarify concepts and create consensus.

**Process: artificial intelligence for creating smartness in OGD**

In the motivation analysis, we found that the usage of OGD is lagging behind and there is a lack of evidence of creating value from OGD. With the rise of cloud computing and artificial intelligence (AI) technology, governments tend to employ data analytics through AI to create more value from open data (Gao and Janssen, 2020). Data analytic tools are often employed to gain new ideas from OGD. The emerging field of AI provides the potential to create even more value from OGD. This is creating a new research area of AI in open data and research questions on what AI techniques can be used and how value can be created from OGD using AI.

The publication of data needs to be accompanied by an infrastructure that is able to handle the data in a user-friendly way in order to lower the threshold for users (Janssen and Zuiderwijk, 2014). What should such an infrastructure look like?
and how should it be designed? Moreover, the internet of things (IoT) is opening a
large amount of data collected by sensors. How can IoT be employed for open
data? Open data infrastructures might integrate mechanisms to open data in real
time and provide tools for the public to process data.

Result: innovating with OGD

Innovation can be defined as ‘the implementation of a new or significantly
improved product (good or service)’ (OECD and Statistical Office of the
European Communities, 2005: 46). As such, innovating with OGD can be
described as users employing open data to create new applications or new insights
into the current situation. The motivation analysis showed that usage lagged
behind. To address this problem, some scholars focused on the adoption behaviour
of open data users (e.g. Zuiderwijk et al., 2015). Some other researchers studied the
open data platform and innovations after the adoption of open data (Ojo et al.,
2015, 2016).

However, the issues about the open data innovation process or the outcome and
impact of open data innovation have not been fully discussed. It is generally antic-
ipated that OGD will contribute to economic growth and social development.
However, the economic and social impact and success factors are yet to be
reported. Future research is needed to understand the economic and social con-
tributions of open data better. What are successful business models for open data?
How do they create products and social value?

On the other hand, innovating with OGD needs high-quality data. Political
decisions direct or mandate agencies to disclose data but bureaucrats may decide
to release less useful or less valuable data (Peled, 2011; Zuiderwijk and Janssen,
2014b). Future research may also pay attention to how to change bureaucrats’
willingness to open more data by gamification (Kleiman, 2019; Kleiman et al.,
2020). In addition, innovation contests like hackathons could help create new ideas
that can be transformed into applications or services. This does raise further
research questions, such as how to support the later stage of open data innovation
(Juell-Skielse et al., 2014). How can governments meet the needs of users by cre-
ating an open data market while at the same time ensuring privacy?

Theory: an integral approach of OGD

As mentioned earlier, the case-study method was frequently adopted to obtain
empirical knowledge in the initial stage of OGD. After that, theoretical models
from other domains were adopted to answer research questions. For example, the
multi-level perspective was applied to explore the nature of the barriers currently
faced by the OGD agenda (Martin, 2014), and the TAM and its extended model
were used to analyse the adoption behaviour of users (Zuiderwijk et al., 2015). In
addition, innovation diffusion theory (Ojo et al., 2015) and social capital theory
(O’Hara, 2012) are also used to study open data issues. All these researchers focused on a single area, whereas Harrison et al. (2012) advocated an ecosystem view. There is no integral theory or framework to understand the whole process of open data, including the purpose of open data, the innovation process and its outcome or impact. Therefore, we expect more integral approaches towards OGD to rise.

**Direction: sustainable OGD initiatives**

We also need more investigation to ensure the sustainability of open data initiatives. OGD might not bring the value as expected, and governments might decide not to update their data continuously. OGD sustainability means that governmental data continue to be released and used regularly with at least the same or improved quality and quantity. Only when data are relevant, up to date and accurate can they generate more value. Furthermore, sustainability is essential for companies that are creating OGD-based business models or for transparency efforts. If OGD are not available or changes, then their business model might fail. Yet, there is decreasing evidence of sustained scaled-up open data examples, particularly in developing countries (World Bank Group, 2016). What makes OGD initiatives sustainable? Who is already using the data? Who will pay for the release of these data?

**Conclusion**

This study has systematically investigated the evolution of OGD research and proposed future research directions. We analysed the evolutionary trajectory of OGD and the underlying motivation. The themes represent four major phases of evolution. The first phase focused on creating transparency, participation, innovation and economic value, as characterized by the publication of OGD initiatives (e.g. the PSI Directive). In the second phase, researchers realized that the discussion on OGD projects should go beyond the technical and national level, and focus on finding more evidence for creating value from data. In this phase, benefits and barriers, the infomediary business model, and the open data ecosystem approach of OGD initiatives were introduced and studied. The third phase was mainly motivated by the limited usages of OGD. The academic research focus shifted to open data adoption and usage, and the evaluation of open data policies. The fourth stage focused on the implementation of open data initiatives and learning through comparisons of initiatives among countries. The next phase is expected to focus on ensuring the sustainability of OGD initiatives and the use of IoT and AI to create smartness, which should result in the creation of more value from open data. Finally, we derived future OGD research directions, which include re-understanding the purpose and benefits of OGD, creating smartness by
employing AI and data marketplaces, innovating with OGD, developing a more integral approach, and ensuring the sustainability of OGD initiatives.

This article provides some implications for researchers and practitioners. First, the global main path concisely describes the major knowledge flow and helps researchers quickly obtain a clear understanding of OGD research through its basic knowledge framework. Second, the main path analysis method can be applied to conduct comprehensive reviews of other important subfields. Third, the article provides a future research agenda framework that can be used to predict trends of OGD development for both OGD researchers and practitioners in the near future. Future recommendations for OGD practitioners are to re-examine the purpose of the expected benefits of OGD and apply the latest technology to open data to create a smart and open government.

Although our study was as comprehensive as possible, there are still some limitations. First, citation motivation may influence the accuracy of our main path analysis. A theory called ‘remote citation’ argues that an article cites others not because of a close connection with the main subject, but merely because of a connection in a broad sense (Liu et al., 2013). This type of citation may influence the main path analysis result. In the future, multiple main paths could be investigated. Next, even though the WoS is one of the most comprehensive sources of scientific articles, it still cannot guarantee the coverage of all possible OGD publications. Future research could include a broader range of articles, including those written in other languages.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the NSFC project of China (No. 71734002). Furthermore, Yingying Gao was sponsored by the China Scholarship Council. Part of this work is funded by the European Commission within the Erasmus+ programme in the context of the CAP4CITY project (see: https://www.cap4city.eu/home/) under grant agreement No. 598273 EPP 1 2018 1 AT EPPKA2 CBHE JP.

ORCID iDs

Yingying Gao https://orcid.org/0000-0001-9944-9059
Marijn Janssen https://orcid.org/0000-0001-6211-8790
Congcong Zhang https://orcid.org/0000-0002-6047-5513

Supplemental material

Please see online Figure and Table at http://journals.sagepub.com/home/ras
Notes
1. All the figures and tables are available online at: https://journals.sagepub.com/home/ras
2. For more detail about main path analysis, please refer to previous studies (e.g. Liu and Lu, 2012; Liu et al., 2019).

References
Ahmadi Zeleti F, Ojo A and Curry E (2016) Exploring the economic value of open government data. Government Information Quarterly 33(3): 535–551.
Arun R, Suresh V, Veni Madhavan CE, et al. (2010) On finding the natural number of topics with latent Dirichlet allocation: Some observations. In: Zaki MJ, Yu JX, Ravindran B, et al. (eds) Advances in Knowledge Discovery and Data Mining. Berlin, Heidelberg: Springer, pp. 391–402.
Attard J, Orlandi F, Scerri S, et al. (2015) A systematic review of open government data initiatives. Government Information Quarterly 32(4): 399–418.
Barbieri N, Ghisetti C, Gilli M, et al. (2016) A survey of the literature on environmental innovation based on main path analysis. Journal of Economic Surveys 30(3): 596–623.
Batagelj V (2003) Efficient algorithms for citation network analysis. arXiv preprint cs/0309023. Available at: http://arxiv.org/abs/cs/0309023
Batagelj V and Mrvar A (1998) PAJEK: Program for large network analysis. Connections 21(2): 47–57.
Bindu N, Sankar CP and Kumar KS (2019) From conventional governance to e-democracy: Tracing the evolution of e-governance research trends using network analysis tools. Government Information Quarterly 36(3): 385–399.
Blei DM and Jordan MI (2003) Modeling annotated data. In: SIGIR Conference on Research and Development in Information Retrieval. New York, NY: Association for Computing Machinery, pp. 127–134.
Bowen G (2009) Document analysis as a qualitative research method. Qualitative Research Journal 9(2): 27–40.
Briscoe G and Mulligan C (2014) Digital innovation: The hackathon phenomenon. Available at: http://www.creativeworkslondon.org.uk/wp-content/uploads/2013/11/Digital-Innovation-The-Hackathon-Phenomenon1.pdf
Brooks TA (2008) Watch this: LOD linking open data. Information Research 4(13): TB0812. Available at: http://InformationR.net/ir/13-4/TB0812.html
Cañares M and Shekhar S (2016) Open data and sub-national governments: Lessons from developing countries. The Journal of Community Informatics 12(2): 99–119.
Chen C (2004) Searching for intellectual turning points: Progressive knowledge domain visualization. Proceedings of the National Academy of Sciences of the United States of America 101(Suppl. 1): 5303–5310.
Chen Y-B, Liu JS and Lin P (2013) Recent trend in graphene for optoelectronics. Journal of Nanoparticle Research 15(2): 1–14.
Chen K, Zhang Y and Fu X (2019) International research collaboration: An emerging domain of innovation studies? Research Policy 48(1): 149–168.
Coglianese C (2009) The transparency president? The Obama administration and open government. Governance 22(4): 529–544.
Conradie P and Choenni S (2014) On the barriers for local government releasing open data. Government Information Quarterly 31(Suppl. 1): 10–17.
Dahlander L and Gann DM (2010) How open is innovation? Research Policy 39(6): 699–709.

Evans AM and Campos A (2013) Open government initiatives: Challenges of citizen participation. Journal of Policy Analysis and Management 32(1): 172–185.

Ganapati S and Reddick CG (2012) Open e-government in U.S. state governments: Survey evidence from chief information officers. Government Information Quarterly 29(2): 115–122.

Gao Y and Janssen M (2020) Generating value from government data using AI: An exploratory study. In: Viale Pereira G, Janssen M, Lee H, et al. (eds) Electronic Government 2020. Cham: Springer, pp. 319–331.

Harrison TM, Pardo TA and Cook M (2012) Creating open government ecosystems: A research and development agenda. Future Internet 4(4): 900–928.

Horsley JP (2006) Introduction on open government information implementation. Government Information Quarterly 1(23): 5–10.

Hossain MA and Chan C (2015) Open data adoption in Australian government agencies: An exploratory study. In: ACIS 2015 Proceedings 60. Available at: https://aisel.aisnet.org/acis2015/60

Hossain MA, Dwivedi YK and Rana NP (2016) State-of-the-art in open data research: Insights from existing literature and a research agenda. Journal of Organizational Computing and Electronic Commerce 26(1–2): 14–40.

Huijboom N and Van den Broek T (2011) Open data: An international comparison of strategies. European Journal of ePractice 12(1): 4–16.

Hummon NP and Dereian P (1989) Connectivity in a citation network: The development of DNA theory. Social Networks 11(1): 39–63.

Hummon NP, Dereian P and Freeman LC (1990) Analyzing the structure of the centrality-productivity literature created between 1948 and 1979. Knowledge 11(4): 459–480.

Janssen K (2011) The influence of the PSI directive on open government data: An overview of recent developments. Government Information Quarterly 28(4): 446–456.

Janssen M and Zuiderwijk A (2014) Infomediary business models for connecting open data providers and users. Social Science Computer Review 32(5): 694–711.

Janssen M, Charalabidis Y and Zuiderwijk A (2012) Benefits, adoption barriers and myths of open data and open government. Information Systems Management 29(4): 258–268.

Jetzek T, Avital M and Bjørn-Andersen N (2012) The value of open government data: A strategic analysis framework. Paper presented at: 2012 Pre-ICIS Workshop, Orlando, Florida, United States, 15 December 2012.

Jetzek T, Avital M and Bjørn-Andersen N (2014) Data-driven innovation through open government data. Journal of Theoretical and Applied Electronic Commerce Research 9(2): 100–120.

Johnson P and Robinson P (2014) Civic hackathons: Innovation, procurement, or civic engagement? Review of Policy Research 31(4): 349–357.

Juell-Skielse G, Hjalmarsson A, Johannesson P, et al. (2014) Is the public motivated to engage in open data innovation? In: Janssen M, Scholl HJ, Wimmer MA, et al. (eds) Electronic Government 2014. Berlin, Heidelberg: Springer, pp. 277–288.

Kaasenbrood M, Zuiderwijk A, Janssen M, et al. (2015) Exploring the factors influencing the adoption of open government data by private organisations. International Journal of Public Administration in the Digital Age 2(2): 75–92.
Kalampokis E, Tambouris E and Tarabanis K (2011) Open government data: A stage model. In: Janssen M, Scholl HJ, Wimmer MA, et al. (eds) Electronic Government 2011. Berlin, Heidelberg: Springer, pp. 235–246.

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.

Kleiman F (2019) Engaging governments in open data policies through gaming. In: Dhaou SB, Carter L and Gregory M (eds), Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance. New York, NY: Association for Computing Machinery, pp. 461–464.

Kleiman F, Janssen M, Meijer S, et al. (2020) Changing civil servants’ behaviour concerning the opening of governmental data: Evaluating the effect of a game by comparing civil servants’ intentions before and after a game intervention. International Review of Administrative Sciences, epub ahead publication, 30 September, DOI: 10.117/00208523209/62211.

Kothari CR (2004) Research Methodology: Methods and Technique. New Delhi: New Age International.

Kulk S and van Loenen B (2012) Brave new open data world? International Journal of Spatial Data Infrastructures Research 7: 196–206.

Lakomaa E and Kallberg J (2013) Open data as a foundation for innovation: The enabling effect of free public sector information for entrepreneurs. IEEE Access 1: 558–563.

Liang H, Wang JJ, Xue Y, et al. (2016) IT outsourcing research from 1992 to 2013: A literature review based on main path analysis. Information and Management 53(2): 227–251.

Liu JS and Lu LYY (2012) An integrated approach for main path analysis: Development of the Hirsch Index as an example. Journal of the American Society for Information Science and Technology 63(3): 528–542.

Liu JS, Lu LYY, Lu WM, et al. (2013) Data envelopment analysis 1978–2010: A citation-based literature survey. Omega (United Kingdom) 41(1): 3–15.

Liu JS, Lu LYY and Ho MHC (2019) A few notes on main path analysis. Scientometrics 119(1): 379–391.

Lu LYY and Liu JS (2014) The knowledge diffusion paths of corporate social responsibility – from 1970 to 2011. Corporate Social Responsibility and Environmental Management 21(2): 113–128.

Martin C (2014) Barriers to the open government data agenda: Taking a multi-level perspective. Policy & Internet 6(3): 217–240.

Maseh E and Katuu S (2017) The Kenyan judiciary’s open government initiative: Prospects and challenges. Journal of Science and Technology Policy Management 8(1): 76–94.

Matheus R and Janssen M (2019) A systematic literature study to unravel transparency enabled by open government data: The window theory. Public Performance and Management Review 43(3): 1–32.

McDermott P (2010) Building open government. Government Information Quarterly 27(4): 401–413.

Mergel I, Kleibrink A and Sörvik J (2018) Open data outcomes: U.S. cities between product and process innovation. Government Information Quarterly 35(4): 622–632.

Nugroho RP, Zuiderwijk A, Janssen M, et al. (2015) A comparison of national open data policies: Lessons learned. Transforming Government: People, Process and Policy 9(3): 286–308.
Obama B (2007) Connecting and empowering all Americans through technology and innovation. Available at: https://www.wired.com/images_blogs/threatlevel/2009/04/obama
techplan.pdf

OECD (Organisation for Economic Co-operation and Development) and Statistical Office of the European Communities (2005) Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data. Paris and Luxembourg: OECD/Euro-stat.

O’Hara K (2012) Transparency, open data and trust in government: Shaping the infosphere. In: Proceedings of the 4th Annual ACM Web Science Conference, WebSci’12. New York, NY: Association for Computing Machinery, pp. 223–232.

Ojo A, Curry E and Zeleti FA (2015) A tale of open data innovations in five smart cities. In: Proceedings of the Annual Hawaii International Conference on System Sciences. 26 March, IEEE, pp. 2326–2335.

Ojo A, Porwol L, Waqar M, et al. (2016) Realizing the innovation potentials from open data: Stakeholders’ perspectives on the desired affordances of open data environment. In: Afsarmanesh H, Camarinha-Matos L and Lucas Soares A (eds) IFIP Advances in Information and Communication Technology. Cham: Springer, pp. 48–59.

Peled A (2011) When transparency and collaboration collide: The USA Open Data Program. Journal of the American Society for Information Science and Technology 62(11): 2085–2094.

Ruijer E, Grimmelikhuijsen S and Meijer A (2017) Open data for democracy: Developing a theoretical framework for open data use. Government Information Quarterly 34(1): 45–52.

Safarov I, Meijer A and Grimmelikhuijsen S (2017) Utilization of open government data: A systematic literature review of types, conditions, effects and users. Information Polity 22(1): 1–24.

Saxena S (2017) Open data in Israel, Georgia and Uzbekistan: Nature and scope. Information and Learning Science 118(7–8): 406–419.

Saxena S (2018) National open data frames across Japan, the Netherlands and Saudi Arabia: Role of culture. Foresight 20(1): 123–134.

Saxena S and Janssen M (2017) Examining open government data (OGD) usage in India through UTAUT framework. Foresight 19(4): 421–436.

Sayogo DS, Pardo TA and Cook M (2014) A framework for benchmarking open government data efforts. In: Proceedings of the Annual Hawaii International Conference on System Sciences, IEEE, pp. 1896–1905.

Schwegmann C (2012) Open data in developing countries. European Public Sector Information Platform topic report, Germany. Available at: http://www.data.gov/open
datasites (accessed 21 January 2020).

Shadbolt N, O’Hara K, Berners-Lee T, et al. (2012) Linked open government data: Lessons from Data.gov.uk. IEEE Intelligent Systems 27(3): 16–24.

Stadler C, Lehmann J, Höffner K, et al. (2012) LinkedGeoData: A core for a web of spatial open data. Semantic Web 3(4): 333–354.

Susha I, Zuiderwijk A, Janssen M, et al. (2015) Benchmarks for evaluating the progress of open data adoption. Social Science Computer Review 33(5): 613–630.

Tang J, Meng Z, Nguyen X, et al. (2014) Understanding the limiting factors of topic modeling via posterior contraction analysis. In: Xing EP and Jebara T (eds) International Conference on Machine Learning. JMLR.org, pp. 190–198.

Van Zyl A (2014) How civil society organizations close the gap between transparency and accountability. Governance 27(2): 347–356.
World Bank Group (2016) World development report: Digital dividends. World Bank Publications. Available at: https://www.worldbank.org/en/publication/wdr2016
Yang TM and Wu YJ (2016) Examining the socio-technical determinants influencing government agencies’ open data publication: A study in Taiwan. Government Information Quarterly 33(3): 378–392.
Zhao Y and Fan B (2018) Exploring open government data capacity of government agency: Based on the resource-based theory. Government Information Quarterly 35(1): 1–12.
Zuiderwijk A and Janssen M (2014a) Barriers and development directions for the publication and usage of open data: A socio-technical view. In: Gascó-Hernández M (ed.) Open Government. New York, NY: Springer, pp. 115–135.
Zuiderwijk A and Janssen M (2014b) Open data policies, their implementation and impact: A framework for comparison. Government Information Quarterly 31(1): 17–29.
Zuiderwijk A, Janssen M and Dwivedi YK (2015) Acceptance and use predictors of open data technologies: Drawing upon the unified theory of acceptance and use of technology. Government Information Quarterly 32(4): 429–440.
Zuiderwijk A, Shinde R and Janssen M (2019) Investigating the attainment of open government data objectives: Is there a mismatch between objectives and results? International Review of Administrative Sciences 85(4): 645–672.

Yingying Gao is a PhD student at Huazhong University and Technology, China, and a guest PhD student at Delft University of Technology, The Netherlands. Her research interest is in the leading technologies applied to governance.

Marijn Janssen is a Full Professor in ICT & Governance in the Information and Communication Technology Research Group of the Technology, Policy and Management Faculty, Delft University of Technology, The Netherlands. He was ranked as one of the leading e-government researchers and has published over 500 refereed publications. For more information, see: www.tbm.tudelft.nl/mariijn

Congcong Zhang is a PhD student in the College of Public Administration, Huazhong University and Technology, China. Her main ongoing research is about data sharing.