

Towards A Taxonomy of Emerging Topics in Open Government Data: A Bibliometric Mapping Approach

Ahmad Nadzri Mohamad*
Faculty of Information Management
Universiti Teknologi MARA
[nazri.mohamad@gmail.com](mailto:nadzri.mohamad@gmail.com)

Allan Sylvester
School of Information Management
Victoria University of Wellington
Allan.Sylvester@vuw.ac.nz

Jennifer Campbell-Meier
School of Information Management
Victoria University of Wellington
jennifer.campbell-meier@vuw.ac.nz

Abstract

The purpose of this paper is to capture the emerging research topics in Open Government Data (OGD) through a bibliometric mapping approach. Previous OGD research has covered the evolution of the discipline with the application of bibliometric mapping tools. However, none of these studies have extended the bibliometric mapping approach for taxonomy building. Realizing this potential, we used a bibliometric tool to perform keyword analysis as a foundation for taxonomy construction. A set of keyword clusters was constructed, and qualitative analysis software was used for taxonomy creation. Emerging topics were identified in a taxonomy form. This study contributes towards the development of an OGD taxonomy. This study contributes to the procedural realignment of a past study by incorporating taxonomy building elements for taxonomy creation. These contributions are significant because there is insufficient taxonomy research in the OGD discipline. The taxonomy building procedures extended in this study are applicable to other fields.

1. Introduction

Open Government Data (OGD) is a philosophy with a set of policies to encourage the publication of government data for transparency, accountability, and value creation [1]. The rise of open data technologies and increasing public demands on government data inspire OGD initiatives worldwide. Typically, public sector institutions publish a tremendous amount of data in various areas such as transportation, health, and education. The utilization of government data stimulates economic, social, and governance benefits. Accessing government data allows the creation of new products and services [2]. Releasing high-valued data increases government-citizen engagement while promoting problem-solving through co-creation and innovation activities [3]. More than seventy countries participate in the Open Government Partnership (OGP).

Other international movements that advocate OGD initiatives are The Asian Open Data Partnership and European Union (EU).

Bibliometrics is the use of mathematical and statistical analysis to explore the development of a discipline in a period of time[4]. Recent developments in data visualization techniques to visualize bibliographic data have sparked interest in bibliometric research. Earlier bibliometric studies used fewer graphical representations in their findings [5]. In recent years, bibliometric mapping studies use visual maps and networks to enhance readers' understanding. Bibliometric mapping research can generate visual maps for research related to citation analysis [6], keyword co-occurrences [7], performance, and research trends [8]. Bibliometric mapping studies can be performed with the utilization of one or more tools such as Science Mapping Analysis Tool (SciMAT), CitNetExplorer, and CiteSpace.

There are two types of bibliometric studies in previous OGD studies. First, bibliometric studies that explore the development of OGD as a primary discipline [9], [10]. Second, bibliometric studies investigate the advancement of OGD with other fields such as freedom of information [11] and entrepreneurship [12]. Bibliometric mapping in these studies generates a cluster of topical maps from utilizing bibliometric mapping tools. However, these bibliometric maps are not explored deeper to construct a topical taxonomy. Bibliometric studies in OGD are based on the extracted bibliographic data from scholarly publications. Bibliographic data such as abstract, keywords, and author information is loaded into a bibliometric tool to facilitate a macro-level analysis. Hence, a further examination is required to look at the articles more closely. This can be achieved by reading the article abstracts or full article contents for context clarification.

*Ahmad Nadzri Mohamad is currently a PhD student at Victoria University of Wellington

Hence, we construct a topical taxonomy by integrating a bibliometric mapping tool (VOSviewer) with qualitative analysis software (NVivo) to address the following research question:

RQ: How can a taxonomy be developed to capture the emerging topics in OGD through a bibliometric mapping approach?

The structure of this paper is as follows. Section two presents the method for taxonomy construction. Section three explains the categories, topics, and subtopics of the taxonomy. Section four is the discussion, followed by research contributions, limitations, and conclusion.

2. Method

Three stages of the research approach were 1) data collection and synthetization, 2) cluster visualization and exploration, 3) cluster refinement and taxonomy development, illustrated in Table 1. These stages were adapted from the procedures in exploring the large body of literature with a bibliometric mapping technique [13].

2.1. Data collection and synthetization –Scopus is one of the major indexing and citation databases with around 78 million records from 5000 publishers worldwide [14]. This database provides a substantial amount of publication data for bibliometric studies [15]. Three types of bibliographic data extracted from Scopus were citation information (*author, document, title, year, source, citation count, document type*), bibliographical information (*affiliation, publisher*), and abstract/keywords information (*abstract, author, and index keywords*). Our search coverage was limited to five years, from 2016 to 2020. More coverage of years would take a longer time for analysis since we looked at the abstract level of the articles for taxonomy development. Non-English documents were removed as the translation would be time-consuming. The search term performed was “Open Government Data”, with two document types selected were journal and conference article. Then, bibliographical data and the full articles were downloaded. Non-retrievable articles were discarded. A total of 524 documents were retrieved.

2.2. Cluster visualization and exploration – This step aims to create a visual map to explore concepts within the literature. Visualization of Similarities (VOSviewer) is a bibliometric mapping tool. This software is freely available designed to visualize bibliometric data networks. The software allows interactive visualization of various data items such as journals, organization, geographic locations, and keywords. In VOSviewer, relations of these bibliographic data are visualized according to the citation, bibliographic coupling, co-citation, or co-authorship network. Unlike other computer software, VOSviewer functions are centralized

on producing an interactive and large amount of data in graphical forms [5].

Table 1. Research approach for taxonomy building

Scopus	VOSviewer		NVivo
Data collection & synthetization	Cluster visualization & exploration		Cluster refinement & taxonomy development
	Visualize	Explore	Organize, code, refine
Familiarization of Scopus database	Select corpus file in VOSviewer	Familiarization of keyword clusters	Analyze article abstracts based on VOSviewer clusters
Specify search criteria	Specify type of analysis	Explore cluster items	Arrangement of codes & nodes
Download of bibliographic data	Set threshold for minimum occurrences of a term	Use zooming & scrolling functions	Integration of NVivo analysis with VOSviewer clusters
Clean text file (delete unnecessary bibliographic data fields)	Choose number of keywords	Explore alternative map by changing the visualization settings	Propose a taxonomy
Download full articles (optional)	Visualize		Discussion with co-authors for refinement

Bibliographic data retrieved from Scopus were loaded into VOSViewer. Graphical maps in VOSViewer were created based on a co-occurrence matrix. This co-occurrence matrix used the uploaded bibliographical data as inputs. A similarity matrix was then constructed based on the keywords co-occurrences. The function of this similarity index is to group similar keywords retrieved from the uploaded bibliographic data. A visual map was created automatically through the bibliometric tool. In general, VOSViewer uses similarity measures to generate nodes and a group of clusters. These clusters are then coded in a specific color scheme. Relatedness of items is closer in the visual map when keywords in the publications are identical. Some published work explains the logic behind graphical maps in VOSViewer [5], [16]. A minimum threshold of five keyword occurrences was selected. Author keywords were used as the unit analysis.

VOSviewer generated a visual map with fifty-four keyword items. Several keyword items are displayed in Diagram 1. These items were organized automatically by VOSviewer into ten clusters, known as cluster one to ten. Each cluster has associated keywords grouped based on the VOSViewer algorithm explained in the previous section. These keywords were cross-checked with the bibliographic data file in CSV format downloaded from Scopus. Visualization features such as zooming in and scrolling function were applied to familiarize with the clusters generated by the bibliometric tool.

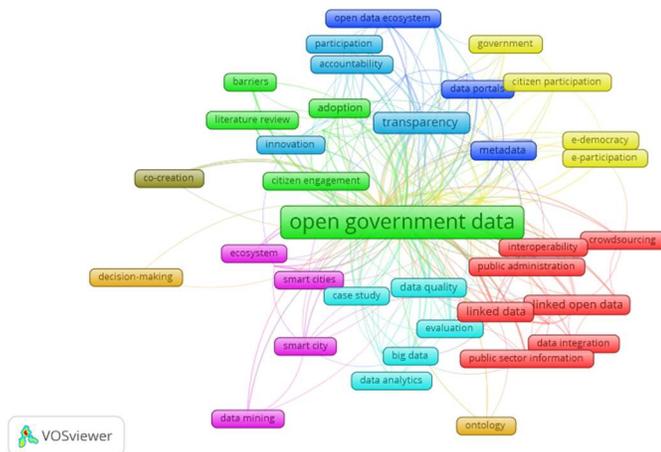


Diagram 1. Keyword co-occurrences in VOSviewer

2.3. Cluster refinement and taxonomy development –

Ten keyword clusters were generated by VOSviewer in a graphical map. These clusters became the foundation for exploration to form a taxonomy. However, such a graphical map lacks the contexts of the articles as it merely captures author keywords. Therefore, we analyzed the abstracts of those publications. Reading the publication abstracts was sufficient for context clarification. In some instances, articles were read to give some degree of understanding. NVivo 12 Plus software was applied to look at the article abstracts. Codes and nodes were created and refined iteratively throughout the coding process. We integrate the clusters generated by VOSviewer together with the analysis outcome from NVivo. A discussion was carried out with co-authors to rectify any inaccurate taxonomy items. Based on this exercise, a taxonomy was formed and proposed.

3. Findings

We found four main categories that capture the emerging topics and subtopics of OGD. These categories are a) OGD implementation and management, b) OGD architecture, c) OGD users and utilization, and d) OGD benefits. Each of these categories is aligned with several main topics and subtopics shown in Table 2. Seven emerging topics and thirty-four subtopics were identified. In the following section, each category will be discussed to demonstrate the emerging topics in the last five years.

3.1. OGD Implementation and Management

The nature of an OGD project involves government personnel in an extensive network of public agencies. In a government department, existing staff can be assigned as an OGD representative. These representatives are often called data stewards. Open data committees can be tasked to expedite the data disclosure process. The organizational hierarchy that connects public agencies with an OGD initiative varies between countries. Nevertheless, data stewards or open data committees are responsible for selecting and disclosing data according to local policies.

Meetings for data disclosure are organized at various organizational levels. In a broader context, there is a national central agency that manages the overall OGD implementation. Negotiations on which and what kind of data suitable for disclosure are ongoing dilemmas for sustainable OGD implementation. This multifaceted characteristic of OGD implementation and management has become an interest for many studies. We found three subtopics under this category: adoption factors and barriers in the public sector, assessment of OGD portals, and OGD ecosystems.

3.1.1. Adoption factors and barriers in the public sector

Factors associated with the decision to support OGD initiatives have become a subject of ongoing inquiry. This is because OGD implementation requires the collaboration and commitment of government personnel across public institutions. However, there is a variability of cooperation and support of government personnel towards OGD adoption [17]. Therefore, scholars are interested in exploring the reasons why some public personnel are more forthcoming than others. Negative factors that hinder OGD adoption are also examined. These studies analyze the adoption factors at various levels. Some studies are conducted at a country level in assessing the factors impacting OGD projects [18]. Other studies concerned the adoption factors at the organizational level [19]. Adoption factors influencing local governments are also performed [20]. The factors that influence OGD adoption is varied from technological, organizational, and environmental reasons. Other aspects that play a role are the infrastructural and cultural role within an organization. Research models to explain these OGD adoptions are illustrated using theories such as Technology-Organization-Environmental (TOE).

Barriers that hinder OGD acceptance in the public sector are also presented. Our initial observation shows that literature on the barriers faced in the public sector is broad. Barriers in the public sector revolve around the process of publishing government data. Barriers can be internal and external barriers to any public institution. For example, internal barriers for a government department relate to the insufficient data-related skills [21] and behavioral resistance [22] of its employees. External barriers faced are external pressure from higher authority and public pressure for data disclosure [23]. Scholars recommend strategies in overcoming barriers, such as policy enhancement [24].

3.1.2. Assessment of OGD portals – One of the most popular aspects of the evaluation of OGD portals is data quality. Strategies are proposed to support the process of releasing quality data. These strategies include using an open-source application [25] and developing a publishing tool for data cleaning processes [26]. In one study, a critical assessment of accessible government data was performed. Suggestions were made to enhance the selection and publication of high-valued data [27]. Another popular interest is the development of frameworks or models to assess OGD portals. These frameworks/models are designed with quality indicators such as accuracy, accessibility, and completeness [28]. User-oriented indicators are used to propose an evaluation framework [29]. Real-time and automated assessment methods are discussed [30]. These frameworks are proposed as a tool to assess the overall quality of OGD portals.

Scholars identified user requirements of OGD platforms [31]. Strategic directions for future improvement are also considered [32]. Other assessments of OGD portals are infrastructure readiness [33], functionalities [34], objectives [35], trust and OGD usage [36]. Evaluating research under this subtopic suggests the idea that data quality is still a primary concern. The development of assessment models reflects the interest and needs to monitor OGD implementations for quality assurances.

3.1.3. OGD Ecosystem – Scholars explore the building blocks of an ecosystem. In a systematic review, five components of an OGD ecosystem are identified [37]. These components are political and legal framework, actors, technological infrastructure, data and standards, and tools to support data interoperability. Scholars also explore how these individual components can collectively communicate well to support an ecosystem. One example of such a concern is in multiple-case research that explores two OGD ecosystems [38]. This research explores how public agencies can facilitate collaboration with external parties. One of the viable mechanisms is to structure formally a collaborative governance for public-private partnerships. Other supportive components such as infrastructure and legal framework enhance the collaboration further. OGD initiatives require continuous cooperation between public agencies and data communities. The importance of having a supportive ecosystem is to ensure that an OGD initiative can generate benefits while maintaining its sustainability over time.

Table 2. A Proposed taxonomy for emerging topics in OGD

Category	Topics	Sub-topics							
OGD Implementation & Management	Adoption factors & barriers in the public sector	Adoption factors & barriers at country level		Adoption factors & barriers at organizational level		Adoption factors & barriers at local governments		Adoption factors & model creations	Strategies to mitigate barriers
	Assessment of OGD portals	Cross-country assessment	Data quality	Readiness of OGD portals	Functionality	Objectives & attainments	Frameworks	Usage	Trust
	OGD ecosystem	Ecosystem across countries			Ecosystem components & models				
OGD Architecture	Metadata	Ontology	Linked open data	Semantic web	Data mining & visualization	Big data analytics	Data integration & interoperability		
OGD Users & Utilization	Usability & barriers	Citizens	NGOs & businesses	Academic community	Innovators				
	Motivation & awareness								
OGD Benefits	Economic benefits	Governance benefits	Social benefits	Transparency & accountability	Value creation	Co-creation	Citizen participation	Innovation	

3.2. OGD Architecture

There are numerous subtopics under OGD architecture. The main interest is to enhance the discoverability and management of large data sources. Scholars explore how OGD platforms can be enhanced with the application of emerging technologies. They explain their experiences and the relevance of their work to advance the quality of OGD platforms. There are nine technological-related terms under this category. These terms are grouped into metadata, ontology, linked open data and semantic web, data mining and visualization, big data analytics, data integration, and interoperability. These terms may intertwine with each other. We will not explain all technical terms in detail. The approach is to explain briefly the fundamental discussions that primarily appear in the literature.

3.2.1. Metadata – Data discoverability depends partly on the metadata assigned in each data item. Accurate metadata enhances the success probability in data searching. That is why metadata management is critical for OGD portals. Studies in this subtopic deal primarily with metadata management and accuracy [39]. Metadata issues [40] and the techniques to improve metadata accuracy are suggested [41]. Metadata schemes are also developed.

3.2.2. Ontology – Ontological techniques are presented to enhance data discoverability. Scholars demonstrated how the ontological approach could be embedded with other technologies to improve searching capabilities [42]. For instance, [43] integrates semantic and machine learning to devise ontology-based semantic search for an OGD platform. Ontological developments are demonstrated to support businesses, public procurement, legal, and shipping operations.

3.2.3. Linked Open Data (LOD) & Semantic Web - Research under this subtopic uncover the recognition of LOD and semantic web applications for OGD architecture. LOD is a design principle that integrates data sources to enhance data accessibility. When browsing for a data source, other relevant data would appear. Hence, the application of LOD would increase the usability of data sources combined. LOD as a term appears in a wide range of studies. Research on LOD intersects with the application of semantic web technologies, semantic models, and semantic vocabulary for data discoverability. Strategies to support linked open data are investigated [44]. Experiences on LOD practices in some countries are shared [45]. Research on the semantic web involves integrating machine learning and cloud computing for data recommendations on OGD platforms [46].

3.2.4. Data mining and visualization – Data mining is a process of analyzing and making sense of data. Data-mining techniques support organizational decision-making and problem-solving. Data visualization covers the graphical representation of a large amount of data and information. Visualization helps to understand data patterns, outliers, and other important information through visual objects such as time-series graphs. Scholars have shown some techniques in data mining and visualization. One example is from [47] to visualize taxpaying information for taxpayers and government officials. Another study [48] extracted local government data to predict vehicle accidents. The effectiveness of data visualization is examined to improve data interpretability [49].

3.2.5. Big data analytics – Big data analytics in the public sector analyzes massive data for useful information and knowledge. This subtopic discovers the incorporation of big data analytics and OGD to enhance public sector services [50]. Scholars identified the concepts, potentials, and challenges of such integration. One particular interest under this subtopic is the use of analytics in public procurement and contracts for transparency.

3.2.6. Data integration and interoperability – This subtopic concerns the application of various technologies on OGD platforms. Data integration from multiple data sources is discussed within those interoperability concerns. Scholars explored the interoperability of emerging technologies with OGD, such as crowdsourcing and the semantic web. Interoperability success is significant to ascertain technological applicability in OGD implementations. Interoperability frameworks are established to explain context-specific applications [51]. In addition, scholars investigated the significance of integrating OGD with other data sources to achieve specific goals [52]. Alternative data sources can be obtained from public and private data repositories. Open data communities can also provide data sources. As a result, data extracted from a variety of sources lead to the development of data-enriched products.

3.3. OGD users and utilization

Scholars are interested to understand how government data are being used. They are also interested to see the impacts of using government data. We found three subtopics under this category: usability and barriers in various user groups, current status on OGD usage, user motivation, and awareness.

3.3.1. Usability and barriers of OGD users – There is a heterogeneous group of OGD users such as citizens, NGOs, innovators, students, academicians, and businesses. The primary attention in this subtopic is to investigate various user groups and their OGD experiences. In specific, user characteristics [53], usage, and barriers [54] are explored. Another ongoing interest sits on the access of government data and the usability of these data to different user groups [55]. The interplay of user factors such as readiness and data skills are highly relevant to ensure that data can be used effectively. To this end, user understanding is essential to provide critical inputs to OGD implementations. As a result, data stewards can prioritize the disclosure of data valued most by users. Such an understanding may lead to better data usage. However, understanding users is complex as OGD initiatives are in various stages worldwide. Differences exist within OGD communities, such as data sharing culture in the public sector. The degree of data literacy is varied amongst data users. Users also have different motivations for accessing government data. To a certain extent, there is a continuous need to examine the wide range of user communities interacting with OGD initiatives. The complexity of data usage amongst varied data users across OGD initiatives might be one of the reasons why this subtopic is highly anticipated in the last five years.

3.3.2. Motivation and awareness – Research in this subtopic contains the rationales of why certain users engage OGD in their activities. We found limited studies that explain user motivations in making use of government data. One case that falls under this subtopic is citizens' intrinsic and extrinsic motivations to engage with OGD [56]. Another case [57] discovers citizen motivations for participating in an open data hackathon. Contrary, there is a strong presence of research that examines the motivations of disclosing data from the public sector [58].

3.4. OGD benefits

OGD implementations are known for economic, social, and governance benefits. Open data can generate data-enriched products and services. Releasing government data leads to job creation and economic opportunity. Data developers and start-up communities can build apps and web-based services based on open data [59]. Governments engage citizens more efficiently with data disclosure for public use. By doing so, governments are encouraging better government-citizen interaction for transparency and democratic values. Other OGD benefits are value-creation, co-creation, and innovation.

Researchers focus on how OGD benefits can be achieved. Digital products from the use of OGD are invented, and benefits or values are reported. For instance, [60] developed a mobile platform for linked legislative data to engage citizens by combining open data and open services. This mobile monitoring platform generates accountability and transparency value to the citizens. However, achieving OGD benefits is not an easy task. One of the key aspects is the cooperation between stakeholders in an ecosystem environment. To nurture innovation as a benefit requires the interdependency of actors in an OGD ecosystem. Data publishers need to provide valuable data for innovation purposes. Innovators from the public and private sectors need to work together to create innovative products. Data hackathons, for example, connect innovators with data providers [61]. A similar situation happens to the public sector in generating co-creation products for problem-solving situations. In such a situation, citizens are collaborators rather than the customers that benefit from data products [62].

4. Discussion

Is the whole process of developing a taxonomy easier through bibliometric mapping? We need to look at some methods in developing a taxonomy to answer this question. First, taxonomy can be built through the use of algorithms [63]. Tools are developed to construct a taxonomy [64]. A critical human interpretation is still required to develop a comprehensive taxonomy. However, this strategy can be difficult for non-technical researchers. Second, researchers can build a taxonomy by manually classify the large source of literature through several procedures [65]. Manual classification of the literature may take a lot of time. Third, a taxonomy can be built based on systematic reviews [66]. In general, a hybrid approach in using some tools with human interpretation is still largely practiced in any of the abovementioned methods.

We acknowledge the diverse approaches to taxonomy building. This research exemplified the application of a bibliometric tool as a basis for taxonomy building based on keyword analysis. Qualitative analysis software was used to rearrange the group of clusters suggested by the bibliometric tool. Codes and nodes were arranged iteratively for concept building. Such a technique helps to shorten the time in developing a taxonomy from a large corpus of data. Still, human interpretation is required. Developing a taxonomy by using a bibliometric mapping tool is not easier than other methods. Nevertheless, using two software allows us to categorize taxonomy items, and concepts more efficiently over a long period of time. Furthermore, the visualization features provided by the bibliometric tool enable familiarization of the literature.

A taxonomy is proposed with four main categories. These categories consist of seven emerging topics and thirty-four subtopics. The first category is the implementation and management of OGD initiatives. The concept of OGD is not new. However, the complexity of data disclosure in the public sector is still an ongoing interest. Emerging topics under this category are adoption and barriers of OGD implementation, assessment of OGD portals, and the emphasis on an ecosystem. The second category is OGD architecture. An importance has been given to explore applicable technologies to enhance OGD platforms. Some of these technologies are open-linked data and the semantic web. The third category is OGD users and utilization. One of the topics under this category is the usability and barriers in using OGD sources. Scholars are not only interested to study the complexity of releasing datasets from the public sector. They are also interested to comprehend the usability of government datasets to different user groups. The usability of data has a connection to the fourth category in the taxonomy which is OGD benefits. An emerging interest in this category is to grasp the benefits and values of utilizing government data. General benefits are recognized such as economic and social benefits. We also discovered OGD studies that explore specific benefits such as innovation and co-creation from the use of government data.

5. Contributions

We need to look at the existing bibliometric studies in the same area to rationalize research contribution. We found four bibliometric mapping studies in OGD that explore the progress of the discipline [9]–[12]. None of these studies extended the bibliometric mapping techniques to develop a taxonomy. We recognized [13] suggestion to do a follow-up analysis using qualitative analysis software for literature discovery. This suggestion allows us to look deeper at the article abstracts for taxonomy construction. As such, we have two research contributions.

First, this research contributes towards the development of a taxonomy based on a bibliometric mapping technique. Such contribution is significant because there is limited study for taxonomy development in OGD. The only taxonomy work specifically on OGD was conducted by [67]. Other taxonomy studies are related to open data [68]. Taxonomy creation is significant for researchers to understand the interrelationship among concepts of a particular domain [69]. Further, researchers can hypothesize, challenge, and investigate the divergence of domain concepts through taxonomy creation [70]. For this reason, conceptual knowledge in taxonomy building is a foundation for comprehending a domain for theory-building and theorizing [71].

Second, this research extends the procedures introduced by [13] by incorporating the taxonomy building elements. In that study, the focus is to identify interesting research questions for theorizing through bibliometric mapping. We applied the same guideline for a different purpose which is to develop a taxonomy. Hence, we realigned the guideline to illustrate the potential of developing a taxonomy based on a bibliometric mapping approach in Table 1. This involved a minor addition to the overall guideline structure. Yet, this procedural contribution is significant to students and researchers in other fields to apply a similar technique for taxonomy creation.

6. Limitations

We retrieved bibliographical data sources from Scopus in five years. Future studies may include bibliographical data from several electronic library databases. Integrating bibliographical data from multiple sources may take some time. First, data providers have different data standards. Second, some library databases do not permit the automatic extraction of bibliographic data. Third, bibliometric mapping tools have different data requirements. Bibliographical data from various sources can be compiled. However, the compiled bibliographical data do not necessarily compatible with all bibliometric mapping tools. Conversion of data format helps to bridge the differences in data standards and requirements. Another limitation is that we used one bibliometric tool due to familiarization. Ease of use of the chosen bibliometric tool is another reason. Future studies may consider the integration of several bibliometric mapping tools. Familiarizing with bibliometric mapping tools is necessary to understand their features and abilities to answer certain types of research questions. Expert reviews can be obtained to add more inputs for taxonomy building.

7. Conclusion

We developed an OGD taxonomy based on a bibliometric mapping technique. Four categories of emerging research interest are aligned with associated topics and subtopics. We acknowledged the potential of taxonomy development by using a bibliometric mapping approach. The keyword clusters obtained from a bibliometric mapping tool were enhanced by using qualitative analysis software. The use of each software requires some technical knowledge, but it is still feasible for a non-technical person. Research contributions are also proposed. Limitations are informed, and possible solutions are provided for consideration in future studies. We believed that this study would enrich the discourse on constructing a taxonomy for a wider audience.

8. References

- [1] OECD, "Open Government Data," *Organisation for Economic Co-operation and Development*, 2018. <http://www.oecd.org/gov/digital-government/open-government-data.htm> (accessed Aug. 26, 2019).
- [2] M. Kassen, "Open data and its peers: Understanding promising harbingers from Nordic Europe," *Aslib J. Inf. Manag.*, no. 5, 2020.
- [3] K. McBride, G. Aavik, M. Toots, T. Kalvet, and R. Krimmer, "How does open government data driven co-creation occur? Six factors and a 'perfect storm': Insights from Chicago's food inspection forecasting model," *Gov. Inf. Q.*, vol. 36, no. 1, pp. 88–97, 2019.
- [4] J. M. Reitz, "Online dictionary for library and information science." ABC-Clio, California, 2020, [Online]. Available: https://products.abc-clio.com/ODLIS/odlis_b.aspx.
- [5] N. J. van Eck and L. Waltman, "Software survey: VOSviewer, a computer program for bibliometric mapping," *Scientometrics*, vol. 84, no. 2, pp. 523–538, 2010.
- [6] F. N. Silva, D. R. Amancio, M. Bardosova, L. da F. Costa, and O. N. Oliveira, "Using network science and text analytics to produce surveys in a scientific topic," *J. Informetr.*, vol. 10, no. 2, pp. 487–502, 2016.
- [7] X. Chen, J. Chen, D. Wu, Y. Xie, and J. Li, "Mapping the Research Trends by Co-word Analysis Based on Keywords from Funded Project," *Procedia Comput. Sci.*, vol. 91, no. Itqm, pp. 547–555, 2016.
- [8] C. Olmeda-Gómez and F. de Moya-Anegón, "Publishing Trends in Library and Information Sciences Across European Countries and Institutions," *J. Acad. Librariansh.*, vol. 42, no. 1, pp. 27–37, 2016.
- [9] Corrales-Garay, M. Ortiz-de-Urbina-Criado, and E. M. Mora-Valentín, "Knowledge areas, themes and future research on open data: A co-word analysis," *Gov. Inf. Q.*, vol. 36, no. 1, pp. 77–87, 2019.
- [10] H. Lv, "Performance assessment and major trends in open government data research based on Web of Science data," *Data Technol. Appl.*, vol. 53, no. 3, pp. 286–303, 2019.
- [11] E. Afful-Dadzie and A. Afful-Dadzie, "Liberation of public data: Exploring central themes in open government data and freedom of information research," *Int. J. Inf. Manage.*, vol. 37, no. 6, pp. 664–672, 2017.
- [12] D. Corrales-Garay, E. M. Mora-Valentín, and M. Ortiz-de-Urbina-Criado, "Entrepreneurship through open data: An opportunity for sustainable development," *Sustain.*, vol. 12, no. 12, 2020.
- [13] N. Sinkovics, "Enhancing the foundations for theorising through bibliometric mapping," *Int. Mark. Rev.*, vol. 33, no. 3, pp. 327–350, 2016.
- [14] J. Baas, M. Schotten, A. Plume, G. Côté, and R. Karimi, "Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies," *Quant. Sci. Stud.*, vol. 1, no. 1, pp. 377–386, 2020.
- [15] N. S. Butt, A. A. Malik, and M. Q. Shahbaz, "Bibliometric analysis of statistics journals indexed in web of science under emerging source citation index," *SAGE Open*, vol. 11, no. 1, 2021.
- [16] N. J. Van Eck, L. Waltman, R. Dekker, and J. Van Den Berg, "A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS," *J. Am. Soc. Inf. Sci. Technol.*, vol. 61, no. 12, pp. 2405–2416, 2010.
- [17] M. S. Altayar, "Motivations for open data adoption: An institutional theory perspective," *Gov. Inf. Q.*, vol. 35, no. 4, pp. 633–643, 2018.
- [18] H.-J. Wang and J. Lo, "Factors influencing the Adoption of open government data at the firm level," in *IEEE Transactions on Engineering Management*, 2020, vol. 67, no. 3, pp. 670–682.
- [19] A. Alhujaylan, L. Car, and M. Ryan, "An investigation of factors influencing private technology organizations' intention to adopt open government data in Saudi Arabia," in *2020 10th Annual Computing and Communication Workshop and Conference, CCWC 2020*, 2020, pp. 654–661.
- [20] S. I. Haini, N. Zairah, and A. Rahim, "Adoption of open government data in local government context: Conceptual model development," in *ICCTA 2019: Proceedings of the 2019 5th International Conference on Computer and Technology Applications*, 2019, pp. 193–198.
- [21] Y. Zhao and B. Fan, "Exploring open government data capacity of government agency: Based on the resource-based theory," *Gov. Inf. Q.*, vol. 35, no. 1, pp. 1–12, 2018.
- [22] B. W. Wirtz, R. Piehler, M. J. Thomas, and P. Daiser, "Resistance of public personnel to open government: A cognitive theory view of implementation barriers towards open government data," *Public Manag. Rev.*, vol. 18, no. 9, pp. 1335–1364, 2016.
- [23] B. Fan and Y. Zhao, "The moderating effect of external pressure on the relationship between internal organizational factors and the quality of open government data," *Gov. Inf. Q.*, vol. 34, no. 3, pp. 396–405, 2017.
- [24] M. Kassen, "Adopting and managing open data: Stakeholder perspectives, challenges and policy recommendations," *Aslib J. Inf. Manag.*, vol. 70, no. 5, pp. 518–537, 2018.
- [25] D. Pirozzi and V. Scarano, *Syntactical heuristics for*

- the open data quality assessment and their applications*, vol. 339. 2019.
- [26] W. D. Sunindyo and I. D. P. D. K. Amrita, "The development of data publishing tool for Indonesian open government data," in *Proceedings of the International Conference on Electrical Engineering and Informatics*, 2019, pp. 30–34.
- [27] P. Utamachant and C. Anutariya, "An analysis of high-value datasets: A case study of Thailand's open government data," 2018.
- [28] H. Zhang and J. Xiao, "Quality assessment framework for open government data: Meta-synthesis of qualitative research, 2009-2019," *Electron. Libr.*, vol. 38, no. 2, pp. 209–222, 2020.
- [29] K. Y. Dahbi, H. Lamharhar, and D. Chiadmi, "Toward an evaluation model for open government data portals," in *Smart Innovation, Systems and Technologies*, 2019, vol. 111, pp. 502–511.
- [30] V. Raça, N. Veljković, G. Velinov, L. Stoimenov, and M. Kon-Popovska, "Real-time monitoring and assessing open government data: A case study of the western Balkan countries," in *Communications in Computer and Information Science*, 2020, vol. 1316, pp. 189–201.
- [31] K. Y. Dahbi, H. Lamharhar, and D. Chiadmi, "Exploring dimensions influencing the usage of Open Government Data portals," 2018.
- [32] J. Ham, Y. Koo, and J. N. Lee, "Provision and usage of open government data: Strategic transformation paths," *Ind. Manag. Data Syst.*, vol. 119, no. 8, pp. 1841–1858, 2019.
- [33] M. A. Osorio-Sanabria, F. Amaya-Fernandez, and M. P. Gonzalez-Zabala, "Developing a model to readiness assessment of open government data in public institutions in Colombia," in *ACM International Conference Proceeding Series*, 2020, pp. 334–340.
- [34] A. T. Chatfield and C. G. Reddick, "A longitudinal cross-sector analysis of open data portal service capability: The case of Australian local governments," *Gov. Inf. Q.*, vol. 34, no. 2, pp. 231–243, 2017.
- [35] A. Zuiderwijk, R. Shinde, and M. Janssen, "Investigating the attainment of open government data objectives: is there a mismatch between objectives and results?," *Int. Rev. Adm. Sci.*, vol. 85, no. 4, pp. 645–672, 2019.
- [36] A. Purwanto, A. Zuiderwijk, and M. Janssen, "Citizens' trust in open government data: A quantitative study about the effects of data quality, system quality and service quality," in *ACM International Conference Proceeding Series*, 2020, pp. 310–318, [Online]. Available: <https://dl.acm.org/doi/10.1145/3396956.3396958>.
- [37] M. A. Osorio-Sanabria, F. Amaya-Fernández, and M. González-Zabala, "Exploring the components of open data ecosystems: A systematic mapping study," in *ACM International Conference Proceeding Series*, 2020, Part F1667.
- [38] J. Linaker and P. Runeson, "Collaboration in open government data ecosystems: Open cross-sector sharing and co-development of data and software," in *Electronic Government*, G. V. Pereira, M. Janssen, H. Lee, I. Lindgren, M. P. R. Bolívar, H. J. Scholl, and A. Zuiderwijk, Eds. Cham, Switzerland: Springer, 2020, pp. 290–303.
- [39] T. Schauppenlehner and A. Muhar, "Theoretical availability versus practical accessibility: The critical role of metadata management in open data portals," *Sustainability*, vol. 10, no. 2, 2018.
- [40] F. Xiao, W. Jeng, and D. He, "Investigating metadata adoptions for open government data portals in US cities," in *Proceedings of the Association for Information Science and Technology*, 2018, vol. 55, no. 1, pp. 573–582.
- [41] A. Zuiderwijk, M. Janssen, and I. Susha, "Improving the speed and ease of open data use through metadata, interaction mechanisms, and quality indicators," *J. Organ. Comput. Electron. Commer.*, vol. 26, no. 1–2, pp. 116–146, 2016.
- [42] P. Křemen and M. Nečaský, "Improving discoverability of open government data with rich metadata descriptions using semantic government vocabulary," *Web Semant. Sci. Serv. Agents World Wide Web*, vol. 55, pp. 1–20, 2019.
- [43] S. Jiang, T. F. Hagelien, M. Natvig, and J. Li, "Ontology-Based Semantic Search for Open Government Data," in *Proceedings - 13th IEEE International Conference on Semantic Computing, ICSC 2019*, 2019, pp. 7–15.
- [44] B. S. Hitz-gamper, O. Neumann, and M. Stürmer, "Balancing control, usability and visibility of linked open government data to create public value," *Int. J. Public Sect. Manag.*, vol. 32, no. 5, pp. 451–466, 2019.
- [45] L. Shi, D. Sukhobok, and N. Nikolov, "Norwegian state of estate report as linked open data," in *On the Move to Meaningful Internet Systems. OTM 2017 Conferences*, 2017, pp. 445–462.
- [46] I. C. Hsu and Y. H. Lin, "Integrated machine learning with semantic web for open government data recommendation based on cloud computing," *Softw. - Pract. Exp.*, vol. 50, no. 12, pp. 2293–2312, 2020.
- [47] T. Cha, "Open government data for machine learning tax recommendation," in *ACM International Conference Proceeding Series*, 2020, pp. 331–333.
- [48] C. H. Wu, S. W. Kuo, and S. C. Kao, "Classification-based data mining applied in vehicle accident prediction," in *Frontiers in Artificial Intelligence and*

- Applications*, vol. 320, IOS Press, 2019, pp. 218–223.
- [49] R. Barcellos, J. Viterbo, L. Miranda, F. Bernardini, C. Maciel, and D. Trevisan, “Transparency in practice: Using visualization to enhance the interpretability of open data,” in *ACM International Conference Proceeding Series*, 2017, pp. 139–148.
- [50] F. Faini and M. Palmirani, “Italian open and big data strategy,” in *Lecture Notes in Computer Science*, 2016, vol. 9831 LNCS, pp. 105–120.
- [51] E. Kalampokis, A. Karamanou, and K. Tarabanis, “Interoperability conflicts in linked open statistical data,” *Inf.*, vol. 10, no. 8, 2019.
- [52] D. Beneventano, S. Bergamaschi, L. Gagliardelli, and L. Po, “Driving innovation in youth policies with open data,” in *Communications in Computer and Information Science*, vol. 631, 2016, pp. 324–344.
- [53] G. Smith and J. Sandberg, “Barriers to innovating with open government data: Exploring experiences across service phases and user types,” *Inf. Polity*, vol. 23, no. 3, pp. 249–265, 2018.
- [54] S. Saxena and I. Muhammad, “Barriers to use open government data in private sector and NGOs in Pakistan,” *Inf. Discov. Deliv.*, vol. 46, no. 1, pp. 67–75, 2018.
- [55] R. Máchová, M. Hub, and M. Lnenicka, “Usability evaluation of open data portals: Evaluating data discoverability, accessibility, and reusability from a stakeholders’ perspective,” *Aslib J. Inf. Manag.*, vol. 70, no. 3, pp. 252–268, 2018.
- [56] B. W. Wirtz, J. C. Weyerer, and M. Rösch, “Citizen and open government: An empirical analysis of antecedents of open government data,” *Int. J. Public Adm.*, vol. 41, no. 4, pp. 308–320, 2018.
- [57] A. Purwanto, A. Zuiderwijk, and M. Janssen, “Citizens’ Motivations for Engaging in Open Data Hackathons,” in *Lecture Notes in Computer Science*, no. July, Panagiotopoulos P. et al. (eds), Ed. Cham: Springer, 2019, pp. 130–141.
- [58] A. Zuiderwijk, C. Volten, M. Kroesen, and M. Gill, “Motivation perspectives on opening up municipality data: Does municipality size matter?,” *Information*, vol. 9, no. 267, pp. 1–26, 2018.
- [59] Jorn Berends, W. Carrara, and C. Radu, “The economic benefits of open data,” 2017. [Online]. Available: <https://www.europeandataportal.eu/en/highlights/economic-benefits-open-data>.
- [60] D. Spiliotopoulos, D. Margaris, and C. Vassilakis, “Citizen engagement for transparent and accountable policy modelling,” in *11th International Conference on Management of Digital EcoSystems, MEDES 2019*, 2019, pp. 158–165.
- [61] F. Kitsios and M. Kamariotou, “Open data hackathons: An innovative strategy to enhance entrepreneurial intention,” *Int. J. Innov. Sci.*, vol. 10, no. 4, pp. 519–538, 2018.
- [62] K. McBride, M. Toots, T. Kalvet, and R. Krimmer, “Open government data driven co-creation: Moving towards citizen-government collaboration,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018, pp. 184–195.
- [63] X. Mountroudou, B. Billings, and L. Mejia-Ricart, “Not just another Internet of Things taxonomy: A method for validation of taxonomies,” *Internet of Things*, vol. 6, no. 2019, 2019.
- [64] B. Carrion, T. Onorati, P. Díaz, and V. Triga, “A taxonomy generation tool for semantic visual analysis of large corpus of documents,” in *Multimedia Tools and Applications*, vol. 78, no. 23, Multimedia Tools and Applications, 2019, pp. 32919–32937.
- [65] D. Buhalis and K. Volchek, “Bridging marketing theory and big data analytics: The taxonomy of marketing attribution,” *Int. J. Inf. Manage.*, vol. 56, 2021.
- [66] L. Versluis and A. Iosup, “A survey of domains in workflow scheduling in computing infrastructures: Community and keyword analysis, emerging trends, and taxonomies,” *Futur. Gener. Comput. Syst.*, vol. 123, pp. 156–177, 2021.
- [67] Y. Charalabidis, C. Alexopoulos, and E. Loukis, “A taxonomy of open government data research areas and topics,” *J. Organ. Comput. Electron. Commer.*, vol. 26, no. 1–2, pp. 41–63, 2016.
- [68] I. Susha, M. Janssen, and S. Verhulst, “Data collaboratives as a new frontier of cross-sector partnerships in the age of open data: Taxonomy development,” in *Proceedings of the 50th Hawaii International Conference on System Sciences (2017)*, 2017, pp. 2691–2700.
- [69] M. Usman, R. Britto, J. Börstler, and E. Mendes, “Taxonomies in software engineering: A Systematic mapping study and a revised taxonomy development method,” *Inf. Softw. Technol.*, vol. 85, pp. 43–59, 2017.
- [70] G. Robert and I. Vessey, “Contemporary application-domain taxonomies,” *Ieee Softw.*, vol. 12, no. 4, pp. 63–76, 1995.
- [71] D. Szopinski, T. Schoormann, and D. Kundisch, “Because your taxonomy is worth it: Towards a framework for taxonomy evaluation,” *27th Eur. Conf. Inf. Syst. - Inf. Syst. a Shar. Soc. ECIS 2019*, pp. 0–19, 2020.