

# Technical aspects for inclusiveness across user domains in data portals

Abdul Aziz<sup>1,2</sup>

<sup>1</sup>Universidad de Zaragoza, Zaragoza, Spain

<sup>2</sup>Aragon Institute of Engineering Research (I3A)

## Abstract

In recent years, open data initiatives have acquired substantial popularity, with governments, organizations, and individuals openly sharing data in an effort to stimulate openness, innovation, and engagement. However, in order to get the most out of open data portals, it is very necessary to make them available to users of all different kinds of backgrounds. This research study investigates and explores the technical components that might encourage inclusion and fulfill the various needs of many user domains. Some examples of these components include user engagement on the data portal, providing feedback on the datasets, and how simple it is to identify the dataset that is relevant to the user looking for. When these technological components of data portals are put into place, the portals become capable of efficiently serving a diverse range of users. This encourages access to essential data and stimulates innovation that is data-driven across all sectors of the economy. Lastly, one of the aspects of my research that I will be focusing on during the whole of this time will be investigating new ways and technologies that may provide open data publishers with actionable feedback based on how open data is used in a variety of user domains.

## Keywords

Open Data, Technical perspective, Inclusiveness, Open Government Data, Open Data Portals

## 1. Introduction

The Open Data (OD) movement is expanding at an astounding rate in today's fast-paced digital environment. The expansion of the Open Data (OD) movement, is made possible by the expanded availability of data on Open Data Portals [1]. Data portals have emerged as crucial hubs for the dissemination and easy accessibility of massive amounts of open data in today's information-driven world. The European Commission[2] argues that The continual distribution of open data through Open Government Data (OGD) portals enhances both the need for and the quality of data that is of a higher level. OGD portals are vital to the release of data, and this improves both the demand for and the quality of data. Specifically, the current trend of disclosing government data through the EU portal [2, 3] is largely driven by the utilization and repurposing of public sector data. Participation from users is generally recognised as being crucial to the growth of open data ecosystems. However, suppliers are the primary force behind

---

*BIR-WS 2023: BIR 2023 Workshops and Doctoral Consortium, 22nd International Conference on Perspectives in Business Informatics Research (BIR 2023), September 13-15, 2023, Ascoli Piceno, Italy*

✉ [abdul.aziz@unizar.es](mailto:abdul.aziz@unizar.es) (A. Aziz)

ORCID [0000-0003-3615-4573](https://orcid.org/0000-0003-3615-4573) (A. Aziz)



© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



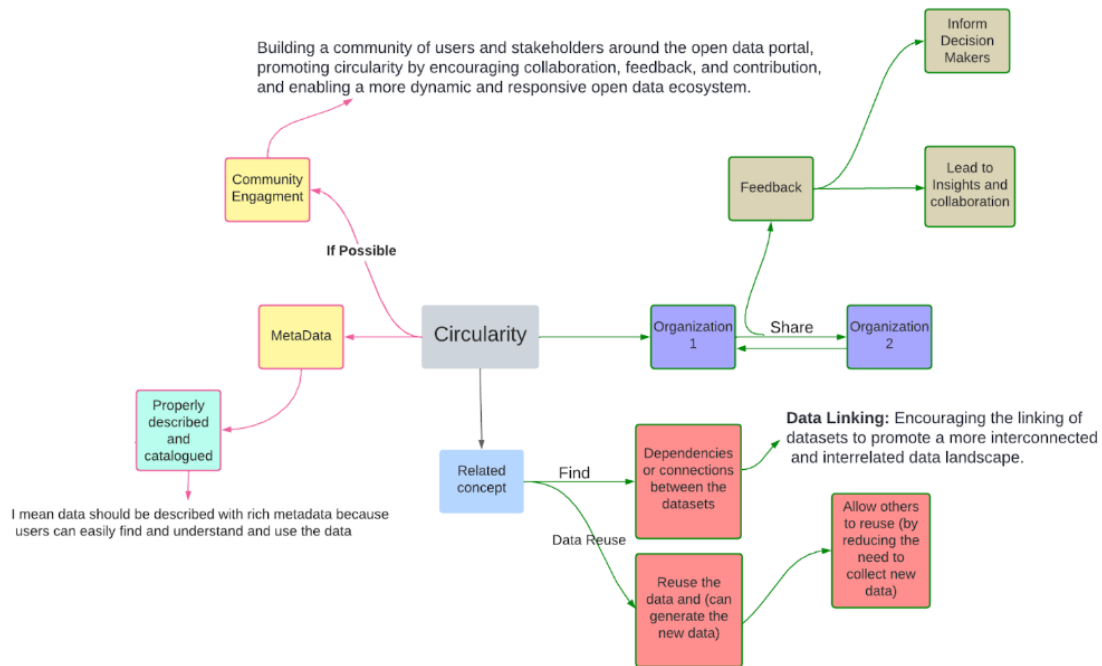
CEUR Workshop Proceedings (CEUR-WS.org)

the emergence of existing ecosystems [4]. The increased availability of open data over the past decade is a direct consequence of open data portals whose objectives include, among others, improved government service delivery, greater openness and accountability, more active citizen participation, and increased economic and social value creation [5, 6]. However, despite the growing prevalence of open data, the genuine notion of open data is still far from being realized. [7]. This lack of active engagement may be linked to a number of things, including a lack of knowledge, experience, and familiarity with the topic. This knowledge gap is the focus of the research study that is reported in this paper. The study intends to investigate innovative approaches for recognizing and addressing feedback gaps between end users, stakeholders, and data providers. The goal is to make it easier to make full use of the potential offered by open data through open data portals while simultaneously involving a variety of diverse disciplines. Because of the disparity in representation, data portals are unable to facilitate the making of decisions that are both fair and driven by data across all industries. As a result, data portals need to identify and solve the technological elements that encourage inclusiveness while also promoting circularity, and they also need to suit the diverse expectations of all user domains. In the context of open data, according to [4], circularity refers to the requirement that data be open, reusable, and accessible in a continuous cycle. For our research method, the new data sets that are available can be value added by utilizing open data portal that can contributed back to the open data community to maintain the cycle. The fundamental principle of circularity is that open data should be accessible and reusable, resulting in a cycle of data production, consumption, and dissemination that is self-sustaining [8]. Figure 1, shows the concept regarding circularity for an open data portal.

Consider an open data portal as a central online library where anyone, including the government, non-profit organizations, and the general public, can discover and use data produced by numerous entities. With a concentration on co-creation, the open data portal incorporates data contributions from a wide variety of sources, such as government agencies, academic institutions, journalists and regular citizens.

The term “co-production” refers to the process by which data consumers (such as academicians and developers) collaborate with the data provider to create new and valuable products (such as applications, visualizations, and insights). An example of this would be a developer creating an app that visualizes real-time environmental data, or a researcher using data from an open data portal to create an interactive map emphasizing regional patterns. These examples illustrate co-production, wherein previously gathered data is used as a starting point to develop novel approaches to existing problems or areas of interest.

To sum up, co-creation emphasizes group information production, drawing on the expertise of multiple parties to enhance the portal’s offerings. Nonetheless, the purpose of co-production is to use this information to generate actionable resources and insights that facilitate intelligent decision-making. The combination of these concepts generates an ecosystem for open data in which cooperation and innovation work in tandem to make data more broadly accessible, usable, and valuable to all members of society.



**Figure 1:** Circularity in Open Data Portals

## 2. Related Work

This section enlightens the previous work in the field of open data, basis of need to explore the open data portals.

According to [9], for “open data” to have a positive influence on the poor and disadvantaged, it must intervene directly to guarantee that aspects presently lacking in the technical environment is made accessible. For the open government portal assessment, [10] proposed that the certain criteria must be met by data portals in order to maximize openness and responsibility. Disclosure of massive datasets alone does not guarantee openness or responsibility. Entity coverage, information kinds, information seeking tactics, and data quality characteristics are all identified as being critical in the study’s conclusion that these goals may be successfully attained. for the transparency of the design for an open data portal, authors in [11] investigated that on the basis of transparency-by-design, the characteristics of open data portals enable transparency, and an enumeration of open data portal categories and criteria is required for the most transparent data portal. Open data portals meet transparency criteria by associating literature properties with stages of the transparency cycle. Moreover, [12] investigated the quality of the Czech Republic open data portal, so they proposed a framework and validated to evaluate the quality of open data portals on the national level. The findings demonstrate that there is a need for quality standards and that open data portals vary in the amount of datasets they provide and the complexity of their services. Likewise, There are several research works in the published literature about the monitoring of the quality of Open Data Portals [13, 14], which are relevant

to have an overall perspective of the current status of Open Data initiatives, their maturity or their commitment to FAIR principles [15]

In terms of the open government data portal usability, [16] conducted a research on a user-centred usability analysis of 41 open government data portals and how the usability of a portal might potentially influence the reuse of data. The authors responded to two questions, the first of which was “How can the usability of open government data portals be evaluated and compared across contexts?” as well as “What are the most commonly missing usability aspects from open government data portals?” In order to find answers to these concerns, a representative sample of 41 national open government data portals was selected for a comprehensive usability evaluation, which included the analysis of 40 user comments. Through this investigation, a paradigm for the usability examination of open government data portals has been validated, and the strengths and shortcomings of portal usability that are consistent across contexts have been identified. For checking the impact of impediments on open government data use, [17] conducted a mixed approach of questionnaires and interviews where they identified several barriers to use, such as brainstorming challenges and a lack of longitudinal data through the analysis of the collected data, and their effects were discussed. Whereas, concerning user-centric Open Government Data Initiatives, Nikiforova and McBrid [16] analyzed and contrasted the many settings of the employment of Open Government Data Portals by users, with a focus on the user-centered features that are most frequently ignored. In order to validate the accessibility of open government data portals from a user’s point of view, they used the questionnaire method. Notably, Zhu and Freeman [18] evaluated various approaches to user interactions with Open Government Data Initiatives and developed a framework that they called the user interaction framework. Using this framework, they evaluated the United States Municipal Open Data Portals and provided the findings regarding user understanding and engagement with the data portals.

Additionally, [19] proposed a methodology for measuring user involvement in open data portals by analysis of Twitter activity of the open data users to determine the correlation between social network activity and the main features characterizing the size, quality, and maturity of Open Data initiatives. They evaluated 27 European Open Government Data portals and their Twitter activity for the year 2021. This enabled them to compare their results with the 2021 Open Data Maturity Report. In a similar manner, Begany and Gil-Garcia [7] analyzed the web analytics behavioral data that was gathered from the New York State open health data portal in order to track the levels of user engagement. In addition to this, they placed a strong focus on the actual usage of open data and, in particular, how users of open data portals interact with open datasets. Lastly, the Open Data Maturity Report [20] has been used as a standard for evaluating the progress that European open data portals have made in the area of open data for a number of years. The policy, the effect, the portal, and the quality are the four aspects that are discussed in the text. A sustainability variable that specifies efforts taken to guarantee the visibility of the portal, including its presence on social media, is included in the portal dimension of this indicator.

### 3. Methods

We use a literature review approach to produce artifacts that back up the overall results of our research method, facilitate collaborative needs, and provide a feedback loop to data publishers, policymakers and other stakeholders. Moreover, we investigated through literature analysis that the theoretical foundations underlying the technical aspects for inclusiveness in open data portals needs to under the tree of FAIR principles, and more specifically on circular ecosystem by comparing technical inclusiveness across different user domains.

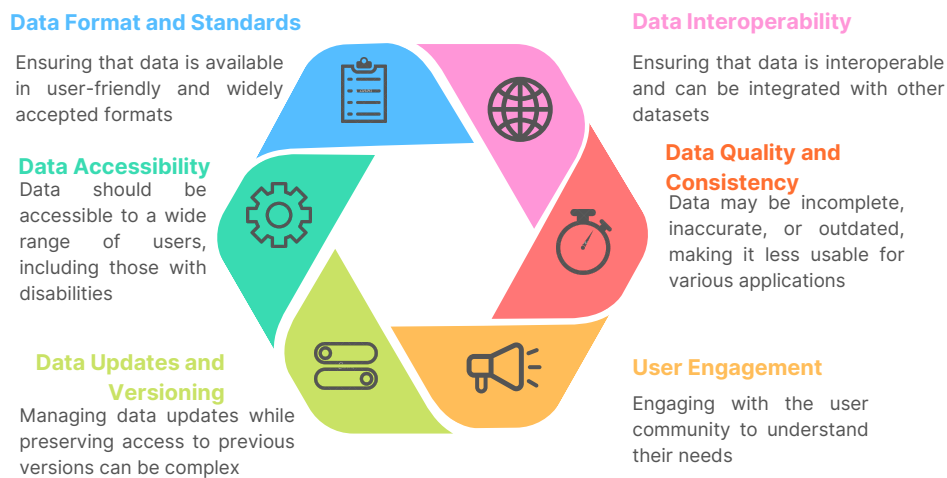
To make open data portals easier to use to all users, accessibility tools and language support are the next targets in our research method because that is how the users with diverse linguistic backgrounds can benefit from the translation possibilities of an open data portal.

Additionally, we also identified problems and tools for automatic thematic annotation of government data where proposed the methodology like the following:

- To determine whether the open government datasets have been correctly annotated, we first conducted a corpus-level data profile and evaluated the validity of thematic categorization, which improves data discoverability and recall.
- The second stage included using the learned models to automatically categorize a corpus that had been manually annotated.
- Lastly, predicting the closest theme of a dataset based on word embeddings

A major feature of open data portals is the ability to offer input and comment on data that is freely accessible to the public. There are a lot of different ways that people may provide input, but they all accomplish the same thing, which is to give the general public a say in how public data are managed. As a preliminary result of our literature analysis, we came to the early results that open data initiatives that integrate feedback for the datasets as well as input for the data publishers into the feedback loop would be more effective for an open data portal. This was one of our initial findings. When an Open Data portal is circular, it encourages the recycling of existing data rather than the acquisition of brand-new information. This is in contrast to linear efforts, which focus on gathering brand-new information. If, for instance, the organizations who are pushing open data portals do not incorporate user input, then it is possible that they will not be aware of any current datasets that need to be updated in order to promote reuse [3, 21]. Creating an online community for inhabitants to share their thoughts and opinions is a second common strategy. Input can also be gathered through surveys, questionnaires, and focus group discussions. In order for the public to have a say in how government data is utilized and made available, it is imperative that open government data portals include a feedback mechanism [22]. User feedback may be collected through online, print, or in-person interviews. User comments are available on reviews, forums, and social media platforms. These remarks disclose the preferences of Open Data Portal users. Indirectly, google analytics may collect open data portal user input. Analytics can quantify Open Data Portal utilization. It is possible to track dataset views, page views, time spent on the open data portal, and search terms etc [23].

The next step in this study is to do research on emerging practices, technologies, and tools that may assist in the process of providing data publishers with actionable feedback and developing a conceptual model for reusability in open data portals. The history of a piece of data is referred



**Figure 2:** Reusability in Open Data Portals

to as its "provenance," and it contains information about where the data originated, who owns it, and any alterations that have been made [24]. The ability of data to be employed for a variety of reasons and by a lot of different users is what we mean when we talk about reusability [25]. Figure 2. shows the initial steps of our conceptual model for the reusability in open data portals. The capacity of open data to be reused is essential to the success of open data. information is possible for data to promote creativity, improve decision-making, and develop new economic prospects when information is easily available and used by a broad range of persons. For instance, open data may be used to create new apps and services, to recognize patterns and trends that can be used to guide policy choices, or to enable study and analysis [25, 26].

In order to ensure the accuracy and reliability of the data that is made accessible via an open data portal, it is essential to take into account both the original source of the data and the extent to which it may be reused. This involves giving explicit documentation and information indicating the origin of the data and any alterations, as well as making the data generally available and useable by a wide variety of users. In addition, this includes providing explicit documentation and information identifying the provenance of the data.

## 4. Results

This section will provide the overall overview of our results till to-date. As we had the road map to carry on the main objectives of this research which actually aims research new methods for identifying gaps in open data portals and open data initiatives related to reuse that difficult user feedback, especially when different user domains are involved.

So we came up with the initial findings as by publishing two research articles, first one as a poster presentation in the Vol. 10 (2022): Actas de la XI Jornada de Jóvenes Investigadores e Investigadoras del I3A namely “Towards a sustainable Open Data Ecosystem First steps for optimizing findability and user feedback” where we produce insights about our research findings, followed by another research article namely “analysing user involvement in Open Government Data initiatives”, a full paper presented in the 26th International Conference on Theory and Practice of Digital Libraries, in which we presented the involvement of the users in open data initiatives. Recently we have completed the experiments and got the desired results for the extension of our previous paper presented in the 26th International Conference on Theory and Practice of Digital Libraries. Currently, we have finished the experiments for the next two articles, out of which one is the extension of our previous conference paper presented in the 26th International Conference on Theory and Practice of Digital Libraries, where we include the temporal analysis of the evolution of Twitter activity generated by Open Data initiatives as well as datasets published on the EU Open Data Portal with maturity scoer of the portal for the years from 2017 to 2021, and proposed the method of self organizing maps to see the insights in the form of clusers for the respective open data portal. And the other paper is about problems and tools for automatic thematic annotation of government data on the data portals, where we investigated that to establish whether open government datasets were appropriately annotated, we did a corpus-level data profile and tested theme categorization, which improves data findability and recall. The second phase applied the learned models to a manually-annotated corpus, and the third used word embeddings to estimate a dataset’s closest topic. Now we are in the writing phase for the said two completed experiments

## 5. Discussion

The study makes a contribution to a better understanding of the gaps that exist in open data portals and programs linked to user feedback and engagement, particularly when many user domains are taken into consideration. The research gives insights into improving findability and user experiences in the open data ecosystem by evaluating user behavior and interaction. These insights may be used to optimize findability. Experiments now being conducted on temporal analysis and theme annotation are helping to further broaden our understanding of the development and curation of open data portals. This understanding is essential in the larger context of open government data efforts and data accessibility. The findings shed light on anticipated and surprising patterns in user engagement with open data, opening the way for enhanced tactics and techniques to create more transparent and accountable governance via open data portals. The results of the study shed light on expected and unexpected patterns in user interaction with open data. Overall, the findings provide light on how open data portals

might better facilitate findability, user involvement, and topic annotation. However, the study only looks at data from 2017–2021, which seems to be short of a time period to capture the most recent changes in the open data landscape, given the frequency with which open data portals update their published datasets.

## 6. Conclusion

The effectiveness and impact of open data portals are strongly reliant on their capacity to serve a diverse range of user domains. This is one of the most important factors determining their influence. The use of a number of different technical variables enables data portals to provide a high level of service to a varied audience of users in an effective manner. According to the findings of this study, there is a significant need for open data portals, which have the potential to enhance the degree to which the government is transparent, as well as user input and output. We were successful in accomplishing this goal as a result of the publication of two research publications that discussed the findability of open data portals, user feedback, and user interaction. The practice of thematically annotating open data that is made accessible via open data portals was also investigated as part of this study. According to the findings of this research, portals that allow users to share datasets have the potential to promote transparency, widen engagement, and kickstart data-driven innovation across a wide variety of user groups. On the other hand, open data portals might be hampered by problems such as an inadequate supply of high-quality data, subjectivity in the feedback provided by users, or technical challenges. In conclusion, the findings expand our understanding of the technology considerations for diversity in data portals and their potential to democratize access to information and empower individuals across a broad variety of businesses and regions if they adopt an inclusive mindset. In addition, the results increase our knowledge of the technological considerations for diversity in open data portals.

## Acknowledgments

I would like to acknowledge the ODECO project, received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 955569

I would also like to acknowledge my co-supervisor: Professor Javier Noguera Iso and Dagoberto José Herrera Murillo from Advanced Information Systems Lab (IAAA), Computer Science and Systems Engineering Department University of Zaragoza.

## References

- [1] T. Jetzek, The value generating mechanisms of open government data, *Geoforum Perspektiv* 12 (2013).
- [2] W. Carrara, W.-S. Chan, S. Fischer, E. Van-Steenbergen, Creating value through open data: Study on the impact of re-use of public data resources. European Commission, 2015.



- [3] A. Simonofski, A. Zuiderwijk, A. Clarinval, W. Hammedi, Tailoring open government data portals for lay citizens: A gamification theory approach, *International Journal of Information Management* 65 (2022) 102511.
- [4] B. Van Loenen, A. Zuiderwijk, G. Vancauwenberghe, F. J. Lopez-Pellicer, I. Mulder, C. Alexopoulos, R. Magnussen, M. Saddiqa, M. D. de Rosnay, J. Crompvoets, et al., Towards value-creating and sustainable open data ecosystems: A comparative case study and a research agenda, *JeDEM-eJournal of eDemocracy and Open Government* 13 (2021) 1–27.
- [5] S. de Juana-Espinosa, S. Luján-Mora, Open government data portals in the European Union: A dataset from 2015 to 2017, *Data in brief* 29 (2020) 105156.
- [6] L. F. Luna-Reyes, J. C. Bertot, S. Mellouli, *Open government, open data and digital government*, 2014.
- [7] G. M. Begany, J. R. Gil-Garcia, Understanding the actual use of open data: Levels of engagement and how they are related, *Telematics and Informatics* 63 (2021) 101673.
- [8] P. Colpaert, S. Joye, P. Mechant, E. Mannens, R. Van de Walle, The 5 stars of open data portals, in: *Proceedings of the 7th International Conference on Methodologies, Technologies and Tools Enabling E-Government (MeTTeG13)*, University of Vigo, Spain, 2013, pp. 61–67.
- [9] M. B. Gurstein, Open data: Empowering the empowered or effective data use for everyone?, *First Monday* 16 (2011). URL: <https://firstmonday.org/ojs/index.php/fm/article/view/3316>. doi:10.5210/fm.v16i2.3316.
- [10] R. P. Lourenço, Open government portals assessment: a transparency for accountability perspective, in: *Electronic Government: 12th IFIP WG 8.5 International Conference, EGOV 2013*, Koblenz, Germany, September 16-19, 2013. *Proceedings* 12, Springer, 2013, pp. 62–74.
- [11] M. Lnenicka, A. Nikiforova, Transparency-by-design: What is the role of open data portals?, *Telematics and Informatics* 61 (2021) 101605.
- [12] R. Máchová, M. Lněnička, Evaluating the quality of open data portals on the national level, *Journal of theoretical and applied electronic commerce research* 12 (2017) 21–41.
- [13] S. Kubler, J. Robert, S. Neumaier, J. Umbrich, Y. Le Traon, Comparison of metadata quality in open data portals using the Analytic Hierarchy Process, *Government Information Quarterly* 35 (2018) 13–29.
- [14] J. Noguerras-Iso, J. Lacasta, M. A. Ureña-Cámara, F. J. Ariza-López, Quality of Metadata in Open Data Portals, *IEEE Access* 9 (2021) 60364–60382.
- [15] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, et al., The fair guiding principles for scientific data management and stewardship, *Scientific data* 3 (2016) 1–9.
- [16] A. Nikiforova, K. McBride, Open government data portal usability: A user-centred usability analysis of 41 open government data portals, *Telematics and Informatics* 58 (2021) 101539.
- [17] J. Crusoe, A. Simonofski, A. Clarinval, E. Gebka, The impact of impediments on open government data use: insights from users, in: *2019 13th international conference on research challenges in information science (rcis)*, IEEE, 2019, pp. 1–12.
- [18] X. Zhu, M. A. Freeman, An evaluation of US municipal open data portals: A user interaction framework, *Journal of the Association for Information Science and Technology* 70 (2019) 27–37.
- [19] D. J. Herrera-Murillo, A. Aziz, J. Noguerras-Iso, F. J. Lopez-Pellicer, Analysing user involve-

- ment in open government data initiatives, in: International Conference on Theory and Practice of Digital Libraries, Springer, 2022, pp. 175–186.
- [20] Publications Office of the European Union, Open data maturity report 2021, Publications Office, LU, 2022. URL: <https://data.europa.eu/doi/10.2830/394148>.
- [21] C. Alexopoulos, E. Loukis, Y. Charalabidis, A platform for closing the open data feedback loop based on web2.0 functionality, *JeDEM-eJournal of eDemocracy and Open government* 6 (2014) 62–68.
- [22] A. Aziz, D. J. H. Murillo, J. N. Iso, F. J. L. Pellicer, Towards a sustainable open data ecosystem: First steps for optimizing findability and user feedback, in: *Jornada de Jóvenes Investigadores e Investigadoras del I3A*, 2022.
- [23] A. Nikiforova, Comparative analysis of national open data portals or whether your portal is ready to bring benefits from open data, in: *IADIS International Conference on ICT, Society and Human Beings*, 2020, pp. 21–23.
- [24] M. De Martino, S. Rosim, A. Quarati, Hydrographic datasets in open government data portals: Mitigation of reusability issues through provenance documentation, in: *Research Conference on Metadata and Semantics Research*, Springer, 2019, pp. 307–319.
- [25] J. Attard, F. Orlandi, S. Scerri, S. Auer, A systematic review of open government data initiatives, *Government information quarterly* 32 (2015) 399–418.
- [26] L. Ding, T. Lebo, J. S. Erickson, D. DiFranzo, G. T. Williams, X. Li, J. Michaelis, A. Graves, J. G. Zheng, Z. Shangguan, et al., Twc logd: A portal for linked open government data ecosystems, *Journal of Web Semantics* 9 (2011) 325–333.