The impact of the activities of non-profit data intermediaries (NPDIs)

A qualitative study on the impact of NPDIs in reducing barriers to using Open Government Data (OGD) in Europe

Master thesis submitted to Delft University of Technology

in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in **Complex Systems Engineering and Management**

Faculty of Technology, Policy and Management

by

Rayyan Mohammed S Alorini

5358175

To be defended in public on August 14th 2023

Graduation committee

Chairperson and First Supervisor: Dr. A.M.G. Zuiderwijk- van Eijk, ICT
Second Supervisor: Dr. M.L.C. de Bruijne, O & G
Advisor: L. (Liubov) Pilshchikova, ICT

TU Delft

Preface

Dear reader,

I hope you enjoy reading my master's thesis, the final milestone of pursuing a master's in Complex Systems Engineering and Management. As I reach the final milestone, it marks the end of my enjoyable journey at TU Delft.

I sincerely thank my supervisors, Anneke Zuiderwijk-van Eijk, Mark de Bruijne, and Liubov Pilshchikova. Anneke, I am grateful for your unlimited support, motivation, and constructive feedback, which have been invaluable throughout my master's thesis journey. Mark, I am grateful for your thorough and insightful feedback, which has broadened my perspectives and immensely helped me in this project. Liubov, your supportive attitude and willingness to provide guidance and suggestions whenever I needed it has been truly appreciated. I would also like to thank the interview participants who generously dedicated their time to make this research possible.

To my loving wife, Bshayer, your boundless love, understanding, patience, and encouragement have been the foundation of my strength, motivating me to get past all challenges in my academic endeavours. To my two wonderful kids, Fahad and Mohammed, your bright smiles and spreading laughter have been a constant reminder that no matter the challenges, the joy of life never dims. The three of you have made my academic journey a fantastic experience filled with joy and beautiful memories. Lastly, I dedicate every achievement to my parents, whose love, support and guidance have shaped the person I am today.

Grateful to all of you for being a part of this journey.

Rayyan Mohammed S Alorini
Delft, July 2023

Summary

The use and value of Open Government Data (OGD) are gaining increasing attention from governments, corporations, and individuals. Governments have developed portals and started initiatives to share, promote and facilitate OGD interactions. The key objectives of the release of OGD by public entities are public responsibility, economic expansion, and operational and technical values. However, the potential of OGD is limited by the ability to reuse public data. Although OGD is becoming more available, the value realisation of OGD is hindered by its limited use. Within the OGD ecosystem, data intermediaries, especially Non-Profit Organizations (NPOs), have been playing a role within the OGD ecosystem, perhaps making OGD more usable for users. We assume that the barriers that are most likely to hinder the potential of OGD are more linked to the barriers to using OGD other than structural or provisioning barriers. Furthermore, the barriers to using OGD might be more influenced by the activities provided by the non-profit data intermediaries (NPDIs). Given the anticipated transformation potential of OGD and the current barriers to reusing OGD, it is necessary to investigate how the existence and activities of the NPDIs affected the barriers to using OGD. This study explores the impact of the NPDIs' activities, particularly their impact on barriers to OGD use. We introduced the following research question:

How do European non-profit data intermediaries reduce the barriers to using OGD?

Adopting an explorative case study approach, we first identified the barriers to using OGD through the use of a literate review method. We introduced three categories of barriers in which these barriers originated. These categories are OGD portals, OGD datasets and OGD users' abilities. Furthermore, we conducted a narrative literature review to analyse and identify the roles and activities of the open data intermediates. We identified seven roles that summarise the set of functions of responsibility that open data intermediaries assume and fourteen activities by specifying certain actions or tasks that open data intermediaries fulfil. Based on the literature review, we could not link the activities to the barriers. However, we conceptualised the barriers to using OGD and the activities of the open data intermediaries separately.

Following this, we conducted desk research by analysing NPDIs' websites, complemented by interviews with the selected NPDIs to analyse their activities and roles in the OGD ecosystem toward reducing the barriers to using OGD. Our findings of the activities of the NPDIs showed that NPDIs have a different scope and objectives, often a social goal, compared to open data intermediaries who might be, for instance, interested in profit. However, the roles and activities of NPDIs, are similar to open

data intermediaries in terms of their operations. NPDIs do not necessarily focus on specific roles or activities but rather provide a wide range of services and conduct various activities. This may be driven by the overarching social goal where they try to offer a complete solution that does not lack in some areas. However, some of the activities we identified might be specific to NPDIs, such as promoting the use of OGD, allowing OGD users to disseminate their OGD-driven insights and improving and facilitating the process of using OGD.

We conducted nine semi-structured interviews with individuals from NPDIs and their users to learn how they perceive NPDIs are reducing the barriers to using OGD. Through coding the interview transcripts, we applied a thematic analysis approach to the data obtained from the interview. We have identified six themes that represent how NPDIs reduce the barriers to using OGD. These themes are; building OGD capacity and expertise, improving OGD accessibility quality and usability, empowering OGD users, OGD process optimisation, promoting and advocating OGD-related activities and policies, and facilitating and improving stakeholders' collaboration and engagement. The findings showed that NPDIs mainly emphasised improving OGD capacity, accessibility, availability, and findability. Then, we conducted a focus group session to triangulate our case study data. Three participants participated in our session, two represented the NPDIs perspective, and one represented the user perspective. We concluded that NPDIs' activities reduce the barriers to using OGD, such as OGD users' ability, OGD accessibility, and quality of OGD datasets and portals. However, quantifying their impact or linking their activities to some of the barriers they reduce is challenging due to the multiple impacts of the NPDIs' activities.

Our study attempted to address the gap in the literature regarding the NPDIs' impact in reducing the barriers to using OGD. Also, our study provided insight into how NPDIs reduce the barriers to using OGD; we identified their characteristics and strategies, which contribute to setting the groundwork for future research exploring the link between NPDIs activities and barriers to using OGD. Our results underline the value of NPDIs to the OGD ecosystem. Policymakers or key NPDIs persons can leverage the results of this study to capitalise on the identified opportunities, such as trying to make NPDIs efforts more proactive in anticipating the barriers of OGD to contribute to better use of OGD ultimately.

Contents

List of Figures

List of Tables

List of abbreviations and acronyms

Abbreviation	Defintion
OGD	Open Government Data
NPDI	Non-profit data intermesiaries
UN	United Nations
WHO	World Health Orginsation
NPO	Non-Profit Organization
API	Application Programming Interface
ASOC	At the School of OpenCohesion
OSF	Open State Foundation
OWiD	Our World in Data
STS	Socio-technical systems
CAS	Complex adaptive systems

1 Introduction

1.1. Problem statement

The use and value of Open Government Data (OGD) is gaining increasing attention from governments, corporations, and individuals. Governments have developed portals and started initiatives to share, promote and facilitate OGD interactions. OGD contains data from the government across various sectors, including transportation, health, education, tourism, social service, and environmental. The key objectives of the release of OGD are public responsibility, economic expansion, and operational and technical values (Janssen et al., 2012). However, the potential of OGD is limited by the ability to reuse public data (Kulk & Van Loenen, 2012). With the release of OGD, individuals, organisations, and the various users of OGD can analyse and use the data for several uses, for instance, for monitoring public funds or innovating and creating new products or services leveraging the available datasets. Although OGD is becoming more available, the value realisation of OGD is hindered by its limited use (Martin, 2014). Several barriers, including technical and statistical skills, limit the user's potential to use, link, and integrate OGD.

1.2. Definition of core concepts

1.2.1. Open Government Data

Taking a broad perspective, let us first define data. The Oxford Dictionary defines data as "facts or information, especially when examined and used to find out things or make decisions ("data noun - Definition, pictures, pronunciation and usage notes | Oxford Advanced Learner's Dictionary at OxfordLearnersDictionaries.com", n.d.)". A more narrow definition would be big data; according to the Oxford Dictionary, big data are "sets of information that are too large or too complex to handle, analyse or use with standard methods". The big data definition is very similar to OGD. However, it should include the "Open" and "Government" concepts to make it more precise. According to Reis et al. (2018), Open Data are datasets that may be freely used, reused, and disseminated by anybody with the bare minimum of credit to the source and sharing under the same conditions. The World Bank (2019) define the openness of data in two dimensions. 1) legally open, which means the data must be published in the public domain; 2) technically open, so that anybody can access and use the data using standard, freely downloadable software tools; the data must be provided in electronic formats that are machine-accessible and non-proprietary. The definition of

Crusoe et al. (2019) will be adopted; they define OGD as interoperable data that public institutions share publicly for anybody to use and redistribute without restriction. Publishers are the public institutions that distribute data, while users who utilise the data in new ways (Crusoe et al., 2019).

1.2.2. OGD ecosystem

OGD is generated and used by a complex network of actors, and each actor has unique interests and purposes. According to Chattapadhyay (2014b), who adopted the perspective of the OGD ecosystem from Rufus Pollock (n.d.), who identified and conceptualised the OGD ecosystem as a network of open data producers and users. They discussed that, in an open data ecosystem, the data is shared by its creator and consumers. In an ecosystem of open data, public institutions distribute open data and also obtain new datasets, analyses, and insights with additional values from various data users. Furthermore, open data ecosystems emerge due to user adaptation, feedback loops, dynamic relationships between suppliers and users, and other interrelated aspects (Zuiderwijk et al., 2014). Zuiderwijk and de Reuver (2021) proposed the following four key elements to be facilitated by the OGD ecosystem. 1) Making the datasets available and published online implies that data users depend on public agencies to release their datasets to be used again. 2) searching, discovering, evaluating, and displaying data, as well as knowing about the licensing relating to the data. 3) It should enable data cleansing, analysis, development, integrating, linking, and visualisation. 4) Allow for data interpretation, discussion, and feedback to the data source and other stakeholders.

Several user groups utilise OGD, Safarov et al. (2017) in their systematic literature review have identified different users of OGD. **Citizens** are one the main users who might be seen sometimes as the primary user of OGD. **Business** users are believed to be motivated by revenue and public value when using public data. **Researchers** who may use OGD for investigating government initiatives in public policy, the educational system, healthcare, and other areas. **Developers** that create professional networks of developers through the use of OGD to assist in creating universal tools and to promote the standardisation of OGD usage practices. **Journalists** who use OGD to conduct inquiries, write articles for newspapers and visualise public data as part of their daily activities. **Non-Governmental organisations** who are creating guidelines, fostering discussion, demanding national governments to release more data, and arranging events to promote and use OGD. In this study, we will have a broad scope of users, and it will not be based on a specific user group's perspective but rather a broader perspective to have a wide spectrum of users' perspectives.

1.2.3. OGD lifecycle

Phases of the OGD lifecycle include data creation, publication, discovery, use, and user feedback. Once the data is published, the users are the ones who are responsible for finding, aggregating, cleaning, and turning the scattered data into useful information. However, this is not an easy task considering the complexity of the tasks and the required advanced technical and statistical capabilities. van Veenstra and van den Broek (2015) has identified a five-phase model of open data lifecycle Figure 1.1. The first phase starts with identifying and selecting the data to be published and setting the strategy to publicly share data by the data providers. Second, the preparation phase involves defining the requirements and getting the data technically ready. The third phase is the publication, which ensures the data is technically findable and the available data is advertised to be re-used. Fourth is the re-use phase, which includes maintaining the data and creating a group of re-users to help collect input so that the data's quality is improved and more re-users utilise the data. The fifth phase is the evaluation phase, which entails evaluating the data proposition and incorporating the strategy into the organisation's processes. OGD users are impacted by the different phases of the OGD lifecycle. However, issues facing OGD users are more visible in phase-4 **re-use** when it comes to using the data that were made available. This does not mean the issue occurs because of shortcomings in the 4th phase alone, but rather, the phases are interconnected, and the phases have implications on other phases either directly or indirectly. Once the data is published, users try to discover the data and build use-cases of the published data for various purposes such as transparency and accountability, economic development or citizen participation (Safarov et al., 2017).

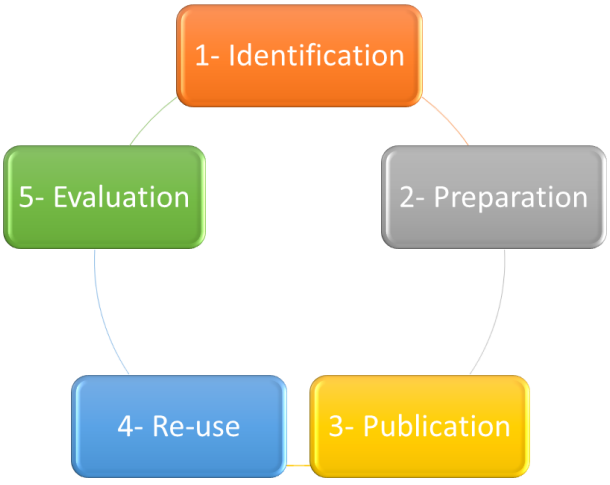

Figure 1.1: Open data life-cycle, adopted from (van Veenstra & van den Broek, 2015)

1.2.3.1. Barriers to using OGD

Several barriers to using OGD in various domains may be related to organisational, legal, political, technological, financial, economic, social, and cultural factors (Saxena

& Muhammad, 2017). Wieczorkowski (2019) defined barriers as specific action, occurrence, or situation that prevents open data from being used effectively. Crusoe and Melin (2018) link the ability to use OGD with the barriers. Crusoe and Melin (2018) refer to different barriers such as data findability, diversity and usability, where the users lose the ability to use the data when they face a combination of the barriers. Benitez-Paez et al. (2018) argued that the provisioning of the data is not necessarily translated into value and reuse unless more has to be done by the publishers and the intermediaries to overcome barriers to using the data and make it more usable. We define barriers to using OGD according to González-Zapata and Heeks (2015), as a lack of some or several crucial resources and competencies required to translate OGD into positive activities and outcomes: an absence of all data, sound data, the technology needed to access ICT-based data, the abilities to change OGD, and the knowledge required to comprehend OGD.

1.2.4. Open data intermediaries

Within the OGD ecosystem, data intermediaries, especially Non-Profit Organizations (NPOs), have been playing a role in making OGD more usable for OGD users. Intermediaries are users of OGD who process the raw data to provide aggregated/processed data services or tools for other OGD users to utilise the data (Zuiderwijk et al., 2014). Some OGD users may prefer to use the processed data and the data services and tools derived from these raw data. In other words, open data intermediaries are participants in the open data ecosystem who link other participants, such as open data producers and data consumers, to one another (Haan, 2018). Furthermore, González-Zapata and Heeks (2015) defined data intermediaries as all parties who support OGD activities by bridging the gap between civil society data consumers and public sector data producers. The two definitions are broad and do not clearly define open data intermediaries and their roles. In addition, da Silva Craveiro and Albano (2015) define intermediaries as actors who work with publicly available data that has been made openly available, whether they are individuals or representatives of governments and civil society groups. da Silva Craveiro and Albano (2015) argued that intermediaries might or might not employ technical or structural artefacts in their operations. By utilising open data, intermediaries increase the value of the data to make it easier to understand and hence more valuable for third parties following their involvement. Therefore, in this study, we will adopt the definition of Van Schalkwyk et al. (2016) as it is more detailed and specifies the roles of the data intermediaries in the OGD ecosystem. They define open data intermediaries as "an agent (i) positioned at some point in a data supply chain that incorporates an open dataset, (ii) positioned between two agents in the supply chain, and (iii) facilitates the use of open data that may otherwise not have been the case" (Van Schalkwyk et al., 2016, p. 4). As depicted in 1.2, we visualise the open data intermediaries as an agent or a group of individuals in between the data providers and the user with a bidirectional line between the two sides. Since our study focuses on the Non-profit data intermediaries hereinafter referred to as (NPDI), we

add to the definition of Van Schalkwyk et al. (2016) " with no commercial interest". In the OGD ecosystem, NPDIs perform several activities for OGD users who are willing to use or already using OGD for various reasons to help reduce the barriers to reusing OGD. NPDIs can assume different roles, such as data demanders, producers, validators, developers, communicators, educators and aggregators (Pilshchikova et al., 2022).

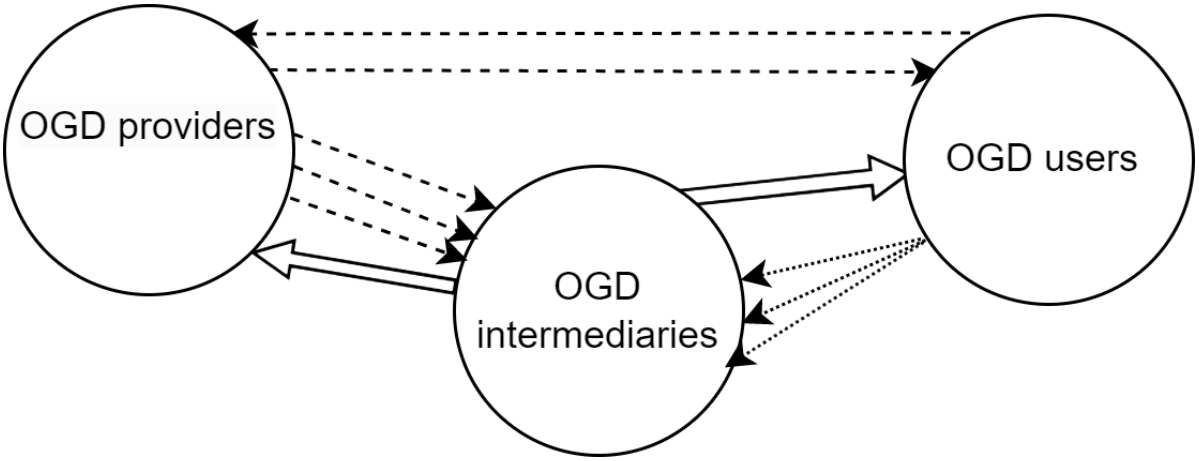

Figure 1.2: Open data intermediaries

1.3. Knowledge gap

To identify the knowledge gap, a systematic literature review was conducted to determine what is known in the literature about the impact of NPOs' activities towards lowering the barriers to using OGD. The methodology of the literature review will be discussed in 2.2.1.

1.3.1. Literature review results and discussion

The literature review revealed several barriers to using OGD and data intermediaries' activities. The anticipated value of OGD, such as producing products and services, opening new economic opportunities, or enhancing public services, will not necessarily be realised by only opening new data (Crusoe & Melin, 2018). Once data is shared, data availability does not guarantee accessibility and variety in data consumption (Erete et al., 2016). This is occurring because of the existence of a variety of barriers, such as the absence of user-friendly portals or interfaces where users can easily navigate and understand available OGD (Brugger et al., 2016; Chattapadhyay, 2014b; da Silva Craveiro & Albano, 2015; Magalhaes et al., 2013; Saxena & Muhammad, 2017, 2018). The published data can be fragmented, incomplete or duplicated in multiple sources (Brugger et al., 2016; da Silva Craveiro & Albano, 2015; Saxena & Muhammad, 2017, 2018) . Magalhaes et al. (2013) also argued that the more data is added to ODG portals, the more overwhelming it becomes for users to find the data and extract useful information. The lack of technical capabilities and infrastructure or statistical skills of the users (Brugger et al., 2016; da Silva Craveiro & Albano, 2015; González-Zapata

& Heeks, 2015; Saxena & Muhammad, 2018). The different data formats; some users require the data in a machine-readable format, while others might need the data in a report or visual format. The absence of meta-data that serves the specific users' needs (Chattapadhyay, 2014b; da Silva Craveiro & Albano, 2015)

In light of the barriers to using OGD, intermediaries, who can potentially lower the barriers to using OGD for users, can assume one or more of five roles as mentioned by (Brugger et al., 2016) and originated by (González-Zapata & Heeks, 2015). Intermediaries can be a) demanders: carry out or consolidate civil society requests, ask for the release of specific statistics, or even advocate for the creation of specific OGD-relevant regulations; b) producers: combining data with already-existing datasets to produce new open government data, c) validators: deciding which datasets or portions of datasets are helpful, who may perform dataset selection, filtering, or evaluation, d) developers: this role extends to the creation of information through the activities of intermediaries to improve the usefulness and accessibility of OGD, e) communicators: that alter OGD and deliver it in formats, channels, and situations that are most suited to the skills and knowledge of ordinary civil society members. It is essential to acknowledge that the different roles of OGD intermediaries can influence barrier reduction. The first three roles are linked with reducing barriers to producing more OGD, and the latter two relate to information creation and use (González-Zapata & Heeks, 2015). Yoon et al. (2018) argued that the primary value of data intermediaries is helping societies make better decisions with data, especially in tackling pressing issues and enhancing the daily lives of community members. For effective use of OGD, data intermediary groups perform four essential tasks: (1) data democratising; (2) enriching the already-existing data; (3) strengthening societies' data awareness and knowledge; and (4) increasing societies' data capability (Yoon et al., 2018).

1.3.2. Identifying the knowledge gap

The studies that have been reviewed have not investigated the impact of the activities of open data intermediaries, specifically non-profit organisations (NPOs), and their impact, particularly on reducing barriers to the use of Open Government Data (OGD). For instance, these activities include connecting multiple parties and facilitating data flow (Yoon et al., 2018), providing resources and skills necessary for creating and transforming OGD (González-Zapata & Heeks, 2015), and offering support and training through manuals and guides(Brugger et al., 2016; Yoon et al., 2018). Therefore, this study aims to explore the impact of the NPDIs' activities, particularly their effect on reducing barriers to OGD use. The main research question of this study is: *RQ: How do European non-profit data intermediaries affect the barriers to using OGD?*

1.4. Social relevance

The expected transformation of OGD is not yet clearly visible. In general, it has been observed that the government objectives behind OGD initiatives are not realised when it comes to the actual implementation of the OGD initiatives in terms of the use of the open data that are released (Davison, 2006). Although OGD is becoming more available, the availability is not sufficient on its own; something must be done with the released data. Only when the data are utilised or turned into something else, such as new insights, new products or services, the value is realised (Janssen et al., 2012). Janssen et al. (2012) also suggested that some sort of support must be provided to OGD users to guarantee that OGD is not only published but used and converted into valuable outcomes. Perhaps, one of the reasons preventing OGD's transformative impact from occurring is the barriers to using OGD. OGD barriers can be divided into 1) Provision barriers, related to the release of datasets by the data providers, and 2) barriers to using related to the use of OGD by different user groups (Martin, 2014). Benitez-Paez et al. (2018) identified six different categories of barriers, 1) usability barriers, such as data is difficult to understand; 2) accessibility barriers, such as issues with the data format; 3) Data quality barriers, such as the absence of row data or metadata; 4) technical barriers such as no API documentation or no advance search functions; 5) legal and policy barriers such as the licence of the data is not clear; 6) currency barriers where the data is not up to date.

We assume that the barriers that are most likely to hinder the potential of OGD are linked to the barriers to using OGD. Once the data is released, and how the data is released could be the start of the users' struggle. Furthermore, the barriers to using might be more influenced by the activities provided by the NPDIs. This can be the case since the NPDIs are not in a hierarchical or structural position where they can order public agencies. Instead, they try to work with the available data and brigade the gaps between the data providers and users. Given the anticipated transformation potential of OGD and the current barriers to reusing OGD, it is necessary to investigate how the existence and activities of the NPDIs contributed to reducing the barriers to using OGD. Hence, this project will focus on exploring the impact of the activities provided by the NPDIs on lowering OGD's use barriers for its users.

1.5. Academic relevance

Several important factors contribute to the academic value of studying the effects of NPDIs in reducing the barriers to using OGD. Filling the knowledge gap in the academic literature, as discussed by Pilshchikova et al. (2022), by exploring the impact of NPDIs' roles and activities in reducing the barriers to using OGD. It will also provide an opportunity to explore their strategies and approaches to promoting the use of OGD and provide evidence-based insights about NPDIs' contribution to reducing the barrier. These insights might inform key actors in the OGD ecosystem to take informed decisions and develop effective OGD use, promotion and adoption

strategies. Furthermore, researchers can provide recommendations to enhance the OGD ecosystem driven by the insights that can contribute to developing a more robust open data ecosystem. Overall, exploring the NPDIs' impact in reducing the barriers to using OGD can contribute to the academic understanding of the OGD ecosystem and provide practical guidance to support more evidence-based decision-making in the OGD domain.

1.6. The CoSEM perspective

NPDIs exist within the OGD ecosystem, and both are considered socio-technical systems (STS). STS refers to systems composed of hardware, software, regulations, procedures and also human factors interconnected and entwined (Fischer & Herrmann, 2011). NPDIs and OGD ecosystems are a perfect manifestation of such a socio-technical nature, as the characteristics of STS are apparent. Our explorative study will acknowledge these aspects by taking a holistic approach when analysing the impact of NPDIs in reducing the barriers to using OGD. We will explore technical, institutional and process-related impacts that the NPDIs are contributing to. Moreover, OGD also crosses various domains, including information technology for technical issues; institutional, where the obligation to share public data openly is set; politics for the information it covers or discloses; and economic expectations for the value it will produce and the commercial interest that might attract companies to use the data for commercial proposes. This will provide the ability to analyse sub-systems with different lenses to better understand the system of interest, which is the NPDIs and their interaction within the OGD ecosystem.

Moreover, the STS aspects of the NPDIs also make NPDIs complex adaptive systems (CAS), which are a dynamic network of interaction consisting of many forces representing governments, corporations or people (Waldrop, 1993). Specific aspects further define CAS, such as the parallel interaction and the reaction of the other players depending on the other player's actions, the decentralised control and competition and cooperation between players, and the overall system behaviour is the result of the decisions and actions of many individual players (Waldrop, 1993). Both STS and CAS concepts are relevant to exploring the impact of NPDIs in the OGD ecosystem. Also, they are CoSEM concepts that are taught and incorporated in most of the CoSEM courses. These concepts and theories will illuminate us and broaden our perspective when analysing NPDIs' impact on reducing the barriers to using OGD. Furthermore, in this research, we have used some of the tools taught in the CoSEM research challenge, such as conducting a systematic literature review to identify the barriers to using OGD.

1.7. Research objectives

The goal of this exploratory study is to explore how NPDIs' activities reduce the barriers to using OGD. The activities of five different NPDIs in Europe will be investigated using a qualitative approach, more specifically, a multiple holistic case study. Constructivism

will be the research paradigm, with a focus on the participants' perspectives on the problem being studied. In order to understand how the activities have (or have not) contributed to reducing the barriers to using OGD, the study will interview representatives from several NPOs and their users. Four sub-research questions are formulated and discussed in the following chapter.

2
Research approach

This chapter discusses the research approach and the different methods applied in answering the research questions.

2.1. Objectives and research questions

This study is an explorative study that aims to explore the impact of the activities of NPDIs on the barriers to using OGD. Qualitative research is a way to investigate and comprehend the significance that different people or groups place on a social or human issue (Creswell, 2009). According to Creswell (2009), case studies entail an in-depth investigation of a project, event, activity, process, or one or more people by the researcher. Hence, the study adopted a case study in a qualitative approach of multiple NPDIs located in Europe. In alignment with the explorative nature of this study, multiple NPDIs will be studied; therefore, the study will follow a multiple holistic case study design, as it will include more than one case with one unit of analysis, which is the NPDIs, to explore the impact of their activities on barriers to using OGD. The holistic approach is followed as it facilitates a comprehensive contextual understanding of the phenomenon, considering its interconnectedness. Yin (2018) argued that having two (or more) cases may have significant analytical advantages. In this study, five different NPDIs will be included. The data from numerous cases are frequently thought to be more convincing. As a result, the multiple-case study is perceived as a more solid approach. (Herriott & Firestone, 1983).

Little is known in the literature about the effect of NPDIs activities on the barriers to using OGD. Therefore, from a constructivist research paradigm, this study depends as much as possible on the participants' opinions on the issue under investigation (Creswell, 2009). Interviews with individuals from several NPDIs and their users were conducted to understand how the activities have(not) contributed to reducing the barriers to using OGD. Interviews are a popular and widely used way of gathering data in qualitative research; therefore, the researcher is positioned as a co-creator of knowledge alongside participants in their quest for understanding and significance (Harrison et al., 2017).

To answer the main research question, the following sub-questions were introduced to contribute to understanding the impact of the NPDIs' activities on reducing the barriers to using OGD. First, the barriers to using OGD and the activities performed by the data intermediaries to reduce the barriers to using OGD were identified. SQ1

was formulated: *SQ1: What are the barriers to OGD use?* Answering this question revealed the barriers to using OGD. Moreover, open data intermediaries participate in various activities to reduce the barriers to using OGD. To identify the open data intermediaries' activities, SQ2 was formulated: *SQ2: What activities are performed by NPDIs?* The finding of sub-questions one and two provided a theoretical background of the current barriers and the various activities of the data intermediaries. Zooming into our research focus, SQ3 was formulated *SQ3: What activities are performed by the selected NPDIs?* to analyze the activities of the selected.

Subsequently, SQ4 was formulated: *SQ4: How do NPDIs and their users perceive NPDIs to reduce the barriers to using OGD?* To understand the impact of the NPDIs' activities on the barriers to using OGD, semi-structured interviews were conducted with key individuals of five NPDIs as well as four of their users to capture how the users and the NPDIs perceive the impact of the NPDIs activities on reducing the barriers to using OGD. Lastly, SQ5 was formulated: *SQ5: How does the participants' feedback support the findings from the interview about the impact of the NPDIs in reducing the barriers to using OGD?* to validate the finding and the interpretation of the interview data, a focus-group session was conducted to validate the outcomes with three participants. The research flow diagram is depicted in Figure 2.1

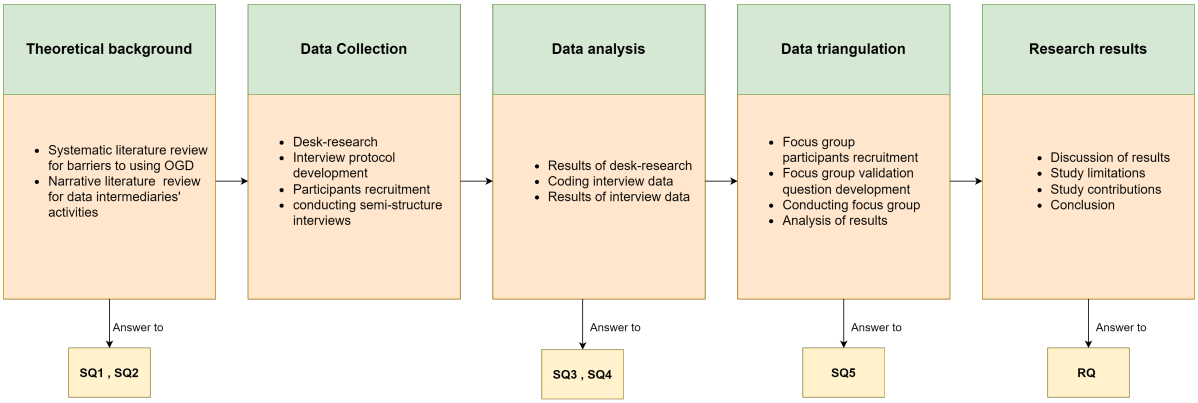

Figure 2.1: Research flow diagram

2.2. Methods of data collection

Multiple research questions were answered using a variety of data collection methods and a combination of different methods. This section discusses the methods used for answering each sub-research question.

2.2.1. Identification of the knowledge gap

As shown in Figure A.1, the search string included terms about Open Government Data, intermediaries, usage, and barriers and these terms were identified as the core concepts to study the literature. Two databases were searched, Scopus and Web of Science. Synonyms and different spelling of key terms were included in the search string to increase the chance of finding relevant studies. After removing duplicates, a

total of twenty-nine articles were initially identified to be relevant. Through reading the article's title, abstract and keyword, fifteen articles were excluded. Fourteen articles were assessed for eligibility; the criteria for inclusion comprised articles written in English, case studies that included interviews were highly desired, literature reviews articles were eliminated, and the article's content focuses on either barrier to using OGD for users, what intermediaries are doing or can do to reduce the barriers to using or intermediaries' activities in OGD ecosystem. Using forward and backward search, three additional articles were included, resulting in ten articles being included in the review; for an overview of selected studies, see TableA.1.

2.2.2. SQ1: What are the barriers to OGD use?

Based on the initial literature review conducted to identify the knowledge gap, the barriers to using OGD were discussed in detail in the literature. The knowledge about the barriers to using OGD may be known in literature but is fragmented. Therefore, a systematic literature review was conducted to identify the barriers hindering users from using OGD. Furthermore, the literature review findings were synthesised to establish and discuss the categorisations of the barriers to using OGD.

2.2.3. SQ2: What activities are performed by NPDIs?

To have an overview of the various activities performed by the data intermediaries, a narrative literature review was also conducted to identify the activities and roles of data intermediaries concerning the OGD ecosystem. In the narrative literature review, we aimed to identify and present the roles and activities of the open data intermediaries. SQ1 and SQ2 provided a theoretical background of the phenomena under investigation. Thus, the theoretical concepts guided us in analysing the selected NPDIS for our case study and informed us about the activities and barriers to using OGD while developing the interview protocol for our interviews.

2.2.4. SQ3: What activities are performed by the selected NPDIs?

After identifying the barriers to using OGD and the roles of the data intermediaries. Furthermore, to get a further zoomed-in detailed perspective of the NPDIs that will be part of the study, the selected NPDIs were researched in depth before conducting the interviews using desk research methods where their websites, relevant reports or other applicable documents were analysed to identify the core activities offered by the NPDIs. Furthermore, during the interviews, the different activities of the NPDIs were discussed to ensure the completeness and up-to-date of the gathered activities. The answer to the third sub-question contributed to helping understand the activities of NPDIs, allowing for an overview of the different NPDIs' activities and deriving a more in-depth analysis based on the activities, scope, location, and mission of the analysed NPDIs. Employing desk research methods might lead to outdated, biased or low-quality findings about the activities and the roles of the NPDIs. During the interviews, the participants were asked questions about the activities of NPDIs. Asking

the participants to reflect on their activities was aimed at addressing this limitation by ensuring the gathered activities of the NPDIs were inclusive and up-to-date.

2.2.5. SQ4: How do NPDIs and their users perceive NPDIs to reduce the barriers to using OGD?

Once barriers to using OGD and activities performed by NPDIs were identified, interviews with selected NPDIs were conducted. The case study comprised a list of five NPDIs selected using a combination of purposeful sampling strategies. Purposeful sampling aims to identify information-rich cases whose investigation will shed light on the research issues (Patton, 1990). Therefore, the strategies that will be used are intensity and criterion sampling. An intensity sample comprises cases rich in information and intense manifestations of the phenomena of interest (Patton, 1990), that is, the NPDIs. Criterion sampling entails selecting examples that satisfy a predetermined criterion (Creswell, 2006), such as geographical location, data intermediaries' scope and role/s the NPDIs play. The selection criteria are also introduced to ensure diversity in the selected cases and reduce representation and selection bias.

The data was collected through semi-structured interviews with NPDIs' subject matter experts and some users of NPDIs services or tools. Semi-structured interviews give an in-depth insight into people's experiences and viewpoints. Therefore they are a useful data collection technique. They enable more freedom in the interview process and allow the discussion of unexpected subjects during the interview (Strauss & Corbin, 1998). Before conducting the interviews, the interview protocol and questions were developed. Several NPDIs were requested to participate in the interviews. The study only included English-speaking participants because of the limitations of the researcher's language and to avoid the need to use automatic translation interviews, which might consist of some errors using automated translators. The case study focused on qualitative data on how the different activities contribute to reducing the barriers to using OGD users.

2.2.5.1. Sample and selection

In this case study, we adopted a multiple holistic case study design to strengthen the findings' validity. Different NPDIs were selected to be studied based on multiple criteria. These considered selection criteria will ensure that the study concentrates on relevant cases that are consistent with the study's objectives. We can be more confident that the study's findings will be valid and reliable by choosing cases that meet the selection criteria. Moreover, in multiple case studies, more than the criteria for selection is required, that is, to select a replication logic (Yin, 2018). According to Yin (2018), each case must be chosen carefully to predict similar (literal replication) or contrasting outcomes (theoretical replication). In alignment with the explorative nature of this study, the theoretical replication logic will be followed by aiming for diversity in the case selection. However, some criteria were added to ensure the cases are also coherent. The cases were selected based on satisfying to following criteria:

1. NPDIs that are actively working and involved in European countries. This criterion is essential since it allows for more consistency in the legal and regulatory systems among the selected cases. Selecting cases that are actively involved in and working in European countries could reduce the differences in aspects such as legal systems, social norms, and political environments. As a result, the cases can be more comparable, making it more straightforward to identify trends and come to conclusions. In addition, this criterion ensures that the examples are relevant to the European context, which is the study's main location of focus.
2. The case focuses on or uses data that are open and produced by publicly funded entities. This criterion is intended to ensure that the selected cases use data that are aligned with our definition of OGD, which refers to OGD as interoperable data that public institutions share publicly for anybody to use and redistribute without restriction. We can ensure that the data being utilized by the selected cases are open, produced by publicly funded institutions and accessible to everyone, which are crucial aspects of OGD.
3. The cases are non-profit-oriented. This criterion is important because the study focuses on the impact of the NPDIs. Selecting cases that are non-profit-oriented will ensure that the organisation are not having commercial interest involved. This criterion also was introduced to add focus on the selection as there are some open data intermediaries that deal with both OGD and non-OGD but open data.
4. Cases are diverse and fit our identified cases diversity dimension. This criterion is necessary to ensure that the study includes a wide range of cases that can accurately represent the wide range of NPDIs. By selecting diverse cases in different dimensions, such as country, scope and activities, we can ensure that the study captures the wide range of experiences, characteristics and impacts of the different NPDIs. This will lead to a comprehensive analysis of the impact of the NPDIs and facilitate identifying differences and similarities among the different NPDIs. This selection criterion requires the selection of multiple cases that are considered diverse (Seawright & Gerring, 2008). Furthermore, the diversity criteria is aligned with the exploitative nature of our research (Seawright & Gerring, 2008). The following diversity dimensions were introduced to be used when selecting the cases:

   (a) Geographical locations, despite the cases being located in Europe, cases were spread over different countries.
   (b) The scope of the cases was considered during the selection; cases were selected based on the diverse scope, including the mission statement and the field where the cases are active.
   (c) The identified roles of the NPDIs influenced the selection. Cases were selected to ensure coverage of most of the roles of the data intermediaries.

2.2.5.2. Coding process

The coding process was applied to the transcripts obtained through the interviews with the nine participants of the case study. The transcripts were obtained using MS Teams; after that, the transcripts were edited and shared with each participant before analysis for their review before the analysis stage. The coding was conducted solely by the thesis author using the CAQDAS ATLAS.ti 23.

There are numerous approaches to qualitative data analysis; the guidelines of Skjott Linneberg and Korsgaard (2019) were followed to execute the coding process as they are specifically made for inexperienced coders. First, prepare for the coding process by examining the research question and objectives, studying the literature on the research topic and preparing the documents that will be used for the coding process (Skjott Linneberg & Korsgaard, 2019). Second, choosing the codding approach, either inductive, deductive or a blended approach (Skjott Linneberg & Korsgaard, 2019). Inductive coding refers to generating codes from the data using phrases or terms used by participants rather than theoretical concepts or vocabularies(Skjott Linneberg & Korsgaard, 2019).In the inductive approach, the researcher will be making plenty of codes, frequently very specific and limited ones, which is beneficial for capturing the complexity and diversity of the data (Skjott Linneberg & Korsgaard, 2019); however, as the number of codes increases, the process may become less workable and consume more time(Gehman et al., 2017), or become too complicated and less focused (Skjott Linneberg & Korsgaard, 2019). The deductive approach implies having a predefined list of codes derived from the literature and created before starting the coding process (Miles et al., 2013). The deductive approach is helpful when the research aim is to test theory (Skjott Linneberg & Korsgaard, 2019) or analytically generalize across cases (Rowley, 2002). The blended approach implies a mix of inductive and deductive approaches, suggesting a cycle between theory and evidence. By doing this, the researcher stays attuned to the current theories and concepts while still being open to unexpected findings in the data (Skjott Linneberg & Korsgaard, 2019).

Next is to execute the code cycles, which usually encompass two or more coding cycles (Skjott Linneberg & Korsgaard, 2019). According to Skjott Linneberg and Korsgaard (2019), first cycle coding can be descriptive, referring to assigning a segment of the text a code capturing what it is about, or attribute coding entails assigning codes that help to structure the data for further analysis such as age, gender or nationality, where the codes at the first cycle are used to create an overview of the data to facilitate further analysis in later cycles. The second cycle coding is more comprehensive as it includes more analytical steps such as classification, synthesis, abstraction or conceptualisation of the data(Saldaña, 2021).

In this study, we adopted a blended approach with a mix of inductive and deductive approaches. The blended approach was followed to have a manageable approach provided the time and the resources for this project as well by the combination of de-

ductive approach to reduce the coding complexity and inductive approach to discover meanings and patterns in our data. We started our coding process by examining our research question, reviewing the state of the art about the barriers to using OGD and the activities performed by the data intermediaries to provide a theoretical background for the research and to influence the interview protocols and our coding process. Moreover, to follow a time bounded approach providing the time and resources available for this master thesis project, we followed the blended approach in two cycles to execute the coding with manageable yet reasonable complexity. In the first coding cycle, we followed a deductive approach based on a pre-defined list of initial codes, including the barriers identified in our literature review in Table 3.2 and the list of actives conducted by the NPDIs in Table 5.2. This coding cycle aimed to label the activities of the NPDIs in the transcripts and the impact the NPDIs create in reducing the barriers we identified in Table 3.2. Then, we conducted a second cycle of inductive coding, where the codes were derived based on the discussion with the participants. This cycle was beneficial and necessary to uncover meaning and codes from the discussion with our participants and not only to limit our analysis to the lists we derived from the literature.

After reviewing the codes of the second cycle, some of the codes were overlapping or too narrow, which yielded the need for further aggregation and integration. Therefore, we conducted a thematic analysis of our data to integrate and conceptualize our data to meaningful findings. Thematic analysis is a technique for finding, analyzing, and reporting patterns (themes) within data by organizing and describing them in detail. Moreover, going further and interpreting the different aspects of the data concerning the research question (Boyatzis, 1998). We followed the guidelines of Braun and Clarke (2006) to implement the thematic analysis into our data. They proposed a six phases process of conducting thematic analysis, phases one and two, which consist of data familiarisation and reading, and the generation of initial codes was already covered in our two-cycle coding. The third phase is about searching for themes that encompass re-focusing the analysis on the themes rather than the codes (Braun & Clarke, 2006). A theme suggests a certain degree of structured responses or meaning within the data set; it also captures a key aspect of the data in connection to the research question (Braun & Clarke, 2006). This phase involved the beginning of the thematic analysis by examining the different codes and finding similarities to combine these codes under a higher level of aggregation of the same subject. This phase ended with a list of initial themes and the relevant codes and quotes organised for the next phase. Fourth, reviewing the themes phase, which encompasses two levels of reviews according to Braun and Clarke (2006) guidelines; first-level involves reviewing the coded data extract under each theme to ensure a coherent pattern of each theme extract. At this review level, some codes were reassigned to more appropriate themes; some codes were merged. For instance, before this review, we had two codes for legal awareness and legal advisors; the data extracts under the legal awareness code were similar and more logical to move under the legal advisor code. Another example is when we had different codes around OGD quality, such as "updating the data" and "validating the data", The codes

were merged under improving OGD quality. Also, one of our themes was removed; the theme was about improving OGD user experience because the extracts under the theme were found to be relevant to other themes. For example, the codes about the "structured approach", "enabling self-organizations", and "best practices" were moved under the OGD process optimization, which resulted in disregarding the theme of improving OGD users' experience. Once we were satisfied with the first level review, we conducted the second-level review that involved reviewing the validity of the themes regarding the entire datasets(Braun & Clarke, 2006). At the end of this phase, we had a clear idea of our themes and their coherence and relation to the entire dataset. Next, the fifth phase of defining and naming the themes encompasses further analysis of each identified theme, capturing its essence, and identifying the interesting aspect for further interpretations about each of our themes (Braun & Clarke, 2006). The sixth phases conclude the process by having the final analysis and the write-up of the thematic analysis into this report; in other words, it involved narrating our complicated story from our data into a convincing story(Braun & Clarke, 2006). The codebook that included the final themes and sub-codes can be found at https://data.4tu.nl/private_datasets/IpXvdNOhF7sLrVixPYlrYor2t_LTuQA83DUU8EVmCBw.

2.2.6. SQ5:How does the participants' feedback support the findings from the interview about the impact of the NPDIs in reducing the barriers to using OGD?

We conducted a focus group with participants who had already been interviewed after analysing the case study data in an effort to triangulate our analysis. Focus group is a method that uses in-depth group interviews with participants that are chosen purposefully (Rabiee, 2004), in our case, our interview participants. However, the sample does not necessarily represent the specific population or group in the research scope (Rabiee, 2004). The focus group allowed us to triangulate our results by gathering three participants in one session, unlike interviews where we would have required three separate interviews(Krueger, 2014). Moreover, the focus group allowed us to triangulate our data we have gathered about how NPDIs affect the barriers to using OGD, using multiple sources to confirm our findings. Focus groups are an effective method that can be used to clarify research findings (Nyumba et al., 2018). We designed our Focus group following the guidelines of Nyumba et al. (2018). We conducted one focus group session with three participants; at the beginning, we shared our previous interview transcripts analysis briefly, and then we prepared discussion slides by asking open-ended questions to allow our participants to reflect on each slide by agreeing or disagreeing about the themes we derived from the interview transcripts.

2.3. Quality of the research design

A study design can be evaluated based on specific logical tests since it is intended to represent a logical collection of statements; the majority of social research has relied on four tests for assessing the research quality (Yin, 2018). In this study, we have

considered the relevant tests of the four quality tests as follows:

Construct validity: According to Yin (2018), this test is particularly challenging for case study researchers to develop objective judgments. One of the steps Yin (2018) suggested to ensure the construct validity was to use multiple sources of evidence. The interviews with the NPDIs and the users were our primary source to explore the impact of the activities of the NPDIs. Furthermore, due to the lack of resources and focus on the impact of NPDIs, we could not assess the impact based on different data sources other than the interview. However, in the interviews, we ensured that we covered different perspectives of the NPDIs and their users, and we combined interviews with desk research to have multiple sources of evidence. Furthermore, we conducted a focus group with three participants that we already interviewed to validate our interruption. The validation through the focus group ensured the objectivity of our analysis and confirmed the validity of our interpretation of the interview data.

Internal validity: This test is, in particular, crucial for explanatory cases Yin (2018). Therefore, following the recommended tactics by Yin (2018) in our explorative study is not possible.

External validity: This test concerns the generalizability of the case study's findings (Yin, 2018). The tactic proposed by Yin (2018) to increase the external validity for multiple case study designs is to follow a replication logic. In the selection criteria for this explorative case study, we have also followed a theoretical replication logic that justifies the case selection diversity and ensures the external validity of the research design.

Reliability: This test proves that a study's processes, such as its data-gathering methods, can be repeated and provide the same outcomes; also, the test aims to reduce errors and biases in the study (Yin, 2018). The interview protocol was developed and used based on the tactic proposed by Yin (2018). Chain of evinces was also maintained as another tactic suggested by Yin (2018) where the interview recordings, transcripts and other data collected for further analysis were securely saved to allow for the reproducibility of the research.

3
Literature review

In this chapter, two literature reviews were conducted to identify the barriers to using OGD and activities conducted by the data intermediaries about the barriers to using OGD.*SQ1: What are the barriers to OGD use?* and *SQ2: What activities are performed by NPDIs?* will be answered in this chapter. First, the systematic literature review of the barriers to using OGD will be discussed. Then, the second narrative literature review about the roles and activities will be discussed. A conceptualisation of the barriers to using OGD and the activities of the NPDIs will be presented at the end of this chapter.

3.1. Barriers to using OGD

In this chapter, first, we aim to identify and understand the barriers faced by different users of OGD, such as researchers, decision-makers, and the general public, while attempting to utilize and disseminate open government data effectively. Open government data can potentially increase accountability, openness and trust in governmental institutions, and several direct and indirect economic benefits from reusing OGD can be harvested ("The benefits and value of open data | data.europa.eu", n.d.). However, using open government data effectively by a wide range of users still faces several barriers (Nikiforova & McBride, 2021; Saxena & Muhammad, 2017; Zuiderwijk & de Reuver, 2021). Therefore, understanding the barriers to using OGD is crucial because it may highlight particular issues requiring attention. Understanding these obstacles will make conducting the interviews for this study straightforward, more relevant and more effective as it will provide more information into the barriers and the context of the ecosystem of OGD. Moreover, a systematic literature review offers a thorough overview of the body of literature on a given subject. A comprehensive literature review can contribute to a more thorough knowledge of a particular phenomenon or problem by combining the results of several pieces of research (Tranfield et al., 2003).

3.1.1. Literature search strategy

As depicted in Figure 3.1, a systematic literature review to identify the barriers to using OGD was done by searching Scopus and Web of Science databases. The search string was developed using three main concepts, Open Government Data, data use, and barriers. Synonyms of the words "use" and "barriers" were introduced to ensure the coverage of a wide range of studies. Each database was queried with the exact search string, and the search was limited to the last five years to ensure researching the most up-to-date studies. Moreover, the search string was used to search the studies' titles, abstracts or keywords. The search was limited to OGD instead of including open data,

as the search for open data as a synonym of OGD resulted in an excessive number of results (nearly 2000 studies). Scopus search resulted in 140, and Web of Science yielded 135 results. Both results were merged, duplicates were removed, which resulted in the identification of 201 that were identified at this stage. Through screening the studies by reading the abstract, introduction discussion and conclusion, 164 studies were eliminated where the studies were not focusing on or discussing barriers or challenges related to OGD. This resulted in 37 studies being eligible for full-text reading. After full-text reading and applying the second criteria for inclusion, where the studies focused on the barriers to using OGD, 16 studies were included in this systematic literature review as shown in Table 3.1.

Table 3.1: Overview of selected studies for OGD barriers literature review

Study Title	Author/s and Year	Geographical scope	Scope
Open Government Data as an Innovation Process: Lessons from a Living Lab Experiment	Ruijer and Meijer (2020)	Netherlands	OGD Usage and Innovation
Barriers to innovating with open government data: Exploring experiences across service phases and user types	Smith and Sandberg (2018)	Sweden	ODG Usage, Transport data
Open government data portal usability: A user-centred usability analysis of 41 open government data portals	Nikiforova and McBride (2021)	Global perspective	OGD portal usability
Roadblocks Hindering the Reuse of Open Geodata in Colombia and Spain: A Data User's Perspective	Benitez-Paez et al. (2018)	Colombia and Spain	Data reuse from a data consumer's perspective
Investigating open government data barriers: A literature review and conceptualization	Crusoe and Melin (2018)	Global perspective	Holistic scope, barriers at all stages
Barriers to use open government data in the private sector and NGOs in Pakistan	Saxena and Muhammad (2017)	Pakistan	Barriers for NGO users

Continued on next page

Table 3.1 – continued from previous page

Study Title	Author/s and Year	Geographical scope	Scope
The impact of impediments on open government data use: Insights from users	Crusoe et al. (2019)	Not specified	OGD barriers faced by university students
Drivers and barriers towards re-using open government data (OGD): a case study of open data initiative in Oman	Saxena (2018a)	Oman	Barriers to using OGD
A rationale for data governance as an approach to tackle recurrent drawbacks in open data portals	Reis et al. (2018)	Global perspective	Barriers to using OGD
Quality assessment framework for open government data Meta-synthesis of qualitative research, 2009-2019	Zhang and Xiao (2020)	Global perspective	Barriers related to OGD quality
Drivers and barriers to re-use Open Government Data (OGD): a case study of open data initiative in Philippines	Saxena (2018b)	Philippines	Barriers to using OGD
Why open government data initiatives fail to achieve their objectives: categorizing and prioritizing barriers through a global survey	Zuiderwijk and de Reuver (2021)	Global perspective	Barriers for sharing and using OGD
Challenges and Supports for Accessing Open Government Datasets Data Guide for Better Open Data Access and Uses	Xiao et al. (2019)	Global perspective	Barriers to using OGD

Continued on next page

Table 3.1 – continued from previous page

Study Title	Author/s and Year	Geographical scope	Scope
Open government data through the lens of universal design	Ferati et al. (2020)	Canada, France and Australia	OGD finding and using
Barriers to using open government data	Wieczorkowski (2019)	Poland	OGD provisioning and use
Toward a user-centred approach to enhance Data discoverability on Open Government Data portals	DAHBI et al. (2019)	Global perspective	OGD Findability

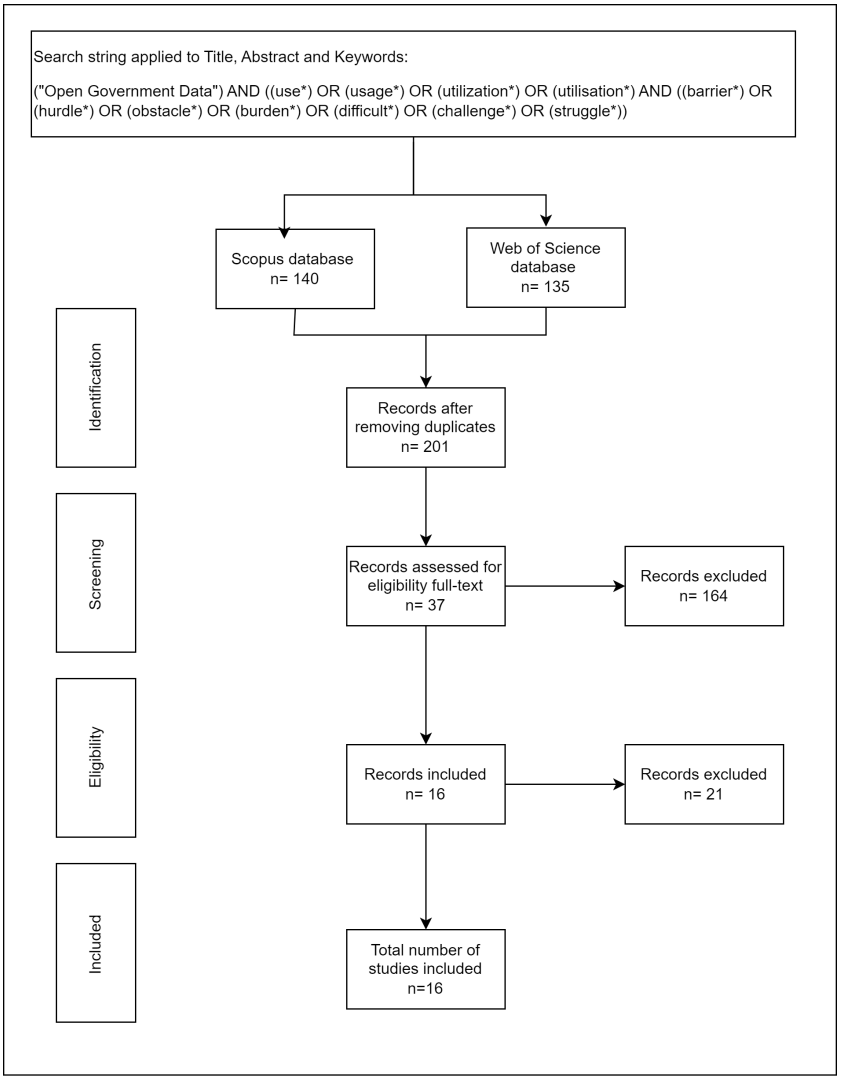

Figure 3.1: Selection process for barriers to using literature review

3.1.2. Results of OGD use barriers

The barriers to using OGD are discussed in literature. Different categorisations of the barriers are used. For instance, Ruijer and Meijer (2020) identified several cultural barriers, such as the inability or lack of understanding to use the data or structural barriers, such as OGD portal unfriendliness or data sets quality. Benitez-Paez et al. (2018) took a broader perspective by identifying seven relevant categories of barriers such as technical, organisational, legal and policy, data quality, financial issues, and cultural and use and participation. Zuiderwijk and de Reuver (2021) suggested to distinguish between barriers affecting open data sharing and barriers facing data users and emphasised that barriers on the user side are more problematic for achieving OGD goals than those on the data providers' side. Hence, in our literature review, we focus on the barriers that hinder OGD data users from using publicly available data, with less focus on the barriers faced by the providers or related to OGD provisioning, in alignment with our assumption that these barriers are more influenced by the NPDIs activities. Therefore, we identified three categories of barriers related to using OGD by the origin of where these barriers have evolved. For the list of identified barriers, please see Table 3.2.

3.1.2.1. Barriers at OGD portals

The first category of barriers facing OGD users is categorized under the OGD portal. In most cases, OGD portals are the first venue for OGD users to interact with the open data. Crusoe et al. (2019) identified four phases of using OGD: motivation, search and evaluation, access and prepare, and aggregate and transform. In this case, the search and preparation phase implies the interaction with the different data sources/portals to identify suitable data sets that serve the motivation of the OGD user. We identified the following list of barriers at OGD portals.

Insufficient tools: For OGD users to find, access and use data, OGD portals or data-sharing techniques to be equipped with or accompanied by analysis and visualisation tools. However, this is not usually the case, (Ruijer & Meijer, 2020; Saxena, 2018b; Saxena & Muhammad, 2017; Zuiderwijk & de Reuver, 2021) discussed that analysis and visualisation tools were missing or limited on some OGD portals. Moreover, some OGD users might need more than superficial or basic analysis and visualisation tools. OGD users might need tools for slightly advanced operations, such as merging different datasets, which are sometimes essential to derive useful information from the data at hand. When missing such tools, it makes using OGD more time-consuming and more challenging (Crusoe & Melin, 2018).

Lack of user-friendly features: Ferati et al. (2020) discussed that one of the challenges faced at OGD portals is related to the lack of some user-friendly features that reduce the efficiency of using these portals. Filtering, search, or dataset previewing were part of the features that some OGD portals required development and improvement, where it does not allow the users to search by keywords or browse the available data

in structured categories (Ferati et al., 2020; Ruijer & Meijer, 2020). Benitez-Paez et al. (2018) have identified usability as one of the categories in their proposed fish-bone diagram for issues leading to low use of OGD. They link user-friendless issues to issues such as lack of reuse examples for the published data. Such issue was witnessed when users were unable to find examples or use cases as a reference to comprehend how to use the data or how the data can be incorporated into other applications.

Issues in finding the data: Findability was one of the most discussed barrier in the reviewed studies. Users had trouble finding the required information, thus, making it challenging to use the data. Benitez-Paez et al. (2018) defined findability as how simple it is for OGD users to discover the information needed. Xiao et al. (2019) have identified findability as the third challenging process among their identified top 5 identified challenging processes to use OGD. Some studies have linked findability issues to not adopting user-centred design techniques and methods when designing OGD portals (DAHBI et al., 2019). It is more difficult to find useful data on OGD portals, according to Crusoe and Melin (2018), when there is a large amount of irrelevant or useless data present. Therefore, it may be challenging to find and utilize the relevant data that is available when there is a large amount of low-quality data. Data findability issues can happen if the infrastructure enabling data provisioning is inadequate. For instance, it may be challenging for users to identify what they need if the data is improperly structured. Therefore, poor infrastructure can cause delays and frustration in finding the required data, Hence, identifying and using data can be more challenging (Crusoe et al., 2019).Zuiderwijk and de Reuver (2021) have identified findability as one of the barriers facing OGD users. Zuiderwijk and de Reuver (2021) have also linked findability to the fragmentation of the different datasets in different portals. Moreover, they recommended that OGD portal developers to develop advanced tools to help users find and use the available data.

Data availability and accessibility: Availability and accessibility of the dataset might slightly overlap with findability. Some authors discussed that accessibility is linked to finding and accessing the data but also to the availability of the data in an online format (Reis et al., 2018). Reis et al. (2018) discussed that one of the barriers to using OGD is related to the OGD portal accessibility, which might occur for various reasons, such as broken links or lack of user awareness. Zhang and Xiao (2020) argue that accessibility and usage of OGD are intertwined; accessing OGD is a fundamental condition to use OGD, and otherwise, it has implications on the use and promotion of the data. Smith and Sandberg (2018) also linked access to data to the use and development of services by various users. Benitez-Paez et al. (2018) discussed and identified in their literature review that data availability is one of the main barriers to using OGD. Furthermore, Zuiderwijk and de Reuver (2021) also connected legislation and access as one of the main barriers to using OGD, and in particular, they identify the category of legislation from the perspective of using open data as a significant barrier.

Data diversity: The diversity of the available datasets was one of the less-discussed barriers. Due to the various roles, problems, and motives of OGD users, their requirements and goals can be diverse (Crusoe & Melin, 2018). For instance, civic monitoring users may need specific indicators and spending details, whereas a transparency advocate may look for patterns and outcomes. Therefore, data on OGD portals must be organized, processed, and presented in a way that fulfils the needs of various users, which may introduce issues in data provisioning. Therefore, OGD portal architectures and tools shall be created to satisfy the various needs of different user groups.

3.1.2.2. Barriers from OGD datasets

The second category of barriers facing OGD users emerges from the available datasets. Crusoe et al. (2019) identified phases of using OGD are relevant in this barriers category. The dataset issues such as its quality and completeness, documentation and metadata, data interoperability and credibility will affect most of the phases of using OGD. The effect on most of the phases is because the data is the central element of the entire use case of OGD. The issues related to the dataset are showstoppers to the users. The magnitude of the issues related to the OGD dataset is evident by the large number of identified barriers related to the dataset category in the reviewed studies, where almost half of the identified barriers were related to the datasets.

Data quality: The highest discussed issue in this category is related to data quality. Sometimes, publicly available government data are incorrect, incomplete, or released with poor quality (Reis et al., 2018; Ruijer & Meijer, 2020; Zhang & Xiao, 2020; Zuiderwijk & de Reuver, 2021). This can happen because of issues in the published data, such as missing data or inconsistent information (Crusoe & Melin, 2018). Using the data and creating helpful conclusions may be challenging if the data is inaccurate or incomplete. To avoid future obstacles and promote more use of OGD, ensuring that the data is of appropriate quality and completeness is crucial. Otherwise, using the data for making decisions could be difficult because it might be incorrect and inaccurate (Benitez-Paez et al., 2018).

According to Nikiforova and McBride (2021), perhaps one of the reasons for the data quality issues is the focus on the supply by the OGD providers. Nikiforova and McBride (2021) discussed that issues with the data quality could arise because of the priority of the data providers, such as government agencies and public institutions, is to develop OGD portals and uploading the data there without investing adequate time and resources in ensuring that the data meets certain quality and completeness standers. Wieczorkowski (2019) also argued that the increase in the amount of published data can lead to a reduction or absence in the data quality. This implies that the quality of the output is only as good as the quality of the input, and the absence of proper quality checks by the data providers will lead to issues in the data quality. The data quality extends beyond the accuracy and integrity of the data; however, it can hinder the users

from being able to use the data if the data is ambiguous or misrepresented, such as using a calendar year instead of a fiscal year in financial data. (Zhang & Xiao, 2020). The data quality is crucial regarding the efficiency and motivation to use OGD. OGD users must invest time and effort before identifying the relevant datasets. Once they discover that the data they selected are not accurate, complete or have quality issues, their time in previous steps might have been wasted (Saxena, 2018a). As a result, it might reduce their motivation to look for another data set or even utilize another OGD dataset.

Lack of documentation: In most cases, allowing access to the dataset is not enough; users need complementary features with the dataset, such as documentation. Clear and concise documentation is essential to help users understand how to access, use, or interpret the dataset correctly. The absence of clear documentation might lead to a struggle in understating the dataset and perhaps can even prevent users from using it (Smith & Sandberg, 2018). Moreover, the documentation might not only be unavailable or incomprehensible, but it might be complex or overly technical for some users, such as the data consisting of technical terms or structured in a complex way. Such issues or limitations in the documentation defeat the primary purpose of the documentation, where it should be a clear guide that explains the how and supports the users in their endeavour to use OGD (Crusoe & Melin, 2018).

Lack of data standardisation and interoperability: Data standardization can be related to the format of the data provided. Depending on how the data will eventually be used, issues in data standardization can have implications on the sustainability of the developed mechanisms to use OGD. For instance, when an Application Programming Interface (API) of a particular OGD provider delivers inconstant data, it becomes less efficient and more challenging to have a consistent data flow from the source to the user (Smith & Sandberg, 2018). Another dimension of interoperability and data standardisation is how the providers gather the data. When the data lacked standardised procedures or practices when it was collected, it will lead to issues in the reliability and accuracy of the data and, eventually, the ability to be reused (Benitez-Paez et al., 2018; DAHBI et al., 2019).

In some cases, the data need to be combined from different sources to generate valuable insights from the publicly available datasets(Crusoe & Melin, 2018). With the issues of interoperability, this step becomes nearly unattainable. This can happen because of the non-unified data format across the different providers, who might not necessarily be related to each other. Furthermore, when the data need to be adjusted, the data must be consistent across the different sources. Sometimes, each provider uses a different classification or vocabulary to describe the data. Zuiderwijk and de Reuver (2021) also discussed and linked the data standardisation and interoperability issues with the compatibility and comparability problems of the different datasets. Moreover, comparability and compatibility issues can be within the dataset or among different

sources.

Data currency: One of the issues related to the publicly available datasets is that it needs to incorporate up-to-date data. The challenge with outdated data occurs when the publishers do not update the published dataset for technical, financial or even bureaucratic reasons. Moreover, it could also be that the data publisher is negligent in updating the published dataset. When the data is outdated, this poses a challenge to OGD users to use the data and derive useful information (Benitez-Paez et al., 2018; Saxena, 2018b).

Furthermore, the issue of timeliness in the data provision poses a challenge for some data users. More than specific pieces of information to be made publicly available are needed, especially when the time aspect related to that information is significant. This can be evident in some cases, for instance, in the health domain, where recently, the information about COVID-19 cases and vaccination progress was published. Such data, and in similar cases in various domains, must reach the users before it spoils and becomes useless (Crusoe & Melin, 2018; Zhang & Xiao, 2020). Outdated data can also happen because of more logical reasons, such as the long embargo period (Crusoe & Melin, 2018) of some classified or privacy-sensitive information. However, the data can not be shared timely with the users for good reasons; this also imposing a challenge on the data users where the data providers need to think of solutions to share the data on time. Moreover, outdated data can impact even the most advanced data users, such as third parties developers, leading them to inaccurate results and eventually impacting the reuse of the available data (Smith & Sandberg, 2018).

Low dataset usability: Not only OGD portals suffer from usability issues, but datasets usability is also an issue that some OGD users face. Not all users have the same technical and infrastructural capability; when data is shared in a large size, this becomes challenging for some users (Ferati et al., 2020). They need additional resources or technical help to download and use the datasets in the first place, which may prevent any future potential of the available datasets. The published dataset loses its usability sometimes when there are no reuse examples of the data (Benitez-Paez et al., 2018); hence, it becomes more challenging for low digital literacy users. The absence of the reuse example does not allow the users to know how the dataset can potentially be used. Even for some users with advanced capabilities, the reuse examples are still necessary to develop integration to read the available data more frequently in an automated manner. In addition, the usability of the dataset could also be linked to the inflexibility of the data's format (Ferati et al., 2020). When the data is provided in multiple formats, it can ensure the data's usability to most users and reduces the challenges or the need for additional steps to convert the provided data into different formats.

Term of use and licence issues: The issues in the terms of use and licensing have two main aspects. First, it is related to legal concerns where the users might need clarification about the terms to use or reuse the data (Benitez-Paez et al., 2018; Crusoe & Melin, 2018). For instance, some data might be related to military operations or intelligence. This can make it difficult for the users to be sure if they are allowed to reuse and disseminate the data and under what conditions. The data providers need to provide clear guidelines on how the data can be reused or distributed, which shall reduce the confusion and misunderstanding some data users have about the terms of use. Second, it is related to the absence of open data licences, which might become a barrier for the data users(Reis et al., 2018). Without the legal mechanism, which is the licence, users might be hesitant to use the data and fear copyright infringement from their use or distribution of the data. On the other hand, when the licence and terms of use are available and clear, users will be encouraged to use the data without the fear of further legal complications.

Absence or low-quality metadata: Similar to other barriers to using OGD, the challenges can have a cascading effect and be interconnected to other barriers. Metadata presence and quality of publicly available OGD are critical. Without it, users might not be able to find, evaluate and use the data (Reis et al., 2018; Saxena, 2018a). DAHBI et al. (2019) argued that poor metadata quality could hinder data discoverability and make it difficult for data reusers. It might also be frustrating and time-wasting when users navigate through datasets without clear information about their content. Crusoe et al. (2019) discussed that once the users are able to find the datasets of interest, metadata are also needed to evaluate the suitability of the datasets. Poor metadata, which can be incomplete, inaccurate or inconsistent, may lead to misunderstanding of the dataset and errors in the subsequent steps in using the data; moreover, it becomes challenging for users to identify the suitability of the identified datasets. Even if the data was discovered and identified as suitable, with the absence of high-quality metadata, it is very challenging to investigate and analyse the available datasets without its metadata (Saxena & Muhammad, 2017).

Data credibility: High credibility and accuracy of published data is critical and can be considered the key to the success of many OGD initiatives (Saxena & Muhammad, 2017). It is essential that the users can trust and rely on the data they use; when the data has credibility or accuracy issues, it could lead to inaccurate conclusions and misguide the decision that was based on the data. When this is the case, users will be less willing to use sources that have had similar issues or when the credibility and accuracy of the data can not be ensured. Lack of data sources tractability was another challenge related to the data credibility (Reis et al., 2018). In some cases, it refers to the need for more information about the origin of the data. The absence of the data source might make it challenging to identify the accuracy and credibility of the data. Without the information about the published data sources, trust might not be easily built, and even when some users use the data, they will be hesitant to widespread it.

3.1.2.3. Barriers related to users' abilities

The third category of barriers is originating from some lack of OGD users' abilities. This does not imply that the issues are happening only because of the lack of the user's ability. However, it can also be related to the level of technical, structural or contextual complexity of the available data. Moreover, it can also be related to the lack of motivation or resources and the lack of awareness of the potential use cases of available data.

Lack of knowledge or capabilities: The barriers to using available OGD due to the user's ability were one of the most discussed barrier types among the three identified categories. It was discussed as an issue related to the user's understandability of the data (Saxena, 2018a, 2018b; Zhang & Xiao, 2020). Some publicly available OGDs may be complex or difficult to understand. This becomes even more apparent when the data users lack some technical background. Data ambiguity and misrepresentation can also cause an issue when less tech-savvy users try to use the data.

Xiao et al. (2019) discussed that barriers regarding the user's ability are seen as challenges in inferring deeper information from the available dataset. Such barriers are also manifested in a way that makes the task of using OGD complex. For instance, when users are willing to develop some simple or advanced digital service that utilizes OGD or even use the available dataset for primary use cases, they are required to have a wide range of technical skills to achieve these objectives (Smith & Sandberg, 2018). For instance, users are expected to know how to use some specialized analytical software in order to be able to analyze the dataset or even have a simple overview of the data (Ferati et al., 2020). Crusoe et al. (2019) discussed that the barriers emerging from the users' ability are seen as an extra burden by the users where they must exert efforts in learning new skills before being able to use the data.

Ruijer and Meijer (2020) and Zuiderwijk and de Reuver (2021) have identified and discussed the barrier related to the users' ability as cultural or inclusiveness barriers. When the data can not be used or understood due to lack from the users' side, this can be seen as a digital divide (Ruijer & Meijer, 2020). This means that the users' ability barrier will lead to the exclusion of some users group who are technically less capable; therefore, the chance to use the data will not be equally given to the different user groups of OGD (Zuiderwijk & de Reuver, 2021).

Lack of motivation and resources: Smith and Sandberg (2018) in their study discussed that the diverse users of OGD have different motivations, such as extrinsic, intrinsic or both. Some users are motivated to use OGD to gain insights into the government, develop new products or services or satisfy their curiosity. Regardless of the potential advantage of using OGD, some users may struggle to motivate to find the time, allocate and finance the required resources to use OGD (Smith & Sandberg, 2018). This might be due to the absence of immediate results when using OGD. It can also be related

to difficulties in understanding or using available datasets, which might result in frustration and demotivate some users to invest time and resources in analyzing the data. Crusoe et al. (2019) have also identified motivation as one of the phases in the users' process of OGD; to utilize OGD, the user must be informed and motivated. Achieving higher-level goals and objectives heavily depends on the motivation to find and use existing datasets.

Inability to identify opportunities: The least discussed barrier in the selected study was related to the inability to identify opportunities in available data. This might also be linked to the limited technical ability of some OGD users, the complexity of available data (Crusoe et al., 2019) or some other challenges such as language or data clarity reasons.

Table 3.2: Overview of identified barriers to using OGD

Origin	No.	Barrier/s	References
OGD portals	1	Insufficient analysis tools for using open data (No data visualisation tools, merging tools)	(Zuiderwijk & de Reuver, 2021) (Crusoe et al., 2019) (Ruijer & Meijer, 2020) (Saxena & Muhammad, 2017) (Saxena, 2018b)
	2	Lack of user-friendly features such as advanced search tools, not able to perform dataset previewing, the absence of a sample for large-size data or no reuse example of provided datasets	(Zuiderwijk & de Reuver, 2021) (Ferati et al., 2020) (Benitez-Paez et al., 2018) (Ruijer & Meijer, 2020)
	3	Findability issues (the ability of the user to discover and identify the data, finding the usable/content-relevant datasets)	(Benitez-Paez et al., 2018) (Xiao et al., 2019) (DAHBI et al., 2019) (Crusoe & Melin, 2018) (Crusoe et al., 2019) (Zuiderwijk & de Reuver, 2021) (Xiao et al., 2019)

Continued on next page

Table 3.2 – continued from previous page

Origin	No.	Barrier/s	References
	4	Availability and accessibility of the data	(Reis et al., 2018) (Smith & Sandberg, 2018) (Zhang & Xiao, 2020) (Benitez-Paez et al., 2018) (Zuiderwijk & de Reuver, 2021)
	5	Diversity barriers to the different users' needs	(Crusoe & Melin, 2018) (Saxena & Muhammad, 2017)
Dataset	6	Limited quality of the data sets including data incompleteness and/or data format issues	(Crusoe & Melin, 2018) (Saxena, 2018a) (Reis et al., 2018) (Zhang & Xiao, 2020) (Zuiderwijk & de Reuver, 2021) (Wieczorkowski, 2019) (Ruijer & Meijer, 2020) (Benitez-Paez et al., 2018) (Nikiforova & McBride, 2021)
	7	Lack of documentation on the available datasets	(Smith & Sandberg, 2018) (Crusoe & Melin, 2018)
	8	Lack of data standardisation and interoperability (data structure and formatting issues, data consistency)	(Smith & Sandberg, 2018) (Benitez-Paez et al., 2018) (DAHBI et al., 2019) (Crusoe & Melin, 2018) (Zuiderwijk & de Reuver, 2021)
	9	Currency (data must be up to date, and data need to reach users before it spoils)	(Benitez-Paez et al., 2018) (Saxena, 2018b) (Zhang & Xiao, 2020) (Crusoe & Melin, 2018) (Smith & Sandberg, 2018)

Continued on next page

Table 3.2 – continued from previous page

Origin	No.	Barrier/s	References
	10	Low dataset usability with the absence of (sample for large size data, reuse example)	(Ferati et al., 2020) (Benitez-Paez et al., 2018)
	11	Terms of use and licenses are not clear or do not foster reuse and distribution.	(Benitez-Paez et al., 2018) (Crusoe & Melin, 2018) (Reis et al., 2018)
	12	Lack or absence of quality metadata	(Saxena & Muhammad, 2017) (Crusoe et al., 2019) (Saxena, 2018a) (Reis et al., 2018) (DAHBI et al., 2019)
	13	Data credibility and accuracy (lack of data source traceability)	(Saxena & Muhammad, 2017) (Zhang & Xiao, 2020) (Reis et al., 2018)
OGD users' abilities	14	Lack of knowledge or capability to use the data such as unable to infer deeper information, difficulty in understanding the dataset or Unable to merge or link OGD	(Zuiderwijk & de Reuver, 2021) (Ruijer & Meijer, 2020) (Smith & Sandberg, 2018) (Saxena, 2018a) (Zhang & Xiao, 2020) (Saxena, 2018b) (Crusoe et al., 2019) (Xiao et al., 2019) Ferati et al. (2020)
	15	OGD users struggled to motivate, allocate time, and finance the required resources	(Smith & Sandberg, 2018) (Crusoe et al., 2019)

Continued on next page

Table 3.2 – continued from previous page

Origin	No.	Barrier/s	References
	16	Hard to identify opportunities and usages	(Crusoe & Melin, 2018)

Since we have identified barriers to using OGD, we also categorised the source of where these barriers originating from. The next section of this chapter will identify the activities of the data intermediaries.

3.2. Open data intermediaries activities

In this chapter, we conduct a second narrative literature review aiming to identify the roles and activities of the open data intermediaries and the NPDIs. Identifying the roles and the activities will also facilitate analyzing the selected NPDIs' roles and activities based on what is known in the literature.

3.2.1. Literature search strategy

The activities of the data intermediaries were identified by conducting a literature review. Since this project is closely related to one of the committee members on-going PhD project, the PhD candidate "L. Pilshchikova", who is an expert in the NPDIs field, was approached to recommend relevant studies about the roles and the activities of the data intermediaries to be included in this literature review. L. Pilshchikova has conducted various literature reviews on the NPDIs' activities and impact; hence, based on these literature reviews, she was able to recommend relevant literature to include. Studies 1-6 were obtained by L.Pilshchikova recommendations. Also, the following search string was used to search Google Scholar for additional studies: **(open data) AND (intermediaries) AND (activities OR roles)**. The search string might have limited coverage regarding the relevant literature on the topic as one of its limitations. Moreover, it might be the reason to overlook other valuable sources where the roles and activities of OGD are discussed along with other aspects, such as the barriers of OGD. The limitations of the representativeness of the selected studies were addressed as the search string was used to further complement the recommended studies by the OGD expert rather than selecting the entire list of studies. The first 50 results in Google Scholar were further scanned to identify additional studies. This step added three more studies to this literature review, 7 - 9; please refer to Table 3.3 for an overview of the selected studies. We conducted a narrative literature review, as proposed by Baumeister and Leary (1997); one of the narrative literature review types aims to examine a given subject's understanding status. For instance, the review might offer

Table 3.3: Overview of selected studies for the activities of data intermediaries

No	Title	Author/s
1	Understanding multiple roles of intermediaries in open government data	(González-Zapata & Heeks, 2015)
2	Open government data intermediaries: a terminology framework	(Magalhaes et al., 2013)
3	Beyond mere advocacy: CSOs and the role of intermediaries in Nigeria's open data ecosystem	(Enaholo, 2017)
4	Empowering Communities with Data: Role of Data Intermediaries for Communities' Data Utilization	(Yoon et al., 2018)
5	Open Data Intermediaries in Developing Countries	(Van Schalkwyk et al., 2016)
6	Promoting the use of open government data: Cases of training and engagement	(Gascó-Hernández et al., 2018)
7	Solving the Democratic Deficit: The Role of Open Data and Intermediaries	(Frank & Waddell, 2014)
8	Towards a Common Definition of Open Data Intermediaries	(Shaharudin et al., 2023)
9	How do Non-profit Open data Intermediaries enhance Open data Usability? A Systematic Literature Review	(Pilshchikova et al., 2022)

helpful summaries and syntheses of a subject, but they are not meant to present new concepts, perspectives, or conclusions (Baumeister & Leary, 1997). These reviews can help compile information on a specific phenomenon. Since they combine numerous pieces of information into a readable format, narrative reviews are useful in research (Green et al., 2006). While narrative reviews can be flexible and fast, on the other hand, they might be more biased and might be less rigorous (Green et al., 2006). Therefore, to address this limitation, we described our methodology in details about searching and including the studies.

3.3. Results of open data intermediaries activities

This section discusses the results of the narrative literature review. We start by discussing the importance of open data intermediaries and NPDIs from the reviewed studies. Then we identify the roles that open data intermediaries can assume within the OGD ecosystem. Lastly, we specify the activities undertaken by open data intermediaries to reduce the barriers to using OGD. For the list of roles, please see Table 3.4; for the list of activities, please see Figure 3.3.

3.3.1. The importance of data intermediaries

The term "intermediaries" was first used in the 1980s within ICT research with an emphasis on the process of intermediation than describing the individuals or organisations that conduct the process (Van Schalkwyk et al., 2016). We define open data intermediaries as "an agent (i) positioned at some point in a data supply chain that incorporates an open dataset, (ii) positioned between two agents in the supply chain, and (iii) facilitates the use of open data that may otherwise not have been the case" (Van Schalkwyk et al., 2016, p. 4). González-Zapata and Heeks (2015) argued that data intermediaries exist to overcome the barriers in OGD. They linked the success of OGD initiatives to the intermediaries' existence and activities. Furthermore, NPDIs are part of the more prominent social objectives of non-profit organisations, which include helping marginalised populations, increasing individual freedom or empowerment, promoting social change, and delivering public services (Salamon & Anheier, 1992). According to Enaholo (2017), NPDIs can also be considered as an emergent setup of civil society organisations who nowadays use OGD to engage with civil societies rather than in the past, using different methods such as mass protest, boycotts or campaigning and lobbying. Enaholo (2017) discussed that the civil society intermediaries are promoting good governance, such as the promotion of increased accountability in the use and expenditure of public funds, justice, fairness, and equality. Moreover, open data intermediaries help to connect complex open data with the people who need such data, in particular in less privileged communities, which helps remove some of the obstacles and challenges to using the data and increase the number of people that benefit from open data (Van Schalkwyk et al., 2016).

3.3.2. Roles of open data intermediaries

Open data intermediaries can assume a wide range of roles, from demanding data to communicating the knowledge of the available OGD. The study of Magalhaes et al. (2013) took a broad perspective by trying to define the intermediaries' roles. They identified three groups of intermediaries civic startups, open data services and infomediaries. **Civic-startups** focus on developing civic-minded digital solutions, which frequently capitalize on the potential of OGD, that seek to address issues on both the governmental and citizen levels (Magalhaes et al., 2013). **Open data services** are defined as new software businesses expanding the software industry by building new innovative solutions on top of the released OGD (Tammisto & Lindman, 2012).**Infomediaries**, a term used in different fields, in the context of OGD, infomediaries offer goods or services based on OGD to the general public or other third parties (O'Hara, 2012).

González-Zapata and Heeks (2015) identified five different roles of the OGD intermediaries concerning the data flow. **Demanders** demand the release of specific datasets or even advocate for the creation of specific OGD policies. **Producers** collect information from the field as well as combine it with already-existing datasets to

produce new open government data. **Validators** who do dataset selection, filtering, or evaluation to overcome the limited usefulness of such data by validating it against real-world conditions or in other ways by ensuring its accuracy. **Developers** focus on forming the information from the available datasets by developing tools such as apps or websites that make manipulation of available OGD more accessible to civil society. **Communicators** that go beyond the primarily technical role of developers to manipulate OGD and present it in formats, channels, and contexts that are best suited to the knowledge and skill of average members of civil society.

The roles identified by González-Zapata and Heeks (2015) did not include or emphasise in essential roles of some data intermediaries, such as the aggregator and the educator roles. Hence, we will adopt the identified roles of the data intermediaries from Pilshchikova et al. (2022) as proposed by Haan (2018) who complimented the five identified roles of the data intermediaries by González-Zapata and Heeks (2015) with the additional two roles. **Aggregators** which collect open data from various sources in order to supply it to users or to extract information. **Educators** address the lack of knowledge and enhance the abilities of the end-user of open data by providing how-to materials, such as manuals or online or offline educational resources. The roles identified by Haan (2018) were adopted as they are more detailed and capture the complexity of the effort of the NPDIS. An overview of the identified role in Table 3.4

Table 3.4: Overview of identified roles of data intermediaries, adopted from (Haan, 2018)

No.	Role	Description
1	Demander	Demand the release of specific datasets or advocate for the creation of OGD policies
2	Aggregator	Collect open data from various sources
3	Producer	Collect information from the field
4	Validator	Evaluate, select and filter available datasets
5	Developer	Developing tools such as apps or websites that make manipulation of available OGD more accessible
6	Communicator	Manipulate OGD and present it in formats, channels, and contexts that are best suited to the knowledge of the end users
7	Educator	Address the lack of knowledge and enhance the abilities of the end-user

3.3.3. Activities of open data intermediaries

Open data intermediaries can assume one or more of the identified roles, and within each of the specified roles, they conduct different tasks and activities to achieve their mission. Frank and Waddell (2014) in a broad perspective argued that the data intermediaries have different agendas which can be achieved in various levels of engagements; **support**: intermediaries can and do assist in making open data more valuable to the

public, **deliberation**: foster a better debate by setting up the necessary infrastructure, letting people know about the opportunity, coordinating and guiding the discussion, and processing the outcomes, **Collaboration**: facilities the interaction and willing to take the burden to solve the identified problem of interest.

Yoon et al. (2018) stated that data intermediaries provide various activities to their communities. For instance, they collect local data, produce reports, analyse and interpret the data, and share data in a processed format. Gascó-Hernández et al. (2018) focused on the training provided by the data intermediaries, which focused on the training of introductory knowledge to use OGD. The training programs of the data intermediaries consists of more than analysis skills to analyse and manipulate OGD, such as leadership or management.Van Schalkwyk et al. (2016) summarised the intermediaries' activities to: help in locating OGD, combine OGD into a specific topic, organise their results into a format that is understandable to interested users, and then disseminate this information to the public.

A recent study from Shaharudin et al. (2023) has identified 14 activities of open data intermediaries. The classification proposed by Shaharudin et al. (2023) is more holistic, and some activities were broken down into sub-activities or are operationally different. For instance, they differentiate between the technical integration of heterogeneous datasets in compiling datasets versus the more sophisticated work of integrating homogeneous datasets in augmenting data. Moreover, their classification can be divided into three main focus areas, preparing and facilitating the missing parts for the users to use OGD, such as activities 1, 2, 3, 4 and 5. The process of using and transforming data into information or knowledge was split into different steps and activities, such as activities numbered 6, 7, 8 and 9. In addition, they included and distinguished the activities that are more related to the process of using the data, such as numbers 10, 11, 12, and 14. In the following paragraphs, we will briefly explain the classification of the activities of the open data intermediaries by Shaharudin et al. (2023). using the same numbering of the activities they used in their literature review.

(1) Compile data, they gather information from many sources, publish it on their platforms, or use it to provide end users with some sort of service or good. **(2) Building data capacity**, which involves organising end-user training and seminar sessions, hackathons, and other open-data-related activities (da Silva Craveiro & Albano, 2017). **(3) Augment data** includes improving the data by integrating open data from various sources and open data with private data. **(4) Contextualize data** includes giving the data a particular context that is important to the audience in order to make the information understandable and meaningful, for instance, converting budget data into easily understood spending stories that can be used and disseminated by a broader range of users(Lämmerhirt et al., 2020). **(5) Curate data** where some data intermediaries might contribute to assembling datasets based on the needs of communities and smaller geographies, such as neighbourhoods (Chan et al., 2016).

(6) Develop product of services most frequently in the form of mobile and web applications to assist civic groups in gathering and integrating data that is relevant to them (Yoon et al., 2018).**(7) Interpreting data** implying transforming OGD into digestible information for their audience. For instance, non-profit organisations can use interpreted data for political work and to translate complex data into easily understandable information for citizens in various social domains such as health and education(Enaholo, 2017). Moreover, **(8) Visualise data** where OGD is displayed as charts, maps, and other visual representations (Dumpawar, 2015). **(9) Validate data** entails validating, updating, and correcting the data to address its inaccuracy and incompleteness.

(10) Demand data involves identifying particular datasets that have to be made accessible, for instance, according to local requirements or needs (Corbett et al., 2018). **(11) Facilitate stakeholders' interactions** by the use of open data events or direct networking (Juell-Skielse et al., 2014). **(12) Channeling feedback** to the appropriate stakeholders on data or issues identified through using the data. **(13) Improving technical openness of OGD** such as transforming OGD into a more machine-friendly format. And lastly, assisting data providers in **(14) identifying potential risks** in opening specific datasets (Davies & Edwards, 2012).

3.4. Discussion

The usage of OGD is not simply about providing more open data; it is beyond the provision of the open data and about a complex process of iterative learning, interacting and networking with the data providers and other actors in the OGD ecosystem. Furthermore, to build an effective OGD ecosystem, the continuity and sustainability of such an ecosystem must be considered. Moreover, the barriers to using OGD are not isolated, but the barriers are interconnected, where barriers at different or similar phases of using OGD can impact each other. In other words, the occurrence of one barrier can be the reason for the occurrence of other barriers at a later stage. For instance, when the data is inaccurate and incomplete, this might lead to demotivating OGD users. In addition, when the data is provided in ambiguous content, format or structure, the data will become less usable, and the users will have less awareness about its potential and how to use it.

It is worth mentioning that most of the studies focused on technical barriers, and fewer focused on non-technical barriers such as institutional or social barriers. For instance, issues with the published data format or structure might happen because of the lack of institutional arrangements to safeguard the data's provision and quality standards. Moreover, it is also not discussed how each barrier is related to the main objectives of OGD. For example, which barriers are more obstructive to transparency or are more linked to operational advancement, Innovation, and economic growth were also less discussed. We found only one study that discussed the barriers of OGD

from an innovation perspective. Furthermore, we conceptualised the barriers to using OGD in Figure 3.2. However, it was challenging to draw links between the barriers. We visualised the barriers as a list of instances or occurrences preventing OGD users from reaching the target of OGD reuse and value creation.

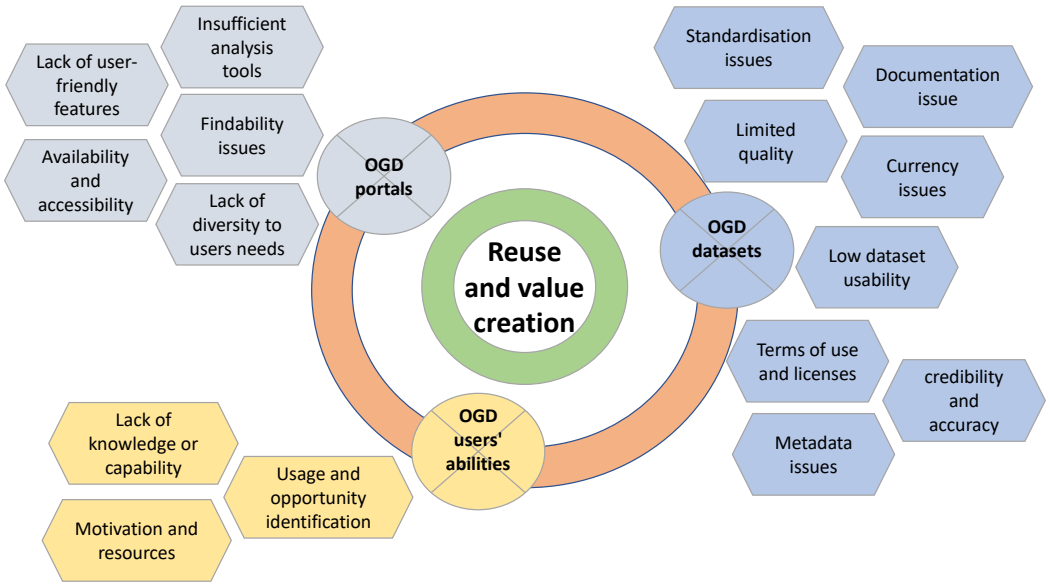

Figure 3.2: conceptualization of the barriers to use of OGD

Moreover, the data intermediaries could assume one or more of the identified seven roles Table 3.4; these roles are higher-level conceptualisations of the work performed by the data intermediaries. In addition, the recent study of Shaharudin et al. (2023) identified 14 activities the data intermediaries perform, which we used to identify the activities of the data intermediaries. We conceptualised the data intermediaries' activities throughout the OGD process. Our conceptualisation illustrated that the process of using OGD is not sequential but rather iterative, meaning some activities are interconnected in specific instances for specific user needs and can lead to further activities.

3.5. Conclusion

A literature review was conducted to answer the first research sub-question, *SQ1: What are the barriers to OGD use?* The main objective of this research is to explore the impact of the activities of NPDIs on the barriers to using OGD. The barriers to using OGD were the focus of this study, having the assumption that the contribution of NPDIs is more related to the barriers to use. Therefore, through the systematic literature review, we have identified 16 different barriers related to the use of OGD, as seen in Table 3.2, and conceptualised in Figure 3.2. We have classified the barriers to using OGD identified in the literature according to our assumptions of where each barrier mostly originated. Most of the barriers were identified as related to the available

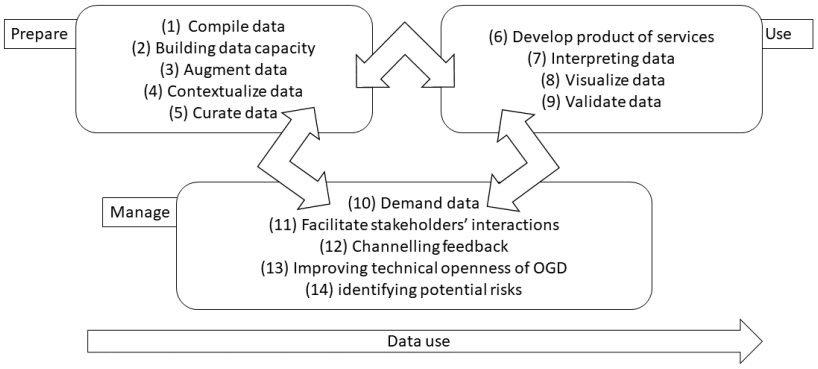

Figure 3.3: conceptualization of the activities of open data intermediaries, list of activities adopted from (Shaharudin et al., 2023)

datasets, such as the limited quality, lack of documentation and data not being up to date. Barriers related to the portal of OGDs or the data publishing mechanism were also identified. For instance, OGD portals may lack some user-friendly features such as advanced search or data previewing, and the data being scattered or duplicated in different data portals/sources, impacting the available datasets' findability. Lastly, barriers originating from the user's ability were identified, such as OGD users lacking knowledge or ability to use the data or perform technical tasks such as merging different datasets. Some OGD users might struggle to motivate to invest time and resources in using OGD due to the learning process and sometime the absence of immediate results.

Furthermore, we conducted a narrative review to answer the second research sub-question, *SQ2: What activities are performed by NPDIs?* The narrative literature review revealed seven roles that open data intermediaries could assume and a list of 14 activities performed by the open data intermediaries. The roles of the data intermediates were a form of conceptualisation of a group of certain activities. Moreover, the sequence of the activities was not discussed in the reviewed studies. The identified barriers to using OGD, and the roles and activities of open data intermediaries, provided context and a better understanding of the phenomena we are investigating. Moreover, the identified roles and activities of the open data intermediaries will be used to analyse the selected NPDIs in the case study chapter.

4
The case study

This chapter discusses the case study background information, and the list of selected NPDIs will be introduced.

4.1. Case study design

This section discusses the case study approach, Also, the mechanism by which the cases were approached and included. Furthermore, the list of selected cases will be introduced briefly.

4.1.1. Motivation of the case study approach

Adopting a case study to explore the impact of NPDIs is essential. Because according to Yin (2018), case studies are suitable to answer research questions of "how", which is the case for our research. Moreover, case studies are also recommended when the behavioural occurrences that constitute the system are out of the researcher's control (Yin, 2018). Furthermore, when the focus of the research is to gain a better understanding of the phenomenon and not to focus on its historical developments (Yin, 2018). The case study approach is used mainly to answer *SQ3: What activities are performed by the selected NPDIs?* and *SQ4: How do NPDIs and their users perceive NPDIs to reduce the barriers to using OGD?* Both will be answered in the preceding chapter 5. Moreover, our study is an explorative study where the impact of the NPDIs in reducing the barriers remains relatively unexplored, as stated in subsection 1.3.2. Furthermore, considering our position as researchers without influence on the phenomena and our focus on understanding the phenomena being the impact of the NPDIs on OGD use barriers confirmed the suitability of the case study approach.

4.1.2. Case study information sources

In our case study, the primary source of information is interviews with individuals from NDPIs and their users. This was decided due to the lack of empirical data that can inform us about the effects of NPDIs on OGD other than the interviewees. We aim to explore multiple cases with less depth on each case. Due to pragmatic reasons with regard to the number of cases willing to participate in our study, we have not been able to specify a specific role we were aiming to interview of individuals from the NPDIs, such as program managers or data scientists. However, in our communication, as seen in section A.4, we ensured that the study scope is clear to increase the suitability of the nominated interview to have background and perception of the NPDIs impact. This was further confirmed by engaging with the proposed interview candidates

before conducting the interview, ensuring the scope alignment, and clarifying the expectations. For instance, one potential NPDI willing to participate was excluded due to the lack of OGD perspective on European countries. In alignment with the explorative scope, we aimed to interview users with diversified roles and interests, such as NPO, academic experts, teachers and students of the NPDI educational programs; for the list of interviewees' information, please see Table 4.1. Moreover, interviews were complemented with reviewing the NPDIs' websites. The selected NPDIs in our case study were analysed through their websites to identify their roles and activities as NPDIs. The different roles identified in the literature review were used to categorise each NPDI. Furthermore, the list of specified activities of data intermediaries in the literature and the interview findings were used to identify the list of activities performed by the NPDIs.

Table 4.1: Interviewee list

Interviewee number	**Perspective**
NPDI-1	Non-profit data intermediaries
NPDI-2	Non-profit data intermediaries
NPDI-3	Non-profit data intermediaries
NPDI-4	Non-profit data intermediaries
NPDI-5	Non-profit data intermediaries
User-NPO	Non-profit organization that used some of the Non-profit data intermediaries activities
User-Researcher	An academic researcher that used some of the Non-profit data intermediaries activities
User-Teacher	A teacher that used some of the Non-profit data intermediaries activities to teach OGD-related courses
User-Student	A student that used some of the Non-profit data intermediaries activities to study OGD-related courses

4.1.3. Recruiting interview candidates (NPDIs and users)

Identifying potential cases started with a list of cases obtained from one of the project team members, "L. Pilshchikova". The initial list consisted of 13 potential NPDIs from 10 different countries. Further to this list, the internet was searched to identify other relevant NPDIs that can be included and contacted to be requested to participate. Several techniques were used, including identifying and browsing identifiable coalitions of NPDIs and searching academic studies for the appearance of NPDIs' names as part of the studies. These techniques resulted in identifying 21 different NPDIs in 12 different countries that were eventually contacted, using an email template as seen in A.4. Four NPDIs have indicated their inability to participate due to human resources or time constraints; 10 NPDIs did not respond to our initial communication and also the reminders

that were sent after two weeks after the initial email. Seven NPDIs were willing to participate, which was anticipated because organisations are sometimes reluctant to participate in academic research, especially for master thesis, as it might offer a limited return on the time invested by the organisations. Two NPDIs were excluded; one was excluded after conducting an interview due to the wide scope of their activities, which included the majority of non-OGD data; also, their activities seemed more related to advocating and proving the resources of open data content creators without focusing on the data to be used or other objectives. The other NPDIs were excluded before conducting the interview because of the lack of a European perspective in the organization. Lastly, cases were also approached pragmatically. Due to the lack of willingness to participate from approached NPDIs, seven showed interest in participating. This has limited the ability to select and prioritize cases that meet the selection and diversity criteria. Our selected cases are satisfactory and aligned with all the criteria we introduced.

Second, we identified some users of the selected NPDIS. After conducting the interviews, the participants from each NPDI were asked to refer us to one of their users or beneficiaries. User interviews were highly desirable as they would reduce the bias and strengthen the findings in our analysis by including a different perspective on the impact of NPDIs in reducing the barriers to using OGD. Also, they were used as a means of triangulation where we consider more than one perspective to explore the impact. Through this process, we have identified six users that use the tools and resources of the different NPDIs for various reasons. Some identified users used the data for academic research purposes, education and teaching, or for citizen participation and achieving diverse societal objectives. After contacting the users, four users agreed to participate in the interview, making the total number of participants in this study nine individuals.

4.1.4. Semi-structure interview design

A total of 9 semi-structured interviews were conducted with key individuals from the identified NPDIs and their users. Two different protocols for each group (NPDIs and the users) were developed and followed during the interviews A.5 A.6. In addition to the different protocols for each group, each interview candidate was asked slightly different questions based on the interview flow. The questions were intentionally formulated in a manner that did not focus on specific barriers to using OGD or activities of NPDIs. However, instead, it aimed to encourage the participants to elaborate and explain how they perceive the impact of the NPDIs. Moreover, adopting a broad perspective in the questions was intended to avoid implying pre-determined answers. When asking participants about the specific contribution towards a specific impact, we risk receiving socially desirable responses instead of the actual reflection on the situation. Therefore, the open-ended questions enabled the participant to express their viewpoints freely without intended or untended influence. The interviews were conducted online using Microsoft Teams, and they were recorded and transcribed. Each interview lasted approximately one hour and consisted of around 20 questions.

For most of the interviewees, the result of the literature review with regard to the identified barriers to using OGD was presented to be further discussed during the interview. The transcript of each interview generated by Microsoft Teams was reviewed and edited; then, it was shared with each interview candidate before being included in the analysis.

4.1.4.1. Introduction of the selected NPDIs

Five different NPDIs participated in this study,Table 4.2 provide an overview of the selected NPDIs.

Table 4.2: Overview of the selected NPDIs

No	NPDIs Name	Location	Scope	Website
1	Monithon	Italy	Civic monitoring of public funds	Click here
2	Openpolis	Italy	Data journalism of social objectives	Click here
3	At the School of Open-Cohesion	Italy	Interdisciplinary education program	Click here
4	Open State Foundation	Netherlands	Digital transparency	Click here
5	Our World in Data	United Kingdom	Knowledge on big problems	Click here

- **Monithon**: Is an Italian Non-profit initiative founded around ten years ago in 2013 that evolved into a methodology and a platform to share the results of monitoring initiatives. The words "monitor" and "marathon" are combined to form the word "Moni-thon," and this is what Monithon aims to support: an organised method of monitoring and reporting on public policy. It was founded as a self-developed program to encourage civic monitoring of government spending. The main objective is achieved by combining open government data with citizen involvement to see how publicly funded projects develop and whether they produce the intended effects. Monithon, as they describe themselves, consist of a small group of journalists, public administrators, and curious citizens (“About - Monithon”, n.d.).
- **Openpolis** is a leading foundation for data journalism based in Italy. Their mission is to promote free access to data and information in order to promote a culture of openness and democratic engagement and therefore give citizens more power. The goal is achieved by gathering relevant data on democracy and society and creating a freely accessible data repository to create and disseminate data-based information. Openpolis keeps track of political power, examines institutions, and designs, develops, and manages data-based ICT platforms in order to generate knowledge, tell stories, conduct research, and raise public awareness about specific social issues(“Documentation - Openpolis”, n.d.).

- **At the School of OpenCohesion (ASOC)** belongs to the OpenCohesion initiative, a national open government programme on cohesion policies in Italy. It is an interdisciplinary educational course programme intended for all types of secondary schools to promote civic monitoring of public funds through open data and information and communication technologies. The project-based course focuses on topics such as civic education, the acquisition of digital, statistical skills and data journalism. The project was established in 2013 in Italy; by 2022, they have spread around ten other countries and are offered in ten languages, such as English, French, and Spanish. Around 35,000 students were part of this education program, which resulted in monitoring around 1,200 publicly funded projects under the European cohesion policy ("Educational path | At the School of Open Cohesion", n.d.)
- **Open State Foundation (OSF)** is a non-profit organisation established in 2012. with the mission To promote democratic transparency, accountability and participation between citizens, politics and government, and all that is directly or indirectly related to or may be connected to these broad objectives. ("ANBI – Open State Foundation", n.d.). OSF is mainly active in the Netherlands. They are also advocating and lobbying for the release of more open data, building online tools to utilise the data and organising open data-related events such as hackathons.
- **Our World in Data (OWiD)** is an online publication established in 2011, based in the UK, and focusing on significant global problems such as poverty, hunger and climate change. It focuses on and uses data mainly from academic research and also from public organisations such as the United Nations (UN), World Health Organization (WHO) and the Institute for Health Metrics and Evaluation (IHME). OWiD's goal is "to make knowledge on the big problems accessible and understandable", in contrast to the news in the media, which focuses on events. The major problems that have affected humanities for centuries or even longer and the significant, long-lasting changes that are gradually reshaping our world are the focus of the publication of OWiD (Max Roser, n.d.).

5
Case study data

In this chapter, first, we identify the roles and activities related to NPDIs by focusing on identifying the roles and activities of the selected NPDIs. We used desk research completed by interviews to answer *SQ3: What activities are performed by the selected NPDIs?* Then, we discuss the results of NPDIs' contribution to reducing the barriers to OGD use to answer *SQ4: How do NPDIs and their users perceive NPDIs to reduce the barriers to using OGD?*

5.1. Activities and roles of selected NPDIs

As summarised in Table 4.2, we have examined each of the selected NPDIs through their websites and available documents, as well as during the interviews with NPDIs representatives, we asked questions about their activities to ensure the completeness and validity of our results. The analysis has helped us identify their various roles in the OGD ecosystem, along with an overview of the individual activities of each NPDI. The list of identified activities on open data intermediaries in subsection 3.3.3 was used to determine the activities of the selected NPDIs, a summary of the identified roles in Table 5.1 and activities in Table 5.2.

5.1.1. Monithon

Monithon can be seen as a comprehensive methodology of civic monitoring of public funds. Their target audience includes university students, the general public and political actors. They assume four roles: demander, aggregator, developer and educator. They offer a structured approach to achieve the monitoring goal. Through their online platform, anyone can start a monitoring project within the geographical scope of Monithon. They are, in particular, building the data capacity for OGD users and complementary skills such as public policies and data journalism. This is achieved through offering standard or tailored training programs. They are, moreover, developing services for OGD users to find, analyse and disseminate the findings of their analysis resulting from the re-use of OGD. They augment data by, for instance, linking local projects to the central datasets, which is a task that requires advanced technical capabilities. They curate data and provide tools and service to the public through their project finder tool that assist users in seeing and selecting the relevant data for the specific geographical location they are willing to monitor.

They are a data demander who advocates for the release of more detailed government data necessary for civic monitoring, such as how the funding decisions were

made, the funding objectives, and which actors are involved in the implementation of the projects. They also work with the respective government organisations in Italy to release additional datasets and enhance the available datasets' technical openness. In addition, they facilitate the interaction between the civil society members in a so-called international monitoring network and the dialogue with institutions such as political decision-makers and public and private entities involved in the monitored projects. In addition to the identified activities proposed by Shaharudin et al. (2023), Monithon does provide its users with services that complement the training program. Before users can publish results in the Monithon reports platform ("Monithon report finder", n.d.), which many other users can access, they validate the process and the findings of the users who used the method to monitor the spending or the implementation of the expenditure. Moreover, Monithon promotes the use of OGD through its activities and assists potential monitorers to self-organise using its personalised platform-like page where users can have their web page and branding.

5.1.2. Openpolis

Openpolis, which targets the general public and journalists, assume most of the identified roles of open data intermediaries. They act as a data demander, aggregator, producer, validators, developers and communicators ("What we do | Openpolis Foundation", n.d.). The aim of Openpolis to make the open data not only available but open to society in a journalistic way. They sometimes collect and make the data readable using advanced data collection techniques such as web scraping. Through the different projects they launch, they augment, validate and curate OGD through the so-called data validation pipelines, where they check the consistency of OGD and detect contextual and data anomalies, data linking pipelines and release and use the data through data analysis pipelines. They also contextualise and improve the openness of OGD, such as in the " Openparlimant" ("Openparliament | Openpolis Foundation", n.d.) project, where the data lacked some contextualisation and was aggregated in a way that did not allow for further interpretation and analysis. Through their process, they gathered the raw data, enhanced and analysed it, and released the treated data to the public.

Moreover, Openpolis develops data analysis products or services and interprets, visualises and distributes their investigations. For instance, they offer non-profit services for third parties of treated data. In addition, they also cooperate with other organisations in a Business-to-Business similar relationship, providing customised APIs of their treated data. They develop technical solutions such as data warehouses and databases where data is integrated into a different format for their internal use. They later develop monitoring solutions to look at and use the data in the processed or aggregated form. They also demand specific datasets from various data providers to analyse social or political phenomena for the general public and businesses who want to know more about the current state of a particular issue or aspect. An example was when Openpolis demanded data about refuges' reception, then analysed and interpreted the released data and made the investigation reports publicly available ("Centers of

Italy", n.d.). Lastly, they promote the use of OGD through the membership programs where their users can sign up for specific themes such as education, emigration, and Parliament monitoring and then receive a weekly newsletter of reports powered by the use of OGD ("Subscribe to openpolis newsletters - Openpolis", n.d.).

5.1.3. At the School of OpenCohesion (ASOC)

ASOC is an initiative part of the broader government initiative of OpenCoesione in Italy (https://opencoesione.gov.it/it/), which is a national open government program on cohesion policies financed by national and European funding for Italy. ASOC is coordinated by the Evaluation Unit of the Department for Cohesion Policies of the Presidency of the Council of Ministers. The project is coordinated with the Ministry of Education in Italy and the Italian EU Commission Representation. ASOC are data aggregator, developer and, more obviously, educator. Moreover, in partnership with other NPDI "Monithon", they build the data capacity of high school students initially in Italy and in almost ten European countries. The capacity building is achieved by developing and offering a comprehensive program taught in high school. The program covers and guides the students throughout the process, from raising their awareness to answering the civic monitoring research question and sharing the outcomes. In other words, they offer an end-to-end hands-on civic education course.

Moreover, their target audience are taught technical data skill, data journalism, and other required skills or knowledge, such as how public administrations function and share data. They also facilitate the interaction between stakeholders, such as students and the institutional representatives of their communities. Furthermore, they expand their scope and facilitate the interaction between different schools for the cross countries' projects monitoring ("The International Initiatives | A Scuola di OpenCoesione", n.d.). Through the interaction with Monithon's tools, which are partly included in ASOC educational program and used to disseminate the analysis results, they validate the results of the minoring reports for accuracy and validity before publishing. Through the tools and structure of activities, they enable students to self-organize and provide them with specific roles to play in the monitoring team, such as data analyst, storyteller, media manager or coder ("UBUNTU | At the School of Open Cohesion", n.d.).

5.1.4. Open State Foundation (OSF)

OSF mainly focuses on digital transparency by opening up public information as open data and making it more accessible for OGD users. OSF assumes most of the roles of data intermediaries and conducts nearly all the identified activities. OSF are a project-based organization that undertakes projects to serve various social goals aimed at the general public, as well as projects in collaboration with several public organizations in partnership with the aim of making OGD more usable. They compile, augment, curate, and contextualize open datasets to make them more accessible. For

instance, they do web scraping of available data but not shared in an open format, then assemble their own datasets, for instance, the scattered election data that were collected and made openly available for everyone(“Slechts kwart van alle gemeenten deelt verkiezingsuitslag als open data – Open State Foundation”, n.d.). Moreover, a particular example will be the elections monitoring project where the data of an open process "election" was collected in a PDF format and turned into a more machine-readable format "CVS" and "JSON" (“Waar is mijn stemlokaal”, n.d.). Once the data was transformed, it was analyzed, interpreted, validated, and promoted to be re-used by citizens or journalists.

OSF demands data that are needed to achieve particular objectives in its agenda. For example, The Open spending project(“Open Spending - Home”, n.d.), first demanded all municipalities release the spending data, then had it through their data preparation pipeline and promoted the data to be used by the general public and journalists who are interested to find out about the spending on each local municipality in the Netherlands. They also work closely with government agencies to ensure that the data they publish follow Open data principles such as "FAIR" (“FAIR Principles - GO FAIR”, n.d.), which simultaneously improves the technical openness of the data. They also promote the use of OGD by making a wide range of users aware of their rights to access public data. For instance, in their recent work of the comparative analysis on how long it takes to receive requested data based from government entities based on the Open government act (Woo) (“Research into the Open Government Act shows: a lot of action, but little result – Open State Foundation”, n.d.), as well as provide consultations on how to officially request data for journalists and the civil society.

5.1.5. Our World in Data (OWiD)

Our World in data acts as a data aggregator, producer, validator, developer and communicator, With the main goal to "make knowledge of the big problems accessible and understandable" (Max Roser, n.d.)]. After gaining high visibility during the COVID pandemic, due to one of their datasets about COVID cases being globally popular, they recently started targeting individuals with high data literacy, such as policymakers, journalists, and people who can affect change. They are based in the UK; however, they have a global perspective where they collect data from sources such as the UN, WHO and various academic repositories. Moreover, Our World in Data simplify the data for the general public and even for journalists, showing how things have evolved over time, which is not a simple task for the general public and even for users with some experience in manipulating datasets. In other words, they collect publicly available data and validate it for accuracy and errors in values or contextually. They clean and process the data, ensuring it is explained well to the users. This is achieved by contextualizing the data rather than just showing it. The data is visualized through charts that present all the details and the units, which means that data is interpreted, not just visualized, to explain what the data tell and make the issue understandable. They also integrate different data sources and periodically update

their datasets that hold global data sets such as the climate change database(Hannah Ritchie et al., n.d.). Their analysts use the treated data once the data is turned into data-science-friendly datasets. Moreover, the treated data is then shared in an open repository, such as GitHub, for anyone to use or even replicate the works ("Our World in Data · GitHub", n.d.). These activities promote the use of OGD and the inclusion of some of their datasets in partnership part of teaching materials that use some of the datasets ("Teaching Hub - Our World in Data", n.d.).

Table 5.1: Summary of identified roles of selected NPDIs

Role	Monithon	Openpolis	ASOC	OSF	OWiD
Demander	X	X		X	
Aggregator	X	X	X	X	X
Producer		X		X	X
Validator		X		X	X
Developer	X	X	X	X	X
Communicator		X		X	X
Educator	X		X		

Table 5.2: Summary of identified activities of selected NPDIs

Activities	Monithon	Openpolis	ACOS	OSF	OWiD
Compile data		X		X	X
Building data capacity	X		X		
Augment data	X	X		X	X
Contextualize data		X		X	X
Curate data	X	X		X	X
Develop product of services	X	X	X	X	X
Interpreting data		X		X	X
Visualise data		X		X	X
Validate data		X		X	X
Demand data	X	X		X	
Facilitate stakeholders' interactions	X		X		
Channeling feedback				X	
Improving technical openness of OGD	X	X		X	
Identifying potential risks of OGD					

5.2. Interview results

This section discusses the interview data analysis results and introduces the identified themes and sub-themes derived from the interview's transcripts analysis about the impact of NPDIs on OGD use.

5.2.1. Interviews findings

The analysis of the interview transcripts resulted in the development of several themes that manifest the NPDIs' contribution to reducing the barriers to using OGD. Some themes we introduced may signal an extended impact of the NPDIs beyond reducing the barriers to using OGD. The list of themes will be discussed in the remainder of this section.

5.2.1.1. Building OGD capacity and expertise

NPDIs contribute to building the data capacity and expertise in many different ways, which contributes to reducing several barriers identified in our literature review; for instance, it reduces the lack of the knowledge or capability to use or even understand OGD, and it can lead users to be more motivated to use OGD and are aware of its potential since users might be more aware of how to use or read OGD. several categories under the capacity and expertise building theme were identified.

Improving OGD users' knowledge and capability: From both perspectives, the users, such as students, teachers, academics and the NPDIs, almost all discussed that NPDIs contribute to improving OGD users' knowledge or capability in different ways. Each OGD requires a variety of technical or contextual knowledge or capabilities. The NPDIs address this for the various needs; NPDI-5 states," *.. the main barrier for young people who are students working on this matter who sometimes have the first impression of data published in the CSV file, you know that you will find a lot of information, and so that is why we explain how not to panic and to use the proper tools to refine and take all these plenty of information and refine them . . .* "[NPDI-5]. Moreover, when asking User-NPO about how the activities of the NPDIs in their work, their response complements how the NPDIs perceive building the user's capacity and capabilities: "*the way to use the data, it was useful the fact that they did us training so two associates of . . . came to my city in . . . , for one week. They explained to us what open data is, how you can modify it, how you can use it, and how to get the metadata and discover the origin or other material desk research that can help you to understand the data better...*" [User-NPO]. Building OGD capacity was extensively discussed during the interviews, although no specific questions were asked about the specific activity; instead, the interview questions were general. The frequency of discussion signalled the significance of improving OGD users' ability for the NPDIs in their agendas and for the users as one of their significant barriers when trying to use, analyse or understand OGD.

Allowing others to build on their previous knowledge: The NPDIs are capitalising on the previous knowledge or expertise they have, where these experiences are availed

to OGD users and also providers in many ways, for instance, through their network, where they have more quick access to the resources or capability. Moreover, sharing their experience publicly and allowing interested individuals to benefit from the knowledge and experience. NPDI-2 stated that they are exposing their technical work openly and freely for technical users to benefit and learn from their experience; NPDI-2 stated, "*We are participating in a Python conference where we are exposing a project we did through Python, so that is another target, not only the general public but also the technical public. . .* "[NPDI-2]. Extending the available knowledge, expertise, experience, or connection of the various NPDIs allows the creation of communities that enables users, providers and citizens to learn in a collaborative way empowered by the previous knowledge or capability that the NPDIs have. In some instances, the NPDIs allows OGD user to leverage the data collected and processed by the NPDIs in a way that allows users to leverage the current state of knowledge known about the research question and expand it without the need to start from scratch, User-Teacher stated: "*. . . we found out that there had been a community of open data researchers, where they made lots of collection of data about. . . , then they analysed the problem, and they exchanged the data with us. They are very generous. They exchange the data. They cooperate to spread the knowledge about the problem...*"[User-Teacher].

Offering diversity in backgrounds and disciplines: One of the significant impacts of NPDIs their role in building OGD users' capacity to offer diversity in backgrounds and disciplines in the activities cornering their users. This is achieved by offering diverse learning opportunities through their methods and/or individuals. This is relevant as, in some cases, OGD users lack a combination of more than one skill or expertise that hinders their use of OGD. As well as the entities they partner with that come from various backgrounds to extend their support to OGD users, such as politicians, academia, and the general public, and offer various disciplines, such as data science, social science, data journalism, policy and technology. The diversity in the perspective also enabled simplification of matters related to the political systems and policies and provided OGD users with access to a wealth of perspectives and expertise which enriched OGD users' understanding and interpretation of OGD, User-Student stated: "*... so the main challenge that we faced while working with the open data, was that public administration documents are not very user-friendly, so we received help from experts ..., where two people from this association . . . . helped us to understand how European structure and investment funds and [the policy under investigation] work because without their help it would be more difficult . . .* " [User-Student]. Moreover, NPDI-1 also stated, "*. . . . [the NPDI] was created exactly as a way to use open government data and leverage specialistic expertise and knowledge of how public policy works in a very detailed way. I think it was important at that time that the data from that portal were very rich, but it was very hard to understand them. . .* " [NPDI-1]. Furthermore, the diversity enabled OGD users to explore OGD from different angles and fostered interdisciplinary collaboration across different fields, NPDI-1 stated, "*Maybe the unique characteristic of . . . . is this, you know, the blending and the mixing of different skills...* " [NPDI-1], enabling OGD users to discover new possibilities,

approach OGD from different perspectives, and make more informed decisions while and after using OGD.

Ensuring to have hands-on experience: Building OGD users' capacity by giving OGD users practical experience is one crucial way NPDIs are lowering the barriers to using OGD. The hands-on approach provides opportunities for users at different levels of expertise to engage through a practical learning approach, which implies that the hands-on approach necessitates users to approach the data rather than only providing a theoretical training approach. This is done by enabling users to access data repositories using tools that enable direct interaction with OGD. The adoption of the hands-on approach by the NPDIs allowed users to develop essential skills in data exploration, manipulation and interpretation, which also advances the non-technical capability of the users to better use and understand OGD and its potential, User-Teacher stated, "*. . . [the NPDI approach] allowed us to follow a project-based learning, to try to explore with our students how they can become active citizens because open data is a pure exercise of active citizenship*" [User-Teacher]. Moreover, this approach allowed users to build their competency and confidence in using OGD by allowing them to navigate and make meaningful use of OGD to address societal challenges; NPDI-5 stated, " *. . . and they really can touch with their hands and see with their eyes, their own eyes, what is the value of the data... students not only learn by studying, but they learn by doing ...* "[NPDI-5]. Furthermore, as the students witness their findings of using OGD start to emerge, this gives the students sense of accomplishment and ownership; User-Teacher stated, "*Students start to see that the research grows in their hands. And, of course, they start being curious. . .* " [User-Teacher]

Reaching different user groups: Exposing OGD to various user groups, such as students, is one of the impacts NPDIs create when building OGD capacity and expertise. These users are becoming more aware of OGD and its potential and are acquainted with the required knowledge and expertise of OGD. Perhaps one of the significant impacts is that the new user groups not only develop technical and non-technical expertise but also become sort of ambassadors or somehow intermediaries who advocate more often their social objective and engage with their community and local government as a result of being able to derive some insights from OGD, NPDI-5 stated, "*We had this kind of impact, civic impact and so the impact is to not only to promote open data but also to spread open data over their territories . . . this is very interesting if you think that this kind of process is starting from young people. . .* " [NPDI-5]. Moreover, this can also be seen as a way of preparing the new generation for the new language of transparency and democracy that is OGD, where these students actually become active monitorers of public funds. Also, in a way to reach the users at a very early age during their education and using OGD as an eye opener to its potential and what can be derived out of it, User-Teacher stated, "*. . . this morning we visited the school in the area where we were conducting our investigation project, they were students of 13 years, we told them that this problem could have been prevented. . . and everybody was very surprised that to know that*

. . . they started saying who told you? OK, who told me? I will show you where it is possible to read this" [User- Teacher].

5.2.1.2. Improving OGD accessibility, quality and usability

NPDIs are actively working to make OGD more accessible to a broader audience by enhancing data accessibility, quality, findability, selection, and contextualization. They develop user-friendly platforms and interfaces that enable OGD users to find and use OGD in an easier and standardized way. By improving OGD accessibility and quality, they enhance the trustworthiness and reliability of OGD for its users. Several impacts under the theme of improving OGD accessibility, quality and usability were identified.

Making OGD open not only available: NPDIs contribute to improving OGD's accessibility by improving the technical openness of the data. The technical openness of the data has connections and implications on many barriers, such as data standardization and interoperability and data quality and accuracy. The effort of NPDIs in making the data open instead of being available stimulated its reuse; NPDI-3 stated, "*Our provinces and other organizations do publish their data through our National Statistics Agency, but it was not being used that much. So, we have a website . . . and that is being used by journalists to dive into the finances of such organizations; the data was open but not reusable. We promoted the use of this data; by the way, before we stepped in, it was not open data. So first, we made it open data, and then we made it so that people actually could use it. So that is the tool we created.* " [NPDI-3]. One way of making the data more open is by simplifying the data or conducting the data processing since NPDIs have the technical infrastructure. Then providing the data in its new open state to interested users: NPDI-2 stated, "*It is the approach the government has it is like I am just giving you the link. I am just giving you the data, and maybe I am providing it in a very fancy way, like fully coloured, but For the journalist, they need to replicate and extract their points of view, so it's hard sometimes, so you just need to give like a very simple data and maybe like a way to read it and I think that's the lack of the government that since we talk about open data, it's just reachable, but it's not open.*" [NPDI-2]. Moreover, not only downstream, NPDIs also improve the technical openness of OGD upstream by ensuring that the data are published in certain standards that ensure its reuse and accessibility once published. Sometimes NPDIs, in collaboration with other entities or users, publish reports to nudge the data providers indirectly and make them and the public aware of the shortcoming in the data they provide, which lead to improving the way the data is provided, User-Researcher stated, "*And then actually I think was helpful because at some point then in the subsequent year from 2019 on, I see that they also the the the Home Office started to collect the data as harmonized ... because this was something that we made in a sense a point of attention in the first report, the one describing the data collection process...* " [User-Researcher].

Gather, curate, integrate and create: The NPDI's ability to compile, curate, integrate and provide meaningful data sources has an impact on lowering the accessibility and usability barriers of OGD. The Data that NPDIs assemble, process and create can be

used by them or also provided to OGD users freely to be reused; NPDI-3 stated, " . . . *we assemble open data database, for example, on elections, we use public data from PDF files, and we transfer it into an open JSON or CSV file, takes a lot of work. But we did it so the data can be used.*" [NPDI-3]. Moreover, by gathering and integrating such data, the NPDIs create rich and comprehensive repositories which contain a wide range of information; NPDI-4 stated, " . . . *so, in theory, anyone could copy our data .. our software developers are working on creating an API which would make this all data completely available, and we have a catalogue which we are trying to build, which would make all of our data catalogues available to anyone.*" [NPDI-4]. Furthermore, these activities that are offered often as a bundle make the process of utilizing OGD less overwhelming for the OGD users, where they are presented only with the information they need to use or process further.

Improving OGD quality: Improving OGD quality and integrity is crucial to instil users' confidence and not waste users' time or provide inaccurate insights when using OGD. NPDIs are playing a critical role in this regard by implementing measures and using techniques to enhance the quality and integrity of OGD, for instance, checking the data consistency, normalisation and accuracy check. Such efforts are crucial to provide OGD users with reliable and trustworthy data; NPDI-4 stated, '*. . . cleaning the data ... typically, some data is just very messy, inherently the data that I am working on now. I think there is just obviously many typos. I think [providers] themselves upload the data to the [repository]. I am not sure how much filtering The [repository owner] does to check if the data makes sense itself, and there are several cases where the data does not seem to make sense. So, I recently been spending a lot of time just making sure certain values are possible. If a rate is higher than the rate of something per 1000 people and it's over 1000 people, then it cannot be true. . .* '[NPDI-4]. Improving the quality of OGD saves OGD users valuable time and effort, which would be spent otherwise on checking or fixing the data quality issues. Therefore, it allows OGD users to focus on conducting analysis or deriving insights from the validated data. As a result of the NPDIs' efforts to improve the quality and integrity of OGD, users can confidently utilise the data without fears of producing misleading or inaccurate conclusions, thus facilitating a more efficient and impactful user journey.

Improving OGD findability: Data findability is one of the frequently discussed barriers to using OGD. Improving the data findability maximizes the utilization of OGD. NPDIs are playing an active role in improving the findability of OGD by developing tools or portals that make the data easily discoverable. For instance, they help data providers leverage their tools or platforms to improve open data findability and, ultimately, its reusability; NPDI-3 stated, "*We work on findability; we work with several governments and organizations on the portal that's called **** it is right now in development with a commercial organization and us together with another governmental organization. Because of the findability of open data that was supposed to be on one platform for the entire government of . . . . They scrapped it because it was getting too big and too complex, and then they said wait a second, did not [The NPDI] already do something like that? And they came*

to us and said, hey, can you build on what you already have and develop it further so that we at least have something open... you get all the data that are already open as open data from municipalities on one platform. It is a very big improvement."[NPDI-3]. Moreover, they do not only, for instance, make the data centralized in one platform but also apply sophisticated techniques that could allow the users to find the data, for example, by using a map to search and locate the data instead of a search engine, NPDI-1 stated, " *The **** [a tool the NPDI developed]. It is a map in which you can find some relevant information from [the OGD portals]. ... citizens could see exactly where the projects are located in their neighbourhoods.*"[NPDI-1]. The improved findability of OGD enables OGD users to access a wide range of data sets and explore different perspectives to uncover new insights. Eventually, NPDIs contribute to fostering a more open and accessible OGD ecosystem.

Improving OGD metadata and documentation: Enhancing the metadata and documentation of OGD is essential for facilitating OGD understandability, usability, and reusability. NPDI-2 stated, "*what is critical for the general public is maybe to understand the data or to have something that explains how data works? that is something challenging for the general public. So we try to be in the middle of accessing many information from different sources*[NPDI-2]. NPDIs play a role in this regard by ensuring the availability or the creation of well-documented metadata and documentation accompanying the dataset they provide or that is already available. The creation and enhancement of the dataset documentation and metadata make the information valuable with regards to its content, structure and context, including the data sources, methods of data collection and formats; NPDI-4 stated, "*... and then as part of this, the metadata is really central, so we spend quite a lot of time and effort for each variable that's in each data set. We try and document the units and a description of it and make it super clear where it is coming from and what the original source is, and that is all documented on GitHub* "[NPDI-4]. Once released, the detailed documentation helps users better understand the scope of the datasets lacking some metadata or documentations, as well as its potential applications and limitations. This comprehensive documentation and metadata creation and assurance enabled data reuse by making it easier for the user to use and understand the data; moreover, it saves the data users' time from trying to understand the structure and the datasets' characteristics.

Tailoring the support to user's needs: Another aspect of the impact of NPDIs in reducing the barriers to using OGD is their ability to provide tailored services or products that cater for the users' specific needs. The NPDIs play an essential role by understanding the diversity in their partner's and user's needs. This is achieved by actively engaging with their users and communities, which equips them with knowledge about the diverse needs and requirements; therefore, they develop tailored services or products to meet these requirements and needs. For instance, they provide specialised consultations or services tailored to goal-oriented products or services when users or other organisations have different goals. NPDI-1 stated, "*they [referring*

to another NPO] use our tools our methods to do this kind of monitoring, monitor the gender gap and try to monitor projects from a gender perspective. So we provided the tools and the methods for their developers to customise their platform. And we worked part of creating the questions about questioner to monitor the projects from a gender perspective" [NPDI-1]. Moreover, they also provide tailored services that are skill-oriented, which ultimately bridge the gap between technical expertise and user requirements enabling users to achieve the maximum value of OGD. NPDI-5 stated, "*Sometimes you can also find a lack of basic knowledge about how to use an excel sheet or a spreadsheet, so on, you can see it as insignificant, but sometimes, very it can be another barrier. So that's what we think is very important to educate young people to use these tools and will find some links you will find very interesting works*" [NPDI-5]. The personalised approach increased user engagement and enabled a deeper understanding and utilisation of OGD.

5.2.1.3. Empowering OGD users

Through their activities, NPDIs are employing OGD users in many ways. For instance, by validating users' insights and allowing the users to disseminate their insights or conclusions from OGD. Also, raising the users' legal awareness regarding their rights to access and request OGD and acting as legal advisors. Also, by enabling the users to identify opportunities to use OGD and by being transparent and open and fostering and welcoming users to reproduce or use their work.

Supervising and validating OGD use: Reviewing and supervising the work of OGD users before it can be published or distributed has significant importance in ensuring the accuracy, reliability and quality of the insights derived from OGD. NPDI5 stated, "*They [referring to their partner NPDI] review the reports, then they are published after a review. Very often they resend the reports to users because there are some parts there are not so clear, and they need to go deeper in the contents, or sometimes there are some lack of contents or lack of information. . .* " [NPDI5] Moreover, it can encourage users to start using OGD and offset some of the burdens on the users as they might be hesitant to start using OGD and derive inaccurate insights or conclusions. The review process is not only about the data and the conclusions but also about the methodologies employed and the data interpretation, which provide holistic feedback to the users and a venue for improvements. It also promoted collaboration and helped to build trust in the OGD ecosystem. Furthermore, it serves as a quality assurance mechanism that safeguards the quality and integrity of OGD insights and empowers advancing knowledge and understanding in many fields.

Allowing users to disseminate findings: The dissemination of the outcomes or insights derived by users from OGD use is critical. For instance, once a group of civilians used OGD to find out about the progress of a project in their neighbourhood. If the outcomes of the analysis were not spread and shared with a wider audience, it could be considered equivalent to not achieving anything. In this case, the dissemination also incentivises some users to use OGD. Because, by doing all the hard work of being

able to use OGD for their specific purpose and derive valuable conclusion and insights, which in most cases was related to social matters, without spreading the work, it might not have the anticipated values and impact. User-NPO stated, "*What we did is to fill a report that is almost a narrative report, then we published it on the website of our European project and on [The NPDI] website. So, the idea is to create a narrative space where people can read what was going on...*" [User-NPO]. Moreover, The NPDIs provide tools or platforms that enable the dissemination of findings and empower OGD users with the required knowledge and techniques to promote their work. NPDI-1 stated, "*The final step is sharing because you can be very good at collecting and analysing data using data, but then there must be something. There must be an output of all that, right? So, if you do not share the report, It is like you did not do anything ... if you are an activist. If you are a university member that wants to do something for your community. You must be very good at promoting your work...* " [NPDI-1]. Allowing OGD users to spread their insights empowers them to contribute to the process of policy-making and drive change in their communities, increase public participation and stimulate finding solutions for social challenges.

Identifying opportunities for OGD use: NPDIs provide guidance and support to users on how to access the data and, more importantly, its potential. Identifying opportunities to use OGD empowers users to unlock the value of the data and encourage meaningful application of OGD across different sectors. NPDIs approach and empower different user groups by informing them about the potential of OGD; NPDI-5 stated, "*The lack is from the user to understand how and what you can do with the open data. So at .... one of the objectives is to improve this knowledge among young people and among teachers and the school community*." [NPDI-5]. One of the participants User-Teacher, seconded this when asked why they collaborated with the NPDIs by stating, "*Before, we didn't have any idea that there was a portal that there was a place where a common citizen could see or could understand how all this money are spent*" [User-Teacher], they also added "*the challenge is above all is to understand and to let other understand that the data is something that really exists. People are not aware that everything that happens around us can be checked and controlled*." [User-Teacher]. So the role of the NPDIs was essential in uncovering the potential of OGD and allowing users to also spread the message and educate others about what impacts or outcomes can be derived from OGD use.

Legal empowerment: NPDIS are promoting and empowering the diverse users of OGD on their rights regarding accessing OGD through their initiatives and resources; NPDI3 stated, "*We work together with journalists on lobbying for transparency, corruption research and spending data that's not open yet. And we are very active in the field of freedom of Information, so we support journalists by advising them, helping them write their requests...*" [NPDI3]. And even for some users who are well equipped with technical and analytical skills, such as researchers, the role of NPDIs as legal advisors was critical to establishing and making their experience of using OGD possible. In other words, the NPDIs increased the legal awareness of OGD users and empowered them through active collaboration on how to obtain the data in the most efficient legal approach.

User – Researcher stated, "*Everyone can go through the Freedom of Information Act and in understanding this proceeding, the contribution of [the NPDI] was crucial. Because I had never heard about Freedom Information Act before doing this project*" [User – Researcher], they added: "*We made this collaboration with .... and saying let just make this as sort of consultancy. Tell us what kind of documents we need to submit and how to write the request in a proper way, and then we will follow the process to see if we are able to go further. . . it took nine months. . . and then at the end, we ended up with the covering 90 ... out of 106 this is how [the NPDI] basically helped us to go through this process*" [User – Researcher]. Empowering OGD users with this knowledge contributed to making the OGD ecosystem more inclusive and diverse, where users are exercising their rights to access OGD with confidence to be able to derive values from OGD.

Adopting a white-box approach: NPDIs are adopting and promoting a whit-box approach, which refers to adopting a transparent approach in the data sources and the methodologies, which is not the default case when it comes to using OGD. Other data intermediaries, for instance, who are concerned with the development of new commercial products or services are more closed in their approach and, for instance, do not share their algorithms publicly. Adopting the white-box approach provides OGD users with visibility into the data sources and methodologies used in the analysis of OGD. Such transparency in the approach improves users' trust and stimulates users' understanding of the data processes. NPDI2 stated, " *If you want to replicate what we have written, you can do it like in the research approach. I have this data. You can get the data, and you can verify if you have a different point of view because, with the same data, you can say the opposite of what I'm writing. So it's like giving the instruments to the general public, but it comes with an active approach..* " [NPDI2]. Moreover, NPDI-4 added, "*. . . our goal is to be as open as possible and everything in theory, if you had enough development skills, you could probably recreate [their entire projects] yourself. . . .*". [NPDI-4]. Therefore, ultimately, the white-box approach empowers users to take decisions based on reliable information and encourages meaningful engagement with the NPDIs and OGD.

5.2.1.4. OGD process optimisation

NPDIs also contribute to the process optimisation of utilising OGD. Their experience and expertise, as well as the resources they provide, ensure that OGD datasets are accessed more efficiently; they also enhance and optimise the process of OGD users' support and engagement.

Recognised as a best practice: Along the way, through the methods, tools or reports of NPDIs, they are being recognized as a best practice that people believe is practical and leads to effective outcomes. Therefore, OGD users are more likely to adopt their methods and tools to use OGD for their various purposes without the need to recreate or search for tools that are proven to be successful; NPDI-1 stated, "*We are recognized as a good methodology that many people already used in the past. So it is better to use something that works than invent something that is not there.*" [NPDI-1]. Also, for the data providers,

NPDIs are advising data providers on how to adopt best practices in data sharing which also results in the optimization of the process of OGD provisioning and usage; NPDI-3 stated, " *... the data was partially open, and right now, we are in the process with them [the data providers] to get every municipality, every decentralized organization to deliver that data in JSON formats because then you do not have to worry about formatting your rows and columns...* " [NPDI-3]. NPDIs' recognition as best practice led to the wide spreading of their tools and methods by OGD users and eliminated the need for additional efforts by OGD users to reinvent or search for successful tools, as well as making OGD provision in a manner that stimulates its reuse.

Establish the process of OGD use: For various reasons, OGD users sometimes are not aware of how to start the OGD use process, which can be related to legal, technical or organisational challenges the users are facing. Unlike the resources or tools to support OGD users along their journey in using OGD, establishing the process is critical as users might also start to learn and adapt along the way. But, when they are stuck or unaware of how to start, this learning process will not even be possible; User–Researcher stated, " *. . . they had this strong impact; their role was fundamental in starting because otherwise, we would get stuck. . .* " [User–Researcher]. In some cases, users might fear starting the process of using OGD due to their non-technical background or being overwhelmed by the complexity of the endeavours. Hence the scaffolding approach of OGD allowed users to overcome the fear and start using and experiencing OGD and its values; User-NPO stated, " *I was really afraid looking at the data because it seems to me work that needed a lot of evaluation, statistics, mathematics competencies. And then it was the [NPDIS] methodology that actually made me sure to say no, even if I do not have big literacy in data works, in how to manage, modify or interpret them and present them and make a difference...* " [User-NPO]. Moreover, the holistic approach of NPDIs also enabled users to recognise the possibilities and the means how to start their initiatives to use OGD. Therefore, NPDIs have an essential role in helping OGD users to establish the process of utilising OGD by providing the necessary guidelines and inspiration for OGD users to start their journey.

Enabling self-organisation: NPDIs also provide mechanisms and tools for the user to self-organise, for instance, by allowing them to have a name and a website to brand themselves and their activities. In most cases, OGD utilisation is happening in the form of a team effort. The comprehensive tools with detailed steps to be followed and different roles to be assumed by each team participant provided OGD users with a way to organise their effort. However, having close monitoring and follow-up with them in case help is needed, User-student stated, " *... [referring to the NPDIs] helped us in organising our work by providing us very strict steps and precise deadlines ... they were always up to date on how our work as students should be carried out ...* ." [User-student]. Moreover, The NPDIs also provided teachers with the methodology and the suggestion of the tool on how to make students engaged and organised; User-teacher stated, "*they also give us suggestions and tools how to make students work so the students have a*

different role, there are the analysts, there are the coders, there is at the end of the research, the storyteller... " [User-teacher]. Hence, NPDIs enable users to self-organise through their comprehensive tools and methodologies that offer effective collaboration and successful utilisation of OGD.

Scalability: NPDIs demonstrated the ability to scale up their data scope and geographical coverage, extending the reach and impact in facilitating OGD usage. They were willing to expand their geographical scope across various countries through their partnership approach. Moreover, the streamlined and sophisticated approach of the NPDIs allowed them to quickly expand to different data sources without redeveloping their technical solutions; NPDI-1 stated, "*we would like to extend to work on this kind of tools to improve them ... the tool we created works for all European projects, more than the almost 2,000,000 projects.* " [NPDI-1]. Furthermore, it was noticeable how efficient and effective NPDIs can extend their knowledge, tools and methodologies in strategies that multiply their impact. For instance, one NPDI adopted the "train the trainer" strategy, where they trained individuals from different nations with the aim that the individuals become trainers once they are back and spread the knowledge to their communities; User-NPO stated, "*.. we had a one-week meeting in which five people from each NPO learned the methodology. Then everybody went back to their nation, decided to look at a project to monitor, and they gathered other people, just the citizens and taught them. So passing the methodology also to a larger group of citizens and using the methodology to check a public project ...* " [User-NPO].

Structured approach: NPDIs did not only provide the required technical and infrastructural support; they also provided OGD users with structured methodologies that helped OGD users realise positive experiences and stimulated using OGD. Also, one of the NPDIs users reflected that the OGD experience was structured as the users might not be aware of how to proceed further, even if they have enough technical knowledge "*When we finished this project, it was really structured, and this was helping because we knew what to do and when to do it. . . .*" [User-Student]. In addition to the generic structure and outline of using OGD, the NPDIs provided technical tools that OGD users can interact with to provide dynamic structure and guidelines based on the progress of each different user, thus allowing each user to receive structure and guidelines based on their circumstances; NPDI-1 stated, "*. . . [the tool] is an online guide that is created every time a user picks a project. So, when they select a project, the tool is created based on the characteristics of the project. With a very long chain of if and then, we could provide some specific suggestions and questions based, for example, on the status of the projects. Is it just started or completed? The suggestion changes because the methodology change just, for example, based on the topic and the theme, so the if a project is related to environment or research and innovation. . .* " [NPDI-1].

Long-term viability: NPDIs support and embrace a sustainable approach to using OGD. Their different activities include compiling a curating the dates and presenting

them to OGD users as insights or new information. Because of their experience, NPDIs develop sustainable solutions that do not rely on entities or individuals but rather streamlined processes and procedures, meaning that their services and tools are guaranteed to be accessible and available in the foreseeable future; NPDI-4 stated, " *... even if everyone stopped working at **** tomorrow. A lot of our data architecture would still exist and still, hopefully, be accessible and work for the foreseeable future. Which I think is probably not the case for a lot of data creators...* " [NPDI-4]. Moreover, through their experience, they also conduct further activities to ensure users have sustainable access to OGD. For instance, when specific data are more likely to be removed from some OGD provider's portal, they anticipate and gather and archive these data sources to ensure they do not disappear and interrupt the services provided to their users; NPDI-4 stated, " *We have a system like an ETL (extract, transform, load) pipeline. We essentially take snapshots of data sources at given points in time, and that allows us to kind of maintain different versions of the data as well. So if the data is updated, occasionally it changes a lot between versions, and the previous version is no longer available at the source, but we have a system in which we keep multiple versions ...* " [NPDI-4].

5.2.1.5. Promotion and advocacy of OGD

Through the promotion and advocacy NPDIs conduct, they increase the availability and accessibility of the data and even OGD quality. They achieve this in different ways, for instance, by demanding the data or lobbying to enhance its quality, using OGD as a proxy for their social goals, raising the awareness of OGD users about its availability or potential and promoting the use of OGD.

Demanding data: Demanding the release of new datasets from data providers exceeds satisfying the NPDIs' needs for these datasets. Since the data is treated and propagated, it stimulates more data use by OGD users. Therefore, the data is not only aimed to be released but also treated and turned into a valuable resource for various uses. NPDIs also lobby to improve the quality and the level of openness of the provided. Furthermore, when some OGD users have a solid use case, they sometimes struggle to realise it as the data is not shared or not fully open. Working with NPDIs and getting the data helped users to realise their goals. User-Researcher stated, " *... we realise this research project on using the data we got access with the help of [the NPDIs]...* " [User-Researcher]. Furthermore, NPDIs anticipate the social challenges or the potential value of particular datasets; therefore, they demand such datasets and ensure they are publicly available, leading to the development of innovative solutions; NPDI-3 stated, "*The information on tram data we started opening up tram data in **** because ... it was not available for reuse, it took hold to all public transport companies, and now you have [a new mobile application] ...* " [NPDI-3]

Promoting the use of OGD: Promoting the use of OGD by shedding light on its availability and potential, and not only the data the provider shares publicly, but also the data that NPDIs gather and treat; NPDI-3 stated, "*. . . we sometimes scrape data that's*

not open, but it's public and assemble our own open databases. We then do analysis on it or warren journalists to look out that we have an open database, you can use it ... " [NPDI-3]. Moreover, NPDIs have initiatives or mechanisms to promote the use and reuse of OGD. For instance, one NPDIs have a membership program and newsletter that users could subscribe to and receive updated insights and highlights about the topic of their interest. Furthermore, indirectly by allowing a wide range of users to access and reuse their work freely and without restrictions, it created a partnership with highly visible partners such as TV channels and journalists; NPDI-2 stated, "*For example, [TV show name] is like a Public TV show on public topics that sometimes use our data and they ask us for analysis of data ...* " [NPDI-2], also NPDI-3 added, "*We did research into voting by mail ... and I could have publicized it on Twitter, on our website and a couple of 1000 people would have read it. We worked together with journalists, and a couple of million people were able to read it...* "[NPDI-3].

Raising awareness: Raising users' awareness about OGD availability and impact is another way of promoting the use and reuse of OGD; User-Teacher stated, "*What they did is to expand and develop this culture of data, to tell this to the school; they helped the school to understand that [OGD] exist and can become a way to teach.*" [User-Teacher]. Also, by conducting data-related campaigns and sharing reports targeting a wide range of users, including politicians and policymakers, through different channels. Also, NPDIs target the providers' awareness; therefore, they not only advocate and demand the data but also advocate and promote using OGD standards and providing consultation and feedback on how the data should be shared, NPDI-3 stated, "*... they [the ministry] called me to ask, Is this data sets conforming open data standards? ...* " [NPDI-3]

OGD-blended education: NPDIs effectively promoted the utilization of OGD by incorporating and blending OGD with civic education at school. Integrating OGD into the curriculum allowed students to use real-world data and gain a deeper understanding of social issues. The NPDIs approach enabled teachers to teach and incorporate OGD into their classrooms; teachers have also gained skills and knowledge to navigate and utilize OGD; User-Teacher stated, "*... we started six years ago to see if it works. And now, it has become a very important project in our school because it is a clear example of civic education and how civic education can enter the curriculum.* "[User-Teacher], also NPDI-5 added, "*... in particular, digital skills are very useful also for the future of the students, and also teachers...* " [NPDI-5]. This blending approach does not only advance students' knowledge about civic knowledge; it also enhances their data literacy and stimulates more use and awareness of OGD and its potential; User-Student stated, "*... by participating in the project, we understood the importance of living in a vigilant and monitoring community; our work was only a small contribution but a great opportunity to become more active, aware and involved citizens ... these lessons on data that we took were really useful to us later, even in our regular educational activities at school...*" [User-Student].

Intertwining OGD and social objectives: All the NPDIs that we interviewed are mainly driven by social objectives; User-NPO stated, "*So what we did with the citizens that participated in our project was to clarify how public money was spent. So this is what we try to understand* " [User-NPO], also NPDI-2 stated, " *... we tend to let people know about how politics work ...* " [NPDI-2],Their integrative approach incorporates OGD into their activities to contribute to the social challenge and improves OGD accessibility and relevance of the data to address real-life challenges. Integrating OGD in their processes, systems and initiatives allowed users to leverage positive change through OGD. Such an approach ensured that OGD is not just a standalone but rather a valuable tool that can be utilised to address social and other challenges. Moreover, their impact is not mainly derived from promoting OGD. However, they aimed to intertwine OGD with the social objective, which entails linking OGD to social challenges and creating a culture of using OGD as a powerful tool to address various social challenges; NPDI-3 stated, "*Open government data and the reuse of open data by the general public on large not a very interesting subject, getting people to realise how this is important for me means getting them to recognise the importance of open data on subjects that matter to them. ... So getting the subject of open data to mean something to the general public is part of our strategy for the coming year in trying to relate it to actual topics like the agricultural debates...* " [NPDI-3].

5.2.1.6. Stakeholders collaboration and engagement

NPDIS impact OGD stakeholders through their facilitation and embracing of engagement and collaboration. This is achieved in different ways, for instance, by channelling OGD users' feedback to OGD providers, through their partnership with other organisations, by being a venue for unique and valuable interactions, and eventually improving OGD usability and applications.

Collaborative feedback loop: NPDIs work closely with OGD providers to implement improvements and solutions to enhance the usability and quality of OGD. Their knowledge, expertise and active approach to their communities enable them to work collaboratively with data providers to address OGD users' feedback and adopt more efficient ways to provide OGD; NPDI-1 stated, "*We try to convince the administrations to do better. And we offer solutions; we offer suggestions to do better . . . We are part of the open government partnership working group. And this is quite useful way to address these problems because you talk directly to people that implement this open data portal, so the people that have all the technical skills to understand this kind of this matter.*" [NPDI-1]. This also illustrates their roles in being in the middle and translating and formulating the issues that OGD users are facing in a more technical way that can be better comprehended by the developers of the OGD portal. Moreover, their knowledge also allows them to anticipate the issue in the provisioning of OGD that might hinder its utilisation. They can act upon these issues collaboratively with OGD providers. NPDI-3 stated, "*. . . they [data providers] published it as a PDF. We called them and said hang on, that is not open date, and no one can use this. Can you do better? And they did. Sometimes it is as easy as picking up the phone and saying, hey, you need to do better, and sometimes it takes longer*

and a lot of talking.." [NPDI-3]

Venue for interaction: NPDIs stimulate OGD stakeholder interactions and create platforms that foster unique and meaningful engagement. Their initiatives and approaches created an environment to encourage and facilitate interaction and collaboration in the OGD ecosystem. One way is by providing a space for diverse stakeholders to exchange knowledge, thoughts, concerns and expertise. NPDI-1 stated, "*Because in those hackathons, this was the first time that those different worlds could meet and exchange their skills and views ... creating this environment where public policy experts met civic hackers and civic communities. And that created a very good cake of these ingredients which could make this magic happen.* " [NPDI-1]. Therefore, this unique nature of interaction facilitated and embraced by NPDIs activities brings added value to the OGD ecosystem. The diversity in perspectives and expertise creates an interdisciplinary collaboration that solves complex societal issues. This is evident in one of the NPDIs where their users were able to resolve a long-lasting social issue through the use of OGD; NPDI-5 stated, " *We have stories of the impact of the territories of the civic monitoring research conducted by our users, for example, re-qualification of the hostel of a building originally owned by Mafia. . . . our users who work on this and they faced with a lot of problems. . . they went through the aim to let this project start again. So they involved the mayor, the local communities, citizens, other NGOs and the Bishop of the city . . . they reached the goal. . .* ." [NPDI-5]

Partnership: NPDIs recognise the importance of partnerships with entities such as other NPOs, data providers and civil societies, and such partnerships maximised their impact and expand their reach. They actively seek and endorse opportunities to partner with other NPOs that share the same or different goals; NPDI5 stated, "*In the third lesson, they have to compose the monitoring report, which is then published into the [other NPDIs] platform is another very important partner of the project.* " [NPDI5]. Moreover, through these partnerships, NPDIs work with individuals who will be equipped with a deep understanding of OGD and develop the skills to communicate its benefits effectively. They invest in these users through specialised knowledge, training, and tools to understand and use OGD. As a result, these users are not only able to leverage OGD for their own goals but also play an important role in promoting its value to others. And without the partnership mindset, the emergence of such intermediaries would unlikely happen; User-NPO stated, "*We are contacted by others as intermediaries, so [NPDI] help us to be in a position like their competitor we were contacted by other associations that want to work on conducting training for them...* " [User-NPO]

5.3. Discussion

Our findings of the activities of the NPDIs in section 5.1 showed that NPDIs have a different scope or objectives, often a social goal, compared to open data intermediaries who might be, for instance, interested in profit or the development of innovative solutions. However, the roles and activities of NPDIs, are similar to open data inter-

mediaries in terms of their operations. NPDIs do not necessarily focus on specific roles or activities but rather provide a wide range of services and conduct various activities. Moreover, NPDIs focusing on roles other than educator roles tend to cover a comprehensive list of roles and activities identified in the literature. This may be driven by the overarching social goal where they try to offer a complete solution that does not lack in some areas.

Some of the activities we have identified in section 5.1 are worth discussing since they were not identified precisely as activities in the reviewed literature. Moreover, perhaps these activities are exclusively applicable to NPDIs. First, the important aspect and activity of promoting the use of OGD. Integrating OGD into the mission of some organisations and embedding OGD in the services encouraged interested users to use the data to achieve the social goal. Promoting OGD and incorporating it into the organisation's functions can maximise the social impact and utilisation of OGD. Moreover, the literature did not discuss the specific activity of NPDIs in validating the results generated by OGD users, sometimes who are under training on how to use OGD. Therefore, the NPDIs played a vital role in ensuring the accuracy, validity and credibility of the insights generated by the users. Furthermore, it helped reduce the potential bias or errors that might affect the interpretation of the findings derived from OGD use. At the same time, reviewing the results might give users more confidence about the accuracy and reliability of results driven by using OGD.

NPDIs also developed solutions and offered services to allow users to disseminate their works, including their findings, which might be driven by the social impact of the NPDIs. Moreover, providing the platform, resources, and support to disseminate the insights enabled users to reach a wider audience, resulting in more interaction, collaboration and use of OGD. Lastly, the provision of the tools and services of NPDIs extended beyond instructions and how to utilise OGD for their social purposes. The NPDIs were critical in enabling OGD users to self-organise within the social initiatives, including various aspects such as team composition, roles assignment, branding and group promotion. This comprehensive approach emphasises the technical aspects of using OGD and the importance of fostering organised teams capable of making a meaningful impact.

Our findings about how the NPDIs reduce the barriers to using OGD in subsection 5.2.1 show that NPDIs strongly emphasised capacity building (discussed 65 times), improving OGD availability and accessibility (discussed 60 times) and enhancing OGD findability (discussed 32 times). These priorities were highlighted constantly during the interviews with the NPDIs, indicating that NPDIs are addressing these barriers at the top of their agendas. Moderate emphases and frequency were seen on dataset quality barriers (discussed 20 times), such as OGD currency, diversity, and standardisation. The insignificant emphasis was found on barriers related to the dataset's metadata and documentation diversity (discussed eight times). This

signalled that the NPDIs sometimes operate more in a reactive mode than a proactive mode about the barriers they reduce. This suggests that The NPDIs are responsive and adaptive to users OGD users' needs. However, it also suggests that there is a potential for the NPDIs to strengthen their approach further in anticipating the barriers and attempting to prevent them from occurring. An example of the proactive approach some NPDIs are adopting is when raising the awareness of the new generation about the availability and value of OGD, moreover building their competencies on how to use such datasets in a meaningful way. Another example was when some NPDIs worked proactively with data providers in anticipating issues that might occur because of how some OGDs were provided; such examples illustrate the potential of the proactive approach.

A notable observation was that NPDIs are motivated mainly by a social goal, often more than one, where OGD is adopted as an instrument in their steps towards the overarching social objective. This approach position the OGD as a mean to an end, that is, the overarching social objective. From an optimistic perspective, the purpose-driven approach can be seen as a reason for OGD being utilised more impactfully and meaningfully. This observation points out that the NPDIs approach extends beyond the use of OGD to offer a comprehensive and inclusive ecosystem of support, including the resources, experience, connections and non-technical skills and resources to derive their social goal. On the other hand, the social drive and objective of NPDIs make it challenging to explore their impact in reducing the barriers to using OGD.

The broad perspective and the less stringent conditions, such as legal and financial conditions to expand, gave NPDIs an advantage to spread over wider geographical locations, social objectives, and in some cases, cross-countries. Moreover, NPDIs also adopt some strategies that might have boosted their chances to expand their reach. Strategies such as localising their initiatives where they are trying to engage with specific communities to use OGD for beneficial use that benefits their neighbourhood or cities, or deriving the issue from a bottom-up approach where users are coming up with the specific issues they would like to investigate is maximising the impact of OGD use. The agility and flexibility complement the apparent no commercial interest and the non-competitive mindset. Such a mindset allows straightforward collaboration with various partners without bureaucratic, financial or legal barriers.

Lastly, from an ecosystem perspective, we have two main observations. First, intermediaries do operate in a circular rather than linear manner. They freely extend their resources and experience to various user groups, which can also be intermediaries or become intermediaries at a specific point in time. For instance, when students try to use OGD for Civic education, they are seen as a user of OGD and NPDIs activities. However, the students also have an objective to explain the data and simplify it to their peers who are less capable of doing so. Thus, students can become intermediaries at a certain point in time or perspective. Second, due to the complex interaction from

the ecosystem perspective, it is challenging to draw the line between the users and the intermediaries. Intermediaries can, at some point, be users and vice versa, thus blurring the line between the two and making it problematic to distinguish users from intermediaries in the OGD ecosystem.

5.4. Conclusion

In this chapter, we answered two of our research questions. Starting with *SQ3: What activities are performed by the selected NPDIs?* Using the desk-research method and complemented with interviews. We used the roles and activities of open data intermediaries we identified in the theoretical background chapter to analyse the roles and activities of the NPDIs. A summary of the identified roles and activities can be seen in Table 5.1 and Table 5.2. Furthermore, we analysed our case study data and identified six themes of the NPDIs' contribution to reducing the barriers to using OGD to answer *SQ4: How do NPDIs and their users perceive NPDIs to reduce the barriers to using OGD?* We have identified and discussed six themes where the impact of NPDIs can be manifested, 1) Building OGD capacity and expertise, 2) Improving OGD accessibility, quality and usability, 3) Empowering OGD users,4) OGD process optimisation, 5) Promotion and advocacy of OGD, and 6) Stakeholders collaboration and engagement. Nevertheless, we were not able to draw a clear link between the activities and the barriers that the NPDIs reduce. In light of the results and our understanding of the effect of the NPDIs activities, we re-conceptualise the barriers to using OGD, the activities of NPDIs and our identified effects of NPDIs' activities on the barriers to OGD use. We position the impact of the NPDIs activities in red oval somehow linked to the activities it is related to. However, we also show how the activities can be linked to the barrier categories but not to each specific barrier, as seen in Figure 5.1.

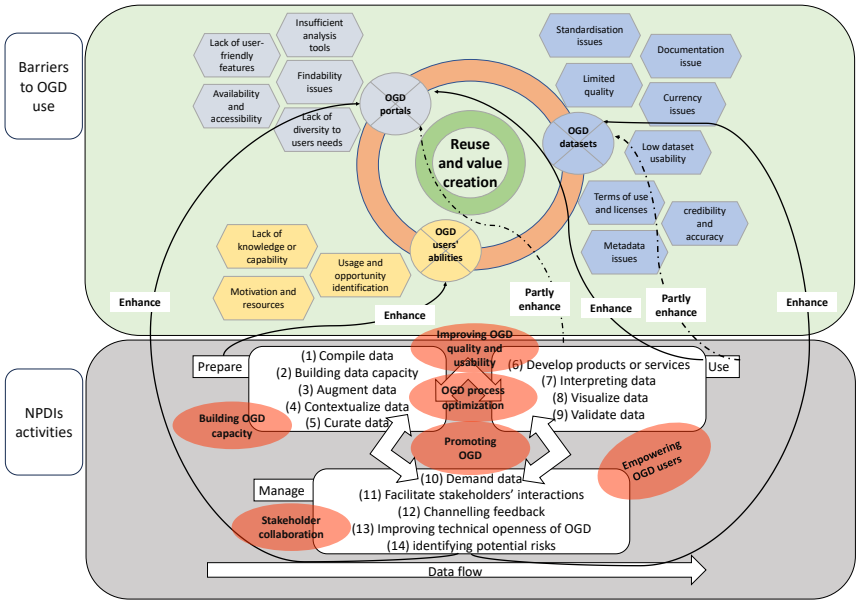

Figure 5.1: Conceptual model of NPDIs activities' impact on OGD use barriers

6
Validation through focus-group

This chapter discusses the last research method used to triangulate our case study data analysis in chapter 5. The focus group method was used to answer *SQ5:How does the participants' feedback support the findings from the interview about the impact of the NPDIs in reducing the barriers to using OGD?* We present the focus group design. Then we discuss our analysis method as well as the results of the focus group session. The chapter will be concluded by answering the sub-research question about reinforcing our analysis and interpretation.

6.1. Focus group session design

We followed the guidelines of Nyumba et al. (2018) to design our focus group. The guidelines of Nyumba et al. (2018) summarise the insights from the focus group applications over the last two decades. This section discusses the focus group design, including the study's objective, developing the focus group questions and identifying and recruiting our participants. Then, we also discuss the data collection and analysis methods.

6.1.1. The objective of the focus group

The first step is to define the focus group session's aim and objectives Nyumba et al. (2018). Aligned with the sub-question we are trying to address, we formulated the purpose of our session: **To discuss and triangulate the findings of our research about the impact of Non-profit data intermediaries (NPDIs) in reducing the barriers to using OGD**. To address the main objective of our focus group session, we prepared presentation slides to facilitate the discussion during the session; please see section A.7.

6.1.2. Participant identification and recruitment

After identifying the objective of our focus group session and developing the discussion points and questions, the next step is to identify the focus group participants (Nyumba et al., 2018). Since our main aim is to discuss and triangulate our findings about the perceived impact of NPDIs in reducing the barriers to using OGD through the focus groups, our pool of participants is limited to the nine participants we have interviewed already. The next step is to recruit our participants (Nyumba et al., 2018). We have requested all of the participants we have interviewed already to participate in our focus group by sending them individual invitation emails that explain our objectives and expectations from their participation. Three participants stated their inability to participate in our focus group discussion due to work-related or health-related

circumstances. Two participants did not respond to our request, and four responded, indicating their willingness to participate. However, a week before the session, one of the agreed participants dropped out for work-related reasons, concluding the list of three participants as shown in Table 6.1.

Table 6.1: Focus group participants list

participant number	**Perspective**
NPDI-P1	Non-profit data intermediaries
NPDI-P2	Non-profit data intermediaries
User-PS	A student that used some of the Non-profit data intermediaries activities to study OGD-related courses

6.1.3. Location and facilitation identification

The next step was to identify a suitable time and location for conducting the session and the required facilitation to conduct the session (Nyumba et al., 2018). Our participants are located in different locations; therefore, the focus group session was aimed to be conducted online. One challenge was identifying a suitable time slot for the group since our participants are unrelated and live in different locations and slightly different time zones. We prepared an email inviting our interview participants and explaining the main objective of our focus group session. Moreover, to ease the process of finding a common time slot, in the invitation email, we included a link to the tool we used to find a suitable time slot, "Rally", which helped us to allow our participant to pick a suitable time over the four days options from 10:00 to 17:00 that we provided.

6.1.4. Focus group data collection

The data collection phase followed the design phase (Nyumba et al., 2018). The data collocation phase encompasses the pre-session preparation and the required facilitation during the session (Nyumba et al., 2018). We developed a script and presentation slides to help us navigate the session. The organizer conducted multiple practice sessions to ensure smooth flow and transition during the session. Moreover, The presentation slides were used to first briefly present the findings of our thematic analysis with regards to the impact of the NPDIs on reducing the barriers to using OGD, and later ask our participants to what extent they agree or disagree with the statement we claimed and allow them to discuss their opinions about it. An example of the statements we asked our participants for their opinions in Table 6.2. The table shows the flow of each of the sections of the focus-group discussion; each section focused on one of the identified themes and included its sub-themes as well. Therefore, in addition to the introduction and the closing sections, we had six sections to discuss our findings; for more details, please refer to section A.7. We planned 1 hour of discussion for the focus group session; we allocated around 8 minutes per section. That left us with

around five minutes for the introduction and around 5 minutes for closing at the end of the session, asking our participants to discuss further or reflect on aspects we did not discuss during the session. The session was conducted using MS Teams, which enabled recording and transcribing the session.

Table 6.2: Focus group discussion statements example

Theme	Statement
Building OGD capacity and expertise	To what extent do you think the NPDI you are (involved with/using data of) are reducing the barriers to using OGD by building OGD capacity and expertise through: • Improving OGD users' knowledge and capability • Allowing others to build on their previous knowledge • Offering diversity in backgrounds (politicians/academia) and disciplines (Data and social, journalism and ICT) to OGD users • Ensuring users follow a practical approach by adopting "hands-on experience" • Expanding OGD user capacity to new user groups, such as students

6.1.5. Focus group data analysis

The next phase was the data analysis phase (Nyumba et al., 2018). Data from focus groups are primarily used to identify areas of agreement and disagreement (Kidd & Parshall, 2000). Although it is recommended to follow a systematic data analysis approach such as coding to analyse focus group data (Nyumba et al., 2018). In our case, the focus group was used as an effective triangulation method that will enable us to validate our interpretation by engaging with our participants again in one session. We prepared the analysis by editing and reviewing our focus group transcript. Then, we examined each section of the transcript to analyse the participants' responses, whether they confirmed or refuted our interpretation of their perceived impact of NPDIs' activities in reducing the barriers to using OGD. In the next section, we will present our results based on the identified themes we developed earlier in subsection 5.2.1.

6.2. Focus group results

This section discusses the findings of our focus group session concerning the identified themes of NPDIs' contribution to reducing the barriers to using OGD. Each theme will be presented with data extracts to triangulate the interpretation of the interview data.

6.2.1. Building OGD capacity and expertise

The results for the first theme that manifested the impact of the NDPIs in reducing the barriers to using OGD showed agreement with our analysis from all participants. It reconfirmed the contribution of NPDIs in building the capacity and expertise of

OGD users. NPDI-P1 stated, "*I think all these things you pointed out are valid points. So, improving the users' knowledge and capability to allow other users to build on the previous knowledge and especially to offer diversity in backgrounds and disciplines data and social science and ICT towards using OGD are one of our main objectives... So we improve the use the reuse of open government data... we do exactly this, and one of the most important things is to make it easier for the final users. ... what I think is significant is to educate to spread kind of culture of open government through the young generation.*" [NPDI-P1].

Moreover, another participant confirmed the important contribution of NPDIs in building the capacity and expertise in a way that simplifies the complexity and educates OGD users using a common sense approach rather than a technical way. NPDI-P2 stated, "*What we do here [referring to the NPDI] is trying to transmit the technical stuff and not to say something very technical because you cannot talk about databases in a very hard way. So you have to transmit a common sense of how data works through articles...* " [NPDI-P2].

From the users' perspectives, our analysis was consistent with their perception. They agreed with the merit NPDIs offer in building the capacity and expertise, especially for the young generation, and also considering the complexity level in using OGD, their objective would not have been achieved without the help of the NPDIs. User-PS stated, "*It is especially important that projects like [the NPDI's project] exist for young people and are implemented in schools. They helped us grow as informed citizens. In my case, we had the opportunity to learn how to interpret data and make it understandable for our peers; data can be complicated without proper training, as it is not always easy to understand.*"[User-PS].

6.2.2. Improving OGD accessibility, quality and usability

The second theme concerns the NPDI's contribution to improving the quality and accessibility of OGD. NPDI-P1 acknowledged all the statements and also appreciated and ascribed the effect of improving OGD quality and usability to the power of the OGD network and not only to themselves as NPDIs. NPDI-P1 stated, "*Again, all these things we are involved in day by day in our projects. So again, this is the power of the network.*" [NPDI-P1].

NPDI-P2 agreed with our statements; moreover, they raised a noteworthy point of discussion by stating that what the NPDIs are doing is crucial to OGD users. However, the activities of the NPDIs are not necessarily sophisticated and are well-known. Therefore, the point of motivation and the trigger for the NPDIs is the "why" in conducting their activities instead of focusing on the "how". NPDI-P2 stated, " *...we are not doing something crazy we are just following some patterns that already exist ... maybe you do not know the perception that is something that impacts you as well. So something that is not very correlated with technical stuff, but how you feel the public, so it's something that's very apart from the technical stuff...* "[NPDI-P2].

The participant representing the users' perspective agreed and acknowledged the value of the NPDIs' contribution in this regard. They expressed the barriers in OGD

quality and usability faced as "fatigue" when using the data. Moreover, they implied that without the NPDI's activities, working with OGD would have seemed far away. User-PS stated, "*We felt the fatigue of working with open government data, but our teachers, thanks to their determination, provided us the motivation to work on civic monitoring because sometimes, working with open data could seem distant for students.*" [User-PS].

6.2.3. Empowering OGD users

All the participants perceived the contribution of NPDIs as empowering OGD users. NPDI-P2 agreed to our interpretation and stressed the validation and supervision of the findings and not only the data (OGD) itself for journalists, subsequently empowering journalists to disseminate the data; moreover, NPDI empowering OGD users to know how to legally request the data as the NPDI recognise OGD as a socially valuable asset. NPDI-P2 stated, "*They [data journalists] contact us, so we have like a peer-approach, they are asking for data, and sharing how they think of the data and if their interpretation could be good. ... this help to disseminate the data to other professionals. Also, we use the Freedom of Information Act that helps us and help the people, the society, to access the data ... So it is something we do, not only us but in a network because we think the data are socially valuable...* " [NPDI-P2].

NPDI-P1 also stressed the significance of providing legal empowerment and awareness to the young generation. The empowerment also acts as an eye opener and provides practical ways for OGD users to overcome barriers in using OGD. For instance, legal empowerment and awareness resulted in a collaborative interaction between young citizens and public organisations to release and enhance OGD. NPDI-P1 stated, "*... we also offered the knowledge to students to understand how to request data. For example, some students . . . talked with some mayors, around 15 mayors in nearby [city name] they put them together, and they sign with them the intention to open data because during the civic monitoring research they were conducting, they found out that the data were sometimes not available in an open format.*" [NPDI-P1].

Lastly, the NPDI user confirmed that NPDIs are empowering their users to achieve their social goal; through the various means, which OGD is one of, users are empowered to fulfil their goal and reach out to a more expansive audience concerning their social goal. User-PS stated, "*I think they empowered us in order to help to raise awareness about sustainable mobility as we chose ... In fact, we could create a video or info-graphics and charts that we spread with our peers and students from other schools ...* " [User-PS].

6.2.4. Optimizing OGD processes

The contribution of NPDIs in optimising the process was also agreed on by the participants. For instance, NPDI-PS agreed and appreciated the structure and guidance in the process of using OGD. Moreover, they pointed out that the teachers that were empowered by the NPDI methodology in teaching how to use OGD were up-to-date and well-equipped with the knowledge of using OGD. This restated that NPDIs are

being and adapting best practices in using OGD. User-PS stated, "*Yes, I agree with the fact that [the NPDI] helped us organise our work by providing strict steps and deadlines. Also, some webinars were organised for the teachers, so they were always up to date on how the work as should be carried. . .* " [User-PS].

NPDI-P1 agreed with the statement and emphasised the importance of the sustainable approach and culture the NPDIs are fostering. This structured approach stimulates the continuity and the value retention of the efforts exerted by the NPDIs, the users and the providers in the OGD ecosystem. This also provided a historical background that strengthens the users' case when using OGD. NPDI-P1 stated, "*We promote a consideration to continue the civic monitoring process. It is a kind of ongoing process in which you will, for example, find the difference between the time in which you conducted the monitoring process; you will go back to the project you were monitoring to see if the data are current and consistent and you will find maybe some lack of information. You will go through other push activities to the public administration...*" [NPDI-P1].

NPDI-P2 agreed with the statements and highlighted a few points. For example, NPDI's contribution in optimising the process is interconnected to the technical infrastructure, in which sometimes NPDIs architect their technical architecture in a way that anticipates the shortcoming in OGD, which allows them to scale up. NPDI-P2 stated, "*For optimising the processes is very based on how you architect your architecture on how you gather and make data speak to each other. If you have a strong base, then you can scale up vertically and horizontally to people to other organisations when the data become bigger and bigger.*" [NPDI-P2]. Moreover, NPDI-P2 also highlighted that NPDIs are in the middle, playing the multi-perspective approach. Such an approach necessitates NPDIs to move into different layers, referring to the requirements of OGD users or providers while focusing on making OGD useable. NPDI-P2 stated, "*So you have to move to different layers, but with the same object. The goal is the same data, but how you transmit them is different, so you have to scale not only about the data but how you approach and transmit your information and your knowledge.*" [NPDI-P2].

6.2.5. OGD promotion and advocacy

NPDIs' contribution to the promotion and advocacy of OGD showed agreement from the participants with our statements. NPDI-P1 agreed with the statements and highlighted the NPDIs' contribution to promoting and building the network of users and experts of OGD. Moreover, they also acknowledge the importance of the collaborative effort with the other partners and OGD stakeholders in promoting the use and reuse of OGD and raising awareness. NPDI-P1 stated, "*From my side, this is exactly the project I work on with the [NPDI], which is promoting the use of open government data. ... through building our network of users, experts and so on year by year. It's our contribution... Also, the contribution to the other people working in the same ecosystem, which is exactly a blended civic education project about the use and reuse of open government data.*" [NPDI-P1].

NPDIs users also agreed with the statements and discussed that the tools and activities of the NPDIs led to achieving social objectives, which is one of the main expected values of OGD. User-PS stated, " *... by participating in this project, we understood the importance of living in a vigilant and monitoring community; our work was only a small contribution but a great opportunity to become more aware and involved citizens and being involved citizens is one of the main aims of civic education ...* " [User-PS].

NPDI-P2 agreed with our statements and mentioned the importance of adopting a bottom-up approach, which means that the activities and the tools offered by the NPDIs promoted the use and reuse of OGD to be triggered by the citizens and not only the intermediaries or other stakeholders in a top-down approach. NPDI-P2 stated, "*... we are trying to make people get the data and use a bottom-up approach ... that it is not only something that we, as data intermediaries from top to bottom, are sharing this data. We are also in some way promoting the general public to do so ...*" [NPDI-P2].

6.2.6. Improving stakeholders' collaboration and engagement

The last theme we discussed in our focus group session that demonstrated the impact of NPDIs in reducing the barriers to using OGD was concerning the contribution of the NPDIs in improving the collaboration and engagement among OGD stakeholders. NPDI-P2 discussed how the NPDIs are fostering interaction also through the partnership. They also added that the NPDIs benefit from this interaction in tackling some of the technical and structural issues they face while using OGD. NPDI-P2 stated, "*So we have this interaction between partners, especially when we are working on sharing data or building new features, and this helps us a lot because, you know, it can be something straightforward and very stupid. But if you are with your head working on the same thing daily, then something you can miss.*" [NPDI-P2].

Moreover, NPDI-P1 discussed the partnership and becoming the venue for the interaction among the stakeholders is also done by sharing the success stories among the network. Sharing such stories also promoted the effectiveness of OGD use, thus stimulating more use of OGD. NPDI-P1 stated, "*I think the most important thing is to foster the partnership ... to use examples, experiences and real examples on how the use of open data can empower the collaboration with other stakeholders because otherwise, it will be very difficult ... because if you see what you can do by reusing open government data and what is the effectiveness and the impact of this reuse, you can also adapt, you can also try to collaborate to the ecosystem, we are constantly doing this.*" [NPDI-P2].

Lastly, the NPDIs user mentioned that the feedback loop facilitated by the NPDI's activity is also bidirectional. They elaborated and gave an example of how they recognised the award they received due to their use case of OGD as feedback from some of the OGD stakeholders, the providers or the public organisations. User-PS stated, "*After doing the other project, we had feedback from the project as we won a European-level award, and we had recognition for our work. So I think this was strong feedback from the*

[NPDI] and other stakeholders." [User-PS].

6.3. Discussion and conclusion

In this chapter, we conducted a focus group session to triangulate our interpretation of the interview data to answer *SQ5:How does the participants' feedback support the findings from the interview about the impact of the NPDIs in reducing the barriers to using OGD?* Our results from the focus group supported and validated our interpretation of the interview data concerning the impact of NPDIs in reducing the barriers to using OGD. It also pointed out some noteworthy points during our discussion with the participants.

Noticeably and in multiple instances, the NPDIs and the users stressed and mentioned that NPDIs are complementing and foster the network approach. The network approach also highlighted the interconnectedness that led to seeing OGD from an ecosystem perspective. NPDIs shared data and sometimes received data or other support from the users or other intermediaries. The interconnectedness in the OGD ecosystem was also witnessed when our participants during the focus group session could not stick to the theme of the discussion; the facilitator kept them on the topic when they were off-topic. However, this occurrence implied that the barriers, the activities and the impact of the NPDIs are interconnected and are not linear; where for instance, one barrier can be to cause of one or multiple other barriers, also the activity can reduce one or multiple at the same time or at later point of time.

The effects were almost always ascribed not only to the NPDIs themselves but to the network of users, providers, or other NPDIs in the OGD ecosystem who collaborate and partner with each other. This signals the nearly no-competitive spirit that aligns well with the social aims of the NPDIs. Nevertheless, one of our participants highlighted the lack of an easy-to-find and trace network of NPDIs and their users. Perhaps once someone is connected and aware of the existence of an NPDI, they will be provided with the support they need. However, when trying to identify NPDIs or travel through this network, it might be challenging, which can be an area of improvement among the NPDIs.

7
Discussion

This chapter discusses the results obtained from our research questions and examines their implications. We will evaluate the results critically and consider their significance and relevance. Next, we acknowledge and discuss the limitations of our research approach and methodologies.

7.1. Results

The literature review primarily focused on the barriers to using OGD. This resulted in identifying 16 barriers to using OGD, which we also categorised based on our assumption of where these barriers originated, as seen in Table 3.2. Almost half of the barriers we identify are related to the quality and credibility of OGD, which means that the reliability and trustworthiness of OGD can be a significant barrier to its adoption and effective use. Moreover, we identified seven roles in the literature that NPDIs could assume, as seen in Table 3.4, adopted from (Haan, 2018). Moreover, we also adopted from Shaharudin et al. (2023) 14 activities exercised by the data intermediaries. Then we abstracted the activities into three dimensions which are "prepare", "use", and "manage", as seen in Figure 3.3. Based on the literature review, We could not identify how much is done in each dimension or activity compared to the other precisely. It is worth noting that we were also not able to link activities to the barriers based on our literature review.

In the case study, the theoretical background results acted relatively as an analytical framework that allowed us to research and identify each selected NPDI's activities and roles in more detail but with no casual relations between the activities and the barriers. During our analysis of the selected NPDIs for our case study, we concluded that NPDIs assumed various roles, sometimes nearly all the roles, by one NPDI, and they covered most of the activities known in the literature. Perhaps one of the things worth reporting is the fact that NPDIs, who are mainly educators, usually place less priority on data-related barriers activities and focus more on process-related activities and NPDIs who focus on other roles than the educator tend to usually cover a wider range of roles and activities at the same time.

Through interviews with NPDIs and their users, we identified themes where the NPDIs' contribution in reducing the barriers to using OGD is manifested. We could not explicitly link the activity and the impact as in which barriers they reduced. The inability to map the effects on the activities of the NPDIs is because of several reasons. Moreover, NPDIs are not concerned with the barriers of OGD as much as with their

societal goals; therefore, they conduct the activities to achieve their social goals without focusing on or measuring their impact on OGD use. Furthermore, the effect of the activities can reduce multiple barriers simultaneously depending on the context, making it challenging to map the activities to the barriers. Nevertheless, we conceptualised how barriers, activities and impact of NPDIs based on our understanding as seen in Figure 5.1

We noticed that NPDIs operate in a network approach that gives them the advantage of increasing their impact and serving their overarching social goals. This observation is also aligned with the study of Van Schalkwyk et al. (2016) with regards to the overlooked dimension of intermediation, where they argue, unlike the impression that intermediaries operate solely in the intermediation layer through the open data supply chain between data providers and users. In practice, multiple intermediaries operate together in the supply chain of open data, thus, intensifying the likelihood of OGD reuse. Moreover, the study of Van Schalkwyk et al. (2016) visualised the data intermediaries as consecutive layers in the middle of the open data supply chain. Their model starts with OGD providers, and then each layer in the middle represents an open data intermediary that builds on the previous intermediary and feeds the following until the data or the services reach the end users. Moreover, drawing the line between intermediaries and users was very challenging, especially in the context of NPDIs. Perhaps this can illustrate that the intermediaries or the intermediation actions are more of a role rather than sometimes a stakeholder type. Depending on the perspective and the point in time, the intermediaries can become users and vice versa.

Regarding the nature of the NPDIS, we noticed they are approachable and agile when approached by partners for collaboration. They require less stringent coordination before OGD users or other stakeholders can interact with them. If we approach the phenomena with different lenses, such as from an economic perspective, NPDIs could be seen as incentivising users to participate in the OGD ecosystem. The dissemination of the findings could be one example of the incentive that stimulates and motivates users to use and disseminate the data. Moreover, the social goal that often connects OGD with the users is another incentive to the users, where the activities and the tools provided by the NPDIs fulfil the social purpose that both the NPDIs and the users have. OGD was seen as one of the means, rather than a driver, of the activities that the NPDIs can employ to achieve their overarching social goal. Therefore, the NPDIs are less focused and concerned with the barriers to OGD and not measuring their effect on OGD use.

Lastly, we triangulated our analysis and interpretation of the interview data through a focus group session that confirmed the validity of our interpretation. During the session, we sensed the collaborative spirit of the NPDIs in the OGD ecosystem and showed the integrative efforts the NPDIs are conducting. Our participants often enthusiastically talked about the network in the OGD ecosystem and how the advancements are ascribed to the network of intermediaries, users and data

providers. It yielded that the collaborative and integrative approach is one of the main characteristics of NPDIs. Despite the positive feedback received from the NPDIs and their users about the valuable contribution the NPDIs make, we must remain cautious. Acknowledging the possibility of unintended consequences resulting from the NPDIs' activities is necessary. There might be hidden issues that were not clear or hard to observe in our cases.

7.2. Limitations

The literature review method employed to answer the first and second sub-research question has some limitations worth addressing. First, the likelihood of selection bias in the selected literature, especially for the narrative review, implies that we might unintentionally ignored studies that were not aligned with our assumptions and point of view. Introducing specific inclusion and exclusion criteria for the systematic literature reviews and seeking the recommendation of studies by an expert in the field reduced our selection bias and made it more objective. The second is the focus of our literature review on the barriers to using OGD. This limitation is connected with our initial choice to focus on the NPDIs' impact on the barriers to using OGD instead of taking a broader approach to identifying the barriers of OGD in general. The focus of our research and literature review on the barriers to using OGD might have prevented some more meaningful insights in the subsequent research phases if we had adopted a wider perspective of the barriers of OGD.

In our case study, also several limitations should be addressed. First, the selection was influenced by some pragmatic reasons where it was challenging, due to the low willingness to participate rate, to select and prioritise the NPDI and the participants to a great extent, such as looking for a specific role or particular years of experience in our participant. Furthermore, three of our cases are based on Italy; this might result in our results applying to the dominant country, in our cases, "Italy", more than the Netherlands and the UK. In addition, we recruited all our users in the case study based on the recommendation of the NPDIs. Participants selected and willing to participate will likely have a beneficial interaction and a positive story about the respected NPDI. During our interviews with the users, we asked the questions in a way that was also applicable to other NPDIs besides the one that nominated the users for the interviews to ensure we could capture any negative experiences with NPDIs. Also, our case study selection focused on the NPDIs and the users to have a comprehensive perspective of the NPDIs' impact. However, it did not consider another essential actor in the ecosystem, the data providers who sometimes partner with the intermediaries. It might have pointed out some valuable hidden impacts of the NPDIs in the OGD ecosystem.

Also, concerning the data analysis in this study, it relied on the participant's responses, which might have provided inaccurate or socially desirable responses. We

conducted desk research to identify the activities and the roles of the NPDIs; however, we could not identify additional means of evidence about the NPDIs' effect on OGD use. Even the NPDIs are not concerned or measuring their impact in this regard. Additionally, since this study followed a qualitative approach and qualitative data coding, this research is not free from researcher bias. The systematic approach in our analysis and the documentation of the methodical steps were employed to reduce the research bias in this study and communicate the objectiveness of our analysis as much as possible. However, this does not guarantee that there is no bias in our results. Moreover, our research approach did not produce a measurable impact but rather an overview of the impact of the NPDIs in reducing the barriers to using OGD. Perhaps this was because we tried to use the barriers to using OGD and the activities of the NPDIs as a starting point for our investigation. It can also be because NPDIs use OGD as a proxy and one of the instruments to fulfil their often social goals. Another limitation of our research is the transferability of our findings to other NPDIs. Since we focused our research on specific European countries, it is essential to acknowledge that the activities and the impact of NPDIs may vary across different regions. Therefore, the findings may have limited transferability to other NPDIs in different regions.

Considering the available time and resources for this research, conducting the focus group and having the participants in one session was efficient. However, it created other limitations regarding the method for triangulating the interview findings. Since most of our findings were primarily positive and praised the NPDIs, this fact might have prevented the NPDIs from invalidating some of our findings even if it was not the case due to providing answers that are influenced by the social desirability bias, which might sometimes outweigh the validity of the response. Moreover, some limitations of the focus group method are relevant to our research, such as participants dominating the discussion due to their characteristics or role. During the session, we acknowledged this and started the discussion of each section with a different participant each time. However, this does not guarantee that the set-up or other participant opinions did not influence the answers.

8
Conclusion

This chapter concludes this thesis report. First, we answer our sub-research questions, and then, we answer the main research question. We will also discuss the scientific and social contributions of our research. Lastly, we provide recommendations for future research.

8.1. Answering sub-research questions

In this section, we answer our sub-research questions that arise from our main research question. We aim to provide a comprehensive insight into exploring the NPDIs' impact in reducing the barriers to using OGD.

8.1.1. SQ1: What are the barriers to OGD use? and SQ2: What activities are performed by NPDIs?

The first and the second research questions were introduced to provide a theoretical background for this research. First, our systematic literature review identified the barriers to using OGD. The systematic literature review identified 16 barriers to OGD use, as seen in Table 3.2. We categorised these barriers under three main categories representing the source of the barriers. These categories are; OGD portals, datasets, and users' abilities' barriers. Half of the barriers we identified in our literature review were related to the datasets, implying that the quality and credibility of OGD are essential to its success and can lead to several barriers to using the available data. Moreover, the barriers to using OGD are complex, meaning that the barriers are interconnected and are not isolated. For instance, barriers related to the dataset quality can be the reason for low motivation, leading to the potential of such datasets being hardly known by novices and sometimes even experienced users.

In the narrative literature review, we identified the roles and activities of the open data intermediaries known in the literature. The roles we adopted in Table 3.4 are; demanders, aggregators, producers, validators, developers, communicators, and educators. Moreover, we identified and adopted 14 activities the open data intermediaries can assume. We conceptualised the sequence or the flow of the activities of the NPDIs in Figure 3.3. We abstracted the activities into three dimensions, "prepare", "use", and "manage", which occurs in an iterative rather than a linear manner. The identified activities and roles facilitated our analysis of the selected NPDIS in our case study.

8.1.2. SQ3: What activities are performed by the selected NPDIs?

We employed desk research method in combination with interviews to research and determine the activities and the roles of the selected NPDIs in our case study. We have identified five NPDIs operating and focusing on various countries with diverse scopes of each NPDI. Unlike other intermediaries, our results showed that NPDIs usually have an overarching social goal. Also, in our analysis, we noticed that NPDIs do not necessarily focus on specific roles or activities but rather provide a wide range of activities. More specifically, intermediaries who focus on roles other than the educator role as the central role tend to offer nearly end-to-end activities and assume most of the roles, showing that the NPDIs use OGD as one of the means to achieve their social objectives. Perhaps some of the work done by the NPDIs can also be identified as activities which might also be specific to the NPDIs. For instance, we noticed that NPDIs are promoting the use and the culture of OGD, which was seen when the NPDIs integrated some OGD benefits into their missions, such as transparency or public responsibility. Another example is the supervision and validation of OGD use by novices and professional users, which might reduce the margin of error and give OGD users more confidence to use and share their insights from OGD use. Moreover, NPDIs incentivise OGD users by providing a free platform to disseminate their insights and reach a wider audience, which benefits users in addressing their social objectives. Lastly, NPDIs' efforts in organising the process of using OGD were noticeable, where they provided specific activities that help OGD users to self-organise and leverage an effective and sustainable use of OGD.

8.1.3. SQ4: How do NPDIs and their users perceive NPDIs to reduce the barriers to using OGD?

The semi-structured interviews with individuals from NPDIs and some users of the NPDIs' services or tools provided us with more context in understanding the impact of the NPDIs in reducing the barriers to using OGD. We conducted nine interviews in this research, five individuals from the NPDIs and four interviews with users of the NPDIs resources and tools. We coded and analysed the interview data and proposed six themes that manifested the impact of the NPDIs in reducing the barriers to using OGD. These themes are; building OGD capacity and expertise, improving OGD accessibility quality and usability, empowering OGD users, OGD process optimisation, promoting and advocating OGD-related activities and policies, stakeholders collaboration and engagement and facilitating and improving stakeholders' collaboration and engagement. The findings showed that NPDIs had mainly emphasised improving the capacity, accessibility and availability of OGD and findability. Therefore, barriers related to the users' abilities, OGD availability, accessibility and findability are the most barriers impacted by the NPDIs activities. This does not mean that other barriers are not impacted or reduced by the activities of the NPDIs, yet it is unclear to what extent these barriers are impacted and reduced. Also, by this, we concluded that the NPDIs' efforts are mainly reactive and less proactive, which signals an opportunity for the

NPDIs to capitalise on, especially those that already have implemented some of the proactive activities. There were no significant differences in the perception between the users and the NPDIs in how the NPDIs are reducing the barriers to using OGD.

The social motive of the NPDIs' activities positioned OGD as a means to an end which is the overarching social objective. This made it challenging to investigate the impact of the NPDIs' activities towards lowering the barriers to using OGD. Moreover, the broad perspective allowed the NPDIs to expand their activities and, subsequently, the impact with more agility and less bureaucracy than other comparable intermediaries. Also, through specific strategies, such as intertwining the use of OGD with the local issues for citizens, the NPDIs maximised their impact. They stimulated the reuse of OGD in a bottom-up approach. Furthermore, the ecosystem perspective was apparent when exploring the effect of NPDIs in reducing the barriers to using OGD. NPDIs were fostering and collaboratively extending their tools and services to make other users or groups become intermediaries without a competitive spirit. It was challenging to draw the line between the users and the intermediaries. This can only be done by knowing the context and specific time points, implying that we can see the intermediaries as a role, not only a stakeholder.

8.1.4. SQ5: How does the participants' feedback support the findings from the interview about the impact of the NPDIs in reducing the barriers to using OGD?

We conducted a focus group session to triangulate the results derived from our interpretation of the case study data. Three participants were able to participate in our session. Two participants represented the NPDIs perspective, and one represented the user perspective. The participants agreed and validated our results. Moreover, reflecting on some of the points we noticed during the session, we noticed the integrative and collaborative approach of the NPDIs from their tones. The importance and the credit were given to other NPDIs, where advancements and improvements were ascribed to the network of NPDIs, providers and users. Furthermore, although the network of OGD actors, specifically users and NPDIs, is growing, there is a lack of an easy-to-find and trace network where users or NPDIs can become aware of the other actors in the network and establish a connection quickly.

8.2. Answering the main research question

The previous answers to our sub-research questions are separate parts of the complex phenomena related to NPDIs' impact in reducing the barriers to using OGD. Here, we answered our main research question, *How do European non-profit data intermediaries affect the barriers to using OGD?* According to our systematic literature review, the barriers to using OGD were mostly related to the OGD itself. However, the NPDIs' activities focused more on the barriers related to the process of using OGD and enhancing OGD users' abilities, such as building the capacity of OGD users and improving

OGD availability, accessibility and findability. This means NPDIs are not explicitly concerned with using OGD but provide the tools and resources to support achieving their social goal by using OGD as one of the available instruments. This does not mean that NPDIs are not contributing to reducing the barriers to using OGD. However, the impact of their activities is not translated into reducing the barriers to using OGD. The NPDIs' approach to reducing the barriers to using OGD is a comprehensive approach which tackles the barriers to using OGD as one of the instruments to achieve the NPDIs' social goal. For example, when the NPDI has a social goal related to the monitoring of public funds, they focus on activities aiming to prepare the potential users of OGDs to be able to do so and contribute to their social goal of monitoring. When the NPDI focuses on data journalism for various social purposes, they adopt a holistic approach and bridge the necessary gaps in most identified activities.

We concluded that the NPDIs' activities contribute to reducing the barriers to using OGD. However, when we tried to map the impact or link the activities to the barrier, it was not possible because each activity can have multiple impacts, which makes it difficult to distinguish and focus on the impact of each activity. For instance, when an NPDI provide training on how to use OGD, they not only reduce the barrier related to the user's ability but also they are helping OGD users to identify opportunities for OGD use, improving the findability and availability of OGD by making the users aware of the existence of such datasets. However, these examples are not exhaustive, and there are many other relations between the NPDIs' activities and reducing the barriers to using OGD.

Moreover, NPDIs are empowering OGD users in many ways, such as legally, by supervising their OGD use or enabling OGD users to widespread the outcomes of their usage of OGD, which reduces the barriers related to the user's ability and the availability and accessibility of OGD. The NPDIs also enhance the process of OGD use by enabling users to establish and self-organise the process of using OGD, as well as employing and fostering best practices which reduce most of the barriers related to the user's ability as well as the datasets and the OGD portals. Promoting OGD is another way of lowering the barriers associated with the users' ability and OGD portals, where users are more aware of OGD availability and potential; thus, they use OGD for different purposes, such as in schools for civic education. Lastly, enhancing the stakeholder collaboration and engagement process creates a collaborative feedback loop that reduces the barriers in the OGD portals and datasets, complemented by maximising their impact through partnership with other NPDIs and adopting the white-box approach.

8.3. Societal and scientific contribution

In this study, We addressed the knowledge gap in the literature about how NPDIs reduce the barriers to using OGD. The anticipated values of using and reusing OGD,

as well as the barriers to use, reuse or provision of OGD, are well addressed in the literature. However, the impact of the NPDIs' activities on the barriers to using OGD is relatively unexplored. This thesis can be seen as one of the initial studies aiming to address the gap in the literature and provide an understanding of how NPDIs are reducing the barriers. Our research also provided insights into some characteristics of the NPDIs and the strategies they adopt to intensify their impact. Our study contributed to the existing literature by exploring the impact of most of the activities conducted by the NPDIs on the barriers to OGD use in a broad perspective, for instance, unlike the study of Brugger et al. (2016) that focused on the impact of the data visualisation. Furthermore, our study has been, to the author's best knowledge, the first study that acknowledges the effect of NPDIs in promoting the use of OGD through their support and use of OGD for social purposes, for instance, recent studies such as González-Zapata and Heeks (2015), Shaharudin et al. (2023), and Yoon et al. (2018) did not recognise the promotion of OGD as one of the specific activity of the data intermediates. Moreover, our study contributed to paving the way for future researchers aiming to explore and investigate the link between the NPDIs activities and the barriers to using OGD, informing them to approach the research design differently, as we have discussed that OGD is one of the instruments rather than one of the objectives in the case of NPDIs.

Furthermore, the results of this thesis highlighted the importance and the value of the NPDIs for the OGD ecosystem and for intensifying the anticipated benefits of OGD, more apparently, the matters related to transparency and efficiency, as suggested by da Silva Craveiro and Albano (2017) where data intermediaries were facilitating to co-creation of budgetary data and enabling the public to understand and utilise OGD related to budgetary data. Even beyond reducing barriers to OGD use, data intermediaries, including NPDIs, can contribute to significant societal challenges through the use of open data, such as helping researchers with the data they require for medical, societal or other fields (N.d., n.d.). Furthermore, by aggregating, simplifying, and promoting the use of OGD, NPDIs can also help governments and societies with global issues such as achieving the target of carbon emission reductions (N.d., n.d.).

Moreover, NPDIs and policymakers can use the results of this thesis to tap into the identified areas of opportunities, such as building strategic decisions to shift NPDIs efforts from being reactive to ultimately being proactive, meaning that they work more in an integrative approach with OGD providers and users to identify and address the issues upstream before they occur. Such a shift in their strategy will also address barriers related to available OGD and their significant contribution to the barriers related to the OGD user's ability.

8.4. Future research recommendations

We propose the following research directions based on our findings and the limitation of our study. Considering the complexity of the OGD ecosystem, future researchers may adopt a broader perspective when exploring the impact of NPDIs on OGD barriers

beyond the barriers to OGD use. Since our selected NPDIs did not focus mainly on barriers to OGD use but rather use OGD as an instrument, the broader perspective might provide valuable insights into how the NPDIs contribute to the OGD ecosystem in general. Moreover, future researchers could broaden the scope of their study by also considering the data providers. The data providers' perspective might reveal some effects of NPDIs activities that also benefit the providers. Furthermore, in this study, we could not link the activities of NPDIs to the barriers they reduce; this could be an opportunity for future researchers to identify which barriers are linked and reduced by NPDIs activities. Identifying the impact of NPDI's activities concerning OGD barriers can provide a better strategy for NPDIs adoption.

Our case study was limited and focused on multiple and diverse NPDIs with less depth for each case. Future researchers may focus on a single NPDI or a specific type of NPDI who are, for instance, dedicated to environmental data, educational data or transparency data. Examining specialised NPDIs can provide insights into the sectoral effects and challenges. Future researchers could also focus on investigating the impact of NPDIs on specific user groups of OGD, such as researchers or journalists, to understand how the NPDIs are adaptive to the specific need of the user groups under investigation. Future researchers may also explore the unintended or unforeseen consequences of NPDIs activities. Examining the NPDIs' activities might uncover unexpected negative or positive outcomes for the stakeholders, OGD users or the OGD ecosystem.

Moreover, for pragmatic reasons, our study had limitations with prioritising the NPDI's selection and the roles of their representatives. Future studies may address this limitation by having a wider list of potential cases with more purposeful selection and prioritisation strategies and identifying specific roles of the case study participants. Lastly, researchers could employ different research methods instead of focusing only on qualitative research methods, such as mixed methods research. Combined with the qualitative approach, the quantitative approach could produce more concrete and measurable evidence of the impact of the NPDIs' activities on OGD barriers.

References

About - Monithon. (N.d.). https://www.monithon.eu/about-english/

ANBI – Open State Foundation. (N.d.). https://openstate.eu/en/about/anbi-en/

Baumeister, R. F., & Leary, M. R. (1997). Writing narrative literature reviews. *Review of General Psychology*, *1*(3), 311–320. https://doi.org/10.1037//1089-2680.1.3.311

Benitez-Paez, F., Degbelo, A., Trilles, S., & Huerta, J. (2018). Roadblocks hindering the reuse of open Geodata in Colombia and Spain: A data user's perspective. *ISPRS International Journal of Geo-Information*, *7*(1). https://doi.org/10.3390/ijgi7010006

Boyatzis, R. E. (1998). Thematic Analysis and Code Development The Search for the Codable Moment. *Transforming Qualitative Information*, (800), 1–200.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706QP063OA

Brugger, J., Fraefel, M., Riedl, R., Fehr, H., Schöneck, D., & Weissbrod, C. S. (2016). Current barriers to open government data use and visualization by political intermediaries. *Proceedings of the 6th International Conference for E-Democracy and Open Government, CeDEM 2016*, 219–229. https://doi.org/10.1109/CeDEM.2016.18

Centers of Italy. (N.d.). https://centriditalia.it/home

Chan, M., Johnson, P. A., & Shookner, M. (2016). Assessing the Use of Government Open Data and the Role of Data Infomediaries The Case of Nova Scotia's Community Counts Program. *JeDEM - eJournal of eDemocracy and Open Government*, *8*(1), 1–27. https://doi.org/10.29379/jedem.v8i1.370

Chattapadhyay, S. (2014a). Access and Use of Government Data by Research and Advocacy Organisations in India: A Survey of (Potential) Open Data Ecosystem. *ACM International Conference Proceeding Series*, *2014-January*, 361–364. https://doi.org/10.1145/2691195.2691262

Chattapadhyay, S. (2014b). Access and use of government data by research and advocacy organisations in India: A survey of (potential) open data ecosystem. *ACM International Conference Proceeding Series*, *2014-January*, 361–364. https://doi.org/10.1145/2691195.2691262

Corbett, J., Templier, M., & Takeda, H. (2018). The Making of a 'Top' Open Data City: A Case Study of Edmonton's Open Data Initiative. *Proceedings of the Annual Hawaii International Conference on System Sciences*, *2018-January*, 2443–2452. https://doi.org/10.24251/HICSS.2018.308

Creswell, J. W. (2006). *Qualitative Inquiry and Research Design: Choosing among Five Approaches*. SAGE Publications Inc.

Creswell, J. W. (2009). *Research design: qualitative, quantitative, and mixed methods approaches*. SAGE Publications Inc.

Crusoe, J., & Melin, U. (2018). Investigating Open Government Data Barriers. https://doi.org/10.1007/978-3-319-98690-6{_}15

Crusoe, J., Simonofski, A., Clarinval, A., & Gebka, E. (2019). The impact of impediments on open government data use: insights from users. *2019 13th international conference on research challenges in information science (rcis)*, 1–12.

DAHBI, K. Y., Lamharhar, H., & Chiadmi, D. (2019). Toward a user-centered approach to enhance Data discoverability on Open Government Data portals. *2019 Third International Conference on Intelligent Computing in Data Sciences (ICDS)*, 1–5.

da Silva Craveiro, G., & Albano, C. (2017). Open data intermediaries: coproduction in budget transparency. *Transforming Government: People, Process and Policy*, *11*(1), 119–131. https://doi.org/10.1108/TG-12-2015-0057/FULL/PDF

da Silva Craveiro, G., & Albano, C. S. (2015). Budgetary Data (in an Open Format) Benefits, Advantages, Obstacles and Inhibitory Factors in the View of the Intermediaries of this System: A Study in Latin American Countries. https://doi.org/10.1007/978-3-319-25013-7{_}18

data noun - Definition, pictures, pronunciation and usage notes | Oxford Advanced Learner's Dictionary at OxfordLearnersDictionaries.com. (N.d.). https://www.oxfordlearnersdictionaries.com/definition/english/data?q=data

Davies, T., & Edwards, D. (2012). Emerging Implications of Open and Linked Data for Knowledge Sharing in Development. *IDS Bulletin*, *43*(5), 117–127. https://doi.org/10.1111/J.1759-5436.2012.00372.X

Davison, R. (2006). Implementing and Managing eGovernment. *The Electronic Journal of Information Systems in Developing Countries*, *23*(1), 1–1. https://doi.org/10.1002/j.1681-4835.2006.tb00154.x

Documentation - Openpolis. (N.d.). https://www.openpolis.it/openpolis-foundation/documentation/

Dumpawar, S. (2015). Open government data intermediaries : mediating data to drive changes in the built environment. https://dspace.mit.edu/handle/1721.1/97994

Educational path | At the School of Open Cohesion. (N.d.). https://www.ascuoladiopencoesione.it/en/node/11443

Enaholo, P. (2017). Beyond mere advocacy: CSOs and the role of intermediaries in Nigeria's open data ecosystem. https://doi.org/10.5281/ZENODO.1117797

Erete, S., Ryou, E., Smith, G., Fassett, K., & Duda, S. (2016). Storytelling with Data: Examining the use of data by Non-Profit organizations. *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*, *27*, 1273–1283. https://doi.org/10.1145/2818048.2820068

FAIR Principles - GO FAIR. (N.d.). https://www.go-fair.org/fair-principles/

Ferati, M., Dalipi, F., & Kastrati, Z. (2020). Open government data through the lens of universal design. *Universal Access in Human-Computer Interaction. Applications and Practice: 14th International Conference, UAHCI 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part II 22*, 331–340.

Fischer, G., & Herrmann, T. (2011). Socio-technical systems: a meta-design perspective. *International Journal of Sociotechnology and Knowledge Development (IJSKD)*, *3*(1), 1–33.

Frank, M., & Waddell, P. (2014). Solving the Democratic Deficit: The Role of Open Data and Intermediaries. *Conference for E-Democracy and Open Governement*, 487.

Gascó-Hernández, M., Martin, E. G., Reggi, L., Pyo, S., & Luna-Reyes, L. F. (2018). Promoting the use of open government data: Cases of training and engagement. *Government Information Quarterly*, *35*(2), 233–242. https://doi.org/10.1016/j.giq.2018.01.003

Gehman, J., Glaser, V. L., Eisenhardt, K. M., Gioia, D., Langley, A., & Corley, K. G. (2017). Finding Theory–Method Fit: A Comparison of Three Qualitative Approaches to Theory Building. *https://doi.org/10.1177/1056492617706029*, *27*(3), 284–300. https://doi.org/10.1177/1056492617706029

González-Zapata, F., & Heeks, R. (2015). *Understanding Multiple Roles of Intermediaries in Open Government Data The influence of political institutions and power on open government data (OGD) View project Digital Start-ups View project UNDERSTANDING MULTIPLE ROLES OF INTERMEDIARIES IN OPEN GOVERNMENT DATA* (tech. rep.). https://www.researchgate.net/publication/290446084

Green, B. N., Johnson, C. D., & Adams, A. (2006). Writing narrative literature reviews for peer-reviewed journals: secrets of the trade. *Journal of Chiropractic Medicine*, *5*(3), 101–117. https://doi.org/10.1016/S0899-3467(07)60142-6

Haan, D. d. (2018). Intermediation as a dissolver of barriers: How intermediaries can help overcome barriers in open data use.

Hannah Ritchie, Max Roser, & Pablo Rosado. (N.d.). CO and Greenhouse Gas Emissions - Our World in Data. https://ourworldindata.org/co2-and-greenhouse-gas-emissions

Harrison, H., Birks, M., Franklin, R., & Mills, J. (2017). Case study research: Foundations and methodological orientations. *Forum qualitative Sozialforschung/Forum: qualitative social research*, *18*(1), 1–17.

Herriott, R. E., & Firestone, W. A. (1983). Multisite Qualitative Policy Research: Optimizing Description and Generalizability. *Educational Researcher*, *12*(2), 14–19. https://doi.org/10.3102/0013189X012002014

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, *29*(4), 258–268. https://doi.org/10.1080/10580530.2012.716740

Juell-Skielse, G., Hjalmarsson, A., Juell-Skielse, E., Johannesson, P., & Rudmark, D. (2014). Contests as innovation intermediaries in open data markets. *Information Polity*, *19*(3-4), 247–262. https://doi.org/10.3233/IP-140346

Kidd, P. S., & Parshall, M. B. (2000). Getting the Focus and the Group: Enhancing Analytical Rigor in Focus Group Research. *Qualitative Health Research*, *10*(3), 293–308. https://doi.org/10.1177/104973200129118453

Krueger, R. A. (2014). *Focus groups: A practical guide for applied research*. Sage publications.

Kulk, S., & Van Loenen, B. (2012). Brave New Open Data World? *. *International Journal of Spatial Data Infrastructures Research*, *7*, 196–206. https://doi.org/10.2902/1725-0463.2012.07.art10

Lämmerhirt, D., Brandusescu, A., Domagała, N., & Enaholo, P. (2020). *Situating Open Data: Global Trends in Local Contexts*. African Minds.

Magalhaes, G., Roseira, C., & Strover, S. (2013). Open government data intermediaries: A terminology framework. *ACM International Conference Proceeding Series*, 330–333. https://doi.org/10.1145/2591888.2591947

Martin, C. (2014). Barriers to the Open Government Data Agenda: Taking a Multi-Level Perspective. *Policy & Internet*, *6*(3), 217–240. https://doi.org/10.1002/1944-2866.POI367

Max Roser. (N.d.). About - Our World in Data. https://ourworldindata.org/about

Miles, M. B., Huberman, A. M., & Saldana, J. (2013). Qualitative Data Analysis: A Methods Sourcebook (H. Salmon, Ed.). *SAGE Publications Ltd (CA)*.

Monithon report finder. (N.d.). https://reports.monithon.eu/

Mutuku, L., & Mahihu, C. (2014). A suggested framework for impactful open data applications in developing countries. *ACM International Conference Proceeding Series*, *2014-January*, 498–499. https://doi.org/10.1145/2691195.2691274

N.d. (N.d.). *Unlocking the value of data: Exploring the role of data intermediaries. An exploration of the role intermediaries could play in supporting responsible data sharing* (tech. rep.). Centre for Data Ethics and Innovation.

Nikiforova, A., & McBride, K. (2021). Open government data portal usability: A user-centred usability analysis of 41 open government data portals. *Telematics and Informatics*, *58*. https://doi.org/10.1016/j.tele.2020.101539

Nyumba, T., Wilson, K., Derrick, C. J., & Mukherjee, N. (2018). The use of focus group discussion methodology: Insights from two decades of application in conservation. *Methods in Ecology and Evolution*, *9*(1), 20–32. https://doi.org/10.1111/2041-210X.12860

O'Hara, K. (2012). Transparency, open data and trust in government: Shaping the infosphere. *Proceedings of the 4th Annual ACM Web Science Conference, WebSci'12*, *volume*, 223–232. https://doi.org/10.1145/2380718.2380747

Open Spending - Home. (N.d.). https://openspending.nl/

Openparliament | Openpolis Foundation. (N.d.). https://openpolis.odoo.com/en/cosa-facciamo/openparlamento

Our World in Data · GitHub. (N.d.). https://github.com/owid

Patton, M. Q. (1990). *Qualitative evaluation and research methods*. SAGE Publications, inc.

Pilshchikova, L., Zuiderwijk, A., & Janssen, M. (2022). How do Non-profit Open data Intermediaries enhance Open data Usability? A Systematic Literature Review. *The 18th International Symposium on Open Collaboration*, 1–3. https://doi.org/10.1145/3555051.3555061

Rabiee, F. (2004). Focus-group interview and data analysis. *Proceedings of the Nutrition Society*, *63*(4), 655–660. https://doi.org/10.1079/PNS2004399

Reis, J. R., Viterbo, J., & Bernardini, F. (2018). A rationale for data governance as an approach to tackle recurrent drawbacks in open data portals. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3209281.3209354

Research into the Open Government Act shows: a lot of action, but little result – Open State Foundation. (N.d.). https://openstate.eu/nl/2023/03/onderzoek-naar-wet-open-overheid-wijst-uit-veel-actie-maar-weinig-resultaat/

Rowley, J. (2002). Using case studies in research. *Management Research News*, *25*(1), 16–27. https://doi.org/10.1108/01409170210782990/FULL/PDF

Rufus Pollock. (N.d.). Building the (Open) Data Ecosystem – Open Knowledge Foundation blog. https://blog.okfn.org/2011/03/31/building-the-open-data-ecosystem/

Ruijer, E., & Meijer, A. (2020). Open Government Data as an Innovation Process: Lessons from a Living Lab Experiment. *Public Performance and Management Review*, *43*(3), 613–635. https://doi.org/10.1080/15309576.2019.1568884

Safarov, I., Meijer, A., & Grimmelikhuijsen, S. (2017). Utilization of open government data: A systematic literature review of types, conditions, effects and users. *Information Polity*, *22*, 1–24. https://doi.org/10.3233/IP-160012

Salamon, L. M., & Anheier, H. K. (1992). In search of the non-profit sector. I: The question of definitions. *Voluntas*, *3*(2), 125–151. https://doi.org/10.1007/BF01397770/METRICS

Saldaña, J. (2021). The coding manual for qualitative researchers. *The coding manual for qualitative researchers*, 1–440.

Saxena, S. (2018a). Drivers and barriers towards re-using open government data (OGD): a case study of open data initiative in Oman. *Foresight*, *20*(2), 206–218. https://doi.org/10.1108/FS-10-2017-0060

Saxena, S. (2018b). Drivers and barriers to re-use Open Government Data (OGD): a case study of open data initiative in Philippines. *Digital Policy, Regulation and Governance*, *20*(4), 358–368. https://doi.org/10.1108/DPRG-08-2017-0045

Saxena, S., & Muhammad, I. (2017). Barriers to use open government data in private sector and NGOs in Pakistan. *Information Discovery and Delivery*, *46*(1), 67–75. https://doi.org/10.1108/IDD-05-2017-0049

Saxena, S., & Muhammad, I. (2018). The impact of open government data on accountability and transparency. *Journal of Economic and Administrative Sciences*, *34*(3), 204–216. https://doi.org/10.1108/jeas-05-2017-0044

Seawright, J., & Gerring, J. (2008). Case Selection Techniques in Case Study Research. *Political Research Quarterly*, *61*(2), 294–308. https://doi.org/10.1177/1065912907313077

Shaharudin, A., Van Loenen, B., & Janssen, M. (2023). Towards a Common Definition of Open Data Intermediaries. *Digital Government: Research and Practice*. https://doi.org/10.1145/3585537

Skjott Linneberg, M., & Korsgaard, S. (2019). Coding qualitative data: a synthesis guiding the novice. *Qualitative Research Journal*, *19*(3), 259–270. https://doi.org/10.1108/QRJ-12-2018-0012/FULL/XML

Slechts kwart van alle gemeenten deelt verkiezingsuitslag als open data – Open State Foundation. (N.d.). https://openstate.eu/nl/2019/04/slechts-kwart-van-alle-gemeenten-deelt-verkiezingsuitslag-als-open-data/

Smith, G., & Sandberg, J. (2018). Barriers to innovating with open government data: Exploring experiences across service phases and user types. *Information Polity*, *23*(3), 249–265. https://doi.org/10.3233/IP-170045

Strauss, A., & Corbin, J. (1998). Basics of qualitative research techniques.

Subscribe to openpolis newsletters - Openpolis. (N.d.). https://www.openpolis.it/iscriviti-alle-newsletter-di-openpolis/

Tammisto, Y., & Lindman, J. (2012). Definition of open data services in software business. *Lecture Notes in Business Information Processing*, *114 LNBIP*, 297–303. https://doi.org/10.1007/978-3-642-30746-1{_}28/COVER

Teaching Hub - Our World in Data. (N.d.). https://ourworldindata.org/teaching

The benefits and value of open data | data.europa.eu. (N.d.). https://data.europa.eu/en/publications/datastories/benefits-and-value-open-data

The International Initiatives | A Scuola di OpenCoesione. (N.d.). https://www.ascuoladiopencoesione.it/en/ASOC-EU-project

The World Bank. (2019). Open Data Essentials | Data. http://opendatatoolkit.worldbank.org/en/essentials.html

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management*, *14*(3), 207–222. https://doi.org/10.1111/1467-8551.00375

UBUNTU | At the School of Open Cohesion. (N.d.). https://www.ascuoladiopencoesione.it/it/team/22-23-en_0012

Van Schalkwyk, F., Canares, M., Chattapdhyay, S., & Andrason, A. (2016). Open Data Intermediaries in Developing Countries. *The Journal of Community Informatics*, *12*(2). https://doi.org/10.15353/joci.v12i2.3219

van Veenstra, A., & van den Broek, T. (2015). A community-driven open data lifecycle model based on literature and practice. *Case studies in e-Government 2.0 - Changing Citizen relationships*, 183. https://repository.tno.nl/islandora/object/uuid%3A823f36f9-b307-4e43-aece-2f718757a175

Waar is mijn stemlokaal. (N.d.). https://waarismijnstemlokaal.nl/

Waldrop, M. M. (1993). *Complexity: The emerging science at the edge of order and chaos*. Simon; Schuster.

What we do | Openpolis Foundation. (N.d.). https://fondazione.openpolis.it/en/cosa-facciamo

Wieczorkowski, J. (2019). Barriers to using open government data. *ACM International Conference Proceeding Series*, 15–20. https://doi.org/10.1145/3340017.3340022

Xiao, F., He, D., Chi, Y., Jeng, W., & Tomer, C. (2019). Challenges and supports for accessing open government datasets: Data guide for better open data access and uses. *CHIIR 2019 - Proceedings of the 2019 Conference on Human Information Interaction and Retrieval*, 313–317. https://doi.org/10.1145/3295750.3298958

Yin, R. K. (2018). Case Study Research Design and Applications. *The Canadian Journal of Program Evaluation*, 319.

Yoon, A., Copeland, A., & McNally, P. J. (2018). Empowering communities with data: Role of data intermediaries for communities' data utilization. *Proceedings of the Association for Information Science and Technology*, *55*(1), 583–592. https://doi.org/10.1002/pra2.2018.14505501063

Zhang, H., & Xiao, J. (2020). Quality assessment framework for open government data: Meta-synthesis of qualitative research, 2009-2019. *Electronic Library*, *38*(2), 209–222. https://doi.org/10.1108/EL-06-2019-0145

Zuiderwijk, A., & de Reuver, M. (2021). Why open government data initiatives fail to achieve their objectives: categorizing and prioritizing barriers through a global survey. *Transforming Government: People, Process and Policy*, *15*(4), 377–395. https://doi.org/10.1108/TG-09-2020-0271

Zuiderwijk, A., Janssen, M., & Davis, C. (2014). Innovation with open data: Essential elements of open data ecosystems. *Information Polity*, *19*, 17–33. https://doi.org/10.3233/IP-140329

A
Appendices

A.1. Identifying the knowledge gap

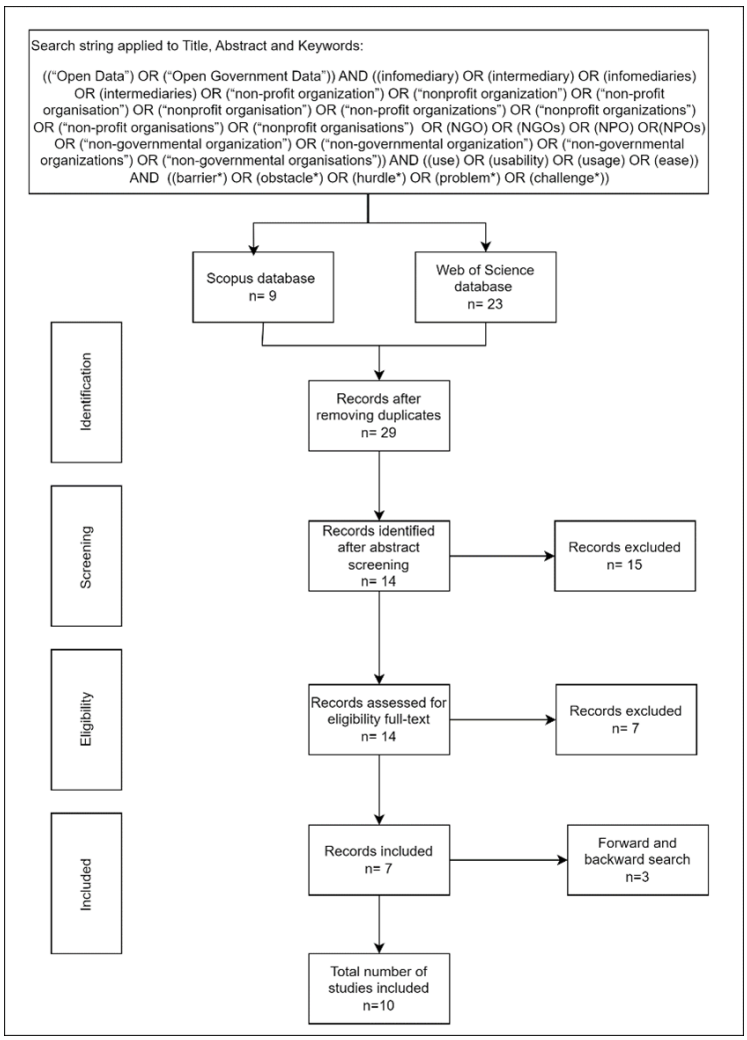

Figure A.1: Selection process for knowledge gap literature review

Table A.1: Overview of selected studies for knowledge gap

No.	Title	Author/s	Year
1	Access and Use of Government Data by Research and Advocacy Organisations in India: A Survey of (Potential) Open Data Ecosystem	(Chattapadhyay, 2014a)	2014
2	Barriers to use open government data in private sector and NGOs in Pakistan	(Saxena & Muhammad, 2017)	2017
3	A Suggested Framework for Impactful Open Data Applications in Developing Countries	(Mutuku & Mahihu, 2014)	2014
4	Budgetary Data (in an Open Format) Benefits, Advantages, Obstacles and Inhibitory Factors in the View of the Intermediaries of this System: A Study in Latin American Countries	(da Silva Craveiro & Albano, 2015)	2015
5	Current Barriers to Open Government Data Use and Visualization by Political Intermediaries	(Brugger et al., 2016)	2016
6	Storytelling with Data: Examining the Use of Data by Non-Profit Organizations	(Erete et al., 2016)	2016
7	The impact of open government data on accountability and transparency	(Saxena & Muhammad, 2018)	2018
8	Empowering Communities with Data: Role of Data Intermediaries for Communities' Data Utilization	(Yoon et al., 2018)	2018
9	Open Government Data Intermediaries: A Terminology Framework	(Magalhaes et al., 2013)	2013
10	Understanding Multiple Roles of intermediaries in Open Government Data	(González-Zapata & Heeks, 2015)	2015

A.2. Approval of HREC at TU Delft

Date 06-Apr-2023
Contact person Dr. Cath Cotton, Policy Advisor Academic Integrity
E-mail c.m.cotton@tudelft.nl

Human Research Ethics Committee TU Delft
(http://hrec.tudelft.nl)

Visiting address
Jaffalaan 5 (building 31)
2628 BX Delft

Postal address
P.O. Box 5015 2600 GA Delft
The Netherlands

Ethics Approval Application: The impact of the activities of Non-Profit data Intermediaries on reducing the barriers to using Open Government Data (OGD)
Applicant: Alorini, Rayyan

Dear Rayyan Alorini,

It is a pleasure to inform you that your application mentioned above has been approved.

In addition to any specific conditions or notes, the HREC provides the following standard advice to all applicants:
• In light of recent tax changes, we advise that you confirm any proposed remuneration of research subjects with your faculty contract manager before going ahead.
• Please make sure when you carry out your research that you confirm contemporary covid protocols with your faculty HSE advisor, and that ongoing covid risks and precautions are flagged in the informed consent - with particular attention to this where there are physically vulnerable (eg: elderly or with underlying conditions) participants involved.
• Our default advice is not to publish transcripts or transcript summaries, but to retain these privately for specific purposes/checking; and if they are to be made public then only if fully anonymised and the transcript/summary itself approved by participants for specific purpose.
• Where there are collaborating (including funding) partners, appropriate formal agreements including clarity on responsibilities, including data ownership, responsibilities and access, should be in place and that relevant aspects of such agreements (such as access to raw or other data) are clear in the Informed Consent.

Good luck with your research!

Sincerely,

A.3. Consent form

INFORMED CONSENT FORM

Study information:

You are invited to participate in this research study titled “The impact of the activities of Non-Profit data Intermediaries on reducing the barriers to using Open Government Data (OGD)”. This study is being done by Rayyan Alorini (MSc student at TU Delft) as part of his master’s thesis project.

This research aims to explore the impact of the activities conducted by the non-profit data intermediaries on reducing the barriers to using Open Government data in Europe. The interview will take you approximately 60 minutes to complete. The data will be used for data collected at the interview will be used as part of the case study analysis. The findings will be included in the master thesis report, which will be published in the TU Delft repository. We will ask you to discuss the impact of the activities of the NPO data intermediaries and how you think they reduce the barriers to using OGD.

As with any online activity, the risk of a breach is always possible. To the best of our ability, your answers in this study will remain confidential. We will minimise risks by securely storing all data containing personal information, including the interview recording, the transcribed interview script, the consent form and any data you provide us in a secure online environment. Access to this information will be restricted only to the project members (myself and my supervisors mentioned below).

Participation in this study is entirely voluntary, and you can withdraw anytime. You are free to omit any questions. If you would like to withdraw from the study, data obtained during the interview, including the recording or data sent by you later, will be deleted within one week of your request.

This study is supervised by:
Dr. A.M.G. (Anneke) Zuiderwijk- van Eijk (a.m.g.zuiderwijk-vaneijk@tudelft.nl)
Dr. M.L.C. (Mark) de Bruijne (M.L.C.deBruijne@tudelft.nl)
L. (Liubov) Pilshchikova (L.Pilshchikova@tudelft.nl)

2: Explicit Consent points

PLEASE TICK THE APPROPRIATE BOXES	Yes	No
A: GENERAL AGREEMENT – RESEARCH GOALS, PARTICPANT TASKS AND VOLUNTARY PARTICIPATION		
1. I have read and understood the study information dated [DD/ MM/ YYYY], or it has been read to me. I have been able to ask questions about the study, and my questions have been answered to my satisfaction.	☐	☐
2. I consent voluntarily to participate in this study and understand that I can refuse to answer questions and withdraw from the study at any time without having to give a reason.	☐	☐
3. I understand that participating in the study involves: interview video recording. I also understand that the interview will be transcribed as text during the study. The recording and the transcription will be deleted one year after the end of this project (August 2024).	☐	☐
4. I understand that the study (the master thesis project) will end by the end of August 2023	☐	☐

PLEASE TICK THE APPROPRIATE BOXES	Yes	No
B: POTENTIAL RISKS OF PARTICIPATING (INCLUDING DATA PROTECTION)		
5. I understand that taking part in the study also involves collecting specific personally identifiable information (PII) [such as names, email address, signature, title, years of experience and current organisation] and associated personally identifiable research data (PIRD) [such as the interview recording and the transcript] with the potential risk of my identity being revealed the case of a security breach to the secure project data storage environment; or in case of anonymisation was incomplete on the interview data that will be part of the thesis report.	☐	☐
6. I understand that the following steps will be taken to minimise the threat of a data breach and protect my identity in the event of such a breach all personally identifiable information (PII) and associated personally identifiable research data (PIRD) will be stored in a secured environment with restricted access to the research team (Rayyan Alorini, Anneke Zuiderwijk, Mark de Bruijne, and Liubov Pilshchikova); the recording and any collected information from the participants will be destroyed one year after the end of this project (by September 2024). The data collected from the interview will be anonymised entirely before being included in the master thesis report.	☐	☐
7. I understand that personal information collected about me that can identify me, such as my Name, signature, current organisation and email address, will not be shared beyond the study team.	☐	☐
C: RESEARCH PUBLICATION, DISSEMINATION AND APPLICATION		
8. I understand that after the research study, the de-identified information I provide will be used for the master thesis report for Rayyan Alorini.	☐	☐
9. I agree that my responses, views, or other input can be quoted anonymously in research outputs	☐	☐
D: (LONGTERM) DATA STORAGE, ACCESS AND REUSE		
10. I give permission for the de-identified information I provide to be archived in the TU Delft repository to be used for future research and learning.	☐	☐
11. I understand that access to the TU Delft repository is open.	☐	☐

Signatures

_____ _____ _____

Name of the participant [printed] Signature Date

Study contact details for further information:
Rayyan Alorini
+31620059814
r.m.s.alorini@student.tudelft.nl

A.4. Participation request email

Subject: Request to participate in a study - master student at TUDelft

Dear XXXXX,

My name is Rayyan Alorini, and I am a second-year student pursuing my master's degree in Complex Systems Engineering and Management at Delft University of Technology in the Netherlands. As part of my thesis, I am conducting a research study to explore the impact of the activities of non-profit data intermediaries on reducing the barriers to using open government data for users.

I am reaching out to your organization as I believe your expertise and experience in this area would make you a valuable participant in my research. The participation type would involve a one-hour interview, where I would discuss your organization's experience and perspectives on the topic with a representative from your team. The interview will be conducted online and recorded for transcription and analysis purposes.

Your participation in this research study would greatly contribute to my thesis and help me gain a deeper understanding of the impact of non-profit data intermediaries on reducing barriers to open government data use. The results of this study will contribute to the academic and practitioner communities, providing valuable insights into the role of non-profit data intermediaries in promoting open government data initiatives.

If your organization is interested in participating in this research study, please let me know your availability in the next two weeks. I will provide you with additional details and arrange a convenient time for the interview.

Thank you for your time and consideration. I look forward to hearing from you soon.

Best regards,

Rayyan Alorini

A.5. Interview protocol for NPDIs

Introduction:

Hello (interviewee name) and thank you for sparing the time for our interview. My name is Rayyan Alorini; I am conducting this research for my master's thesis project at TU Delft. The scope of the research is about the activities of non-profit data intermediaries involved with Open Government Data. My research aims to explore the impact of non-profit data intermediaries' activities in reducing the barriers to using OGD.

I will ask you around 20 questions in this interview, which will last approximately one hour. I would like to have your detailed answers, as they will be helpful for the interview analysis. I may ask irrelevant questions; please interrupt me anytime and ask for more details.

Recording:

As you noticed in the consent form, the interview will be recorded and transcribed using Microsoft Teams transcribe function. Without further ado, I would like to start the recording and the transcription.

**Ensure the interviewee agrees and signs the form before conducting the interview; ask to sign the form after the interview. If the interviewee is unwilling to sign the form, thank them and end the interview.*

Start the recording and transcription!

Participant Background Information:

1. Can you please tell me about yourself, what you do and your role at (the respective NPDI)?
2. Can you please tell us briefly about (the respective NPDI)?

Data intermediaries' activities:

4. What are the main goals of your activities as an organisation?
    a. How do you achieve these goals?
5. Who is your main target audience?
    b. Are there any specific reasons for addressing this target?
6. What are the open data sources you use?
7. How do you engage with your community to promote the use of open government data?
8. Can you describe any efforts your organisation has made to promote the use of open government data in specific sectors or industries?
9. In your opinion, which of the activities you provide are critical for your target audience of OGD? Why? Can you please give an example?

Barriers specific questions:

10. How do you think OGD is currently being promoted to be used?
11. In the way OGD is provided, what do you think is lacking in the provided data to be used?
12. What do you see as the most significant obstacles/challenges to using open government data? (For the users)
    i. Which of these obstacles/challenges are you addressing in (the respective NPDI)?
    ii. Can you please tell us more about the efforts you are making towards reducing these barriers?
13. What challenges did you encounter at (the respective NPDI) when trying to address these obstacles/challenges?
    iii. How did you overcome these challenges?
14. Can you provide examples of specific projects or initiatives inside (the respective NPDI) that have resulted in successfully reducing barriers to using open government data?
15. Do you assess the impact of your activities on using open data? How? Leading question if the impact is measured. Please provide me with some documents (reports, statistics, etc.)
16. Are there plans or initiatives at (the respective NPDI)to further work on the obstacles/challenges to using open government data?

Based on the discussion, the interviewee might be shown the barriers identified in the literature and asked to reflect on some of the relevant barriers to them.

Value creation:

17. To sum it up. how do you think your users are benefiting? In what way?
18. Beyond your users, where else do you see (the respective NPDI) creating an impact?
19. Can you share any reports or documents about measuring or quantifying the impact you create?

Closing:

20. Would you like to further discuss any activities, barriers of OGD or benefits created by (the respective NPDI) for OGD users?
21. I would like to talk to some of your users. Can you provide me with a short list of your users, which I can try to schedule for an interview?
22. Do you have any questions before we conclude the interview?

Thank you for your time; it was an insightful interview. Shortly, I will share with you the transcript of this interview. And if you would like to receive my final report, please let me know. I will conclude my research by the end of July, when I can share my final report. If you have any questions after the interview, please email me or my supervisors – their contacts are part of the consent form.

A.6. Interview protocol for Users

Introduction:

Hello (interviewee name) and thank you for sparing the time for our interview. My name is Rayyan Alorini; I am conducting this research for my master's thesis project at TU Delft. The scope of the research is about the activities of non-profit data intermediaries involved with Open Government Data. My research aims to explore the impact of non-profit data intermediaries' activities in reducing the barriers to using OGD.

I will ask you around 20 questions in this interview, which will last approximately one hour. I would like to have your detailed answers, as they will be helpful for the interview analysis. I may ask irrelevant questions; please interrupt me anytime and ask for more details.

Recording:

As you noticed in the consent form, the interview will be recorded and transcribed using Microsoft Teams transcribe function. Without further ado, I would like to start the recording and the transcription.

**Ensure the interviewee agrees and signs the form before conducting the interview; ask to sign the form after the interview. If the interviewee is unwilling to sign the form, thank them and end the interview.*

Start the recording and transcription!

Participant Background Information:

1. Can you please tell me about yourself, your occupation and what you do?

Before we delve deeper, let me define OGD as "interoperable data that public institutions share publicly for anybody to use and redistribute without restrictions".
And let me provide a clear definition of what we mean by data intermediaries "as an agent (i) positioned at some point in a data supply chain that incorporates an open dataset, (ii) positioned between two agents in the supply chain, and (iii) facilitates the use of open data that may otherwise not have been the case"

2. What kind of open data do you use?
3. Why do you need this data?
4. What do you do with this data?
5. If the interviewee represents an organisation, who is your main target audience?
    a. Are there any specific reasons for addressing this target?

User experience (barriers they face):

6. What type of obstacles/challenges do you face when using Open Government Data?

7. Follow-up questions to dive deep into the issues the users are facing.
8. Which of these obstacles/challenges do you consider significant? And why?

Based on the discussion, the interviewee might be shown the barriers identified in the literature and asked to reflect on some of the relevant barriers to them.

Impact of the NPOs activities:

8. Can you tell me what your relationship with (the respective NPDI), or other data intermediaries is?
9. How did you know about the data intermediaries?
10. What support or services does (the respective NPDI) or other data intermediaries provide to you?

    a. Does (the respective NPDI) provide tools or resources to help you use OGD? (this is Follow-up questions, asking for further details if needed)

11. How does this (the activities, services, or tools) impact you?
12. What would you have done differently without working/using the tools from (the respective NPDI)?
13. Could you provide an example of how (the respective NPDI) or other data intermediaries had an impact on you or your organisation to overcome obstacles/challenges to using OGD?
14. Are there plans or initiatives to further work with (the respective NPDI) or other NPDIs to overcome obstacles/challenges in using OGD?

Value creation:

15. What benefits, such as the production of good or helpful results or effects, do you think (the respective NPDI) or other non-profit data intermediaries create for you/your organisation?

Closing:

16. Would you like to further discuss any challenges of OGD, or benefits created by (the respective NPDI)?
17. Do you have any questions before we conclude the interview?

Thank you for your time; it was an insightful interview. Shortly, I will share with you the transcript of this interview. And if you would like to receive my final report, please let me know. I will conclude my research by the end of July, when I can share my final report. If you have any questions after the interview, please email me or my supervisors – their contacts are part of the consent form.

A.7. Focus group questions

Focus Group :

Validation session for NPDIs impact in reducing the barriers to use OGD

09/06/2023

TU Delft

To discuss and validate the findings of our research about the impact of Non-profit data intermediaries (NPDIs) in reducing the barriers to using OGD

TU Delft

2

OGD use Barriers

OGD Portal	1	Insufficient analysis tools for using open data (No data visualization tools, merging tools)
	2	Lack of user-friendly features such as, advanced search tools,
	3	Findability issues
	4	Availability and accessibility
	5	Diversity barriers *(different needs)*
Datasets	6	Limited quality of the datasets including data incompleteness and/or data format issues
	7	Lack of documentation on the available datasets
	8	Lack of data standardisation and interoperability
	9	Data currency
	10	Low dataset usability *(sample for large size data, reuse example)*
	11	Terms of use and licenses are not clear or do not foster reuse and distribution.
	12	Lack or absence of quality metadata
	13	Data credibility and accuracy
OGD Users	14	Lack of knowledge or capability to use and understand the data
	15	OGD users struggled to motivate, allocate time, and finance the required resources
	16	Unable to identify opportunities and usages

TU Delft

- Building OGD capacity and expertise
- Improving OGD accessibility, quality and usability
- Empowering OGD users
- Optimizing OGD processes
- OGD promotion and advocacy
- Improving stakeholders' collaboration and engagement

TU Delft

1- To what extent do you think the NPDI you are (involved with/using data of) are reducing the barriers to using OGD by building OGD capacity and expertise through:

- Improving OGD users' knowledge and capability
- Allowing others to build on their previous knowledge
- Offering diversity in backgrounds *(politicians/ academia)* and disciplines *(Data and social, journalism and ICT)* to OGD users
- Ensuring users follow a practical approach by adopting "hands-on experience"
- Expanding OGD user capacity to new user groups, such as students

TU Delft

Time limit: 11:45 5

2- To what extent do you think the NPDI you are (involved with/using data of) are reducing the barriers to using OGD by improving OGD accessibility, quality and usability through:

- Making OGD open and not only available
- Gathering, curating, integrating and creating data & knowledge
- Improving OGD quality
- Improving OGD findability
- Improving OGD metadata and documentation
- Tailoring the support to the user's needs *(goal-oriented / skills-oriented)*

TU Delft

Time limit: 11:53 6

3- To what extent do you think the NPDI you are (involved with/using data of) are reducing the barriers to using OGD by empowering OGD users through:

- Supervising and validating OGD use
- Allowing users to disseminate findings
- Identifying opportunities for OGD use
- Legal empowerment *(In requesting the data)*
- Adopting a white-box approach *(clarifying and sharing the inner workings)*

TU Delft

Time limit: 12:03 7

4- To what extent do you think the NPDI you are (involved with/using data of) are reducing the barriers to using OGD by optimising OGD processes through:

- Becoming *(being recognized)* and adopting best practices to use OGD
- Establishing the process of OGD use
- Enabling OGD user to self-organise
- Scaling up their target and scope rapidly and easily
- Embracing a structured approach to OGD use
- Ensuring long-term availability and sustainability

TU Delft

Time limit: 12:12 8

5- To what extent do you think the NPDI you are (involved with/using data of) are reducing the barriers to using OGD by OGD promotion and advocacy through:

- By NPDIs demanding data
- Promoting the use of OGD
- Raising stakeholders' awareness *(users, facilitators, politicians)*
- Implementing OGD-blended civic education
- Intertwining OGD and social objectives

TU Delft

Time limit: 12:18 9

6- To what extent do you think the NPDI you are (involved with/using data of) are reducing the barriers to using OGD by improving stakeholders' collaboration and engagement through:

- Creating a collaborative feedback loop
- Becoming a venue for interaction
- Fostering and embracing partnership *(with other stakeholders)*

TU Delft

Time limit: 12:22 10

Closing:

- Do you think that certain impacts may have been overlooked or that something important may have been missed in the findings?
- Is there any other aspects or topic that you would like to discuss or highlight that we haven't covered in our discussion?

TU Delft

11