Towards a Common Definition of Open Data Intermediaries

Shaharudin, Ashraf; van Loenen, B.; Janssen, M.F.W.H.A.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Towards a Common Definition of Open Data Intermediaries

ASHRAF SHAHARUDIN and BASTIAAN VAN LOENEN, Faculty of Architecture & the Built Environment, Delft University of Technology
MARIJN JANSSEN, Faculty of Technology, Policy & Management, Delft University of Technology

The role of open data intermediaries is considered instrumental in the supply and use of open data. There are various definitions of open data intermediaries in the literature and some of them are quite different from each other. These definitions can benefit from harmonization so knowledge about open data intermediaries can be developed on top of a shared understanding of what open data intermediaries mean. The objective of this article is to propose a common definition of open data intermediaries. We first carried out a systematic literature review and compiled the definitions of open data intermediaries from the literature. We found that each definition can be broken down into four basic components: (i) Who are the actors of open data intermediaries? (ii) What do they do? (iii) Where are they located in the open data lifecycle? and (iv) Why are they needed? We then conducted another round of data gathering and analysis to substantiate the four basic components. We proposed the following common definition of open data intermediaries: Third-party actors who provide specialized resources and capabilities to (i) enhance the supply, flow, and/or use of open data and/or (ii) strengthen the relationships among various open data stakeholders.

CCS Concepts: • **Information systems → Information systems applications** • **Human-centered computing → Collaborative and social computing** • **Social and professional topics → Computing and business**;

Additional Key Words and Phrases: Open data, intermediaries, infomediaries, definition

1 INTRODUCTION

Digitalization brings forth a great volume and range of data. Opening up data allows data to be re-used by various sectors, including businesses, researchers, and civil society groups, which can generate tremendous value for society. The International Open Data Charter defines open data as "digital data that is made available with the technical and legal characteristics necessary for it to be freely used, re-used, and redistributed by anyone, anytime, anywhere" [30] (see also Reference [45]). There are many benefits of open data discussed in the literature.

For example, from an economic point of view, open data is said to stimulate innovation and improve business processes, whereas, from a sociopolitical point of view, open data can enhance government accountability and empower citizens [31, 69]. Open data also has the potential to improve the reproducibility and dissemination of scientific research [43, 63].

Even though open data is available to anyone from anywhere at any time, there are many socio-technical impediments to the meaningful use of open data, such as lack of knowledge, confusion concerning data licenses, and the absence of the appropriate software to process data [71]. Hence, there is a need for open data intermediaries with certain skills and resources to make sense of open data for particular audiences [14]. Their role is considered instrumental not only in the use of open data but also in providing access to open data [10, 25], since there are also challenges faced by data providers [31]. Besides, open data intermediaries can establish crucial linking points between sub-segments of the open data community [39].

Various definitions of open data intermediaries can be found in the literature. Some of the definitions are quite different from each other, for example, definitions by References [32, 33] regard open data intermediaries as those who are actively involved in the processing of open data whereas the definition by Reference [50] considers them as those who connect community members with open data and do not necessarily process the data themselves. These definitions can benefit from harmonization to ensure a shared understanding of open data intermediaries among researchers and practitioners. The absence of a common definition of open data intermediaries may lead to a divergence of understanding of what they entail. A divergence may not necessarily be an issue if it is due to different research findings, opinions, or contexts, which can be debated and deliberated further, but it is avoidable if the divergence is simply because different groups refer to different things when they speak of open data intermediaries. Therefore, it is necessary to take stock of the existing definitions in the literature and synthesize them, so knowledge about open data intermediaries can be developed on top of a mutual and up-to-date understanding of what they mean. To date, no study has compiled the various definitions of open data intermediaries and harmonized them for a common definition. Hence, this article aims to fill the said gap. The objective of this article is to propose a common definition of open data intermediaries.

The organization of this article is as follows: In Section 2, we present the research methods employed, including data gathering and analysis methods. In Section 3, we present the findings from the first round of data gathering, of which the goal is to gather all definitions of open data intermediaries in the literature. We then broke down the definitions gathered into basic components and conducted another round of data gathering and analysis to substantiate and ascertain the basic components identified. We present the findings and analysis from the second round of data gathering and analysis in Section 4. In Section 5, we develop a common definition of open data intermediaries based on the findings from the two rounds of data gathering and analysis. In Section 6, we discuss the definition proposed. In Section 7, we conclude by discussing the contributions of this article.

2 RESEARCH METHODS

2.1 Overview

The objective of this article is to propose a common definition of open data intermediaries. To achieve the said objective, we conducted a **systematic literature review (SLR)**. We chose to employ SLR, as it is a robust method to gather what existing literature says about a certain topic [46]. It is more rigorous than a non-systematic literature review, as it involves following clearly defined protocols and transparent reporting, which allows replication [62].

We followed the eight steps of the SLR process by Reference [65]. First, we formulated the problem that we wanted to achieve from the SLR. In our case, we want to answer: What is the definition of open data intermediaries in the literature? Second, we developed and validated the review protocol. Third, we searched for the literature. Fourth, we screened for inclusion by reviewing the title and abstract. Fifth, we assessed the quality of each publication by reviewing the full text. Sixth, we extracted data from the literature. Seventh, we analyzed the data. Last, we reported the findings. The abovementioned steps are detailed in the subsequent sections.

2.2 Literature Search

We searched for relevant publications in four databases, namely, Scopus, **Web of Science (WoS)**, Google Scholar, and **Open Access Theses and Dissertations (OATD)**. Publications up to June 1, 2022, are included in the search (no start year was set). The search terms used were "open data intermediaries," "open data intermediary," "open data intermediation," "open data infomediaries," "open data infomediary," "open government data intermediaries," "open government data intermediary," "open government data intermediation," "open government data infomediaries," and "open government data infomediary." We included the term "infomediaries" and "infomediary" in our searches, as our initial literature scanning shows that the term is often used as a synonym for data intermediary. In addition, we included the term "intermediation" to capture literature that uses the said term instead of "intermediary"; while linguistically, the former is a participle whereas the latter is a subject, both terms would point to the equally relevant literature as far as our article is concerned. Although the scope of our article is not limited to open government data but open data as a whole, we also included the term "open government data," since the academic sub-area of open government data has gained tremendous interest over the years, resulting in much literature in this area.

The search strategy for each database, including the search query and the number of publications found, is shown in Table A.1 in the Appendix. Note that Google Scholar only allows terms searched either in the title or in the whole publication. Because the latter gives an unmanageable number of publications, which is about 16,900 publications, we searched only in the title for Google Scholar, whereas in the title and abstract for the other databases.

2.3 Literature Filtering

In total, there are 176 publications compiled from the four databases searched. We removed duplicated publications, publications with no authors' information, and inaccessible publications in the first filtering stage, giving us 101 publications. We then removed irrelevant publications (publications that are not about open data) based on the title and abstract and publications in a non-English language, leaving us with 59 publications. Based on the content of each publication, 9 of them are found irrelevant to the objective of this article: 7 publications do not describe anything informational about open data intermediaries except referring to them in passing, 1 publication is an engineering article about a novel method to integrate information from multiple systems, and 1 publication is a two-page conference paper with the research method vaguely described and the findings section of only one paragraph. In the end, 50 publications were selected.[1] Figure 1 visualizes our literature filtering stages.

Most of the publications (23) are journal articles, 13 are conference papers, 4 are book chapters, 7 are dissertations, 2 are reports, and 1 is a working paper. Almost all of those publications employ qualitative methods except 1 that uses a quantitative method and 4 that employ a combination of qualitative and quantitative methods. The earliest publication is from 2011, and the largest share of publications in the pool (12) is from 2017.

2.4 Data Gathering and Analysis

We conducted inductive coding to gather data from the literature. Inductive coding allows "research findings to emerge from frequent, dominant or significant themes inherent in raw data" [61, p. 238]. With inductive coding, we did not start with preconceived ideas about open data intermediaries. Instead, we coded based on what is written in the literature. There are five key features of inductive coding [61], namely, (i) the code that is tagged to the raw text, (ii) the code description, (iii) text or data associated with the code, (iv) links between codes (in our case, we captured links across different publications to see how they cross-reference each other), and (v) the

---

[1]Note that we only count References [52] and [53] as one, since the former, which is a non-academic report, is republished, as the latter in an academic journal.
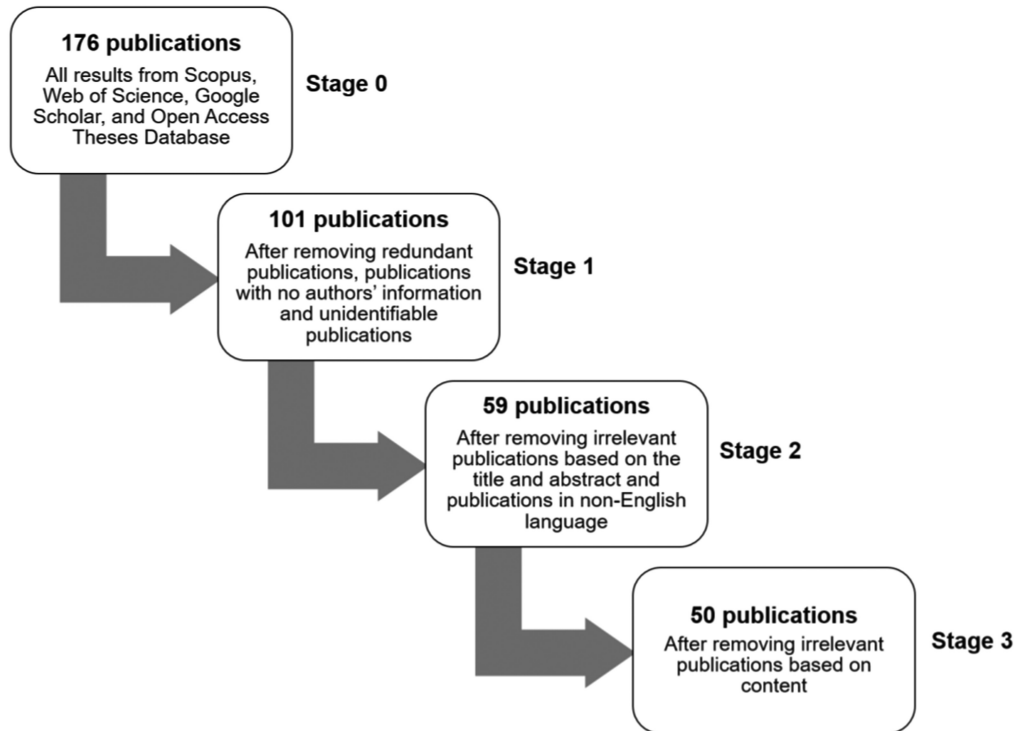
Fig. 1. Filtering stages of literature.

type of model in which the code is embedded (in our case, we took note of the type of research method and context of each publication).

To develop a common definition of open data intermediaries, we conducted two rounds of data gathering and analysis. In the first round, we gathered existing definitions from the literature pool. We then analyzed those definitions by breaking them down into basic components. We then identified the appropriate data that capture each of the basic components and conducted a second round of data gathering and analysis based on the same literature pool. The goal of this second round is to substantiate and ascertain the best description for each of the basic components of the definitions from the first round. Finally, we stitched together the most appropriate description for each of the basic components to produce a common definition of open data intermediaries.

3 FINDINGS: DEFINITIONS OF OPEN DATA INTERMEDIARIES IN THE LITERATURE

3.1 Compilation of Definitions

There are 12 definitions of open data intermediaries found in the literature surveyed. Table A.2 in the Appendix shows the list of the definitions, their source, and publications that adopted or were inspired by the respective definition. "Adopt" here means that the publication follows entirely the definition provided by the source publication, whereas "inspired by" means that the publication builds on the definition in the source publication to propose a new definition.

Even though all of the 50 publications reviewed discuss open data intermediaries, we were careful not to assume that every publication attempts to define open data intermediaries simply based on what the publication associates to them. This is to avoid misrepresenting the viewpoints of the authors by taking things out of context.

Therefore, unless the publication explicitly says something along the line of "open data intermediaries are…" or "open data intermediaries are defined as ….," we did not take them as attempts to define open data intermediaries.

The first attempt to define open data intermediaries was made in 2014 by Reference [10, p. 362], which defined them as "organisations that share data for its access, consumption and re-usage (including re-sharing) by other organisations and individuals." The author further clarified three points, namely, (i) "sharing of open data by such organisations can either be done on a commercial or a non-commercial basis"; (i) "shared data can either be primary (collected by the organisation concerned) or secondary (sourced from an external creator) in nature"; and (iii) "the data intermediary organisation may or may not add value to the data before sharing it further" [10, p. 362]. Reference [17, p. 96] built on Reference [10] to define open data intermediaries as "those who operate within the open data ecosystem by means of their contribution, in one way or the other, to the supply of open data by governments as well as to the demand for such data by citizens," which goes beyond sharing data as in Reference [10].

In 2015, Reference [57, p. 226] defined open data intermediaries as "all the players (in an individual way or representatives of governments and social organizations), who are involved with public data that are released in an open format. They may or may not make use of technological, legal or structural artifacts in their activities. In making use of open data, the intermediaries aggregate value to the data to ensure that they can be understood more easily (and hence have a greater value) [by] third parties after their intervention." Meanwhile, Reference [25, p. 4] defined open government data intermediaries as "all actors that assist OGD [open government data] initiatives by bridging the barriers that separate public sector data producers and civil society data consumers." They emphasized that open government data intermediaries have a two-way relationship, with the government on the supply side and with the civil society on the demand side.

Reference [53, p. 7] defined an open data intermediary 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." Reference [53] noted that an intermediary may neither supply nor access open data but facilitates the flow of data. To distinguish open data intermediaries from internet intermediaries such as internet service providers and cyber cafes, Reference [53] emphasized the degree of "agency" of actors in fulfilling the function of intermediating open data as the differentiating factor. In this regard, according to them, internet service providers and cyber cafes are not considered open data intermediaries, as they do not execute a high degree of involvement in intermediating open data. Note that in the following year, Reference [53] was republished as Reference [52]; while the former is a report of a project funded by the World Wide Web Foundation and Canada's International Development Research Centre, the latter is an article in an academic journal.

In our literature pool, six publications adopted the definition offered by Reference [53] or Reference [52], namely, References [1, 18, 26, 38, 56, 67]. Interestingly, in the same year Reference [53] was published, da Silva Craveiro & Albano in Reference [57] came up with their definition of open data intermediaries but later in Reference [56], they adopted the definition by Reference [53] instead of reiterating their own definition.

According to Reference [6, p. 222], open *government* data intermediaries are "actors who bridge gaps between data producers (governments) and data users (civil society) in that they supply essential resources and capabilities necessary to turn government data into development actions and results." Reference [41, p. xi] defined them as "actor[s] that bridge the gap between marginalized groups and OGD [open government data] by facilitating physical access, technical capacity, and value for use of information," whereas Reference [55, p. 2] defined them as "actors that translate, use, or otherwise mediate communication using data produced by or for government." Meanwhile, Reference [4, p. 133] defined them as "the in-between actor standing between a government and a citizen in the process of data communication."

A term that is often used as a synonym to open data intermediaries is infomediaries. Reference [32, p. 695] considered infomediaries as those involved in "the handling of information between information providers and consumers." This definition was adopted by Reference [51]. Reference [33, p. 10] defined infomediaries as

"specific categories of open data users who extract, aggregate, and transform data, altering it into a format that is seen as valuable, beneficial, and, most importantly, usable to the general public." Reference [21] adopted the definition of infomediaries by Reference [33]. Meanwhile, Reference [50, p. 31] defined a *civic* infomediary as "a person or organization that connects community members with open data so that public value can be derived from the data."

Based on our compilation, it can be seen that some of the definitions are rather different from each other and may result in conceptual confusion about open data intermediaries. For example, while definitions by References [32, 33] consider open data intermediaries to be actively involved in the processing of open data, Reference [50] defined them as those who connect community members with open data. Another aspect, while the definitions by References [33, 53, 57] highlight their function in the use of open data, the definition by Reference [17] highlights their function in the supply and demand of open data.

3.2 Breakdown of the Definitions

Inspired by the 5W1H questions method (what, who, where, when, why, and how), derived from the *Septem Circumstantiae* (elements of circumstances) from the field of philosophy [58], we found that the elements in the 12 definitions gathered from the literature can be categorized into the who, what, where, and why, that we call basic components (see Table 1). Specifically,

  (1) The who: Who are the actors of open data intermediaries?
  (2) The what: What do open data intermediaries do?
  (3) The where: Where are open data intermediaries located in the open data lifecycle?
  (4) The why: Why are open data intermediaries needed?

For *the where*, we followed the open data lifecycle model introduced by Reference [64]. Open data lifecycle is "a conceptualization of the process and practices around handling data, starting from its creation, through the provision of open data to its use by various parties" [9, p. 12]. While there are several open data lifecycle models in the literature such as References [2, 9, 59], we chose to follow the model by Reference [64], because it concisely integrates the activities of both data providers and data users in one lifecycle, instead of separate lifecycles, unlike most of the other data lifecycles in the literature. The model is developed based on synthesizing different open data lifecycle models in the literature and validating and detailing it through a case study of the **Netherlands Organization for Applied Scientific Research (TNO)**. There are five stages in the open data lifecycle model by Reference [64], namely, (i) identification: setting the open data strategy and selecting the data; (ii) preparation: setting requirements for data publication, modeling and describing data, converting data to a machine-readable format, linking data, and storing data; (iii) publication: publication of data and metadata; (iv) re-use: exploiting published data; and (v) evaluation: assessing the value of open data and monitoring and improving data [64].

Naturally, based on the 5W1H, one may ask, do the definitions not describe *the when* and *the how*? *The when*, which one may likely put as "when do open data intermediaries carry out their tasks?" is similar to *the where*, which is, "where are open data intermediaries located in the open data lifecycle?" Meanwhile, from the definitions compiled, it is rather difficult to differentiate *the how*, which one may likely put as "how do open data intermediaries do what they do?" from *the what* which is, "what do open data intermediaries do?." For the said reasons, in our analysis of the definitions, *the when* is equivalent to *the where,* and *the how* is equivalent to *the what.*

Note that care is needed when comparing the components across definitions, as five of the definitions [4, 6, 25, 41, 55] are specific for open government data intermediaries. While these five definitions are still pertinent for our article, we need to acknowledge that they are for the specific context of governments as data providers. Meanwhile, one definition [50] is for civic infomediary, which is specific to the context of open data for civic value.

Table 1. The Breakdown of Open Data Intermediaries' Definitions Gathered from the Literature

No.	Source	The who	The what	The where	The why
1.	[10]	organizations	share data for its access, consumption, and re-usage (including re-sharing) by other organizations and individuals		
2.	[32]		the handling of information	between information providers and consumers	
3.	[57]	all the players (in an individual way or representatives of governments and social organizations), who are involved with public data that are released in an open format	they [i.e., open data intermediaries] may or may not make use of technological, legal, or structural artifacts in their activities in making use of open data, the intermediaries aggregate value to the data		to ensure that they [i.e., data] can be understood more easily (and hence have a greater value) [by] third parties after their intervention
4.	[25]	all actors	assist OGD [open government data] initiatives		bridging the barriers that separate public sector data producers and civil society data consumers
5.	[52, 53]	an agent		positioned at some point in a data supply chain that incorporates an open dataset positioned between two agents in the supply chain	facilitates the use of open data that may otherwise not have been the case
6.	[6]	actors	bridge gaps they supply essential resources and capabilities necessary	between data producers (governments) and data users (civil society)	to turn government data into development actions and results
7.	[41]	an actor	bridges the gap by facilitating physical access, technical capacity, and value for use of information	between marginalized groups and OGD [open government data]	
8.	[55]	extra-institutional actors	translate, use, or otherwise mediate communication using data produced by or for government		
9.	[17]	those who operate within the open data ecosystem			by means of their contribution, in one way or the other, to the supply of open data by governments as well as to the demand for such data by citizens
10.	[33]	specific categories of open data users	extract, aggregate, and transform data		altering it [i.e., data] into a format that is seen as valuable, beneficial, and, most important, usable to the general public
11.	[50]	a person or organization	connects community members with open data		so public value can be derived from the data
12.	[4]	the in-between actor		standing between a government and a citizen in the process of data communication	

Note: Definitions are taken in verbatim from the source, but are arranged based on the four basic components.
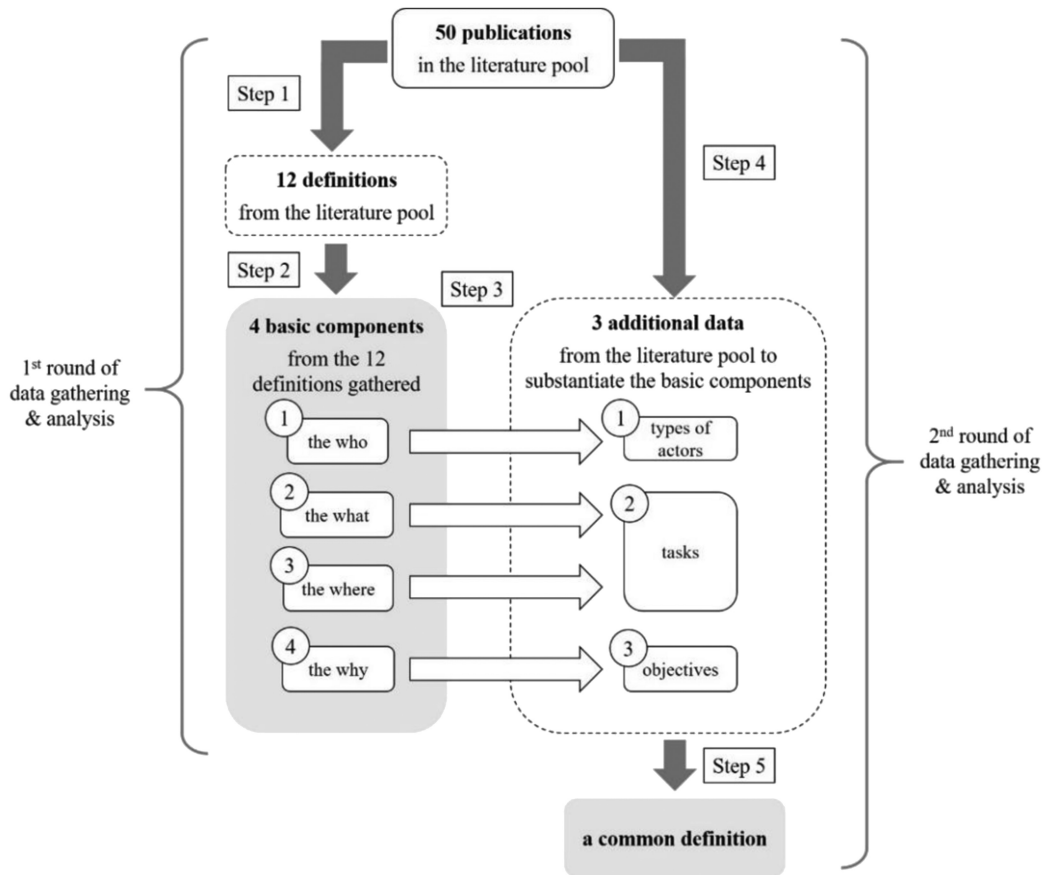
Fig. 2. The development of a common definition based on two rounds of data gathering and analysis.

4 FINDINGS: BASIC COMPONENTS OF THE DEFINITIONS OF OPEN DATA INTERMEDIARIES

4.1 Overview

As the objective of this article is to develop a common definition of open data intermediaries, we investigated further the basic components of the definitions compiled from the literature (i.e., *the who, the what, the where*, and *the why*). This is to substantiate and ascertain the best description for each component and develop a definition of open data intermediaries that is encompassing, guided by the body of literature.

We conducted a second round of data gathering and analysis from the same literature pool. *The who* is captured by looking into the types of actors of open data intermediaries, *the what* and *the where* are captured by looking into their tasks, and *the why* is captured by looking into their objectives. In short, we gathered three more data (types of actors, tasks, objectives) from the 50 publications in the literature pool to substantiate the four basic components of open data intermediaries' definitions. Figure 2 visualizes the steps.

In the following subsections, we describe (i) the types of actors of open data intermediaries (Section 4.2) to answer *the who*, (ii) the tasks of open data intermediaries (Section 4.3) to answer *the what* and *the where*, and (iii) the objectives of open data intermediaries (Section 4.4) to answer *the why*, based on our systematic literature review.

Table 2. Types of Actors of Open Data Intermediaries from the Literature Surveyed

No.	Type of actor	Description	Source
1.	Civil society organizations	Non-profit and non-governmental organizations that offer services for citizens with open data and/or advocate for open data for civic purposes	[6, 7, 17, 18, 23–25, 27, 33, 35, 39, 40–42, 44, 49, 51, 70]
2.	Entrepreneurs/ businesses	For-profit entrepreneurs and private companies that develop products and services for others (e.g., citizens and other businesses) with open data	[1, 7, 13, 23, 24, 27, 32, 33, 35, 51]
3.	Media	Media organizations and individual journalists who use open data to produce news stories	[3, 6, 7, 13, 18, 27, 33, 41, 42, 49, 70]
4.	Public organizations	Public organizations, including public libraries and public research institutes that help citizens to make use of open data	[8, 27, 32, 33, 36, 40, 49, 50]
5.	Researchers	Researchers in universities or research institutes who use open data to conduct research for the benefit of citizens	[13, 27, 33, 36, 40–42, 44, 48]
6.	Multi-partner	Collaborative organizations that are made up of different types of organizations such as public-private partnerships that use open data	[28, 40]

4.2 Types of Actors: *The Who*

4.2.1 Compilation of Types of Actors. There are six types of actors of open data intermediaries identified in the literature surveyed. Table 2 shows the list of the types of actors of open data intermediaries found and the publications that mentioned each type of actor.

The literature described **civil society organizations (CSOs)** as one type of actor of open data intermediaries. Some publications described civic app/platform developers as open data intermediaries CSOs [23, 25, 33, 35, 40, 42, 51], whereas another group of publications identified certain advocacy groups as open data intermediaries CSOs [6, 7, 17, 18].

Entrepreneurs and private businesses are another type of open data intermediaries actor. Based on the literature, entrepreneurs and businesses use open data to develop services in the form of web or mobile applications [23, 35, 51] as well as in the form of advisory services [1].

Several publications described media as another type of actor of open data intermediaries. In some, they are called data journalists [18, 33]. According to Reference [18], the resources and competencies provided by data journalists that act as open data intermediaries mainly involve transforming open data into digestible information in the form of news stories.

Another type of actor of open data intermediaries found in the literature are public organizations such as governments [8, 33, 40, 49], public libraries [50], and public research institutes [40]. Several publications, such as References [33, 49], found that some public organizations play the role of both open data providers and intermediaries.

Other types of actors of open data intermediaries are researchers [13, 33, 40, 42, 48] and multi-partner open data intermediaries such as public-private partnership organizations [40] and living labs made up of universities, development agencies, and private companies [28].

4.2.2 The Who: Who Are the Actors of Open Data Intermediaries? Based on the literature, there are various types of actors of open data intermediaries. They are not necessarily organizations, as some of them are individuals such as entrepreneurs, individual journalists, and researchers. There are also multi-partner open data intermediaries. Although they are mostly users of open data, certain literature also identified those who advocate for open data or facilitate access to open data as open data intermediaries.

Therefore, to capture the multifaceted types of actors of open data intermediaries, References [25, 50, 53] refer to *the who* in their open data definitions generically as "actors," "an agent," or "a person or organization."

Table 3. Tasks of Open Data Intermediaries from the Literature Surveyed

No.	Task	Description	Source
1.	Compile data	Gather data from multiple sources and in multiple formats	[1, 8, 13, 15, 18, 25–27, 36, 40–42, 44, 49, 53, 67, 68, 70]
2.	Build data capacity	Organize training sessions, workshops, hackathons, and other open data-related events as a third party as well as engage with stakeholders to improve open data practices	[14, 17, 38, 42, 49, 50, 56, 67]
3.	Augment data	Enhance the value of data by integrating open and/or private data from different sources	[1, 13–15, 21, 26, 67, 68]
4.	Contextualize data	Add relevant and specific context to the data for it to be relatable and meaningful to the targeted audience	[6, 15, 18, 25, 26, 41, 44, 56]
5.	Curate data	Select and reorganize data based on what is relevant and needed for the targeted audience	[1, 8, 14, 15, 21, 26, 67]
6.	Develop products and services	Use open data to offer products and services such as web-based and mobile applications and advisory services	[1, 8, 13, 18, 22, 25–27, 36, 41, 44, 66, 67]
7.	Interpret data	Turn data into more digestible information for the targeted audience	[13, 15, 17, 26, 36, 41, 44, 53, 67, 70]
8.	Validate data	Checking, updating, and rectifying data in terms of its accuracy, completeness, and timeliness	[13, 15, 18, 25, 26, 36]
9.	Demand open data	Identify datasets that should be opened and/or advocate for the adoption of open data policy	[13, 17, 18, 25, 36, 42]
10.	Visualize data	Represent data in charts, maps, and other visual forms for it to be more comprehensible to the targeted audience	[6, 15, 17, 18, 26, 41, 44]
11.	Facilitate stakeholders' interactions	Connecting stakeholders either through direct engagement of open data-related events	[8, 15, 26, 34, 36, 41]
12.	Channel feedback	Channel feedback regarding data or issues identified based on data to the relevant stakeholders	[8, 17, 26, 27, 44]
13.	Improve technical openness of data	Enhance the technical openness of public data such as by converting it into a machine-readable format	[26, 38, 41, 42, 44]
14.	Identify risks of opening data	Assist data providers in identifying potential risks in opening particular datasets	[14]

Nevertheless, based on our analysis, open data intermediaries mostly take care of the interest of other open data actors, distinguishing them from solely open data providers or end-users. For example, open data intermediaries process open data to deliver products and services benefitting other open data actors, and not only for their own internal benefit, making them different from merely open data end-users. Similarly, open data intermediaries that facilitate access to open data support users to gain access to data, but they are not necessarily the data providers themselves. Therefore, *the who* in defining open data intermediaries can be more accurately described as "third-party actors." According to Oxford Dictionary, a third party is "a person who is involved in a situation in addition to the two main people involved" [47]. Open data intermediaries can be the actor between open data providers and users [1, 42, 68] or between open data users and other users [28, 50].

4.3 Tasks of Open Data Intermediaries: *The What* and *the Where*

4.3.1 Compilation of Tasks. There are 14 tasks of open data intermediaries found in the literature surveyed. Table 3 shows the list of the tasks and the publications that described the respective task. Typically, multiple tasks are needed for them to serve their functions.

One of the most popular tasks of open data intermediaries is compiling data from various sources. They compile data from multiple sources and publish the data on their platforms and/or use the data to offer a product

or a service for end-users. For example, Aclímate Colombia compiles data on commercial crops, station-level daily weather data, and data related to crop yield from various open data sources and makes it all conveniently downloadable through its one-stop platform [68].

Another task of open data intermediaries is building the data capacity of data end-users (e.g., citizens, community organizations, journalists) as well as data providers. For example, BudgIT, a **civil society organization (CSO)** in Nigeria, organizes training sessions and workshops for journalists and individuals on engaging with open data [17]. Some others organize hackathons and other open data-related events [50, 56]. Several others work closely with government officials to improve the ways government data is made open [42].

Open data intermediaries also augment data by integrating open data from multiple sources and open data with private data. For example, Farmerline, a company based in Accra, Ghana, combines open meteorological data sourced from the Ministry of Food and Agriculture with the data the company collects (e.g., weekly market prices) and the data it purchases from non-open sources (e.g., weather forecast data) to provide advisory services to farmers [1]. Another example is Geonext, a company specializing in geodetic engineering, which produces its own data by conducting additional measurements and combines it with open data to enhance the services it offers to clients [26].

Contextualizing data is another task of open data intermediaries, which involves adding relevant and specific context to the data for it to be relatable and meaningful to a targeted audience. In the literature, this task is often carried out by journalists and CSOs. For example, CSOs in Argentina, Mexico, and Uruguay who champion budget transparency translate budget data into spending stories for it to be easily digestible by the public [56]. Meanwhile, journalists use open data to write news reports for public consumption [6, 18].

Another task of open data intermediaries is curating data. The abundance of open data that potential users may need to sift through may put off some of them from meaningfully using it [67]. Hence, this is where some open data intermediaries step in, for example, by curating datasets based on smaller geographies such as neighborhoods [67] and based on communities' needs [8].

Several open data intermediaries develop products and services, typically in the form of web and mobile applications. For example, an open data intermediary in the city of Edmonton, Canada, combines the city's open data and data from other sources to develop an application that simplifies the process of finding a home [13]. Another example is a company named IntellinQ, which builds a spatial database management software called GeollinQ to help users to access open data from various sources and reorganize it in a standard way [26]. There are also CSOs that develop data tools to facilitate community organizations to aggregate and integrate data that is of their interest [67]. Some open data intermediaries offer advisory services; for example, Farmerline and Esoko, which use open data to provide advice to farmers in Ghana [1].

Certain open data intermediaries interpret data to turn it into accessible information for their audience. CSOs interpret complex data into digestible information for citizens in key socio-economic areas such as health and education as well as using the interpreted data for advocacy work [17]. A related task to interpreting data is visualizing data of which open data is represented in charts, maps, and other visual forms [15, 18, 41].

There are open data intermediaries that validate the quality of open data. According to Reference [15], this involves addressing the "inaccuracy, incompleteness and obsolescence" of data by "validating, updating, and correcting" the data. Reference [26] identifies four companies in their study that assess the quality of open data, and in some cases rectify it, for their clients.

Some open data intermediaries also demand open data. This includes identifying specific datasets that should be made open based on local needs [13, 25] and advocating for open data policies to be adopted in general [18]. WhatDoTheyKnow web platform, developed by an open data intermediary in Hong Kong, allows users to request data disclosure from the relevant government authority and track the response [42].

Other tasks of open data intermediaries found in the literature are facilitating stakeholders' interactions either through direct networking or open data events [8, 15, 26, 34], channeling feedback regarding data or issues identified based on data to the relevant stakeholders [8, 17, 26], improving technical openness of data such as by

Table 4. Tasks of Open Data Intermediaries Based on the Open Data Lifecycle Stages

No.	Stage of open data lifecycle [64]	Tasks of open data intermediaries from the literature review
1.	Identification	demand open data, facilitate stakeholders' interactions, identify risks of opening data
2.	Preparation	compile data, build data capacity, augment data
3.	Publication	curate data, improve technical openness of data
4.	Re-use	build data capacity, contextualize data, develop products and services, interpret data, visualize data, facilitate stakeholders' interactions
5.	Evaluation	validate data, channel feedback

turning public data into a machine-readable format [26, 38, 42, 44], and assisting data providers in identifying potential risks in opening particular data [14].

4.3.2 The What: What Do Open Data Intermediaries Do? Based on the literature, open data intermediaries carry out a wide range of tasks, deploying various types of resources and capabilities. Most of the tasks involve active processing of open data such as collecting, augmenting, contextualizing, visualizing data, and developing products and services with open data. However, some tasks do not necessarily require them to actively process open data, for example, building data capacity, facilitating stakeholders' interactions, and channeling feedback.

In this regard, it is a challenge to describe *the what* in defining open data intermediaries without potentially excluding their certain tasks, an issue several existing definitions grapple with. As summarized by Reference [53], open data intermediaries deploy economic capital (e.g., financial resources), cultural capital (e.g., knowledge of local custom), social capital (e.g., existing networking with other stakeholders), symbolic capital (e.g., well-regarded position in society), and/or technical capital (e.g., data processing skills). Different open data intermediaries offer different types of resources and capabilities according to their specialization. These resources and capabilities are not at the disposal of most lay users. Lay users are "users who have limited or no training in a particular area, however, they are likely to have personal interests or special needs in that area" [11, p. 4]. Therefore, inspired by Reference [6] with a modification, we believe that *the what* in defining open data intermediaries can be concisely described as "provide specialized resources and capabilities."

4.3.3 The Where: Where Are Open Data Intermediaries Located in the Open Data Lifecycle? Based on the open data lifecycle stages of Reference [64] as described in Section 3.2, we locate the different tasks of open data intermediaries found in the literature (Table 4). It is clear that they carry out tasks at various stages of the open data lifecycle. Two tasks, facilitating stakeholders' interactions and building data capacity, fall in multiple stages.

For the said reason, we agree with the approach taken by References [10, 17, 25, 33, 50, 55, 56] which are agnostic about *the where* in defining open data intermediaries. Although one may also consider the approach by Reference [53] which describes *the where* as "positioned at some point in a data supply chain," it is not the most concise approach for a definition, because the term "data supply chain" itself may need to be further clarified. Likewise, if we were to use "open data lifecycle" in our proposed common definition, then one might question which model of open data lifecycle are we referring to. Hence, we believe that the better strategy is to silence *the where,* since silencing it does not change the essence of the definition, whereas the alternative, which is to include "positioned at any stage of the open data lifecycle," may raise more questions and render the proposed common definition to be less succinct.

4.4 Objectives of Open Data Intermediaries: *The Why*

4.4.1 Compilation of Objectives. There are seven objectives of open data intermediaries found in the literature. Table 5 shows the list of the objectives and the publications that describe the respective objective.

Many publications described the objective of open data intermediaries to facilitate the use of open data. There are several barriers to open data use that call for interventions by open data intermediaries. They make open

Table 5. Objectives of open data Intermediaries From the Literature Surveyed

No.	Objective	Description	Source
1.	Facilitate use	Help to overcome the barriers to open data use, including socioeconomic gaps, lack of awareness, and lack of data skills of end-users	[10, 18, 22, 38, 42, 49, 50, 60, 66, 67]
2.	Increase the accessibility to open data	Push for transparency by calling for closed/partially closed data to be opened to users and in appropriate formats	[3, 10, 17, 22, 36, 38, 41, 42, 54, 60, 67]
3.	Close the feedback loop	Close the feedback loop between data providers or governments and citizens	[17, 20, 22, 28, 38, 41, 51]
4.	Provide services to citizens	Provide services based on open data to improve livelihoods and day-to-day activities of citizens	[1, 14, 20, 24, 27, 42, 44, 51]
5.	Bring stakeholders together	Bridge relationships between data providers and potential end-users as well as other stakeholders	[1, 28, 34, 38, 39, 67]
6.	Enhance trust between stakeholders	Become the credible trusted partner for data providers and data end-users	[1, 33, 36, 38]
7.	Improve open data practices	Assist data providers to publish open data and carry out open data initiatives	[38, 42, 48]

data more usable and impactful in various ways, such as by directly working with community organizations to understand their data needs [67], building the community's data skills [67], and simplifying complex data into digestible information [18].

Another objective of open data intermediaries found in the literature is to increase the access to open data by end-users. This objective is closely related to the objective of facilitating the use of open data. However, the emphasis of this objective is to push for transparency by calling for closed or partially closed data to be opened to end-users and in appropriate formats [17, 38, 60] to build an informed and empowered citizenry [3, 42, 67].

The literature also describes the objective of open data intermediaries to close the feedback loop between data providers and citizens. According to Reference [20], this can be achieved by open data intermediaries through at least three ways, namely, by coordinating citizens' feedback on community issues based on open data and conveying it to authorities, providing the avenue for citizens to deliberate on issues based on open data, and by working together to solve community's issues using open data. There are open data intermediaries that are created by governments specifically to collect feedback, such as the Toronto Cycling App, which is developed by a company commissioned by the City of Toronto. It offers cycling-related information based on Toronto's open data of which app users have the option to share their cycling trips for the city council to improve the cycling network infrastructure [51].

Providing services to citizens is also one of the objectives of open data intermediaries found in the literature. In general, the services are aimed to improve the livelihood of citizens, for example, to improve farm profit [1], and to help them with their day-to-day activities, for example, to facilitate their mobility [20, 51].

The literature also mentioned bringing stakeholders together as another objective. Not only do open data intermediaries connect open data providers with potential end-users [38], but they may also connect players in a market such as farmers and distributors [1] and multiple innovators that use open data [28, 34].

Other objectives of open data intermediaries found in the literature are to enhance trust between stakeholders by being the trusted partner for data providers and data end-users [1, 33, 38] and to improve open data practices [38, 42, 48] by assisting data providers in publishing open data and carrying out open data initiatives.

4.4.2 *The Why: Why Are Open Data Intermediaries Needed?* There are several reasons why open data intermediaries are needed. In general, the objectives can be grouped into two: (i) to enhance the supply, flow, and/or use of open data; and (ii) to strengthen the relationships among various open data stakeholders. While the first

Table 6. Categorization of the Objectives of Open Data Intermediaries Found in the Literature

No.	General objective	Objective found in the literature
1.	Enhance the supply, flow, and/or use of open data	facilitate use, increase the accessibility to open data, close the feedback loop, provide services to citizens, improve open data practices
2.	Strengthen the relationships among various open data stakeholders	bring stakeholders together, enhance trust between stakeholders

objective involves the active processing of open data by open data intermediaries, the second objective does not necessarily imply so. Open data intermediaries that work towards the second objective facilitate relationships among other open data stakeholders. Table 6 groups the various objectives of open data intermediaries found in the literature into the two general objectives.

Both groups of objectives are described in some ways in the definitions by References [6, 17, 25, 33, 50, 53, 57]. However, the two are not mentioned together in each of the definitions. We propose to mention the two together in describing *the why* in the common definition.

5 DEFINING OPEN DATA INTERMEDIARIES

Based on our findings, the types of actors of open data intermediaries are diverse, and for that reason, it is a challenge to be specific in describing *the who* in defining them. Nevertheless, we found that what makes open data intermediaries distinct from solely open data providers or end-users is that they take care of the interest of other open data actors. For example, when open data intermediaries use open data, they do so to create products and services for the benefit of end-users and not merely for their internal benefit. Likewise, when they help data providers to publish open data—for example, by identifying data that end-users want—they are not the data providers themselves, but they help to enhance the supply of open data for end-users to use. Some open data intermediaries also connect multiple open data end-users to each other to stimulate innovation. Therefore, we propose to describe *the who* as "third-party actors."

From the literature, we also found that in defining open data intermediaries, it is difficult to describe what they do without potentially excluding certain tasks that some of them carry out. As summarized by Reference [53], open data intermediaries may deploy economic capital (e.g., financial resources), cultural capital (e.g., knowledge of local custom), social capital (e.g., existing networking with other stakeholders), symbolic capital (e.g., well-regarded position in society), and/or technical capital (e.g., data processing skills). Different open data intermediaries offer different resources and capabilities based on their specialization. These resources and capabilities are not at the immediate disposal of most open data users. Therefore, we propose to describe *the what* as "provide specialized resources and capabilities."

In terms of determining where open data intermediaries are located in the open data lifecycle [64], we found that their tasks are located at various stages of the open data lifecycle. Therefore, we agree with the approach taken by References [10, 17, 25, 33, 50, 55, 56], which are agnostic about *the where* in defining them. This is a more concise approach than describing *the where* as "positioned at any stage of the open data lifecycle," because the term "open data lifecycle" may need to be further clarified, and clarifying it would make the proposed common definition wordy. However, silencing *the where* does not change the essence of the proposed definition.

Last, based on our literature review, we found that the objectives of open data intermediaries can be generally grouped into two, namely, (i) to enhance the supply, flow, and/or use of open data and (ii) to strengthen the relationships among various open data stakeholders. While References [6, 17, 25, 33, 50, 53, 57] mentioned these two general objectives in some ways in their definitions, none of the said literature mentioned the two general objectives together. We propose to mention the two general objectives together in the common definition.

Ultimately, stitching together *the who, the what, the where*, and *the why* (Figure 3), we develop the following common definition of open data intermediaries: third-party actors who provide specialized resources and
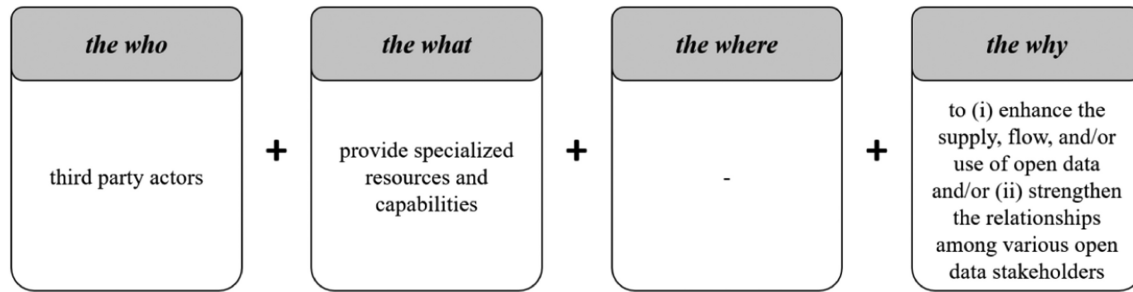
Fig. 3. Stitching together *the who, the what, the where*, and *the why* to develop a common definition of open data intermediaries.

capabilities to (i) enhance the supply, flow, and/or use of open data and/or (ii) strengthen the relationships among various open data stakeholders.

6 DISCUSSION

Definitions have often been the starting point for intellectual inquiry since the time of Socrates [19]. According to Kant, as summarized by Reference [5], to define means "to present the complete concept of a thing within its limits and in its primary or original character." According to Reference [16], as cited by Reference [19, p. 204], there are six rules for definitions, namely: (i) "a definition should give the essence or nature of the thing defined, rather than its accidental properties"; (ii) "a definition should give the genus and differentia of the thing defined"; (iii) "one should not define by synonyms"; (iv) "a definition should be concise"; (v) "one should not define by metaphors"; and (vi) "one should not define by negative terms or by correlative terms." Reference [37] notes that a definition "should include all of those things that fall under it and exclude all of those things that do not."

We assess our proposed common definition according to these six rules. Our definition describes the essence of open data intermediaries and not their accidental properties (rule i). Rule (ii) is not applicable in our context, because our definition is not an *intensional* definition (i.e., a definition that specifies the necessary and sufficient conditions for a thing) [12], instead, it is a theoretical definition, which is a definition that "function[s] as proposals to see or interpret some phenomenon in a certain way" and "since proposals have no truth value, neither do theoretical definitions" [29]. We do not define using synonyms (rule iii). We also make sure that our definition is concise (rule iv). In particular, although one may describe *the where* as "positioned at any point of the open data lifecycle," we opted to instead silence *the where*. This is because if we were to include the term "open data lifecycle" in the definition, one may question which open data lifecycle model are we referring to. Thus, silencing *the where* produces a more concise definition while retaining the essence of the definition, whereas the alternative, which is to include the term "open data lifecycle," may raise more questions and render the definition to be less concise. We do not define using metaphors (rule v), negative terms, or correlative terms (rule vi). Overall, our definition abides by the rules of Reference [16]. Our definition is also inclusive enough while excluding actors that do not play the function of open data intermediation, thus in accordance with Reference [37].

The most notable component in our proposed definition is *the why* component, which is "(i) to enhance the supply, flow, and/or use of open data and/or (ii) to strengthen the relationship among various open data stakeholders." While the first part implies active processing of open data by open data intermediaries, it may not necessarily be the case for the second part. From our literature review, an open data intermediary does not necessarily have to be actively involved in processing open data, as it can also be an actor who strengthens relationships among open data stakeholders without actively processing open data. Of course, it can also do both: actively processing open data while strengthening relationships among open data stakeholders.

The common definition we proposed can serve as a guide to judge whether an actor can be considered an open data intermediary. This is, however, not as straightforward. For example, one can argue that an internet service provider enables open data flows, hence, should it be considered an open data intermediary? To answer this, we follow the argument of Reference [53], that is, to decide whether an actor can be considered an open data intermediary, we have to look at the degree of active involvement of the actor in carrying out the open data intermediation function. Hence, Reference [53] justifies that an internet service provider is not considered an open data intermediary, because it has a low degree of active involvement in processing open data. Of course, one can further ask, how to measure the degree of active involvement? This question is, however, beyond the scope of this article.

From the literature that we reviewed, an example of an open data intermediary that enhances the use of open data but does not actively work on strengthening the relationships among open data stakeholders is Farmerline, a company in Ghana that uses open data to provide advisory services to farmers on agricultural practices, weather forecasts, market prices, and financial matters [1]. However, an example of an open data intermediary that strengthens relationships among open data stakeholders but does not actively process open data is the Edmonton Public Library, which regularly holds hackathons and other data-related events [50]. An example of an open data intermediary that processes open data and strengthens the relationships among open data stakeholders simultaneously is The City Life, an organization that works on housing justice in Boston, which collects data from the Massachusetts Trial Court Electronic Case Access (MassCourts) platform, generates a comprehensive aggregated eviction dataset, and uses it for its advocacy work including by engaging city officials to channel feedback from citizens [44].

7 CONCLUSION

There are various definitions of open data intermediaries in the literature. They can benefit from harmonization to ensure a common understanding among researchers, as some of them are quite different from each other. Based on a systematic review, we proposed the following common definition of open data intermediaries: third-party actors who provide specialized resources and capabilities to (i) enhance the supply, flow, and/or use of open data and/or (ii) strengthen the relationships among various open data stakeholders.

The fact that there is no common definition of open data intermediaries to date means that there may be literature that looks into open data actors performing intermediation functions but does not label these actors as open data intermediaries. While this implies a limitation in our literature search, it also reaffirms the need for a common definition of open data intermediaries moving forward. This common definition can be used by researchers and practitioners to mutually identify open data intermediaries and build knowledge about them on top of a mutual understanding of who are open data intermediaries.

In the process of developing the common definition, we compiled various types of actors, tasks, and objectives of open data intermediaries, which in itself contributes to further crystallizing the understanding of them. We identified six types of actors of open data intermediaries. One of the notable insights from looking at the types of actors is that apart from businesses and civil society organizations, which are commonly known to play the role of open data intermediaries, there are also open data intermediaries among public organizations. In terms of their tasks, we identified 14 of them, which take place at various stages of the open data lifecycle. We also identified seven objectives of open data intermediaries, which can be grouped into objectives that relate to the processing of open data and objectives that relate to relationship building among open data stakeholders. This compilation provides a comprehensive picture of the diversity of open data intermediaries, which is important not only for researchers but also for practitioners in designing policies or business models for open data intermediaries. Besides, it can also support open data networks of actors to account for "who is currently doing what?" and by extension "what else needs to be done?"

A shared understanding of open data intermediaries that this article offers through the proposed common definition as well as the compilation of types of actors, tasks, and objectives can serve as a foundation for researchers

and practitioners to collectively advance the role of open data intermediaries in generating value from open data. From this understanding, the needs of open data intermediaries as well as the expectations from them by other open data actors can be better identified and acted upon. This, ultimately, can help ensure the sustainability of not only open data intermediaries but also the open data network of actors as a whole.

A APPENDIX

Table A.1. The Search Strategy for Each Database

Database	Search query	Search in	Results	Notes
Scopus	TITLE-ABS-KEY (("open data" OR "open government data") AND ("intermediary" OR "intermediation" OR "infomediary"))	title, abstract, keywords	76	Using the singular form of a word in the search in Scopus gives the singular, plural, and possessive forms of most words
WoS	TS = (("open data" OR "open government data") AND ("intermediaries" OR "intermediary" OR "intermediation" OR "infomediaries" OR "infomediary"))	title, abstract, author keywords, and Keywords Plus	47	
Google Scholar	allintitle: open + data + (intermediaries OR intermediary OR intermediation OR infomediaries OR infomediary)	title	44	Google Scholar only allows terms searched either in the title or in the whole publication. The latter will give about 16,900 publications, hence, the search is only done in the title.
OATD	abstract: ("open data" AND intermediaries) OR abstract: ("open data" AND intermediary) OR abstract: ("open data" AND intermediation) OR abstract: ("open data" AND infomediaries) OR abstract: ("open data" AND infomediary) OR abstract: ("open government data" AND intermediaries) OR abstract: ("open government data" AND intermediary) OR abstract: ("open government data" AND intermediation) OR abstract: ("open government data" AND infomediaries) OR abstract: ("open government data" AND infomediary)	abstract	9	In OATD, we cannot conduct the search based on abstract and title at the same time. However, based on our check, in our case, conducting the search based on the abstract will include results if we were to conduct the search based on the title.

Table A.2. Definitions of Open Data Intermediaries from the Literature Reviewed

No.	Literature	Definition	Inspired* or adopted by
1.	[10, p. 362]	Open data intermediaries are "organizations that share data for its access, consumption, and re-usage (including re-sharing) by other organizations and individuals"	[17]*
2.	[32, p. 695]	Infomediaries are those involved in "the handling of information between information providers and consumers"	[51]
3.	[57, p. 226]	Open data intermediaries are "all the players (in an individual way or representatives of governments and social organizations), who are involved with public data that are released in an open format. They may or may not make use of technological, legal or structural artifacts in their activities. In making use of open data, the intermediaries aggregate value to the data to ensure that they can be understood more easily (and hence have a greater value) [by] third parties after their intervention]"	Nil
4.	[25, p. 4]	Open government data intermediaries are "all actors that assist OGD [open government data] initiatives by bridging the barriers that separate public sector data producers and civil society data consumers"	Nil
5.	[53, p. 7] & [52, p. 12]	An open data intermediary is "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"	[1, 18, 26, 38, 56, 67]
6.	[6, p. 222]	Open government data intermediaries are "actors who bridge gaps between data producers (governments) and data users (civil society) in that they supply essential resources and capabilities necessary to turn government data into development actions and results"	Nil
7.	[41, p. xi]	A government data intermediary is "an actor that bridges the gap between marginalized groups and OGD [open government data] by facilitating physical access, technical capacity, and value for use of information"	Nil
8.	[55, p. 2]	Open government data intermediaries are "extra-institutional actors that translate, use, or otherwise mediate communication using data produced by or for government"	Nil
9.	[17, p. 96]	Open data intermediaries are "those who operate within the open data ecosystem by means of their contribution, in one way or the other, to the supply of open data by governments as well as to the demand for such data by citizens"	Nil
10.	[33, p. 10]	Infomediaries are "specific categories of open data users who extract, aggregate, and transform data, altering it into a format that is seen as valuable, beneficial, and, most important, usable to the general public"	[21]
11.	[50, p. 31]	A civic infomediary is "a person or organization that connects community members with open data so that public value can be derived from the data"	Nil
12.	[4, p. 133]	An open government data intermediary is "the in-between actor standing between a government and a citizen in the process of data communication"	Nil

Note: "Inspired" means the publication develops a new definition based on the definition offered by the source publication whereas "Adopted by" means the publication follows entirely the definition offered by the source public.

REFERENCES

[1] Alexander Andrason and François van Schalkwyk. 2017. *Opportune Niches in Data Ecosystems: Open Data Intermediaries in the Agriculture Sector in Ghana.* Social Science Research Network, Rochester, NY. DOI : https://doi.org/10.2139/ssrn.2949722

[2] Judie Attard, Fabrizio Orlandi, and Sören Auer. 2016. Value creation on open government data. In *49th Hawaii International Conference on System Sciences (HICSS).* 2605–2614. DOI : https://doi.org/10.1109/HICSS.2016.326.

[3] Stefan Baack. 2015. Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism. *Big Data Soc.* 2, 2 (Dec. 2015), 2053951715594634. DOI : https://doi.org/10.1177/2053951715594634

[4] Annemarie Balvert and Gijs van Maanen. 2019. Open for whom? The role of intermediaries in data publication. *Rev. Publicum* 5, 2 (Dec. 2019), 129–159. DOI : https://doi.org/10.12957/publicum.2019.47205

[5] Lewis White Beck. 1956. Kant's Theory of Definition. *Philos. Rev.* 65, 2 (1956), 179–191. DOI : https://doi.org/10.2307/2182830

[6] Jérôme Brugger, Marianne Fraefel, Reinhard Riedl, Hansjakob Fehr, Daniel Schöeneck, and Christoph Stäehli Weissbrod. 2016. Current barriers to open government data use and visualization by political intermediaries. In *Conference for E-Democracy and Open Government (CeDEM)*. 219–229. DOI : https://doi.org/10.1109/CeDEM.2016.18

[7] Michael Parmisano Cañares. 2014. Opening the local: Full disclosure policy and its impact on local governments in the Philippines. In *8th International Conference on Theory and Practice of Electronic Governance (ICEGOV'14)*. Association for Computing Machinery, New York, NY, 89–98. DOI : https://doi.org/10.1145/2691195.2691214

[8] Mavis Chan, Peter Johnson, and Malcolm Shookner. 2016. Assessing the use of government open data and the role of data infomediaries the case of Nova Scotia's Community Counts program. *JeDEM - EJ. EDemoc. Open Gov.* 8, 1 (July 2016), 1–27. DOI : https://doi.org/10.29379/jedem.v8i1.370

[9] Yannis Charalabidis, Anneke Zuiderwijk, Charalampos Alexopoulos, Marijn Janssen, J Höchtl, and Enrico Ferro. 2018. *The World of Open Data: Concepts, Methods, Tools and Experiences*. Springer International Publishing AG.

[10] Sumandro Chattapadhyay. 2014. Access and use of government data by research and advocacy organisations in India: A survey of (potential) open data ecosystem. In *8th International Conference on Theory and Practice of Electronic Governance*. ACM, 361–364. DOI : https://doi.org/10.1145/2691195.2691262

[11] Abdusselam Selami Cifter and Hua Dong. 2009. User characteristics: Professional vs. lay users. In *Include 2009*. Royal College of Art, London. Retrieved from https://bura.brunel.ac.uk/bitstream/2438/3471/1/layusers_include09.pdf .

[12] Roy Cook. 2009. *Dictionary of Philosophical Logic*. Edinburgh University Press, Edinburgh.

[13] Jacqueline Corbett, Mathieu Templier, and Hirotoshi Takeda. 2018. The making of a "Top" Open Data City: A case study of Edmonton's open data initiative. In *Open Data, Information Processing, and Datification in Government*. DOI : https://doi.org/10.24251/HICSS.2018.308

[14] Tim Davies and Duncan Edwards. 2012. Emerging implications of open and linked data for knowledge sharing in development. *IDS Bull.* 43, (Sept. 2012). DOI : https://doi.org/10.1111/j.1759-5436.2012.00372.x

[15] Suruchi Dumpawar. 2015. *Open Government Data Intermediaries: Mediating Data to Drive Changes in the Built Environment*. Master's thesis. Massachusetts Institute of Technology. Retrieved from https://dspace.mit.edu/handle/1721.1/97994 .

[16] Paul Edwards. 1967. *The Encyclopedia of Philosophy*. Macmillan, New York.

[17] Patrick Enaholo. 2017. Beyond mere advocacy: CSOs and the role of intermediaries in Nigeria's open data ecosystem. In *The Social Dynamics of Open Data*, François van Schalkwyk, Stefaan Verhulst, Gustavo Magalhaes, Juan Pane and Johanna Walker (Eds.). African Minds, Cape Town, South Africa, 89–108. Retrieved from http://www.africanminds.co.za/the-social-dynamics-of-open-data/ .

[18] Patrick Enaholo and Doyinsola Dina. 2020. Journalists and the intermediation of open data: A Nigerian perspective. In *Situating Open Data: Global Trends in Local Contexts. African Minds*, Cape Town, South Africa, 31–49. Retrieved from https://www.africanminds.co.za/situation-open-data-global-trends-in-local-contexts/#:~:text=Open%20data%20and%20its%20effects,focus%20on%20its%20situated%20use.

[19] John Flowerdew. 1992. Definitions in science lectures. *Appl. Linguist.* 13, 2 (June 1992), 202–221. DOI : https://doi.org/10.1093/applin/13.2.202

[20] Mark Frank and Phil Waddell. 2014. Solving the Democratic Deficit: The Role of Open Data and Intermediaries. In *CeDEM14 Proceedings of the International Conference for E-Democracy and Open Government, Krems an der Donau*, Austria, 487–493. Retrieved from https://www.donau-uni.ac.at/dam/jcr:87536a6c-cf87-4841-bc18-21952a6223bc/cedem14_proceedings.pdf .

[21] Yingying Gao and Marijn Janssen. 2022. The open data canvas—Analyzing value creation from open data. *Digit. Gov. Res. Pract.* 3, 1 (Mar. 2022), 5:1–5:15. DOI : https://doi.org/10.1145/3511102

[22] Elisabeth Gebka, Jonathan Crusoe, and Karin Ahlin. 2021. Methods to capture user information needs: Design principles for open data intermediaries and data providers. *Data Sci. J.* 20, 1 (Dec. 2021), 37. DOI : https://doi.org/10.5334/dsj-2021-037

[23] Edson Carlos Germano, Cesar Alexandre de Souza, and Violeta Sun. 2016. Business models adopted by intermediaries in the use of open government data. *REBRAE* 9, 1 (July 2016), 79–94. DOI : https://doi.org/10.7213/rebrae.09.001.AO05

[24] Olivier Glassey. 2017. Proactive transparency and open data: A tentative analysis. In *Electronic Government (Lecture Notes in Computer Science)*. Springer International Publishing, Cham, 314–323. DOI : https://doi.org/10.1007/978-3-319-64677-0_26

[25] Felipe González-Zapata and Richard Heeks. 2015. Understanding Multiple Roles of Intermediaries in Open Government Data. In *Proceedings of the 13th International Conference on Social Implications of Computers in Developing Countries (IFIP 9.4)*, Negombo, 698–707. Retrieved from https://www.researchgate.net/profile/Solomon-Demissie-6/publication/297760744_CLOUD_COMPUTING_ADOPTION_ISSUES_FOR_ETHIOPIAN_PUBLIC_AND_PRIVATE_ENTERPRISES/links/56e3215a08ae68afa10ca64a/CLOUD-COMPUTING-ADOPTION-ISSUES-FOR-ETHIOPIAN-PUBLIC-AND-PRIVATE-ENTERPRISES.pdf .

[26] Davey den Haan. 2018. *Intermediation as a Dissolver of Barriers: How Intermediaries Can Help Overcome Barriers in Open Data Use*. Master's thesis. Wageningen University. Retrieved from https://studenttheses.uu.nl/handle/20.500.12932/31148 .

[27] Jesse Hablé. 2019. *Value Drivers and Inhibitors in Municipal Open Government Data Ecosystems: An In-depth Analysis of Stakeholder Perceptions on Values, Barriers and Success Factors in Open Government Data Initiatives in Dutch Municipalities.* Master's thesis. Delft University of Technology. Retrieved from https://repository.tudelft.nl/islandora/object/uuid%3A1347feeb-dce7-44c7-b8b7-3f58f2dd254b.

[28] Hendrik Hielkema and Patrizia Hongisto. 2013. Developing the Helsinki Smart City: The role of competitions for open data applications. *J. Knowl. Econ.* 4, 2 (June 2013), 190–204. DOI : https://doi.org/10.1007/s13132-012-0087-6

[29] Patrick J. Hurley. 2014. *A Concise Introduction to Logic* (12th ed.). Cengage Learning, Stamford, CT.

[30] IODC. 2015. Principles. International Open Data Charter. Retrieved from https://opendatacharter.net/wp-content/uploads/2015/10/opendatacharter-charter_F.pdf.

[31] Marijn Janssen, Yannis Charalabidis, and Anneke Zuiderwijk. 2012. Benefits, adoption barriers and myths of open data and open government. *Inf. Syst. Manag.* 29, 4 (Sept. 2012), 258–268. DOI : https://doi.org/10.1080/10580530.2012.716740

[32] Marijn Janssen and Anneke Zuiderwijk. 2014. Infomediary business models for connecting open data providers and users. *Soc. Sci. Comput. Rev.* 32, 5 (Oct. 2014), 694–711. DOI : https://doi.org/10.1177/0894439314525902

[33] Peter A. Johnson and Sarah Greene. 2017. Who are government open data infomediaries? A preliminary scan and classification of open data users and products. *J. URISA* 28, 1 (Dec. 2017), 9–18.

[34] Gustaf Juell-Skielse, Anders Hjalmarsson, Elea Juell-Skielse, Paul Johannesson, and Daniel Rudmark. 2014. Contests as innovation intermediaries in open data markets. *Inf. Polity* 19, 3–4 (Jan. 2014), 247–262. DOI : https://doi.org/10.3233/IP-140346

[35] Maxat Kassen. 2018. Open data and its intermediaries: A cross-country perspective on participatory movement among independent developers. *Knowl. Manag. Res. Pract.* 16, 3 (July 2018), 327–342. DOI : https://doi.org/10.1080/14778238.2018.1471377

[36] DeeDee Kim. 2018. *Data + Disasters: Rethinking the Role of Metro Boston's Data Intermediaries in Disaster Management.* Master's thesis. Massachusetts Institute of Technology. Retrieved from https://dspace.mit.edu/handle/1721.1/117828.

[37] Steve LeMay, Marilyn M. Helms, Bob Kimball, and Dave McMahon. 2017. Supply chain management: The elusive concept and definition. *Int. J. Logist. Manag.* 28, 4 (Jan. 2017), 1425–1453. DOI : https://doi.org/10.1108/IJLM-10-2016-0232

[38] Arthur Glenn Maail. 2017. The relational impact of open data intermediation: Experience from Indonesia and the Philippines. In *The Social Dynamics of Open Data*, François van Schalkwyk, Stefaan Verhulst, Gustavo Magalhaes, Juan Pane and Johanna Walker (Eds.). African Minds, Cape Town, South Africa, 153–166. Retrieved from http://www.africanminds.co.za/the-social-dynamics-of-open-data/.

[39] Viktor Mayer-Schönberger and Zarino Zappia. 2011. Participation and power: Intermediaries of open data. *Alexander von Humboldt Institut für Internet und Gesellschaft (HIIG)*, Berlin. Germany. Retrieved from https://www.hiig.de/wp-content/uploads/2017/08/participation_and_power.pdf.

[40] Albert Meijer and Suzanne Potjer. 2018. Citizen-generated open data: An explorative analysis of 25 cases. *Gov. Inf. Q.* 35, 4 (Oct. 2018), 613–621. DOI : https://doi.org/10.1016/j.giq.2018.10.004

[41] Amanda Meng. 2016. *The Social Impact of Open Government Data.* PhD thesis. Georgia Institute of Technology. Retrieved from https://smartech.gatech.edu/handle/1853/56292.

[42] Amanda Meng, Carl DiSalvo, Lokman Tsui, and Michael Best. 2019. The social impact of open government data in Hong Kong: Umbrella Movement protests and adversarial politics. *Inf. Soc.* 35, 4 (Aug. 2019), 216–228. DOI : https://doi.org/10.1080/01972243.2019.1613464

[43] Jennifer C. Molloy. 2011. The open knowledge foundation: Open data means better science. *PLOS Biol.* 9, 12 (Dec. 2011), e1001195. DOI : https://doi.org/10.1371/journal.pbio.1001195

[44] Chenab Navalkha. 2021. *Data for Housing Justice: Examining Activists' Use of Open Government Data for Housing Justice in Boston, MA, and New York, NY.* Master's thesis. Massachusetts Institute of Technology. Retrieved from https://dspace.mit.edu/handle/1721.1/138973

[45] OKF. 2013. Defining Open Data—Open Knowledge Foundation blog. Retrieved from https://blog.okfn.org/2013/10/03/defining-open-data/.

[46] Chitu Okoli and Kira Schabram. 2010. A guide to conducting a systematic literature review of information systems research. *SSRN Electron. J.* (2010). DOI : https://doi.org/10.2139/ssrn.1954824

[47] Oxford University Press. 2022. Definition of third party noun. *Oxford Advanced Leaner's Dictionary.* Retrieved from https://www.oxfordlearnersdictionaries.com/definition/english/third-party?q=third+party.

[48] Sora Park and J. Ramon Gil-Garcia. 2017. Understanding transparency and accountability in open government ecosystems: The case of health data visualizations in a state government. In *18th Annual International Conference on Digital Government Research*. Association for Computing Machinery, New York, NY, 39–47. DOI : https://doi.org/10.1145/3085228.3085318

[49] Luigi Reggi and Sharon S. Dawes. 2022. Creating open government data ecosystems: Network relations among governments, user communities, NGOs and the media. *Gov. Inf. Q.* (Jan. 2022), 101675. DOI : https://doi.org/10.1016/j.giq.2022.101675

[50] Pamela Robinson and Lisa Ward Mather. 2017. Open data community maturity: Libraries as civic infomediaries. *URISA J.* 28, 1 (Dec. 2017), 31–39.

[51] Suthee Sangiambut and Renee Sieber. 2017. The civic open data and crowdsourcing app ecosystem: Actors, materials, and interventions. *J. URISA* 28, 1 (2017), 49–62.

[52] François van Schalkwyk, Michael Cañares, Sumandro Chattapdhyay, and Alexander Andrason. 2016. Open data intermediaries in developing countries. *J. Communit Inf.* 12, 2 (June 2016). DOI : https://doi.org/10.15353/joci.v12i2.3219

[53] François van Schalkwyk, Sumandro Chattapadhyay, Michael Cañares, and Alexander Andrason. 2015. *Open Data Intermediaries in Developing Countries*. World Wide Web Foundation, Washington, D.C. Retrieved from https://idl-bnc-idrc.dspacedirect.org/handle/10625/56288.

[54] François van Schalkwyk, Michelle Willmers, and Maurice McNaughton. 2016. Viscous open data: The roles of intermediaries in an open data ecosystem. *Inf. Technol. Dev.* 22, sup1 (Aug. 2016), 68–83. DOI:https://doi.org/10.1080/02681102.2015.1081868

[55] Andrew Schrock and Gwen Shaffer. 2017. Data ideologies of an interested public: A study of grassroots open government data intermediaries. *Big Data Soc.* 4, 1 (June 2017), 2053951717690750. DOI:https://doi.org/10.1177/2053951717690750

[56] Gisele da Silva Craveiro and Claudio Albano. 2017. Open data intermediaries: Coproduction in budget transparency. *Transform. Gov. People Process Polic* 11, 1 (Jan. 2017), 119–131. DOI:https://doi.org/10.1108/TG-12-2015-0057

[57] Gisele da Silva Craveiro and Cláudio Sonáglio Albano. 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. In *Open and Big Data Management and Innovation (Lecture Notes in Computer Science)*. Springer International Publishing. Cham, 223–235. DOI:https://doi.org/10.1007/978-3-319-25013-7_18

[58] Michael C. Sloan. 2010. Aristotle's Nicomachean ethics as the original locus for the septem circumstantiae. *Class. Phil.* 105, 3 (2010), 236–251. DOI:https://doi.org/10.1086/656196

[59] Megan K. Sutherland and Meghan E. Cook. 2017. Data-driven smart cities: A closer look at organizational, technical and data complexities. In *18th Annual International Conference on Digital Government Research*. Association for Computing Machinery, New York, NY, 471–476. DOI:https://doi.org/10.1145/3085228.3085239

[60] Piyushimita (Vonu) Thakuriah, Lise Dirks, and Yaye Mallon Keita. 2017. Digital infomediaries and civic hacking in emerging urban data initiatives. In *Seeing Cities through Big Data: Research, Methods and Applications in Urban Informatics*, Piyushimita (Vonu) Thakuriah, Nebiyou Tilahun and Moira Zellner. (Eds.). Springer International Publishing, Cham, 189–207. DOI:https://doi.org/10.1007/978-3-319-40902-3_11

[61] David R. Thomas. 2006. A general inductive approach for analyzing qualitative evaluation data. *Am. J. Eval.* 27, 2 (June 2006), 237–246. DOI:https://doi.org/10.1177/1098214005283748

[62] Antônio Márcio Tavares Thomé, Luiz Felipe Scavarda, and Annibal José Scavarda. 2016. Conducting systematic literature review in operations management. *Prod. Plan. Control* 27, 5 (Apr. 2016), 408–420. DOI:https://doi.org/10.1080/09537287.2015.1129464

[63] Paul F. Uhlir and Peter Schröder. 2007. Open data for global science. *Data Sci. J.* 6, (2007), OD36–OD53. DOI:https://doi.org/10.2481/dsj.6.OD36

[64] Anne Fleur van Veenstra and Tijs van den Broek. 2015. A community-driven open data lifecycle model based on literature and practice. In *Case Studies in e-Government 2.0: Changing Citizen Relationships*, Imed Boughzala, Marijn Janssen and Saïd Assar (Eds.). Springer International Publishing, Cham, 183–198. DOI:https://doi.org/10.1007/978-3-319-08081-9_11

[65] Yu Xiao and Maria Watson. 2019. Guidance on conducting a systematic literature review. *J. Plan. Educ. Res.* 39, 1 (Mar. 2019), 93–112. DOI:https://doi.org/10.1177/0739456X17723971

[66] Ayoung Yoon and Andrea Copeland. 2020. Toward community-inclusive data ecosystems: Challenges and opportunities of open data for community-based organizations. *J. Assoc. Inf. Sci. Technol.* 71, 12 (2020), 1439–1454. DOI:https://doi.org/10.1002/asi.24346

[67] Ayoung Yoon, Andrea Copeland, and Paula Jo McNally. 2018. Empowering communities with data: Role of data intermediaries for communities' data utilization. *Proc. Assoc. Inf. Sci. Technol.* 55, 1 (2018), 583–592. DOI:https://doi.org/10.1002/pra2.2018.14505501063

[68] Andrew Young and Stefaan Verhulst. 2017. *Aclímate Colombia: Open Data to Improve Agricultural Resiliency*. GovLab - Open Data's Impact. DOI:https://doi.org/10.1007/978-3-319-63743-3_6

[69] Zhe Zhu, Michael A. Wulder, David P. Roy, Curtis E. Woodcock, Matthew C. Hansen, Volker C. Radeloff, Sean P. Healey, Crystal Schaaf, Patrick Hostert, Peter Strobl, Jean-Francois Pekel, Leo Lymburner, Nima Pahlevan, and Ted A. Scambos. 2019. Benefits of the free and open Landsat data policy. *Remote Sens. Environ.* 224, (Apr. 2019), 382–385. DOI:https://doi.org/10.1016/j.rse.2019.02.016

[70] Mária Žuffová. 2020. *Governments in the Spotlight?: On the Use and Impacts of Freedom of Information Laws and Proactive Publication of Government Data*. PhD thesis. University of Strathclyde.

[71] Anneke Zuiderwijk, Marijn Janssen, Sunil Choenni, Ronald Meijer, and Roexsana Sheikh Alibaks. 2012. Socio–technical Impediments of Open Data. *Electron. J. E-Gov.* 10, 2 (Dec. 2012), 156–172.