DATA FOR POLICY PROCEEDINGS PAPER

Exploring the contributions of open data intermediaries for a sustainable open data ecosystem

Ashraf Shaharudin[1] ⓘ, Bastiaan van Loenen[1] and Marijn Janssen[2] ⓘ

[1]Department of Urbanism, Faculty of Architecture and the Built Environment, Delft University of Technology, Delft, the Netherlands
[2]Department of Engineering, Systems and Services, Faculty of Technology, Policy and Management, Delft University of Technology, Delft, the Netherlands
Corresponding author: Ashraf Shaharudin; Email: a.a.binahmadshaharudin@tudelft.nl

Abstract

Open data promises various benefits, including stimulating innovation, improving transparency and public decision-making, and enhancing the reproducibility of scientific research. Nevertheless, numerous studies have highlighted myriad challenges related to preparing, disseminating, processing, and reusing open data, with newer studies revealing similar issues to those identified a decade prior. Several researchers have proposed the open data ecosystem (ODE) as a lens for studying and devising interventions to address these issues. Since actors in the ecosystem are individually and collectively impacted by the sustainability of the ecosystem, all have a role in tackling the challenges in the ODE. This paper asks what the contributions of open data intermediaries may be in addressing these challenges. Open data intermediaries are third-party actors providing 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. They are critical in ensuring the flow of resources within the ODE. Through semi-structured interviews and a validation exercise in the European Union context, this study explores the potential contribution of open data intermediaries and the specific ODE challenges they may address. This study identified 20 potential contributions, addressing 27 challenges. The findings of this study pave the way for further inquiry into the internal incentives (viable business models) and external incentives (policies and regulations) to direct the contributions of open data intermediaries toward addressing challenges in the ODE.

Policy Significance Statement

Practitioners, policymakers, and designers may consider the findings from this paper to examine the viability of the potential contributions suggested, deliberate policies around open data intermediaries, and develop their organizational and business models.

1. Introduction

Open data are "data that can be freely used, shared, and built on by anyone, anywhere, for any purpose" (OKF, 2013). Such data promise various benefits, including stimulating innovation, improving transparency, and enhancing the reproducibility and dissemination of scientific research (Molloy, 2011; Janssen et al., 2012; Zhu et al., 2019). Open data has demonstrated its value in various domains. For instance, open

data provided by the Global Forest Watch has been used by environmental advocates in the Philippines to support the enactment of the National Mangrove Forests Conservation and Rehabilitation Act (Camero, 2015), by indigenous groups in Peru to track and report forest loss in the Amazon (Moloney, 2021), and by international researchers to study global forest fire trends over time (Tyukavina et al., 2022; Johnson, 2023). The implementation of an open procurement data policy in the European Union (EU) has led to more competitive bidding, which has been attributed to increased scrutiny by nongovernmental organizations and investigative journalists, as well as learning by national procurement regulators (Duguay et al., 2023).

Nevertheless, various challenges exist in preparing, disseminating, processing, and reusing open data (Conradie and Choenni, 2014; Johnson et al., 2017; Sugg, 2022). These include those faced by open data providers, such as scattered data management (Linders, 2013; Ma and Lam, 2019), resource constraints (Janssen et al., 2012; Nikiforova and Zuiderwijk, 2022) and the lack of a convincing business case for publishing open data (Barry and Bannister, 2014; Cahlikova and Mabillard, 2020), and those faced by open data users, such as poor or inconsistent open data quality (Crusoe et al., 2019; Huijboom and Van den Broek, 2011), underdeveloped data skills (Janssen et al., 2012; Zuiderwijk and Reuver, 2021), and limited complementary technologies for using open data (Temiz et al., 2022; Zuiderwijk et al., 2012). More recent studies seem to reveal similar challenges to those identified a decade ago persist (to illustrate this, we cite studies published before/in 2013 and after 2013 for each of the aforementioned examples). Therefore, it is imperative to investigate how to address these challenges. As such, this paper asks what the contributions of open data intermediaries may be in this regard.

Open data intermediaries can be defined as "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" (Shaharudin et al., 2023). Notably, they serve a pivotal role in the effective distribution of resources in the open data ecosystem (ODE) (Chattapadhyay, 2014). Open data intermediaries could be a public organization, a for-profit company, a nonprofit organization, a public–private partnership, or various other types of entities undertaking a wide range of tasks for different objectives (Andrason and van Schalkwyk, 2017; Dumpawar, 2015; Enaholo and Dina, 2020; González-Zapata and Heeks, 2015; Schrock and Shaffer, 2017; Shaharudin et al., 2023). They include software providers that integrate preprocessed open data in their systems, nonprofit organizations that compile, contextualize, and visualize open data for public consumption, and crowdsourcing platforms that facilitate the contribution and reuse of open data.

Despite the crucial role of open data intermediaries, research on their potential contributions to addressing persistent challenges in the ODE remains lacking. Most studies have only investigated their existing activities. While this baseline understanding is undoubtedly important, the next necessary step is to explore what could be done by open data intermediaries for the benefit of the entire ODE. One pathway toward this end is to explore the connections between challenges in the ODE and potential contributions by open data intermediaries to address them. Hence, the research question of this paper is presented as follows: What are the potential contributions of open data intermediaries to address various challenges in the ODE? We answer this question within the context of the European Union (EU). The legal regime of open data in the EU is rather specific since many organizations are legally obligated to publish and facilitate the reuse of open data. Failure to comply may result in EU member states facing legal consequences brought by the European Commission (EC), including heavy financial penalties.

The structure of this paper is as follows. The next section, Section 2, outlines the conceptual framework, particularly on the ODE and open data intermediaries. Section 3 presents the research methodology. Section 4 presents the findings. Section 5 reflects on the findings. Lastly, Section 6 concludes the paper.

2. Conceptual framework

This section lays out the conceptual framework to motivate the need to explore the potential contributions of open data intermediaries and situate this study within the broader discourse on open data.

Subsection 2.1 discusses the ODE as an analytical tool describing relationships between actors involved in open data. We do this by reviewing claims around the ODE, as advocated by open data researchers and practitioners, and comparing them with theoretical concepts from the broader field of science and technology studies (STS), specifically those associated with actor–network theory (ANT). Once we have established the scope of an ODE, Subsection 2.2 discusses the "sustainability" of an ODE. In particular, we review the features of sustainable open data ecosystems and the challenges faced by different open data actors that could disrupt the sustainability of an ODE. In Subsection 2.3, we discuss open data intermediaries as a category of actors in an ODE. We review their types, activities, and roles toward a sustainable ODE.

2.1. Open data ecosystem (ODE)

Based on empirical studies of the U.K. government's open data initiatives and the International Aid Transparency Initiative, Davies (2011) argued that successful value generation from open data relies on more than just the dataset; it also relies on the "mobilisation of a wide range of technical, social and political resources, and on interventions […] [to] support coordination of activity around datasets" (Davies, 2011, p. 1). Thus, he advocated for the ODE as an analytical lens through which "the emergent, autonomous and self-organising components" are "linked together in local and global feedback loops and developing according to local specialisations and adaptation rather than top-down design" (Davies, 2011, p. 3). According to Davies, the "ecosystem" metaphor is useful in describing the interrelations of open data actors and devising interventions to increase cooperation and interaction among actors, just as the metaphor is used in other fields such as economics and political science. This dual function of the metaphor was also emphasized by Harrison et al. (2012, p. 907): "Although ecosystems are naturally occurring phenomena and the metaphor may be applied to any existing socio-technical domain, they can also be seeded, modelled, developed, managed, that is, intentionally cultivated for the purpose of achieving a managerial and policy vision."

Csáki (2019) succinctly defined an ODE as a "*way of looking* [emphasis added] at how participating actors and groups create shared meaning and generate value around open data and how the structural properties of their interactions shape this process, which in turn enables or constrains the growth and health of the ecosystem itself" (Csáki 2019, p. 19). This definition evokes similarities with how ANT is described by one of its cofounders, which is "a disparate family of material-semiotic *tools, sensibilities and methods of analysis* [emphasis added] that treat everything in the social and natural worlds as a *continuously generated effect of the webs of relations* [emphasis added] within which they are located" (Law, 2008, p. 141). Thus, an ODE can be considered a specific application of the ANT, rooted in the larger field of STS. From this perspective, an ODE is not an isolated nor necessarily novel analytical approach to a socio-technical network but one that can draw theoretically and empirically grounded insights from other types of networks, including information systems [e.g., Díaz Andrade and Urquhart (2010); Nehemia-Maletzky et al. (2018); Walsham (1997)].

A fundamental notion underpinning ANT is the participation of humans and nonhumans in a network, including technological artifacts (in the case of open data, e.g., data portals, application programming interfaces, and data standards), not only passively as a resource or constraint, but also actively influencing the dynamic interactions in the network (Callon and Law 1997). This mirrors the conceptual elements of the ecosystem identified by Oliveira and Lóscio (2018), namely resources, roles, actors, and relationships. Resources can be datasets, data-based software, and hardware, which may be exchanged individually or in combination, through relationship transactions. A role is a function of an actor in the ecosystem. Actors are autonomous entities such as businesses, public institutions, and individuals serving one or more specific roles. Relationships are the interactions among actors in the ecosystem. Notably, Oliveira and Lóscio (2018) made a distinction between roles and actors: actors are not wedded to any particular roles.

Poikola et al. (2011) drew attention to the autonomy yet interdependency of actors in an ODE. As they put it, the "ecosystem evokes an image of the well-being of the entity and, on the other hand,

fulfilling one's own needs through the richness and vitality of the ecosystem" (Poikola et al., 2011, p. 14). They also highlighted the ever-changing nature of an ecosystem: "With the ecosystem, we wish to highlight not only the technological systems and institutionalized organisations but also the living, *dynamically changing network of interaction* [emphasis added]" (Poikola et al., 2011, p. 14). This hybridity of self-interested yet interdependent actors is an essential notion of ANT asserting that "every stable social arrangement is simultaneously a point (an individual) and a network (a collective)" (Callon and Law 1997, p. 165).

2.2. Sustainable ODEs and current challenges

One of the empirical focuses of ANT is "to trace and explain the processes whereby relatively stable networks of aligned interests are created and maintained, or alternatively to examine why such networks fail to establish themselves" (Walsham, 1997, p. 469). In the same manner, researchers are investigating factors that contribute to a sustainable ODE, where stable networks of open data actors' aligned interests are created and maintained (adopting the words of Walsham). Notably, van Loenen et al. (2021) suggested the key features of a sustainable ODE, which can be summarized as user-driven (open data supply matches the demands of users of different types and domains), circular (all actors mutually create and capture value), inclusive (all actors, not only government organizations, but are also incentivized to contribute open data and participate in ecosystem processes), and skills-based (appropriate data skills and competencies are applied) [comparable to Heimstädt et al. (2014)]. To achieve a sustainable ODE, Harrison et al. (2012) recommended strategic ecosystem thinking, which involves identifying the actors in the ecosystem, understanding the relationship between these actors, identifying the resources that each actor must engage with others, and observing the indicators that represent ecosystem health.

Numerous empirical studies and reviews have highlighted myriad challenges related to the preparation, dissemination, and reuse of open data (Bonina, 2013; Conradie and Choenni, 2014; Johnson et al., 2017; Sugg, 2022; Toogood, 2021). These challenges endanger the sustainability of ODEs. Some of them are encountered by open data providers, such as scattered data management (Linders, 2013; Ma and Lam, 2019), resource constraints (Janssen et al., 2012; Nikiforova and Zuiderwijk, 2022), and the lack of a convincing business case for publishing open data (Barry and Bannister, 2014; Cahlikova and Mabillard, 2020), while other challenges are experienced by open data users, such as poor or inconsistent open data quality (Crusoe et al., 2019; Huijboom and Van den Broek, 2011), underdeveloped data skills (Janssen et al., 2012; Zuiderwijk and Reuver, 2021), and limited complementary technologies for using open data (Temiz et al., 2022; Zuiderwijk et al., 2012). Research in this area is already quite saturated, with new studies highlighting similar challenges to those identified a decade prior. Therefore, the next step is to investigate how to overcome these challenges in practice.

From an ecosystem perspective, all actors, not limited to open data providers, can and should contribute to addressing challenges in ODEs. Actors in an ODE are individually and collectively affected by each other; hence, solving issues in an ODE is not a "one-person job." This understanding lays the groundwork for our research question: What could open data intermediaries contribute toward a sustainable ODE? We adopt the features that van Loenen et al. (2021) proposed (i.e., user-driven, circular, inclusive, and skills-based features) as our benchmark for a sustainable ODE. This is because, to the best of our knowledge, the paper is the first to identify the features of a sustainable ODE at the conceptual level. It is similar to—but more refined than—the features proposed by Heimstädt et al. (2014), who drew inspiration from the business ecosystem literature. However, we must first determine what the open data intermediaries are.

2.3. Open data intermediaries

Based on a systematic literature review, Shaharudin et al. (2023) proposed a theoretical definition of open data intermediaries as "third-party actors who provide specialised 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 intermediary is a crucial concept in ANT. However, ANT differentiates intermediaries, a messenger that "transports meaning or force without transformation" from mediators that "transform, translate, distort, and modify the meaning or the elements they are supposed to carry" (Latour 2007, p. 39). Our conceptualization of open data intermediaries does not make such distinction (i.e., both ANT's intermediaries and mediators are termed "open data intermediaries" in an ODE) since this is how the term is conventionally used in the literature as an umbrella term for heterogeneous types [for a review, see Shaharudin et al. (2023)]. Regardless, ANT highlights the importance of intermediaries/mediators as match-makers for situations in which "people, goods and services are brought together" (Goodchild and Ferrari 2024, p. 107). In the same vein, Chattapadhyay (2014) argued that a vital aspect of the ODE lies in the efficient circulation of resources, with open data intermediaries serving a pivotal role in this regard. The role of open data intermediaries is instrumental in the access to and use of open data (González-Zapata and Heeks 2015; Neves et al. 2020) and in connecting open data actors (Mayer-Schönberger and Zappia 2011; Yoon et al. 2018). Furthermore, since the ODE emphasizes the self-organization of actors (Davies 2011; Oliveira and Lóscio 2018), open data intermediaries are crucial in mitigating information asymmetry between actors.

As a point of clarification, open data intermediaries should certainly include "open data" in their activities despite their end product not necessarily being open [the same emphasis was notably made by van Schalkwyk et al. (2016, p. 12)]. While this may seem obvious, being involved with data shared on a case-to-case basis through individual arrangements and not with open data as it is conventionally defined, as in the case of certain forms of data collaboratives (Susha et al., 2017) or boundary organizations (Perkmann and Schildt, 2015), does not make such actors open data intermediaries. As similarly highlighted by scholars of ANT (Klecuń, 2004; Lee and Hassard, 1999), a delimitation based on concrete practices and functioning for empirical study is necessary to avoid dealing with an "endless chain of association" (Müller, 2015, p. 30).

Various types of actors serve the role of open data intermediaries, including civil society organizations (Meng et al., 2019; Sangiambut and Sieber, 2017), companies (Andrason and van Schalkwyk, 2017; Germano et al., 2016), the media (Enaholo and Dina, 2020; Johnson and Greene, 2017), researchers (Corbett et al., 2020; Park and Gil-Garcia, 2017), and government organizations (Hablé, 2019; Meijer and Potjer, 2018). They undertake various tasks, deploying different types of resources based on their specialization, such as augmenting data by integrating open data with non-open data (Andrason and van Schalkwyk, 2017; Corbett et al., 2018), developing open data-based products such as mobile apps (Chan et al., 2016; Kim, 2018), contextualizing open data into digestible information (Corbett et al., 2018; Meng, 2016), and providing training to end-users (Enaholo, 2017; Meng et al., 2019). Although the actors referred to as open data intermediaries do not necessarily undertake activities solely related to open data, open data are certainly part of their intermediation activities [see the review by Shaharudin et al. (2023)].

To date, the relevant literature has primarily investigated the existing activities of open data intermediaries. Recognizing that persistent challenges related to the preparation, dissemination, and use of open data must be resolved, and building upon the ecosystem understanding that all actors can and should contribute to addressing these challenges, the following question is relevant: What are the potential contributions of open data intermediaries in addressing challenges in the ODE? This insight is currently missing from the literature and thus represents what this paper aims to offer.

3. Research methodology

This paper focuses on the EU context due to the shared legal regime concerning open data, which enables a more contextual interpretation of the findings. The EU Open Data and the Re-use of Public Sector Information Directive (Open Data Directive) was published in 2019, and EU countries were required to transpose this directive into national laws, regulations, and administrative provisions by July 2021 (European Parliament and Council of the European Union Directive (EU), 2019). This directive is a recast to the earlier Reuse of Public Sector Information (PSI) Directive 2003, and it establishes a set of

minimum rules (including technical aspects) governing the reuse of data held by public bodies and "public undertakings" (including nonpublic entities operating in the water, energy, transport, and postal services sectors), as well as research data. The directive also lists specific high-value datasets across six thematic categories, four of which constitute spatial datasets (geospatial, earth observation and environment, meteorological, and mobility), while the remaining two are statistics and companies and company ownership. EU countries that fail to comply with this directive may face legal consequences brought by the EC, including heavy financial penalties. In fact, the EC referred several countries to the Court of Justice of the EU last year due to noncompliance with the Open Data Directive (European Commission 2023). This legal framework makes open data in the EU context rather specific since many organizations are legally obligated to publish and facilitate the reuse of open data.

Open data intermediaries are *not* to be confused with the "data intermediaries" established within the EU Data Governance Act framework (European Parliament and Council of the European Union Regulation (EU), 2022). As written (e.g., Item 10 of the Preamble), this Act does not apply to open data that falls under the Open Data Directive. Data intermediaries in the Data Governance Act are not allowed to use the data they intermediate (e.g., by developing products based on these data) and can only facilitate data sharing between parties. This is different from open data intermediaries, which can be users of open data (even for financial profit) because the data are already open.

This study adopted a qualitative research approach. It does not seek to offer generalizable quantitative sample-to-population findings nor suggest any causal relationships. Instead, it aims to pave the way for new areas of inquiry and interventions regarding the potential contributions of open data intermediaries for a sustainable ODE. There are two stages in the methodology. In Stage 1, we gathered data through semi-structured interviews involving ten interviewees from eight organizations representing open data providers, intermediaries, and a data standard body. From these interviews, we derived challenges in the ODE and the potential contributions of open data intermediaries (Section 3.1). In Stage 2, we explored the links between the individual potential contributions of open data intermediaries and specific challenges in the ODE that they can address. We validated these links with four organizations we interviewed in Stage 1 and nine additional open data practitioners and researchers (Section 3.2). Guided by the ecosystem perspective, where actors are interrelated, we involved representatives of diverse roles (open data providers, intermediaries, users, and researchers) in Stages 1 and 2.

3.1. Stage 1: semi-structured interviews

We conducted semi-structured interviews with ten experts from eight organizations based in Denmark, Germany, the Netherlands, and Spain, representing open data providers, intermediaries, and a data standard body between May and July 2023 (Table 1). These organizations were selected based on

Table 1. *List of interviewee organizations*

No.	Organization	Country	Role related to open data
1.	Agency for Data Supply and Infrastructure (SDFI)	Denmark	Provider & intermediary
2.	Federal Agency for Cartography and Geodesy (BKG)	Germany	Provider & intermediary
3.	Berlin Senate Chancellery	Germany	Intermediary
4.	Cadastre, Land Registry, and Mapping Agency (Kadaster)	Netherlands	Provider & intermediary
5.	Geonovum	Netherlands	Standard body
6.	National Centre for Geographic Information (CNIG)	Spain	Provider
7.	Red.es (a public business entity under the Ministry of Economic Affairs and Digital Transformation)	Spain	Intermediary
8.	Department of Urban Planning, Environment and Mobility, Madrid City Council	Spain	Provider

purposive sampling, in which we aimed to obtain insights from organizations that could provide information on the challenges in ODE and the potential contributions of open data intermediaries. Six out of the eight organizations are spatial data organizations because four out of six thematic categories of the high-value datasets listed under the EU Open Data Directive framework constitute spatial datasets.

The individual interviewees are involved in formulating open data legislation at the national and EU levels, fulfilling the requirements of open data legislation, and coordinating open data stakeholders' engagements. Seven of the ten interviewees have at least 10 years of experience in dealing with open data. The minimum open data experience among all of the interviewees is 5 years. All interviews were conducted online except for two, of which one was conducted in person, and another one in writing. We shared the informed consent form (reviewed by our university's Human Research Ethics Committee and in compliance with the EU General Data Protection Regulation) and the tentative questions for each interviewee several days before the interview.

We asked questions across four areas (i) the background of the interviewee and their organization in relation to open data; (ii) their perception of the value of open data and its benefits and costs to their own organization; (iii) their perceptions of the sustainability of the ODE, challenges in the ODE, and potential remedies; (iv) their perceptions of the current and potential contributions of open data intermediaries. However, the questions were not exactly the same for each interview since they were semi-structured. In semi-structured interviews, interviewers ask practical, customized, open-ended questions in a conversation-like manner without expecting specific answers, thereby allowing interviewees to speak freely and fully (Magnusson and Marecek, 2015). Since the ODE is a rather abstract concept, we offered its working definition to interviewees as "a network of interdependent yet self-interested open data actors" to guide the interviews. The interview guide can be found in the Supplementary Appendix and can also be found here (with de-identified interview transcripts and the informed consent form template): https://doi.org/10.4121/d7dd11e0-7c6c-49db-946a-ffe71520f8fd.v1. As an additional privacy protection measure, the order of the de-identified transcripts in the repository differs from that in Table 1.

We coded the interview data inductively (Linneberg and Korsgaard, 2019) with the aid of ATLAS.ti software. There are two categories of codes (i) challenges in ODEs and (ii) potential contributions of open data intermediaries. The coding results can be found here: https://doi.org/10.4121/d7dd11e0-7c6c-49db-946a-ffe71520f8fd.v1. We opted for the inductive approach to derive the challenges in ODEs and the potential contributions of open data intermediaries close to the context the interviewees referred to, without being cognitively restricted by preconceived vocabularies. This contextual understanding is crucial in making informed decisions when linking the two to avoid speculating at a very abstract level.

3.2. Stage 2: analysis & validation

In Stage 2, we linked the individual contributions of open data intermediaries to the challenges in ODEs based on the data derived in Stage 1. To validate our analysis, we requested feedback from all of our interviewees; four out of eight organizations responded with their input (organizations 1, 6, 7, and 8 in Table 1). We also obtained feedback from nine other open data practitioners and researchers (Table 2). None of those listed in Table 2 are the authors of this study. In total, we received 13 inputs during our validation exercise in February 2024. Eleven of those were done in writing, where we shared the initial version of Table 5 (with the description for each item) and validators edited it by suggesting new links between the challenges of ODEs and potential contributions of open data intermediaries where they considered appropriate and removing initial links where they considered otherwise. They also added new challenges in ODEs and the potential contributions of open data intermediaries that were not listed in the initial version but deemed relevant. The other two validations (with organization 1 in Table 1 and organization 6 in Table 2) were performed through online meetings, where we discussed the table with the validators.

Based on the input of the 13 validators, we evaluated the individual link suggestions. We considered each suggestion by speculating a case where the suggestion could be materialized. If

Table 2. *List of additional validators' organizations (aside from those interviewed)*

No.	Organization	Country	Role related to open data
1.	Aalborg University	Denmark	Researchers (x 2) on open data use
2.	National Institute of Geographic and Forestry Information (IGN)	France	Provider & intermediary
3.	Leibniz Institute of Ecological Urban and Regional Development (IOER)	Germany	User
4.	HuffPost	Greece	User and researcher on open data use
5.	Italian National Institute of Statistics (ISTAT)	Italy	Provider
6.	Dataninja	Italy	Intermediary
7.	Delft University of Technology (TU Delft)	Netherlands	(i) Researcher on open data governance & (ii) Researcher who uses open data

there is such a case, we would accept the suggestion unless there is a counterargument against it (i.e., the contribution unlikely addresses the challenge). Where appropriate, we merged the new challenges and contributions proposed by validators (e.g., we merged *develop common standards* proposed by a validator with *transform data into open standards* as one of the potential contributions). We also added new links that were not directly suggested by validators but are appropriate and consistent with the general insights and reasoning offered by validators. Finally, we generated the mapping of potential contributions of open data intermediaries and the ODE challenges that they could address (Table 5).

Most of the experts engaged in this study (in Stages 1 and 2) are from the public sector, which may represent a limitation. However, these experts have rich experience in open data accumulated over many years in various capacities, including in coordinating open data stakeholder engagements. The organizations they represent are not only open data providers but also open data intermediaries and users. Notably, this paper does not claim to offer exhaustive lists of ODE challenges and the potential contributions of open data intermediaries (if that is even possible), nor does it seek to establish causal relationships.

4. Findings

Section 4.1 discusses the challenges in ODEs, while Section 4.2 discusses the potential contributions of open data intermediaries and connects them with the challenges they may address in ODEs. We derived the challenges in the ODE and the potential contributions of open data intermediaries from the interviews. Then, we validated the connections proposed between the two with input from some of the interviewees and additional experts.

4.1. Challenges in the ODE

Table 3 presents challenges in the ODE identified from the interviews. We attempted to categorize them based on the key features of a sustainable ODE suggested by van Loenen et al. (2021), namely user-driven (open data supply matches the demand of users of different types and domains), circular (all actors mutually create and capture value), inclusive (all actors, not only government organizations, but are also incentivized to contribute open data and participate in ecosystem processes), and skills-based (appropriate data skills and competencies are applied). This form of categorization helps to show which challenges threaten which specific features of a sustainable ODE. However, we found that some challenges do not fit well into any of the four proposed features. Those that do not fit are either very foundational challenges around data management systems (categorized as "foundational") or those associated with broader

Table 3. *Challenges in the ODE identified from the interviews*

ID	Challenges in ODE	Description
	User-driven	
C01	Different data/metadata standards	Open data are provided under different data/metadata standards.
C02	Different open data licenses	Open data are provided under different licenses (e.g., CC0, CC BY, CC BY-SA, OdBL, and OGL).
C03	Siloed open data domains (e.g., across sectors)	Open data domains (e.g., across sectors) exist in silos, impeding cross-domain data integration.
C04	High technical threshold to use open data	Using open data may require complex technical know-how or tooling by users (including experts on a particular domain, but not the other).
C05	Unfulfilled user needs	Open data users' needs do not match what and how open data are provided.
C06	Limited feedback from lay users (i.e., nonexpert users)	Limited feedback from lay users (i.e., nonexpert users) on open data initiatives.
	Circular	
C07	Loss of open data providers' revenue	Open data providers lost revenue from data that they previously charged users for.
C08	Limited value return from data reuse	Open data providers are not compensated well for the costs they bear to develop and maintain open data initiatives.
C09	Limited use case visibility	Limited visibility of open data use cases (which open data are used, by whom, for what purposes, and how are they used).
	Inclusive	
C10	Limited open data from nongovernment sectors	Limited open data are provided by sectors other than the government sector.
C11	Lack of incentives for voluntarily publishing open data	Open data providers (including government agencies) have limited incentives for voluntarily publishing open data without legal force.
C12	Requiring viable business models	Running initiatives or businesses based on open data requires a viable business model.
C13	Overlooked nongovernment open data	Open data provided by organizations apart from the government (e.g., businesses and nonprofit organizations) may not be visible to open data users.
C14	Practical constraints in multistakeholder engagement	Open data stakeholders face practical constraints to conducting multistakeholder engagements, such as a lack of time, resources, and commitment.
C15	Lack of data awareness*	Some segments of society have a limited general understanding of the role of data, different types of data, the application and value of open data, ethical issues around data, and privacy aspects.
	Skills-based	
C16	Limited knowledge of open data providers	Certain open data providers lack knowledge regarding the legal requirements and best practices for publishing open data.
C17	Limited knowledge of (potential) open data users	Potential or current open data users have limited knowledge regarding open data.

(Continued)

Table 3. *Continued*

ID	Challenges in ODE	Description
	Foundational	
C18	Poor open data quality*	Certain types of open data have low or inconsistent quality.
C19	Incurring maintenance costs for open data providers	Maintaining the infrastructure and human resources to publish and update open data and engage stakeholders incurs costs for data providers.
C20	Incurring development costs for open data providers	Developing the infrastructure and human resources to publish and update open data and engage stakeholders incurs costs for data providers.
C21	Technical difficulties in establishing open data management systems	Certain open data providers face significant technical bottlenecks or difficulties in establishing open data management systems.
C22	Complex and/or rigid open data standards to comply with	Some open data standards are too complex and/or rigid for open data providers to comply with.
C23	Heterogeneous data administration	Open data are provided by different data administrations that operate under different systems/jurisdictions.
C24	Privacy concerns	Open data providers have to address privacy concerns that may not be easily identifiable and addressable.
	Broad	
C25	Inflexible/unclear government–market boundary	Open data providers among government agencies are uncertain or restricted in terms of the extent to which they can offer value-added products using their open data.
C26	Reliance on a political agenda	Some open data initiatives (not only by government agencies) rely on the political agendas of their funders.
C27	Inflexible governance/law (esp. with evolving technology)	Laws and bureaucracy related to open data publication may take a long time to change, thus being unable to keep up with rapidly evolving technologies and data applications.

Note. * were suggested by validators.

political factors beyond open data (categorized as "broad"). A sample quote from the interviews for each of the challenges is presented in the Supplementary Appendix. Challenges indicated by an asterisk (*) in Table 3 were suggested by validators.

Challenges related to user-drivenness: Open data users must deal with different open datasets published under various standards and licenses. As highlighted by an interviewee, the EU Directive on Establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) states that spatial data published by national mapping agencies in the EU should follow ISO standards. However, generic administrative open data are rarely published in those standards; instead, they are published in other standards (e.g., DCAT). Simultaneously, to use and possibly integrate open data from different sources for specific purposes, users must check the conditions of the relevant data licenses, and whether they are compatible. This task is not necessarily straightforward, especially for users with limited legal literacy. These technical and legal interoperability issues are exacerbated by siloed open data domains (geospatial, demographic statistics, etc.), which can be characterized as loosely connected ecosystems within the larger ODE.

Using open data may require complex technical know-how and tooling, which particularly impacts nonspecialist users (including experts on a particular domain but not others). Simultaneously, with the growing (potential) applications of open data, the diverse needs of users (e.g., in terms of the data and the format they require) are not entirely met. While some open data providers have taken the initiative to seek

user feedback, limited input is obtained from nonexpert users—especially among small and medium enterprises and individual citizens.

Challenges related to circularity: Open data providers that previously sold their data lost (one of) their source(s) of income upon providing it as open data. In the EU, many organizations are legally required to publish open data even though it would not have been in line with their managerial or strategic decisions. However, open data providers receive limited value returns (particularly monetary) from the open data they provide. Meanwhile, open data use cases are not always visible. Thus, open data providers cannot fully assess the value of their open data to decide what data they (do not) need to provide and prove its value to seek (more) funding.

Challenges related to inclusivity: Nongovernment sectors (i.e., the private and civil sectors) provide limited open data, which is partly attributed to the lack of incentives to voluntarily publish open data. As noted by at least two interviewees, the legal obligation in the EU entails a major push for open data; however, this mainly affects the public sector at present. Running initiatives based on open data—as providers, users, and intermediaries—requires viable business models, which remain underdeveloped. One interviewee also noted that open data from nongovernment sectors are available in some cases but simply not visible enough. Multistakeholder engagements are necessary to encourage nongovernment sectors to contribute open data and participate in decision-making related to open data. However, individual actors face practical constraints in leading such engagements due to limited resources, time, and commitment.

Challenges related to skills: Some existing or potential open data providers have limited knowledge of the best practices for publishing open data. In the context of the EU, certain open data providers struggle to meet the requirements set by the law due to the lack of technical expertise in the organization. Conversely, potential or current open data users may have limited knowledge and skills related to using open data in a meaningful manner.

Foundational challenges: Publishing open data comes with considerable costs for providers linked to developing and maintaining the relevant infrastructure and expertise. This is especially the case for organizations that had underdeveloped data management systems before they had/decided to publish open data. In the development stage, open data providers must deal with major technical undertakings to build open data infrastructure and processes. Additionally, some open data providers are legally required to publish data in specific data models that are rather complex, such as in the case of the EU INSPIRE Directive, which was noted by at least two interviewees. Certain open data providers may also need to integrate data from different administrations before publishing it as open data, which involves significant coordination efforts. Furthermore, as the volume of open data grows, privacy concerns may become more challenging to address.

Broad challenges: Certain challenges in ODEs are an extension of broader political factors beyond open data alone, particularly surrounding policies around market competition, overarching digital strategies, and technological governance. Different governments and societies may have different preferences and approaches to these aspects, which ultimately affect the ODE. In the EU, decisions in these areas are also negotiated and made at the supranational level.

4.2. *Potential contributions of open data intermediaries*

Table 4 presents the potential contributions of open data intermediaries derived from the interviews. We categorized the contributions into technical, nontechnical, and combination contributions. Technical contributions may include the direct processing of open data and developing tools to facilitate its supply or use. Nontechnical contributions do not necessarily involve directly processing open data or developing tools but are more geared toward relationship building, stakeholder engagement, and financial support. Combination contributions are those that involve a mix of technical and nontechnical activities. Moreover, these potential contributions may be interrelated and some of them are not necessarily new. For instance, fostering collaborations is a known contribution of open data intermediaries, as reviewed by Shaharudin et al. (2023). However, the word "potential" refers to the potentiality of these contributions in

Table 4. *Potential contributions of open data intermediaries*

ID	Potential contributions of open data intermediaries	Description
	Technical	
P01	Implement federated architecture	Develop and maintain federated architecture open data systems.
P02	Integrate data (e.g., across sectors/ administration)	Integrate open data from different sectors and administrations.
P03	Transform data into open standards (esp. web standards)	Transform open data in certain standards into open standards to proliferate its reuse.
P04	Customize data (based on use cases)	Transform open data into formats familiar to certain professionals, domains, or use cases.
P05	Offer process automation	Offer applications/software that provide process automation in data publication and/or use.
P06	Develop open-source tooling	Develop open-source tooling to be used by open data providers or users.
P07	Provide direct technical services*	Provide direct technical services to open data providers or users (beyond offering tooling or preprocessed data).
P08	Offer freemium data platform*	Offer a one-stop platform for data providers to provide open data together with their paid data.
	Nontechnical	
P09	Foster public–private collaboration	Initiate and coordinate collaboration between the government and businesses.
P10	Foster public–civic collaboration	Initiate and coordinate engagement between the government and nongovernmental organizations (NGOs).
P11	Implement multistakeholder collaboration	Initiate and coordinate the collaboration between the government, businesses, and NGOs.
P12	Perform open data advocacy*	Conduct advocacy/lobbying activities related to open data issues.
P13	Invest in open data-based civic tech	Invest in businesses or initiatives that offer open data-based applications or programs that benefit the community, in line with societal goals.
P14	Showcase open data value	Demonstrate open data use cases and the value of open data.
P15	Promote open nongovernmental data	Identify open data from organizations outside the government sector and promote their reuse.
	Combination	
P16	Provide consultancy	Advise and help open data providers and/or users with aspects related to the publication and use of open data.
P17	Streamline cross-administrative processes	Advise and coordinate efforts to streamline data management processes across different administrations.
P18	Facilitate internal reuse of open data	Identify and recommend the internal reuse of open data.
P19	Facilitate feedback on open data	Help users provide feedback to open data providers.
P20	Provide education on data literacy and skills*	Provide training and other education activities to improve data literacy and skills.

Note. * were suggested by validators.

addressing specific challenges in the ODE. Table 5 links the potential contributions of open data intermediaries with the ODE challenges they may address.

Technical contributions: By implementing federated architecture, open data intermediaries may help address user-driven challenges (different data standards, different open data licenses, and siloed open data domains) and the foundational challenges of data management systems. Integrating data may also help resolve issues of different data standards, siloed open data domains, the high technical thresholds for nonspecialist users, and unfulfilled user needs (due to fragmented data). It may also help address the limited skills of data providers and users in combining different datasets and the problem of open data being provided by heterogeneous administrations. Transforming open data into open standards, especially web standards that are already used widely across many domains, may address issues similar to integrating data. It may also help to overcome issues where data are published according to standards set by the law, but the standards are not adaptive to changing technology and user needs. Likewise, customizing data based on the common use cases in a specific industry or domain may also help to address the same problem of misaligned legal development vis-à-vis technological progress. It may also address some challenges related to skills and user-drivenness (different data standards, siloed open data domains, high technical threshold, and unfulfilled user needs) but not foundational challenges.

Open data intermediaries may also offer automation for certain open data publishing and use processes. This may help to overcome challenges around user-drivenness (siloed open data domains, high technical threshold, and unfulfilled user needs), skills, and foundational challenges (including by enhancing privacy protection). Offering open-source tooling may reduce the high technical threshold for nonspecialist users through an affordable means. This would enable the implementation of certain business models by open data actors that do not have the technical expertise and proprietary in-house tooling or financial capacity to acquire such resources. Open data intermediaries may also provide direct technical services in certain parts of open data processing (instead of providing tooling and pre-processed data). This may resolve challenges linked to the high technical threshold for using open data, limited skills among open data providers and users, and several foundational challenges (poor open data quality, technical difficulties in establishing data management systems, and complex data standards to comply with). Additionally, open data intermediaries may offer a one-stop platform for providers to offer their open data together with paid data, thereby encouraging (nongovernment) actors to provide open data through the freemium business model (addressing inclusivity challenges).

Nontechnical contributions: By fostering public–private, public–civic, and multistakeholder collaboration, open data intermediaries may help to address inclusivity, user-drivenness, circularity, and broad foundational challenges. While most links between these three contributions and their challenges are quite straightforward, the connections between public–private/multistakeholder collaborations and the challenges of development/maintenance costs incurred by the open data providers may not be very clear. The idea is that the public, private, and civil sectors may pool resources and jointly coordinate efforts to build and maintain technical infrastructure to publish open data (whether the open data are from the public, private, or civil sectors or a combination of them), thereby reducing the costs incurred for a single organization. This form of collaboration may work if all parties are interested in publishing specific open data, such as data showing the locations of electric vehicle charging stations. Open data intermediaries may facilitate this type of collaboration.

Open data intermediaries may undertake open data advocacy, among other activities, to raise awareness of the loss of open data providers' revenue and lobby for additional funding based on the socioeconomic value of open data. Open data intermediaries may also invest in and provide business model design support to civic technology companies producing applications that reduce the technical threshold for (lay) users to benefit from open data. They may return a share of the profits to open data providers (e.g., through joint ventures) and encourage more open data releases. Furthermore, open data intermediaries may contribute to showcasing the (critical) value of open data to address circularity, inclusivity, foundational, and broad political challenges. By promoting open nongovernment data, open data intermediaries may help address most of the inclusivity challenges as well as unfulfilled user needs (user-drivenness) and the limited knowledge of (potential) open data users (skills-based).

Table 5. *Potential contributions of open data intermediaries and the ODE challenges they may address*

Potential contributions of open data intermediaries (column) / Challenges in the ODE (row)	Technical contributions								Nontechnical contributions							Combination contributions				
	P01: Implement federated architecture	P02: Integrate data	P03: Transform data into open standards	P04: Customize data	P05: Offer process automation	P06: Develop open-source tooling	P07: Provide direct technical services	P08: Offer freemium data platform	P09: Foster public-private collaboration	P10: Foster public-civic collaboration	P11: Implement multistakeholder collaboration	P12: Perform open data advocacy	P13: Invest in open data-based civic tech	P14: Showcase open data value	P15: Promote open non-governmental data	P16: Provide consultancy	P17: Streamline cross-administrative processes	P18: Facilitate internal reuse of open data	P19: Facilitate feedback on open data	P20: Provide education on data literacy and skills
User-drivenness challenges																				
C01: Different data/metadata standards	X	X	X	X					X		X						X			
C02: Different open data licenses	X										X									
C03: Siloed open data domains (e.g., across sectors)	X	X	X	X	X				X		X						X			
C04: High technical threshold for users		X	X	X	X	X	X						X			X				
C05: Unfulfilled user needs		X	X	X	X				X	X	X			X		X			X	
C06: Limited feedback from lay users (i.e., nonexpert users)									X	X	X					X			X	
Circularity challenges																				
C07: Loss of open data providers' revenue												X	X	X						
C08: Limited value return from data reuse													X	X			X	X	X	
C09: Limited use case visibility									X	X	X			X		X	X			X

Inclusivity challenges

Challenge	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
C10: Limited open data from nongovernment sectors								X	X	X	X	X		X						
C11: Lack of incentives for voluntarily publishing open data								X	X	X	X	X	X	X			X			
C12: Requiring viable business model						X		X	X	X		X	X		X					
C13: Overlooked nongovernment open data								X	X	X			X	X						
C14: Practical constraints in multistakeholder engagement								X	X	X				X	X			X		
C15: Lack of data awareness								X	X	X		X	X	X						X

Skills-based challenges

Challenge	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
C16: Limited knowledge of open data providers		X		X	X		X							X	X		X		X	X
C17: Limited knowledge of (potential) open data users		X	X	X	X		X							X	X		X			X

Foundational challenges

Challenge	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
C18: Poor open data quality						X								X				X		
C19: Incurring maintenance costs for open data provider	X			X	X			X		X		X		X	X					
C20: Incurring development costs for open data provider	X			X	X			X		X		X		X	X					
C21: Technical difficulties in establishing data management systems	X			X		X								X	X					
C22: Complex and/or rigid open data standards to comply with	X			X		X								X						
C23: Heterogeneous data administration	X	X	X	X						X				X	X					
C24: Privacy concerns				X					X	X				X					X	X

Broad challenges

Challenge	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
C25: Inflexible/unclear government–market boundary											X			X						
C26: Reliance on a political agenda								X	X	X	X	X		X	X		X	X		
C27: Inflexible governance/law (esp. with evolving technology)			X	X							X			X						

Combination contributions: Open data intermediaries may provide various types of consultancy to open data providers and users, such as those related to technical, managerial, economic, and political aspects. The technical and managerial types of consultancy may help open data providers and users to overcome user-drivenness, skills-based, circularity, and foundational challenges. The economic and political types of consultancy may help to address the challenges faced by open data providers and users linked to requiring viable business models and mitigating broad political challenges. Open data intermediaries may also help streamline cross-administrative data management processes to tackle various issues across all dimensions, particularly foundational issues.

Open data intermediaries may facilitate the internal reuse of open data to address shortcomings of limited value return from open data, lack of incentives for publishing open data voluntarily, reliance on a political agenda, and limited knowledge of open data providers and users. For this contribution, the open data intermediary may likely be a unit within the same organization as the open data provider to have comprehensive knowledge of the data produced and used by the organization. Open data intermediaries may also facilitate feedback on open data, which may tackle some challenges across all dimensions. This facilitation of feedback may involve technical aspects (e.g., introducing a feedback feature in a data platform) or nontechnical aspects (e.g., organizing user group meetings). Additionally, open data intermediaries may provide education on data literacy and skills. Apart from tackling skills-based challenges, this education may aim to improve general public data awareness, such as on different types of data, the potential applications of open data, ethical ways of using data, and privacy aspects.

5. Discussion

This study affirms the role of open data intermediaries in addressing challenges in the ODE. Our findings show that not only could they help the ecosystem strengthen the four features of a sustainable ODE as proposed by van Loenen et al. (2021), but they could also help address foundational issues around data management systems and mitigate the broad political factors impacting ODEs. However, it is important to note that these potential contributions addressing the various challenges in ODEs would not magically materialize. To reiterate, in the ODE, actors are considered self-interested and open data intermediaries require internal incentives (e.g., through viable business models) and/or extrinsic incentives (e.g., through policies and regulations) to drive them to offer those contributions. Therefore, this study emphasizes further inquiry into the business models of open data intermediaries and other external conditions of the ODE that could encourage open data intermediaries to act in ways that support the sustainability of the ODE, in which their interests and other actors are aligned. Additionally, this study also encourages further investigations of whether some of the identified potential contributions of open data intermediaries already exist *and* address the ODE challenges linked to them (or others) and, if so, what the success factors are (and vice versa). Moreover, further studies could identify other potential contributions of open data intermediaries beyond those listed in this paper.

This study highlights the need to distinguish a "role" from an "actor" in the ecosystem, as elucidated by Oliveira and Lóscio (2018) and Shaharudin et al. (2023). Diverse ODE actors can undertake open data intermediation, including public organizations, for-profit companies, and civil society organizations. It is especially worth emphasizing that open data intermediaries do not only exist outside of the public sector, as some have implied (Balvert and van Maanen, 2019; Schrock and Shaffer, 2017). Several public organizations in our research, such as SDFI (Denmark), BKG (Germany), and IGN (France), have identified some of their tasks as being those of open data intermediaries. One of the interviewees mentioned, "We do not necessarily produce the data, but we [are] intermediaries ourselves. So we get the data from others, in particular from the official mapping agencies [redacted], and we process that data, combine it, and provide it." Similarly, during the validation exercise, a public organization representative noted how they provide a paid service to help smaller public agencies fulfill the technical data requirements set by the law. Having said that, what a government organization can do as an open data intermediary may differ from what a for-profit company or a civil society organization can do in that role. This is due to the different legal obligations, societal expectations, resources, and other factors that these

organizations have. Hence, the contributions of diverse types of open data intermediaries that are possible in practice are also different.

Our findings also suggest that an open data provider or user could benefit from the contributions of multiple open data intermediaries simultaneously, in parallel and/or sequentially. As van Schalkwyk et al. (2016, p. 22) argued, "No single intermediary is likely to possess all the types of capital required to unlock the full [open data] value." This also implies that the beneficiary of an open data intermediary's contributions could also be another open data intermediary. This insight highlights what Oliveira and Lóscio (2018) emphasized: value is not created in a chain but instead in a network in which actors can participate in multiple ecosystems. A chain refers to linear pathways where value is transferred from one actor to the next in a linear sequence, whereas a network refers to complex, nonlinear pathways where value is transferred from and to multiple actors in various directions. This also means that there could be multiple orderings of open data providers, intermediaries, and users in the ecosystem. Thus, apart from provider–intermediary–user relationships, they could also take the form of provider–intermediary–intermediary–user or provider–parallel intermediaries–user relationships, among others. Thus, open data intermediaries are not merely the "bridge" between open data providers and users.

Some of the challenges of the ODE identified from our interviews are very foundational issues around open data management systems or related to broader political factors and do not necessarily fit into the four features of a sustainable ODE suggested by van Loenen et al. (2021). This implies that the four features are inadequate to assess the sustainability of an ODE. Thus, additional layers of criteria may be necessary and deserve future attention. At the very least, the four features noted by van Loenen et al. (2021) must be clarified or refined to readily incorporate those foundational and broader political issues around open data. Some proposals for an ODE sustainability assessment framework have been made in the past (Vancauwenberghe, 2018; Welle Donker and van Loenen, 2017) but they were largely limited to the context of open government data and, more importantly, the object of assessment is usually the data, whereas an ODE centers the relationships between actors and the value flows instead of the data per se.

Our findings on the challenges of ODEs and the potential contributions of open data intermediaries call the boundaries of ODEs into question. Notably, the ODE challenges identified include those associated with external factors beyond just open data, and the contributions of open data intermediaries may involve many more activities than directly handling open data. Drawing from ANT, Latour (2007, p. 29) asserted that "it's not the sociologist's duty to decide *in advance* [emphasis added] and in the member's stead what the social world is made of" since "social aggregates are not the object of ostensive definition—like mugs and cats and chairs that can be pointed at by the index finger—but only of a performative definition" (p. 34). Latour emphasized tracing connections "instead of being constantly bogged down in the impossible task of deciding once and for all what is the right unit of analysis" (p. 34). In other words, the boundaries of an ODE do not have to (or cannot) be defined beforehand. However, for a specific assessment, research inquiry, or intervention, one should identify relevant actors, trace their associations, and analyze their interactions, which would involve making and remaking boundaries (Harrison et al., 2012; Lee and Hassard, 1999). Similarly, van Schalkwyk et al. (2015) argued that to determine whether an actor can be considered an open data intermediary, one should assess its "degree of agency" in fulfilling the open data intermediation function. Having said that, striking a balance between being pragmatic and reductionist is undoubtedly something that open data researchers and practitioners would have to constantly grapple with.

6. Conclusion

This paper has explored the potential contributions of open data intermediaries in addressing various challenges in ODEs. Through interviews and a validation exercise conducted with 19 individuals from 15 organizations in the EU, it has been shown that open data intermediaries could help overcome challenges that are detrimental to ODE sustainability through various technical and nontechnical contributions. For example, they may implement a federated architecture data platform, offer process automation, develop open-source tooling, invest in open data-based civic technology companies, promote

open nongovernmental data, streamline cross-administrative processes, and facilitate the internal reuse of open data. This paper also demonstrated the potential efficacy of adopting the ecosystem perspective, inspired by ANT, as the analytical lens to generate insights toward improving value generation from open data.

This paper contributes to the body of knowledge on ODEs by exploring the links between the potential contributions of open data intermediaries and the challenges in ODEs. These contributions of open data intermediaries would not automatically resolve the challenges in ODEs linked to them by default; instead, they have to be designed for that purpose. As stressed in this paper, actors in ODEs are self-interested; hence, they require intrinsic incentives (e.g., through viable business models) and/or external conditions (e.g., through policies and regulations) that drive them toward acting in a particular manner. Hence, research into business models and policies relevant to open data intermediaries is necessary while considering the diverse types of actors serving the role of open data intermediaries (e.g., government organizations, companies, and civil society organizations). Notably, different types of actors would require different sets of incentives. Additionally, further research is required to investigate whether some of the contributions of open data intermediaries identified in this study already exist, address the ODE challenges linked to them (or others), and which conditions allow (or prevent) such scenarios. In conclusion, this paper paves the way for further inquiry and intervention related to open data intermediaries so that they can contribute toward sustainable ODEs.

References

Andrason A and van Schalkwyk F (2017) *Opportune Niches in Data Ecosystems: Open Data Intermediaries in the Agriculture Sector in Ghana* (SSRN Scholarly Paper No. 2949722). Rochester: Social Science Research Network. Retrieved from https://papers.ssrn.com/abstract=2949722

Balvert A and van Maanen G (2019) Open for whom? The role of intermediaries in data publication. *Revista Publicum 5*, 129–159. 10.12957/publicum.2019.47205.

Barry E and Bannister F (2014) Barriers to open data release: A view from the top. *Information Polity 19*, 129–152. 10.3233/IP-140327.

Bonina C (2013) *New business models and the value of open data: definitions, challenges and opportunities*. London: EPSRC, New Economic Models in the Digital Economy (NEMODE+). Retrieved from https://zenodo.org/record/1322706

Cahlikova T and Mabillard V (2020) Open data and transparency: opportunities and challenges in the Swiss context. *Public Performance & Management Review*. 10.1080/15309576.2019.1657914 (accessed 14 November 2023).

Callon M and Law J (1997) After the individual in society: lessons on collectivity from science, technology and society. *Canadian Journal of Sociology/Cahiers Canadiens de Sociologie 22*, 165. 10.2307/3341747.

Camero J (2015) House Okays National Mangrove Forests Conservation and Rehabilitation Act. *Republic of the Philippines House of Representatives Press Releases.* https://www.congress.gov.ph/press/details.php?pressid=8604 (accessed 13 November 2023).

Chan M, Johnson P and 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*, 1–27. 10.29379/jedem.v8i1.370.

Chattapadhyay S (2014) Access and use of government data by research and advocacy organisations in India: a survey of (potential) open data ecosystem. In *Proceedings of the 8th International Conference on Theory and Practice of Electronic Governance*. Guimaraes Portugal: ACM, 361–364. 10.1145/2691195.2691262.

Conradie P and Choenni S (2014) On the barriers for local government releasing open data. *Government Information Quarterly 31*, S10–S17. 10.1016/j.giq.2014.01.003.

Corbett J, Templier M and Takeda H (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*. 10.24251/HICSS.2018.308.

Corbett J, Templier M, Townsend H and Takeda H (2020) Integrating across sustainability, political, and administrative spheres: A longitudinal study of actors' engagement in open data ecosystems in three Canadian cities. *Communications of the Association for Information Systems 47*(1). 10.17705/1CAIS.04728.

Crusoe J, Simonofski A, Clarinval A and Gebka E (2019) The impact of impediments on open government data use: insights from users. In *2019 13th International Conference on Research Challenges in Information Science (RCIS)*, pp. 1–12. 10.1109/RCIS.2019.8877055.

Csáki C (2019) Open data ecosystems: a comparison of visual models. In Kő A, Francesconi E, Anderst-Kotsis G, Tjoa AM, and Khalil I (eds.), *Electronic Government and the Information Systems Perspective*. Cham: Springer, 16–30. 10.1007/978-3-030-27523-5_2.

Davies T (2011) *Open Data: Infrastructures and Ecosystems.* Presented at the ACM Web Science Conference, Evanston, Illinois.

Díaz Andrade A and Urquhart C (2010) The affordances of actor–network theory in ICT for development research. *Information Technology & People 23*, 352–374. 10.1108/09593841011087806.

Duguay R, Rauter T and Samuels D (2023) The impact of open data on public procurement. *Journal of Accounting Research 61* (4), 1159–1224. 10.1111/1475-679X.12479.

Dumpawar S (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

Enaholo P (2017) Beyond mere advocacy: CSOs and the role of intermediaries in Nigeria's open data ecosystem. In van Schalkwyk F, Verhulst S, Magalhaes G, Pane J and Walker J (eds.), *The Social Dynamics of Open Data.* Cape Town, South Africa: African Minds, pp. 89–108. http://www.africanminds.co.za/the-social-dynamics-of-open-data/

Enaholo P and Dina D (2020) Journalists and the intermediation of open data: a Nigerian perspective. In *Situating Open Data: Global Trends in Local Contexts*. Cape Town: African Minds, pp. 31–49. https://www.africanminds.co.za/situation-open-data-global-trends-in-local-contexts/#:~:text=Open%20data%20and%20its%20effects,focus%20on%20its%20situated%20use

European Commission (2023) Press Release: 4 Member States referred to the Court of Justice of the EU. European Commission. https://ec.europa.eu/commission/presscorner/detail/en/IP_23_706 (accessed 21 February 2024)

European Parliament and Council of the European Union Directive (EU) (2019) (2019/1024 of the European Parliament and of the Council of 20 June 2019 on Open Data and the Re-use of Public Sector information., 172 OJ L. http://data.europa.eu/eli/dir/2019/1024/oj/eng (accessed 10 August 2022).

European Parliament and Council of the European Union Regulation (EU) (2022) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European Data Governance and Amending Regulation (EU) 2018/1724 (Data Governance Act)., 152 OJ L. http://data.europa.eu/eli/reg/2022/868/oj/eng (accessed 1 March 2024).

Germano EC, de Souza CA and Sun V (2016) Business models adopted by intermediaries in the use of open government data. *REBRAE 9*(1), 79–94. 10.7213/rebrae.09.001.AO05.

González-Zapata F and Heeks R (2015) *Understanding Multiple Roles of Intermediaries in Open Government Data.* Presented at the 13th International Conference on Social Implications of Computers in Developing Countries, Negombo, Sri Lanka.

Goodchild B and Ferrari E (2024) Intermediaries and mediators: an actor–network understanding of online property platforms. *Housing Studies 39*, 102–123. 10.1080/02673037.2021.2015297.

Hablé J (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

Harrison TM, Pardo TA and Cook M (2012) Creating open government ecosystems: a research and development agenda. *Future Internet 4*, 900–928. 10.3390/fi4040900.

Heimstädt M, Saunderson F and Heath T (2014) From toddler to teen: Growth of an open data ecosystem. *JeDEM – eJournal of eDemocracy and Open Government 6*, 123–135. 10.29379/jedem.v6i2.330.

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

Janssen M, **Charalabidis Y and Zuiderwijk A** (2012) Benefits, adoption barriers and myths of open data and open government. *Information Systems Management 29*, 258–268. 10.1080/10580530.2012.716740.

Johnson P (2023) The Latest Data Confirms: Forest Fires Are Getting Worse. https://www.wri.org/insights/global-trends-forest-fires (accessed 13 November 2023).

Johnson PA and Greene S (2017) Who are government opendata infomediaries? A preliminary scan and classification of open data users and products. *Journal of the Urban and Regional Information Systems Association 28*, 10.

Johnson PA, **Sieber R**, **Scassa T**, **Stephens M and Robinson P** (2017) The cost(s) of geospatial open data. *Transactions in GIS 21*, 434–445. 10.1111/tgis.12283.

Kim D (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

Klecuń E (2004) Conducting critical research in information systems: can actor–network theory help? In Kaplan B, Truex DP, Wastell D, Wood-Harper AT, and DeGross JI (eds.), *Information Systems Research: Relevant Theory and Informed Practice*. Boston: Springer, 259–274. 10.1007/1-4020-8095-6_15.

Latour B (2007) *Reassembling the Social: An Introduction to Actor–Network-Theory*. Oxford: Oxford Universtiy Press.

Law J (2008) Actor network theory and material semiotics. In Turner BS (ed.), *The New Blackwell Companion to Social Theory*. Chichester: Wiley-Blackwell, pp. 141–158.

Lee N and Hassard J (1999) Organization unbound: actor–network theory, research strategy and institutional flexibility. *Organization 6*(3), 391–404. 10.1177/135050849963002.

Linders D (2013) Towards open development: Leveraging open data to improve the planning and coordination of international aid. *Government Information Quarterly 30*, 426–434. 10.1016/j.giq.2013.04.001.

Linneberg MS and Korsgaard S (2019) Coding qualitative data: a synthesis guiding the novice. *Qualitative Research Journal 19*, 259–270.

Ma R and Lam PTI (2019) Investigating the barriers faced by stakeholders in open data development: a study on Hong Kong as a "smart city". *Cities 92*, 36–46. 10.1016/j.cities.2019.03.009.

Magnusson E and Marecek J (2015) *Doing interview-based qualitative research: a learner's guide*. Cambridge: Cambridge University Press.

Mayer-Schönberger V and Zappia Z (2011) Participation and power: Intermediaries of open data. *Presented at the 1st Berlin Symposium on Internet and Society*, Berlin: Alexander von Humboldt Institut für Internet und Gesellschaft (HIIG). https://www.hiig.de/wp-content/uploads/2017/08/participation_and_power.pdf (accessed 28 January 2022).

Meijer A and Potjer S (2018) Citizen-generated open data: an explorative analysis of 25 cases. *Government Information Quarterly 35*, 613–621. 10.1016/j.giq.2018.10.004.

Meng A (2016) *The Social Impact of Open Government Data* (PhD thesis). Georgia Institute of Technology. Retrieved from https://smartech.gatech.edu/handle/1853/56292

Meng A, **DiSalvo C**, **Tsui L and Best M** (2019) The social impact of open government data in Hong Kong: umbrella movement protests and adversarial politics. *The Information Society 35*, 216–228. https://doi.org/10.1080/01972243.2019.1613464.

Molloy JC (2011) The open knowledge foundation: open data means better science. *PLOS Biology 9*, e1001195. 10.1371/journal.pbio.1001195.

Moloney A (2021) Peru's Indigenous Tribes Use Tech to Track Amazon Deforestation. Thomson Reuters Foundation News. https://news.trust.org/item/20210712174822-u31bt/ (accessed 13 November 2023).

Müller M (2015) Assemblages and actor–networks: rethinking socio-material power, politics and space. *Geography Compass 9*, 27–41. 10.1111/gec3.12192.

Nehemia-Maletzky M, **Iyamu T and Shaanika I** (2018) The use of activity theory and actor network theory as lenses to underpin information systems studies. *Journal of Systems and Information Technology 20*, 191–206. 10.1108/JSIT-10-2017-0098.

Neves FT, **de Castro Neto M and Aparicio M** (2020) The impacts of open data initiatives on smart cities: a framework for evaluation and monitoring. *Cities 106*, 102860. 10.1016/j.cities.2020.102860.

Nikiforova A and Zuiderwijk A (2022) Barriers to openly sharing government data: towards an open data-adapted innovation resistance theory. In *Proceedings of the 15th International Conference on Theory and Practice of Electronic Governance*. New York: Association for Computing Machinery, pp. 215–220. 10.1145/3560107.3560143.

OKF (2013) Defining Open Data – Open Knowledge Foundation blog. https://blog.okfn.org/2013/10/03/defining-open-data/ (accessed 28 June 2022).

Oliveira MIS and Lóscio BF (2018) What is a data ecosystem? In *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*. Delft: ACM, pp. 1–9. 10.1145/3209281.3209335.

Park S and Gil-Garcia JR (2017) Understanding transparency and accountability in open government ecosystems: The case of health data visualizations in a State Government. In *Proceedings of the 18th Annual International Conference on Digital Government Research*. New York: Association for Computing Machinery, pp. 39–47. 10.1145/3085228.3085318.

Perkmann M and Schildt H (2015) Open data partnerships between firms and universities: the role of boundary organizations. *Research Policy 44*, 1133–1143. 10.1016/j.respol.2014.12.006.

Poikola A, **Kola P** and **Hintikka** (2011) *PublicDdata: An Introduction to Opening Information Resources*. Helsinki: Ministry of Transport and Communications. https://core.ac.uk/download/pdf/198191501.pdf (accessed 9 August 2022).

Sangiambut S and Sieber R (2017) The civic open data and crowdsourcing app ecosystem: actors, materials, and interventions. *Journal of the Urban & Regional Information Systems Association 28*, 49–62.

Schrock A and Shaffer G (2017) Data ideologies of an interested public: a study of grassroots open government data intermediaries. *Big Data & Society 4*, 2053951717690750. 10.1177/2053951717690750.

Shaharudin A, **van Loenen B and Janssen M** (2023) Towards a common definition of open data intermediaries. *Digital Government: Research and Practice 4*, 6:1–6:21. 10.1145/3585537.

Sugg Z (2022) Social barriers to open (water) data. *WIREs Water 9*, e1564. 10.1002/wat2.1564.

Susha I, **Janssen M and Verhulst S** (2017) Data collaboratives as a new frontier of cross-sector partnerships in the age of open data: taxonomy development. In *Hawaii International Conference on System Sciences 2017 (HICSS-50)*. https://aisel.aisnet.org/hicss-50/eg/open_data_in_government/4

Temiz S, **Holgersson M**, **Björkdahl J and Wallin MW** (2022) Open data: lost opportunity or unrealized potential? *Technovation 114*, 102535. 10.1016/j.technovation.2022.102535.

Toogood S (2021) *Barriers to Implementation of Open Data Strategies in Small, Medium, and Large Municipalities in British Columbia* (Thesis). Retrieved from https://dspace.library.uvic.ca/handle/1828/13641

Tyukavina A, **Potapov P**, **Hansen MC**, **Pickens AH**, **Stehman SV**, **Turubanova S**, **Parker D**, **Zalles V**, **Lima A**, **Kommareddy I**, **Song X-P**, **Wang L and Harris N** (2022) Global trends of forest loss due to fire from 2001 to 2019. *Frontiers in Remote Sensing 3*, 1–20. https://www.frontiersin.org/articles/10.3389/frsen.2022.825190 (accessed 24 November 2023).

van Loenen B, **Zuiderwijk A**, **Vancauwenberghe G**, **Lopez-Pellicer FJ**, **Mulder I**, **Alexopoulos C**, **Magnussen R**, **Saddiqa M**, **Rosnay MD de**, **Crompvoets J**, **Polini A**, **Re B and Flores CC** (2021) Towards value-creating and sustainable open data ecosystems: A comparative case study and a research agenda. *JeDEM – eJournal of eDemocracy and Open Government 13*, 1–27. 10.29379/jedem.v13i2.644.

van Schalkwyk F, **Cañares M**, **Chattapdhyay S and Andrason A** (2016) Open data intermediaries in developing countries. *The Journal of Community Informatics 12*. 10.15353/joci.v12i2.3219.

van Schalkwyk F, **Chattapadhyay S**, **Caňares M and Andrason A** (2015) *Open Data Intermediaries in Developing Countries*. Washington, DC: World Wide Web Foundation. Retrieved from http://hdl.handle.net/10625/56288

Vancauwenberghe G (2018) Assessing open data. In van Loenen B, Vancauwenberghe G and Crompvoets J (eds.), *Open Data Exposed*. The Hague: T.M.C. Asser Press, 149–172. 10.1007/978-94-6265-261-3_8.

Walsham G (1997) Actor–network theory and IS research: current status and future prospects. In Lee AS, Liebenau J and DeGross JI (eds.), *Information Systems and Qualitative Research: Proceedings of the IFIP TC8 WG 8.2 International Conference on Information Systems and Qualitative Research, 31st May–3rd June 1997, Philadelphia, Pennsylvania, USA.* Boston: Springer, pp. 466–480. 10.1007/978-0-387-35309-8_23.

Welle Donker F and van Loenen B (2017) How to assess the success of the open data ecosystem? *International Journal of Digital Earth 10*, 284–306. 10.1080/17538947.2016.1224938.

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

Zhu Z, **Wulder MA**, **Roy DP**, **Woodcock CE**, **Hansen MC**, **Radeloff VC**, **Healey SP**, **Schaaf C**, **Hostert P**, **Strobl P**, **Pekel J-F**, **Lymburner L**, **Pahlevan N and Scambos TA** (2019) Benefits of the free and open Landsat data policy. *Remote Sensing of Environment 224*, 382–385. 10.1016/j.rse.2019.02.016.

Zuiderwijk A, **Janssen M**, **Choenni S**, **Meijer R and Alibaks RS** (2012) Socio-technical impediments of open data. *Electronic Journal of E-Government 10*, pp. 156–172.

Zuiderwijk A and 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*, 377–395. 10.1108/TG-09-2020-0271.