



Article

Linked Open Government Data: Still a Viable Option for Sharing and Integrating Public Data?

Alfonso Quarati * and Riccardo Albertoni

Institute for Applied Mathematics and Information Technologies, National Research Council, 16149 Genoa, Italy; albertoni@ge.imati.cnr.it

* Correspondence: quarati@ge.imati.cnr.it

Abstract: Linked Data (LD) principles, when applied to Open Government Data (OGD), aim to make government data accessible and interconnected, unlocking its full potential and facilitating widespread reuse. As a modular and scalable solution to fragmented government data, Linked Open Government Data (LOGD) improve citizens' understanding of government functions while promoting greater data interoperability, ultimately leading to more efficient government processes. However, despite promising developments in the early 2010s, including the release of LOGD datasets by some government agencies, and studies and methodological proposals by numerous scholars, a cursory examination of government websites and portals suggests that interest in this technology has gradually waned. Given the initial expectations surrounding LOGD, this paper goes beyond a superficial analysis and provides a deeper insight into the evolution of interest in LOGD by raising questions about the extent to which the dream of LD has influenced the reality of OGD and whether it remains sustainable.

Keywords: Open Government Data; Linked Data; data friction



Citation: Quarati, A.; Albertoni, R. Linked Open Government Data: Still a Viable Option for Sharing and Integrating Public Data? *Future Internet* **2024**, *16*, 99. <https://doi.org/10.3390/fi16030099>

Academic Editor: Gianluigi Ferrari

Received: 7 February 2024

Revised: 11 March 2024

Accepted: 13 March 2024

Published: 15 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Open Government Data (OGD), traditionally sourced from governments, encompass public records in areas like transportation, infrastructure, education, health, and the environment [1]. These data are intended to be reused and redistributed, either for free or at marginal cost, to create new business opportunities, increase government transparency and accountability, foster citizen engagement, promote economic growth, reduce costs and efficiencies, and support innovation [2–5]. Since the launch of the Data.gov portal in the United States in 2009, which stands out as one of the pioneering and influential OGD platforms globally [4,6], several countries have followed suit and launched their OGD portals in the following years, resulting in an enormous amount and variety of datasets available worldwide [7] (p. 151). Thanks to these endeavors, a Deloitte-conducted study commissioned by the European Commission foresees a remarkable surge in the overall direct economic value of public sector information—from EUR 52 billion in 2018 to a whopping EUR 215 billion in 2028 [8]. OGD has applications in diverse sectors, including urban planning, environmental protection, security, mobility, and agriculture [9]. Numerous success stories highlight promising aspects, encouraging the dissemination of best practices that can be applied in similar contexts [10–12].

Despite the consistent growth in the publication of datasets and the increasing acknowledgment of the potential advantages of open and accessible government data among both governments and citizens, there are evident indications that the utilization of OGD remains constrained [13,14] and presents challenges [15], with stakeholders expressing concern that only a small proportion of their datasets are actively used [16]. This concern is supported by our recent research [17,18], which confirms that the vast majority of datasets published by government portals, whether international, national, or regional, are largely ignored by users who focus their attention on a limited number of government datasets.

Following Edward's metaphor of "data friction" which expresses "the costs in terms of time, energy and attention required simply to collect, control, store, move, receive and access data" [19] (p. 84), we can speak of "OGD friction" referring to the obstacles, difficulties, or resistance encountered in the access, use, and sharing of OGD. These frictions can "have both physical and social aspects" [19]. Regarding the first, OGD quality is essential for effective utilization by users, and it is considered a potential technical factor that can hinder or facilitate its exploitation [20–22]. In particular, metadata quality is particularly vital because it enables users to search and access dataset descriptions, ultimately enhancing the speed and ease of OGD use [18,23,24]. Low levels of open data literacy [25] are a significant social friction, indicating citizens' limited familiarity with the design and use of technology [26], and difficulties in accepting technology [27], a misalignment between user needs and the capabilities of available datasets [26,28].

To address these challenges and promote the reuse of OGD, several authors have suggested applying Linked Data (LD) principles to OGD [2,29–32]. Linked Open Government datasets offer the potential to "replace isolated data silos with larger, interconnected datasets built on top of the Web architecture" [33,34], fostering synergy among disparate data sources [35]. This approach allows users to gain a better understanding of the data's context by exploring related datasets [36,37]. As proposed by Tim Berners-Lee (<https://www.w3.org/DesignIssues/LinkedData.html> (accessed on 1 February 2024)) in 2010, the principles of LD offer a modular and scalable solution to counteract fragmentation in government data. This method not only enhances citizen awareness about government functions but also improves administrative efficiency. Furthermore, Linked Open Government Data (LOGD) enable access to and retrieval of data from authoritative and reliable sources. This approach can significantly reduce the inconsistencies found in legacy datasets, leading to more accurate and reliable results when querying open databases [38].

Under these auspices, in the early 2010s, a few government agencies began adopting LOGD principles and technologies [6,39], a movement that was bolstered by academic and research studies [40], prototypes [41–43], and methodological proposals [32,44,45].

However, despite the initially promising expectations, our recent research activities on OGD and LD, involving the analysis of various government portals in different capacities [18,25,46], have occasionally identified what seems to be a significant decline in interest in LOGD. This impression is corroborated by Penteado et al. [33], who recently observed, "Although the open government data movement is still producing large amounts of data worldwide, linked data still represents a tiny portion". Attard et al. also noted "Yet, the use of Linked Data in open government initiatives is still quite low" [34]. In 2020, Hogan also observed that despite the LD community's success in "convincing various stakeholders to publish data with the implicit promise that applications would justify the cost, these applications did not emerge, leading to the removal of many datasets and related services" [47]. In the adjacent field of Open Science, considering its role in supporting and improving "the discoverability, accessibility, shareability, reusability, reproducibility, and monitoring of data-driven research results on a global scale", the decision by OpenAire to abandon Linked Open Data (LOD) technologies is indicative of a similar trend. As reported on their website (<https://www.openaire.eu/pausing-our-lod-services> (accessed on 30 January 2024)), "starting from Monday, 8 May 2023, the SPARQL endpoint will shut down and that no new OpenAIRE LOD Dump versions will be released".

To illuminate the current state of LD practices adoption in OGD, our objective is to determine whether such adoption is unequivocally declining or if emerging practices and trends, even after nearly 15 years since its inception, still substantiate its foundational assumptions. Within the evolving landscape of the OGD movement, this paper seeks to address the titular question: *Are LOGD still a viable option for sharing and integrating public data?* To tackle this inquiry, we will examine three crucial snapshots: first, an analysis of LOGD adoption practices by national portals and the scientific community; second, a systematic literature survey probing potential factors impeding LOGD development; and third, an overview of current adoption practices—ranging from foundational to advanced—that

align with the expectations set forth in the early 2010s. This review aims to redirect the focus on the use of LOGD, assessing how they align with the current requirements of the OGD movement and potentially reshaping certain initial assumptions that may have proven challenging to implement.

2. Background

2.1. Open Government Data

According to the Open Knowledge Foundation (<https://opendefinition.org/> (accessed on 1 February 2024)), Open Data (OD) refers to data that is freely available for anyone to access, use, and share. OD core principles, including availability, accessibility, and reusability, promise economic, operational, political, and social benefits [11]. OD encompasses a wide range of information and is characterized by being machine-readable and non-discriminatory.

OGD is a subset of Open Data that focuses specifically on data generated or collected by governments [7,48]. It adheres to the principles of government transparency, accountability, and innovation. OGD aims to make government activities and decisions more visible to the public, encourage citizen engagement, and promote accountability.

The inception of government portals in the United States, initiated by President Obama's Open Government Initiative in 2009, marked a significant transformation in how government agencies at local, regional, and national levels developed their digital presence [4]. This movement led to the widespread adoption of OGD portals, which were established as the primary online platforms to act as authoritative sources of information about government activities [49]. These portals are designed to facilitate effective communication with citizens and various stakeholders, thereby enhancing the transparency and accessibility of government information. Integral to these efforts, OGD portals, supported by dedicated software platforms, such as CKAN (<https://ckan.org/> (accessed on 1 February 2024)) and Socrata (<https://dev.socrata.com/> (accessed on 1 February 2024)), play a vital role in adhering to data release policies set by governmental administrations. Portal managers utilize these platforms not only to publish datasets but also to employ specific metadata standards to categorize and organize information systematically. This organization allows users to benefit from government datasets, furnished with advanced search and browsing functionalities, ensuring they can access the data they need efficiently and effectively.

To assess the effectiveness of countries' OGD initiatives in adhering to best practices in publishing government datasets, various indices have been developed by governmental organizations and research institutions. These indices utilize a range of indicators, such as data accessibility, quality, and transparency, to provide a snapshot of how closely countries are following established governmental Open Data practices. Such comparative tools are crucial for benchmarking the open data achievements of different nations, offering insights into global trends and facilitating the sharing of best practices in OGD implementation [50]. The genesis of these indices began in 2013 with the introduction of the Global Open Data Index (<https://okfn.org/en/who-we-are/our-history/global-open-data-index/> (accessed on 1 February 2024)) by the Open Knowledge Foundation and the Open data Barometer (https://opendatabarometer.org/?_year=2017&indicator=ODB (accessed on 1 February 2024)) by the World Wide Web Foundation. Subsequently, more indices emerged, including the OURdata Index (<https://www.oecd.org/gov/digital-government/policy-paper-ourdata-index-2019.htm> (accessed on 1 February 2024)) by the Organisation for Economic Co-operation and Development (OECD), the Open data Inventory (<https://odin.opendatawatch.com/> (accessed on 1 February 2024)) by Open Data Watch, and the Open data maturity report (<https://data.europa.eu/en/publications/open-data-maturity> (accessed on 1 February 2024)) by the European Union, all in 2015, culminating in the most recent Open Government Development Index (OGDI) by the UN Department of Economic and Social Affairs in 2020 [7,51]. While these indices share some commonalities, they also exhibit significant differences in their objectives and dimensions. Additionally,

their geographical coverage varies considerably [50]. The OGDI is particularly notable for its broad scope, having been applied to benchmark 193 countries.

2.2. Linked (Open) Data

The LD paradigm, as conceptualized by Tim Berners-Lee in his Web architecture note [52], encompasses best practices for publishing and interlinking structured data on the Internet [53]. Key principles include the following:

- Use of URIs as Identifiers: Employ Uniform Resource Identifiers (URIs) as names for entities or things.
- HTTP URIs for Accessibility: Utilize HTTP URIs to ensure that others can easily look up and access the identified entities.
- Provide Useful Information on Lookup: When a URI is looked up, furnish relevant and valuable information utilizing standard technologies like RDF (Resource Description Framework) and SPARQL (SPARQL Protocol and RDF Query Language).
- Incorporate Links to Other URIs: Enhance discoverability by including links to additional URIs, thereby enabling users to explore related information.

LD align with the proven success of the web's architecture, extending its principles to the global sharing of structured data. They leverage standard practices for effective data sharing, building upon the web's decentralized and open information space. Several key standards, including RDF [54], OWL [55], SPARQL [56], and IRI (<https://www.rfc-editor.org/rfc/rfc3987> (accessed on 2 February 2024)) contribute to the interoperability and structure of LD. RDF represents information in a machine-readable manner, OWL facilitates the axiomatization of complex relationships, and SPARQL serves as a query language for RDF data; IRIs extend the capabilities of URIs by allowing a broader range of characters, including those from the Universal Character Set, to support internationalization.

Tim Berners-Lee operationalized the LD principles in the five-star LOD model, aiming to improve data quality progressively. The model encourages data publishers to make data not only accessible but also easily reusable and linkable. The five stars represent levels of adherence to specific principles, promoting a more interconnected and valuable web of data.

The adoption of LD principles empowers publishers to represent conflicting information. Entities are seamlessly connected through RDF links, forming a global data graph that spans sources and enables dynamic discovery. Data publishers enjoy flexibility in vocabulary choice, promoting diversity. Linked Data's self-describing nature allows applications to resolve unfamiliar vocabularies by dereferencing URIs. Utilizing HTTP as a standardized data access mechanism and RDF as a standardized data model simplifies data access compared to Web APIs. Standardized mechanisms enhance consistency and ease of interaction, streamlining data access and utilization in the LD paradigm [53].

3. Methodology

Given the introductory assumptions, to answer our question "Linked Open Government Data: Still a viable option for sharing and integrating public data?" we defined the following research questions (RQs):

RQ1: *What is the current state of Linked Open Government Data?*

This RQ sets a broad foundation for our research by seeking to understand the overall current landscape of LOGD. The answer to this question is provided by the combination of two sub-questions.

RQ1.1: *What is the prevalence of RDF and SPARQL endpoint distributions in national OGD portals?*—This sub-question narrows down the focus to specific technical aspects (RDF formats and SPARQL endpoints), which are crucial for understanding the implementation and accessibility of LOGD.

RQ1.2: *What are the relations between OGD and Linked Open Data found in the literature?*—This sub-question aims to explore the relationship between OGD and LOD, based on the researchers' publication practices.

RQ2: *What factors are holding back the spread of LOGD?*

This question addresses potential challenges and barriers in the adoption of LOGD, which is essential for understanding what might be impeding its broader utilization.

RQ3: *What valuable examples of LOGD adoption can be found today?*

This RQ seeks to identify successful case studies or instances where LOGD has been effectively implemented, which can provide insights into best practices and the benefits of LOGD.

After the subsequent Section 4, the paper endeavors to address the three research questions in the following sections, as outlined in Figure 1.

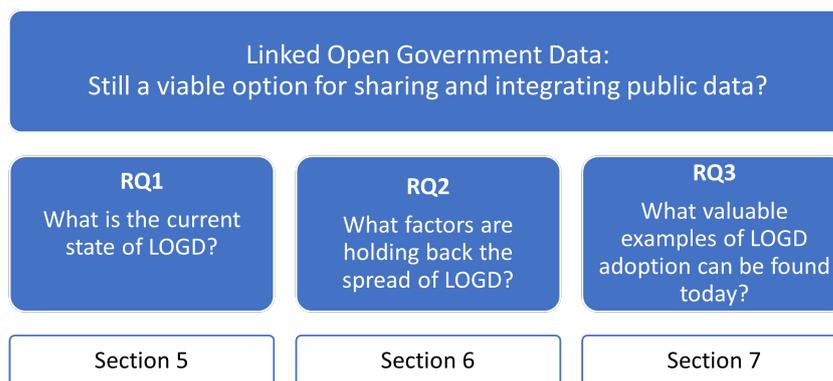


Figure 1. Research questions and paper organization.

4. Related Works

In addressing our first research question, “*What is the current state of Linked Open Government Data (LOGD)?*”, it is important to recognize the concerns raised by several scholars about the suboptimal adoption of LD and LOD dissemination practices in the government sector, often centered around the limited use of RDF distributions [57–60]. Despite these concerns, there is a notable scarcity of studies that provide a quantitative analysis of these issues in specific contexts or scenarios [33,61–63]. Among these, Ibanez et al. [62], in their analysis of a diverse sample of regional and local European institutional websites, observed that RDF is not widely used. They found that “RDF is still a minority format”, not only when compared to CSV formats (constituting less than 5% of the data formats used) but also significantly “less common than non-tabular structured formats like XML and JSON”, being approximately five to six times less prevalent. In 2018, Pawełoszek et al. [61] focused on exploring the potential for creating business models based on open data and also briefly examined four national portals: the US, UK, Germany, and Poland. This study yielded percentages of RDF distributions that were similar to the findings in our own research. Penteado et al. [33], drawing on various studies of national portals from countries like the US, Brazil, Italy, Colombia, and Greece, also highlight the extremely limited proliferation of RDF datasets. In their analysis of LOD challenges and opportunities for data-driven government initiatives in Russia, Aitkin et al. [63], as of May 2020, noted that the Russian Open Data Portal showed minimal growth in open data, with only 23,775 datasets. Over 60% of these data were in CSV format, indicating compliance with the third level of the five-star open data model, but only five had an RDF distribution, highlighting a significant gap in LOD adoption in Russia.

Compared to previous research, our study provides two significant contributions. Firstly, it presents a comprehensive and current overview of national portals worldwide, including a detailed quantitative analysis of RDF distribution publication and the availability of national SPARQL endpoints. Secondly, in addressing RQ1.2, our study concurrently sheds light on the scientific community’s interest concerning the intersection of OGD and LOD.

Concerning the second research question “*What factors are holding back the spread of LOGD?*”, the literature review we carried out shows a clear lack of systematic investiga-

tions into the obstacles that impede the development of LD within OGD. This situation contrasts sharply with the extensive body of research that examines, from various angles, the challenges and barriers faced in the widespread implementation of OGD. One of the earlier studies focusing on OGD issues is by Zuiderwijk et al. [64] in 2012. They specifically explored socio-technical barriers to the use of open data. Their approach combined a literature review, four workshops, and six interviews to categorize these barriers into ten distinct categories. Notably, they found that the impediments identified through empirical research differed from those documented in the existing literature, providing a unique and comprehensive perspective on the challenges faced in the realm of OGD. Attard et al. [65], in 2015, conducted a thorough analysis of the existing literature, focusing specifically on the processes of data publishing and consuming within OGD initiatives. They also aimed to identify key challenges and issues that prevent these initiatives from achieving their full potential. Based on their findings, they categorized the challenges into five distinct groups. We adopted their categorization as a framework to guide and structure our specific investigation of literature in the LOGD field. In the same year, Verma et al. [66] carried out a comprehensive survey across various government agencies in India. By employing statistical methods to analyze the responses to the questionnaires, they were able to identify five key factors that influence the implementation and effectiveness of government initiatives. These factors, as determined through principal component factor analysis, include governance, resource constraints, capacity building, technology, and lack of awareness. These elements were found to be significant in shaping the outcomes and effectiveness of government policies and initiatives within the surveyed agencies. Roa et al.'s 2018 study [67] presented a systematic analysis of OGD literature, identifying six dimensions of data friction. Their categorization of data frictions overlaps significantly with the findings of our study. However, they differentiate between data quality and technical aspects, whereas, in our research, following the framework established by Attard et al. [65], these two dimensions are merged into one.

Building on the foundations laid by the mentioned studies, our survey highlights the opportunity to systematically examine the challenges associated with LOGD. This area, which has not been fully explored in existing research, is where our survey aims to fill the gap. Through a thorough analysis of obstacles in LOGD, we aim to contribute to a more effective approach for the dissemination and implementation of LOGD practices. This is essential for overcoming the current stagnation we have highlighted in the field. Indeed, although many previous studies have addressed barriers in the LOGD space, they often carry this out superficially, identifying a problem only to immediately suggest a technological solution. This approach tends to overlook the complexity and interconnectedness of the challenges. However, a few of the studies we reviewed offer somewhat more elaborate insights into several LOGD barriers. Portisch et al. [68] delve into the intricate task of connecting organizational information from public datasets to established entities in knowledge graphs such as Wikidata and DBpedia. They undertook this by manually establishing links between datasets from the Open Data Portal Watch [24] and their respective publishing organizations in these knowledge graphs. Through this meticulous process, they uncovered a series of interconnected challenges. These ranged from the dynamic nature of organizations, which frequently change, to the complexities arising from the lack of a uniform base ontology that would facilitate standardized link creation. Furthermore, they grappled with the variable quality of metadata, which is crucial for establishing accurate links, and the complexities introduced by multilingual datasets. Another notable challenge was distinguishing between similar public sector organizations, a task requiring precise disambiguation. To navigate these obstacles, the authors not only proposed targeted solutions for each identified challenge but also advocated for a community-driven approach. They suggested a hand-search service that would enable the collaborative efforts of the data science community and dataset publishers in annotating and refining dataset-level links. This collaborative method highlights the importance of human input and collective effort in enhancing the accuracy and reliability of data linking in the public sector. In their

2020 study, Geci et al. [69] investigated the use of LOGD to improve budget transparency in Kosovo. Their research, which included desk reviews and interviews with government and NGO officials, revealed key challenges: poor metadata quality (e.g., temporality, formats, provenance), difficulties in data linking by government employees, and a lack of educational programs for effective data management. The main conclusion drawn from the interview responses indicates that the application of LOGD in Kosovo remains limited. Despite this, there is a general belief in the potential for implementing LOGD in managing Kosovo's budgetary data. This highlights the need for focused efforts to increase public understanding and engagement with open data initiatives. Additionally, there is a crucial requirement for targeted training of staff.

As highlighted in Section 6.2, several barriers identified in the context of LOGD are also found in general discussions about OGD, such as those mentioned above, albeit with varying nuances and interpretations. However, our review of the literature on the nature of LOGD issues shows that the implementation of LOGD presents a unique set of complex challenges that significantly differ from general data practices [70]. Collectively, these challenges underscore the specialized nature and complexity of LOGD, hinting at why its deployment faces significant obstacles.

5. What Is the Current State of Linked Open Government Data?

To address research question RQ1, we implemented two complementary methods of investigation. Initially, we surveyed the portals of the countries best ranked by OGD. The objective of this survey was to gauge the extent of the diffusion of LOD technology within OGD initiatives. Specifically, we examined how many of these portals publish their datasets in RDF format, including the proportion of RDF datasets relative to their total datasets. Additionally, we assessed the availability of SPARQL endpoints in these portals.

To enhance our understanding and provide a more nuanced view, we complemented this empirical analysis with a systematic literature review. This review focused on exploring the scientific community's interest in the application of LD technology within the realm of OGD. This dual approach enabled us not only to assess the current state of LOD technology in OGD initiatives but also to comprehend the academic perspective and interest in this technology.

5.1. What Is the Prevalence of RDF and SPARQL Endpoint Distributions in National OGD Portals?

To verify the prevalence of LOD practices, we examined the presence of two key indicators, i.e., the number of datasets with RDF distribution(s) and the availability of SPARQL endpoints alongside the national portals of a large number of countries, selected according to the OGD index adopted by the UN [7]. The decision to concentrate on national portals, as opposed to other governmental portals such as geo-portals or statistical portals, as well as regional or transnational platforms, stems from the pivotal role these national portals play. As highlighted by the United Nations report [7], users prioritize identifying the "official" national government site among the multitude of potentially available government sites, recognizing it as the gateway or starting point for national users. Bulazel et al. [49] emphasize that a government agency's official web presence serves as the authoritative source of information about its activities, distinguishing itself from unofficial sources like Wikipedia, the news media, or social networks, while [7] notes that national portals tend to be more advanced than those operating at the local level. According to Cyganiak et al. [43], national portals function as central one-stop platforms, offering interested public access to data published by government bodies. These portals play a crucial role in providing "visibility to the process of translating policy into reality".

Considering the extensive coverage of the OGD (Open Government Data Index), which assesses 193 countries, we utilized its ranking system that categorizes countries into four tiers. For our analysis of the adoption of LOD practices, we focused on the 78 countries rated as "Very High" and "High" by the OGD [7] (see Table A1 in Appendix A). These

countries' portals were selected as they are expected to be the trendsetters, being the highest scored by the index. For the selected portals, we checked the presence of RDF distributions and the presence of SPARQL endpoints. The selection of these indicators is directly influenced by the third principle of LD, as proposed by Tim Berners-Lee in 2006, which states: "When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)" (<https://www.w3.org/DesignIssues/LinkedData.html> (accessed on 1 February 2024)). Berners-Lee outlined this principle as part of a broader framework to guide the effective use of the web for linking data, also slightly revising it, as the fourth star in the five-star scheme for LOD: "use open standards from W3C (RDF and SPARQL) to identify things so that people can point at your stuff". Given this context, it is reasonable to infer that the presence or absence of these two parameters (RDF and SPARQL) in a national OD portal is indicative of the degree to which the portal adheres to LOD principles. This adherence is crucial for ensuring that, when users access a URI (Uniform Resource Identifier), they are met with information that is both useful and standardized, thereby fulfilling a key aspect of LOD. The use of RDF distributions as a gauge to measure the prevalence of LOD practices is a concept endorsed by multiple researchers. These authors have highlighted, in both recent and occasionally unsystematic studies, the rather limited adoption of these practices [33,57–59,61–63]. Kumar et al. [58] underscore that "RDFs are central to Linked Data and LOD", encapsulating their crucial role. Meanwhile, Penteadó et al. [33] emphasize the significance of RDF as a benchmark: "Although RDF is not the exclusive format for serializing linked data, its widespread recognition as the most popular format makes it a valuable proxy for gauging the implementation of LOGD".

The systematic analysis of the 78 countries that ranked highest in the GODI, conducted between June and September 2023, serves both as an update and a complement to the results of previously cited studies. This analysis confirms the continued very low adoption of RDF in the distribution of OGD datasets. Out of these 78 portals, only 26 feature at least one RDF distribution and only 21 of them have more than one. As illustrated in Figure 2, only 10 of these 26 portals have more than a marginal one percent of RDF distributions. However, the overall percentage of RDF usage remains low, with only four portals exceeding a 4 percent usage rate. Among the portals analyzed, the Italian portal notably stands out with a modest yet comparatively higher adoption rate of RDF, at 6.5%.

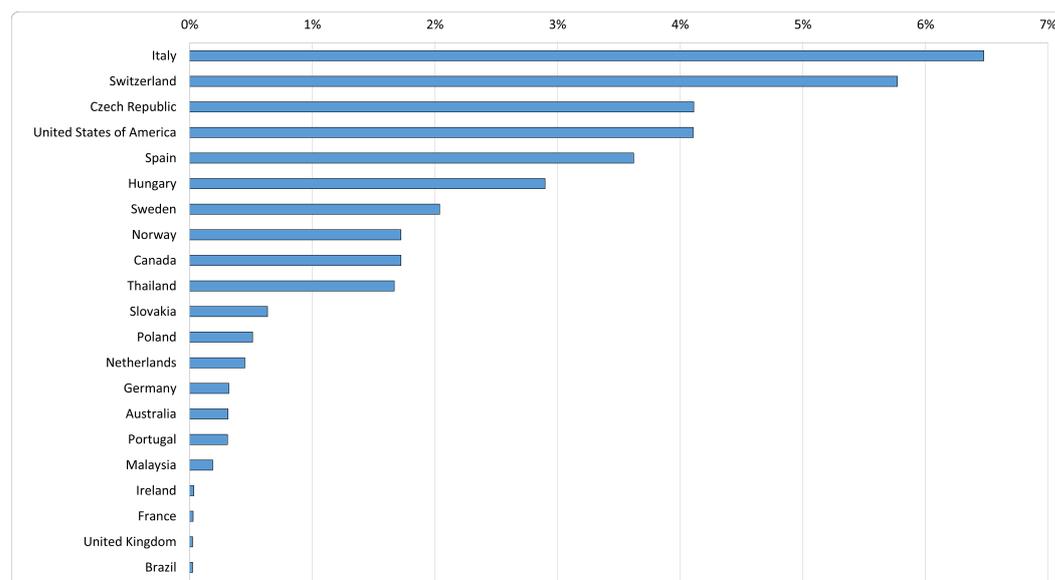


Figure 2. Percentage of datasets with an RDF distribution per country.

Additionally, a closer examination in absolute terms further highlights the limited extent of RDF adoption in OGD. Among the 21 portals that feature multiple RDF distributions, they collectively account for just 19,255 out of an extensive total of 801,765 datasets,

which equates to a mere 2.4 percent. This figure, however, is significantly skewed by the contributions of a few countries. Notably, the United States, Italy, and Spain have a disproportionate impact on this percentage, with their respective counts of 10,297 RDF datasets out of 250,717, 3854 out of 59,516, and 2530 out of 69,879. In contrast, countries like Australia, Germany, the UK, and France display a markedly lower number of RDF representations, despite their substantial volumes of published datasets. For instance, Australia and Germany, with 105,647 and 82,845 datasets, respectively, maintain only 331 and 265 RDF distributions each. The situation is even more pronounced in the UK and France, where, despite having 51,502 and 44,486 datasets, respectively, each country has a surprisingly low number of 13 RDF distributions. This discrepancy highlights the uneven adoption of RDF across different national portals, and underscores the need for a broader and more balanced advancement in the use of RDF in OGD initiatives. Interestingly, the distribution between the “Very High” and “High” categories by the GODI is evenly split, with each category comprising 39 countries. However, a closer examination of the 21 countries that maintain more than one RDF dataset, as shown in Figure 2 and in Table A1, reveals that 16 of them, accounting for over three-quarters, are rated as “Very High” by GODI. This trend suggests that a strong commitment to good OGD practices may be somewhat conducive to the adoption of LOD practices. However, considering the overall limited adoption of LOD, as evidenced by our analysis, this potential facilitation appears modest.

The examination of the second key indicator, the presence of national SPARQL endpoints, reveals an even more challenging scenario in the realm of LOD practices. Among the 78 countries evaluated, a mere five—Italy, the Czech Republic, Spain, Germany, and Croatia—have established a SPARQL endpoint. Notably, Italy, Spain, and the Czech Republic are also among those countries that demonstrated a higher presence of RDF distributions.

This limited deployment of SPARQL endpoints, essential for querying RDF data, further underscores the global disparity in the adoption of advanced LOD technologies within government data initiatives. The scarcity of such endpoints reflects a significant hurdle in realizing the full potential of LOD, especially in facilitating efficient data interoperability and access across diverse government platforms. This result resonates with the literature, which on several occasions reports that most of the SPARQL endpoints were unavailable or almost permanently down [47,71]. Meanwhile, Mouzakitis et al. [70] notes, “Unfortunately, in recent years, a significant number of data providers have ceased supporting and maintaining public SPARQL endpoints, thereby damaging the trust between consumers and open Linked Data providers”.

5.2. What Are the Relations between OGD and LOGD Found in the Literature?

To address Research Question 1.2, we conducted a comprehensive exploration of Open Government Data-related topics by querying the Scopus and Web of Science digital libraries. Keywords such as “open government data” and “linked data” were employed in the search (see Table 1). The search, limited to English-language papers spanning the period from 2010 to June 2023, aimed to discern trends and select recent studies for analysis. Notably, the initial date retrieved by the digital library search engine aligns with Tim Berners-Lee’s promotion of the “Linked Open Data 5 Star” concept, which specifically targets “especially government data owners”. To determine the extent to which LOD research is integrated with OGD research, we used two search queries, namely Q1 and Q2 (see Table 1). The first included only “open government data” and the second combined “open government data” with terms related to LD. Both libraries were searched using title content, abstract content, and keywords.

In total, the first search returned 1290 papers, while the second returned 193 papers. These figures and Figure 3 highlight two facts. Approximately one-sixth (15%) of the research studies on the OGD phenomenon also cover LOD technology. Additionally, although the number of articles on OGD experienced growth in the first decade followed by stabilization in the last three years (probably due to the COVID-19 pandemic, as suggested by Wirtz et al. [48]), the studies that simultaneously address LOD, except for the

initial four-year period (2010–2013), have remained relatively constant throughout the considered timeframe.

Table 1. The results of the analysis of digital libraries.

Query	Sources	Search Text	All
Q1	Scopus	TITLE-ABS-KEY ("open government data")	1236
	WoS	(TI = ("open government data") OR AB = ("open government data") OR AK = ("open government data"))	790
	Scopus ∪ WoS		1290
Q2	Scopus	TITLE-ABS-KEY ("open government data") AND TITLE-ABS-KEY ("linked data" OR "linked open data" OR "LOD" OR "linked open government data")	180
	WoS	(TI = ("open government data") OR AB = ("open government data") OR AK = ("open government data")) AND (TI = ("linked data" OR "linked open data" OR "LOD" OR "linked open government data") OR AB = ("linked data" OR "linked open data" OR "LOD" OR "linked open government data") OR AK = ("linked data" OR "linked open data" OR "LOD" OR "linked open government data"))	105
	Scopus ∪ WoS		193
Q3	Scopus	TITLE-ABS-KEY ("open government data") AND TITLE-ABS-KEY ("linked data" OR "linked open data" OR "LOD" OR "linked open government data") AND TITLE-ABS-KEY (challenge* OR issue* OR obstacle* OR barrier* OR hinder*)	52
	WoS	(TI = ("open government data") OR AB = ("open government data") OR AK = ("open government data")) AND (TI = ("linked data" OR "linked open data" OR "LOD" OR "linked open government data") OR AB = ("linked data" OR "linked open data" OR "LOD" OR "linked open government data") OR AK = ("linked data" OR "linked open data" OR "LOD" OR "linked open government data")) AND (TI = (challenge* OR issue* OR obstacle* OR barrier* OR hinder*) OR AB = (challenge* OR issue* OR obstacle* OR barrier* OR hinder*) OR AK = (challenge* OR issue* OR obstacle* OR barrier* OR hinder*))	35
	Scopus ∪ WoS		58

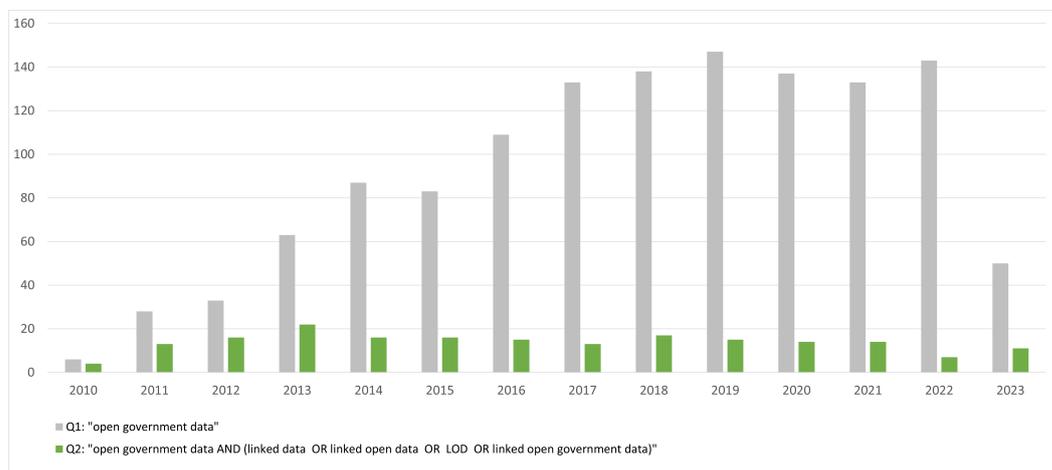


Figure 3. RQ1.2: Relations between OGD and LOGD. Results of the bibliographic search between 2010 and June 2023.

Based on these figures and trends, one could reasonably conclude that research on OGD has gradually shifted its focus away from LOD. Conversely, it may suggest that researchers consider LOD less compelling for the advancement of OGD than it appeared to be in the early years of the past decade. On the other hand, a less unfavorable interpretation could be considered, suggesting that the consolidation of LOD within the OGD research field, having reached a plateau in the number of studies after the initial four years, has also attracted researchers’ curiosity to a lesser extent.

6. What Factors Are Holding back the Spread of LOGD?

Based on the analysis of statistical and bibliographical results regarding the first research question, it appears plausible to assume that the momentum of LOGD has indeed decelerated, if not come to a halt, compared to the initial enthusiasm. To comprehend the reasons behind this slowdown, our focus shifts to addressing the second research question, RQ2: “What factors are holding back the spread of LOGD?”

6.1. Bibliographic Analysis

To address RQ2, we extended the earlier literature search conducted for RQ1 by introducing a new search query, namely Q3 (see Table 1). This query specifically seeks papers on LOGD that discuss various issues related to their dissemination and implementation. In total, the search returned 58 documents across Scopus and WoS, as shown in Figure 4, which also shows the results of the Q2 query for comparison. These figures indicate that slightly less than a third (30%) of the publications related to LOGD potentially touch on issues critical to its dissemination. However, a systematic survey of these papers is required to comprehensively assess the depth of analysis regarding these challenges.

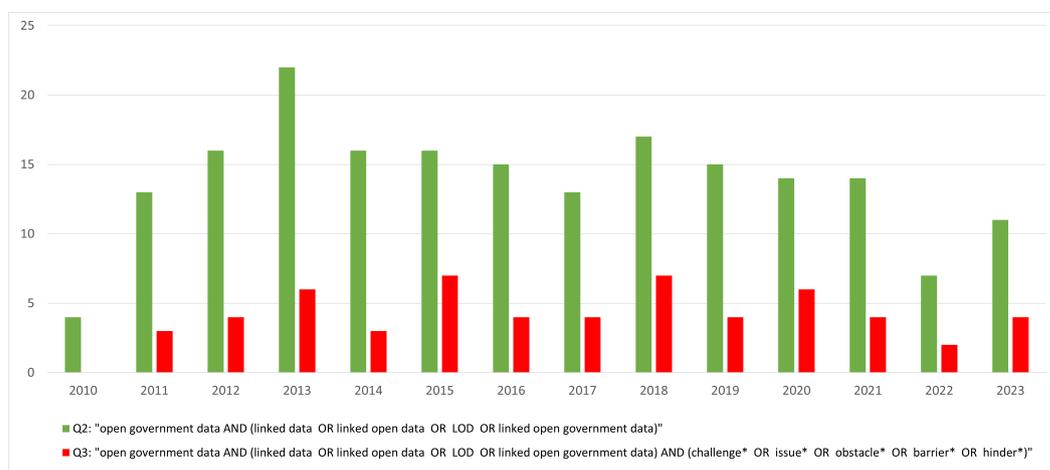


Figure 4. RQ2: LOGD data frictions. Results of the bibliographic search between 2010 and June 2023.

After an initial screening of the abstracts and titles of the 58 papers, conducted independently by the two authors, it became evident that only a very small proportion of the papers (16) focused, albeit with different and often marginal emphasis, on LOGD issues. For instance, several articles identified through searches based on abstracts, titles, and keywords as potential discussion papers on issues in LOGD are actually proposals where LD is presented as a solution to problems associated with OGD. These problems include, for example, the fragmentation of datasets and issues related to their quality [38,72]. However, this number is insufficient for a sufficiently detailed analysis of the phenomenon under consideration. Therefore, all 193 papers resulting from the Q2 query were included in the second review phase. This phase encompassed all studies on LOGD, along with a few additional papers we considered relevant though not included in the Q2 results. These papers were individually reviewed by the two authors, who assigned ratings (ranging from 0 to 10) and added semantic tags indicating the type of issues addressed in each paper. Following this analysis, 40 papers emerged after further parsing. These papers were then organized based on the assigned tags, from which key concepts related to LOGD issues were extracted.

6.2. LOGD Data Friction

Extending the concept of OGD data friction, borrowed from Edwards [19], we can speak of LOGD data friction, to define all the phenomena, both technical and social, that hinder the diffusion of LOGD. Using a bottom-up approach, the analysis of the

literature allows us to systematize these phenomena following the framework proposed by Attard et al. in [65,73]. This framework identifies five dimensions that impact the creation of value through OGD, which we believe can also apply to LOGD, albeit with challenges specific to the LD paradigm, namely Technical, Organizational, Policy/Legal, Social, and Economic. We classified the literature findings of Section 6.1 according to these five dimensions, as shown in Table 2, also highlighting the main issues arising for each one.

Table 2. Data friction dimensions that impact on the creation of Linked Open Government Data.

Dimension	Description	Issues
Technical	deals with challenges like diverse data sources, vocabulary alignment for semantic interoperability, managing LD complexity, and ensuring data quality for LOGD production	<i>Heterogeneity</i> [2,33,41,49,61,68,69,74–79] <i>Vocabularies</i> [2,33,35,41,61,62,76,80–84] <i>LD lifecycle complexity</i> [33,35,41,62,74,79,80,85,86] <i>Data quality</i> [33,62,68,69,77–79,87,88]
Organizational	emphasizes strategic data management in government structures, directing attention and enhancing the skills of public servants, fostering a culture that understands and effectively manages LOGD, aligning with broader public sector objectives	<i>Cultural change</i> [32,33,68,70,80,89,90] <i>Lack of technical expertise</i> [33,35,41,62,69,70,74,80,83,90–93]
Policy/Legal	deals with crucial aspects of data management and usage, requiring a thorough understanding of the legal and regulatory landscape	<i>Data ownership and licensing</i> [73,94,95] <i>Legal regulation</i> [69,78,89] <i>Data privacy and security</i> [73,78]
Social	discusses challenges related to stakeholder awareness and motivation, as well as public misconceptions about open data	<i>User engagement</i> [9,33,64,68,69,89,93,96–98]
Economic/Financial	underlines the importance of financial commitment for valuable data release	<i>Sustainability</i> [2,43,97] <i>Maintenance</i> [9,33,88]

6.2.1. Technical Dimension

The technical dimension encompasses several critical facets that collectively shape the interconnected data landscape. *Heterogeneity*, a key concern in this dimension, arises from the diverse nature of data sources, structures, and formats, necessitating strategies for improved integration. *Vocabularies* play a crucial role in achieving semantic interoperability, with efforts focused on aligning existing vocabularies and mitigating heterogeneity. *LD lifecycle complexity* highlights the complexity of establishing meaningful relationships between entities, emphasizing the need to address data linkage issues for the reliability of LOGD. Finally, *Data quality*, characterized by accuracy, completeness, and consistency, is a fundamental aspect of the production and release of valid LOGD.

Heterogeneity

Heterogeneity is widespread in the realm of Open Government, exerting a significant influence on both the publication and consumption of OGD. The data are disseminated by disparate providers, often lacking coordination, resulting in a diversity of APIs and formats (referred to as *data source heterogeneity*). Furthermore, publishers exhibit heterogeneity in their use of vocabulary, combining and reusing distinct data models and terminologies (referred to as *vocabulary heterogeneity*). Additionally, they create and employ their own institutionally minted and managed identifiers, contributing to a further layer of heterogeneity in the form of unique identifiers (*identifier heterogeneity*) [2,41,74].

Adopting LD practices eases the data source heterogeneity and vocabulary heterogeneity. As elucidated by Heath and Bizer [53], adhering to LD principles establishes a standard approach for accessing information through the dereferencing of HTTP URIs into RDF descriptions. Accessing such data is possible through various channels, including querying SPARQL endpoints or obtaining data in the form of RDF dumps. The utilization of RDF and HTTP dereferenceable Internationalized Resource Identifiers (IRIs) establishes a consistent method for expressing data references, facilitating the exploration of additional data sources through the traversal of RDF links, and ensures a seamless integration with the vocabularies that organize such data. Utilizing self-explanatory HTTP-dereferenceable IRIs for vocabulary terms fosters the integration and combination of domain-specific and cross-cutting third-party vocabularies.

However, the *alignment of vocabularies* remains imperative to deal with vocabulary heterogeneity when aggregating and independently provided data. As highlighted by Theocharis [75], the diversity of terms employed across public sector bodies has already led to confusion during data searches and interconnections. Distinct vocabularies depict data at varying levels of detail, covering overlapping yet not identical facets of the domain, bringing to a *granularity mismatch*. This affects the alignment and mapping among both controlled vocabularies and data models. *Multilinguality* further exacerbates vocabulary heterogeneity, especially concerning controlled vocabularies. The demand for automated cross-language semantic matching methods has been underscored by Narducci, emphasizing their role in supporting local and national administrations in linking their service catalogs [79]. Consequently, achieving vocabulary alignment can be a challenging task, as noted by Penteado et al. [33], Bulazel et al. [49], and Vert et al. [77].

Another significant hurdle in handling OGD arises from the need to effectively manage, connect, and comprehend the relationships between references to entities across various datasets [77]. *Entity consolidation and interlinking* are crucial but complex activities required to address identifier heterogeneity, enhancing the coherence and accessibility of information within diverse datasets. In particular, Portisch et al. [68] address the organization's consolidation as a pivotal case of consolidation in the context of LOGD, also highlighting multilingual issues. After the interlinks are established, they may not align with the expectations implied by the joint utilization of the interlinked datasets. The *quality of the links* affects the usability of the aggregated data obtainable by exploiting the interlinking [69] and, as it has also been noticed in a more general application of LD technologies, it might seriously affect the suitability of aggregated data for specific downstream applications [99,100].

Vocabularies

Several hurdles are related to the development, reuse, and maintenance of vocabularies [101,102]. OGD covers a wide set of specialized domains and requires specialized and cross-cutting vocabularies. The variety of covered domains implies that "*No One Vocabulary Fits All*": a single vocabulary cannot accommodate all datasets [81]. Consequently, combining different ontologies and vocabularies becomes necessary for comprehensive data representation [76]. When well-established shared vocabularies are unavailable, *creating new vocabularies is a challenge* [80]. The creation of ontologies is a time-consuming process that demands technical and domain expertise [83]. The *absence of widely adopted shared vocabularies* further complicates data interoperability [82]. Establishing standardization in common metadata and thesauri is imperative for seamless interoperability [2,84]. In addressing these challenges, there is a recognized need for initiatives that focus on developing standardized models, as outlined in [33]. This signifies the importance of collaborative efforts and standardized approaches to overcome the obstacles associated with transforming, selecting, and creating vocabularies in the context of data interoperability.

Even in cases where vocabularies for a specific dataset are available, challenges arise in *vocabulary reuse*. Reusing third-party vocabularies adds an additional layer of complexity to the transformation process, as emphasized by Akatkin et al. [80]. Selecting appropriate vocabularies to represent diverse datasets accurately is a non-trivial endeavor [41,61]. Despite the existence of some well-established vocabularies, their limited reuse is a notable issue [62]. On the other hand, distributing the effort among publishers requires clear guidelines for vocabulary choice and use, leaving room for interpretation [62].

LD Lifecycle Complexity

The development and dissemination of LD involve a multifaceted and time-consuming lifecycle [103], demanding concerted efforts in both creating and integrating semantic assets [74,80]. The ability to *construct SPARQL queries* becomes paramount when deploying and interlinking datasets [33]. RDF itself poses challenges, requiring substantial effort to capture knowledge from data owners and domain specialists, and transform the data into a suitable RDF format based on appropriate vocabularies [35].

Transforming datasets from disparate formats, such as CSV into RDF, brings its own set of hurdles, involving non-machine-readable sources and ambiguous model semantics [86]. The process of transforming data into shared vocabularies is acknowledged as a challenging task [35]. The quest for *discovering links* between services described in different languages, elucidated by [79], underscores the significant human effort required due to the multitude of service descriptions and linguistic and cultural barriers. To overcome the difficulties in converting existing OGD datasets into RDF, some authors argued for the necessity of automatic triplification tools [41] and provide solutions that free publishers from the intricacies of LD technology [85], although Ibanez et al. [62] raise concerns about potential quality issues, such as vocabulary dereferenceability. To cope with these issues, *centralizing efforts* at the portal or meta-portal level, as observed in initiatives like EDP, can potentially create economies of scale and ensure more homogeneous results [62].

Data Quality

Government datasets managed by different departments and agencies present significant challenges in ensuring data quality, consistency, and standardization, which can hinder the adoption of LOGD [68]. Several key issues highlight these challenges. As noted by Ibanez et al. [62], when assessing the uptake and quality of linked datasets belonging to several European portals indexed by the European Data Portal, there is a problem of *vocabulary dereferenceability*: the datasets analyzed referred to more than 3000 proprietary vocabularies but over than 95% of them were non-dereferenceable, undermining efforts to establish standardized and consistent terminologies. Wieczorkowski [78] emphasizes the *tension between volume and quality* in big OGD, pointing out that it is difficult, also economically, to guarantee satisfactory quality when the main priority is to open up large amounts of (linked) data, even at the risk of publishing incorrect, obsolete, and non-integrated data. Vert et al. [77] observe that the open world perspective underlying LD, according to which anyone can say anything about anything, leads to challenges such as *inconsistencies in data* and *contrasting views* arising from different sources about the same entity, raising concerns about timeliness, completeness, and accuracy. Apart from the already mentioned problems with the quality of links, Penteadó et al. [33] note that *lack of linking* in published datasets compromises overall data quality and has a global impact on the reproducibility of LOGD processes. Furthermore, despite the authority of data providers, datasets are often not uniformly accurate or consistent, as noted in [87], where inconsistencies between source systems, often due to *different domain foci*, increase the complexity of data integration. Acknowledging the significance of enhancing data availability, quality, and understandability, Scholl et al. [69] underscore these factors as essential prerequisites for the successful implementation of LOGD. In the same vein, Deng et al. [88] highlight that the inconsistency of content in OGD poses a significant challenge to LD integration.

6.2.2. Organizational Dimension

The Organizational Dimension focuses on the strategic aspects of data management within the structures of government bodies. It underscores the importance of developing integrated and efficient workflows that are capable of navigating and overcoming the bureaucratic and cultural hurdles often encountered in government settings. This dimension also emphasizes the need for directing the attention and enhancing the skills of public servants specifically towards LOGD. The goal is to foster an organizational culture that not only understands the value and potential of LOGD but is also equipped to effectively manage, share, and utilize these data resources in a way that aligns with the broader objectives of the public sector.

Cultural Change

Government organizations often face *challenges in adapting* to new data-sharing practices and technologies, primarily due to their inherent bureaucratic processes and strict regulations. Moreover, as highlighted by Shadbolt et al. [32], semantic web technologies

and protocols pose challenges for government agencies with intricate data requirements and constrained resources. Governments prioritize expanding the volume of open data in various forms rather than solely maximizing LD. As Akatkin et al. have highlighted [80], there is a crucial need for an integrated approach that combines the use of semantic methods with expert collaboration. This approach aims not only to enhance legal regulation but also to facilitate knowledge building, starting from existing processes and practices. Additionally, ref. [89] discuss a significant issue: the *lack of awareness* in identifying valuable datasets for reuse. This common challenge impedes the full realization of open data's potential. Public officials need to be adept at recognizing which datasets can be valuable for use and reuse, and which should not be allocated further resources. This discernment requires an in-depth understanding of the possible applications of various datasets and the specific needs of potential users interested in accessing and reusing this data. They also advocate for public institutions to adopt a mindset of being data reusers right from the beginning of the data publication process. In the same vein, Portisch et al. [68] identify an *organizational shortcoming* in the realm of data publishing, where data publishers frequently lack essential contextual information. This deficiency hinders their ability to accurately determine which other entities their data should be linked to. *Operational and process barriers* are the focus of other authors. According to Mouzakitis et al., public sector organizations encounter numerous hurdles in incorporating LD technologies into their routine operations. The bureaucratic nature and intricate legislation typical of public services often result in these bodies adopting new technologies like LD more slowly than private entities. Training consultants for public servants must comprehend these difficulties and their impact on the public sphere, and develop solutions to address them [70]. Penteado et al. [33] point out that most literature on publishing LOGD focuses on "less complex, non-operational datasets". There is a pressing need for an engineering perspective, along with the identification of practical challenges and considerations of "organizational limitations". Similarly, Tamburis et al. stress the need to thoroughly examine and understand the current processes used in statistical offices and other data-publishing entities. They also highlight the importance of establishing and implementing relevant policies [90].

Lack of Technical Expertise

Implementing LD principles demands technical expertise in areas such as RDF modeling, ontology development, and SPARQL query writing [70], posing a challenge for government agencies lacking the necessary skills and resources for effective adoption. Klein [35] emphasizes the significant effort required to capture knowledge and transform data into the RDF format, underscoring the need for technical skill. The need for support in adopting LOGD is multifaceted. Tambouris [90] advocates for *supporting guidance specific to LOGD*, considering the entire process and the audience, including complex examples and non-functional requirements. Abida et al. [91] highlight the clear *lack of automatically supported, integrated solutions* for the end-to-end process of LOD production and publishing in an e-Government context, necessitating user-friendly tools. Trinh et al. [93] echo the sentiment, emphasizing the challenge of exploiting LD due to the dataset knowledge and technical expertise required. Additional tools such as *user-friendly and visually oriented tools* [35], templates for open government portals [92], and approaches like CVSIntoRDF [62] are suggested to alleviate the technical burden. Managing large volumes of RDF data is identified as challenging by Kaoudi et al. [74], while Akatkin et al. [80] stress the complexity and multidimensionality of developing and disseminating LD. The creation of domain-specific ontologies is deemed time-consuming and to require expert personnel by Hocht et al. [83]. Penteado et al. [33] note that, despite guidelines for publishing LD, many producers lack sufficient knowledge of these practices. Scholl et al. [69] connect the expected results in LOGD initiatives to the lack of ability to link available data.

6.2.3. Policy/Legal Dimension

The Policy/Legal Dimension, addressing critical issues surrounding data management and usage, demands a comprehensive and articulate understanding of the legal and regulatory landscape related to data. This dimension encompasses several key areas, concerning Data Ownership and Licensing, Legal Regulation, and Data Privacy and Security.

Data Ownership and Licensing

Navigating data ownership and licensing agreements in the context of LOGD is complex and fraught with challenges. Different government agencies have varying policies, making *the use and reuse of LOGD under these diverse policies* a difficult task. Attard et al. [73] highlight a key issue: *the lack of standardized and pluggable license agreements*. This results in a plethora of open licenses that, despite their open nature, often contain conflicting terms. Such discrepancies make it difficult to merge data from different sources, as the licenses may be incompatible. Furthermore, sharing data across various public entities can lead to unclear dataset ownership. This ambiguity around the rightful owner of the data can lead to copyright inconsistencies, impeding the publication of the data. Zhu et al. [94] underscore the importance of license compatibility or interoperability, especially for LOGD applications. They observe that many portals apply licenses to some datasets but not all, creating a patchwork of usage terms. Additionally, they note the problematic practice of using legal notices instead of open licenses. These notices, often found in the U.S.'s federal OGD portal, lack standardization and are not machine-readable, posing challenges for LOGD applications. Morando [95] also points out a practical issue: the complexity of understanding what can legally be carried out with mixed datasets. He argues that it is essential for reusers to have a clear understanding of their rights and limitations regarding data usage, ideally "without asking their lawyers" or navigating through numerous licenses. This clarity is crucial for the effective and lawful reuse of LOGD.

Legal Regulation

The legal and regulatory frameworks that govern data sharing, intellectual property rights, and access to public information are intricate and vary across different jurisdictions. Wieczorkowski [78] criticized *the overly general nature of data reuse laws*, observing that they are often vague and do not prescribe the use of proven and effective methods for data publication. He highlights the crucial *role of the state in eliminating data-related and legal obstacles*, which is essential for harnessing the economic potential of OGD applications. In line with the European Union's report [89], it is important to ensure that open data publication efforts are in sync with initiatives aimed at enhancing the quality of policymaking, particularly with emerging "predictive" policy tools. A notable example is the American Foundations for Evidence-Based Policymaking Act of 2018. This Act mandates all agencies to clearly define their data requirements with planned regulatory activities. This legislation has led to a structured approach to data needs assessments by public agencies, facilitating a more systematic connection with data publication initiatives. Similarly, measures such as the EU open data directive are instrumental in promoting the reuse of public sector information. While stressing the importance of training for both officials and civil stakeholders in the field of LOGD, Geci et al. [69] argue that the effectiveness of such training and the overall success of LOGD practices are strictly dependent on a proactive legal framework based on a national strategy.

Data Privacy and Security

Government datasets often include sensitive or personally identifiable information, which poses a significant challenge in the context of open data initiatives. As Attard et al. [73] have pointed out, data providers are tasked with finding a delicate balance. They must make data as freely available as possible while also respecting individuals' right to privacy. This balance is crucial to *maintain public trust and ensure legal compliance*, especially when dealing with datasets that contain personal information. Wieczorkowski [78] brings attention

to additional complexities in the collection and processing of private and, in particular, personal and sensitive data. Under the current EU General Data Protection Regulation (<https://gdpr-info.eu/> (accessed on 1 February 2024)) (GDPR), any information relating to an identified or identifiable natural person is subject to protection. This regulation imposes stringent technological and organizational requirements for data processing, including the need for pseudonymization and restrictions on profiling individuals. These requirements, although designed to protect personal data, can potentially discourage or complicate the reuse of such data in LOGD projects.

6.2.4. Social Dimension

LOGD initiatives face challenges in *effectively engaging users*, necessitating heightened awareness, education, and support for developers, researchers, and the public to drive adoption and ensure meaningful data utilization. This challenge is underscored by several factors, including a lack of a centralized body responsible for open data portals and user ecosystems, as highlighted in the European Union's report on creating LOGD initiatives [89]. To address this, strategies such as hackathons and contests have been employed to boost the usage of OGD by civil society [96]. The significance of stakeholder engagement and co-creation in enhancing data usefulness is emphasized, with examples like Fusepool incorporating predictive labeling and user annotations [9]. The identification of data needs is a crucial aspect, and involving target users is essential to enhance the social and commercial value of OGD [97]. Recent initiatives and directives from the European Commission on Open Data, specifically focusing on "High Value Datasets" (HVDs) [104,105], which are defined as datasets "re-use of which is associated with important benefits for society, the environment and the economy", are aligning with this perspective [106]. Combining HVDs and LD has the potential to amplify the impact of their adoption. Unfortunately, as highlighted by Zuiderwijk et al., integration practices for datasets face challenges due to limitations in cross-domain and interdisciplinary knowledge and skills [64]. *Trustworthiness* of LD is a recurring concern, with challenges related to the trust of the data itself and trade-offs between authoritative content and ease of use [98]. Moreover, the complex nature of LD requires applications on top of the layer to cater to users' different information needs, as highlighted by Trinh et al. [93]. Portisch et al. [68] recognized the importance of community engagement in establishing LOGD, and specific education programs, including training for both public servants and civil stakeholders, improved technological capabilities, and tools, are identified as a need for successful LOGD initiatives by Geci et al. [69]. However, as recently noted by Penteadó et al. [33] *the lack of tools and formalized practices for community engagement* is still evident.

6.2.5. Economic/Financial Dimension

In the economic/financial domain, emphasizing *financial commitment* is crucial for the successful release of valuable data. The absence of a clear funding and governance model can render initiatives unsustainable. Coupled with a lack of ongoing commitment to maintaining LOGDs, this can lead to outdated or inaccessible data. Together, these factors play a key role in ensuring the successful implementation and long-term sustainability of LOGD projects.

According to Ding et al. [2], one of the major challenges for OGD lies in "the costly integration of government data across domains and political boundaries". This is attributed to datasets being published in various formats, utilizing different vocabularies, and accompanied by varying-quality metadata. Cyganiak et al. [43] propose a "self-service approach" to address the cost problem of LOGD, shifting the burden of data conversion towards the data consumer. However, realizing this potential comes with its own set of costs, as observed in the pioneering efforts in the US and UK, where creating high-quality LD requires considerable investment. Alexopoulos et al. [97] acknowledge the potential barrier to improving OGD provision in Greece due to the *reduced budgets of government agencies* amid an economic crisis. Despite the acknowledged benefits of opening government data,

agencies might prioritize other ICT projects deemed to be of higher visibility and priority. The economic constraints pose a challenge to realizing the full potential of OGD.

Maintenance also emerges as a significant concern. Penteado et al. [33] highlight maintenance tasks such as updating information in the graph, preserving links, and checking service availability. The critical nature of the maintenance phase is underscored, requiring further development to ensure the validity of produced data in a decentralized context. Kaschesky [9] emphasizes the importance of data stewardship and curation for ensuring the effectiveness of data. Addressing the governance aspect, Deng et al. [88] emphasize the challenges of managing large and diverse datasets from government agencies, requiring significant time and effort. The difficulty in using human effort for governing such datasets, especially when they are distributed across multiple agencies, poses a substantial hurdle.

7. What Valuable Examples of LOGD Adoption Can Be Found Today?

To answer RQ3, the collected literature was screened to identify remarkable success cases for LOGD. The rationale for collecting success cases was prompting a more congruous view as listing problems per se is pivotal but can provide only a part of a complex story, and success cases complete the picture serving as inspiration or models for others to learn from and emulate. We consider remarkable success cases that refer to an exceptional example of achievement, progress, or positive outcomes. We sought situations or stories that stand out due to their impressive and noteworthy nature. Unfortunately, the cases discussed in the literature focused more on very interesting use cases and proof of concept systems demonstrating the potential behind the LOGD adoption rather than its actual measured impact (e.g., [107,108]).

The discussion of proof of concept systems is quite natural when considering the initial phases of emerging technologies, but we expected to find in the scientific literature some deeply organized discussion of the LOGD take-up and that was not the case. Besides the systematic search for relevant publications, we also used a less formal method to complement the collected resources with reports and institutional websites, mainly through backward and forward searches in the references of the papers we identified in the first phase.

Through a comprehensive examination of resources, several adoptions of LD technologies within the realm of open government have come to light. These applications exhibit distinctive features that make them worthy of being regarded as noteworthy cases.

7.1. DCAT and Open Data Catalog Interoperability

DCAT, based on the LD best practice, has an impact on the way data is shared and documented in the context of OGD. DCAT, developed through W3C standardization, facilitates unambiguous and shared data exchange across systems. Initially designed at DERI, it underwent refinement by W3C groups, becoming a Recommendation in 2014. A second version of DCAT [109] was published as a W3C Recommendation in 2020 and the third version is under development as a response to new use cases. DCAT harmonizes diverse community approaches, extending the core for profiles to ensure the semantic uniformity crucial for lossless interoperability.

DCAT leverages LD best practices, adopting a metadata schema based on the Open-World Assumption (OWA) and defined using the RDF data model [54]. The OWA, when applied to the metadata model, signifies that the metadata schema remains open, allowing extension with types and relationships from other schemes. RDF facilitates a machine-actionable approach by assigning each term in the schema a unique identifier, enabling the retrieval of term semantics and the joint use of terms from different vocabularies. These assumptions prove effective in uncoordinated open environments like the Web [110].

In the LOGD context, DCAT's success is evident in widespread adoption, with DCAT-AP as a notable example. Developed as part of Interoperable Europe (<https://joinup.ec.europa.eu/interoperable-europe> (accessed on 3 January 2024)), DCAT-AP [111] serves as a profile for sharing catalog information in Europe, with maintenance by SEMIC (<https://joinup.ec.europa.eu/interoperable-europe>).

[//joinup.ec.europa.eu/collection/semic-support-centre/welcome](https://joinup.ec.europa.eu/collection/semic-support-centre/welcome) (accessed on 6 February 2024)). It has been used across Europe since 2014, primarily for government and scientific data catalogs, demonstrating broad geographic coverage and support in platforms like the European Data Portal (<https://data.europa.eu/> (accessed on 3 January 2024)) and CKAN (<https://ckan.org/>). Over the past decade, DCAT-AP has evolved into a comprehensive ecosystem, featuring interconnected specifications, GeoDCAT-AP [112], StatDCAT-AP [113], and mobilityDCAT-AP (<https://mobilitydcat-ap.github.io/mobilityDCAT-AP/releases/1.0.0/index.html> (accessed on 18 January 2024)), that are domain-specific extensions for geospatial, statistical, and mobility data, and share the DCAT-AP's geographic coverage. Besides the European profile, extra-European profiles such as DCAT-US (<https://resources.data.gov/resources/dcat-us/> (accessed on 18 January 2024)) and national profiles have emerged (e.g., DCAT-AP-IT for Italy, DCAT-AP-NL for The Netherlands, DCAT-AP-NO for Norway, DCAT-AP-SE for Sweden, and DCAT-AP-DK for Denmark).

7.2. Vocabularies

There is evidence of widespread adoption of Linked Data (LD) technologies to reduce data friction around vocabularies, facilitating the emergence of standardized vocabularies, and their alignment and collaborative use. This section presents some notable cases where the use of RDF/OWL and SKOS vocabularies is being concretely reused in practical government settings. In the context of government data, long-lasting efforts have been made to help standardize the vocabularies used to describe concepts [114,115], ensuring that information is organized and retrieved accurately, for example, AGROVOC, initiated in the early 1980s, and GEMET, initiated around 1997. Publishing controlled vocabularies such as SKOS [116,117] following the LD principle has become a common practice (see AGROVOC [118], EU vocabularies (<https://op.europa.eu/en/web/eu-vocabularies/controlled-vocabularies> (accessed on 1 February 2024)) including Eurovoc (<https://op.europa.eu/s/y9pP> (accessed on 18 January 2024))). SKOS makes it easier for organizations and systems to collaborate and share their controlled vocabulary, and the adoption of HTTP dereferenceable URIs for thesaurus terms enabled the possibility of stating machine-actionable relations among terms in different thesauri. The adoption of LD technology led to the creation of open, flexible, and exploitable environments of thesauri and code lists, making the interoperability and the joint exploitation of vocabularies easier [119].

Ontologies as vocabulary meant in terms of data schemes are adopted at national and international levels to improve the interoperability of government data. Internationally, noteworthy examples are exemplified by the Core Vocabularies. The European Core Vocabularies (<https://joinup.ec.europa.eu/collection/semic-support-centre/core-vocabularies> (accessed on 15 January 2024)) serve as versatile domain data models or information exchange data models. For instance, they can serve as the basis for a standardized publication format for data across base registries like cadasters, business registers, and service catalogs. These domain data models also offer a default starting point for conceptualizing and structuring data models in newly developed information systems. The Core Vocabularies can form the foundation for context-specific information exchange data models used to share data between information systems. An illustration is the application of Core Vocabularies in defining information exchange data models, such as a "Business Activity Registration Request", designed for registering new business activities for a foreign branch of a legal entity in another EU Member State. Another instance is the Italian National Catalog of Data Semantics (<https://schema.gov.it/> (accessed on 15 January 2024)), conceived as part of projects outlined in the National Recovery and Resilience Plan (NRRP). Developed collaboratively by the Italian Department for Digital Transformation and other institutions, this catalog streamlines the search and reuse of semantic assets, including ontologies, data schemes, and controlled vocabularies. It provides a valuable resource for those developing semantically and syntactically interoperable APIs.

7.3. European Data Portal

The official portal for European data (<https://data.europa.eu> (accessed on 1 February 2024)) acts as a center for accessing open data in Europe via a variety of geodata national, regional, local, and international portals. It collects more than one and a half million datasets harvested from 182 national and institutional catalogs hosted in 36 countries. The European Data Portal and the former EU Open Data Portal are combined in the official portal for European data to encourage people, companies, and organizations to reuse and easily access European open data. Notably, the European portal leverages LD principles and artifacts in diverse ways. It builds upon established controlled vocabularies like Eurovoc. The recommended and preferred method for data provision involves utilizing a DCAT-AP metadata model, facilitating effective catalog harvesting from source catalogs. Notably, all metadata on the portal are structured as RDF triples, allowing easy querying using the SPARQL query language at the specified endpoint. Using SPARQL search empowers advanced users to locate datasets; SPARQL proves valuable in extracting specific information from extensive RDF datasets, even when they are organized in a complex manner. Queries can be executed seamlessly through the machine-readable endpoint (<https://data.europa.eu/sparql> (accessed on 1 February 2024)). For example, via the European Data portal's SPARQL endpoint, we can discover that 35 data catalogs expose at least a dataset distribution in OWL, RDF, or SPARQL (see SPARQL query Listing 1 (<https://api.triplydb.com/s/Qd3vDgKDu> (accessed on 1 February 2024))).

Listing 1. Number of catalogs on data.europa.eu providing OWL, RDF, or SPARQL distributions.

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX dct: <http://purl.org/dc/terms/>
PREFIX dcat: <http://www.w3.org/ns/dcat#>

SELECT (count(distinct ?catalog) as ?numberOfCatalogs) {
  {
    ?distribution dct:format ?format . filter(regex(str(?format), "RDF", "i")). }
  UNION {
    ?distribution dct:format ?format . filter(regex(str(?format), "SPARQL", "i")).}
  UNION {
    ?distribution dct:format ?format . filter(regex(str(?format), "OWL", "i")) }
  ?catalog dcat:dataset ?dataset .
  ?dataset dcat:distribution ?distribution .
}
```

This piece of information can be obtained via a SPARQL query but it is not as straightforward through the data portal interface, as the primary focus of the portal interface is on locating datasets based on format, producer, and themes rather than facilitating a comprehensive examination of features and statistics derived from the entire data collection.

For additional guidance and details, refer to the documentation (https://dataeuropa.gitlab.io/data-provider-manual/pdf/documentation_data-europa-eu_V1.0.pdf (accessed on 1 February 2024)).

7.4. National and International LOGD Take-Up

Numerous countries have demonstrated a keen interest in, or have actively embraced, LD principles as part of their governmental open data initiatives. An explicit endorsement of LD is found in the Data Strategy of the Federal German Government, which articulates its future-looking vision: "Data records are to be made usable through (semantic) connections with the aid of LD and application programming interfaces for the application of artificial intelligence. (<https://www.bundesregierung.de/resource/blob/992814/1950610/5f386401845da99721f9faa082f415cf/datenstrategie-der-bundesregierung-englisch-download-bpa-data.pdf> (accessed on 15 January 2024))". The Agency for Digital Italy (AGID), tasked with achieving the objectives of the Italian digital agenda and promoting the widespread adoption of information and communication technologies, advocates for a systematic approach towards the native production of LOD (<https://www.agid.gov.it/sites/default/>

[files/repository_files/lg-open-data_v.1.0_1.pdf](#) (accessed on 15 January 2024)). Emphasizing the importance of adhering to LOD best practices based on W3C standards, including various serializations of RDF, AGID underscores that such adherence is crucial for attaining the fourth and fifth levels in the five-star model for open data. In Spain, practical guidelines are offered as part of the Aporta Initiative, a program developed by the Ministry of Economic Affairs and Digital Transformation through the Public Business Entity Red.es. These guidelines, accessible at <https://datos.gob.es/en/documentacion/practical-guide-publication-linked-data> (accessed on 15 January 2024), aim to assist organizations interested in converting their tabular data (commonly found in open data portals) into RDF. The guide comprehensively describes and consolidates best practices, tips, and workflows, empowering those managing open data portals and individuals preparing data for portal publication to create RDF datasets efficiently and sustainably over time. Other notable examples of LD practices in the context of governmental data are constituted by the Italian parliament. Both the wings of the Italian parliament, Camera and Senato, publish data about the parliamentary work in RDF and provide SPARQL endpoints (<https://dati.camera.it/sparql> (accessed on 15 January 2024), <https://dati.senato.it/sparql> (accessed on 15 January 2024)). The European Patent Office publishes information about European patent applications using LD. By means of LD, an HTTP dereferenceable URI is associated with each patent application and its related data can be queried via SPARQL endpoint (<https://data.epo.org/linked-data/sparql.html> (accessed on 1 February 2024)). Further bodies in the UK, such as the Office of National Statistics and the Scottish Government, have adopted SPARQL endpoints as access points for information on statistical geographies (<https://statistics.data.gov.uk/sparql> (accessed on 1 February 2024), <https://statistics.data.gov.uk/sparql> (accessed on 1 February 2024)).

8. Discussion and Conclusions

To address the paper's title question, *Are LOGD still a viable option for sharing and integrating public data?* we explored the three distinct yet complementary research questions outlined in Figure 1. This approach provided insights into the challenges and opportunities of adopting LD practices within the OGD context.

RQ1 is divided into RQ1.1 and RQ1.2, providing an overall snapshot of the current landscape of LOGD. RQ1.1 reveals that the number of datasets served natively as RDF is low compared to the vast amount of government data produced. Only 2.4% of data is available in RDF across the 78 national portals analyzed and only five of them provide an active SPARQL endpoint. Question RQ1.2 suggests a deceleration of LOGD momentum. However, this fact can have a double reading. On the one hand, it seems to suggest a possible decrease for scholars in the perceived importance of LODs for OGD progress compared to the early years of the last decade. Alternatively, a more favorable interpretation may suggest that the consolidation of LODs within OGD research has reached a plateau, potentially leading researchers to explore other areas.

In the opposite direction, in RQ3, we have shown evidence that LD practices have a noteworthy impact, especially on how governmental data is aggregated, and made searchable and accessible by catalogs, as primarily witnessed by the aggregation served by the European Data Portal (more than 1,600,000 datasets from 36 nations and 183 catalogs, made available in a harmonized access point). Apart from the impact on data catalogs, LD practices have affected national guidelines and have improved data interoperability, consolidating practices for sharing controlled vocabulary (i.e., via publication of SKOS terminology and thesauri) and fostering the adoption of transversal data models such as in the case of the Italian National Catalog of Data Semantics and the European Core Vocabularies. Notably, the influence of guidelines varies across nations. In the instances of Italy and Spain, the impact is discernible through the availability of datasets adhering to LD principles, as illustrated in Figure 2. Conversely, in the case of the German Federal Government, the impact is characterized more as a visionary pursuit rather than an immediate, tangible outcome, as evidenced by the low number of RDF distributions in the German data

portal. The LD approach enables the retention of the original semantics and carries forward the choices made in data modeling. This makes data consolidation and interpretation more reliable, as information can be handled more consciously and coherently.

However, the contrast emerging from RQ1 and RQ3 finds an explanation in the frictions still present and analyzed in RQ2 and tensions that need to be balanced. The first tension is between voluntary efforts vs well-funded coordinated efforts. *Data activism*, as advocated by some LD enthusiasts, must be encouraged, but the generation of high-quality data cannot rely solely on activism. Some public administrations already allocate resources for producing and processing quality data; the key is to adopt technologies that efficiently leverage existing funding, unlocking the data's full potential. The second tension revolves around the effort required for *publishing versus consuming data*. The potential of LD to break down data silos, integrate information from diverse sources, and establish value chains that enhance the access and reuse of public data clashes with two conflicting interests. On one hand, providing LD semantically has the potential to simplify user searches and the comprehension of government datasets, thereby encouraging their reuse without the need for creating more complex mashup applications. On the other hand, if the complexities and constraints—both cultural and economic—associated with data publication are deemed excessive or burdensome, administrations may either refrain from publishing or release data in a suboptimal manner (e.g., using flat formats, ad hoc terminology, or exhibiting inconsistency). Consequently, users stand to lose most, if not all, of the benefits offered by OGD. Simplicity in publishing should not compromise the provision of easily interpretable data. For instance, sidestepping the responsibility of offering precise semantics through shared vocabularies imposes additional challenges on consumers when it comes to data integration—perhaps by resorting to guesswork regarding the original data semantics and misusing the information. Striking the right balance between data publishers and consumers is crucial. Linked data, in this sense, provides the tool for preserving the original semantics of data, playing on the side of fostering the adoption of standard vocabularies and explicit links between datasets provided by distinct providers. It eases splitting some of the efforts required for ETL among publishers and consumers, but the extent of the split for each of the different players needs to be balanced differently depending on the nature and origin of the data.

Although the analysis in response to RQ2 (Table 2) indicates that the majority of LOGD works primarily to address technical and organizational issues, it is crucial not to underestimate the influence of social, legislative, and economic factors. The challenges identified are frequently addressed either through specific, ad hoc solutions or, as noted by some authors, solutions deemed too general. These solutions primarily pertained to technical and methodological aspects, with occasional considerations for legislative or social factors. However, these solutions only partially succeeded in advancing LOGD practices. In our assessment, both research and government agencies often lacked the holistic perspective necessary to unlock the potential of LD for OGD. Despite its challenges, adopting an integrated approach that comprehensively addresses various dimensions of OGD is indispensable. Any integrated approach in the OGD domain, however, should keep in mind the tensions discussed and decide the appropriate trade-off to the target open data initiative. Our analysis of LOGD attrition and success stories, coupled with broader insights derived from OGD best practices, suggests the following recommendations as key components of such an approach.

Establish robust data governance to foster a culture of openness and elaborate clear policies, by championing organizational structures that streamline LOGD workflows, ensuring seamless navigation through the bureaucratic and cultural challenges prevalent in government contexts.

Identify High Value Datasets. As recognized by the EU Commission, certain subsets of government data are more strategic than others. In this context, identifying priority data and establishing a suitable budget is paramount to ensure cohesive approaches and avoid financial inefficiencies. Given their relevance, it is reasonable to expect that the

quality of these datasets would be particularly well cared for, facilitating their seamless integration into LOGD. Moreover, the capability to integrate High Value Datasets from various thematic fields through LD can be a powerful driver for maximizing their value.

Cultivate stakeholders engagement also by anchoring the *development of targeted guidelines* in real-world use cases. Guidelines should be specifically crafted for LOGD. This includes the provision of tangible examples and an intricate consideration of non-functional requirements, reinforcing a comprehensive approach directly applicable to day-to-day operations. Tailoring guidelines to address specific user needs and challenges ensures practical relevance and effectiveness. This approach aligns organizational efforts with the practicalities of user interactions, fostering a more meaningful and user-centric approach.

Longer-term maintenance activities must be explicitly factored into the equation. Maintaining up-to-date data and the continuous operation of servers, including those supporting SPARQL endpoints, incurs costs that must be carefully considered. This financial aspect must be accounted for to prevent the squandering of the already invested efforts.

Coordinated approaches should also consider a *pay-as-you-go perspective*, as not every dataset requires the same extent of LD. For example, not every dataset needs to be interlinked with the others to be reused: the adoption of URIs and dereferenceable standard data terms provided by shared LD vocabularies is a step forward in balancing the effort between publisher and consumer.

While the inherently decentralized nature aligns with the spirit of the Web of Data, the heterogeneity and complexity within public administrations and their, often intricate, organizational practices suggest a need for *centralized, domain-focused solutions*. These solutions can effectively handle the intricate processes of data transformation and integration—from raw to structured and interlinked data—thus enhancing the efficiency of published data. When recurrent issues are addressed in a centralized hub, organizations can implement a more cohesive and systematic approach to problem-solving. This centralization fosters a comprehensive understanding of recurring issues, paving the way for the implementation of more effective and standardized solutions. Instead of mandating an LD expert in each administration, a coordinated approach should prioritize shared and domain-targeted tools accessible even to smaller administrations, allowing them to participate without requiring specialized expertise. This strategy empowers public administrations to contribute effectively, emphasizing the integration of user-friendly tools to address this essential need.

The tensions and frictions under discussion serve as a catalyst for re-evaluating certain assumptions commonly held in the LD framework, which may not seamlessly align with the LOGD. Unlike the broader LD perspective, where data is often disseminated in an entirely unregulated manner without an explicit public mandate for data production, the governmental context demands a distinct approach. In the realm of Government Data, the combined governance of LD and OGD practices within the collaborative network of stakeholders and initiatives becomes crucial and, in our view, inevitable, in order to best benefit from the potential offered by LD. Unlike the more generalized LD viewpoint, the LOGD framework necessitates a deliberate and explicit commitment to a public mandate in data production. This shift underscores the significance of orchestrating effective governance mechanisms among the diverse actors involved in LOGD initiatives, acknowledging their pivotal role in fulfilling the public mandate for data accessibility and transparency in governmental operations. Establishing a seamlessly interconnected global data space for OGD on a worldwide scale might be deemed unattainable and could potentially lead to disillusionment. Nevertheless, the endeavor to align with LD principles and technologies should be viewed as a strategic investment. This investment is aimed at maximizing the efficiency of data production efforts, refining the streamlined data production processes and unlocking future multiplication factors within the data value chain to extract enhanced value. In essence, the commitment to LD principles becomes a proactive approach to optimize the generation and utilization of data, paving the way for greater efficiency and value realization in the broader data ecosystem. From this perspective alone, we posit that

the question presented in the title of this paper can be answered in the affirmative: *yes, Linked Open Government Data remains a viable option for sharing and integrating public data.*

Author Contributions: The authors contributed in the following way to the work reported in this paper and to its writing: Conceptualization, A.Q.; methodology, A.Q. and R.A.; investigation, A.Q. and R.A.; writing A.Q. and R.A.; project administration, A.Q.; funding acquisition A.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Table A1 presents data for the 78 countries included in our study, all of which were rated as “High” or “Very High” according to the OGD Index ranking [7]. For each country, we provide the URLs of the national portal when reachable or known, information about the published datasets, the number of RDF distributions, and the corresponding percentage relative to the total. Additionally, we indicate the presence or absence of a SPARQL endpoint. In some instances, particularly with several African countries, although a national portal was present and accessible, obtaining quantitative data on the datasets was not possible.

Table A1. International portals ordered by percentage of RDF distributions. Extends and modifies [7] Annex Table 18. Open Government Data Index OGD Index (pp. 317–321). Legend: n.a, portal/URL not accessible/available.

Country	Region	OGDI	Portal URL (accessed on 1 February 2024)	#Datasets	#RDF	%RDF	SPARQL
Italy	Europe	Very High	https://www.dati.gov.it	59,516	3854	6.5%	YES
Switzerland	Europe	High	https://opendata.swiss	9910	572	5.8%	NO
Czech Republic	Europe	Very High	https://data.gov.cz	5474	225	4.1%	YES
USA	Americas	Very High	https://catalog.data.gov	250,717	10,297	4.1%	NO
Spain	Europe	High	https://datos.gob.es	69,879	2530	3.6%	YES
Hungary	Europe	High	https://www.opendata.hu/dataset	69	2	2.9%	NO
Sweden	Europe	Very High	https://www.dataportal.se	8585	175	2.0%	NO
Norway	Europe	Very High	https://data.norge.no	1683	29	1.7%	NO
Canada	Americas	Very High	https://search.open.canada.ca/opendata	37,724	650	1.7%	NO
Thailand	Asia	Very High	https://data.go.th/en/dataset	9172	153	1.7%	NO
Slovakia	Europe	High	https://data.gov.sk	3306	21	0.6%	NO
Poland	Europe	High	https://dane.gov.pl	2336	12	0.5%	NO
The Netherlands	Europe	Very High	https://data.overheid.nl	15,326	69	0.5%	NO
Germany	Europe	Very High	https://www.govdata.de	82,845	265	0.3%	YES
Australia	Oceania	Very High	https://data.gov.au	105,647	331	0.3%	NO
Portugal	Europe	Very High	https://dados.gov.pt	4205	13	0.3%	NO
Malaysia	Asia	Very High	https://data.gov.my	12,227	23	0.2%	NO
Latvia	Europe	High	https://data.gov.lv	793	1	0.1%	NO
Luxembourg	Europe	High	https://data.public.lu	1796	1	0.1%	NO
Uruguay	Americas	Very High	https://catalogodatos.gub.uy	2397	1	0.0%	NO
Ireland	Europe	Very High	https://data.gov.ie	14,758	5	0.0%	NO
France	Europe	Very High	https://www.data.gouv.fr	44,486	13	0.0%	NO
United Kingdom	Europe	Very High	https://www.data.gov.uk	51,502	13	0.0%	NO
Brazil	Americas	Very High	https://dados.gov.br/dados/conjuntos-dados	12,398	3	0.0%	NO
Slovenia	Europe	High	https://podatki.gov.si	4549	1	0.0%	NO
Austria	Europe	Very High	https://www.data.gv.at	44,418	1	0.0%	NO
Argentina	Americas	Very High	https://www.datos.gov.ar	1175	0	0%	NO
Bulgaria	Europe	Very High	https://data.egov.bg/data	11,216	0	0%	NO
China	Asia	Very High	n.a	-	-	-	NO
Colombia	Americas	Very High	https://www.datos.gov.co/en	7373	0	0%	NO
Cyprus	Asia	Very High	https://www.data.gov.cy	1278	0	0%	NO
Denmark	Europe	Very High	https://www.opendata.dk	700	0	0%	NO
Estonia	Europe	Very High	https://avaandmed.eesti.ee	1769	0	0%	NO
Finland	Europe	Very High	https://www.avoindata.fi	2096	0	0%	NO
Greece	Europe	Very High	http://geodata.gov.gr	248	0	0%	NO
India	Asia	Very High	https://data.gov.in	601,498	0	0%	NO
Indonesia	Asia	Very High	https://www.satupemerintah.net	-	-	-	NO
Japan	Asia	Very High	https://data.e-gov.go.jp	22,126	0	0%	NO

Table A1. Cont.

Country	Region	OGDI	Portal URL (accessed on 1 February 2024)	#Datasets	#RDF	%RDF	SPARQL
Kazakhstan	Asia	Very High	https://data.egov.kz	3757	0	0%	NO
Mexico	Americas	Very High	https://datos.gob.mx	10,014	0	0%	NO
New Zealand	Oceania	Very High	https://catalogue.data.govt.nz	32,082	0	0%	NO
Philippines	Asia	Very High	https://data.gov.ph	171	0	0%	NO
Republic of Korea	Asia	Very High	https://www.data.go.kr/en/index.do	77,980	0	0%	NO
Republic of Moldova	Europe	Very High	https://date.gov.md	1176	0	0%	NO
Russian Federation	Europe	Very High	n.a	-	-	-	NO
Saudi Arabia	Asia	Very High	https://od.data.gov.sa	6860	0	0%	NO
Singapore	Asia	Very High	https://data.gov.sg	1945	0	0%	NO
United Arab Emirates	Asia	Very High	n.a	-	-	-	NO
Uzbekistan	Asia	Very High	https://data.egov.uz	8017	0	0%	NO
Belarus	Europe	High	n.a	-	-	-	NO
Peru	Americas	High	https://www.datosabiertos.gob.pe	3270	0	0%	NO
Belgium	Europe	High	https://data.gov.be	>10,000	0	0%	NO
Ghana	Africa	High	https://data.gov.gh	315	0	0%	NO
Mauritius	Africa	High	https://data.govmu.org/dkan	464	0	0%	NO
Romania	Europe	High	https://data.gov.ro	3604	0	0%	NO
Turkey	Asia	High	n.a	-	-	-	NO
Albania	Europe	High	n.a	-	-	-	NO
Panama	Americas	High	https://www.datosabiertos.gob.pa	4191	0	0%	NO
South Africa	Africa	High	https://southafrica.opendataforafrica.org	-	-	-	NO
Ukraine	Europe	High	https://data.gov.ua	29,552	0	0%	NO
Burkina Faso	Africa	High	https://burkinafaso.opendataforafrica.org	-	-	-	NO
Croatia	Europe	High	https://data.gov.hr/ckan/dataset	2491	0	0%	YES
Georgia	Asia	High	n.a	-	-	-	NO
Qatar	Asia	High	https://www.data.gov.qa	173	0	0%	NO
Uganda	Africa	High	https://uganda.opendataforafrica.org	-	-	-	NO
Azerbaijan	Asia	High	https://www.opendata.az	510	0	0%	NO
Kenya	Africa	High	https://kenya.opendataforafrica.org	-	-	-	NO
Kuwait	Asia	High	n.a	-	-	-	NO
North Macedonia	Europe	High	n.a	-	-	-	NO
Serbia	Europe	High	https://data.gov.rs/sr	2198	0	0%	NO
Dominican Republic	Americas	High	n.a	-	-	-	NO
Bahrain	Asia	High	https://www.data.gov.bh	373	0	0%	NO
Ecuador	Americas	High	https://www.datosabiertos.gob.ec	1013	0	0%	NO
Mongolia	Asia	High	https://opendata.gov.mn/en/dataset	839	0	0%	NO
Montenegro	Europe	High	https://data.gov.me/datasets	197	0	0%	NO
Sri Lanka	Asia	High	https://data.gov.lk	144	0	0%	NO
Costa Rica	Americas	High	n.a	-	-	-	NO
Guatemala	Americas	High	n.a	-	-	-	NO

References

- Davies, T.; Walker, S.B.; Rubinstein, M.; Perini, F. *The State of Open Data: Histories and Horizons*; African Minds and IDRC: Cape Town, South Africa; Ottawa, ON, Canada, 2019. [CrossRef]
- Ding, L.; Peristeras, V.; Hausenblas, M. Linked open government data. *IEEE Intell. Syst.* **2012**, *27*, 11–15. [CrossRef]
- Kapoor, K.; Weerakkody, V.; Sivarajah, U. Open Data Platforms and Their Usability: Proposing a Framework for Evaluating Citizen Intentions. In *Open and Big Data Management and Innovation*; Janssen, M., Mäntymäki, M., Hidders, J., Klievink, B., Lamersdorf, W., Van Loenen, B., Zuiderwijk, A., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2015; Volume 9373, pp. 261–271. [CrossRef]
- Krishnamurthy, R.; Awazu, Y. Liberating data for public value: The case of Data.gov. *Int. J. Inf. Manag.* **2016**, *36*, 668–672. [CrossRef]
- Berends, J.; Carrara, W.; Radu, C. *The Economic Benefits of Open Data*; Analytical Report 9; European Union: Brussels, Belgium, 2017.
- Hendler, J.; Holm, J.; Musialek, C.; Thomas, G. US government linked open data: Semantic.data.gov. *IEEE Intell. Syst.* **2012**, *27*, 25–31. [CrossRef]
- Aquaro, V. *Digital Government in the Decade of Action for Sustainable Development*; Number 2020 in United Nations e-Government Survey; Department of Economic and Social Affairs, United Nations: New York, NY, USA, 2020.
- Barbero, M.; Bartz, K.; Linz, F.; Mauritz, S.; Wauters, P.; Chrzanowski, P.; Graux, H.; Hillebrand, A.; de Vries, M.; Innessi, A.; et al. Study to support the review of Directive 2003/98. *Reuse Public Sect. Inf.* **2018**. [CrossRef]
- Kaschesky, M.; Selmi, L. 7R Data Value Framework for Open Data in Practice: Fusepool. *Future Internet* **2014**, *6*, 556–583. [CrossRef]
- Compton, S. Success Stories: Issue 2, Global Open Data for Agriculture and Nutrition (GODAN). 2017. Available online: <https://www.godan.info/v2.pdf> (accessed on 2 February 2024).
- Stagars, M. Promises, Barriers, and Success Stories of Open Data. In *Open Data in Southeast Asia: Towards Economic Prosperity, Government Transparency, and Citizen Participation in the ASEAN*; Springer International Publishing: Cham, Switzerland, 2016; pp. 13–28. [CrossRef]

12. Eggers, W.D.; Datar, A. Connecting data to residents through data storytelling. In *The Chief Data Officer in Government: A CDO Playbook*; Shah, S., Eggers, W.D., Eds.; Deloitte: Seoul, Republic of Korea, 2018; pp. 1–4.
13. Sadiq, S.; Indulska, M. Open data: Quality over quantity. *Int. J. Inf. Manag.* **2017**, *37*, 150–154. [[CrossRef](#)]
14. Safarov, I.; Meijer, A.J.; Grimmelikhuijsen, S. Utilization of open government data: A systematic literature review of types, conditions, effects and users. *Inf. Polity* **2017**, *22*, 1–24. [[CrossRef](#)]
15. Lassinantti, J.; Ståhlbröst, A.; Runardotter, M. Relevant social groups for open data use and engagement. *Gov. Inf. Q.* **2019**, *36*, 98–111. [[CrossRef](#)]
16. Stone, A. Are Open Data Efforts Working? 2018. Available online: <http://www.govtech.com/data/Are-Open-Data-Efforts-Working.html> (accessed on 2 February 2024).
17. Quarati, A.; De Martino, M. Open government data usage: A brief overview. In Proceedings of the 23rd International Database Applications & Engineering Symposium, IDEAS 2019, Athens, Greece, 10–12 June 2019; pp. 28:1–28:8. [[CrossRef](#)]
18. Quarati, A. Open Government Data: Usage trends and metadata quality. *J. Inf. Sci.* **2023**, *49*, 887–910. [[CrossRef](#)]
19. Edwards, P.N. *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*; The MIT Press: Cambridge, MA, USA, 2010.
20. Janssen, M.; Charalabidis, Y.; Zuiderwijk, A. Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Inf. Syst. Manag.* **2012**, *29*, 258–268. [[CrossRef](#)]
21. Science, D.; Hahnel, M.; Fane, B.; Treadway, J.; Baynes, G.; Wilkinson, R.; Mons, B.; Schultes, E.; Bonino da Silva Santos, L.O.; Arefiev, P.; et al. *The State of Open Data Report 2018*; Digital Science: London, UK, 2018. [[CrossRef](#)]
22. Zuiderwijk, A.; Janssen, M.; Susha, I. Improving the speed and ease of open data use through metadata, interaction mechanisms, and quality indicators. *J. Organ. Comput. Electron. Commer.* **2016**, *26*, 116–146. [[CrossRef](#)]
23. Reiche, K.; Hofig, E. Implementation of metadata quality metrics and application on public government data. In Proceedings of the 2013 IEEE 37th Annual Computer Software and Applications Conference Workshops, Kyoto, Japan, 22–26 July 2013; pp. 236–241. [[CrossRef](#)]
24. Neumaier, S.; Umbrich, J.; Polleres, A. Automated Quality Assessment of Metadata Across Open Data Portals. *J. Data Inf. Qual.* **2016**, *8*, 2:1–2:29. [[CrossRef](#)]
25. Santos-Hermosa, G.; Quarati, A.; Loria-Soriano, E.; Raffaghelli, J.E. Why Does Open Data Get Underused? A Focus on the Role of (Open) Data Literacy. In *Data Cultures in Higher Education: Emergent Practices and the Challenge Ahead*; Raffaghelli, J.E., Sangrà, A., Eds.; Springer International Publishing: Cham, Switzerland, 2023; pp. 145–177. [[CrossRef](#)]
26. Jarke, J. Open government for all? Co-creating digital public services for older adults through data walks. *Online Inf. Rev.* **2019**, *43*, 1003–1020. [[CrossRef](#)]
27. Zuiderwijk, A.; Janssen, M.; Dwivedi, Y.K. Acceptance and use predictors of open data technologies: Drawing upon the unified theory of acceptance and use of technology. *Gov. Inf. Q.* **2015**, *32*, 429–440. [[CrossRef](#)]
28. Ruijter, E.; Grimmelikhuijsen, S.; van den Berg, J.; Meijer, A. Open data work: Understanding open data usage from a practice lens. *Int. Rev. Adm. Sci.* **2020**, *86*, 3–19. [[CrossRef](#)]
29. Dadzie, A.S.; Rowe, M. Approaches to visualising linked data: A survey. *Semant. Web* **2011**, *2*, 89–124. [[CrossRef](#)]
30. Lnenicka, M.; Komarkova, J. Big and open linked data analytics ecosystem: Theoretical background and essential elements. *Gov. Inf. Q.* **2019**, *36*, 129–144. [[CrossRef](#)]
31. Archer, P.; Dekkers, M.; Hazard, N.; Loutas, N.; Karalopoulos, A.; Peristeras, V.; Wigard, S. Business Models for Linked Open Government Data: What Lies Beneath? 2013. Available online: <https://www.w3.org/2013/share-psi/workshop/krems/papers/LinkedOpenGovernmentDataBusinessModel> (accessed on 1 February 2024).
32. Shadbolt, N.; O'Hara, K. Linked Data in Government. *IEEE Internet Comput.* **2013**, *17*, 72–77. [[CrossRef](#)]
33. Penteado, B.E.; Maldonado, J.C.; Isotani, S. Methodologies for publishing linked open government data on the Web: A systematic mapping and a unified process model. *Semant. Web* **2022**, *14*, 585–610. [[CrossRef](#)]
34. Attard, J.; Orlandi, F.; Auer, S. Value creation on open government data. In Proceedings of the 2016 49th Hawaii International Conference on System Sciences (HICSS), Koloa, HI, USA, 5–8 January 2016; pp. 2605–2614. [[CrossRef](#)]
35. Klein, E.; Gschwend, A.; Neuroni, A. Towards a linked data publishing methodology. In Proceedings of the 6th International Conference for E-Democracy and Open Government, CeDEM 2016, Krems, Austria, 18–20 May 2016. [[CrossRef](#)]
36. DiFranzo, D.; Graves, A.; Erickson, J.S.; Ding, L.; Michaelis, J.; Lebo, T.; Patton, E.; Williams, G.T.; Li, X.; Zheng, J.G.; et al. The Web is My Back-end: Creating Mashups with Linked Open Government Data. In *Linking Government Data*; Wood, D., Ed.; Springer: New York, NY, USA, 2011; pp. 205–219. [[CrossRef](#)]
37. Kalampokis, E.; Karacapilidis, N.; Tsakalidis, D.; Tarabanis, K. Understanding the Use of Emerging Technologies in the Public Sector: A Review of Horizon 2020 Projects. *Digit. Gov. Res. Pract.* **2023**, *4*, 1–28. [[CrossRef](#)]
38. Futia, G.; Melandri, A.; Vetrò, A.; Morando, F.; De Martin, J.C. Removing Barriers to Transparency: A Case Study on the Use of Semantic Technologies to Tackle Procurement Data Inconsistency. In *The Semantic Web*; Blomqvist, E., Maynard, D., Gangemi, A., Hoekstra, R., Hitzler, P., Hartig, O., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 623–637. [[CrossRef](#)]
39. Sheridan, J.; Tennison, J. Linking UK Government data. In Proceedings of the CEUR Workshop Proceedings, Heraklion, Greece, 31 May 2010.

40. Galiotou, E.; Fragkou, P. Applying Linked Data Technologies to Greek Open Government Data: A Case Study. *Procedia Soc. Behav. Sci.* **2013**, *73*, 479–486. [CrossRef]
41. Breitman, K.; Viterbo, J.; Salas, P.; Saraiva, D.; Magalhães, R.; Gama, V.; Casanova, M.; Chaves, M.; Franzosi, E. Open government data in Brazil. *IEEE Intell. Syst.* **2012**, *27*, 45–49. [CrossRef]
42. Ding, L.; Lebo, T.; Erickson, J.; Difranzo, D.; Williams, G.; Li, X.; Michaelis, J.; Graves, A.; Zheng, J.; Shangguan, Z.; et al. TWC LOGD: A portal for linked open government data ecosystems. *J. Web Semant.* **2011**, *9*, 325–333. [CrossRef]
43. Cyganiak, R.; Maali, F.; Peristeras, V. Self-service linked government data with dcat and gridworks. In *Proceedings of the 6th International Conference on Semantic Systems; ACM International Conference Proceeding Series; ACM: New York, NY, USA, 2010.* [CrossRef]
44. Janev, V.; Milošević, U.; Spasić, M.; Milojković, J.; Vraneš, S. Linked open data infrastructure for public sector information: Example from Serbia. In *I-SEMANTICS (Posters & Demos); CEUR Workshop Proceedings; ACM: New York, NY, USA, 2012.*
45. Kalampokis, E.; Tambouris, E.; Tarabanis, K. On publishing linked open government data. In *Proceedings of the 17th Panhellenic Conference on Informatics; ACM International Conference Proceeding Series; ACM: New York, NY, USA, 2013.* [CrossRef]
46. Quarati, A.; De Martino, M.; Rosim, S. Geospatial Open Data Usage and Metadata Quality. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 30. [CrossRef]
47. Hogan, A. The Semantic Web: Two decades on. *Semant. Web* **2020**, *11*, 169–185. [CrossRef]
48. Wirtz, B.W.; Weyerer, J.C.; Becker, M.; Müller, W.M. Open government data: A systematic literature review of empirical research. *Electron. Mark.* **2022**, *32*, 2381–2404. [CrossRef] [PubMed]
49. Bulazel, A.; DiFranzo, D.; Erickson, J.; Hendler, J. The importance of authoritative uri design schemes for open government data. In *Information Retrieval and Management: Concepts, Methodologies, Tools, and Applications; IGI Global: Hershey, PA, USA, 2018.* [CrossRef]
50. Lnenicka, M.; Luterek, M.; Nikiforova, A. Benchmarking open data efforts through indices and rankings: Assessing development and contexts of use. *Telemat. Inform.* **2022**, *66*, 101745. [CrossRef]
51. Zheng, L.; Kwok, W.M.; Aquaro, V.; Qi, X.; Lyu, W. Evaluating Global Open Government Data: Methods and Status. In *Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance, ICEGOV '20, Athens, Greece, 23–25 September 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 381–391.* [CrossRef]
52. Berners-Lee, T. Linked Data. 2006. Available online: <https://www.w3.org/DesignIssues/LinkedData.html> (accessed on 1 February 2024).
53. Heath, T.; Bizer, C. *Linked Data: Evolving the Web into a Global Data Space*, 1st ed.; html version ed.; Synthesis Lectures on the Semantic Web: Theory and Technology; Morgan & Claypool: San Rafael, CA, USA, 2011; Volume 1, pp. 1–136. [CrossRef]
54. Cyganiak, R.; Wood, D.; Lanthaler, M. RDF 1.1 Concepts and Abstract Syntax. W3C Recommendation, W3C. 2014. Available online: <https://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/> (accessed on 1 February 2024).
55. OWL 2 Web Ontology Language Document Overview (Second Edition). W3C Recommendation, W3C. 2012. Available online: <https://www.w3.org/TR/2012/REC-owl2-overview-20121211/> (accessed on 1 February 2024).
56. SPARQL 1.1 Overview. W3C Recommendation, W3C. 2013. Available online: <https://www.w3.org/TR/2013/REC-sparql11-overview-20130321/> (accessed on 1 February 2024).
57. Mahmud, S.; Hossin, M.; Hasan, M.; Jahan, H.; Noori, S.; Ahmed, M. Publishing CSV Data as Linked Data on the Web. In *Proceedings of ICETIT 2019; Springer: Cham, Switzerland, 2020; Volume 605, pp. 805–817.* [CrossRef]
58. Kumar, B.P. 9—Open Data for smart cities. In *Solving Urban Infrastructure Problems Using Smart City Technologies; Vacca, J.R., Ed.; Elsevier: Amsterdam, The Netherlands, 2021; pp. 185–211.* [CrossRef]
59. Guha, R.V.; Brickley, D.; MacBeth, S. Schema.Org: Evolution of Structured Data on the Web: Big Data Makes Common Schemas Even More Necessary. *Queue* **2015**, *13*, 10–37. [CrossRef]
60. Velitchkov, I. Linked Data Uptake. 4 April 2021. Available online: <https://www.strategicstructures.com/?p=2193> (accessed on 12 January 2024).
61. Pawełoszek, I.; Wiczorkowski, J. Open government data and linked data in the practice of selected countries. In *European Conference on e-Digital Government; Academic Conferences International Limited: Reading, UK, 2018.*
62. Ibanez, L.; Millard, I.; Glaser, H.; Simperl, E. An Assessment of Adoption and Quality of Linked Data in European Open Government Data. In *The Semantic Web—ISWC 2019; Springer: Cham, Switzerland, 2019; Volume 11779, pp. 436–453.* [CrossRef]
63. Akatkin, Y.; Yasinovskaya, E. Data-Driven Government in Russia: Linked Open Data Challenges, Opportunities, Solutions. In *Communications in Computer and Information Science; Springer: Cham, Switzerland, 2020.* [CrossRef]
64. Zuiderwijk, A.; Janssen, M.; Choenni, S.; Meijer, R.; Alibaks, R.S. Socio-technical Impediments of Open Data. *Electron. J.-Gov.* **2012**, *10*, 156–172.
65. Attard, J.; Orlandi, F.; Scerri, S.; Auer, S. A systematic review of open government data initiatives. *Gov. Inf. Q.* **2015**, *32*, 399–418. [CrossRef]
66. Verma, N.; Gupta, M. Challenges in publishing open government data: A study in Indian context. In *Proceedings of the 2015 2nd International Conference on Electronic Governance and Open Society: Challenges in Eurasia; ACM International Conference Proceeding Series; ACM: New York, NY, USA, 2015.* [CrossRef]

67. Roa, H.N.; Loza-Aguirre, E.; Flores, P. A Survey on the Problems Affecting the Development of Open Government Data Initiatives. In Proceedings of the 2019 Sixth International Conference on eDemocracy & eGovernment (ICEDEG), Quito, Ecuador, 24–26 April 2019; pp. 157–163. [\[CrossRef\]](#)
68. Portisch, J.; Fallatah, O.; Neumaier, S.; Jaradeh, M.Y.; Polleres, A. Challenges of Linking Organizational Information in Open Government Data to Knowledge Graphs. In *Knowledge Engineering and Knowledge Management*; Keet, C.M., Dumontier, M., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2020; Volume 12387, pp. 271–286. [\[CrossRef\]](#)
69. Geci, M.; Csáki, C. The Potential of BOLD in National Budget Planning: Opportunities and Challenges for Kosovo. In *Electronic Government*; Scholl, H.J., Gil-Garcia, J.R., Janssen, M., Kalampokis, E., Lindgren, I., Rodríguez Bolívar, M.P., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2021; Volume 12850, pp. 178–189. [\[CrossRef\]](#)
70. Mouzakitis, S.; Papaspyros, D.; Petychakis, M.; Koussouris, S.; Zafeiropoulos, A.; Fotopoulou, E.; Farid, L.; Orlandi, F.; Attard, J.; Psarras, J. Challenges and opportunities in renovating public sector information by enabling linked data and analytics. *Inf. Syst. Front.* **2017**, *19*, 321–336. [\[CrossRef\]](#)
71. Buil-Aranda, C.; Hogan, A.; Umbrich, J.; Vandebussche, P.Y. SPARQL Web-Querying Infrastructure: Ready for Action? In *The Semantic Web—ISWC 2013*; Alani, H., Kagal, L., Fokoue, A., Groth, P., Biemann, C., Parreira, J.X., Aroyo, L., Noy, N., Welty, C., Janowicz, K., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 277–293.
72. de Oliveira, E.F.; Silveira, M.S. Open Government Data in Brazil a Systematic Review of Its Uses and Issues. In *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*; dg.o '18; Association for Computing Machinery: New York, NY, USA, 2018. [\[CrossRef\]](#)
73. Attard, J.; Orlandi, F.; Auer, S. Data driven governments: Creating value through open government data. In *Transactions on Large-Scale Data- and Knowledge-Centered Systems XXVII*; Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer: Berlin/Heidelberg, Germany, 2016. [\[CrossRef\]](#)
74. Kaoudi, Z.; Manolescu, I. Triples in the clouds. In Proceedings of the 2013 IEEE 29th International Conference on Data Engineering (ICDE), Brisbane, QLD, Australia, 8–12 April 2013. [\[CrossRef\]](#)
75. Theocharis, S.; Tsihrintzis, G. Ontology development to support the Open Public data—The Greek case. In Proceedings of the IISA 2014—5th International Conference on Information, Intelligence, Systems and Applications, Chania, Greece, 7–9 July 2014. [\[CrossRef\]](#)
76. Araújo, I.; Reis, A.; Mariano, A.; Oviedo, V. Design and Application of the AHP-TOPSIS-2N to Evaluate (Linked) Open Government Data from the Electricity Datasets. In *Intelligent Sustainable Systems*; Lecture Notes in Networks and Systems; Springer: Singapore, 2023. [\[CrossRef\]](#)
77. Vert, S.; Vasiu, R. Relevant Aspects for the Integration of Linked Data in Mobile Augmented Reality Applications for Tourism. In *Information and Software Technologies*; Dregvaite, G., Damasevicius, R., Eds.; Communications in Computer and Information Science; Springer International Publishing: Cham, Switzerland, 2014; Volume 465, pp. 334–345. [\[CrossRef\]](#)
78. Wiczorkowski, J. *Barriers to Using Open Government Data*; ACM International Conference Proceeding Series; ACM: New York, NY, USA, 2019. [\[CrossRef\]](#)
79. Narducci, F.; Palmonari, M.; Semeraro, G. Cross-Language Semantic Retrieval and Linking of E-Gov Services. In *The Semantic Web—ISWC 2013*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8219, pp. 130–145.
80. Akatkin, Y.; Laikam, K.; Yasinovskaya, E. The Concept and the Roadmap to Linked Open Statistical Data in the Russian Federation. In *Electronic Governance and Open Society: Challenges in Eurasia. EGOSE 2021*; Communications in Computer and Information Science; Springer: Cham, Switzerland, 2022. [\[CrossRef\]](#)
81. Subramanian, A.; Garg, A.; Poddar, O.; Srinivasa, S. Towards semantically aggregating indian open government data from data.gov.in. In *ISWC (Posters, Demos & Industry Tracks)*; CEUR Workshop Proceedings; ACM: New York, NY, USA, 2017.
82. Espinoza-Arias, P.; Fernandez-Ruiz, M.; Morlan-Plo, V.; Notivol-Bezares, R.; Corcho, O. The Zaragoza's Knowledge Graph: Open Data to Harness the City Knowledge. *Information* **2020**, *11*, 129. [\[CrossRef\]](#)
83. Höchtl, J.; Reichstädter, P. Linked open data—A means for public sector information management. In *Electronic Government and the Information Systems Perspective. EGOVIS 2011*; Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer: Berlin/Heidelberg, Germany, 2011. [\[CrossRef\]](#)
84. Brys, C.; Aldana-Montes, J. A semantic model for electronic government and its enforcement in the Province of Misiones, Argentina. *Electron. Gov.* **2016**, *12*, 337–356. [\[CrossRef\]](#)
85. Buranarach, M.; Ruengittinun, S.; Krataithong, P.; Supnithi, T.; Hinsheran, S. A scalable framework for creating open government data services from open government data catalog. In Proceedings of the 9th International Conference on Management of Digital EcoSystems, MEDES 2017, Bangkok, Thailand, 7–10 November 2017. [\[CrossRef\]](#)
86. Lebo, T.; Erickson, J.; Ding, L.; Graves, A.; Williams, G.; DiFranzo, D.; Li, X.; Michaelis, J.; Zheng, J.; Flores, J.; et al. Producing and Using Linked Open Government Data in the TWC LOGD Portal. In *Linking Government Data*; Springer: Berlin/Heidelberg, Germany, 2011; p. 72. [\[CrossRef\]](#)
87. Shi, L.; Sukhobok, D.; Nikolov, N.; Roman, D. Norwegian State of estate report as linked open data. In *On the Move to Meaningful Internet Systems. OTM 2017 Conferences. OTM 2017*; Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer: Cham, Switzerland, 2017. [\[CrossRef\]](#)

88. Deng, D.; Mai, G.; Shiau, S. Construction and Reuse of Linked Agriculture Data: An Experience of Taiwan Government Open Data. In *Semantic Technology. JIST 2018*; Springer: Cham, Switzerland, 2018; Volume 11341, pp. 367–382. [\[CrossRef\]](#)
89. Publications Office of the European Union; European Commission, Directorate General for Informatics. *Creating Public Sector Value through the Use of Open Data: Insights and Recommendations from the data.europa.eu Campaign: Summary Paper 2023*; Publications Office of the European Union: Luxembourg, 2023.
90. Tambouris, E. Multidimensional open government data. *EJournal EDemocracy Open Gov.* **2016**, *8*, 1–11. [\[CrossRef\]](#)
91. Abida, R.; Belghith, E.; Cleve, A. An End-to-End Framework for Integrating and Publishing Linked Open Government Data. In *Proceedings of the Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises, WETICE, Bayonne, France, 10–13 September 2020*. [\[CrossRef\]](#)
92. Sinif, L.; Bounabat, B. *Approaching an Optimizing Open Linked Government Data Portal*; ACM International Conference Proceeding Series; ACM: New York, NY, USA, 2018. [\[CrossRef\]](#)
93. Trinh, T.D.; Do, B.L.; Wetz, P.; Anjomshoaa, A.; Tjoa, A. Linked widgets: An approach to exploit open government data. In *Proceedings of the 2nd International Conference on Smart Digital Environment*; ACM International Conference Proceeding Series; ACM: New York, NY, USA, 2013. [\[CrossRef\]](#)
94. Zhu, X.; Thomas, C.; Moore, J.; Allen, S. Open Government Data Licensing: An Analysis of the U.S. State Open Government Data Portals. In *Diversity, Divergence, Dialogue. iConference 2021*; Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer: Cham, Switzerland, 2021. [\[CrossRef\]](#)
95. Morando, F. Legal Interoperability: Making Open (Government) Data Compatible with Businesses and Communities. *JLIS.IT* **2013**, *4*, 441–452. [\[CrossRef\]](#)
96. Matheus, R.; Ribeiro, M.; Vaz, J. Brazil towards government 2.0: Strategies for adopting open government data in national and subnational governments. In *Case Studies in E-Government 2.0: Changing Citizen Relationships*; Springer: Cham, Switzerland, 2015. [\[CrossRef\]](#)
97. Alexopoulos, C.; Loukis, E.; Mouzakitis, S.; Petychakis, M.; Charalabidis, Y. Analysing the Characteristics of Open Government Data Sources in Greece. *J. Knowl. Econ.* **2018**, *9*, 721–753. [\[CrossRef\]](#)
98. Lebo, T.; Wang, P.; Graves, A.; McGuinness, D. Towards Unified Provenance Granularities. In *Provenance and Annotation of Data and Processes. IPAW 2012*; Springer: Berlin/Heidelberg, Germany, 2012; Volume 7525, pp. 39–51.
99. Albertoni, R.; Martino, M.D.; Podestà, P. Quality measures for skos: ExactMatch linksets: An application to the thesaurus framework LusTRE. *Data Technol. Appl.* **2018**, *52*, 405–423. [\[CrossRef\]](#)
100. Albertoni, R.; Gómez-Pérez, A. Assessing linkset quality for complementing third-party datasets. In *Proceedings of the Joint EDBT/ICDT 2013 Workshops*; ACM: New York, NY, USA, 2013; pp. 52–59.
101. Quarati, A.; Albertoni, R.; Martino, M.D. Overall quality assessment of SKOS thesauri: An AHP-based approach. *J. Inf. Sci.* **2017**, *43*, 816–834. [\[CrossRef\]](#)
102. Albertoni, R.; De Martino, M.; Quarati, A. Documenting Context-Based Quality Assessment of Controlled Vocabularies. *IEEE Trans. Emerg. Top. Comput.* **2021**, *9*, 144–160. [\[CrossRef\]](#)
103. Ngomo, A.C.N.; Auer, S.; Lehmann, J.; Zaveri, A. Introduction to Linked Data and Its Lifecycle on the Web. In *Reasoning Web. Reasoning on the Web in the Big Data Era: 10th International Summer School 2014, Athens, Greece, September 8–13, 2014. Proceedings*; Koubarakis, M., Stamou, G., Stoilos, G., Horrocks, I., Kolaitis, P., Lausen, G., Weikum, G., Eds.; Springer International Publishing: Cham, Switzerland, 2014; pp. 1–99. [\[CrossRef\]](#)
104. European Commission; Directorate-General for Communications Networks, Content and Technology. *Identification of Data Themes for the Extensions of Public Sector High-Value Datasets—Final Study*; Publications Office of the European Union: Luxembourg, 2023. [\[CrossRef\]](#)
105. European Commission; Directorate-General for Communications Networks, Content and Technology. *Directive (EU) 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 (Recast)*; Publications Office of the European Union: Luxembourg, 2019.
106. Nikiforova, A.; Rizun, N.; Ciesielska, M.; Alexopoulos, C.; Miletić, A. Towards High-Value Datasets Determination for Data-Driven Development: A Systematic Literature Review. In *Electronic Government*; Lindgren, I., Csáki, C., Kalampokis, E., Janssen, M., Viale Pereira, G., Virkar, S., Tambouris, E., Zuiderwijk, A., Eds.; Springer Nature: Cham, Switzerland, 2023; pp. 211–229.
107. Gomes, J., Jr.; Bernardino, H.S.; de Souza, J.F.; Rajabi, E. Indexing, enriching, and understanding Brazilian missing person cases from data of distributed repositories on the web. *AI Soc.* **2023**, *38*, 565–579. [\[CrossRef\]](#)
108. Karamanou, A.; Kalampokis, E.; Tarabanis, K.A. Linked Open Government Data to Predict and Explain House Prices: The Case of Scottish Statistics Portal. *Big Data Res.* **2022**, *30*, 100355. [\[CrossRef\]](#)
109. Albertoni, R.; Browning, D.; Cox, S.; González Beltrán, A.; Perego, A.; Winstanley, P. Data Catalog Vocabulary (DCAT)-Version 2. W3C Recommendation, W3C. 2020. Available online: <https://www.w3.org/TR/vocab-dcat-2/> (accessed on 1 February 2024).
110. Albertoni, R.; Browning, D.; Cox, S.; Gonzalez-Beltran, A.N.; Perego, A.; Winstanley, P. The W3C Data Catalog Vocabulary, Version 2: Rationale, Design Principles, and Uptake. *Data Intell.* **2023**, 1–37. [\[CrossRef\]](#)
111. van Nuffelen, B. *DCAT Application Profile for Data Portals in Europe, Version 2.1; Technical Specification*; European Commission: Brussels, Belgium, 2020.

112. Perego, A.; van Nuffelen, B. *GeoDCAT-AP-Version 2.0.0: A Geospatial Extension for the DCAT Application Profile for Data Portals in Europe*; Semic Recommendation; European Commission: Brussels, Belgium, 2020.
113. Dragan, A.; Sofou, N. *StatDCAT-AP-DCAT Application Profile for Description of Statistical Datasets, Version 1.0.1*; Technical Specification; European Commission: Brussels, Belgium, 2019.
114. Interian, R.; Mendoza, I.; Bernardini, F.; Viterbo, J. Unified vocabulary in Official Gazettes: An exploratory study on procurement data. In *Proceedings of the 15th International Conference on Theory and Practice of Electronic Governance, ICEGOV '22*, Guimarães, Portugal, 4–7 October 2022; Association for Computing Machinery: New York, NY, USA, 2022; pp. 195–202. [[CrossRef](#)]
115. Alvarez-Rodríguez, J.M.; Labra-Gayo, J.E.; Rodríguez-González, A.; De Pablos, P.O. Empowering the access to public procurement opportunities by means of linking controlled vocabularies. A case study of Product Scheme Classifications in the European e-Procurement sector. *Comput. Hum. Behav.* **2014**, *30*, 674–688. [[CrossRef](#)]
116. Baker, T.; Bechhofer, S.; Isaac, A.; Miles, A.; Schreiber, G.; Summers, E. Key choices in the design of Simple Knowledge Organization System (SKOS). *J. Web Semant.* **2013**, *20*, 35–49. [[CrossRef](#)]
117. Bechhofer, S.; Miles, A. SKOS Simple Knowledge Organization System Reference. W3C Recommendation, W3C. 2009. Available online: <https://www.w3.org/TR/2009/REC-skos-reference-20090818/> (accessed on 1 February 2024).
118. Coll, I.S.; Kolshus, K.; Turbati, A.; Stellato, A.; Mietzsch, E.; Martini, D.; Zeng, M. AGROVOC: The linked data concept hub for food and agriculture. *Comput. Electron. Agric.* **2022**, *196*, 105965. [[CrossRef](#)]
119. Albertoni, R.; Martino, M.D.; Podestà, P.; Abecker, A.; Wössner, R.; Schnitter, K. LusTRE: A framework of linked environmental thesauri for metadata management. *Earth Sci. Inform.* **2018**, *11*, 525–544. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.