

# Synchronization and Standardization of Open Data Platforms: A Systematic Literature Review

Besart Hyseni<sup>1</sup>, Lejla Abazi Bexheti<sup>1</sup>

<sup>1</sup>South East European University, Faculty of Contemporary Sciences and Technologies,  
Tetovo, North Macedonia

**Abstract** – This study is a systematic review that intends to improve our knowledge of the ideas of open data synchronization and standardization by reviewing the methodology and instruments utilized in previous studies through a careful examination of accessible literature. The methodology of this study was developed in several key phases. Firstly, a search was conducted in electronic databases to identify relevant studies in the field for Open Government Data (OGD). After identifying the studies, they were carefully reviewed through an initial abstract-based screening process. In the next phase, the research design and sample used in each identified study were analyzed to assess their quality and reliability. Subsequently, data from the studies were analyzed and categorized into various categories, including the usability of open data, synchronization, scalability, and reliability. The gained results highlight the importance of open data in our information-driven society. Open data positively impact transparency, anti-corruption efforts, and free access to information. To achieve these benefits, it is crucial to standardize and license open data, giving priority to their usability. The study underscores the importance of open data for transparency, innovation, and citizen engagement. Despite the disadvantages, this technology has the potential to change how governments and society use data.

DOI: 10.18421/TEM132-40

<https://doi.org/10.18421/TEM132-40>


**Corresponding author:** Besart Hyseni,  
South East European University, Faculty of Contemporary  
Sciences and Technologies, Tetovo, North Macedonia  
**Email:** [bh29738@seeu.edu.mk](mailto:bh29738@seeu.edu.mk)

Received: 07 January 2024.

Revised: 25 March 2024.

Accepted: 01 April 2024.

Published: 28 May 2024.

 © 2024 Besart Hyseni & Lejla Abazi Bexheti; published by UIKTEN. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 License.

The article is published with Open Access at  
<https://www.temjournal.com/>

**Keywords** – Open data synchronization, open data standardization, open data platforms, systematic literature review.

## 1. Introduction

In recent years, open government data has gained significant attention and has become a common topic among transparency advocates, non-governmental organizations (NGOs), and public officials. Numerous academic studies now focus on Open Government Data (OGD) initiatives and policymaking to explain variations in OGD provisions across government organizations [1]. Indeed, there has been extensive scholarly interest in OGD and its provision by governments [2]. The open knowledge meaning (OKD) clarifies the meaning of "open," which focuses around the ability to access, reuse, and redistribute data without limitations, with the exception of credit and share-alike requirements. Any further constraints would classify an item as closed knowledge. It also underlines the need of guaranteeing the overall usability and accessibility of the dataset or knowledge work by making it available at a fair cost, preferably via free internet downloads, and in a handy and editable format [3]. This is especially important in the context of scientific data, where access varies and license for reuse might be unclear [4]. Supplementary data linked with research publications is sometimes open even if the article itself is not, however this is not always explicitly mentioned. Clear labeling and licensing are critical to save scientists time wasted establishing the openness of datasets, especially as computerized analysis of scientific literature becomes more common [5].

Open government data (OGD) has the potential to be a powerful instrument for societal and economic progress. It can improve public governance by encouraging citizen-driven approaches to service design, increasing public sector efficiency, and fostering honesty and accountability [6].

Governments may design evidence-based policies, foster innovation both inside and outside the public sector, and empower citizens to make well-informed decisions by assuring OGD availability, accessibility, and reuse by multiple stakeholders. Recognizing the benefits of OGD, several international instruments, including the G8 Open Data Charter in 2013, the International Open Data Charter (IoDC) in 2015, and the G20 Anti-Corruption Open Data Principles, were adopted in recent years to encourage policies that promote access to government data [7].

Different fields have distinct definitions of data. Information systems and knowledge management fields often define data within the data-information-knowledge-wisdom (DIKW) framework [8]. In contrast, the legal field often considers data and information as synonymous. For instance, personal data is defined as "any information in respect of commercial transactions" in the Malaysian Personal Data Protection Act 2010 [9]. While the rigid distinction between data and information may be useful in certain contexts, it may have limited utility outside those fields [10]. Data typically becomes meaningful when it can be transformed into information. Data and information are interconnected, with data serving as the building blocks of information. Open data standardization involves the establishment and implementation of consistent rules, formats, and protocols for structuring and presenting openly accessible data to the public. This standardization aims to ensure that open datasets are organized uniformly and formatted in an interoperable manner, facilitating easy access, sharing, and analysis by a diverse range of users and applications [11], [12].

### ***1.1. Open Data Usability***

Open data has gained significant attention in recent years as governments, organizations, and individuals increasingly share data with the public. Usability is a critical aspect of open data, as it determines how effectively users can access, understand, and utilize the available information [13]. This literature review explores the concept of open data usability, its importance, challenges, and best practices. Usability is a key factor in ensuring that open data achieves its intended objectives. When data is usable, it becomes a valuable resource for various stakeholders, including researchers, policymakers, businesses, and the general public [9]. Usable open data can facilitate evidence-based decision-making, foster innovation, and promote transparency [14], [15]. The usability of open data is critical to realize the full potential of open data projects. Addressing challenges related to data quality, accessibility, and metadata while implementing best practices can improve the

usability of open data, thereby enabling more effective decision-making, innovation, and transparency [16].

### ***1.2. Open Data Standardization***

Standardization is an important part of open data, since it ensures that data is structured, presented, and expressed consistently in order to improve interoperability and usability. This study of the literature delves into the notion of open data standards, its importance, obstacles, and best practices in this sector [17]. Open data standards are critical to accomplishing open data projects' aims of openness, accessibility, and data reuse. Standardized data formats and metadata enable users to easily discover, access, and integrate data from various sources, facilitating data-driven decision-making and innovation [18]. Open data standardization is a critical element in realizing the potential of open data initiatives. By addressing challenges related to diverse data sources, evolving standards, and interoperability, and by implementing best practices, open data can be made more accessible, reusable, and valuable to a wide range of users and applications [19].

### ***1.3. Open Data Synchronization***

Open data synchronization is a critical aspect of maintaining the timeliness, accuracy, and reliability of open datasets. This literature review explores the concept of open data synchronization, its significance, challenges, and best practices [20]. Open data, by its nature, is dynamic and subject to changes over time. Synchronization ensures that the published datasets remain up-to-date, reflecting real-world developments. Timely and accurate open data is essential for informed decision-making, research, and application development [21]. Open data synchronization is essential for ensuring that open datasets remain accurate, timely, and reliable. By addressing challenges related to data source diversity, volume, and schema evolution, and by implementing best practices such as real-time updates, versioning, data provenance, and automated monitoring, open data can better serve the needs of users and applications [22].

### ***1.4. Open Data Scalability***

Scalability is a fundamental consideration when dealing with open data initiatives, as it determines how well data platforms and infrastructures can grow to accommodate increasing data volumes, users, and demands [23].

This literature review explores the concept of open data scalability, its significance, challenges, and best practices [24]. Open data initiatives aim to make data accessible to a wide audience, and as these initiatives gain popularity, the volume of data and users can increase rapidly [25]. Scalability is crucial to ensuring that data platforms can handle this growth without compromising performance or accessibility [26]. Open data scalability is crucial for ensuring that open data platforms can handle growing data volumes and user demands [27]. By addressing challenges related to data volume, diversity, and user load, and by implementing best practices such as cloud infrastructure, distributed processing, data partitioning, and caching, open data initiatives can continue to provide accessible and performant data resources [28].

### **1.5. Open Data Reliability**

Reliability is a critical aspect of open data, as it ensures that the data is accurate, trustworthy, and dependable. This literature review also explores the concept of open data reliability, its significance, challenges, and best practices [29]. Reliability is essential for open data because users, including researchers, policymakers, and businesses, rely on accurate and consistent information for decision-making, analysis, and innovation [30]. Unreliable data can lead to incorrect conclusions and decisions, eroding trust in open data initiatives [31]. The dependability of open data is critical to the success of open data efforts [32]. Open data can provide users with reliable and trustworthy information for a variety of applications by addressing challenges related to data quality, source diversity, and data updates, as well as implementing best practices such as data quality assurance, metadata and provenance tracking, data governance, and user feedback mechanisms [33].

### **1.6. Open Data Advantages**

Open data, characterized by its accessibility and openness for public use, has gained widespread attention due to its potential to bring about numerous benefits. This study also thoroughly investigates the advantages of open data as identified in academic research and practical applications [18]. Open data offers a multitude of advantages, ranging from transparency and accountability to fostering innovation, improving decision-making, enhancing citizen engagement, and supporting research and education. These benefits are well-documented in academic literature and have practical implications for governments, businesses, and society as a whole [34].

### **1.7. Open Data Disadvantages**

While open data has numerous advantages, it also comes with its share of disadvantages and challenges [35]. This literature review explores the disadvantages of open data as identified in academic research and practical applications [36]. Privacy concerns, data quality issues, security risks, potential misuse, the digital divide, and resource constraints are among the key disadvantages associated with open data. Recognizing and addressing these drawbacks is essential for the responsible management and use of open data [37], [38].

## **2. Methodology**

The methodology of this study has been constructed in several distinct phases to encompass the review, analysis, and synthesis of open government data (OGD) as presented in Figure 1. For data collection, the study commences with a search in electronic databases. This involves the examination of various electronic sources to identify relevant studies in the field of OGD. Once identified, these studies undergo an initial screening process based on their abstracts. This process aids in discovering studies related to the OGD topic and deemed suitable for deeper analysis. Following the initial review process, a comprehensive examination phase of the research design and sample used in each identified study is conducted. This includes a detailed analysis of how the studies are designed and how samples are selected for research. This thorough investigation assists in assessing the quality and reliability of the studies. The final stage of data collection involves a comprehensive review of the text of each identified study. This includes a deep analysis of all parts of the study, including methodology, results, and conclusions. This process aids in understanding the details of each study and helps establish context for a broader analysis. After data collection is complete, the study proceeds to the data analysis phase. The data is divided into numerous areas, including those pertaining to open data usability, synchronization, scalability, and dependability. Additionally, the effects, conditions, advantages, and disadvantages of OGD usage are analyzed, as well as the platforms and users involved. Following the data analysis, the synthesis phase begins. The data is grouped and classified according to types of usage, effects, conditions, and users. In this phase, the aim is to establish a framework for OGD usage and clarify the limitations that may shape further research in this field.

Table 1. Conceptual model of research

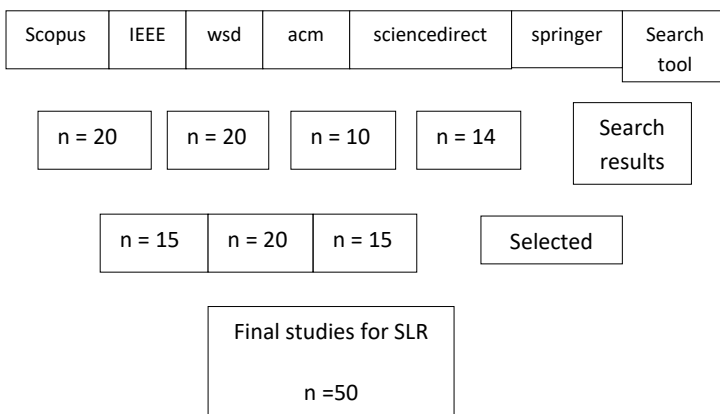
DATA COLLECTION	DATA ANALYSIS	DATA SYNTHESIS
->Electronic database search ->Initial review based on abstract ->Full review of research design and sample ->Full review of text	->Classification of usability open data ->Classification of synchronization, scalability, reliability ->Classification of effects, conditions, advantages and disadvantages ->Classification of platforms ->Classification of user	->Identification of types of utilization, effects, condition and users ->Establishment of OGD utilization framework ->Clarify limitations for future research

Data collection database

The literature was gathered using the following high-quality scientific databases.

- www.scopus.com Scopus
- www.ieee.org – Advanced Technology and Humanity
- www.apps.webofknowledge.com – Web of Science
- www.dl.acm.org – Association for Computing Machinery database
- www.sciencedirect.com – ScienceDirect
- www.springer.com - Springer

Table 2. Selected primary studies



The study search was carried out between September 1, 2020, and August 31, 2023. "Open data usability," "open data synchronization," "open data scalability," "open data reliability," "open data advantages," and "open data disadvantages" are significant phrases for data collecting.

2.1. Selection Criteria

In order to ensure updated and relevant article selection, the following inclusion criteria were applied:

- Articles written between September 1, 2022, and August 31, 2023.
- Articles written in the English language.
- Articles using at least one of the following key terms: "open data usability," "open data synchronization," "open data scalability," "open data reliability," "open data advantages," and "open data disadvantages."

The research divides the usage of accessible government data into groups based on their goals and outcomes. Usage can serve to examine, analyze, and utilize open government data for specific purposes, such as scientific research or the development of concrete practices. The effects of usage are intertwined with the potential consequences of using this data and focus on the impact it can have on society, the economy, or government improvement. Conditions relate to the broader environment and context in which open government data is used, including technical, social, and political aspects of public data usability. Users, on the other hand, are individuals or groups who use this data to achieve their specific objectives and gain value from it, possibly through products, benefits, or practices. In order to support the process of analysis and synthesis of data, the study has created a structured form that gathers important information for each identified study. This information includes details about the study's publication, the methodology used, research questions, abstract content, achieved results, the way open government data is used, identified effects, conditions under which the data is used, the user groups involved, the research field, and the study's location. In the final stage of analysis, the data is synthesized to identify key trends and elements that make up each factor of open government data usage. To better comprehend the usage of open government data, this procedure organizes and categorizes aspects. The data use findings and classifications are part of a larger framework that includes the relationships between components and how they impact one another. Thus, the study aims to provide a clear and structured overview of open government data usage.

3. Results

This study examines 50 recent academic papers (2021-2023) to investigate the most recent research in Open Data, highlighting the dynamic nature of AI.

The findings will be provided according to important study areas such as usability, standardization, scalability, and others.

### **3.1. Descriptive Analysis**

This study includes 50 scholarly works from the period between 2021 and 2023, with the aim of identifying the latest scientific research in the field of Open Data. This approach is taken because AI is undergoing remarkable changes within a short period, and relying on opinions from previous years may not be appropriate. Structurally, we will present the findings according to the hierarchy of the research conducted, starting with open data usability, data standardization, synchronization, scalability, reliability, advantages, and disadvantages. In other words, the categorization of the reviewed works is based on the keywords of this study.

### **3.2. Accessibility and User Experience of Open Data**

The following piece looks at various aspects of open data, highlighting its benefits and limits, as well as providing remedies and future prospects for its use. It emphasizes the importance of open data in terms of data access, economic value, policymaking, and knowledge expansion, while also acknowledging drawbacks such as privacy concerns and data synchronization issues. It also looks at the usage of open data in the public financial sector, as well as the advantages of greater public policy engagement and anti-corruption measures. Finally, the book emphasizes the relevance of open data in improving global government services and community engagement.

Open data, as outlined in research [16], offers several advantages, including enhancing data access, economic value, policy-making, and knowledge advancement. However, drawbacks like privacy concerns, information source limitations, and data synchronization issues exist. Despite these challenges, open data holds substantial potential for economic development, improved governance, and democratic accountability. It is vital to address these challenges to harness its benefits effectively. In another study [39], the SODAS platform was introduced as a solution for efficient open data sharing. The platform, based on CKAN and DCATv2, addresses five key issues and presents three core strategies to tackle them. It comprises four components defined in the second stage, along with nine functional blocks implemented as an open-source resource. SODAS has been successfully applied in various real fields, establishing its efficiency as a real-time, high-speed, large-volume Internet data center. It is a key step toward using unconstrained data by leveraging extended CKAN

and DataMap to serve as a comprehensive platform for real-time Internet data storage, administration, analysis, and dissemination in fields such as healthcare, energy, and smart cities [40].

In addition, there are intentions to identify general concerns with open data platforms in the future and provide an updated version of SODAS as a solution to these issues. Furthermore, because SODAS has a modular structure, the use of high-speed communication technologies such as Remote Direct Memory Access (RDMA) will be considered to improve overall performance. Later, user performance on the upgraded SODAS will be analyzed, as well as its efficacy in real-world services. The use of open government data evaluation using a theoretical lens dubbed "dynamic capability" is the study's main topic. To do so, they employ a framework known as the "Open Data Roadmap" and undertake a particular examination in Denton, Texas, a medium-sized city in the United States. They also include data from two additional public open data models, as well as an appraisal by municipal residents. They conclude that the ODR framework has the potential to assist governments and organizations in increasing community participation, introducing creative new services, and presenting open data as a component of their competitiveness. This study appears to provide useful insights into how open data might help cities and organizations throughout the world enhance services and community participation [41].

Whereas open data in the public financial sector is characterized as publicly available, it is viewed as a strong resource related to other types of data presentation, such as linked data and big data. There should be at least one software solution for working with open data that is completely free to reuse, utilize, and distribute. Most European governments, as well as others across the world, that have launched open data programs distinguish between more than 10 open data groupings. The benefits of open data in the public financial sector are substantial, but openness and anti-corruption initiatives are critical [42]. It is beneficial to make income and spending data public in order to stimulate active engagement in public policies and negotiation processes, as well as to achieve a high degree of openness and anti-corruption efforts through different open data programs [43]. The usage of open data in this study investigates the benefits and problems they bring for the Saudi government. The Saudi government has launched an open data site, and the research recommends an enlarged framework that employs linked open data technologies to construct the "Saudi Linked Open Data Cloud." This technique gathered and connected data from numerous and disparate sources [44].

The usage of open data open government data is a subject of study aiming at increasing support for the use of public government data. Based on the synthesis of the evaluated works, this essay makes some recommendations for future academics interested in open data visualization. In general, the synthesis indicates that the quality of evidence in the examined research might be improved. Most studies discuss scenarios and prototypes without systematically analyzing their value to open data users [45]. Even when usability assessments were conducted, research techniques were frequently not presented in a form that could be repeated or applied to different situations. Concrete recommendations for future research include:

- Using qualitative research and needs assessments to guide the development of future open data visualization tools.
- Incorporating usability testing at various phases of development while designing overarching structures and specialized open government data visualizations.
- Enhancing usability evaluations by testing visualizations with a varied set of open data consumers.
- Investigating the use of innovative visualization approaches such as data storytelling and privacy-preserving representations.

These researches have policy and practice consequences. To begin, government agencies may improve their open data visualizations by working with university research institutes to perform user surveys and by incorporating open government data visualization creation into ordinary work practices rather than special initiatives. Furthermore, government agencies may create more extensive open data policies that prioritize citizen interaction and the usage of open government data platforms with integrated visualization and analytical tools. Furthermore, given the results that data quality was a prevalent problem in developing effective open government data visualizations, data publishers may increase the quality of open data by releasing more comprehensive data in structured forms, including data vocabularies. Finally, publishers should improve the quality of open data visualizations by engaging users with data through strategies such as storytelling and collaborating with community organizations to learn about specific citizen concerns and how to involve them in the creation and improvement of open government data visualizations. Using these techniques can improve the usability of publicly available data for a wide range of consumers [45].

### ***3.3. Enhancing Open Data Utility***

These studies focus on the use and standardization of open government data (OGD) and open information (OI) in e-government. The recommendations include creating a public information strategy, establishing awareness campaigns, and enhancing collaboration between public organizations and academia.

Authors developed 16 dashboard design principles in the context of Open Government Data (OGD) and then implemented these principles to construct the "Namur (Belgium) Budget Dashboard (NBDash)." This was done in order to evaluate the effectiveness of well-designed dashboards versus individual visualizations using a practical situation. The findings of this 108-person experimental study show that the usage of well-designed dashboards might increase the use of data on the open government site. It has also been shown that choosing significant metrics and employing suitable visualizations structured in a clear presentation are critical variables in dashboard performance [46]. These findings highlight the significance of suitable dashboard design in increasing citizen interaction with accessible government data [47].

Regarding the standardization of open data (OI) in electronic government (e-government) in Tanzania, it aims to achieve several goals and address specific challenges. Standardizing open data is a process that aims to make data accessible, usable, and shareable across all public institutions and the public sector. During this study, it was identified that awareness and understanding of OI in e-government in Tanzania are still relatively low in public organizations. However, the study has revealed a sense of enthusiasm and interest in organizations to implement OI. This is an important step toward standardizing OI practices in these organizations [48]. To fulfill the goals of standardizing OI in e-government, the study proposes several recommendations, including drafting a public information (PI) policy, conducting widespread awareness campaigns for OI in e-government, training and development of an incentivization scheme, strengthening collaboration among public organizations and between the public sector and academia [48].

### ***3.4. Strategies for Enhanced Interoperability and Utility***

The present piece examines the necessity of data synchronization in wireless communications, as well as its function in signal interference management.

It also emphasizes usability difficulties observed on Indian government websites, highlighting the importance of improving accessibility and mobile friendliness. The article recommends that the Indian government provide equitable access for all users and utilize the findings to build rules for useable and accessible e-government websites. The article also presents Kubric, a platform that allows for data synchronization while constructing synthetic datasets for machine learning model training, with the goal of simplifying the process and encouraging developer cooperation.

Data synchronization is an important process in wireless communications, especially when the radio spectrum is crowded with various devices without wires. This spectrum load significantly creates interference among different wireless signals operating on the same frequency, and it is unusual for the separation of such captured signals to become a risky part of communications [49]. According to the findings, the majority of the websites under consideration have low usability standards in terms of HTML and CSS problems. The navigation on these sites is poor, with several broken connections. The accessibility is minimal, and the majority of the pages are inaccessible to persons with disabilities and cannot be browsed using assistive technology such as screen readers. Furthermore, the results show that these pages are not mobile-friendly. In a developing nation like India, where government portals are not yet technically mature, accessibility problems, broken links, and a lack of mobile preparedness can result in a disastrous user experience, forcing consumers to avoid government websites [50]. The government should guarantee that all users, including those with and without impairments, have equal access by reviewing current pages for usability, accessibility, and mobile ready standards. Policy-making bodies, such as the Indian Ministry of Electronics and Information Technology, might utilize this study to develop policies for usable and accessible e-government on future websites. The study has a significant limitation in that the evaluation of websites was done only with automated tools; in the future, a more comprehensive assessment can be conducted using a hybrid of automated tools and expert and user analysis for individuals with disabilities [7].

In this application, data synchronization refers to the process of coordinating and interacting with data from diverse sources in order to build synthetic datasets suitable for machine learning model training. Kubric offers a single platform as well as a standard export data format (SunDS) that allows data to be utilized directly in training operations. This makes it easy to create the synthetic datasets required for computer vision applications. According to the text,

Kubric lowered the engineering work required to create training data and promoted reuse and cooperation among developers from other projects [51]. The author also notes certain restrictions and Kubric's future work. Some Blender and PyBullet capabilities are missing, such as volumetric effects like fog or fire, models of soft materials and clothes, and complex camera effects like depth of field and motion blur. They intend to enhance and consolidate resources from several datasets, including the ABC dataset and Amazon Berkeley Objects, in the future. Data synchronization is essential for building synthetic datasets, and Kubric is described as a tool that can help with this. This platform intends to lower the obstacles to generate synthetic data required by machine learning model developers, as well as to facilitate pipeline and dataset cooperation and sharing [52].

### ***3.5. Scalability Challenges and Solutions in Open Data Ecosystems***

Although scalability is not explicitly discussed, this article examines it in terms of system and technological capability to handle massive amounts of data or increasing use with a distributed user base. It investigates how narrative-based search interfaces might increase open data adoption, particularly among non-technical users, thereby improving data inclusion. The article also investigates scalability difficulties in large-scale internet search engines, highlighting unsolved concerns and ongoing obstacles for commercial search engine companies.

Scalability is a term that can be used to describe the ability of a system or technology to increase its capacity to handle and process large volumes of data or to increase its usage with a large and distributed user base. In the context of the given text, there is no direct mention of scalability. However, some elements of the existing research suggest that this technique may have the potential to increase the usage and inclusion of open data, especially for users with less technical expertise [53]. For example, the study analyzes the prospect of creating narrative-based search interfaces to assist overcome obstacles to open data adoption, particularly for less technical users. This has the potential to increase data inclusion in various aspects. Furthermore, the author emphasizes that data curation is an ongoing activity and can occur at different stages of the data lifecycle, including data collection, initial use, and subsequent decline. This approach supports history-based research and introduces narrative-based data exploration, especially when supported by recommendation systems. This approach has the potential to support all open data users, but especially those with less technical expertise [54].

Current scalability issues in large-scale online search engines are described in Cambazoglu and Baeza-Yates' study [55]. They also discuss several unresolved concerns in this sector. The authors expect that scalability in online search engines will continue to be a concern for some time. They point out that, in contrast to previous studies, the present scalability study is driven largely by the demands of commercial search engine businesses. The authors admit that the lack of large-scale infrastructure and datasets makes scaling research problematic, particularly for university researchers. They do, however, provide some suggestions for people working in this field:

- To discover true scalability difficulties, it is critical to regularly monitor online trends, user bases, and device specifications.
- New strategies aiming primarily at enhancing search quality should be implemented, and their consequences for efficiency should be thoroughly understood.
- Solutions from previously addressed problems in comparable research domains, such as databases, computer networks, and distributed computing, should be modified to address scaling challenges.
- Because the expense of using genuine distributed systems is quite high, evaluating novel solutions may include modeling large-scale systems and using synthetic data.

A different study by J. Cho uses keyword network analysis to identify disparities in the content of accessible public data handled by the central government, local governments, public institutions, and the education office in Korea. The routing network was analyzed by extracting particular keywords from 1,200 examples of free data accessible on Korea's Public Data Portals. For each form of government, subject groupings were developed, and their usefulness was assessed using download statistics. Several topic groups were established for public institutions to provide specialized information on national concerns such as Health Care and Real Estate, while 15 subject groups were established for the central government to provide national administrative information such as Crime, Police, and Security. Local governments and education offices, respectively, were allotted 16 and 11 topic groupings of data concentrating on regional life, such as Factories and Local Production, Resident Registration, and Lifelong Education. The usability of specialized information at the national level was greater in public and central governments than in regional governments. It was also revealed that theme groupings such as Health Care, Real Estate, and Crime received a lot of attention.

There is also a considerable disparity in data utilization due to the availability of popular data with high usage.

### 3.6. *Ensuring the Reliability of Open Data*

The paper that follows underlines the significance of dependability in open data (OGD), stressing its vital role in quality and availability. It addresses how trustworthy open data may lead to better decision-making and the discovery of hidden values, especially in areas such as construction and infrastructure. The document also emphasizes the need of validating and using open data for commercial reasons, recommending that governments create policies to improve access and quality. Furthermore, it investigates the dependability of portals in the GCC area for smart city infrastructure, identifying frequent flaws and offering enhancements such as the usage of sophisticated visualization tools. Open data problems, such as quality evaluation, standards, security, and privacy concerns, are also explored, as are the implementation and utilization challenges of open data in smart city environments.

In the article, reliability is highlighted as a critical attribute linked with open data (OGD) and is related to its quality and availability. The essay highlights that open data must be dependable and usable within the open movement in numerous sectors such as open resources, open content, open access, and open knowledge. The reliability of open data implies that it should be validated and safe, allowing confidence to be formed on it. Reliability is especially important in the building and infrastructure industries [56]. The collection and use of open data may assist in making better decisions by providing knowledge, generating data-driven results, and uncovering hidden values. The capacity to uncover meaningful insights from varied sets of data supplied by government entities and to match data-driven techniques with new technical breakthroughs to boost efficiency, productivity, and creativity [57] is critical to the industry's future success.

Meanwhile, just because there is a lot of free data does not automatically imply that it is accurate, dependable, and valuable. Open data must be validated and useful for commercial reasons, and the government must encourage open public data rules that boost access and quality. This necessitates the commitment of valuable data and the implementation of a value creation mechanism, as well as the development of a data management method and a data usage guide based on dependability and data growth [57]. The reliability of the portal is an effort to assess how effectively this portal can contribute to the transition to smart city infrastructure.



In a research investigation, it was found that although all the portals in the GCC region are suitable for their purposes, they have some common weaknesses, such as the lack of advanced visualization tools and effective mechanisms for involving stakeholders in portal development. To enhance the assistance that portals can provide, it is suggested that they should more frequently use visualization tools and promote public awareness of common tools that can help in using open data [58]. Portal reliability is crucial in empowering citizens and improving their agreements and decisions with open data [59]. The study shows that GCC countries are committed to using open data to develop smart cities. However, there is room for improvement and international collaboration to help improve portals and increase public awareness of open data. Portal reliability is a key aspect of making this transition to a more efficient and rapid smart city infrastructure [59].

In an overview, some of the obstacles and weaknesses observed in the realm of open data are deficiencies in quality evaluation, inadequate standards, security and privacy issues, unclear regulations, and usability challenges [60]. In another study [61], the ways in which open data is implemented and used in the smart city environment are associated with several challenges and deficiencies. One of these challenges is the lack of quality assessment of open data, leading to the risk of using poor-quality data for citizens. Additionally, another challenge is the variation in open data standards, which hampers interoperability and data usage [62]. The issues of incomplete metadata and the establishment of user environments for open data are also identified challenges. These challenges complicate the integration of open data with other technologies and may hinder their use in creating intelligent city services. Open data users, such as application developers, may also have difficulty finding suitable environments and services for using the data in their projects [63].

#### 4. Conclusion

The research that has been presented in this study highlights the growing importance of open government data (OGD) in our information-driven society. OGD has become a crucial tool for transparency advocates, non-governmental organizations, and government entities alike.

The discussion rightly emphasizes the need for standardized definitions and clear licensing of open data to promote accessibility and usability. One of the key takeaways from this discussion is the idea that open data goes beyond mere availability; it must also be usable. The challenges that scientists face in accessing and reusing data, even when it is technically available, underscore the need for improved usability standards. Clear labeling, long-term user-centered design, and thorough documentation are critical components in ensuring that open data fulfills its purpose efficiently. The inclusion of international instruments and charters, such as the Group of Eight (G8) Open Data Charter and the International Open Data Charter, demonstrates the worldwide acknowledgment of open data's potential advantages. These charters act as powerful motivators for governments all around the globe to implement laws that encourage data accessibility and openness.

In conclusion, the thematic analysis of various texts related to open data provides us with an in-depth overview of the advantages and disadvantages of this concept in different contexts. The advantages of open data include transparency, the potential to combat corruption, ease of access, and assistance to citizens worldwide. This technology has international impact and is crucial for the development of smart civilizations. The downsides of open data, on the other hand, include quality and standards issues, security and privacy concerns, usability issues for non-technical users, and difficulty in integrating with other technologies.

The study of open data is an important research field that has deepened our understanding of the significance of transparency and sustainability in our information-driven society. The research conclusions emphasize the need for standardized definitions and clear licenses for open data, highlight their usability as a key element for success, and examine the challenges that users and organizations face. Furthermore, future research can address issues such as the effectiveness of using open data, their security and privacy, their impact on global development, the applications of new technologies, the use of open data to improve governance and citizen-government interaction, and legal changes that affect their availability.

**References:**

- [1]. Schwoerer, K. (2019). Designing Open Government Data Programs For Usability: The Impact Of Usability Attributes On Open Data Use. *Book Of Abstracts*, 81.
- [2]. Shaikh, S. A., Kitagawa, H., Matono, A., Mariam, K., & Kim, K. S. (2022). GeoFlink: an efficient and scalable spatial data stream management system. *IEEE Access*, 10, 24909-24935.
- [3]. Chinthavali, S., Hasan, S. S., Yoginath, S., Xu, H., Nugent, P., Jones, T., ... & Polsky, Y. (2022). An Alternative Timing and Synchronization Approach for Situational Awareness and Predictive Analytics. In *2022 IEEE 23rd International Conference on Information Reuse and Integration for Data Science (IRI)* (pp. 172-177). IEEE.
- [4]. Herrera-Murillo, D. J., Aziz, A., Noguera-Iso, J., & Lopez-Pellicer, F. J. (2022). Analysing user involvement in open government data initiatives. In *International Conference on Theory and Practice of Digital Libraries*, 175-186. Cham: Springer International Publishing.
- [5]. Shi, Y. (2022). Advances in big data analytics. *Adv Big Data Anal.*
- [6]. Gemmi, G., Cerdà-Alabern, L., & Maccari, L. (2022). A realistic open-data-based cost model for wireless backhaul networks in rural areas. In *2022 18th International Conference on Network and Service Management (CNSM)*, 55-63. IEEE.
- [7]. Agrawal, G., Kumar, D., & Singh, M. (2022). Assessing the usability, accessibility, and mobile readiness of e-government websites: a case study in india. *Universal Access in the Information Society*, 21(3), 737-748.
- [8]. Bessie, Y., & Jura, D. (2022). Analysis of Christian Education on Mori Tribe's Rice Harvesting Tradition in North Morowali, Central Sulawesi. In *1st Virtual Workshop on Writing Scientific Article for International Publication Indexed SCOPUS*, 292-297. Sciendo.
- [9]. Syerov, Y., Klynina, T., Sokil, M., Sverdlyk, Z., Shepichak, V., Kravets, R., ... & Bilyk, O. (2023). Information Management Processes: Review of the State Open Data Web Portal. *Developments in Information and Knowledge Management Systems for Business Applications*, 6, 189-210.
- [10]. Qonita, F., Budiman, M. F., Sari, V. M., & Limantara, N. (2023). Analysis of User Experience on The Government Application of Indonesian Higher Education Institutional Information Systems Using Usability Method. In *2023 4th International Conference on Innovative Trends in Information Technology (ICITIT)*, 1-6. IEEE.
- [11]. Alogaiel, N. F., & Alrwais, O. A. (2023). An Assessment of The Quality of Open Government Data in Saudi Arabia. *IEEE Access*.
- [12]. El Khatib, M. M., Alzoubi, H. M., Ahmed, G., Kazim, H. H., Al Falasi, S. A., Mohammed, F., & Al Mulla, M. (2022). *Digital transformation and SMART-the analytics factor*. *2022 International Conference on Business Analytics for Technology and Security (ICBATS)*, 1-11.
- [13]. Samakit, T., Anutariya, C., & Buranarach, M. (2023). QUALYST: Data Quality Assessment System for Thailand Open Government Data. *2023 20th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 196-201.
- [14]. Tshering, Y., & Anutariya, C. (2022). Enabling Semantic Interoperability in Bhutan's E-Government: An Ontology-based Framework. *2022 19th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 1-6.
- [15]. Bouchelouche, K., Ghomari, A. R., & Zemmouchi-Ghomari, L. (2022). Enhanced analysis of Open Government Data: Proposed metrics for improving data quality assessment. *2022 5th International Symposium on Informatics and its Applications (ISIA)*, 1-6.
- [16]. Hyseni, Besart; Bexheti, Lejla Abazi. (2022). Perspective of open data application in Kosovo, challenges, and advantages. *2022 11th Mediterranean Conference on Embedded Computing (MECO)*, 1-8.
- [17]. Akgül, Y. (2022). Evaluating the performance of websites from a public value, usability, and readability perspectives: a review of Turkish national government websites. *Universal Access in the Information Society*, 1-16.
- [18]. Hyseni, B., & Bexheti, L. A. (2023). The Impact of Open Data Standardization on the Successful Management of e-Government. *2023 12th Mediterranean Conference on Embedded Computing (MECO)*, 1-8.
- [19]. Hou, J., & Knaze, J. (2022). Exchange rate regimes and business cycle synchronization. *Open Economies Review*, 33, 523-564.
- [20]. Leng, Z., Li, G., Liu, C., Cubuk, E. D., Sun, P., He, T., . . . Tan, M. (2023). LiDAR augment: Searching for scalable 3D LiDAR data augmentations. *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 7039-7045.
- [21]. Ashahril, S. M., Isa, A. M., & Anwar, N. (2022). Challenges in Open Government Data Implementation: A Systematic Literature Review. *2022 International Conference on Artificial Intelligence of Things (ICAIoT)*, 1-6.
- [22]. Buryakov, D., Kovacs, M., Kryssanov, V., & Serdült, U. (2022). Using open government data to facilitate the design of voting advice applications. *International Conference on Electronic Participation*, 19-34.
- [23]. Ma, H., Rong, Y., & Huang, J. (2022). Graph Neural Networks: Scalability. *Graph Neural Networks: Foundations, Frontiers, and Applications*, 99-119.
- [24]. Vujic, M., Dedic, L., Furjan, M. T., & Pihir, I. (2022). The benefits of open data in urban traffic network. *5th EAI International Conference on Management of Manufacturing Systems*, 267-282.
- [25]. Schneider, M., & Körner, M. (2022). Harnessing administrative data inventories to create a reliable transnational reference database for crop type monitoring. *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium*, 5385-5388.

- [26]. Debauche, O., Nkamla Penka, J. B., Hani, M., Guttadauria, A., Ait Abdelouahid, R., Gasmi, K., . . . others. (2023). Towards a Unified Architecture Powering Scalable Learning Models with IoT Data Streams, Blockchain, and Open Data. *Information, 14*, 345.
- [27]. Ilhami, R., Endah Marlovia, E. M., & Achmad, W. (2022). Smart government policy implementation for smart city concept realization. *International Journal of Health Sciences Scopus coverage years: from 2021 to Present*, 8379–8389.
- [28]. Cui, L., Xie, G., Qu, Y., Gao, L., & Yang, Y. (2018). Security and Privacy in Smart Cities: Challenges and Opportunities. *IEEE Access, 6*, 46134–46145. Doi:10.1109/access.2018.2853985
- [29]. Soldan, M., Pardo, A., Alcázar, J. L., Caba, F., Zhao, C., Giancola, S., & Ghanem, B. (2022). Mad: A scalable dataset for language grounding in videos from movie audio descriptions. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5026–5035.
- [30]. Kijanovic, S., & Komosar, A. (2023). Research on Open Data Practice Within Education. *2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH)*, 1–6.
- [31]. Gao, T., Yu, T., & Yan, J. (2022). Research on Performance Evaluation of Government Data Open Platform. *2022 8th Annual International Conference on Network and Information Systems for Computers (ICNISC)*, 943–948.
- [32]. Ghiasi, G., Gu, X., Cui, Y., & Lin, T.-Y. (2022). Scaling open-vocabulary image segmentation with image-level labels. *European Conference on Computer Vision*, 540–557.
- [33]. Mohammadabbasi, M., Jafari, S. M., Almasizadeh, S., Rabiee, S., & Safari, M. (2022). Identifying the Applications and Consequences of Using the Internet of Things (IoT) in the Smart Government. *2022 International Conference on Artificial Intelligence of Things (ICAIoT)*, 1–5.
- [34]. De Capitani di Vimercati, S., Facchinetti, D., Foresti, S., Oldani, G., Paraboschi, S., Rossi, M., & Samarati, P. (2021). Scalable Distributed Data Anonymization. *2021 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*. IEEE. Doi:10.1109/percomworkshops51409.2021.9431063
- [35]. Burian, J., & Kočvarová, B. (2022). Open Data and Its Role in Geoparticipation. In *Geoparticipatory Spatial Tools* 59–86. Springer.
- [36]. Duboc, L., Bahsoon, R., Alrebeish, F., Mera-Gómez, C., Nallur, V., Kazman, R., . . . Buyya, R. (2022). Systematic scalability modeling of QoS-aware dynamic service composition. *ACM Transactions on Autonomous and Adaptive Systems, 16*, 1–39.
- [37]. Srivastava, S., & Singh, A. K. (2023). Fraud detection in the distributed graph database. *Cluster Computing, 26*, 515–537.
- [38]. Huang, H., Kong, W., Zhou, S., Zheng, Z., & Guo, S. (2023). Blockchain Scalability Fundamentals. In *Blockchain Scalability*, 1–49. Springer.
- [39]. Won, H., Han, J., Gil, M.-S., & Moon, Y.-S. (2023). SODAS: Smart Open Data as a Service for Improving Interconnectivity and Data Usability. *Electronics, 12*, 1237.
- [40]. Kassen, M. (2022). *Open data governance and its actors*. Springer.
- [41]. Habib, A., Prybutok, V. R., & Philpot, D. (2022). Assessing and building municipal open data capability. *Information Systems and e-Business Management, 20*, 1–25.
- [42]. Darari, F. (2022). Open Government Data as a Digital Infrastructure for Public Information Services: A Case Study of the City of Denpasar. *2022 International Conference on Electrical Engineering and Informatics (ICELTICs)*, 136–142.
- [43]. Komosar, A., & Kijanovic, S. (2023). Research on the Application of Open Data in the Public Finance Sector. *2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH)*, 1–6.
- [44]. Alsukhayri, A. M., Aslam, M. A., Khan, I. H., Abbasi, R. A., & Babour, A. (2022). Toward Building a Linked Open Data Cloud to Predict and Regulate Social Relations in the Saudi Society. *IEEE Access, 10*, 50548–50561.
- [45]. Ansari, B., Barati, M., & Martin, E. G. (2022). Enhancing the usability and usefulness of open government data: A comprehensive review of the state of open government data visualization research. *Government Information Quarterly, 39*, 101657.
- [46]. Rojek, K., & Wyrzykowski, R. (2022). Performance and scalability analysis of AI-accelerated CFD simulations across various computing platforms. *European Conference on Parallel Processing, 223–234*.
- [47]. Chokki, A. P., Simonofski, A., Frénay, B., & Vanderose, B. (2022). Engaging citizens with open government data: The value of dashboards compared to individual visualizations. *Digital Government: Research and Practice, 3*, 1–20.
- [48]. Mkude, C. (2023). Open innovation in e-government: Exploring its practices in Tanzania. *Digital Government: Research and Practice, 4*, 1–13.
- [49]. Lancho, A., Weiss, A., Lee, G. C., Tang, J., Bu, Y., Polyanskiy, Y., & Wornell, G. W. (2022). Data-driven blind synchronization and interference rejection for digital communication signals. *GLOBECOM 2022-2022 IEEE Global Communications Conference*, 2296–2302.
- [50]. Wu, D., Xu, H., Yongyi, W., & Zhu, H. (2022). Quality of government health data in COVID-19: definition and testing of an open government health data quality evaluation framework. *Library Hi Tech, 40*, 516–534.
- [51]. Moll, P., Patil, V., Wang, L., & Zhang, L. (2022). SoK: The evolution of distributed dataset synchronization solutions in NDN. *Proceedings of the 9th ACM Conference on Information-Centric Networking*, 33–44.
- [52]. Greff, K., Belletti, F., Beyer, L., Doersch, C., Du, Y., Duckworth, D., . . . others. (2022). Kubric: A scalable dataset generator. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 3749–3761.

- [53]. Fang, Y., Zhao, X., Huang, P., Xiao, W., & de Rijke, M. (2022). Scalable representation learning for dynamic heterogeneous information networks via metagraphs. *ACM Transactions on Information Systems (TOIS)*, 40, 1–27.
- [54]. Wolff, A., Tylosky, N., & Hasan, T. (2022). Open data inclusion through narrative approaches. *Proceedings of the 2022 ACM/IEEE 44th International Conference on Software Engineering: Software Engineering in Society*, 125–129.
- [55]. Cambazoglu, B. B., & Baeza-Yates, R. (2011). Scalability challenges in web search engines. In *Advanced topics in information retrieval*, 27–50. Springer.
- [56]. Tu, Y., Urata, M., Endo, M., & Yasuda, T. (2022). Proposals and practices for open image data processing and promotion in local governments. *2022 IEEE 11th Global Conference on Consumer Electronics (GCCE)*. IEEE. Doi:10.1109/gcce56475.2022.10014180
- [57]. Park, J. J., Kim, E., & Jun, H. J. (2022). Development of an Open Government Data (OGD) Evaluation Framework for BIM. *Buildings*, 12, 490. Doi:10.3390/buildings12040490
- [58]. Adje, K. D., Letaifa, A. B., Haddad, M., & Habachi, O. (2023). Smart City Based on Open Data: A Survey. *IEEE Access*, 11, 56726–56748. Doi:10.1109/access.2023.3283436
- [59]. Kozievitch, N. P., Fonseca, K. V., Rosa, M. d., Berardi, R., Gutierrez, M. B., Belizario, M., . . . Liu, E. (2022). Assessment of Open Data Portals: A Brazilian case study. *2022 IEEE International Smart Cities Conference (ISC2)*. IEEE. Doi:10.1109/isc255366.2022.9921914
- [60]. Rao, S. U., & Lakshmanan, L. (2023). Securing communicating networks in the age of big data: an advanced detection system for cyber attacks. *Optical and Quantum Electronics*, 56. Doi:10.1007/s11082-023-05715-7
- [61]. Conde, J., Munoz-Arcenales, A., Choque, J., Huecas, G., & Alonso, A. (2022). Overcoming the Barriers of Using Linked Open Data in Smart City Applications. *Computer*, 55, 109–118. Doi:10.1109/mc.2022.3206144
- [62]. Yusuf, M., Sophan, M. K., Satoto, B. D., Anamisa, D. R., Alamsyah, N., & Syarif, I. (2022). Development of Village Open Data using DevOps Approach. *2022 6th International Conference on Informatics and Computational Sciences (ICICoS)*, 36–41.
- [63]. Anthopoulos, L. G., & Janssen, M. (2022). Business model canvas for big and open linked data in smart and circular cities: Findings from Europe. *Computer*, 55, 119–133.