Journal of Information Systems and Informatics
Vol. 6, No. 2, June 2024 e-ISSN: 2656-4882 p-ISSN: 2656-5935
DOI: 10.51519/journalisi.v6i2.862 Published By DRPM-UBD

Data Quality Analysis on Open Government Data Portals: A Qualitative Study Using ISO/IEC 25012:2008 Standards

Emigawaty[1], Andri Syafrianto[2]*

[1]Informatics Study Program, Faculty of Computer Science, Universitas Amikom Yogyakarta, Yogyakarta, Indonesia
[2]Informatics Study Program, STMIK EL RAHMA Yogyakarta, Yogyakarta, Indonesia
Email: [1]emigawaty@amikom.ac.id, [2]andrisyafrianto@gmail.com
Corresponding author email: [2]andrisyafrianto@gmail.com

Abstract

This study evaluates the data quality on Open Government Data (OGD) portals using the ISO/IEC 25012:2008 standard, which categorizes data quality into two main groups: inherent data quality and system-dependent data quality. This standard encompasses dimensions such as accuracy, completeness, consistency, and relevance. Using a qualitative approach, interviews were conducted with data providers and users from the government, industry, and academia. The findings indicate that while some datasets are adequate, there are issues with semantic consistency, completeness, timeliness, and currency of the data. These findings highlight the importance of strict and continuous application of data quality standards in OGD management. Recommendations for improvement include training for data managers and enhancing validation mechanisms before data is published. This study supports government efforts to improve transparency and accountability by providing high-quality data that can be reliably used by various stakeholders.

Keywords: Data Quality, OGD Portal, Data Completeness, ISO/IEC 25012:2008

1. INTRODUCTION

Data quality in the context of Open Government Data (OGD) is a key element that significantly determines the extent to which the provided data can be effectively used by the public, academics, and the private sector or industry [1]. High-quality data must meet several criteria, such as accuracy, relevancy, timeliness, and consistency [2]. When government-provided data is inaccurate or outdated, data users will find it challenging to conduct reliable analyses, ultimately diminishing trust in the OGD initiative itself [3]. Therefore, it is crucial for the government to ensure that data shared through OGD portals undergoes rigorous validation and treatment processes [4].

Furthermore, data quality has direct relevance to government transparency and accountability. Data that is easily accessible, accurately interpreted, and

contextually relevant allows the public or data users to effectively monitor public policies and government performance [5]. Thus, in the context of OGD initiatives, data quality is not only about providing data or information on data portals but also about strengthening the relationship between the government and the public, encouraging public participation, and enhancing democracy through greater openness and engagement [5], [6].

The urgency of maintaining data quality on OGD portals is exceptionally high, as poor-quality data can lead to various detrimental risks [7]. One such risk is the spread of misinformation, which can result in incorrect decision-making by data users [8]. For instance, if economic data published on a portal is inaccurate or outdated, analysts and policymakers may make erroneous decisions regarding fiscal policy or budget allocation, ultimately negatively impacting the overall economy [8]. The consequence could be a loss of public trust in the government, particularly if the public frequently encounters unreliable data on OGD portals [9], [10].

A concrete example of this risk is the use of inaccurate public health data in response to a health crisis, such as the recent COVID-19 pandemic. Inconsistent data on disease spread or the availability of healthcare facilities could lead to ineffective resource allocation, exacerbating the crisis. Moreover, reliance on poor-quality data can also stifle innovation in the private and academic sectors, as research and development based on erroneous data will not yield effective or sustainable solutions. Therefore, maintaining data quality on OGD portals is critical to avoid these risks and ensure that data can be optimally utilized for the public good [11], [12].

ISO/IEC 25012:2008 serves as an international standard that provides a framework for assessing and improving data quality, including the data available on OGD portals [13], [14]. This standard outlines at least 12 data quality characteristics divided into two main categories: inherent data quality and system-dependent data quality [14], [15]. The inherent data quality category includes aspects such as accuracy, completeness, and consistency, focusing on the quality of the data itself, independent of the system that manages it. Meanwhile, the system-dependent category covers characteristics such as timeliness, accessibility, and usability, which are more related to how the data is presented and managed within the OGD portal system.

By adopting the ISO/IEC 25012:2008 standard, governments can more systematically identify and address data quality issues on OGD portals. If data on the portal is frequently inconsistent or incomplete, this standard can guide specific steps to improve that quality, such as through stricter data validation processes or enhanced data maintenance mechanisms. Implementing this standard also enables

continuous monitoring of data quality, ensuring that OGD portals consistently provide reliable and useful data for users, while simultaneously strengthening government transparency and accountability.

The objective of this study is to evaluate the data quality on OGD portals using the ISO/IEC 25012:2008 standard as the assessment reference, focusing on identifying the strengths and weaknesses of the data in meeting the quality characteristics established by the standard. The results of this evaluation are expected to provide concrete recommendations for improving data quality, thereby enabling OGD portals to more effectively support transparency, accountability, and enhanced public participation.

This study employs a semi-structured qualitative approach by conducting interviews with various stakeholders who serve as data providers and users from the government, academia, and industry/business sectors. The goal is to gain in-depth insights into their experiences and challenges in using data from OGD portals. The findings from these interviews will be analyzed to identify the most significant data quality issues and to develop relevant and applicable recommendations within the context of open government data management. The research questions formulated are: (1) What are the perspectives of data providers and users regarding the data quality on OGD portals? (2) How do the perspectives of data providers and users correlate with the data quality standards based on ISO/IEC 25012:2008? With these research questions, this study aims to contribute not only academically but also practically in terms of improving data quality on the side of government data providers.

2. METHODS

2.1. Research Flow

This study follows five sub-stages of the process as a guide for conducting the research, as illustrated in Figure 1. Generally, this research consists of six stages: (1) Identifying the Research Problem, (2) Review of Literature, (3) Designing the Qualitative Study, (4) Data Collection, (5) Data Analysis, and (6) Interpretation and Findings. The purpose of these six steps in qualitative research is to develop a deep understanding of a phenomenon through a systematic and structured approach. The process begins with identifying a specific research problem, aiming to find relevant knowledge gaps. The literature review is conducted to connect the research to a broader scientific context and to establish a strong theoretical foundation. Following this, the study is designed by selecting appropriate methods and participants, as well as setting up data collection procedures to ensure that the data obtained is relevant and of high quality.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

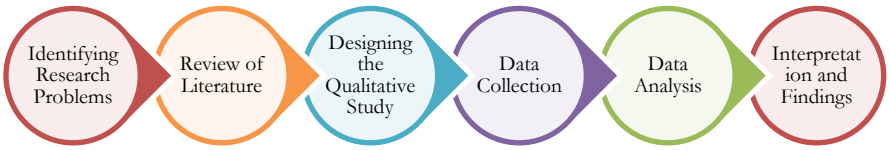

Figure 1. Research Flows and Agendas

1) **Identifying Research Problems**
   In this step, the researcher determines the problem or issue to be studied. The researcher identifies research questions that are relevant and significant to explore regarding the data quality in the area of open data portals. The main focus is to understand the context of the problem and to find gaps in knowledge that need to be addressed through the research. The outcome of this step is a clear and focused formulation of the research problem.
2) **Review of Literature**
   The researcher conducts a literature review to understand the existing research related to the topic under study. This step helps the researcher identify theories, concepts, and previous findings that can serve as a theoretical foundation for the research. The literature review also helps clarify the position of the current research, particularly in data quality at the OGD portals within a broader context and identify gaps or shortcomings in previous studies.
3) **Designing the Qualitative Study**
   At this step, the researcher designs the qualitative study to be conducted. This includes selecting the research method (e.g., case study: Open Data Portals), determining the sample or participants to be studied, and choosing appropriate data collection methods such as interviews, observations, or document analysis. The researcher also considers ethical aspects of the research, including how to protect participants' privacy and ensure their informed consent.
4) **Data Collection**
   The researcher collects data from participants or invited experts in this study using the methods designed earlier. This process involves direct interaction with participants through interview sections to gather rich and in-depth information. The data collected is usually descriptive and aims to deeply understand the participants' experiences or perspectives with regards to the use and provide the data for OGD portals.
5) **Data Analysis**
   After collecting the data, the researcher analyzes it by organizing, categorizing, and identifying emerging themes or patterns. This process is

often iterative, with the researcher going back and forth between the data and analysis to refine their understanding. The goal is to uncover deep meanings and relevant insights from the collected data (interview sections).

6) **Interpretation and Findings**
In this final step, the researcher interprets the analyzed findings in the context of the research questions and existing literature. The researcher seeks to understand the implications of the findings and how they contribute to broader knowledge. The results of this interpretation are then compiled into a report or presentation that outlines the main discoveries of the research and offers new insights or solutions to the problem identified at the beginning of the study.

2.2. Data Collection

This research employs a semi-structured interview technique, designed to combine open-ended questions with a predefined framework to guide the discussion. The interviews were conducted from September to October 2023, involving a total of 23 respondents selected based on their expertise and experience in using data from OGD portals. The respondents consisted of 5 government officials from the district/city and provincial levels, 10 academics from various universities in Indonesia, and 8 representatives from industry and businesses, covering small, medium, and large-scale enterprises. The sample of OGD portals represents several government data portals from regional, city, and district levels in the Sumatra and Java islands.

The semi-structured interview mechanism began with a series of key questions designed to identify the respondents' perceptions and experiences related to data quality on OGD portals, including aspects such as accuracy, consistency, and timeliness of the data. Additionally, the interviews provided space for respondents to express additional views and experiences that might not have been covered by the initial questions, allowing for deeper and richer information gathering. Each interview lasted 45 to 60 minutes and was recorded with the respondents' consent, followed by transcription for further analysis. The data obtained from these interviews will be analyzed using thematic analysis to identify key patterns and themes related to data quality on OGD portals, which will then be compared with the ISO/IEC 25012:2008 standard as a benchmark for evaluation.

2.3. Interview Instrument Design

Table 1 below presents the interview instrument design for evaluating data quality on OGD portals. The interview instrument is classified based on the ISO/IEC 25012:2008 standard, specifically into inherent data quality and system-dependent

data quality. Each parameter consists of two main questions. In the inherent data quality section, there are six parameters: accuracy, consistency, completeness, precision, syntactic consistency, and semantic consistency. These parameters focus on the characteristics of the data quality itself, assessing the extent to which the data reflects actual conditions, consistency across datasets, and the correctness of format and terminology usage.

Meanwhile, in the system-dependent data quality section, there are six evaluation parameters: timeliness, accessibility, representation, security, relevancy, and punctuality. These parameters evaluate aspects related to how data is presented and managed within the OGD portal system, including the frequency of data updates, ease of access and usability, and the level of data protection. This instrument aims to provide comprehensive insights into various aspects of data quality, helping to identify the strengths and weaknesses of OGD portals and offering specific and applicable recommendations for improvement.

Table 1. Interview Instrument Design (ISO/IEC 20512:2008)

ID	Classification	Parameter	Interview Questions
1	Inherent Data Quality	Accuracy	1. How do you assess the accuracy of the data available on the OGD portal? 2. Have you ever encountered data that does not reflect the actual conditions or incorrect data?
2		Consistency	1. Have you ever encountered inconsistent data between different datasets on the OGD portal? 2. How do you assess the consistency of data within the same dataset on the OGD portal?
3		Completeness	1. To what extent is the data provided on the OGD portal complete according to your needs? 2. Is there any important information that you consider missing or not included in the dataset you use?

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

ID	Classification	Parameter	Interview Questions
4		Precision	1. How do you assess the level of precision of the data available on the OGD portal? 2. Is the data detailed enough for the analysis or decisions you need to make?
5		Syntactic Consistency	1. Have you ever encountered errors in the format or data entry on the OGD portal? 2. How do you assess the uniformity of data formatting across the datasets?
6		Semantic Consistency	1. Are the terms and concepts used in the data on the OGD portal consistent across all datasets? 2. Is the use of terms and concepts consistent with the common understanding in your field?
7	System-Dependent Data Quality	Timeliness	1. What is your opinion on the timeliness of the data available on the OGD portal? 2. Do you feel that the data on the OGD portal is updated regularly and according to your needs?
8		Accessibility	1. What has been your experience in accessing data on the OGD portal? 2. Is the data available for download in a format that meets your needs?
9		Representation	1. Is the format and structure of the data on the OGD portal easy to understand? 2. Do the metadata and documentation provided help you in understanding the data?
10		Security	1. How do you perceive the data security efforts on the OGD portal?

ID	Classification	Parameter	Interview Questions
			2. Do you feel that the data you access is secure from unauthorized access or manipulation?
11		Relevancy	1. Is the data available on the OGD portal relevant to your needs? 2. Do you feel that there is important data missing from the OGD portal?
12		Availability	1. Is the data on the OGD portal always available when you need it? 2. Have you ever experienced access disruptions to the data on the OGD portal?

3. RESULTS AND DISCUSSION

3.1 Results of the Study

The following is a summary of the interview results with respondents, considering the groups with positive (pro) and negative (con) responses regarding the condition of the OGD portal based on their knowledge and experience.

Table 2. Interview Results Based on Inherent Data Quality Parameters

Parameter	Group	Positive	Negative
Accuracy	Government	1. Data is usually accurate and supports effective policymaking. 2. Rare data errors do not significantly affect overall decision-making.	1. There are significant errors in the data that could impact policy decisions. 2. Frequent data errors lead to less effective policy decisions.
	Academics	1. The data is generally accurate, supporting valid research outcomes.	1. Minor errors in the data often affect research outcomes and analysis.

Parameter	Group	Positive	Negative
		2. The data accuracy enables consistent and reliable replication of studies.	2. The need to verify data from other sources reduces research efficiency.
	Industry / Business Practitioners	1. Accurate data facilitates effective business analysis. 2. The quality of data accuracy supports confidence in reports and strategic decisions.	1. Data inaccuracies occasionally disrupt the analysis process and business decisions. 2. Data inaccuracies can lead to incorrect business decisions.
Consistency	Government	1. Consistency of data across datasets facilitates integrated decision-making. 2. Data is consistent in format and information, supporting accurate analysis.	1. Inconsistencies between datasets often hinder effective data integration. 2. Differences in data formats can lead to difficulties in processing and analysis.
	Academics	1. Data consistency supports more stable and reliable research outcomes. 2. Uniformity of data facilitates the process of analysis and interpretation of research results.	1. Minor inconsistencies in the data can affect the validity of research results. 2. Inconsistent data formatting across some datasets can disrupt data analysis.

Parameter	Group	Positive	Negative
	Industry / Business Practitioners	1. Consistent data supports effective analysis and comparison across datasets. 2. Consistency across datasets simplifies the preparation of reports and business strategies.	1. Data inconsistencies hinder integration and the business analysis process. 2. Data inconsistency often leads to errors in business reports and analysis.
Completeness	Government	1. The data is complete, covering all the information needed for public policy. 2. Data completeness supports the planning and implementation of government programs.	1. Some datasets are incomplete, making comprehensive policymaking difficult. 2. The lack of information in some datasets can reduce the effectiveness of policies.
	Academics	1. The data is complete and includes various variables necessary for research. 2. Complete data enables thorough and comprehensive analysis.	1. There is a lack of important details that can reduce the depth of analysis. 2. Incompleteness of data in some areas necessitates searching for additional data sources.
	Industry / Business Practitioners	1. Complete data meets the needs for in-depth business analysis. 2. Data completeness	1. The lack of critical data can hinder business decisions and operational processes.

Parameter	Group	Positive	Negative
		supports informed strategic decision-making.	2. Incomplete data often requires additional data collection, slowing down the analysis.
Precision	Government	1. High-precision data enables more accurate and effective policies. 2. Precise details support decisions based on solid data.	1. Lack of precision in some datasets can result in less accurate policies. 2. Imprecise data can lead to inaccuracies in policy decisions
	Academics	1. Precise data supports more detailed analysis and valid outcomes. 2. Data precision allows for study replication and reliability of research results.	1. Lack of precision in some data can significantly affect research outcomes. 2. Imprecise data can lead to misinterpretation and errors in analysis.
	Industry / Business Practitioners	1. High-precision data supports accurate business analysis and better decision-making. 2. Data precision helps in creating more accurate and reliable reports.	1. Data imprecision often leads to errors in analysis and decision-making. 2. Imprecise data can disrupt business strategy and planning.
Syntactic Consistency	Government	1. Data format is consistent, facilitating data usage and integration.	1. Data formatting errors often disrupt the process of

Parameter	Group	Positive	Negative
		2. Format consistency supports more efficient data processing.	analysis and data usage. 2. Inconsistent data formats often require additional processing.
	Academics	1. Data format is generally consistent, supporting accurate analysis. 2. Syntactic consistency facilitates data processing in research.	1. Inconsistent data formats across some datasets affect the analysis and interpretation process. 2. Some data with inconsistent formats complicate integration and analysis.
	Industry / Business Practitioners	1. Data format is consistent across most datasets, supporting business analysis. 2. Format consistency facilitates data usage across various systems.	1. Inconsistent data formats hinder efficient integration and analysis. 2. Data format mismatches often cause issues in data processing.
Semantic Consistency	Government	1. Terms and concepts are consistent, making data interpretation easier for policy-making. 2. Semantic consistency supports clear communication and	1. Inconsistent terminology often confuses data users and affects its usage. 2. Variation in terms within the data can lead to difficulties in interpretation and application.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Parameter	Group	Positive	Negative
		understanding of the data.	
	Academics	1. Consistent terms and concepts support valid analysis and study replication. 2. Data with semantic consistency facilitates the research process and understanding.	1. Some semantic inconsistencies can affect the validity of research results. 2. Semantic misalignment in the data can disrupt analysis and interpretation.
	Industry / Business Practitioners	1. Consistent terms and concepts support clear business analysis and reporting. 2. Semantic consistency aids in the creation of business reports and strategies.	1. Terminology misalignment often affects the understanding and use of data for business decisions. 2. Semantic inconsistencies can lead to difficulties in interpreting business data.

The results of the data quality evaluation for inherent data quality parameters on the OGD portal, as presented in Table 2, show varying perspectives from different groups of data providers and users. From the government's perspective, the data is generally accurate and consistent, supporting policymaking and decision-making. However, there are some errors and format inconsistencies that affect data integration. The data is also considered complete, but there is a lack of information in some datasets that requires improvement. The precision and syntactic consistency of the data are rated as good, but there is a need to enhance precision and align semantic terms to avoid confusion or ambiguity.

From the academic perspective, the data provided is generally accurate and supports valid research. However, minor errors and data format inconsistencies can affect the results of analysis and interpretation. The data is considered complete for comprehensive analysis, although the lack of detail in some datasets

affects the depth of research. The precision of the data supports detailed outcomes, but there are some precision deficiencies that need to be addressed. The syntactic and semantic consistency of the data supports research, but semantic misalignment can disrupt interpretation.

Business practitioners view the data on the OGD portal as accurate and supportive of strategic decision-making, though inaccuracies occasionally disrupt analysis. Data consistency aids in analysis and integration, but there are issues with format misalignment that affect the efficiency of analysis. Complete data supports business decisions, but the lack of critical information can hinder the decision-making process. Good data precision is crucial for business analysis, yet there are deficiencies that impact outcomes. Syntactic consistency supports efficient analysis, although improvements are needed in addressing semantic misalignment to enhance understanding and data utilization.

Table 3. Interview Results Based on System-Dependent Data Quality Parameters

Parameter	Group	Positive	Negative
Timeliness	Government	1. Data is updated regularly, ensuring that information is always relevant and current. 2. Timeliness of data supports timely and responsive policymaking.	1. Data updates are occasionally delayed, causing information to not always be relevant. 2. Data is sometimes not updated according to schedule, reducing its relevance for current policies.
	Academics	1. Data is generally up-to-date, supporting relevant and timely research. 2. Up-to-date data enables studies based on the most current and accurate information.	1. Delays in data updates can impact the relevance of research findings. 2. Delays in updates can sometimes affect the accuracy and relevance of

Parameter	Group	Positive	Negative
			research findings.
	Industry / Business Practitioners	1. Data is updated in a timely manner, supporting swift and accurate business decisions. 2. Data timeliness supports business strategies based on the latest information.	1. Delays in data updates sometimes disrupt business and operational decisions. 2. Data that is not always up-to-date can lead to inaccurate business decisions.
Accessibility	Government	1. Data is easily accessible and downloadable from the portal, facilitating its use by the public. 2. Data accessibility supports transparency and public participation.	1. Some datasets are difficult to access or download, hindering widespread use of the data. 2. Inconsistent data access reduces the ease of obtaining the required information.
	Academics	1. Data is easily accessible in a format suitable for analysis and research. 2. Data accessibility supports the effectiveness of research and analysis.	1. Some data formats are difficult to access or do not meet research needs. 2. Difficulties in accessing data can impact the research and analysis process.
	Industry / Business Practitioners	1. Data is easily accessible and download	1. Data access is sometimes limited, or

Parameter	Group	Positive	Negative
		formats are suitable for business needs. 2. Good accessibility supports efficient business operations.	download formats do not meet business needs. 2. Data access issues can impede business operations and decision-making.
Representation	Government	1. Data is presented in a clear and standardized format, supporting accurate interpretation. 2. Good data representation facilitates understanding and use of information by relevant stakeholders.	1. Data format and structure are often confusing, complicating user interaction with the data. 2. Data with poor representation can hinder the understanding and use of information.
	Academics	1. The data format is adequate, supporting effective analysis and interpretation. 2. Good data representation facilitates the research and analysis process.	1. Some data formats could be clearer, but they are generally understandable with additional effort. 2. Data formats and structures sometimes complicate the interpretation of research results.
	Industry / Business Practitioners	1. Data is presented in a way that supports	1. The format of data representation is often inadequate

Parameter	Group	Positive	Negative
		effective business analysis. 2. The data structure facilitates integration and use in business processes.	for in-depth business analysis. 2. Poor data representation can disrupt the data processing and analysis processes.
Security	Government	1. Data is well-protected, reducing the risk of unauthorized access. 2. Data security supports public trust in the use of government data.	1. There are still concerns about potential data security breaches. 2. There are some concerns about data protection and potential risks of breaches.
	Academics	1. Data is secure and data protection is sufficient for research purposes. 2. Data security supports the integrity of research and the validity of results.	1. Concerns regarding data protection and potential unauthorized access may impact research. 2. Data protection needs improvement to mitigate the risk of unauthorized access.
	Industry / Business Practitioners	1. Data protection is adequate, supporting compliance with regulations. 2. Secure data supports reliable and efficient	1. Data protection sometimes needs enhancement to avoid security breaches. 2. Weaknesses in data protection can pose risks to

Parameter	Group	Positive	Negative
		business operations.	business integrity.
Relevance	Government	1. The data is relevant to the research topic, encompassing the necessary variables. 2. Data relevance supports deeper understanding and research.	1. Some necessary data is unavailable, reducing relevance for current policies. 2. Irrelevant data often impacts the effectiveness of government policies.
	Academics	1. Data is relevant to the research topic and includes the necessary variables. 2. Data relevance supports deeper understanding and research.	1. There is a lack of relevant data in some research areas, affecting analysis. 2. Limitations in data relevance can reduce the quality of research outcomes.
	Industry / Business Practitioners	1. Data relevant to business needs supports accurate strategic decision-making. 2. Data relevance supports business strategies based on current information.	1. Lack of relevant data can hinder business decision-making and operational analysis. 2. Irrelevant data often impacts the planning and execution of business strategies.
Timeliness	Government	1. Data is available in a timely manner when needed for policy-making.	1. Occasional disruptions in data access can impact the

Parameter	Group	Positive	Negative
		2. Timeliness of data supports swift and accurate decision-making.	timeliness of its usage. 2. Timeliness is often disrupted, reducing the effectiveness of policies and decisions.
	Academics	1. Data is available in a timely manner to support efficient research processes. 2. Timeliness of data supports the validity and relevance of research results.	1. Disruptions in data access sometimes affect the timeliness of research. 2. Delays in data access can impact the research schedule and publication of results.
	Industry / Business Practitioners	1. Data is available in a timely manner, supporting swift business decision-making. 2. Timeliness of data supports efficient business operations.	1. Delays in data access often disrupt business and operational decisions. 2. Data access disruptions affect the timeliness of business decisions and operations.

Based on the responses from the interviewees, as shown in Table 3 regarding data quality on the OGD portal for the System-Dependent Data Quality parameters, several reflections can be made:

First, from the government's perspective, the quality of data on the OGD portal is somewhat varied. In terms of data timeliness, the government feels that the data on the OGD portal is generally up-to-date, which is very helpful in the policy-making process that requires current data. However, there are some cases where delays in data updates have hindered decision-making processes, particularly in

dynamic situations that require a quick response. Regarding accessibility, the government acknowledges that the data on the OGD portal is fairly easy to access, but there are challenges in understanding the data format and structure, which sometimes necessitates additional training for government employees to use it effectively.

Additionally, the government also assesses that the representation of data on the OGD portal is quite good, with visualizations and data formats that facilitate interpretation. However, there are some concerns regarding data security; although security protocols are considered adequate, there are concerns about potential security breaches, especially concerning sensitive data. On the other hand, the relevance of the data is rated as quite high, as the data provided generally aligns with the needs of the policies being formulated. Nevertheless, the government has identified some datasets that are either irrelevant or outdated, which should be updated or removed. Lastly, the timeliness of data in supporting administrative processes and policies is also recognized as important, but there are still datasets that are occasionally not available according to the required schedule, thereby disrupting the decision-making process.

Second, from the academic perspective, the quality of data on the OGD portal also presents several challenges and opportunities. Regarding data timeliness, academics often find that the data on the OGD portal is not up-to-date, which can hinder research requiring current data. This underscores the need for a more responsive and regular data updating system. In terms of accessibility, although the data on the OGD portal is generally easy to access, academics frequently encounter difficulties in obtaining well-formatted data or data available in formats compatible with the analytical tools they use. This highlights the need for better data format standardization.

In terms of data representation, academics tend to appreciate the available data visualizations, but they also emphasize that additional context is often needed to better understand the data, especially when it is used for in-depth analysis. Data security is not as much of a primary concern for academics compared to the government, though it is still recognized as important, particularly when the data used involves sensitive information. Data relevance is a major concern, with academics often finding that some data does not meet their research needs, indicating the need for better data curation. Lastly, regarding data timeliness, academics feel that delays in obtaining the necessary data can postpone publication or the completion of research, making timely access crucial.

Third, from the perspective of industry and business practitioners, the quality of data on the OGD portal has a direct impact on strategic business decisions. Regarding data timeliness, business practitioners emphasize the importance of

having always up-to-date data for market analysis and operational decisions. Delays in data updates can negatively affect business opportunities and responsiveness to market changes. Data accessibility on the OGD portal is also rated as fairly good by business practitioners, but they emphasize the need for more user-friendly data formats that can be directly used in business analysis tools without requiring additional conversion or adjustments.

Regarding data representation, business practitioners appreciate clear and concise visualizations, but they also desire more customization options in data visualizations to better suit the specific needs of their industry. In terms of data security, the industry places a high priority on data protection, especially data that could affect their competitive position in the market. This concern highlights the importance of strong security protocols and transparency in how data is protected. Data relevance is also a key focus, as business practitioners often require highly specific and relevant data for market analysis or trend prediction. Data that is less relevant or outdated is seen as a barrier in the decision-making process. Finally, the timeliness of data availability is crucial for business practitioners, as any delays can result in missed opportunities or delayed reactions to market dynamics.

Furthermore, this research also provides a perspective on the relationship between each evaluation parameter, both from the Inherent Data Quality and System-Dependent Data Quality sides, and the three groups of data providers and users. The correlation between the opinions of various groups—government, academics, and industry/business practitioners—serves to identify similarities and differences in perspectives related to certain data quality parameters. This helps to understand the extent to which their assessments are similar or different from each other. The correlation also reveals potential issues and areas for improvement, particularly when there are significant differences in assessments, which may indicate certain aspects of the data that do not sufficiently meet the needs of specific users. By examining these correlations, policymakers or OGD portal managers can prioritize improvements in data quality, especially in areas showing low correlation.

The following are the correlation results generated from the data quality evaluation on the OGD portal for the two main parameters.

Table 4. Correlation Results of Data Providers' and Users' Perspectives on Inherent Data Quality

Parameter	Government	Academics	Industry / Business Practitioners	Correlation
Accuracy	4 (Good)	3 (Adequate)	3 (Adequate)	Moderate correlation:

Parameter	Government	Academics	Industry / Business Practitioners	Correlation
				Industry and government agree that accuracy is good, while academics consider it sufficient.
Consistency	3 (Adequate)	3 (Adequate)	3 (Adequate)	Moderate correlation: Government and academia agree that consistency is sufficient, whereas the industry is more satisfied.
Completeness	3 (Adequate)	2 (Poor)	3 (Adequate)	Moderate correlation; perspectives are fairly similar, but academics find completeness to be lower compared to government and industry.
Precision	4 (Good)	3 (Adequate)	3 (Adequate)	Moderate correlation: The government views precision as good, while academics and industry consider it adequate.

Parameter	Government	Academics	Industry / Business Practitioners	Correlation
Syntax Consistency	3 (Adequate)	4 (Good)	3 (Adequate)	Moderate correlation: Academics are more satisfied with syntax consistency compared to the government and industry.
Semantic Consistency	3 (Adequate)	3 (Adequate)	4 (Good)	Moderate correlation: The industry is more satisfied with semantic consistency compared to the government and academics.

The correlation results of data providers' and users' perspectives on Inherent Data Quality, as presented in Table 4, show that the government and industry/business practitioners generally rate the accuracy and consistency of the data as good, while academics are more critical, rating it as adequate. Data completeness is rated lower by academics, whereas precision is considered good by the government and adequate by others, indicating a need for improvement in precision and completeness to meet the expectations of all users. Syntactic consistency is rated good by academics, but the government and industry rate it as adequate, while semantic consistency is rated good by the industry but less satisfactory by the government and academics.

Table 5. Correlation Results of Data Providers' and Users' Perspectives on System-Dependent Data Quality

Parameter	Government	Academics	Industry / Business Practitioners	Correlation
Timeliness of Data	4 (Good)	3 (Adequate)	2 (Poor)	Low correlation; differing

Parameter	Government	Academics	Industry / Business Practitioners	Correlation
				opinions regarding data timeliness. Academics and industry stakeholders feel that the data is not sufficiently up-to-date.
Accessibility	3 (Adequate)	4 (Good)	4 (Good)	High correlation; data accessibility is generally rated well by academia and industry, while the government still perceives some challenges.
Data Representation	4 (Good)	3 (Adequate)	4 (Good)	Moderate correlation; both the government and industry are satisfied with data representation, whereas academia feels improvements are needed.
Data Security	3 (Adequate)	3 (Adequate)	4 (Good)	Moderate correlation; the industry is more satisfied with data security

Parameter	Government	Academics	Industry / Business Practitioners	Correlation
				compared to the government and academia.
Relevance of Data	4 (Good)	3 (Adequate)	4 (Good)	High correlation; data is considered relevant by the government and industry, but academics feel that some data does not fully align with research needs.
Timeliness	3 (Adequate)	2 (Poor)	3 (Adequate)	Moderate correlation; the timeliness of data is rated as adequate by government and industry, but academics often feel that the data is not available in a timely manner.

Rating Scale: 5 = Very Good, 4 = Good, 3 = Adequate, 2 = Poor, 1 = Very Poor

Meanwhile, the results of the correlation between user perspectives and System-Dependent Data Quality, as presented in Table 5, reveal some common viewpoints regarding data accessibility and relevance. All user groups generally rate these aspects as good or sufficient, reflecting the effectiveness of the OGD portal in providing useful data. However, data timeliness emerges as a significant issue, with business users frequently finding the data outdated, while government representatives rate it as adequate, and academics often encounter data that is not up-to-date. Data security is rated more favorably by business users compared to government and academic users, indicating a need for enhanced security

protocols. Data timeliness is rated as sufficient overall, but there is a perceived delay, particularly by academics, highlighting the need for improvements to ensure data is available in a timely manner. In summary, while there is agreement on several aspects, the differences in perspective are crucial for directing efforts to improve data quality, ensuring that the OGD portal can more effectively meet the diverse needs of its users.

3.2 Discussion

The study reveals several strengths and limitations in the quality of data on the OGD portal from different perspectives. From the government's standpoint, the timeliness of the data is largely appreciated, as it provides current information necessary for effective policy-making. This timely data is seen as a significant advantage, facilitating informed decisions. Additionally, the portal is generally regarded as user-friendly in terms of data accessibility, and both the government and business practitioners find the data visualizations to be helpful for interpreting and utilizing the data. The alignment of data with policy needs further underscores the portal's relevance for governmental purposes. Moreover, the study's cross-sectoral insights offer a comprehensive view of data quality, highlighting how the portal serves various user groups.

However, the study also identifies several drawbacks. Academics and business practitioners frequently find the data to be outdated, which impedes their research and business decision-making. This delay in data availability points to a need for more regular updates to keep pace with the dynamic nature of research and market conditions. Although the data is generally accessible, the formats provided often require additional adjustments, posing challenges for effective use. Academics, in particular, find that the lack of contextual information in data visualizations limits their depth of analysis. Business practitioners also express a need for more customizable visualization options to better meet industry-specific requirements. Concerns about data security, particularly for sensitive information, are also highlighted by both government and industry users. Finally, issues of data relevance are noted, as some datasets are either outdated or not aligned with user needs. These findings reveal significant differences in how data quality is perceived across user groups, indicating areas where the OGD portal could be improved to better meet the diverse needs of its users.

In light of these findings, it is recommended that the OGD portal enhance its data update frequency to address timeliness issues, improve data format standardization to ease accessibility, and provide more contextual information and customizable visualization options. Strengthening data security protocols and curating datasets to ensure their relevance will also be crucial. By addressing these

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

concerns, the portal can better serve its various users and improve overall satisfaction and utility.

4. CONCLUSION

The perspectives and evaluations of data quality on the OGD portal vary among data providers and users—government, academia, and industry/business. The government generally rates the data quality moderately, providing positive assessments on accuracy, consistency, and precision, but considers aspects such as completeness and timeliness as needing improvement. Academics are more critical, particularly concerning data completeness and timeliness, which are often deemed insufficient for their research needs. Meanwhile, industry/business generally offers more positive evaluations, especially regarding data accessibility, relevance, and security, though they also believe that data timeliness needs enhancement. Overall, despite some commonalities in evaluation aspects, each group of data providers and users has distinct needs and expectations that influence their views on the data quality of the OGD portal.

The correlation between data quality evaluation parameters and the perspectives of data providers and users reveals significant similarities and differences. Inherent data quality parameters such as accuracy, consistency, and precision are often rated positively by the government and industry, while academics tend to be more critical, particularly regarding data completeness and precision. Regarding system-dependent data quality, parameters such as accessibility and relevance are generally rated positively by all user groups, but there are notable differences in the evaluations of data timeliness and punctuality. Business users find the data frequently outdated, whereas government and academic views are more varied. Data security is rated positively by industry but considered adequate by government and academia. The common viewpoints on some parameters indicate the effectiveness of the OGD portal in certain aspects, while differences highlight areas needing further attention and improvement. This correlation analysis helps identify improvement priorities and adjust data to better meet the needs of various user groups.

In response to the first research question, the study found that perspectives on data quality varied significantly among government data providers, academic researchers, and industry practitioners. Government officials generally viewed the data on the OGD portal as timely and accessible, though they noted issues with data updates and security. Academics often encountered challenges with outdated data and insufficient contextual information, which affected their research. Industry practitioners emphasized the need for up-to-date data and user-friendly formats to support business decisions. These findings highlight that while the portal is valued for its current data and accessibility, there are considerable gaps in

data relevance, timeliness, and format that impact different user groups in distinct ways.

Addressing the second research question, the study compared these perspectives against the ISO/IEC 25012:2008 data quality standards, which include attributes such as accuracy, completeness, consistency, and timeliness. The analysis revealed that while the OGD portal meets some of these standards, particularly in terms of accessibility and data representation, it falls short in areas such as timeliness and relevance. For instance, delays in data updates and challenges with data formats underscore a misalignment with the standards' requirements for timeliness and usability. Furthermore, issues related to data security and contextualization reflect gaps in meeting the completeness and contextual relevance aspects of the ISO/IEC standards.

For future research, several directions to consider include the development of strategies to improve data quality, particularly in terms of completeness, timeliness, and precision. Research could also focus on analyzing the impact of data quality improvements on user satisfaction and data utilization effectiveness. Additionally, developing a more comprehensive evaluation tool to assess data quality on the OGD portal could be a primary focus. International comparative studies may provide insights into best practices that could be applied locally. Understanding the differences in perspectives between data providers and users and how to meet their specific needs, as well as exploring the impact of integrating data from various sources, are also important research areas. Finally, research could assess the usability and accessibility aspects of the portal and strengthen data security strategies to protect against unauthorized access.

REFERENCES

[1] A. Vetrò, L. Canova, M. Torchiano, C. O. Minotas, R. Iemma, and F. Morando, "Open Data Quality Measurement Framework: Definition and application to open government data," *Gov. Inf. Q.*, vol. 33, no. 2, pp. 325–337, 2016, doi: 10.1016/j.giq.2016.02.001.

[2] J. Merino, I. Caballero, B. Rivas, & Serrano, M., and M. Piattini, "A data quality in use model for big data," *Futur. Gener. Comput. Syst.*, vol. 63, pp. 123–130, 2016, doi: 10.1016/j.future.2015.11.024.

[3] M. Janssen, P. Brous, E. Estevez, L. S. Barbosa, and T. Janowski, "Data governance: Organizing data for trustworthy Artificial Intelligence," *Gov. Inf. Q.*, vol. 37, no. 3, pp. 1–8, 2020, doi: 10.1016/j.giq.2020.101493.

[4] M. A. A. Alryalat, N. P. Rana, G. P. Sahu, Y. K. Dwivedi, and M. Tajvidi, "Use of social media in citizen-centric electronic government services: A literature analysis," *Crowdsourcing Concepts, Methodol. Tools, Appl.*, vol. 13, no. 3, pp. 925–977, 2019, doi: 10.4018/IJEGR.2017070104.

[5] M. Beno, K. Figl, J. Umbrich, and A. Polleres, "Open data hopes and fears: determining the barriers of open data," in *Conference for E-Democracy and Open Government (CeDEM*, 2017, pp. 69–81.
[6] I. Cantador, M. E. Cortés-Cediel, and M. Fernández, "Exploiting Open Data to analyze discussion and controversy in online citizen participation.," *Inf. Process. Manag.*, vol. 57, no. 5, 2020, doi: 10.1016/j.ipm.2020.102301.
[7] S. Kubler, J. Robert, S. Neumaier, J. Umbrich, and Y. Le Traon, "Comparison of metadata quality in open data portals using the Analytic Hierarchy Process," *Gov. Inf. Q.*, vol. 35, no. 1, pp. 13–29, 2018, doi: 10.1016/j.giq.2017.11.003.
[8] J. Chen, "The Dangers of Accuracy: Exploring the Other Side of the Data Quality Principle," *Eur. Data Prot. Law Rev.*, vol. 4, no. 1, pp. 36–52, 2018, doi: 10.21552/edpl/2018/1/7.
[9] Y. Sun, B. Wang, Z. Sun, and X. Yang, "Does Every Data Instance Matter? Enhancing Sequential Recommendation by Eliminating Unreliable Data," in *The Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21)*, 2021, pp. 1579–1585.
[10] Z. Zhao *et al.*, "RAD: On-line Anomaly Detection for Highly Unreliable Data," 2019, doi: 10.1109/DSN.2019.00068.
[11] M. Yi, "Exploring the quality of government open data: Comparison study of the UK, the USA and Korea," *Electron. Libr.*, vol. 37, no. 1, pp. 35–48, 2019, doi: 10.1108/EL-06-2018-0124.
[12] A. Quarati, "Open government data: usage trends and metadata quality," *J. Inf. Sci.*, vol. 49, no. 4, pp. 1–24, 2023, doi: 10.1177/01655515211027775.
[13] C. Guerra-García, A. Nikiforova, S. Jiménez, and L. Perez-Gonzalez, H. G., Ramírez-Torres, M., Ontañon-García, "ISO/IEC 25012-based methodology for managing data quality requirements in the development of information systems: Towards Data Quality by Design," *Data Knowl. Eng.*, vol. 145, 2023, doi: 10.1016/j.datak.2023.102152.
[14] D. Williams and H. Tang, "Data quality management for industry 4.0: A survey," *Softw. Qual. Prof.*, vol. 22, no. 2, pp. 26–35, 2020.
[15] H. Fadlallah *et al.*, "Context-aware big data quality assessment: a scoping review," *ACM J. Data Inf. Qual.*, vol. 15, no. 3, pp. 1–33, 2023, doi: 10.1145/3603707.