buildings

MDPI

*Article*

# Determinants of Data Quality Dimensions for Assessing Highway Infrastructure Data Using Semiotic Framework

**Chenchu Murali Krishna [1,*], Kirti Ruikar [2] and Kumar Neeraj Jha [1]**

[1] Department of Civil Engineering, Indian Institute of Technology Delhi, Delhi 110016, India
[2] School of Architecture, Building and Civil Engineering, Loughborough University, Epinal Way, Loughborough LE113TU, Leicestershire, UK
* Correspondence: cez198622@iitd.ac.in

**Abstract:** The rapid accumulation of highway infrastructure data and their widespread reuse in decision-making poses data quality issues. To address the data quality issue, it is necessary to comprehend data quality, followed by approaches for enhancing data quality and decision-making based on data quality information. This research aimed to identify the critical data quality dimensions that affect the decision-making process of highway projects. Firstly, a state-of-the-art review of data quality frameworks applied in various fields was conducted to identify suitable frameworks for highway infrastructure data. Data quality dimensions of the semiotic framework were identified from the literature, and an interview was conducted with the highway infrastructure stakeholders to finalise the data quality dimension. Then, a questionnaire survey identified the critical data quality dimensions for decision-making. Along with the critical dimensions, their level of importance was also identified at each highway infrastructure project's decision-making levels. The semiotic data quality framework provided a theoretical foundation for developing data quality dimensions to assess subjective data quality. Further research is required to find effective ways to assess current data quality satisfaction at the decision-making levels.

**Keywords:** highway data quality assessment; data quality dimensions; semiotic framework; decision-making



**Citation:** Krishna, C.M.; Ruikar, K.; Jha, K.N. Determinants of Data Quality Dimensions for Assessing Highway Infrastructure Data Using Semiotic Framework. *Buildings* **2023**, *13*, 944. https://doi.org/10.3390/buildings13040944

Academic Editors: Ming-Hung Hsu, Osama Abudayyeh, Zheng-Yun Zhuang and Ying-Wu Yang

Received: 30 December 2022
Revised: 27 March 2023
Accepted: 29 March 2023
Published: 2 April 2023



## 1. Introduction

Highway agencies devote significant resources to collecting, storing, and maintaining many forms of data, ranging from preliminary survey data to pavement condition data, throughout the life cycle of a highway project. For instance, the National Highway Authority of India launched Data Lake, a project monitoring tool to track and monitor the progress of projects and to act as the central repository of documents across the project life cycle [1]. According to the FMI's (2019) report titled "Big Data Equals Big Questions for the Engineering and Construction Industry," some of the most significant infrastructure projects require an average of 130 million emails, 55 million documents, and 12 million workflows. At the same time, 95.5% of all data collected in the engineering and construction industry is unutilised because many firms cannot manage and process vast amounts of data for decision-making [2]. According to a 2018 industry report titled "Construction Disconnected" by FMI, 48% of all reworks in infrastructure projects in the United States are caused by poor data and miscommunication, resulting in an annual cost of over USD 31.3 billion. Globally, an average of 52% of rework was caused by poor data and communication, amounting to USD 280 billion. The primary cause of poor data and information was that 34.4 percent of reworks were caused by incorrect project data, meaning it was out-of-date or otherwise flawed data, while 28.8 percent of reworks were caused by difficulty gaining access to necessary project data [3]. Despite the significant investment, data utilisation to users' needs for extracting information, knowledge, and support decisions

has become debatable [4]. Data collection is becoming an increasingly significant asset for today's highway arena within highway management and operation. Several systems and technologies have created significant infrastructure data in recent years [5].

Data have been widely used to manage system operations and provide information on highway conditions. However, public and private users discovered that utilising and operating the data is becoming increasingly complex. Data are collected with varying degrees of precision and resolution, and data formats are often incompatible [6].

Technological advancements in data collection result in the real-time monitoring of data and a massive volume of data, Such as data collected in the structural health monitoring of a bridge [7] and data collected during the degradation process of concrete material [8]. In addition, the issue intensifies as the volume of data continues to increase [9–15]. Ghasemaghaei and Calic [16] discussed the role of data quality and diagnosticity in the firm's decision-making, considering the effect of big data processing. However, there is substantial evidence that data quality issues are pervasive in practice and that relying on poor or uncertain data results in less effective decision-making. It also increases the cost of correcting the data in the decision-making process of highway projects [17,18].

Data quality has been extensively studied in various disciplines for several decades [19]. It has become a professional field, emphasising organisational strategy and effective decision-making [20,21]. In addition, data quality is considered a multi-dimensional concept in the literature [22–24]. In the last two decades, scholars and practitioners have proposed several classifications of data quality dimensions, many of which have overlapping and occasionally contradictory meanings concerning respective disciplines (e.g., [14,24–26]). Despite the different classifications, few investigations have attempted to integrate these perspectives of data quality dimensions to assess the quality of highway data for effective decision-making. For instance, Coleman [27] gave an insightful examination of the various current classifications of data quality dimensions and identified sixteen mutually incompatible dimensions.

Although numerous studies have found the significance of data quality for decision-making based on various frameworks and methodologies, not much focus has been given to assessing data quality at different decision-making levels of highway projects [5,28,29]. Samitsch et al. [30] provided a guide for companies seeking to improve organisational performance by improving data quality, with a combination of 16 dimensions. Addressing this issue necessitates a method for comprehending data quality, followed by methods for enhancing data quality and decision-making based on data quality information. This research proposes a semiotic-based framework for comprehending highway infrastructure data quality, consisting of four levels: syntactic (form), empiric (connection), semantic (meaning), and pragmatic (use) [29]. The semiotic-based framework assesses and understands data quality based on the semiotic theory's application. Semiotic theory concerns using signs and symbols to convey data, information, and meaning [31]. A review of data quality frameworks applied in various fields was also carried out. Such as the semiotics framework, AIMQ methodology, data quality assessment (DQA), the observe-orient-decide-act methodology (OODA DQ), and the Canadian Institute for health information methodology (CIHI) framework are used in the healthcare industry for data quality assessment [32–36], while the total data quality management (TDQM) framework, comprehensive methodology for data quality management (CDQ), data quality practical approach (DQPA), task-based data quality method (TBDQ), and data quality assignment framework (DQAF) are used in the IT industry to deliver high-quality information products (IP) to information consumers [9,37–41]. A DQMos model and DQMes methodology are used for evaluating data quality in software engineering experiments data [42]. A questionnaire survey identified the critical data quality dimensions of the proposed semiotic framework levels from the National Highway stakeholders for decision-making. The survey helps the National highway stakeholders understand the parameters or dimensions of data quality to assess the quality of data stored in the data lake. The study investigated identifying the framework for

analysing data quality and determining the appropriate framework for assessing highway infrastructure data. Currently, there are no specialised studies of data quality dimensions for evaluating highway infrastructure data.

A literature review was conducted first for the study, followed by identifying data quality frameworks. The second step identified data quality dimensions within the four levels of the semiotic data quality framework. In the third step, an interview and questionnaire were conducted in two stages. Initially, an interview survey was undertaken to develop a list of data quality characteristics that reflect the opinions of data consumers regarding data quality. For the second stage, a questionnaire was developed from the identified dimensions through an interview study. The questionnaire survey was conducted to gather information on the importance of each of these dimensions to data consumers at the individual level of decision-making, followed by a ranking of the dimensions within the categories of semiotic frameworks to comprehend stakeholders' priorities for each characteristic data quality.

The paper is structured as follows: The next section focuses on the literature review of frameworks and data quality dimensions and identifies the most effective framework for evaluating highway infrastructure data. The subsequent section addresses the research methodology, the findings, and an analysis of the findings. Finally, conclusions and future work scope are presented.

## 2. Objectives

The study's main objective was to investigate highway infrastructure data quality dimensions and the framework for assessing data quality. According to the 2018 FMI report, the cost of reworks caused by poor data quality and accessibility of data in the United States was USD 31.3 billion, while in Australia and New Zealand, it was USD 8.4 billion, and in the United Kingdom, it was USD 10.2 billion [3]. The literature shows that poor data quality negatively impacts the time and cost to make a decision and decision-making performance in the highway infrastructure project lifecycle. Hence, assessing data quality is critical for organisations and creates importance for identifying the dimensions to define data quality. The objective of the study was divided into three key research objectives as follows:

- To establish the data quality dimensions necessary for determining the data quality of highway infrastructure data to facilitate effective decision-making.
- To determine the importance of data quality dimensions at each level of decision-making.
- To determine the priority of dimensions within the semiotic framework categories.

## 3. Literature Review

### *3.1. Data and Data Quality*

Before going to the concept of data quality dimensions, let us review the first-order questions that arise from the history of the data quality domain. What is data, and what is data quality? Liebenau and Backhouse [43] defined data as "linguistic, mathematical or other symbolic representation that is universally accepted to represent people, things, events, and ideas." Data represent objects or processes in the actual world in their most basic form. Thus, while addressing data quality, we may argue that poor data quality results from an inaccurate depiction of the real world [44]. Abedjan et al. [45] addressed the tools used for detecting data errors. The study of data quality assessment began in the 1950s, particularly regarding the quality of products and services. Several researchers published several definitions, though no universally accepted definition of data quality exists. Wang and Strong [46] defined data quality as information usable by data consumers, and Crosby [47] defined it as "conformance to requirements." The General Administration of Quality Supervision, 2008, defined data quality as "the degree to which a set of inherent characteristics fulfil the requirements" [15]. At the same time, Fu and Easton [48] explained that data quality is commonly referred to as a collection of "characteristics" of data, such as precision, exhaustiveness, consistency, and timeliness. Most of these characteristics dictate the various dimensions along which data quality may be represented. A low degree

of data quality can significantly influence the overall effectiveness of the associated data applications [49].

### *3.2. Data Quality Assessment Framework*

Researchers define various frameworks and approaches for data quality assessment. For example, Madnick and Zhu [50], English [51], and Redman [52] explored strategies for increasing data quality, Batini et al. [53] provided a thorough and comparative description of data quality techniques for assessing and improving data quality, Gao et al. [54] proposed a fussing attributes approach for improving uncertain data quality, and Madnick et al. [55] reviewed current practices and research in the field. The research literature describes or defines data quality from simple lists of data quality dimensions to comprehensive frameworks (for example, [24,25,29,56]).

Hassenstein and Vanella [57] presented a data quality encyclopedia for the data life cycle. It describes the data quality dimensions, the data quality evaluation procedure, and the data quality context and practices in various fields. At the same time, Gabr et al. [58] comprehensively defined each traditional and big data quality dimension, metrics, and handling approach with specific definitions. They examined the metrics and methodologies used to monitor and manage each dimension and how they are monitored and managed. The study also examined the most-used data quality dimensions of traditional and large data sets.

Svetlana [59] presented the findings of an expert survey on data quality concerns to demonstrate that it is not required to employ all the numerous dimensions of data quality provided by researchers. However, the essential data quality criteria may be blended for a particular application. The study equips data users and producers with the knowledge necessary to effectively address application-specific data quality issues. In addition to the Svetlana findings, Eliza et al. [60] provided a methodology that allows users to manage data quality and make decisions based on data quality. It eliminates the requirement to fully integrate insufficient data by considering the operational context of the user to enhance a specific element of data quality.

Different approaches from the literature review were summarised to review the well-known and established frameworks for assessing and improving data quality for different data types. Table 1 lists fourteen data quality frameworks identified from the literature.

**Table 1.** Frameworks identified from the literature review.

| S. No. | Framework | Dimensions | References |
|---|---|---|---|
| 1 | TDQM: Total Data Quality Management | Accuracy, objectivity, believability, reputation, access, security, relevance, value-added, timeliness, completeness, amount of data, interpretability, ease of understanding, concise representation, and consistent representation. | [38] |
| 2 | TIQM: Total Information Quality Management | Definition conformance, completeness, validity, accuracy, precision, non-duplication, the equivalence of redundant or distributed data, accessibility, timeliness, contextual clarity, derivation integrity, usability, usability, and rightness. | [51] |
| 3 | COLDQ: Cost-effect of Low Data Quality | *Data model:* Clarity of definition, comprehensiveness, flexibility, robustness, essentialness, attribute granularity, the precision of domains, homogeneity, naturalness, identifiability, obtainability, relevance, simplicity, and semantic and structural consistency. *Data values:* Accuracy, completeness, consistency, currency, null values, and timeliness. Information Policy: Accessibility, metadata, privacy, redundancy, security, and unit cost. *Presentation:* Appropriateness, correct interpretation, flexibility, format precision, portability, consistent representation, representation of null value, and use of storage. | [61] |

**Table 1.** *Cont.*

| S. No. | Framework | Dimensions | References |
|---|---|---|---|
| 4 | AIMQ: A Methodology for Information Quality Assessment | Accessibility, appropriate amount, believability, completeness, concise representation, consistent representation, ease of operation, free-of-error, interpretability, objectivity, relevancy, reputation, security, timeliness, and understandability. | [34] |
| 5 | DQA: Data Quality Assessment | Accessibility, appropriate data, objectivity, believability, reputation, security, relevancy, value-added, timeliness, completeness, interpretability, ease of manipulation, understandability, concise representation, consistent representation, and free-of-error. | [35] |
| 6 | HIQM: Hybrid Information Quality Management | Accuracy, completeness, consistency, and timeliness. | [62] |
| 7 | CDQ: Comprehensive Methodology for Data Quality Management | Accuracy, completeness, and currency, Unstructured: Currency, relevance, and reliability. | [63] |
| 8 | DQPA: A Data Quality Assessment Framework | Accuracy, completeness, consistency, timeliness, uniqueness, and volatility. | [39] |
| 9 | SPDQM: Square-Aligned Portal Data Quality Model | Accuracy, traceability, correctness, expiration, completeness, consistency, accessibility, compliance, confidentiality, efficiency, precision, and understandability. Availability, accessibility, verifiability, confidentiality, portability, and recoverability. Validity, value-added, relevancy, specialisation, usefulness, efficiency, effectiveness, traceability, compliance, precision, concise representation, consistent representation, attractiveness, and readability. | [64] |
| 10 | HDQM: A Data Quality Methodology for Heterogeneous Data | Accuracy and currency. | [65] |
| 11 | DQAF: Data Quality Assessment Framework | Completeness, timeliness, validity, consistency, and integrity. | [40] |
| 12 | TBDQ: Task-Based Data Quality Method | Accuracy, completeness, consistency, and timeliness. | [41] |
| 13 | OODADQ: The Observe-Orient-Decide-Act Methodology | Speed and volume. | [36] |
| 14 | Semiotic Approach Data Quality-SESP model | Accuracy, consistency representation, unbiased, accessibility, up-to-date, traceability, security, believability, interpretability, ease of manipulation, understandability, completeness, appropriate amount of information, relevancy, concise representation, value-added, and reputation. | [32] |

According to the analysis of the frameworks listed in Table 1, the data quality dimensions considered by each framework vary considerably. Some data quality dimensions are recognised by only one framework, whereas specific dimensions appear frequently. For example, the HDQM and OODADQ frameworks considered only two dimensions for assessment, while the frameworks DQA and HIQM considered more than four dimensions. The dimensions varied according to the field of applications and perspective of the application, such as the health care industry, information technology, and business management. For example, let us consider how the accuracy dimension has been used in the HDQM and HIQM frameworks. In the HDQM framework in the IT industry, dimension accuracy is defined as the proximity between a value "v" and another value "v." of the domain D in the user interface development. This is regarded as the correct representation of the real-world phenomenon value "v" seeks to represent. At the same time, the HIQM framework in the business management sector defines accuracy as the value difference between two

databases containing the same value as the correct representation of the real-world value. To understand the most critical dimensions applied in the various fields, the frequency of usage by different data quality dimensions was considered and is shown in Figure 1. Only dimensions used more than once are considered in the figure. The study of Figure 1 helps finalise the dimensions from the literature review perspective to be identified in the data quality in the semiotic framework for assessing highway infrastructure data. The semiotic approach data quality framework is the most applicable of the 14 frameworks mentioned above for evaluating highway infrastructure data. The reason for the selection is explained in the semiotic framework section.

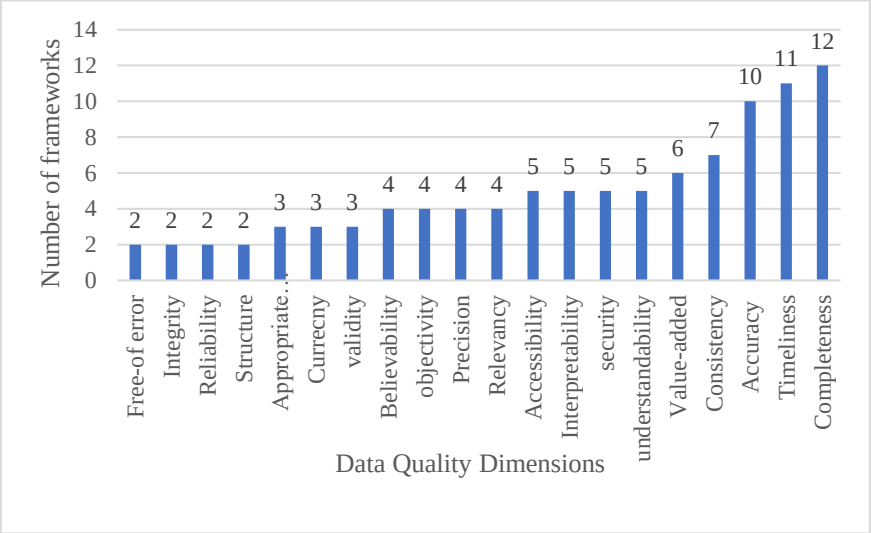


**Figure 1.** Number of frameworks that used specific data quality dimensions.

### *3.3. Semiotic Framework*

Semiotics is the study of signs and symbols used to convey meaning to various users. Data quality researchers have also adopted the semiotic perspective of data; for instance, Price and Shanks [29] identified three data quality levels: syntactic, semantic, and pragmatic. Semiotic theory concerns using symbols to convert knowledge and define levels in the framework for analysing structure, physical form, meaning, and data usage. A thorough examination of the various levels of semiotics would reveal that the pragmatic level is associated with knowledge, the semantic level with information, and only the syntactic level with data. In other words, the dimensions operating at the pragmatic, semantic, and syntactic levels pertain to the quality of knowledge, information, and data.

According to Falkenberg et al. [66], data are meaningful symbolic creations consisting of a limited arrangement of signs and symbols. Thus, the semiotic framework was used in this study to define data quality dimensions. The semiotic framework consists of four levels: empiric, syntactic, pragmatic, and semantic. Each level of the semiotic framework facilitates data quality evaluation from several perspectives, including structure, data, information, and knowledge for assessing highway infrastructure data for decision-making at various levels of the highway decision-making hierarchy, for instance, while selecting a treatment technique for damaged pavement in a highway construction project.

Each decision-making level bases its decisions on the raw data, information, and knowledge available at that level. The strategic level is the top level of an organisation and is responsible for strategic planning. This involves making long-term, big-picture decisions and establishing policies that impact the organisation. For the decision of treatment technique, the system performance (policymaking) policies are established, requiring knowledge to make policies. Similarly, at the network level, the fund distribution (planning) decisions are made, i.e., allocating funds according to project requirements. At the program level, the decision of pavement evaluation and prioritisation is considered for each project.

At the project selection level, the project is selected according to the prioritisation made at the program level, and treatment selection is made at the project level.

Kahn et al. [67] addressed the relationships between semiotic levels, the data-information-knowledge (DIK) hierarchy, and associated data-quality issues, as shown in Figure 2. The relationship between semiotic levels and structure, data, information, and knowledge facilitate the identification of unique data quality issues that may necessitate the application of specialised skills to resolve. Knoke and Yang [68] claimed that information originates with data and is transferred to knowledge in the DIK hierarchy. Depending on how data's meaning, structure, and operation are communicated at different semiotic levels of the DIK hierarchy, such transference could increase or decrease data's meaningfulness, transferability, and applicability.

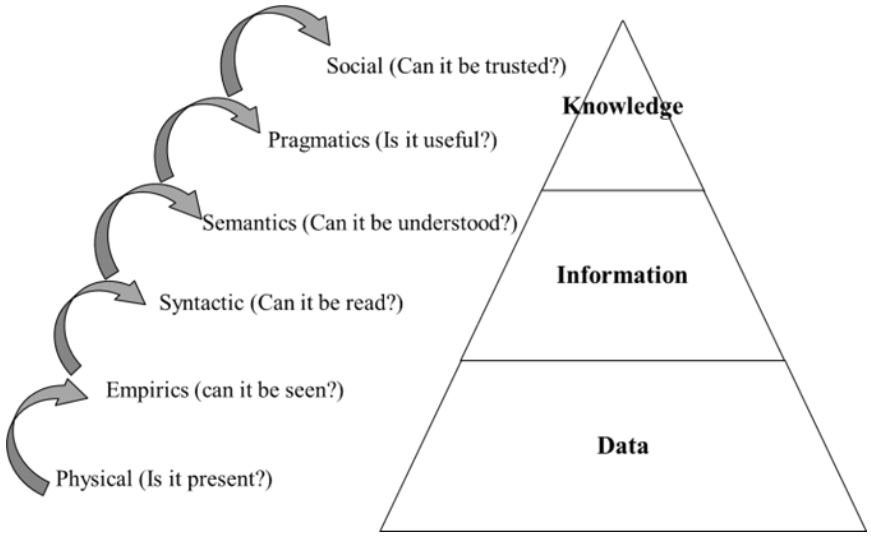


**Figure 2.** Semiotic levels and the data-information-knowledge hierarchy, adapted from [32]. 2018, Huang.

The empiric level focuses on the quality aspect of data access and the means of communication. It considers how much and in what way raw data are available for stakeholders for decision-making. In highway projects, decision-makers at each project phase, such as preconstruction, construction, and post-construction phases, consider data availability essential for effective decision-making. At the empirical level, accessibility, security, and timeliness (currentness) are considered to evaluate the data communication and access perspective of the raw data stored in the data lake [68]. For example, the dimension accessibility of highway data could be the availability of real-time traffic data on a particular highway. If the data are easily accessible through an open data portal such as a data lake, API, or mobile app, they would have a high level of accessibility. On the other hand, if the data are only available through a difficult-to-navigate website or requires complex technical skills, they would have a low level of accessibility.

On the other hand, the syntactic level concentrates on the forms and structure of data, or, more accurately, their physical form instead of their content. After assessing the accessibility criteria of raw data, the second crucial limitation for decision-makers is the kind and format of accessible data. To quantify the structure of raw data stored in a data lake, the syntactic level considers accuracy, concise presentation, ease of operation, consistency, integrity, and completeness as data quality dimensions [49]. For instance, the accuracy dimension in highway infrastructure data could be the precision of the measurements taken for the width of a particular road lane. Inaccurate measurements could lead to too narrow lanes, potentially causing safety issues or impeding traffic flow.

The semantic level of data quality is concerned with the meaning of data for information generation rather than the data [69]. The decision-makers at the program and project

selection decision-making levels require information regarding project performance for decisions such as budget allocation and project prioritisation. The dimensions at the semantic level are credibility, interpretability, and understandability for assessing the interpretation of data that provides meaning. For example, dimension interpretability refers to the ease with which stakeholders can understand and use data. In the context of highway data, interpretability could be the use of visualisations or dashboards that make it easier for stakeholders to understand complex data sets. This could include interactive maps or charts that allow users to explore different aspects of highway infrastructure data, such as traffic volume or accident rates.

The pragmatic level is concerned with the relationship between data, information, and behaviour in a specific context of decision-making [69]. The generation of knowledge from the available data and information for making the policies and planning at the strategic and network levels of decision-making of highway infrastructure projects requires data utilisation quality. Dimensions of data quality associated with the pragmatic level include appropriateness, value-addition, reputation, relevancy, and usefulness [68]. Contextual features of pragmatic concerns are related to dimensions of relevance and utility of data and information for making decisions. As a dimension, reputation focuses on the user's expectations of data utility. The value-addition dimension aims to comprehend the user intent. These facets concern the data's compatibility with the challenging job. Related data quality dimensions are concerned with the intended application, i.e., how data would be utilised in connection to the current issue [70], for instance, value addition as a data quality dimension that refers to the extent to which data are valuable and add value to the organisation or individual stakeholders using it. In the context of highway data, it could use data analytics and machine learning algorithms to identify patterns and trends in data that are not immediately apparent. This could help highway agencies to identify areas of the highway system that require additional investment or maintenance and to prioritise their efforts accordingly.

Consequently, each semiotic level handles certain data quality and communication concerns. Understanding the overall data utilisation of highway infrastructure data stored in the data lake for making decisions at each decision-making level depends on the quality dimensions of the semiotic levels [32]. Within each semiotic level, it is crucial to identify the data quality requirements of decision-makers at their respective decision-making levels. For instance, strategic-level decision-makers focus on the utility of data and information for making effective policies throughout the organisation. Similarly, the other decision-making levels also required their specific data quality according to the requirement of decision-makers. Table 2 shows the data quality dimensions and the perspectives of dimensions along with the semiotic framework categories.

Applying a semiotic framework can be considered one of the philosophical approaches to studying data and its quality. In a semiotic framework, a top-down approach involves starting with high-level concepts or theories and breaking them into their constituent parts to understand how they work. In terms of the decision-making hierarchy, NHAI also follows a top-down approach. The higher officials make the authority's decisions at the top of the organisational structure and then communicate to the lower-level employees for implementation. Overall, by using a semiotic framework for data quality assessment, NHAI can ensure that its decision-making processes are informed by high-quality data that are relevant, accurate, and consistent. This can help to improve the efficiency and effectiveness of NHAI's operations and ensure that its highway and road networks are developed and maintained to the highest standard. However, the semiotic perspective has not become popular among researchers and practitioners to date [71]. The present study uses semiotic categories to describe the highway infrastructure data quality, specifically to identify the data quality dimensions to assess the data quality for effective decision-making [29]. Presently, no research has been reported to comprehend the link between data quality dimensions and highway infrastructure data about the semiotic levels that represent them.

**Table 2.** Data quality (DQ) dimensions and perspectives as per the semiotic framework levels.

| Semiotic Levels | DQ Dimensions | DQ Dimensions Perspective |
| --- | --- | --- |
| Empiric = It addresses issues that arise when data are utilised repeatedly. This level focuses on developing means of communication and data handling. | Accessibility | Accessibility implies that data must be accessible, obtainable, or retrievable when necessary for data to be accessible. |
| | Timeliness | Timeliness is concerned with the age of data and whether data are current. It is achieved if the recorded value is not out of date. |
| | Security | As a dimension, security involves securing data and limiting access to it. |
| Syntactic = It focuses on the structures and formats of data. It deals with the physical form of data rather than their content. | Accuracy | The accuracy dimension is concerned with the conformity of the recorded value with the actual value. It implies that data are accurate, flawless, trustworthy, and error-free. |
| | Completeness | Completeness concerns capturing all values for a specific variable and preventing data loss. It implies that the data must have adequate breadth, depth, and scope for the given task. |
| | Conciseness | Conciseness is a well-organised, concise, and condensed representation of data. |
| | Consistency | Consistency is achieved when data are represented in the same format, are compatible with previous data, and are represented consistently. |
| | Ease of operation | Ease of operation implies that data are manipulatable, integrated, customised, and utilised for multiple purposes. It is similar to flexibility. |
| | Integrity | Integrity measures correctness and consists of semantic and physical integrity. Semantic integrity measures consistency and completeness concerning the rule of the description language. Physical integrity measures the correctness of implementation details. |
| | Structure | Format or structure implies that data are in the correct format and structure. |
| Semantic = At the semantic level, dimensions are connected with information rather than data. Information is selected data to which meaning has been assigned in a particular context. It is concerned with meaning. | Ambiguity | Ambiguity arises due to improper representation and is when data can be interpreted in more than one way. |
| | Believability | Believability is concerned with whether data can be believed or regarded as credible. |
| | Interpretability | Interpretability means that data should be interpreted; that is, it should be defined clearly and represented appropriately. |
| | Definition | Meaningfulness or definition is concerned with the interpretation of data. The failure of this dimension results in meaningless data. |
| | Reliability | Reliability in terms of concepts drawn from the field of quality control. |
| | Understandability | Understandability concerns whether data are clear, readable, unambiguous, and easily comprehendible. |
| | Validity | Data are valid when verified as genuine and satisfying appropriate standards related to other dimensions. |
| Pragmatic = It focuses on how individuals use information. It concerns the relationship between data, information, and behaviour in each context. | Appropriateness | Appropriateness as a data quality dimension means that data must be appropriate to the task at hand. |
| | Relevant | Relevancy is concerned with the applicability of data to the task at hand. It is a crucial dimension if the data do not address the customer's needs and when the customer finds the data inadequate. |
| | Value | Value is added as a dimension that addresses the benefits and advantages of using data. |

## 4. Methodology

In order to meet the research objectives, this study was carried out in three steps. The first step was to identify the data quality dimensions of highway infrastructure using the semiotic framework. Most appropriate dimensions that were applicable to the highway infrastructure project were identified. In the second step, the questionnaire was prepared to the selected data quality dimensions finalised in step one. The responses were collected for the questionnaire from the highway infrastructure stakeholders. Finally, the responses were analysed in the third step to identify the critical dimensions and to rank them according to their mean value. These steps are described in detail in the following sub-sections.

Step 1: Identification of data quality dimensions of highway infrastructure data using the semiotic framework.

The semiotic framework consists of 43 data quality dimensions, as defined by Tejay. G. et al. [72]. These data quality dimensions are defined in the context of information system security. For the study of highway infrastructure projects' data quality, the dimensions were reduced to 20 out of 43 data quality dimensions, according to the relevant literature sources. A few dimensions have synonyms dimensions, and those were combined and considered a single dimension. The dimension accessibility, portability, and locatability have a similar meaning in the context of data quality; thus, we considered accessibility the primary dimension for assessing data quality. The established data quality dimensions were used to determine the data quality of highway infrastructure data. The 20 dimensions were personally reviewed with the three highway stakeholders; one chief general manager from the headquarters office responsible for network-level decision-making, one regional officer from the regional office responsible for the program and project selection level decision-making, and one project director from the project-implementing unit responsible for project-level decision-making were selected to verify the exhaustiveness/comprehensiveness of the selected data quality dimensions. Among the professionals, the chief general manager had more than ten years of experience, the project director had eight years of experience, and the regional officer had six years of experience in highway construction projects. The responses were not uniform, and the experience of the stakeholders was considered a limitation. Hence, all 20 dimensions were considered for the questionnaire survey for a comprehensive understanding of highway data quality for the effective use of data for effective decision-making.

Step 2: Data Collection

The questionnaire was designed based on the 20 data quality dimensions identified in Step 1. The survey targeted the National Highway of India decision-makers who utilised these data in decision-making. A pilot study was undertaken with 40 responses to test the language and understanding of the questionnaire. The responses are from the site engineers, deputy engineers, and managers from the project implementing units and regional offices. According to the suggestions from the pilot study, some significant changes were made to the questionnaire to make it more understandable for the stakeholders. The questionnaire was then shared via google forms with the 220 stakeholders. The stakeholders included the members, chief general manager, managers, regional officers, deputy general managers, and project directors. A total of 105 experts participated in the survey, which is a 48% response rate. The stakeholders with significant experience deal with the critical decisions from the National Highway Authority of India (NHAI), representing the strategic, network, program, project selection, and project levels, respectively. The questionnaire consists of three parts. Part 1 deals with the basic contact details, role, responsibility, and decision-making level in the decision-making hierarchy. The second part evaluates each attribute's importance at each decision-making level for the available data. The third part deals with ranking the dimensions, which states the priority of dimensions required in decision-making within the category of the semiotic framework.

A five-point Likert scale of 1 to 5 was used to record the decision-makers' level of importance of the data quality attribute. Here, '1' refers to "no importance," '2' refers to "low importance," '3' refers to "somehow important," '4' refers to "important," and '5' refers to "high importance" [73].

Step 3: Data Analysis

The data were analysed by using the software package SPSS 25. The analysis was carried out in two parts. The first part analysed the data's reliability using Cronbach's alpha test. It was found to be 0.875 at a 5% significance level greater than 0.5. Hence, it confirmed the reliability of the data. The dimensions were ranked according to their mean value to measure the consensus in the experts' opinions. However, when the mean values of two or more dimensions were identical, the dimensions with the lowest standard deviation were

placed higher [74]. The ranking of the dimensions based on the data collected through the questionnaire survey is shown in Table 3.

**Table 3.** Ranking of data quality dimensions.

| S. No. | Data Quality Dimensions | Mean | Std. Deviation | Rank |
|---|---|---|---|---|
| 1 | Accuracy | 4.52 | 0.64 | 1 |
| 2 | Accessibility | 4.40 | 0.70 | 2 |
| 3 | Completeness | 4.36 | 0.76 | 3 |
| 4 | Consistency | 4.28 | 0.67 | 4 |
| 5 | Timeliness | 4.27 | 0.68 | 5 |
| 6 | Structure | 3.90 | 0.90 | 6 |
| 7 | Ambiguity | 3.90 | 0.98 | 7 |
| 8 | Integrity | 3.83 | 0.85 | 8 |
| 9 | Value | 3.72 | 0.88 | 9 |
| 10 | Validity | 3.63 | 1.04 | 10 |
| 11 | Reliability | 3.58 | 1.12 | 11 |
| 12 | Appropriateness | 3.58 | 1.12 | 12 |
| 13 | Relevant | 3.58 | 0.85 | 13 |
| 14 | Definition | 3.50 | 0.81 | 14 |
| 15 | Interpretability | 3.38 | 0.96 | 15 |
| 16 | Understandability | 3.38 | 1.10 | 16 |
| 17 | Believability | 3.36 | 1.01 | 17 |
| 18 | Ease of Operation | 3.35 | 0.99 | 18 |
| 19 | Security | 3.35 | 0.90 | 19 |
| 20 | Conciseness | 3.29 | 0.83 | 20 |

### *4.1. Identification of Critical Data Quality Dimensions of Highway Infrastructure Data*

The descriptive statistical analysis did not yield a whole number for the mean value of the responses. Therefore, for the purpose of interpretation, the impact of each dimension on data quality can be considered to lie between the midpoints of two adjacent scales [75]. The importance of the dimensions about the mean value ($\mu$) greater than or equal to 4.5 was deemed to have a very high impact on the important data quality dimension. Similarly, the range of mean values $4.5 > \mu \geq 3.5$ was treated as having high importance; $3.5 > \mu \geq 2.5$ was treated as having moderate importance; $2.5 > \mu \geq 1.5$ was treated as having low importance; and mean values less than 1.5 were treated as having very low importance on data quality. In the study, the key data quality dimensions for assessing highway infrastructure data for effective decision-making dimensions were deemed to be those that were both very high and of high importance.

### *4.2. Importance of Data Quality Dimensions at Respective Decision-Making Levels*

Based on the questionnaire results, consideration was also given to the importance of dimensions. The data quality requirement may not be the same at all levels of decision-making. For instance, the project level focuses on the primary data collection and format. Hence, the dimensions critical at the project level are not critical at the remaining decision-making levels. Hence, the importance of dimensions at all decision-making levels was considered. The significance of data quality dimensions is determined at the strategic, network, program, project selection, and project levels of highway projects. Based on the ratings for the importance of dimensions at the decision-making level, decision-makers believe that all data quality attributes defined under the semiotic model are considered critical in data usage for information generation at all decision-making levels, with a rating of 4 out of 5. The context of data quality differs at each level of decision-making; consequently, data quality dimensions were determined, and the ranking of data quality dimensions was also calculated at each level of the semiotic framework, i.e., at the syntactic, pragmatic, empirical, and semantic levels.

### *4.3. Ranking of Data Quality Dimensions within the Semiotic Framework*

Along with the level of importance, the decision-makers also prioritise data quality dimensions in each category of the semiotic framework. The priority of data quality requirements has changed from stakeholder to stakeholder at each decision-making level. The semiotic framework comprised the syntactic, empiric, semiotic, and pragmatic categories, which deal with the structure, meaning, information, and knowledge of data characteristics [32]. The prioritisation of dimensions was also taken in the questionnaire survey. The responses to dimensions given by the respondents were converted into a rank using Henry Garrett's ranking technique [76]. This technique provides the change of orders of problems into numerical scores. The prime advantage of this technique over simple frequency distribution is that the dimensions are arranged based on their priority from the point of view of decision-makers. Garrett's formula for converting the ranks into the per cent position is shown below as Equation (1):

$$\text{Percent position} = 100 \times (R_{ij} - 0.5)/N_j \quad (1)$$

where $R_{ij}$ = rank given for ith dimension by jth decision-maker

$N_j$ = number of dimensions ranked by the jth individual.

The per cent position of each rank was converted into sources referring to the table given by Garrett and Woodworth [77]. For each factor, the scores of individual stakeholders were added together and divided by the total number of respondents for whom scores were added. These mean scores for all the dimensions were arranged in descending order; the dimensions were accordingly ranked.

## 5. Results and Discussion

This study identified and evaluated the key data quality dimensions for assessing the data quality of highway infrastructure for decision-making effectiveness. For this purpose, the study considered the critical dimensions throughout the highway infrastructure project, as well as the criticality of dimensions at each level of decision-making, as the data quality requirement varies at each level of decision-making. Using the ranking of dimensions shown in Table 3, the overall critical dimensions and preference of the dimensions for the overall project data were determined. Table 4 illustrates the significance of dimensions at each level of decision-making. From the analysis of Table 4, it is clear that the requirements for decision-makers are no longer the same but vary according to the respective hierarchical levels of decision-making.

The data quality dimensions listed in Table 3 are relevant to ensure the overall quality of data for highway infrastructure projects. Effective decision-making relies on the availability of high-quality data, and addressing each of these dimensions can help to ensure that the data used in decision-making are accurate, complete, consistent, and timely. Based on the mean scores in Table 3 and the analysis of Figure 3, the top five data quality dimensions are accuracy, accessibility, completeness, consistency, and timeliness. Ensuring that data are accurate involves verifying that they are correct and error-free. Accessibility involves making the data available and easily retrievable to authorised stakeholders. Completeness ensures that all required data elements are present and accounted for. Consistency involves verifying that the data are consistent with other data elements within the project. Timeliness ensures that the data are available when needed and up to date. Other dimensions listed in Table 3, such as relevance, interpretability, and believability, are also crucial for effective decision-making. Relevant data are essential to the decision-making process because they ensure that the data are related to the project's objectives. Interpretability ensures that data are presented in a way that is easy to understand. At the same time, believability involves ensuring that the data can be trusted and are not biased.

**Table 4.** Decision-maker's level of importance of dimensions.

| S. No. | Data Quality Attributes | Strategic Level | Network Level | Program Level | Project Selection Level | Project Level |
|---|---|---|---|---|---|---|
| 1 | Accuracy | 4.6 | 4.4 | 4.7 | 4.8 | 4.4 |
| 2 | Consistency | 4.8 | 4.6 | 4.1 | 4.3 | 4.2 |
| 3 | Completeness | 4.8 | 4.4 | 4.2 | 4.6 | 4.3 |
| 4 | Structure | 4.8 | 4.3 | 3.6 | 4.2 | 3.7 |
| 5 | Integrity | 4.8 | 4.0 | 3.6 | 4.1 | 3.7 |
| 6 | Conciseness | 3.0 | 3.4 | 3.4 | 3.3 | 3.2 |
| 7 | Ease of Operation | 3.6 | 3.3 | 3.3 | 3.2 | 3.4 |
| 8 | Accessibility | 4.2 | 4.6 | 4.7 | 4.4 | 4.3 |
| 9 | Timeliness | 4.4 | 4.6 | 4.4 | 3.9 | 4.2 |
| 10 | Security | 3.8 | 3.4 | 3.4 | 3.2 | 3.3 |
| 11 | Definition | 4.2 | 3.8 | 3.6 | 3.8 | 3.3 |
| 12 | Ambiguity | 4.8 | 4.4 | 4.2 | 4.1 | 3.5 |
| 13 | Believability | 3.6 | 3.0 | 3.2 | 4.0 | 3.4 |
| 14 | Interpretability | 3.4 | 3.4 | 3.2 | 3.7 | 3.4 |
| 15 | Reliability | 4.2 | 3.4 | 3.5 | 2.9 | 3.7 |
| 16 | Understandability | 2.8 | 3.3 | 3.4 | 3.4 | 3.4 |
| 17 | Validity | 3.0 | 4.0 | 3.6 | 3.0 | 3.7 |
| 18 | Relevant | 4.2 | 3.8 | 3.6 | 3.7 | 3.4 |
| 19 | Value | 4.2 | 3.8 | 3.9 | 4.0 | 3.6 |
| 20 | Appropriateness | 3.8 | 3.4 | 3.6 | 3.9 | 3.6 |

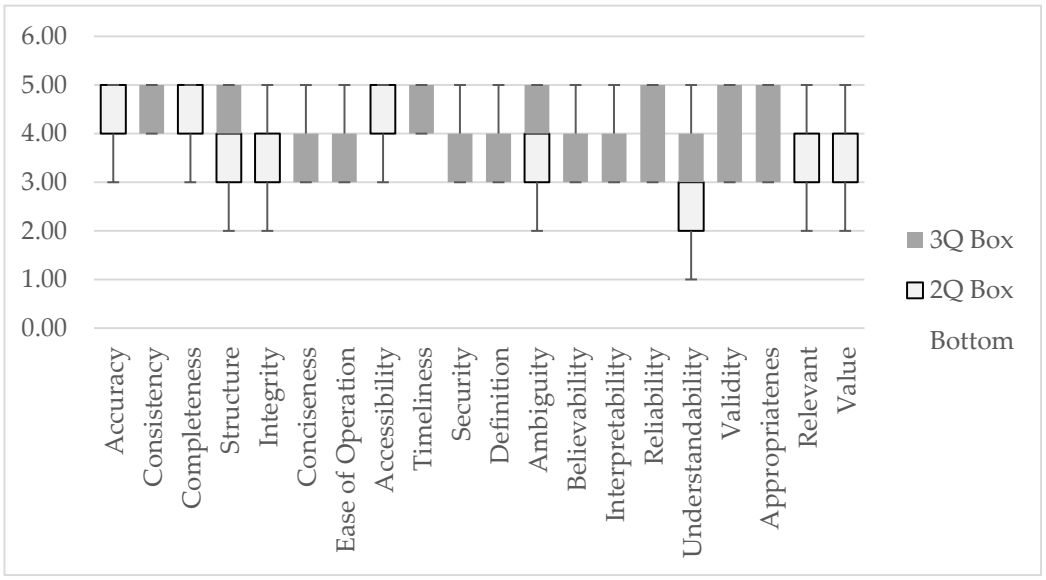


**Figure 3.** Semiotic levels and the data-information-knowledge hierarchy.

At the same time, the stakeholders consider conciseness, ease of operation, and security as the lowest priorities with low mean values. This might be due to the dimension conciseness that implies the compact representation of data, which would create a problem of understanding for all stakeholders for data usage in decision-making. The dimension ease-of-operation implies that data are manipulated and easily customised, which stakeholders feel could cause problems in decision-making if the data are easily manipulated. The dimension security implies keeping data secure and restricting access to the data, and the stakeholders feel the restricting of data would cause issues with the decision-making.

Figure 3 shows the box-and-whiskers plot of the data quality dimensions for assessing overall highway project data quality. It shows that the range of most of the dimensions is between 3 and 5, i.e., the responses from the decision-makers range from somehow important to high importance. Based on the data, it seems that accuracy, completeness, and

accessibility are considered to be the most critical dimensions of data quality by more than half of the decision-makers, as they have the highest median value of 5. The dimensions consistency, structure, integrity, timeliness, ambiguity, relevant, and value also have a relatively high median value of 4, indicating that they are still considered essential by many decision-makers. On the other hand, dimensions such as conciseness, ease of operations, security, definition, believability, understandability, validity, and appropriateness have a median value of 3, indicating that they are considered less critical by decision-makers. It is important to note that these findings are based on decision-makers' responses and reflect the objective of data quality measures. Nonetheless, they provide valuable insights into the perceived importance of different dimensions of data quality in the context of highway project data.

### *5.1. Critical Dimensions at Each Level of Decision-Making Hierarchy*

The importance of data quality dimensions at each decision-making level, such as strategic, network, program, project selection, and project decision levels of highway infrastructure projects, was also identified, along with key data quality dimensions for assessing the over-project data. Table 4 shows the description of assessment measures and a survey result on the level of importance for semiotic framework data quality attributes obtained from highway decision-makers, respectively. Based on the level of importance, decision-makers think that all data quality dimensions described within the semiotic framework are crucial for generating information at all levels of the decision-making hierarchy for highway infrastructure. However, the results indicate that the conciseness, ease of operational ability, and data understandability dimensions do not significantly influence decision-making processes at all levels of highway infrastructure decision-making, i.e., strategic, network, program, project selection, and project. This may be due to the absence of a system that facilitates the understanding of collected data at these levels. The collected data could be in various formats, including text, images, or numbers, and they could be used as input for decision-making. For data to be used as input, they must be clearly understood according to the judgment of the highway engineers. This may result from the insignificant use of project data at these levels or the continuation of decision-making processes due to limited project scope in the early stages of a project. Therefore, the dimensions with a rating of 4 out of 5 are regarded as crucial for generating information at all levels of the decision-making hierarchy.

#### 5.1.1. Strategic Level

At the strategic level, decision-makers are higher-level authorities, such as the chairman and division heads of NHAI. They deal with policies, guidelines, and the distribution of funds. At the strategic level of the decision hierarchy, the accuracy, consistency, completeness, structure, and integrity dimensions from the syntactic category; the accessibility and timeliness dimensions from the empiric category; the definition, reliability, and ambiguity dimensions from the semantic category; and the relevance and value dimensions from the pragmatic category are crucial decision-making dimensions.

#### 5.1.2. Network Level

At the network level, decision-makers, such as chief general managers, are responsible for determining priorities, developing programs, and determining project objectives. According to the analysis, the critical dimensions at the network level of the decision hierarchy are accuracy, consistency, completeness, structure, and integrity from the syntactic category; accessibility and timeliness from the empirical category; ambiguity from the semantic category; and validity from the pragmatic category. In addition, the network level is subdivided into two decision-making levels, including the program level and project selection level.

#### 5.1.3. Program Level

At the program level of decision-making, the critical dimensions are accuracy, consistency, and completeness from the syntactic category; accessibility and timeliness from the empiric category; and ambiguity from the semantic category. No dimensions from the pragmatic category are crucial for program-level decision-making. Program-level decision-making deals with the programming of the projects. For making decisions at the program level, the decision-makers focus on the form and structure of data, establishing means of communication and data handling and information on the data. The pragmatic category dimensions deal with the knowledge generated from the data, which is not much focused on decision-making at the program level.

#### 5.1.4. Project Selection Level

The project selection level addresses project selection, safety improvement, and traffic control studies at the regional office level. From the syntactic category, the critical dimensions are accuracy, consistency, completeness, structure, and integrity. From the empiric category, the critical dimension is only accessibility. From the semantic category, the critical dimensions are ambiguity and believability. From the pragmatic category, the dimension value is only critical in analysing the questionnaire data.

#### 5.1.5. Project Level

Project-level decisions involve the project director, designers, maintenance engineers, schedulers, and many other engineers responsible for project implementation at the project-implementing unit. For effective decision-making, dimensions such as accuracy, consistency, and completeness from the syntactic level; and accessibility and timeliness from the empiric level are considered critical out of all 20 data quality dimensions. At the project level, its primary concern is data generation, the physical form of data generation, and storage for information generation. Therefore, the dimensions in the semantic and pragmatic categories that deal with information and knowledge generation are not of as high importance as the syntactic and empirics level at the project level.

### *5.2. Ranking of Dimensions within the Semiotic Framework Categories*

Garrett's ranking technique was used to analyse various dimensions for ranking the dimensions within the semiotic framework levels. The decision-makers were asked to rank the dimensions within the framework to understand their preferences for data quality dimensions within the semiotic framework. The semiotic framework comprised syntactic, empiric, semiotic, and pragmatic categories, which deal with the structure, meaning, information, and knowledge of data characteristics [32]. Before ranking the dimensions within the semiotic framework levels, the percentage position for the ranks and their corresponding Garrett value were calculated using Equation (1), as shown in Table 5. The total score was calculated for factors by multiplying the number of stakeholders ranking that dimension (Garrett and Woodworth [76]).

**Table 5.** Percentage position and Garrett value for rank 1 to 7.

| Ranks | Percentage Position | Garret Score |
|---|---|---|
| 1 | 7.14 | 79 |
| 2 | 21.43 | 66 |
| 3 | 35.71 | 57 |
| 4 | 50.00 | 50 |
| 5 | 64.29 | 43 |
| 6 | 78.57 | 34 |
| 7 | 92.86 | 22 |

#### 5.2.1. Syntactics Category

The syntactic category focused on data structures and formats, i.e., the physical form of data rather than its content. In order to understand the data quality requirements in terms of the syntactic category of the data being used by decision-makers, the Garrett ranking technique was used, and the dimensions were ranked as shown in Table 6. Based on the Garrett mean values, stakeholders ranked dimensions as accuracy, consistency, completeness, structure, integrity, conciseness, and ease of operation. This is because the syntactic category is primarily concerned with the physical form rather than the data content; the decision-makers prioritised accuracy over the ease of operation dimension [35]. It is shown that the priority of data quality requirements changed from stakeholder to stakeholder. Hence, we considered most of the responder's ranking as the topmost ranked and followed for other dimensions, as shown in Table 6.

**Table 6.** Ranking of dimensions within the syntactic category of the semiotic framework.

| S. No. | Factors | Rank | | | | | | | Total Number of Stakeholders | Total Score | Total Mean | Rank |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | | | |
| 1 | Accuracy | 52 | 16 | 6 | 6 | 11 | 9 | 5 | 105 | 6695 | 63.76 | 1 |
| 2 | Consistency | 12 | 40 | 18 | 19 | 3 | 6 | 7 | 105 | 6051 | 57.63 | 2 |
| 3 | Completeness | 12 | 25 | 42 | 6 | 5 | 5 | 10 | 105 | 5897 | 56.16 | 3 |
| 4 | Structure | 6 | 5 | 15 | 45 | 24 | 6 | 4 | 105 | 5233 | 49.84 | 4 |
| 5 | Integrity | 8 | 6 | 9 | 16 | 49 | 10 | 7 | 105 | 4942 | 47.07 | 5 |
| 6 | Conciseness | 8 | 7 | 5 | 6 | 8 | 44 | 27 | 105 | 4113 | 39.17 | 6 |
| 7 | Ease of Operation | 7 | 6 | 10 | 7 | 5 | 25 | 45 | 105 | 3924 | 37.37 | 7 |

#### 5.2.2. Empiric Category

The empiric category dealt with the issues that arise when data are utilised repeatedly. This category focused on developing means of communication and data handling. Based on the percentage position and Garrett's mean value, the dimensions were ranked as accessibility, timeliness, and security within the empiric category. The dimension accessibility of data was given the highest priority over the security of the data dimension. Table 7 shows the percentage position and Garret score for the ranks as per Equation (1), while Table 8 shows the ranking of the dimensions based on the Garrett mean value. Accessibility refers to how easily users can access data. This includes factors such as the availability of the data, the ease of retrieving them, and the format in which they are presented. Timeliness refers to how up-to-date and relevant the data are. This includes factors such as the frequency of updates and how quickly they are made available. Security refers to data protection from unauthorised access, modification, or disclosure. This includes factors such as the level of encryption used, the strength of access controls, and the measures in place to prevent data breaches. By ranking these dimensions based on their importance, organisations can prioritise their efforts to improve information quality. However, it is essential to note that the relative importance of each dimension may vary depending on the specific context and the users' needs.

**Table 7.** Percentage position and Garrett value for rank 1 to 3.

| Ranks | Percentage Position | Garret Score |
|---|---|---|
| 1 | 16.67 | 69.00 |
| 2 | 50.00 | 50.00 |
| 3 | 83.33 | 31.00 |

#### 5.2.3. Semantic Category

The semantic category deals with the dimensions connected with information rather than data. Information is selected data to which meaning has been assigned in a particular context. It is concerned with meaning. Within the semantic category, the dimensions were

ranked as ambiguity, definition, believability, interpretability, reliability, understandability, and data validity, as shown in Table 9. The dimension ambiguity was prioritised over other data dimensions within the category. This might be because the data should be clear for understanding if any ambiguity in data can lead to significant challenges in decision-making for highway projects, potentially resulting in poor design, construction, and long-term consequences for the environment and public safety. The Garrett mean values that were calculated using Equation (1) and the percentage position values are shown in Table 5.

**Table 8.** Ranking of dimensions within the empiric category of the semiotic framework.

| S. No. | Factors | Rank | | | Total Number of Stakeholders | Total Score | Total Mean | Rank |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | | | | |
| 1 | Accessibility | 46 | 33 | 26 | 105 | 5630 | 53.62 | 1 |
| 2 | Timeliness | 38 | 47 | 20 | 105 | 5592 | 53.26 | 2 |
| 3 | Security | 21 | 25 | 59 | 105 | 4528 | 43.12 | 3 |

**Table 9.** Ranking of dimensions within the semantic category of the semiotic framework.

| S. No. | Factors | Rank | | | | | | | Total Number of Stakeholders | Total Score | Total Mean | Rank |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | | | |
| 1 | Ambiguity | 51 | 16 | 7 | 6 | 10 | 9 | 6 | 105 | 6652 | 63.35 | 1 |
| 2 | Definition | 12 | 25 | 41 | 7 | 5 | 5 | 10 | 105 | 5890 | 56.10 | 2 |
| 3 | Believability | 7 | 6 | 10 | 7 | 6 | 24 | 45 | 105 | 3933 | 37.46 | 3 |
| 4 | Interpretability | 8 | 7 | 9 | 16 | 48 | 10 | 7 | 105 | 4965 | 47.29 | 4 |
| 5 | Reliability | 12 | 39 | 17 | 19 | 5 | 6 | 7 | 105 | 6014 | 57.28 | 5 |
| 6 | Understandability | 8 | 7 | 5 | 6 | 8 | 45 | 26 | 105 | 4125 | 39.29 | 6 |
| 7 | Validity | 7 | 5 | 16 | 44 | 23 | 6 | 4 | 105 | 5276 | 50.25 | 7 |

#### 5.2.4. Pragmatic Category

The pragmatic category focused on how individuals use information. It concerns the relationship between data, information, and behaviour in each context. For ranking the dimensions within the pragmatic category, the percentage position of the ranks was calculated using Equation (1), as shown in Table 7. The dimension value of data was given the highest priority over the other dimensions, such as relevant and appropriateness. In the context of highway stakeholders, the dimension value is crucial because it determines the extent to which the data can inform decision-making about highway infrastructure projects, budgeting, and maintenance. Although the dimension appropriateness is critical, it was ranked third in this context because it is a prerequisite for both relevance and value. As per the Garrett ranking technique, the dimensions were ranked as value, relevant, and appropriateness, respectively, as shown in Table 10.

**Table 10.** Ranking of dimensions within the pragmatic category of semiotic framework.

| S. No. | Factors | Rank | | | Total Number of Stakeholders | Total Score | Total Mean | Rank |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | | | | |
| 1 | Relevant | 33 | 49 | 23 | 105 | 5440 | 51.81 | 2 |
| 2 | Value | 44 | 30 | 31 | 105 | 5497 | 52.35 | 1 |
| 3 | Appropriateness | 28 | 26 | 51 | 105 | 4813 | 45.84 | 3 |

The dimensions were ranked to understand the decision-makers' data quality requirements for decision-making at the individual decision-making levels [30]. As the level of decision-making in the organisation changes, the priority of data quality also changes. At the strategic level, decision-makers focus on policymaking, which could be implemented

throughout the organisation. Hence, the data quality requirements at the strategic level differ at the network and project levels. It is important to note that this study utilised semiotic-based quality dimensions to assess data quality at different decision-making levels from the data users' perspective. This proactive assessment of the highway management decision-making hierarchy allows data collectors to determine the level of data quality requirements of highway infrastructure managers and potential decision-makers in a more integrated manner. It allows highway agencies' data management teams to identify the causes behind minimal data usage to improve the quality of generating information and supporting decisions.

## 6. Conclusions

This research was conducted in a multidisciplinary framework that included three primary fields: data quality, big data, and highway infrastructure project data. Even though data quality has been a well-studied topic for the past two decades, the precise terminology for data quality aspects is still lacking. Digitalisation and data management in construction, particularly highway infrastructure, is a developing topic in India, with a scant prior study focusing on data quality. Using data quality dimensions as part of data governance projects is undoubtedly crucial, as it ensures that data users and stakeholders may derive the most significant benefit from data usage. The research discussed in this paper aims to investigate a framework in which data quality dimensions could be more important within the context of highway infrastructure projects in the construction sector. The semiotic framework was adopted from the literature review of various data quality frameworks for this study to establish data quality dimensions for highway infrastructure data. The systematic literature review, semiotic framework, and Garrett ranking were chosen as research methods because of the increasing novelty of vast quantities of data quality and highway infrastructure data, as well as the impracticality of implementing other research methods due to geographical, legal, ethical, and organisational constraints.

Accuracy, accessibility, and consistency are well-discussed data quality dimensions that are supported by the results. Based on this research, the data quality dimensions of completeness and timeliness were added to the three previously mentioned data quality dimensions to produce a list of the five most appropriate data quality dimensions for highway infrastructure data in the construction industry. Considering the results of the semiotic framework of the hierarchical data quality dimensions for the overall highway project data, the contextual category of data quality dimensions was considered to be the most crucial for evaluating data quality. This is easily explained by the breadth of the three domains involved (i.e., data quality, big data, and highway infrastructure data), where thousands of unique data applications used in the highway infrastructure database are possible. Thus, each application's probability of selecting different data quality dimensions increases.

The current research study provides a ranking of the most critical data quality dimensions in the specific context of highway infrastructure projects, as shown in Table 3. This is one of the first studies within this field to use the semiotic framework to achieve this. This research study also considered the level of importance at each decision-making level of the hierarchy, as shown in Table 4. Considering the very contextual nature of data quality, different contexts would be expected to produce a different list of the most critical data quality dimensions. Thus, the study also provided the ranking of the dimensions within the semiotic framework categories using the Garrett raking technique to understand the priorities of the stakeholders.

The comparatively little amount of literature, and more significantly, publications with the perspective of highway infrastructure data, is one of the most significant limitations of this study. Planned are additional research methods that could be applied to the same corpus of literature, with the primary objective of reducing the amount of author bias introduction when evaluating the significance of the other data quality frameworks.

This study serves as a foundation for further research by the authors in highway infrastructure to assess overall data usage in terms of significant data quality using data quality dimensions as features for assessing the current data quality satisfaction levels at decision-making levels from the data users' perspective. There is a need for agencies and data management teams to assess the root cause of the minimal usage of data to improve the quality of generating information and supporting decisions, and they are also required to show the interdependency of various decisions in the final output of a project and address the potential data users' requirements. In ongoing research, the semiotic framework provides a theoretical foundation for developing an instrument, i.e., data quality dimensions, to access the subjective quality of highway project data. The development of quantitative indices for each data quality dimension to quantify the quality would eventually help to develop the decision-making competency of decision-makers. This would help the organisation in the effective execution of projects without delaying the projects and avoid losses due to wrong decisions. By using data quality dimensions as features for machine learning algorithms, further work will distinguish quality data from non-quality data from very large streams of highway datasets. Finally, the ten main data quality dimensions identified serve as a foundation for determining which machine learning algorithms might identify data usage more effectively. Following this, a computationally efficient method for optimum data usage will be designed to use data effectively.

*Practical Engineering and Real-World Applications of Semiotic Framework*

The semiotic framework for assessing data quality is a theoretical framework that analyses data in terms of its essential components: syntactics, pragmatics, empirics, and semantics. This strategy has several real-world and practical engineering applications, such as data integration, business intelligence, data mining, data governance, and data visualisation. In the construction sector context, the semiotic framework of data quality assessment is used in evaluating building designs. Architects and engineers may use this framework to evaluate the accuracy and completeness of their building designs by analysing the signs and symbols used to represent the different design aspects. This may help them uncover design inconsistencies or errors and make the necessary adjustments before construction begins. In engineering applications of the construction industry, the semiotic framework of data quality may be used in several ways, including quality assurance, risk assessment, and compliance. Throughout the project lifecycle of a construction project, a substantial quantity of data must be gathered and evaluated for quality assurance purposes. The semiotic framework may be used to verify that the obtained data are correct and trustworthy, therefore guaranteeing that the project is on track and satisfies all objectives. The semiotic framework may be used to evaluate the risk associated with specific construction activities. Engineers can make informed decisions and reduce the likelihood of accidents or errors by evaluating the data quality used to evaluate risk. The construction industry is highly regulated, and businesses must adhere to various standards and regulations. The semiotic framework can ensure the accuracy and dependability of the data used to demonstrate compliance, thereby reducing the risk of fines.

In conclusion, the semiotic data quality framework has numerous practical engineering applications in the construction industry. Specific to highways, data quality dimensions are indispensable for planning and design, asset management, safety and emergency response, performance measurement, and policy and decision-making. By ensuring the quality of their data, transportation agencies can make more informed decisions, allocate resources more efficiently, and provide more effective transportation systems. Using this framework, architects, engineers, and other construction professionals can guarantee that the data they use is error-free, resulting in improved project outcomes and reduced risk.

**Author Contributions:** Conceptualisation, methodology, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, C.M.K.; writing—review, K.N.J. and K.R.; writing—editing, C.M.K.; supervision, K.N.J. and K.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References


1. Press Information Bureau. NHAI Becomes the First Construction Sector Organisation to Go Fully Digital. 2020. Available online: https://pib.gov.in/indexd.aspx (accessed on 12 June 2020).
2. Snyder, J.; Menard, A.; Spare, N. *Big Data = Big Questions for the Engineering and Construction Industry*; White Paper; First Myanmar Investment (FMI): Yangon, Myanmar, 2019.
3. Thomas, E.; Schott, P.; Bowman, J.; Synder, J.; Spare, N. *Construction Disconnected: Rethinking the Management of Project Data and Mobile Collaboration to Reduce Costs and Improve Schedules*; Plan Grid; First Myanmar Investment (FMI): Yangon, Myanmar, 2018.
4. Deibe, D.; Amor, M.; Doallo, R. Big Data Geospatial Processing for Massive Aerial LiDAR Datasets. *Remote Sens.* **2020**, *12*, 719. [CrossRef]
5. Pierce, L.M.; McGovern, G.; Zimmerman, K.A. *Practical Guide for Quality Management of Pavement Condition Data Collection*; FHWA: Washington, DC, USA, 2013.
6. Oh, E.; Lee, H. An Imbalanced Data Handling Framework for Industrial Big Data Using a Gaussian Process Regression-Based Generative Adversarial Network. *Symmetry* **2020**, *12*, 669. [CrossRef]
7. Zhang, Y.; Kim, C.-W.; Zhang, L.; Bai, Y.; Yang, H.; Xu, X.; Zhang, Z. Long Term Structural Health Monitoring for Old Deteriorated Bridges: A Copula-ARMA Approach. *Smart Struct. Syst. Int. J.* **2020**, *25*, 285–299.
8. Zhang, Z.; Liu, M.; Liu, X.; Wang, X.; Zhang, Y. Model Identification of Durability Degradation Process of Concrete Material and Structure Based on Wiener Process. *Int. J. Damage Mech.* **2021**, *30*, 537–558. [CrossRef]
9. Batini, C.; Rula, A.; Scannapieco, M.; Viscusi, G. From data quality to bid data quality. *J. Database Manag.* **2015**, *26*, 60–82.
10. Lee, I. Big Data: Dimensions, Evolution, Impacts, and Challenges. *Bus. Horiz.* **2017**, *60*, 293–303. [CrossRef]
11. Sadiq, S.; Papotti, P. Big Data Quality-Whose Problem Is It? In Proceedings of the IEEE 32nd International Conference on Data Engineering (ICDE), Helsinki, Finland, 16–20 May 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1446–1447.
12. Saha, B.; Srivastava, D. Data Quality: The Other Face of Big Data. In Proceedings of the IEEE 30th International Conference on Data Engineering, Chicago, IL, USA, 31 March–4 April 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 1294–1297.
13. Taleb, I.; el Kassabi, H.T.; Serhani, M.A.; Dssouli, R.; Bouhaddioui, C. Big Data Quality: A Quality Dimensions Evaluation. In Proceedings of the 2016 Intl IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld), Toulouse, France, 18–21 July 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 759–765.
14. Elouataoui, W.; el Alaoui, I.; el Mendili, S.; Gahi, Y. An Advanced Big Data Quality Framework Based on Weighted Metrics. *Big Data Cogn. Comput.* **2022**, *6*, 153. [CrossRef]
15. Cai, L.; Zhu, Y. The Challenges of Data Quality and Data Quality Assessment in the Big Data Era. *Data Sci. J.* **2015**, *14*, 2. [CrossRef]
16. Ghasemaghaei, M.; Calic, G. Can Big Data Improve Firm Decision Quality? The Role of Data Quality and Data Diagnosticity. *Decis. Support Syst.* **2019**, *120*, 38–49. [CrossRef]
17. Haug, A.; Zachariassen, F.; van Liempd, D. The Costs of Poor Data Quality. *J. Ind. Eng. Manag.* **2011**, *4*, 168–193.
18. Laranjeiro, N.; Soydemir, S.N.; Bernardino, J. A Survey on Data Quality: Classifying Poor Data. In Proceedings of the IEEE 21st Pacific Rim International Symposium on Dependable Computing (PRDC), Zhangjiajie, China, 18–20 November 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 179–188.
19. Sadiq, S.; Yeganeh, K.; Indulska, M. Cross-Disciplinary Collaborations in Data Quality Research. *ECIS Proc.* **2011**, *78*, 1–13.
20. Sidi, F.; Ishak, I.; Affendey, L.S.; Jaya, M.I.; Suriani Affendey, L.; Jabar, M.A. A Review of Data Quality Research in Achieving High Data Quality Within Organization. *J. Theor. Appl. Inf. Technol.* **2017**, *30*, 12. [CrossRef]
21. Yonke, C.L.; Walenta, C.; Talburt, J.R. *The Job of the Information/Data Quality Professional*; International Association for Information and data Quality (IAIDQ): Baltimore, MD, USA, 2011.
22. Ballou, D.P.; Pazer, H.L. Modeling Data and Process Quality in Multi-Input, Multi-Output Information Systems. *Manag. Sci.* **1985**, *31*, 150–162. [CrossRef]
23. Ballou, D.; Wang, R.; Pazer, H.; Tayi, G.K. Modeling Information Manufacturing Systems to Determine Information Product Quality. *Manag. Sci.* **1998**, *44*, 462–484. [CrossRef]
24. Wand, Y.; Wang, R.Y. Anchoring Data Quality Dimensions in Ontological Foundations. *Commun. ACM* **1996**, *39*, 86–95. [CrossRef]
25. English, L.P. *Information Quality Applied: Best Practices for Improving Business Information, Processes and Systems*; Wiley Publishing: Hoboken, NJ, USA, 2009; ISBN 047013447X.
26. Redman, T.C. *Data Quality for the Information Age*; Artech House, Inc.: Norwood, MA, USA, 1997; ISBN 0890068836.
27. Coleman, C. *Managing Information Quality: Increasing the Value of Information in Knowledge-Intensive Products and Processes*; Springer: Berlin/Heidelberg, Germany, 2007.

28. Tan, S.G.; Cheng, D. Quality Assurance of Performance Data for Pavement Management Systems. In *Design, Analysis, and Asphalt Material Characterization for Road and Airfield Pavements*; ASCE: Reston, VA, USA, 2014; pp. 163–169.
29. Price, R.; Shanks, G. Chapter 4 Data Quality and Decision Making. In *Handbook on a Decision Support System*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 65–82.
30. Samitsch, C. *Data Quality and Its Impacts on Decision-Making: How Managers Can Benefit from Good Data*; Springer: Berlin/Heidelberg, Germany, 2014; ISBN 3658082003.
31. Krogstie, J. A Semiotic Approach to Data Quality. In Proceedings of the Lecture Notes in Business Information Processing; Springer: Berlin/Heidelberg, Germany, 2013; Volume 147, pp. 395–410.
32. Huang, H. Big Data to Knowledge–Harnessing Semiotic Relationships of Data Quality and Skills in Genome Curation Work. *J. Inf. Sci.* **2018**, *44*, 785–801. [CrossRef]
33. Long, J.A.; Seko, C.E. A New Method for Database Data Quality Evaluation at the Canadian Institute for Health Information (CIHI). In Proceedings of the 7th International Conference on Information Quality (IQ 2002), Tempe, AZ, USA, 24–28 February 2002; pp. 238–250.
34. Lee, Y.W.; Strong, D.M.; Kahn, B.K.; Wang, R.Y. AIMQ: A Methodology for Information Quality Assessment. *Inf. Manag.* **2002**, *40*, 133–146. [CrossRef]
35. Pipino, L.L.; Lee, Y.W.; Wang, R.Y.; Yang, R.Y. Data Quality Assessment. *Commun. ACM* **2002**, *45*, 211–218. [CrossRef]
36. Sukumar, S.R.; Natarajan, R.; Ferrell, R.K. Quality of Big Data in Health Care. *Int. J. Health Care Qual. Assur.* **2015**, *28*, 621–634. [CrossRef]
37. Jankalová, M.; Jankal, R. How to Characterise Business Excellence and Determine the Relation between Business Excellence and Sustainability. *Sustainability* **2020**, *12*, 6198. [CrossRef]
38. Wang, R.Y. A Product Perspective on Total Data Quality Management. *Commun. ACM* **1998**, *41*, 58–65. [CrossRef]
39. Del Pilar Angeles, M.; García-Ugalde, F. A Data Quality Practical Approach. *Int. J. Adv. Softw.* **2009**, *1*, 259–299.
40. Sebastian-Coleman, L. *Measuring Data Quality for Ongoing Improvement: A Data Quality Assessment Framework*; Elsevier: Walthman, MA, USA, 2012; ISBN 0123977541.
41. Vaziri, R.; Mohsenzadeh, M.; Habibi, J. TBDQ: A Pragmatic Task-Based Method to Data Quality Assessment and Improvement. *PLoS ONE* **2016**, *11*, e0154508. [CrossRef] [PubMed]
42. Valverde, C.; Marotta, A.; Panach, J.I.; Vallespir, D. Towards a Model and Methodology for Evaluating Data Quality in Software Engineering Experiments. *Inf. Softw. Technol.* **2022**, *151*, 107029. [CrossRef]
43. Liebenau, J.; Backhouse, J. *Understanding Information: An Introduction*; Palgrave Macmillan: London, UK, 1990; ISBN 0333536800.
44. Azeroual, O.; Jha, M.; Nikiforova, A.; Sha, K.; Alsmirat, M.; Jha, S. A Record Linkage-Based Data Deduplication Framework with DataCleaner Extension. *Multimodal Technol. Interact* **2022**, *6*, 27. [CrossRef]
45. Abedjan, Z.; Chu, X.; Deng, D.; Fernandez, R.C.; Ilyas, I.F.; Ouzzani, M.; Papotti, P.; Stonebraker, M.; Tang, N. Detecting Data Errors: Where Are We and What Needs to Be Done? *Proc. VLDB Endow.* **2016**, *9*, 993–1004. [CrossRef]
46. Wang, R.Y.; Strong, D.M. Beyond Accuracy: What Data Quality Means to Data Consumers. *J. Manag. Inf. Syst.* **1996**, *12*, 5–33. [CrossRef]
47. Crosby, P.B. *Quality Is Free: The Art of Making Quality Certain*; Signet Book: West Bengal, India, 1980; Volume 2247, ISBN 0451622472.
48. Fu, Q.; Easton, J.M. Understanding Data Quality: Ensuring Data Quality by Design in the Rail Industry. In Proceedings of the IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 11–14 December 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 3792–3799.
49. Ramasamy, A.; Chowdhury, S. Big Data Quality Dimensions: A Systematic Literature Review. *J. Inf. Syst. Technol. Manag.* **2020**, *17*. [CrossRef]
50. Madnick, S.; Zhu, H. Improving Data Quality through Effective Use of Data Semantics. *Data Knowl. Eng.* **2006**, *59*, 460–475. [CrossRef]
51. English, L.P. *Improving Data Warehouse and Business Information Quality: Methods for Reducing Costs and Increasing Profits*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1999; ISBN 0471253839.
52. Redman, T.C. *Data Quality: The Field Guide*; Digital Press: Oxford, UK, 2001; ISBN 1555582516.
53. Batini, C.; Cappiello, C.; Francalanci, C.; Maurino, A. Methodologies for Data Quality Assessment and Improvement. *ACM Comput. Surv.* **2009**, *41*, 1–52. [CrossRef]
54. Gao, B.; Zhou, Q.; Deng, Y. BIM-AFA: Belief Information Measure-Based Attribute Fusion Approach in Improving the Quality of Uncertain Data. *Inf. Sci.* **2022**, *608*, 950–969. [CrossRef]
55. Madnick, S.; Wang, R.; Dravis, F.; Chen, X. Improving the Quality of Corporate Household Data: Current Practices and Research Directions. *SSRN Electron. J.* **2000**, 365180. [CrossRef]
56. Redman, T.C. Improve Data Quality for Competitive Advantage. *MIT Sloan Manag. Rev.* **1995**, *36*, 99.
57. Hassenstein, M.J.; Vanella, P. Data Quality—Concepts and Problems. *Encyclopedia* **2022**, *2*, 498–510. [CrossRef]
58. Gabr, M.I.; Helmy, Y.M.; Elzanfaly, D.S. Data Quality Dimensions, Metrics, and Improvement Techniques. *Future Comput. Inf. J.* **2021**, *6*, 25–44. [CrossRef]
59. Jesiļevska, S. Data Quality Dimensions to Ensure Optimal Data Quality. *Rom. Econ. J.* **2017**, *20*, 63.
60. Gyulgyulyan, E.; Ravat, F.; Astsatryan, H.; Aligon, J. Data Quality Impact in Business Inteligence. In Proceedings of the 2018 Ivannikov Memorial Workshop, (IVMEM), Yerevan, Armenia, 3–4 May 2018; IEEE: Piscataway, NJ, USA, 2019; pp. 47–51.

61. Loshin, D. *Enterprise Knowledge Management: The Data Quality Approach*; Morgan Kaufmann: London, UK, 2001; ISBN 0124558402.
62. Cappiello, C.; Ficiaro, P.; Pernici, B. HIQM: A Methodology for Information Quality Monitoring, Measurement, and Improvement. In Proceedings of the International Conference on Conceptual Modeling; Springer: Berlin/Heidelberg, Germany, 2006; pp. 339–351.
63. Batini, C.; Cabitza, F.; Cappiello, C.; Francalanci, C.; di Milano, P. A Comprehensive Data Quality Methodology for Web and Structured Data. In Proceedings of the 2006 1st International Conference on Digital Information Management, Bangalore, India, 6–8 December 2006; IEEE: Piscataway, NJ, USA, 2006; pp. 448–456.
64. Moraga, C.; Moraga, M.Á.; Caro, A.; Calero, C. SPDQM: SQuaRE-Aligned Portal Data Quality Model. In Proceedings of the 9th International Conference on Quality Software, QSIC, Jeju, Republic of Korea, 24–25 August 2009.
65. Carlo, B.; Daniele, B.; Federico, C.; Simone, G. A Data Quality Methodology for Heterogeneous Data. *Int. J. Database Manag. Syst.* **2011**, *3*, 60–79. [CrossRef]
66. Falkenberg, E.D. *A Framework of Information System Concepts*; The FRISCO Report (Web Edition); University of Leiden, Department of Computer Science: Leiden, The Netherlands, 1998; ISBN 3901882014.
67. Kahn, M.G.; Raebel, M.A.; Glanz, J.M.; Riedlinger, K.; Steiner, J.F. A Pragmatic Framework for Single-Site and Multisite Data Quality Assessment in Electronic Health Record-Based Clinical Research. *Med. Care* **2012**, *50*, S21–S29. [CrossRef]
68. Knoke, D.; Yang, S. *Social Network Analysis*; SAGE Publication: Thousand Oaks, CA, USA, 2019.
69. Lee, Y.W.; Strong, D.M. Knowing-Why about Data Processes and Data Quality. *J. Manag. Inf. Syst.* **2003**, *20*, 13–39. [CrossRef]
70. Alshikhi, O.A.; Abdullah, B.M. Information Quality: Definitions, Measurement, Dimensions, And Relationship with Decision Making. *Eur. J. Bus. Innov. Res.* **2018**, *6*, 36–42.
71. Jayawardene, V.; Sadiq, S.; Indulska, M. *An Analysis of Data Quality Dimensions*; The University of Queensland: St Lucia, Australia, 2015; pp. 1–31.
72. Tejay, G.; Dhillon, G.; Goyal Chin, A. Data Quality Dimensions for Information Systems Security: A Theoretical Exposition. In *Security Management, Integrity, and Internal Control in Information Systems*; IFIP TC-11 WG 11.1 &WG 11.5 Joint Working Conference 7; Springer: Berlin/Heidelberg, Germany, 2005; pp. 21–39.
73. Tobler, E. A Needs Assessment of Arizona Agricultural Education Equine Science Curriculum. Ph.D. Dissertation, Utah State University, Logan, UT, USA, 2018.
74. Johari, S.; Jha, K. Determinants of Workmanship: Defining Quality in Construction Industry. In *Proceedings of the 35th Annual Conference*; Leeds Beckett University: Leeds, UK, 2019; p. 761.
75. Tripathi, K.K.; Jha, K.N. An Empirical Study on Performance Measurement Factors for Construction Organizations. *KSCE J. Civ. Eng.* **2018**, *22*, 1052–1066. [CrossRef]
76. Assistant Librarian, D.S. Application of Garret Ranking Technique: Practical Approach. *Int. J. Libr. Inf. Stud.* **2016**, *6*, 135–140.
77. Garrett, H.E.; Woodworth, R.S. *Statistics in Psychology and Education*; Vakils, Feffer and Simons Private Ltd.: Bombay, India, 1969; p. 329.