Factors Influencing Master Data Quality: A Systematic Review

Azira Ibrahim[1], Ibrahim Mohamed[2], Nurhizam Safie Mohd Satar[3]
Faculty of Information Science and Technology
Universiti Kebangsaan Malaysia
Bangi, Malaysia

Abstract—Master data refers to the data that represents the core business of the organization, shared among different applications, departments, and organizations and most valued as the important asset to the organization. Despite the outward benefit of master data mainly in decision making and organization performance, the quality of master data is at risk. This is due to the critical challenges in managing master data quality the organization may expose. Hence the primary aim of this study is to identify factors influencing master data quality from the lens of total quality management while adopting the systematic literature review method. The study proposed 19 factors that inhibit the quality of master data namely data governance, information system, data quality policy and standard, data quality assessment, integration, continuous improvement, teamwork, data quality vision and strategy, understanding of the systems and data quality, data architecture management, personnel competency, top management support, business driver, legislation, information security management, training, change management, customer focus, and data supplier management that can be categorized to five components which are organizational, managerial, stakeholder, technological, and external. Another important finding is the identification of the differences for factors influencing master data compared to other data domain which are business driver, organizational structure, organizational culture, performance evaluation and rewards, evaluate cost/benefit tradeoffs, physical environment, risk management, storage management, usage of data, internal control, input control, staff participation, middle management's commitment, the role of data quality and data quality manager, audit, and personnel relation. It is expected that the findings of this study will contribute to a deeper understanding of the factors that will lead to an improved master data quality.

Keywords—Quality management; total quality management; data quality; data quality management; master data; master data quality; master data quality management; systematic literature review

I. INTRODUCTION

The evolution of digital transformation and a data-driven economy requires the formulation of new strategies to ensure the organization stays relevant and competitive. An organization is expected to face various issues as the effect of development that requires proactive management action [1]. Taking into account that data is an important element for every organization [2]–[4], the massive amount of data that are created and stored in response to digitalization possess new challenges in the management of data quality.

In particular, the organization is normally held responsible to manage a few types of data namely master data, transaction data, and reference data, to name a few. Master data is ranked as having the highest priority to be managed due to the valuable information it holds about the organization [5] and should be considered as an important asset to the organization [1], [6]. Master data represents the organization's core business objects that form the foundation of the main business process and must therefore be used unambiguously across the entire related application, department, and organization. Typical master data classes are supplier, customer, material, product, employee, and asset [7]–[9]. In the public sector context, master data composed of data about service providers, customers, and services or products offered [10].

The importance of master data requires it of high quality in supporting the organization to perform roles such as planning and decision making [11] and ensuring compliance with the regulatory and legal provision [12]. While the increasing demand for information system initiatives evidenced that high-quality master data is one of the important elements in the successfulness of the implementation [13], [14]. According to [12] current, accurate, and complete master data is required.

Studies in academic and industry highlighted that data quality is an urgent issue. The impact of poor data quality can be manifested across the operational, tactical, and strategic levels of the organization [15]. In the specific context of master data, poor master data quality incurred additional costs to the organization which involves a cost in assuring the quality of master data and cost affected by poor data quality [16]. On a similar tone, The Data Warehousing Institute (TDWI) calculated that data quality problems cost U.S. businesses about USD 600 billion a year [17]. Similarly, a study conducted in 2016 by Royal Mail [18] showed that poor quality of customer contact data costs, on average, 5.9% of the annual revenue to UK companies.

Despite the benefit and impact of poor master data quality, improving master data quality is still an issue. The industry is struggling in trusting the quality of the data and the implementation of data quality measures. A recent survey evidenced that only 40% of the respondent confident in the quality of data in their company and also their organization's data quality management practices [19]. According to [20] poor master data quality is one of the biggest challenges faced by the organization in managing the complexity of digitalization apart from standardization and governance.

Furthermore, according to [1], 80% of companies acknowledged the impact of poor master data quality to be high or very high for their performance, 82% of the company engaged in data quality initiative but not using the systematic or established method and only 15% of the companies know the established method for improving master data quality.

Undoubtedly, the importance of master data, the effect of poor quality master data, and the lack of adequate master data quality management underline the importance to initiate a study that revolves around the establishment of systematic master data quality management in ensuring the improvement of master data quality. However, considering that master data appears to have different characteristics compared to other domains of data and featuring distinct challenge and requirement, such as organizational, people, process, and technology [7], [21]–[23], thus a deeper understanding of the aspect related to master data is required before commencing any improvement initiative.

Xu [24] highlighted the importance to investigate, understand, and explain the factors influencing data quality, before proceeding with data quality improvement. However, a study that systematically explores factors influencing master data quality is scarce. Fortunately, the progress in the data quality management discipline by [24]–[30] made a substantial contribution in investigating factors influencing data quality.

The theoretical foundation for data quality management studies was originated from the Total Quality Management (TQM) discipline. TQM originally focused on quality improvement in the manufacturing domain [31], [32]. TQM provides an established organizational-wide foundation in identifying factors that contribute to data quality in the organization namely stakeholder, quality management, teamwork, process management, and top management support [33]. Based on TQM, [34] introduced the Total Data Quality Management (TDQM) approach in managing data quality, with the analogy of data as a manufactured product. The contribution by [34], is regarded as an important milestone for the advancement in data quality study.

In response, this paper aims to identify factors influencing master data quality from the lens of TQM based on the current and rigorous work in data quality management. The identification of the factors influencing master data quality will support the ongoing study in developing a framework for managing master data quality. Therefore, yields two research questions which are 1) what are the factors influencing master data quality in the organization?, and 2) how do the factors influencing master data quality differ from other data domains?. This paper employs systematic literature review approach in answering both research questions.

The remainder of the paper is structured as follows: Section II reviews the literature on data quality and master data quality. Section III describes the method for conducting a systematic literature review. Section IV presents the finding of the study. Section V discusses the finding. The paper ends with conclusions in Section VI.

II. RELATED WORK

A. Data Quality

Data quality is a complex construct composed of multiple dimensions [35]–[39]. Although previous scholars agree that there is no definite definition for data quality, however, it was acknowledged that data quality must meet user requirements for specific usage context or fitness for use [40]–[42]. Seminal literature such as [37] operationalized the term data quality using dimensions namely accuracy, timeliness, completeness, and consistency.

While defining data quality is an issue, the same goes for identifying the factors influencing data quality. Grounded on the theory of TQM, the studies in data quality progressively contribute to a deeper understanding of issues related to data quality. Besides, data quality can be considered as a sub-discipline of TQM. Several researchers show the advancement in discussing factors influencing data quality in various contexts [24]–[28]. Based on the theory of TQM, factors influencing data quality can be classified into five components which are organizational, managerial, stakeholder, technological, and external [25], [43], [44].

The works by [24], [25] focusing on the quality of accounting data that resides in AIS were among the most cited work in understanding factors influencing data quality. The theoretical foundations of the study are based on four area which are TQM, just-in-time (JIT), data quality, and accounting.

In getting a deeper insight into the factors influencing accounting data quality, [25] applied a qualitative methodology involving multiple case studies. The author suggested 26 factors that were classified by five categories, namely 1) AIS characteristics (nature of system), 2) data quality characteristics (data quality policies and standards, data quality approach, role of data quality, internal control, input control, understanding of the system and data quality, and continuous improvement), 3) stakeholders (top management's commitment, middle management's commitment, roles of data quality manager/manager group, customer focus, personnel relations, information supplier management, audit and review, and personnel competency), 4) organizational (training, organizational structure, organizational culture, performance evaluation and rewards, manage change, evaluate cost/benefit tradeoffs, teamwork, physical environment, and risk management), and 5) external factor.

Complementing the study by [25], the three most important factors influencing accounting data quality suggested by [24] through quantitative study namely 1) top management commitment, 2) the nature of the systems, and 3) input controls. Further, in the context of health data, [26] suggested six factors influencing data quality which are 1) top management support, 2) resources, 3) regulatory capability, 4) business-IT alignment, 5) staff participation, and 6) data/system integration.

In contrary to the previous studies, [27], [28] explored factors related to data quality management regardless of specific data domain, where the findings support a higher

generalization. Also rooted in TQM theory based on the study by [33], [27] suggested information quality management (IQM) framework which consists of 11 interdependent factors which are 1) IQM governance, 2) continuous IQM improvement, 3) training, 4) information quality requirements management, 5) information quality risk management, 6) information quality assessment/monitoring, 7) continuous information quality improvement, 8) information product lifecycle management, 9) storage management, 10) information security management, and 11) information architecture management. Furthermore, [28] enriched the work of [24], [25], [27] by suggesting the top three factors influencing data quality management, namely, 1) data governance, 2) management commitment and leadership, and 3) continuous data quality management improvement.

In the conclusion, the advancement of data quality study, ranging from specific data domain to general data domain provides a sound foundation in understanding and having deeper insight on issues related to data quality.

B. Master Data Quality

Acknowledged as an important asset and representing the core business process, assuring high-quality master data has gained extensive attention in the literature [39], [40], [45]. Concerning the improvement of master data quality, understanding factors influencing the quality of the data is a pre-requisite. Even though literature focusing on factors influencing master data quality is scarce, partial contribution by a few scholars such as [1], [21], [46], [47] providing a good starting point.

The first serious discussion and analyses of factors influencing master data quality were performed by [46], emphasizing that issues related to master data quality not only confined to technological aspects but more to organizational. Grounded on previous data quality theoretical foundation study by [9], [48]–[51], the author empirically validated five factors influencing master data quality which are the 1) delegation of responsibilities, 2) rewards, 3) data control, 4) employee competencies, and 5) information system.

As the continuation, a substantial work performed by [47] proposing 12 factors influencing master data quality which are 1) responsibilities for specific types of master data, 2) roles concerning data creation, use and maintenance, 3) organizational procedures, 4) management focus concerning data quality, 5) data quality measurements, 6) reward and reprimand about data quality, 7) training and education of data users, 8) written data quality policies and procedures, 9) emphasis on the importance of data quality by managers, 10) IT system for data management, 11) possibilities for input in existing IT system, and 12) usability of IT system. The identified factors were empirically validated using a survey mechanism that involved 787 Danish manufacturing company. The main difference in the work by [46] and [47] is the latter reclassified the factors identified in the previous literature to enable a more systematic understanding of the issues related to master data quality in ensuring the right improvement strategy.

A more specific perspective has been adopted by [1] that explored the challenges and requirements in managing master data quality in the context of digitalization. The author has adopted the SLR approach in getting a deeper insight into the current state of master data quality study and further validated the finding using 33 semi-structured interviews. In assuring the quality of master data during information sharing, the author suggested functional requirement for master data quality management (MDQM) tool that composed of six modules which are 1) analysis, 2) cockpit, 3) data model, 4) rules engine, 5) software architecture, and 6) software ergonomics. The functionality of each module can assist the organization in developing a tool for managing master data quality.

On another note, the study by [21] provides an understanding that different class of master data, exhibit distinct data quality challenges and requirements. The finding demonstrated the need to consider the development of a master data quality management approach based on the individual classes of master data. The author proposed a data quality assessment and improvement model that consists of eight elements which are 1) data quality assessment and improvement process, 2) technology, 3) protocol, 4) performance, 5) policy, 6) data standard, 7) data governance, and 8) data quality dimension.

Overall, although extensive research has been carried out in the field of master data quality supported by empirically validated finding, no single study exists that adopt both TQM as a theoretical lens and SLR as methodology. Theory helps in providing a systematic understanding of the real-world phenomenon, particularly provides a focus for the research [52]–[55]. In the case of data quality study, the wide adoption of TQM theory in understanding issues related to data quality is evidenced in many seminal works but, deficient in the context of master data quality. In the context of SLR methodology adoption, only evidenced in [1]. Nevertheless, the study by [21], [46], [47] does not systematically review all the relevant literature in discussing factors related to master data quality.

As a result of the lack of theoretical lens and systematic methodology, only partial contribution can be found in master data quality studies. In particular, finding by [1] emphasized on technological factors, while [21], [46], [47] unable to provide adequate and sufficient explanation on the master data quality challenges.

III. METHOD

SLR is a research method that provides a more structured and rigorous process in identifying and analyzing previous literature based on the specified research question. Normally, SLR-based study required the adaptation of established standards in guiding the researcher to perform the related and necessary process that will enable them to evaluate and examine the quality and rigor of a review. Therefore, this study is performed based on the guideline proposed by [56] that is designed particularly for Information System research, which consist of four main stages namely 1) planning, 2) selection, 3) extraction, and 4) analyses of findings. Each stage will be described further in the next section.

A. Planning

The planning stage emphasizes the identification of the research questions based on the study objective that acts as a frame in scoping the literature search. The main objective of this study is to investigate the factors influencing master data quality at the organizational level. Thus, this study formulated research questions which are 1) what are the factors influencing master data quality in the organization?, and 2) how do the factors influencing master data quality differ from other data domains?.

B. Selection

The selection stage identifies several relevant articles for the current study consist of three main processes. The first process is identifying the source of articles, followed by the construction of keywords, and lastly identification of inclusion criteria.

1) Source: The searching process covers seven main database sources, namely, 1) Web of Science, 2) Scopus, 3) ACM Digital Library, 4) Emerald, 5) Science Direct, 6) Springer Link, and 7) IEEE. Additionally, the study also includes Google Scholar to find more related articles on master data quality topics. The selection of databases was based on its coverage relating to information management source, expert recommendation, and accessibility of the database. The title, abstract, and keywords were used to conduct searches for journals, and proceedings, books, book chapters, and industry research.

a) Keywords: Construction of search keywords involves the process of 1) identification of alternative spellings and synonyms for major terms based on the thesaurus, dictionaries, encyclopedia, and past researches, 2) identification of keywords in relevant papers or books, and 3) usage of the Boolean OR to incorporate alternative spellings and synonyms [57]. Search keywords were constructed to retrieve as many articles as possible related to master data quality, the topic of interest in this study.

The search keywords are formulated by mentioning both the terms "master data quality" and "master information quality" due to the previous research in data management used both terms interchangeably. Search keyword also includes the term "master data management", in reflection to the previous literature that referred master data management in relevance to the approach in managing master data. Thus, based on the search keywords, the initial search strings are ("master data quality"), ("master information quality"), and ("master data management"). Then, the search strings were joint using "OR" Boolean. The search strings were then used as the input to each electronic database to retrieve the articles based on the titles, abstracts, contents, and keywords, depending on the advanced search facility.

2) Inclusion criteria: The inclusion criteria are defined as means to reduce the number of studies to a certain amount that is reasonable to the author. There are three inclusion criteria formulated which are 1) language, 2) literature type, and 3) timeline as per Table I. In the first criteria, this study only focuses on the article that is written in the English language.

The second criteria, limit the articles that are categorized only under journal, proceedings, books, and book chapters. Moreover, only articles between 2015 and 2020 are selected. Overall, a total of 2117 articles were found during the initial search, and 1285 articles were excluded based on exclusion criteria.

C. Extraction

A total of 832 articles were extracted for the third stage known as the study extraction. The manual searching process from Google Scholar is performed, in the case where the articles were not indexed in the selected database. The manual search resulted in additional two articles making the total articles 834. The metadata for the selected article include 1) title of the article, 2) publication year, 3) author, 4) abstract, 5) keywords, 6) article type, and 7) DOI/ISBN/ISSN Number is extracted. Then, the deduplication process is performed to remove the duplicated copies of the identified articles that exist across electronic repositories [58]. From this exercise, a total of 111 articles were removed during the checking of duplication, while 723 articles were further screened based on quality assessment criteria decided by the researcher.

At this stage, quality assessment was conducted by performing the practical screening against the 723 identical articles. Practical screening is the activity of screening the title and abstract of the articles based on quality assessment criteria to check the relevance of the articles [56]. The quality assessment criteria are 1) focus of the article, 2) mentioning any factor influencing master data quality, and 3) adequately describe the factors involved as per Table II. Consequently, a total of 708 articles were excluded because they are not fulfilling the quality assessment criteria. Finally, a total of 15 remaining articles are ready to be analyzed.

D. Analyses

This stage further analyzed 15 selected articles in answering the research questions. The detailed analyses are presented in the following Section IV.

TABLE I.   INCLUSION AND EXCLUSION CRITERIA

Criteria	Inclusion	Exclusion
Language	English	Non-English
Article type	Research article, conference proceeding, book chapter, and book	Not categorized as a research article
Timeline	Between 2015 and 2020	Less than 2015

TABLE II.   QUALITY ASSESSMENT CRITERIA

Code	Criteria
QA1	Is the main focus of the article is master data quality?
QA2	Are the articles describing any factor influencing master data quality?
QA3	Are the factors influencing master data quality adequately defined?

IV. RESULT

The systematic review process produced 15 related studies as presented in Table III. Regarding the credibility of the source, eight studies are from indexed journals [1], [14], [21], [59]–[63], four studies are from established conferences [13], [64]–[66], and three studies are from book publications [45], [67], [68]. In the case of present study, four articles were published in 2019 [14], [62], [63], [66], two articles in 2018[13], [65], four articles in 2017 [1], [59]–[61], two articles in 2016 [21], [64] [24, 86], and three articles in 2015 [45], [67], [68].

TABLE III.   LIST OF RELATED ARTICLE BY YEAR

Year	Author	Source
2015	[67]	Apress
2015	[45]	epubli GmbH
2015	[68]	Morgan Kaufmann
2016	[21]	International Journal of Business Information Systems
2016	[64]	24th European Conference on Information Systems (ECIS 2016)
2017	[59]	Studies in Health Technology and Informatics
2017	[60]	Journal of Theoretical and Applied Information Technology
2017	[1]	Lecture Notes in Business Information Processing, Springer, Cham.
2017	[61]	Journal of Enterprise Information Management
2018	[13]	26th European Conference on Information Systems (ECIS 2018)
2018	[65]	International Conference on Information Management and Technology (ICIMTech)
2019	[14]	International Journal of Information Management
2019	[66]	International Conference on Smart Applications, Communications and Networking (SmartNets 2019)
2019	[62]	International Journal of Business Information Systems
2019	[63]	International Journal of Information Management

The detailed finding of the study is described based on the research questions.

A. RQI: What are the Factors Influencing Master Data Quality in the Organization?

Further analyses of the finding produced a total of 19 factors influencing master data quality, then the identified factors are further classified into five components which are organizational, managerial, stakeholder, technological and external as suggested by [25], [43], [44]. The theoretical perspective of the classification is useful to group the factors into specific components to have a broader overview of their effect on master data quality and allowing systematic analysis of the finding. As exhibited in Table IV, the five components are organizational (five factors), managerial (six factors), stakeholder (four factors), technological (two factors), and external (2 factors). Based on Table IV, the most frequently discussed factor is data governance which is mentioned in 11 out of 15 studies, followed by information system and data quality policy and standard which is discussed in more than half of the studies. It is then followed by data quality assessment, integration, continuous improvement, teamwork,

data quality vision and strategy, understanding of the systems and data quality, data architecture management, and personnel competency with the occurrence between 4 and 7.

Lastly, with a frequency of less than 4, the factors are top management support, business driver, legislation, information security management, training, change management, customer focus, and data supplier management.

1) Organizational: Organizational is one of the components that have a major influence on master data quality. In particular, an organization does not only provide strategic direction to enable the implementation of a feasible road map in improving master data quality but also in many ways materialized the commitment in ensuring the achievement of data quality goals. In this case, a total of 11 studies were found focusing on an organizational component in improving master data quality. The discussed factors are data governance [14], [21], [68], [45], [59], [61]–[65], [67], teamwork [59], [61], [64], [67], data quality vision and strategy [45], [62], [63], [67], training [59], and change management [64].

TABLE IV.   FACTORS INFLUENCING MASTER DATA QUALITY

Component	Factor	Author
Organizational	Data governance	[14], [21], [68], [45], [59], [61]–[65], [67]
Organizational	Teamwork	[59], [61], [64], [67]
Organizational	Data quality vision and strategy	[45], [62], [63], [67]
Organizational	Training	[59]
Organizational	Change management	[64]
Managerial	Data quality policy and standard	[14], [21], [59], [61], [62], [65], [67], [68]
Managerial	Data quality assessment	[21], [45], [59], [62], [64], [65], [68]
Managerial	Continuous improvement	[21], [45], [59], [62], [64]
Managerial	Understanding of the systems and data quality	[45], [59], [61], [64]
Managerial	Data architecture management	[45], [65]–[67]
Managerial	Information security management	[14], [68]
Stakeholder	Personnel competency	[14], [59], [64], [66]
Stakeholder	Top management support	[14], [64], [66]
Stakeholder	Customer focus	[62]
Stakeholder	Data supplier management	[1]
Technological	Information system	[1], [13], [14], [21], [45], [62], [63], [66]–[68]
Technological	Integration	[13], [62], [63], [66]–[68]
External	Business driver	[14], [67], [68]
External	Legislation	[14], [61], [67]

a) Data governance: Data governance involves the establishment of an organizational structure for managing master data quality that can be either a newly formed committee or reoccupied existing formal organizational structure. The latter is preferred to avoid any bureaucracy [45]. The core component of effective data governance is explained by the enactment of roles, responsibilities, and decision areas related to master data quality management [14], [21], [68], [45], [59], [61]–[65], [67]. Roles and responsibilities can be defined based on three organizational levels which are strategic, managerial, and operational [67]. The strategic level involves the role and responsibilities of the business sponsor, chief information officer (CIO), and chief operating officer (COO) which are the head of the IT and business department, and the leader for data governance. While managerial level includes the roles and responsibilities of the program manager and solution architect for the respective master data quality management initiative. Lastly, the operational level comprises the roles and responsibilities of the technical and business team.

Another strategy in defining roles and responsibilities in managing master data quality is through the concept of ownership [61]–[64]. Vilminko-Heikkinen and Pekkola [63] further explained the approach using three-level of master data quality management which are managerial level that involves the concept owner role responsible for the whole master data quality management initiative, support function role involving technical task and data domain level role consist of data owner task responsible for the data domain as a whole.

Apart from that, the roles and responsibilities must be determined not only for the internal decision area but must include the external process especially when the organization is involved with outsourcing activity [45]. Furthermore, in assuring the continuous quality of master data, roles and responsibilities at every stage of the data life cycle such as data creation, modification, access, and deletion should be defined [45], [59]. [45], [59].

b) Teamwork: Sufficient communication, understanding, and involvement between technical and business employees across the department are the provision of effective teamwork [24]–[26], [28]. The effective teamwork can be facilitated using business friendly approach [69]. The management of master data quality must involve both technical and business people to ensure fair and equal accountability [59], [61], [64], [67]. Furthermore, it is essential to strengthen the alignment of responsibilities between both parties [59] especially at the high-level coordination [61] involving the enforcement of policies to support business activities and also compliance to regulation [67]. Lack of teamwork effort normally leaves the task of managing master data to the technical people [64] or worst still to no man island and could potentially compromise the quality of master data.

c) Data quality vision and strategy: Data quality vision and strategy provide the direction in ensuring the quality of master data [45], [62], [63], [67]. Data quality vision is developed in line with the organization's vision, providing the

key business initiatives to support the organization's vision [67]. While data quality strategy is the detailed component of the data quality vision, elaborating the business case, and roadmap for the implementation of data quality initiative [63], [67].

d) Training: Effective and adequate training is essential in ensuring the employees are equipped with sufficient knowledge and skills in managing master data quality [24], [25], [27], [28], [70]. According to [59], in the case of the industry that deals with specific technical data such as the health field, employees need to have sufficient training not only on how to properly perform data entry and data processing, but capable to perform data quality checking especially involving semantic data quality checking to ensure the quality of the master data.

e) Change management: Change management refers to the organization's capability in managing internal and external change such as merging, technology transformation, government regulations, and market shift [24], [25], [28]. According to [64], change can be managed using a top-down approach or bottom-up approach, with the ultimate goal is to ensure the commitment and involvement of the employee in taking up new responsibilities.

2) Managerial: Improving master data quality requires effective and efficient management involving the provision of comprehensive data quality guidelines in ensuring the process of managing master data quality properly performed. As previously mentioned, a total of 11 studies were found to focus on managerial component related to master data quality improvement. Nevertheless, the analyses for this component has resulted in a total of six factor, namely data quality policy and standard [14], [21], [59], [61], [62], [65], [67], [68], data quality assessment [21], [45], [59], [62], [64], [65], [68], continuous improvement [21], [45], [59], [62], [64], understanding of the systems and data quality [45], [59], [61], [64], data architecture management [45], [65]–[67] and information security management [14], [68].

a) Data quality policy and standard: Policy and procedure act as a frame to enable the improvement of data quality that includes the data that is managed internally by the organization [71], [72]. On top of that, data quality policy and standards provide managerial level guidance in implementing master data quality management initiatives. According to [24], [25], [28], the guideline normally has two main parts namely, what to achieve concerning data quality goal and how to achieve the stipulated goal [24], [25], [28].

In particular, data quality policy and standards spell out the detailed definition of master data [59], [62] and master data quality management taxonomy [61]. The definition includes the structure of the data, a business process that uses specific master data, the reason the master data is created, and governance of the master data [59], [61], [62], [67]. Other than that, the document also contains business rules for managing master data quality, guidelines in responding to data quality issues, and service level agreement (SLA) that act as data quality indicators [21], [59]. Ultimately, a well-written data quality policy and standard must support the business process

of the organization including compliance to regulation, fulfilling customer needs, and providing consistent reporting [14], [67], [68].

b) Data quality assessment: Data quality assessment is a pre-requisite step before proceeding with any data quality improvement initiative [27]. Several seminal authors emphasized that "only what can be measured can be improved" [39], [40], [49]. Therefore, measurement of data quality has to be done to determine the level of data quality over time. Data quality assessment consists of four main phases namely, definition, measurement, analysis, and improvement, which involve various stakeholders such as data collector, data custodian, and data user [34], [73].

During the definition phase, analysis of the current state of data quality is performed to discover any problem related to data quality [21], [65], then data quality requirement and measurement metric is determined based on a key performance indicator (KPI) for data quality and also business process performance [45], [64], [68]. Later, the data quality dimension is identified [21], [62], which acts as a facet for data quality that will be used in the subsequent phase. In the case of identifying data quality dimensions, the seminal work by [73], [74] are most frequently cited.

Then, measurement of data quality is performed which involves quantitative and qualitative strategy [28] based on identified KPI. The measurement can include syntax and semantic checking utilizing current technology such as rules engine and fuzzy search [45], [59]. As for the analysis phase, the assessment result should be compared to the earlier defined data quality requirement which is based on KPI, perform benchmarking, and prioritize data quality improvement strategy [21], [45], [59]. Finally, the improvement of data quality is implemented with adequate monitoring in place [21].

c) Continuous improvement: Continuous improvement of master data quality is normally driven by the changes in the internal and external environment such as technology and regulation that requires a proper response by the organization [24], [25]. Continuous improvement is an ongoing process for assuring the quality of master data, which involves the implementation of a preventive measure focusing more on the business process betterment [26], [59]. In particular, involving the installation of data quality elements in every phase of the data life cycle is required that include data collection, processing, deletion, and archiving [21], [45].

Another aspect of continuous improvement involves the assessment of the maturity level of master data quality management practice in the organization [64]. On top of that, regular data quality examination is also required, not only to gauge the level of master data quality but to enable the employee to see the effect of their work on data quality [59]. Finally, the result of the data quality improvement initiative must be integrated into the organization's operations for reporting and monitoring purposes [45], [59].

d) Understanding of the systems and data quality: Improvement of master data quality requires the understanding of how the application works, the importance of data quality and the relation with the business objective, and also the usefulness of the data [24], [25], [28]. Employees should understand the effect of poor data quality [45], [59] on the organization and be aware that the management of master data quality is an enterprise-wide initiative, which does not only affect specific business units but the organization as a whole [61], [64]. The understanding is important in motivating the employees to give full commitment in improving master data quality.

e) Data architecture management: Data architecture management involves the coordination of business process, application, data, and integration process [27] that includes the definition of global and local data, retention, and distribution of data [66], [67]. Master data quality requirement provides a basis for the data architecture management [65] include the identification of required tools and technology to build the solution [67].

f) Information security management: Information security management is referred to as the extent of the process and practice in the organization to safeguard the confidentiality of the master data [27], [28]. According to [14], [68], the privacy and security of master data include the protection from unauthorized access and the provision of reliable and secure communication means during data sharing.

3) Stakeholder: In this section, it is important to gain a deeper understanding of the influence of stakeholders as one of the critical components in managing master data quality. As previously mentioned, a total of six studies focused on master data quality related to stakeholders. The current study, managed to further categorized the component into four factors namely personnel competency [14], [59], [64], [66], top management support [14], [64], [66], customer focus [62], and data supplier management [1].

a) Personnel competency: Managing master data quality requires the employees to be equipped with sufficient skills and knowledge in both technical and business areas. According to [64], the organization should have a clear definition of the knowledge and skills that are needed for managing master data quality to ensure the right people are employed for the right task. Furthermore, according to [14], [59], [64], [66], the organization should appoint well-trained, experienced, and qualified personnel in both technical and business areas representing all departments in the organization.

b) Top management support: Awareness, competency, and leadership on master data quality possessed by top management is another important factor that is frequently discussed in the previous studies. Top management must be aware of the importance of master data quality and support activity related to master data quality management [64]. Apart from that, managing master data quality requires well-trained personnel, hence, top management should provide sufficient resources in improving skills and knowledge [14]. According to [66], top management also should focus on rewards and recognition programs for employees within an organization.

c) Customer focus: Focus on the user's needs is important to ensure the quality of data satisfies the defined requirement [24], [25], [27], [28]. Users must be involved during the data quality requirement elicitation, to ensure the correct requirement is captured [62]. In the context of system development process, data model that uses business metadata such as Source-Transaction-Agent (STA) model can be utilized to assist business and IT person to communicate and participate effectively and efficiently in business data modelling [69].

d) Data supplier management: Data supplier refers to the party, either internal and external to the organization that provides raw, unorganized data, while data supplier management is defined as having an effective relationship with the data provider by having an agreement about the acceptable level of data quality supplied and provide regular data quality reporting and technical assistance to data suppliers [24], [25], [28]. According to [1], sata provider is responsible to provide quality master data with fewer errors.

4) Technological: The technological component is considered as the operational level of the master data quality management initiative [75]. As previously mentioned, a total of 10 studies focused on master data quality related to technology. The present study managed to further classified the component into two factors which are information system [1], [13], [14], [21], [45], [62], [63], [66]–[68] and integration [13], [62], [63], [66]–[68].

a) Information system: In supporting the effective management of master data quality, the information system should provide sufficient functions, cutting edge architecture, and also adequate ergonomics features. The system should have the capability in assuring the quality of master data such as data profiling, data cleansing, data matching, data merging, data synchronization, and data consolidation [62], [63], [67], [68]. On top of that, [1] highlighted that the system should be capable to perform testing and simulation of data quality measurement and provide data quality monitoring in the form of a cockpit.

Furthermore, a system architecture is developed based on business process architecture [62] and adheres to the modular principle, adaptability, and reconfigurability [1]. Lastly, [1] explained that the system also must possess adequate ergonomic features such as easy to use, understandable, and comprehensive.

b) Integration: Since master data can exist in multiple sources, a certain degree of data integration is needed to preserve the quality of master data. Data integration implementation depends on the requirement such as data volume, data latency, nature of data, and the number of staging layers needed [67]. According to [68], the Entity Identity Information Management (EIIM) approach can be used to maintain the integrity of the master data, which is the fundamental element for Master Data Management (MDM).

In addition to that [13] proposed the usage of a federated approach to integrate the data based on shared attributes metadata to overcome the problem caused by a single source of truth approach.

5) External: External component refers to the factor that affects the master data quality which is not within the organization's control but somehow needs to be faced by the organization to stay competitive or comply with the regulation. As previously mentioned, a total of four studies focused on the master data quality related to the external component. The present study managed to further classify the component into two factors which are business driver [14], [67], [68] and legislation [14], [61], [67].

a) Business driver: In order to stay competitive and relevant in the data-driven economy, requires the organization to effectively and efficiently adapt to the ever-changing business need that in many ways require changes in how master data is managed. Poorly managed data, affect the data quality, hence influence the organization's performance. The related business driver that should be considered by organizations includes consumer demand for higher quality product or service, capability in offering new product and services in less time, single view reporting to enable more informed decision making, data integration from multiple sources, return on investment and ensuring data security especially during data sharing process [14], [67], [68].

b) Legislation: Every organization operates in an environment that is governed by certain rules and regulations that have to adhere. Such legislation includes the data protection act in ensuring the confidentiality and privacy of customer data is assured [14], [61], [67]. Failure to comply with the stipulated legislation not only affects the reputation of the organization, but to make it worse, is the possibility to face a lawsuit.

B. RQ2: How do the Factors Influencing Master Data Quality Differ from other Data Domain?

Table V summarized the differences for the factor that influence master data quality compared to other domains of data namely accounting data, health data, and general data.

Based on Table V, there are a total of 34 factors that influence accounting data, health data, general data, and master data with some differences. In particular, the factor that discussed only in master data is business driver, while in contrast 15 factors are discussed in other data domain but not explicitly in the master data domain namely 1) organizational structure, 2) organizational culture, 3) performance evaluation and rewards, 4) evaluate cost/benefit tradeoffs, 5) physical environment, 6) risk management, 7) storage management, 8) usage of data, 9) internal control, 10) input control, 11) staff participation, 12) middle management's commitment, 13) role of data quality and data quality manager, 14) audit, and 15) personnel relation. On top of that, it is worth highlighting that only 6 factors are discussed across all data domains, which include 1) teamwork, 2) data quality policy and standard, 3) continuous improvement, 4) top management support, 5) information system, and 6) integration.

TABLE V. FACTORS INFLUENCING THE QUALITY OF ACCOUNTING DATA, HEALTH DATA, GENERAL DATA, AND MASTER DATA

Factor	A	B	C	D
Business driver				/
Organizational structure	/			
Organizational culture	/		/	
Performance evaluation and rewards	/		/	
Evaluate cost/benefit tradeoffs	/		/	
Physical environment	/		/	
Risk management	/		/	
Storage management			/	
Usage of data	/			
Internal control	/			
Input control	/		/	
Staff participation		/		
Middle management's commitment	/			
Role of data quality and data quality manager	/			
Audit	/		/	
Personnel relation	/		/	
Teamwork	/	/	/	/
Data quality policy and standard	/	/	/	/
Continuous improvement	/	/	/	/
Top management support	/	/	/	/
Information system	/	/	/	/
Integration	/	/	/	/
Training	/		/	/
Change management	/		/	/
Data quality assessment	/		/	/
Understanding of the systems and data quality	/		/	/
Personnel competency	/		/	/
Customer focus	/		/	/
Data supplier management	/		/	/
Data governance			/	/
Data architecture management			/	/
Information security management			/	/
Data quality vision and strategy	/			/
Legislation	/			/

(A) Accounting data B) Health data C) General data D) Master data

While other factors namely data governance, data architecture management, information security management, data quality vision and strategy training, change management, data quality assessment, understanding of the systems and data quality, personnel competency, customer focus, personnel relation, data supplier management, and legislation are mentioned in master data quality study but discussed partially in another study. Overall, the similarity and dissimilarity of the finding provide a good justification in pursuing further study on master data quality.

V. DISCUSSION

The result of the analyses in Table IV gives further insight into the potential factor that could impact master data quality. It should also be noted that the first three factors which are data governance, data quality policy and standard, and information system clearly stand out from the rest. The high occurrence of data governance is noteworthy, 11 studies in master data discuss the factor, while only two studies in other data domains highlight the factor [27], [28]. This outcome is probably due to the increasing importance of data to organizations, particularly in the context of digital transformation, which has given rise to the need of establishing the roles and responsibilities in managing master data quality such as data ownership [61], [63], [64], among other.

However, establishing roles and responsibilities for managing master data quality exhibits complex challenges, since master data do not belong to a specific department but is an asset for the organization as a whole. Hence, managing master data requires an enterprise-wide approach compared to other data domains that are more compartmentalized to a specific business unit. The responsibilities are huge, where employees are reluctant to carry such accountability. The organization also might find difficulties in shifting the data management approach from department-based to enterprise wide-based. Hence, the complexity creates the need to further study in getting more insight and understanding of the phenomena.

Based on Table V, it can be summarized that there are differences in the factors influencing master data quality compared to other data domains. As highlighted by [21], different data domains, possess different challenges and requirements, hence requiring a more tailored suit management approach in ensuring data quality. In particular, the business driver factor is discussed in three master data studies domains but none in other data domains. The reason is might due to the effect of digital transformation to the organization that requires more proactive action in managing master data quality, while other data domain does not consider it as a threat.

Another interesting result to explore is regarding the total of 15 factors that are not explicitly discussed in the master data domain but mentioned in another data domain namely organizational structure, organizational culture, performance evaluation and rewards, evaluate cost/benefit tradeoffs, physical environment, risk management, storage management, usage of data, internal control, input control, staff participation, middle management's commitment, the role of data quality and data quality manager, audit, and personnel relation. As for the organizational structure, the factor was consolidated under the data governance factor due to the relevancy and suggestion by [28], whereas expanding the definition of data governance. Hence, the rest of the factor needs further investigation regarding the relevancy in the master data context.

Overall, the analyses suggested a total of 19 factors are relevant in the context of master data quality. The cross-reference of the identified 19 factors against the study in other

domain shows that teamwork, data quality policy and standard, continuous improvement, top management support, information system, integration, training, change management, data quality assessment, understanding of the systems and data quality, personnel competency, customer focus, and data supplier management are considered as established factors as were discussed in a minimum of three data domains of the study including master data domain. Meanwhile, as for the data governance factor, important to note that, even though lack of discussion of the factor in other domains of study, but the emphasis is so obvious in the master data domain, making it the most dominant factor. Other than that, data architecture management, information security management, data quality vision and strategy, and legislation discussed in two domains including the master data domain, which surely require a deeper understanding. The overall result, reflect the need and further justifying the need to pursue a study in master data quality to get a deeper insight into the influencing factors.

VI. CONCLUSION

A better understanding of the factors influencing master data quality will enable practitioners to improve master data quality. There is evidence that the effect of digital transformation requires the organization to change how it manages master data. Poorly managed master data, produce low-quality data and affect organization performance in term of fulfilling increasing customer demand, providing 360-degree single view reporting and integration of multiple data sources, to name a few.

Even though master data is acknowledged as an asset to the organization and a core element to the business process, the comprehensive study on factors influencing master data quality is very limited. With the aim to diminish the gap, this study can be considered as one of the first attempts to thoroughly review factors influencing master data quality.

The significant findings that transpired from this review study are that 19 factors of master data quality have been identified and categorized into five components which are organizational, managerial, stakeholder, technological, and external. The top 10 most influential factors are data governance followed by information system, data quality policy and standard, data quality assessment, integration, continuous improvement, teamwork, data quality vision and strategy, understanding of the systems and data quality, and data architecture management.

Interestingly, the analyses show that there are some dissimilarities for factors influencing master data quality, compared to other data quality domains. In the domain of accounting, the differences include organizational structure, organizational culture, performance evaluation and rewards, evaluate cost/benefit tradeoffs, physical environment, risk management, usage of data, internal control, input control, middle management commitment's, role of data quality and data quality manager, audit, personnel relation, business driver, data governance, data architecture management, and information security management.

While for the health data, the contradict factors involve staff participant, business driver, training, change management, data quality assessment, understanding of the systems and data quality, personnel competency, customer focus, data supplier management, data governance, data architecture management, information security management, data quality vision and strategy, and legislation. Eventually, for general data, the differences include organizational culture, performance evaluation and rewards, evaluate cost/benefit tradeoffs, physical environment, risk management, storage management, input control, audit, personnel relation, business driver, data quality vision and strategy, and legislation. It is recommended to further investigate these factors using an in-depth interview to better understand the phenomenon.

REFERENCES

[1] T. Schäffer and C. Leyh, "Master data quality in the era of digitization - Toward inter-organizational master data quality in value networks: A problem identification," Piazolo F., Geist V., Brehm L., Schmidt R. Innov. Enterp. Inf. Syst. Manag. Eng. ERP Futur. 2016. Lect. Notes Bus. Inf. Process. Springer, Cham, vol. 285, pp. 99–113, 2017.

[2] Z. Mohammad Yusof, Pengurusan rekod dan maklumat: Isu dan cabaran. Penerbit Universiti Kebangsaan Malaysia, Bangi, Malaysia, 2015.

[3] S. E. Madnick, R. Y. Wang, Y. W. Lee, and H. Zhu, "Overview and framework for data and information quality research," ACM J. Data Inf. Qual., vol. 1, no. 1, pp. 1–17, 2009.

[4] M. Mukred and Z. M. Yusof, "The Delone–McLean information system success model for electronic records management system adoption in higher professional education institutions of Yemen," Lect. Notes Data Eng. Commun. Technol., vol. 5, pp. 812–823, 2018.

[5] S. Nelke, M. Oberhofer, Y. Saillet, and J. Seifert, "U.S. Patent 2015/0066987 A1: Method and system for accessing a set of data tables in a source database," 2015.

[6] M. Y. Choi, C. J. Moon, K. S. Park, and D. K. Baik, "An enterprise master data model based on the data taxonomy based on their origin," in International Conference on Enterprise Information Systems and Web Technologies (EISWT), 2010.

[7] A. Dreibelbis, E. Hechler, I. Milman, M. Oberhofer, P. van Run, and D. Wolfson, Enterprise master data management: An SOA approach to managing core information. Pearson Education, 2008.

[8] D. Loshin, Master data management. Morgan Kaufmann OMG Press, 2009.

[9] H. A. Smith and J. D. Mckeen, "Developments in practice XXX: Master data management: Salvation or snake oil?," Commun. Assoc. Inf. Syst., vol. 23, no. 4, pp. 1–11, 2008.

[10] F. Haneem, N. Kama, and A. Azmi, "Master data identification in public sector organisations," Adv. Sci. Lett., vol. 22, no. 10, pp. 2999–3003, 2016.

[11] J. S. Arlbjørn, C. Y. Wong, and S. Seerup, "Achieving competitiveness through supply chain integration," Int. J. Integr. Supply Manag., vol. 3, no. 1, pp. 4–24, 2007.

[12] B. Otto, K. M. Hüner, and H. Österle, "Toward a functional reference model for master data quality management," Inf. Syst. E-bus. Manag., vol. 10, no. 3, pp. 395–425, 2012.

[13] T. Dahlberg, A. Lagstedt, and T. Nokkala, "How to address master data complexity in in-formation systems development – A federative approach," in 26th European Conference on Information Systems (ECIS 2018), 2018, pp. 1–15.

[14] F. Haneem, N. Kama, N. Taskin, D. Pauleen, and N. A. Abu Bakar, "Determinants of master data management adoption by local government organizations: An empirical study," Int. J. Inf. Manage., vol. 45, pp. 25–43, 2019.

[15] T. C. Redman, "The impact of poor data quality on the typical enterprise," Commun. ACM, vol. 41, no. 2, pp. 79–82, 1998.

[16] A. Haug, F. Zachariassen, and D. van Liempd, "The costs of poor data quality," J. Ind. Eng. Manag., vol. 4, no. 2, pp. 168–193, 2011.

[17] W. W. Eckerson, "Data quality and the bottom line: Achieving business success through a commitment to high quality data," 2002.

[18] Royal Mail and DataIQ, "How better customer data drives marketing performance and business growth," 2016.

[19] C. Lehmann, K. Roy, and B. Winter, "The state of enterprise data quality: 2016: Perception, reality and the future of DQM," 2016.

[20] Deloitte, "Complexity: Overcoming obstacles and seizing opportunities: The Deloitte global CPO survey 2019," 2019.

[21] R. Silvola, J. Harkonen, O. Vilppola, H. Kropsu-Vehkapera, and H. Haapasalo, "Data quality assessment and improvement," Int. J. Bus. Inf. Syst., vol. 22, no. 1, pp. 62–81, 2016.

[22] A. Hayler, "Ten years on, master data management software market comes of age," 2014. .

[23] M. Treder, "Chapter 5- Masterdata management," in The chief data officer management handbook: Set up and run an organization's data supply chain, Apress, 2020, pp. 61–78.

[24] H. Xu, "What are the most important factors for accounting information quality and their impact on AIS data quality outcomes?," J. Data Inf. Qual., vol. 5, no. 4, pp. 1–22, 2015.

[25] H. Xu, A. Koronios, and N. Brown, "Managing data quality in accounting information systems," in IT-based management: Challenges and solutions, Idea Group Publishing: Hershey PA, 2002, pp. 277–299.

[26] C. Liu, D. Zowghi, and A. Talaei-Khoei, "An empirical study of the antecedents of data completeness in electronic medical records," Int. J. Inf. Manage., vol. 50, pp. 155–170, 2020.

[27] S. Baškarada and A. Koronios, "A critical success factor framework for information quality management," Inf. Syst. Manag., vol. 31, no. 4, pp. 276–295, 2014.

[28] A. Lucas, "Critical success factors for corporate data quality management," in World Conference on Information Systems and Technologies, 2019, pp. 630–644.

[29] E. Hassan, Z. M. Yusof, and K. Ahmad, "Factors affecting information quality in the Malaysian public sector," Int. J. Adv. Sci. Eng. Inf. Technol., vol. 9, no. 1, pp. 32–38, 2019.

[30] E. Hassan, Z. M. Yusof, and K. Ahmad, "Modeling of information quality management in malaysian public sector: A pls-sem approach," J. Theor. Appl. Inf. Technol., vol. 96, no. 19, pp. 6361–6375, 2018.

[31] J. M. Juran and A. B. Godfrey, Juran's quality handbook, Fifth edit. McGraw-Hill, 1998.

[32] W. E. Deming, Out of the crisis. MIT Press, 1986.

[33] J. Motwani, "Critical factors and performance measures of TQM," TQM Mag., vol. 13, no. 4, pp. 292–300, 2001.

[34] R. Y. Wang, "A product perspective on total data quality management," Commun. ACM, vol. 41, no. 2, pp. 58–65, 1998.

[35] L. Bertossi and M. Milani, "Ontological multidimensional data models and contextual data quality," J. Data Inf. Qual., vol. 9, no. 3, pp. 1–36, 2018.

[36] D. Ballou, R. Wang, H. Pazer, and G. K. Tayi, "Modeling information manufacturing systems to determine information product quality," Manage. Sci., vol. 44, no. 4, pp. 462–484, 1998.

[37] D. P. Ballou and H. L. Pazer, "Modeling data and process quality in multi-input, multi-output information systems," Manage. Sci., vol. 31, no. 2, pp. 150–162, 1985.

[38] T. C. Redman, Data quality for the information age. Artech House, 1996.

[39] Y. Wand and R. Y. Wang, "Anchoring data quality dimensions in ontological foundations," Commun. ACM, vol. 39, no. 11, pp. 86–95, 1996.

[40] R. Y. Wang and D. M. Strong, "Beyond accuracy: What data quality means to data consumers," J. Manag. Inf. Syst., vol. 12, no. 4, pp. 5–34, 1996.

[41] G. K. Tayi and D. P. Ballou, "Examining data quality," Commun. ACM, vol. 41, no. 2, pp. 54–57, 1998.

[42] S. Watts, G. Shankaranarayanan, and A. Even, "Data quality assessment in context: A cognitive perspective," Decis. Support Syst., vol. 48, no. 1, pp. 202–211, 2009.

[43] H. Xu, "Data quality issues for accounting information systems' implementation: Systems, stakeholders, and organizational factors," J. Technol. Res., pp. 1–11, 2009.

[44] H. Xu, J. H. Nord, G. D. Nord, and B. Lin, "Key issues of accounting information quality management: Australian case studies," Ind. Manag. Data Syst., vol. 103, no. 7, pp. 461–470, 2003.

[45] B. Otto and H. Österle, Corporate data quality: Prerequisite for successful business models. epubli GmbH, 2015.

[46] A. Haug and J. S. Arlbjørn, "Barriers to master data quality," J. Enterp. Inf. Manag., vol. 24, no. 3, pp. 288–303, 2011.

[47] A. Haug, J. S. Arlbjørn, F. Zachariassen, and J. Schlichter, "Master data quality barriers: An empirical investigation," Ind. Manag. Data Syst., vol. 113, no. 2, pp. 234–249, 2013.

[48] A. Umar, G. Karabatis, L. Ness, B. Horowitz, and A. Elmagardmid, "Enterprise data quality: A pragmatic approach," Inf. Syst. Front., vol. 1, no. 3, pp. 279–301, 1999.

[49] L. P. English, Improving data warehouse and business information quality: Methods for reducing costs and increasing profits. Wiley & Sons, 1999.

[50] H. Xu, J. H. Nord, N. Brown, and G. D. Nord, "Data quality issues in implementing an ERP," Ind. Manag. Data Syst., vol. 102, no. 1, pp. 47–58, 2002.

[51] Y. W. Lee, L. L. Pipino, J. D. Funk, and R. Y. Wang, Journey to data quality. The MIT Press, 2006.

[52] C. Kivunja, "Distinguishing between theory, theoretical framework, and conceptual framework: A systematic review of lessons from the field," Int. J. High. Educ., vol. 7, no. 6, pp. 44–53, 2018.

[53] S. Reeves, M. Albert, A. Kuper, and B. D. Hodges, "Qualitative research: Why use theories in qualitative research?," BMJ, vol. 337, no. 7670, pp. 631–634, 2008.

[54] A. Tashakkori and C. Teddlie, Handbook of mixed methods in social and behavioral research. Sage Publications, 2010.

[55] L. Varpio, E. Paradis, S. Uijtdehaage, and M. Young, "The distinctions between theory, theoretical framework, and conceptual framework," Acad. Med., pp. 989–994, 2020.

[56] C. Okoli, "A guide to conducting a standalone systematic literature review," Commun. Assoc. Inf. Syst., vol. 37, no. 1, pp. 879–910, 2015.

[57] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," 2007.

[58] Q. He, Z. Li, and X. Zhang, "Data deduplication techniques," in International Conference on Future Information Technology and Management Engineering, FITME, 2010, vol. 1, pp. 430–433.

[59] K. Arthofer and D. Girardi, "Data quality- and master data management - A hospital case," Stud. Health Technol. Inform., vol. 236, pp. 259–266, 2017.

[60] F. Haneem, A. Azmi, and N. Kama, "Co-dependence relationship between master data management and data quality: A review," J. Theor. Appl. Inf. Technol., vol. 95, no. 22, pp. 6323–6335, 2017.

[61] R. Vilminko-Heikkinen and S. Pekkola, "Master data management and its organizational implementation: An ethnographical study within the public sector," J. Enterp. Inf. Manag., vol. 30, no. 3, pp. 454–475, 2017.

[62] R. Silvola, A. Tolonen, J. Harkonen, H. Haapasalo, and T. Mannisto, "Defining one product data for a product," Int. J. Bus. Inf. Syst., vol. 30, no. 4, p. 489, 2019.

[63] R. Vilminko-Heikkinen and S. Pekkola, "Changes in roles, responsibilities and ownership in organizing master data management," Int. J. Inf. Manage., vol. 47, pp. 76–87, 2019.

[64] R. Vilminko-Heikkinen, P. Brous, and S. Pekkola, "Paradoxes, conflicts and tensions in establishing master data management function," in 24th European Conference on Information Systems (ECIS), 2016, pp. 1–16.

[65] Z. Murti, A. Andarrachmi, A. N. Hidayanto, and S. B. Yudhoatmojo, "Master data management planning: (Case study of personnel information system at XYZ Institute)," in International Conference on

Information Management and Technology (ICIMTech), 2018, pp. 160–165.

[66] T. E. Hutang and B. M. Kalema, "The effects of demographic variables on master data quality management to improve service delivery," in International Conference on Smart Applications, Communications and Networking (SmartNets), 2019, pp. 1–6.

[67] S. Chaki, "Pillar No. 4: Master information management," in Enterprise information management in practice: Managing data and leveraging profits in today's complex business environment, Apress, 2015, pp. 63–78.

[68] J. R. Talburt and Y. Zhou, "The value proposition for MDM and big data," in Entity information life cycle for big data: Master data management and information integration, Morgan Kaufmann, 2015, pp. 1–16.

[69] I. Mohamed and M. F. Noordin, "STA data model for effective business process modelling," Procedia Technol., vol. 11, pp. 1218–1222, 2013.

[70] H. Xu, "Managing accounting information quality: An Australian study," in International Conference on Information Systems (ICIS), 2000, pp. 628–634.

[71] M. Basri, Z. Mohammad Yusof, and N. A. Mat Zin, Dasar maklumat nasional di Malaysia. Penerbit Universiti Kebangsaan Malaysia, Bangi, 2013.

[72] N. Abdul Halim, Z. M. Yusof, and N. A. M. Zin, "The requirement for information governance policy framework in Malaysian public sector," Int. J. Eng. Technol., vol. 7, no. 4.15, pp. 235–239, 2018.

[73] Y. W. Lee, D. M. Strong, B. K. Kahn, and R. Y. Wang, "AIMQ: A methodology for information quality assessment," Inf. Manag., vol. 40, no. 2, pp. 133–146, 2002.

[74] B. K. Kahn, D. M. Strong, and R. Y. Wang, "Information quality benchmarks: Product and service performance," Commun. ACM, vol. 45, no. 4, pp. 184–192, 2002.

[75] J. Xiao, K. Xie, and X. Wan, "Factors influencing enterprise to improve data quality in information systems application - An empirical research on 185 enterprises through field study," in 16th International Conference on Management Science and Engineering, IEEE, Moscow, Russia, 2009, pp. 23–33.