IPTEK The Journal of Technology and Science, 33(3), 2022 (e/pISSN:2088-2033, 0853-4098)

DOI: 10.12962/j20882033.v33i3.14198 | Received 20 August, 2022; Revised 9 Sept, 2022; Accepted 22 Sept, 2022

ORIGINAL RESEARCH

DATA QUALITY MANAGEMENT STRATEGY TO IMPROVE REMOTE SENSING DATA QUALITY: A CASE STUDY ON NATIONAL REMOTE SENSING DATA BANK

Agnes Sondita Payani* | Alia Mutia Mayanda | Adenia Adiresta | Yova Ruldeviyani

Faculty of Computer Science, Universitas Indonesia, Jakarta, Indonesia

Correspondence

*Agnes Sondita Payani, Faculty of Computer Science, Universitas Indonesia, Jakarta, Indonesia. Email: agnes.sondita@ui.ac.id

Present Address

Magister Teknologi Informasi, Fakultas Ilmu Komputer, Kampus UI Salemba, Jl. Salemba Raya No.4, DKI Jakarta, Indonesia, 10430

Abstract

National Remote Sensing Data Bank (Bank Data Penginderaan Jauh Nasional; BDPJN) is the operational implementer in the field of remote sensing to meet the needs of national data on remote sensing images. Remote sensing images are essential in the Indonesian government for various needs, such as in regional (spatial planning, city boundaries), disaster (forest fires, floods, landslides, earthquakes, volcanic eruptions), plantation, natural resources, and environment sectors. Quality management of BDPJN's satellite images becomes challenging due to the increasing number of products owned and users annually. For this reason, a data quality management strategy is needed to guarantee and improve the quality of BDPJN data. To develop such a strategy, an assessment of the maturity of BDPJN's data quality management was conducted from the aspect of data processing by implementing Loshin's Data Quality Management Maturity Model (DQM3) to find out the characteristics that were lacking. The results were then mapped based on Data Quality Management (DQM) activities in DAMA-DMBOK as a recommendation for data quality management strategies. This study applies quantitative research where data collection was done by distributing questionnaires to 24 respondents who are data stewards of a medium, high, and very high-resolution mosaic images. Based on the assessment, BDPJN is in the maturity level of Defined to Managed. The recommendations are 21 DQM activities that can be carried out to improve BDPJN data quality.

KEYWORDS:

BDPJN, Data Management, Data Quality Management Maturity Model, Loshin

1 | INTRODUCTION

Under the obligations of Law Number 21 of 2013 and Presidential Instruction Number 11 of 2012, the Indonesian National Institute of Aeronautics and Space (LAPAN) is an agency that processes and distributes remote sensing data while controlling quality[1]. The data management and distribution activities are carried out through the National Remote Sensing Data Bank (BDPJN). LAPAN has been integrated with the National Research and Innovation Agency (BRIN). The implementation of BDPJN operational management was transitioning from the Center of Remote Sensing Technology and Data Center (Pusat Teknologi dan Data Penginderaan Jauh; Pustekdata) to the Center for Data and Information (Pusat Data dan Informasi; Pusdatin) of BRIN. Remote sensing data is crucial as it covers decisions and policy makings by the Indonesian government. Some aspects that require the use of remote sensing data include disasters (forest fires, floods, landslides, earthquakes, volcanic eruptions), environment, forestry, agriculture, fisheries, taxes, security, regional (spatial planning, city limits), and many others.

BDPJN is the implementation of remote sensing operations covering the functions of (1) collecting, maintaining, updating, and dissemination of remote sensing metadata and data in Indonesia territory; (2) providing remote sensing data with minimal cloud cover; (3) presenting information on the quality of remote sensing data in the form of metadata and/or data history; (4) supplying supervision of data utilization; (5) producing inputs to the government regarding the mastery of remote sensing technology and data; (6) building the access system of the National Spatial Data Network (JDSN); and (7) providing access to geospatial data licensed by the Indonesian government[2]. This license is covering all ministries, local governments, Indonesian National Armed Forces (TNI), Indonesian National Police (POLRI), and educational institutions.

Management of BDPJN data quality remains a challenge since the amount of data and the number of satellite image products continues to increase. BDPJN products distributed to users are quite diverse based on the types of satellite and the processed image products from certain satellite images. Currently, there are 22 satellite images, each of which has its own data character[3]. In addition, the types of processed satellite image products have also increased over the years to improve data services and accommodate user needs. For example, in the past two years, BDPJN has started to produce Analysis Ready Data (ARD) satellite images mosaic with minimal clouds in medium, high, and very high resolution. This data has been processed and is ready to use.

BDPJN has implemented quality management based on ISO 9001:2015. The organization has created operational standards and guidelines to produce the best data quality. Nevertheless, not all activities have been covered by the implementation of ISO 9001:2015, especially regarding products that have just passed the experimental stage but are already being used by users. This may impact the data quality of certain products due to the absence of Quality Control (QC) or handover forms between workgroups. Without QC or data handover forms between workgroups, there is also a risk of data leaking to unauthorized parties.

Poor Data Quality Management (DQM) may cause a strained relationship between the organization and the users and affect the inaccurate information generated from the data. This can also impact the precision of decisions and policy makings derived from using or applying BDPJN data[4–6]. The increasing number of BDPJN data users will also increase the urgency to ensure that the DQM of BDPJN is running effectively. Well-documented data will facilitate the search for data, and the results can be obtained as expected. In addition, accurate or qualified data will be scientifically maintained.

A study on BDPJN's service and data quality measurement has been conducted previously using the Community Satisfaction Index (CSI) survey method[7]. Data Quality is 1 of 8 elements of assessment based on the Decree of the Minister of the Administrative and Bureaucratic Reform of the Republic of Indonesia No. 25 of 2004 (KEP/25/M.PAN/2/2004). However, the results were only obtained from the users' perspective. This study did not discuss the evaluation of DQM in BDPJN and provided no recommendations for improving DQM. These research gaps underlie the Data Quality Management Maturity (DQMM) measurement carried out in the present study to measure how well BDPJN's DQM is. For this purpose, a DQM model developed by Loshin was implemented[8]. Previous studies used Loshin's DQM model to measure DQMM. The studies include case studies on Statistics Indonesia (BPS) in Kaur Regency, Bengkulu Province[9], Indonesian Agency for Meteorology, Climatology, and Geophysics (Badan Meteorologi, Klimatologi, dan Geofisika; BMKG)[10], and STIS Polytechnic of Statistics[11]. DQM is one of the functions of data management based on the DAMA-Data Management Body of Knowledge (DAMA-DMBOK). Data management is a business function that starts from the planning, controlling, and delivering of data and information assets. Disciplinary functions include coaching, implementing, and overseeing plans, policies, practices, procedures, processes, programs, and projects that manage, protect, convey, and enhance the value of the information and data assets. Meanwhile, DQM is

the planning, implementation, and control activity that applies quality management techniques to assess, ensure, improve, and measure the suitability of the data used[12].

Many studies have discussed image quality, but no one has discussed image quality management, whereas data management will also affect the results of image data quality. In addition, previous research has only examined the aspect of user satisfaction regarding BDPJN data quality[7]. It has yet to discuss evaluation in terms of internal management processes and has yet to provide recommendations to improve the data quality. Therefore, this study was conducted to determine the maturity of data quality management to obtain a strategy of DQM to increase the quality of BDPJN data. To develop this strategy, a maturity assessment of BDPJN's data quality management was done using Loshin's Data Quality Management Maturity Model (DQM3) to determine the lacking characteristics. The results were then mapped based on DQM activities in DAMA-DMBOK as the data quality management strategy, which is the best theory for improving data quality[10, 13–20]. This study is expected to produce recommendations that BDPJN management units can use to maintain and improve BDPJN data quality.

2 | PREVIOUS RESEARCHES

2.1 | Data Management Body of Knowledge (DMBOK)

DMBOK is an accessible reference book for professionals in the field of data management who support DAMA's missions, including[12]:

- Provide a framework for the implementing enterprise data management practices
- Define general data management concepts vocabulary and serve as the basis of data management best practices
- Serve as a fundamental guide for the Certified Data Management Professional (CDMP) and other certifications.

DMBOK compiles 11 knowledge areas, one of which is DQM. The data quality knowledge area includes the planning and implementing management techniques of quality to assess, improve, and measure the suitability of data for use within an organization[12].

2.2 | Data Quality

Studies on data quality have been conducted since the 1990s[21]. According to DAMA[12], data quality is related to high-quality data characteristics and the processes used to improve or measure data quality. In addition, high-quality data must meet user expectations and needs. Various dimensions of data quality have been proposed in several previous studies. DAMA UK (2013) proposed six main dimensions of data quality and six other dimensions, while Cai and Zhu[21] submitted five dimensions with 14 elements of data quality in the era of big data. Meanwhile, data quality dimensions or geographic information can also refer to ISO 19157:2013 and ISO 19157:2015 SNI on Geographic Information – Data Quality, where there are six dimensions with 15 indicators[22]. According to Law No. 21 of 2013 concerning Outer Space. There are five dimensions that reflect the quality of image data[1]. A comparison of the data quality dimensions can be seen in Table 11 .

Table 1 is used to identify various dimensions of data quality from several sources. The dimensions proposed by DAMA UK do not adequately cover the characteristics of image data which are unstructured data, and those dimensions are more suitable for structured data. While the Cai and Zhu[21] dimensions are built for the big data era with rich data types, they are still too general. So, this study will use a combination of the dimensions of ISO 19157:2013 and Law No. 21/2013, which are more relevant for remote sensing image data quality. These dimensions will be used and adjusted for this research instrument.

2.3 | Data Quality Management (DQM)

According to DMBOK, the objectives of DQM are to (1) develop a regulated approach to make data suitable for user needs, (2) establish data quality control of requirements, specifications, and standards as a part of the data lifecycle, (3) define and implement the processes for measuring, monitoring, and reporting data quality levels, and (4) identify and advocate for opportunities of data quality improvement through process and system improvements[12].

DAMA-DMBOK also provides six activities that can be a strategy to increase DQM as follows:

TABLE 1 Comparison of data quality dimensions.

References	Data Quality Dimensions
Henderson[12]	Six main dimensions: Completeness, Uniqueness, Timeliness, Validity, Accuracy, Consistency Other dimensions: Usability, Timing Issues, Flexibility, Confidence, and Value.
Cai and Zhu[21]	Five dimensions with a total of 14 elements of data quality in the era of big data: Availability (accessibility, timeliness, authorization), Usability (definition or documentation, credibility, and metadata), Reliability (accuracy, integrity, consistency, completeness, auditability), Relevance (fitness), and Presentation Quality (readability, structure).
ISO 19157:2013 Geographic Information – Data Quality and SNI ISO 19157:2015 Geographic Information – Data Quality	Six dimensions of geospatial information quality with a total of 15 indicators: Completeness (commission, omission), Logical Consistency (conceptual, domain, format, topology), Positional Accuracy (external, internal, grid data), Thematic Accuracy (classification accuracy, non-quantitative attribute accuracy, quantitative attribute accuracy), Temporal Accuracy (time measurement, temporal consistency, temporal validity), and Usability.
Law No 21/2013[1]	Five dimensions of data quality: Geometric Correction Level, Radiometric Correction Level, Acquisition Time, Percentage of Cloud Cover, and Intellectual Property Rights.

1. *Determine high-quality data (DQM11).* This includes understanding business needs, determining terms, identifying organizational difficulty points, and building consensus on drivers and priorities for data quality improvement.
2. *Determine data quality strategies (DQM21).* This activity discusses the strategies needed to improve data quality.
3. *Determine the scope of the initial assessment.* This activity addresses the need for priorities based on existing factors. This activity has two sub-activities, namely: (a) identify critical data (DQM31), and (b) identify existing rules and patterns (DQM32).
4. *Display the initial data quality assessments.* This activity discusses assessing the quality of data by understanding the context and relationship of the data and comparing the actual data with rules and expectations. The sub-activities of this activity are (a) identify and prioritize problems (DQM41) and (b) perform an analysis of the root cause of the problems (DQM42).
5. *Identify and prioritize improvements.* In this activity, stakeholders are interviewed to find data problems that affect them, followed by an analysis of the business impact of these problems. This activity has the following sub-activities: (a) prioritize actions based on business impact (DQM51), (b) develop preventive and corrective measures (DQM52), and (c)confirm planned actions (DQM53)
6. Develop and deploy data quality operations This activity is related to the implementation of a plan that enables the teams to manage data quality rules and standards, monitor the continuous conformity of data with the rules, identify and manage data quality issues, and report quality levels. The four sub-activities are as follows: (a) develop data quality operational procedures (DQM61), (b) fix data quality defects (DQM62), (c) measure and monitor data quality (DQM63), and (d) report data quality level and findings (DQM64).

2.4 | Data Quality Management Maturity Model (DQM3)

Prior studies on the DQM3 include those by Loshin[8] and Bitterer[23], Information Quality Management-Capability Maturity Model (IQM-CMM)[24], and the Total Information Quality Management-Capability Maturity Model (TIQM-CMM)[25]. The comparison of the DQM3 is described in Table 2 , while the comparison of components and previous studies using the DQM3 can be seen in Table 3 .

After conducting the literature study, the present study implemented Loshin's DQM3 since the model can reduce the impacts of poor data quality, such as those on organizational trust, investment risk, and productivity[8]. In addition, Loshin's DQM3 also measures DQM from data processing[9]. Thus, the evaluation results on DQM characteristics from poor aspects of the data processing can be identified.

TABLE 2 Comparison of data quality management maturity model (DQM3).

Model	Level				
	1	2	3	4	5
TIQM-CMM[25]	**Initial** – DQM is usually temporary and cluttered, and there is no Key Process Area (KPA). Organizational success depends on the ability of its members, not on the success of its business processes.	**Repeatable** - Project management is based on the success of repeating similar projects. KPA focuses on implementing the basic controls of project management.	**Defined** - KPA aims to develop effective and efficient management processes for the entire project implementation.	**Quantitatively Managed** - KPA focuses on a quantitative understanding of the processes and the products built. Improvement criteria use the measurement results.	**Optimizing** - a quantitative understanding of the organization's business objectives and performance becomes the focus for continuous improvement processes.
Gartner[23]	**Aware** - The concept of data quality (DQ) understanding could be much higher. Any issues are largely ignored and no initiatives or incentives to increase DQ. No responsibility or accountability of DQ.	**Reactive** - Prefers to wait for problems instead of taking proactive steps. DQ issues are still considered the sole responsibility of the IT department.	**Proactive** - DQ is gradually becoming a part of IT. DQ tools come into use. The main problem is documented but still needs to be completely fixed.	**Managed** - DQ becomes a major concern, demonstrated by a wide implementation of DQ software and regular DQ assessment and impact analysis. Several data management roles are assigned, and a metric-based DQ dashboard is already used.	**Optimizing** - Maintain the best possible DQ through a continuous quality-level monitoring process and the attachment of data quality metrics. DQ becomes a sustainable strategic initiative. Subjective DQ assessment.
IQM-CMM[24]	**Chaotic** - There needs to be documentation and standards, consistent implementation of DQM, and awareness of any information quality problems.	**Reactive** - Start to be more aware of information quality issues. However, organizations only react to information quality issues that have occurred. Organizations have begun documenting, standardizing, and implementing basic data management processes.	**Measuring** - Organizations begin to manage the data they have as a lifecycle product, data changes that occur are recorded and can be canceled.	**Managed** - Organizations have information governance, ensure data accountability, and provide rewards and incentives. Organizations have also developed and documented information architectures.	**Optimizing** - Organizations continuously monitor the performance of information quality management through internal or external benchmarking.
Loshin[8]	**Initial** - The DQM processes are mostly temporary, and most work deals with data quality issues.	**Repeatable** - There has been a basis for data quality management, and the identification of the best practices has begun.	**Defined** - There has already been a data quality team and best practice documentation, and the adaptation throughout the organization is started.	**Managed** - There has been an ability to analyze the business impact to express DQ expectations and measure compliance. DQ is proactive, and data defects are identified at the beginning of an information workflow.	**Optimizing** - There has been a data quality governance framework for measuring and identifying systemic data quality improvement opportunities.

3 | MATERIAL AND METHOD

3.1 | Research Flow

There are six stages in this study, namely (1) problem identification, (2) literature study, (3) determination of the framework or DQM3, (4) data collection through the distribution of questionnaires, (5) DQMM calculation, and (6) recommendations for DQM improvement strategies. The research flow can be seen in Fig. 1 . The problem identification stage aims to identify problems in BDPJN, namely the data quality management issues that have yet to be implemented thoroughly in every BDPJN activity. The literature study stage was carried out by studying the DQM3 proposed in previous studies as well as the assessment indicators used, and comparing the DQM3 levels applicable in this study. The next stage was to determine the framework or the DQM3 to be used in this study. At this stage, the Loshin model was selected to be implemented in the DQMM measurements. Furthermore, data collection was done by distributing questionnaires to the BDPJN technical team that performs data management in Pustekdata. The next stage is the calculation of DQMM based on Loshin's model. After that, the maturity measurement

TABLE 3 Model components and previous studies on data quality management maturity model (DQMM).

Model	DQMM Components	Previous Studies
TIQM-CMM[25]	Goal, Policy, Process, Structure, Training	Fidler and Lavbic[26], Schäffer et al.[27]
Gartner[23]	Vision, Strategy, Governance, Organization, Process, Enabling Infrastructure, Metrics	M. Miftahul Akbar[28]
IQM-CMM[24]	Measurement and Analysis (MA), Organizational Process Performance (OPP), and Quantitative Project Management (QPM)	Schäffer et al.[27], Shrayner and Vladimir[29]
Loshin[8]	Data Quality Management Expectation, Dimension of Data Quality, Policies, Procedures, Governance, Standard, Technology, Performance Management	Sabtiana et al.[9], Wibisono et al.[10], Wilantika and Wibowo[11], Rahmawati and Ruldeviyani[14], Indriany et al.[15], Gunawan and Ruldeviyani[30], Setiadi et al.[31]

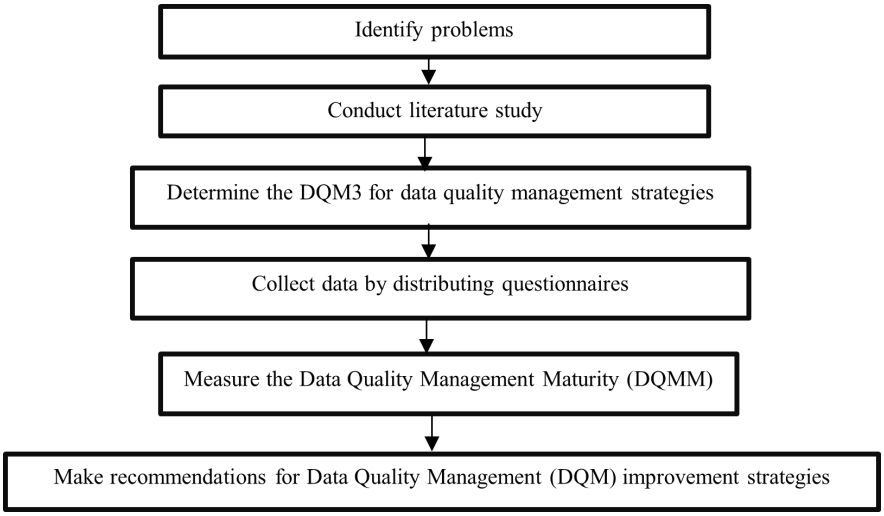

FIGURE 1 The research flow.

of DQM was conducted. In the last stage, several recommendations were proposed to improve the DQM of BDPJN based on Loshin's lacking data quality management characteristics, which were then mapped to DQM DMBOK activities.

3.2 | Research Instruments

A total of 115 questions were built based on Loshin's DQM3 as the research instrument of this study. The questionnaires comprise closed questions used to measure the maturity of data quality management. As mentioned previously, Loshin's DQM3 has eight components. To facilitate the data processing, each component was given an ID based on adaptations from previous studies[8]. The components and IDs are Expectation (E), Dimension (D), Policy (K), Procedure (P), Governance (G), Standardization (S), Technology (T), and Performance Management (M). Each component has several maturity levels, and each level has some characteristics. The codes for maturity levels are Initial (I), Repeatable (R), Defined (D), Managed (M), and Optimized (O). Meanwhile, the characteristic ID is a combination of component ID, maturity level, and characteristic number. The number of questions per component can be seen in Table 4 .

The questionnaire was divided into eight sections according to the number of components. It consists of closed questions where the answer is one of the two options, as shown in Table 5 . After the questionnaire was compiled, a readability test was conducted by one of the staff at Pustekdata, one of the implementers of BDPJN mosaic data management. After passing the readability test, the questionnaire was distributed to all respondents at Pustekdata before Pustekdata became a Remote Sensing Research Center.

TABLE 4 Number of questionnaire questions and category.

Question Category	#Question
Expectation (E)	17
Dimension (D)	12
Policy (K)	17
Procedure (P)	19
Governance (G)	17
Standardization (S)	15
Technology (T)	11
Performance Management (M)	7
Total	**115**

TABLE 5 Description of questionnaire answer scale.

Scale	Description
0	Not yet implemented
1	Has been implemented

3.3 | Data Collection

Data was collected by disseminating questionnaires to 24 employees who perform data processing and data management of medium, high, and very high-resolution image mosaics for BDPJN. Selected respondents are employees who know the evidence and information to answer the questions on the questionnaire. The questionnaires are online forms with an estimated time of filling out of 15-20 minutes. All respondents were asked to fill out the questionnaires in their free time so that they could answer each question slowly and carefully.

3.4 | DQMM Measurements

Calculations at one level and one component are the average values of all answers to the questionnaire. This means that the results of the calculations at one level and one component are 0 to 1. Meanwhile, the calculations per component are the total values of each level at 1 component. There are five maturity levels, meaning that if each level gets 1, the maximum value for 1 component is 5.

For example, there are three characteristics of the Expectation component at the Initial level, namely EI1, EI2, and EI3. If EI1 and EI2 have already been implemented, the value of each of them is 1. Meanwhile, EI3 has not been implemented, so it is worth 0. Then, the EI value can be seen in equation (1). Furthermore, the EI, ER, ED, EM, and EO values are added up to get the value of the Expectation (E) component, as formulated in Eq. 2.

$$EI = \frac{EI_1 + EI_2 + EI_3}{3} = \frac{1 + 1 + 0}{3} = 0.6 \tag{1}$$

$$E = EI + ER + ED + EM + EO \tag{2}$$

3.5 | Recommendation to Improve DQM

The data collection and calculation results of DQMM showed the characteristics of BDPJN DQM that still need to be implemented. These characteristics were mapped into activities that need to be done based on DQM activities in DMBOK. Through this mapping, organizations can determine the necessary steps to take as a DQM strategy to improve the quality of remote sensing data in BDPJN. According to DMBOK, there are 6 DQM activities with several sub-activities. Therefore, DQM encoding, activity number, and sub-activity number were used to facilitate the mapping. For example, the code DQM63 represented activity six and sub-activity 3, which measures and monitors data quality.

TABLE 6 Current and target maturity level of BDPJN DQM.

Components	Maturity Level	
	Current	Target
Expectation (E)	4.10 Managed	5 Optimizing
Dimension (D)	4.19 Managed	5 Optimizing
Policy (K)	3.84 Defined	5 Optimizing
Procedure (P)	3.54 Defined	5 Optimizing
Governance (G)	3.55 Defined	5 Optimizing
Standardization (S)	4.18 Managed	5 Optimizing
Technology (T)	3.21 Defined	5 Optimizing
Performance Management (M)	3.73 Defined	5 Optimizing

4 | RESULTS

The results of the maturity level evaluation are displayed in Table 6 . The maturity levels were obtained based on questionnaires filled up in November 2021 by 24 respondents who are Group Leaders, Leaders, and Operational Staff processing and managing medium, high, and very high-resolution remote sensing mosaic data. Meanwhile, the maturity target is the best level, which is the Optimizing level. According to the questionnaires, the maturity levels based on current conditions were as follows.

Expectation. Based on the results of the maturity measurement, the Expectation component has reached the Managed level, which means that Pustekdata has identified the dimensions of image data quality, including geometric correction level, radiometric correction level, data acquisition time, and cloud cover percentage. SNI standards have established geometric and radiometric correction procedures for some data resolutions. In addition, Pustekdata also understands the users' expectations for qualified BDPJN data. The shortcomings in this component are that Pustekdata needs to conduct periodic data quality evaluations and business impact analyses resulting from flawed data.

Dimension. Based on the maturity measurement in BDPJN, this component has reached the Managed level. The dimensions of BDPJN image data quality have been stated in Law No. 21 of 2013. Some details have been explained in the SNI standard, while others are still in the process of standard design (RSNI). Several employees understand the general dimensions of geographic data quality according to SNI ISO 19157:2015 Geographic Information – Data Quality. Data quality SLAs already exist according to ISO 9001, but some of their implementations still need to be properly monitored. In addition, there have been no reports in the form of a data quality matrix.

Policy. The result of the maturity measurement on the Policy component shows the Defined level. DQM activities have existed through the implementation of ISO 9001 at Pustekdata. The shortcoming of this component is that there is no automatic notification to discover inappropriate data. In addition, socialization at the staff level regarding the data governance guidelines used to achieve management objectives is extremely needed.

Procedure. The maturity of the procedure component is at the defined level. The data-checking procedure has been done automatically/semi-automatically. However, several obstacles in identifying data errors remain due to data structure errors or because data needs to be completed. The root cause of the problem affecting the data has yet to be fully identified. Errors or weaknesses in data are sometimes known after the data is needed or used by the users. Data quality control has yet to be fully implemented in all working groups. Besides, the results of BDPJN data quality measurements still need to be transparent.

Governance. The Governance component has a Defined level. At Pustekdata, there has been communication to discuss data quality management. Every BDPJN data has a person in charge, and data governance policies are designed and recommended by working group representatives. However, some staff considers data quality issues as IT problems.

Moreover, best practices regarding data quality management have yet to be thoroughly socialized, and there has been no establishment of special committees and regular meetings that discuss data management. In addition, the provision of rewards for the performance of data governance in the operational field has yet to be implemented.

Standardization. Based on the results of the maturity measurement, the Standardization component has reached the Managed level. BDPJN data already has SNI or RSNI standards. All data migration or exchange activities follow the established procedures, and although some staff assesses that data duplication still exists, some data do not have metadata.

TABLE 7 Mapping of not achieving Loshin's characteristics to DMBOK DQM activities.

Component Level	ID	Current Condition	DMBOK	DQM Activities
Expectation (E)	Managed	EM2	Pustekdata has yet to analyze the resulting business impact if the BDPJN data is flawed.	DQM51
		EM4	Evaluation of data quality has not been regularly scheduled	DQM63
Dimension (D)	Managed	DM2	There have been no reports in the form of a data quality matrix	DQM64
Policy (K)	Optimizing	KO1	There are no automatic notifications to find out inappropriate data	DQM52
Procedure (P)	Repeatable	PR1	Data errors due to incomplete data are difficult to identify	DQM52
		PR2	Data errors due to data structure errors are difficult to identify	DQM52
	Managed	PM3	Some of the weaknesses of BDPJN data are known after the users receive the data.	DQM62
	Optimizing	PO2	Every employee has been unable to publish the results of BDPJN data quality measurements.	DQM64
Governance (G)	Initial	GI2	Some staff consider the data quality issues as IT problems	DQM42
	Repeatable	GR2	Best practices on data quality management have yet to be socialized throughout the organization.	DQM21
	Managed	GM1	A committee has yet to be formed from various representatives of working groups that specifically handle data management.	DQM61
		GM2	Regular meetings have yet to be made by a committee of the working group representatives to discuss data management.	DQM61
	Optimizing	GO2	There are no rewards or awards for staff with good BDPJN data governance performance.	DQM21
Standardization (S)	Initial	SI2	Some staff believe that there are duplications of data	DQM42
	Repeatable	SR1	Some staff judge that there are data that do not have metadata	DQM42
Technology (T)	Initial	TI1	Ad hoc work is done routinely	DQM61
	Repeatable	TR1	There are no tools/technologies to measure BDPJN data quality	DQM52
		TR2	Data standardization technology is not yet available	DQM52
	Managed	TM1	There are no dashboards or reporting applications for operational monitoring and analysis.	DQM64
	Optimized	TO1	Non-technical users have not been able to process and utilize BDPJN data	DQM61
Performance Management (M)	Defined	MD1	Pustekdata still needs to establish a framework for analyzing the impact of BDPJN data quality.	DQM51

Technology. BDPJN is at a Defined level on the Technology component. Pustekdata embraces the application of new technology; there are already applications in the form of data catalogs for data retrieval, which have become a standard in data services for users. However, ad hoc work related to the use or development of technology is still done regularly, there is no technology to standardize data and measure data quality, and there needs to be a dashboard for operational reporting or monitoring. In addition, no technology facilitates non-technical users to take advantage of BDPJN data.

Performance Management. The maturity measurement on the Performance Management component shows that this component is at a defined level. According to most respondents, Pustekdata is aware of the impact of poor data quality on certain users. Nevertheless, no framework has yet been established for analyzing the impact of BDPJN data quality.

5 | DISCUSSION

Some of the statements or characteristics of Loshin's questionnaires need to be implemented optimally if most respondents answered that they had not been implemented. To increase the maturity level of BDPJN, a mapping was carried out between the current conditions that have yet to be achieved and the DQM DMBOK activities. The results of the mapping can be seen in Table 7 .

According to Tabel 7 , the recommendation is as follows:

1. **Expectation**. Pustekdata has been holding FGD annually to gather specifications of stakeholder data needs. Through these activities, Pustekdata or Pusdatin, which will later manage the BDPJN, can formulate expectations and dimensions of data quality needed by its stakeholders. However, based on the questionnaires, many staff still need to understand the business impact of low-quality data or tolerance of data quality. Therefore, each data steward team needs to define data quality based on the inputs from stakeholders at the annual event. Each representative of the data management team needs to hold regular meetings to evaluate the quality of BDPJN data at least once every quarter.

2. **Dimension**. Data quality management should be applied in all lines. A data quality matrix containing information on data quality dimensions should be listed for all data types. This information is presented in the metadata, and this is needed as information that data users often need.

3. **Policy**. There is still duplication of data according to the Standardization component. Developing a semi-automatic/automatic program that can identify duplicated data, especially data sourced from more than one ground station, is necessary. Duplication can be found in old data where the database still needs to be integrated.

4. **Procedure**. Data operations have implemented a Service Level Agreement (SLA) for data quality management. Each member of staff needs to be disciplined to always refer to the existing SLA. The SLA can also include the dimensions of image data quality. In addition, the resolution of the time required for each job can be determined so that the compliance of each staff member with the rules of data quality management with business expectations can be achieved properly.

5. **Governance**. To increase the maturity level, regular meetings between teams are needed to discuss data governance so that the occurring issues on data quality can be socialized and resolved together. Thus, each person will understand the best practices that should be done. It is necessary to establish a data quality control committee from representatives of team members and provide rewards to employees with good data governance performance.

6. **Standardization**. Documentation of data quality in metadata also needs to be considered. Based on the results of the assessment, it is found that some data still need metadata. Scene-based data usually already has standard ground station metadata, while mosaic data sometimes needs metadata that provides detailed information regarding data quality. The built metadata may contain information on data quality dimensions, i.e., geometric correction, radiometric correction level, acquisition time, and percentage of cloud cover. Information on data quality must be accessible to all users, either through catalogs or websites available at Pustekdata or Pusdatin.

7. **Technology**. Data quality measurement and monitoring must be included in the operational reports submitted by each team. It is necessary to create a dashboard to monitor operational data to maintain data quality. Ad hoc work that is done routinely needs to be done as a special team, especially for technology development. The scene base data already has a default standardization technology, while the mosaic product data needs to be developed further to comply with the standards. In addition, it is necessary to develop procedures or technology that can facilitate non-technical users.

8. **Performance Management**. According to Loshin, the impact classification scheme of improving data quality is financial, user satisfaction, productivity, risk, and compliance[8]. In the case of BDPJN, the impact classification on user satisfaction and productivity can be identified and explored to be used as a performance improvement matrix.

To determine the most critical issues, the questionnaires were distributed and discussed with staff representatives who have worked for over ten years, five years, and 1-5 years, carry out the operational activities and know the data quality issues to be considered as representatives of various segments. From the discussion, the most critical issues were identified, namely (1) the need for regular meetings between related operational groups to discuss data management issues that will have an impact on data quality; (2) the creation of notifications to discover inappropriate data; (3) the improvement of data duplication; (4) the provision of complete metadata; and (5) the development of technology or procedures for the utilization of BDPJN data that facilitate non-technical users. In addition, socialization on the best practices of data governance for all working groups is incredibly necessary.

6 | CONCLUSION

Based on the measurement results, the maturity levels of the eight components of BDPJN data quality management were obtained as follows: Expectation is 4.10 (Managed), Dimension is 4.19 (Managed), Policy is 3.84 (Defined), the procedure is 3.54 (Defined), Governance is 3.55 (Defined) Standardization is 4.18 (Managed), technology is 3.21 (Defined), and Performance Management is 3.73 (Defined). The average maturity level is 3.79, meaning it has reached the Defined Managed level. The Defined level indicates that BDPJN already has a team responsible for data quality and documentation of best practices and has begun to adapt DQM in all working groups. In addition, based on the results of the maturity level measurement, there are

21 characteristics of Loshin's DQM3 that need to be done to increase the maturity of BDPJN data quality management towards the Optimizing level. Technology, Governance, and Procedure become the components with the lowest maturity levels.

Recommendations that need to be made based on the results of the discussions on the most critical issues are: (1) the need to hold regular meetings to discuss the data management issues; (2) the creation of notifications to discover inappropriate data; (3) the improvement of data duplication; (4) the provision of complete metadata; (5) the development of technology or procedures to make it easier for non-technical users to utilize BDPJN data, and (6) the socialization of the best practices of data governance to all working groups. Data standardization technology can be developed and potentially be research material in future studies in this field.

The results of this study prove that Loshin's DQM3 can be used as a data quality management strategy by measuring the maturity of data quality management. In this study, the scope of measurement was limited to the operation of cloud-free mosaic data in low, medium, high, and very high resolution. Further studies in the future can carry out measurements of the entire working group to assess the maturity of data quality management. In addition, calculations using GIS maturity models to measure the maturity in organizations and the use of GIS by remote sensing data users can also be considered.

CREDIT

Agnes Sondita Payani: Conceptualization, Writing - original draft preparation, Formal investigation, Formal analysis, Methodology, and Resources. **Alia Mutia Mayanda:** Conceptualization, Methodology, Writing - original draft preparation, Formal investigation, and Resources, Writing - editing, and Formal analysis. **Adenia Adiresta:** Conceptualization, Methodology, Formal investigation, Writing – review, Supervision, and Resources. **Yova Ruldeviyani:** Conceptualization, Writing – review, Formal investigation, Resources, and Supervision.

References

1. Pemerintah Indonesia, Undang-Undang Republik Indonesia Nomor 21 Tahun 2013 Tentang Keantariksaan. LN.2013/No. 133, TLN No. 5435, LL SETNEG: 44 HLM; 2013.

2. Payani AS, Yudha GD, Wahyuningsih SD, Kannia N, Kusapy DS, Simanjuntak D, et al. Uji Keberhasilan Implementasi Master Plan Teknologi Informasi dan Manajemen BDPJN Menggunakan Metode DeLone dan McLean. Jurnal Edukasi dan Penelitian Informatika 2020;6(2):258–266. https://jurnal.untan.ac.id/index.php/jepin/article/view/37757.

3. Brahmantara RP, Hutapea DY, Ruldeviyani Y. Evaluation on Data Operations Management using CMMI and DMBOK: BDPJN Case Study. In: Proceedings of The 9th International Conference on Cyber and IT Service Management Bemgkulu, Indonesia: IEEE; 2021. p. 69–75. https://ieeexplore.ieee.org/document/9588888.

4. Al-Salim W, Darwish ASK, Farrell P. Analysing data quality frameworks and evaluating the statistical output of United Nations Sustainable Development Goals' reports. Renewable Energy and Environmental Sustainability 2022;7(17):1–12. https://ui.adsabs.harvard.edu/abs/2022REES....7...17A/abstract.

5. Bowo WA, Suhanto A, Naisuty M, Ma'mun S, Hidayanto AN, Habsari IC. Data Quality Assessment: A Case Study of PT JAS Using TDQM Framework. In: Proceedings of The 4th International Conference on Informatics and Computing Semarang, Indonesia; 2019. p. 1–6. https://ieeexplore.ieee.org/abstract/document/8985896.

6. Cichy C, Rass S. An overview of data quality frameworks. IEEE Access 2019;7:24634–24648. https://ieeexplore.ieee.org/document/8642813.

7. Izmaya H, Purnamasari D. Pengukuran Indeks Kepuasan Masyarakat Terhadap Pelayanan Dan Kualitas Data Penginderaan Jauh. Jurnal Informatika: Jurnal Pengembangan IT 2021;6(2):93–98. http://ejournal.poltektegal.ac.id/index.php/informatika/article/view/497/1519.

8. Loshin D. The Practitioner's Guide to Data Quality Improvement. Elsevier Inc.; 2011. http://www.sciencedirect.com/science/article/pii/B9780123737175000191.

9. Sabtiana R, Yudhoatmojo SB, Hidayanto AN. Data Quality Management Maturity Model: A Case Study in BPS-Statistics of Kaur Regency, Bengkulu Province, 2017. In: Proceedings of The 6th International Conference on Cyber and IT Service Management Parapat, Indonesia: IEEE; 2018. p. 1–4. https://ieeexplore.ieee.org/abstract/document/8674323.

10. Wibisono SB, Hidayanto AN, Nugroho WS. Data Quality Management Maturity Measurement of Government-Owned Property Transaction in BMKG. CommIT (Communication and Information Technology) Journal 2018;12(2):59. https://journal.binus.ac.id/index.php/commit/article/view/4470.

11. Wilantika N, Wibowo WC. Data Quality Management in Educational Data. Jurnal Sistem Informasi 2019;15(2):52–67. https://jsi.cs.ui.ac.id/index.php/jsi/article/view/848.

12. Henderson D, editor. DMBOK - Data Management Body of Knowledge 2nd Edition. New Jersey, USA: Technics Publication; 2017.

13. Putri DT, Ruldeviyani Y. Data Quality Assessment for Suspicious Report Using Loshin'S Framework and Dmbok : a Case Study of Financial Transaction Analysis. Jurnal Penelitian Teknologi Informasi dan Komunikasi 2021;12(1):17–26.

14. Rahmawati SD, Ruldeviyani Y. Data Quality Management Strategy to Improve the Quality of Worker's Wage and Income Data: A Case Study in BPS-Statistics Indonesia, 2018. In: Proceedings of The 4th International Conference on Informatics and Computing Semarang, Indonesia: IEEE; 2019. p. 1–6. https://ieeexplore.ieee.org/document/8985803.

15. Indriany HS, Hidayanto AN, Wantania LJ, Santoso B, Putri WU, Pinuri W. Data Quality Management Maturity: Case Study National Narcotics Board. In: Proceedings of The 10th IEEE International Conference on Communication, Networks and Satellite Purwokerto, Indonesia: IEEE; 2021. p. 206–212. https://ieeexplore.ieee.org/document/9530824.

16. Pratiwi M, Ruldeviyani Y. Data Governance to Improve Data Quality for Statistical Process Control (SPC): A Case Study PT. XYZ. In: Proceedings of The 4th International Conference on Software Engineering and Information Management Yokohama, Japan: Association for Computing Machinery; 2021. p. 99–105. https://dl.acm.org/doi/10.1145/3451471.3451488.

17. Agung Budi Prasetyo, Muhamad Irfan Darmawan RM. Analisis Dan Perancangan Tata Kelola Data Sistem Pemerintahan Berbasis Elektronik Domain Data Quality Management Pada DAMA DMBOK v2. Journal eProceedings of Engineering 2019;6(2):7775–7786. https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/10880.

18. Wibowo NS, Utami E, Fatta HA. Perancangan Struktur Tata Kelola Data di Pemerintah Daerah Menggunakan Framework Data Management Body Of Knowledge. Journal of Computer, Information System & Technology Management 2021;4(1):45–52. http://e-journal.unipma.ac.id/index.php/RESEARCH/article/view/7049/0.

19. Rianto, Shofa RN, Yusuf E. Implementasi Kualitas Data dalam Peran Tata Kelola Data. Jurnal Siliwangi 2020;6(1):44–52. https://jurnal.unsil.ac.id/index.php/jssainstek/article/view/2508.

20. Widjaya W, Sutedja I, Hartono AW. Key aspects of data management framework for early adopter: A systematic literature review. ICIC Express Letters 2019;13(9):761–771.

21. Cai L, Zhu Y. The challenges of data quality and data quality assessment in the big data era. Data Science Journal 2015;14:1–10. https://datascience.codata.org/articles/10.5334/dsj-2015-002/.

22. BSN, Badan Informasi Geospasial, Informasi Geografi/Geomatika, editor, SNI ISO 19157:2015 Informasi geografis - Kualitas data. BSN; 2015.

23. Bitterer A, Gartner' s Data Quality Maturity Model. Gartner; 2007. https://www.gartner.com/en/documents/500906.

24. Baskarada S. IQM-CMM: Information Quality Management Capability Maturity Model. Wiesbaden GmbH, Wiesbaden: Vieweg+Teubner Verlag Wiesbaden; 2008. https://link.springer.com/book/10.1007/978-3-8348-9634-6.

25. Suhardi, Gunawan IGNAR, Dewi AY. Total Information Quality Management-Capability Maturity Model (TIQM-CMM): An information quality management maturity model. In: Proceedings of 2014 International Conference on Data and

Software Engineering Bandung, Indonesia; 2014. p. 1–6. https://ieeexplore.ieee.org/abstract/document/7062675.

26. Fidler M, Lavbic D. Improving information quality of Wikipedia articles with cooperative principle. Online Information Review 2017;41(6):797–811. https://www.emerald.com/insight/content/doi/10.1108/OIR-01-2016-0003/full/html.

27. Schäffer T, Leyh C, Bley K, Schimmele M. Towards an open ecosystem for maturity models in the digital era: The example of the data quality management perspective. In: Proceedings of The Americas Conference on Information Systems 2018: Digital Disruption; 2018. p. 1–10.

28. M Miftahul Akbar. Evaluasi Tingkat Kematangan E-government Dan Partisipasi Masyarakat dalam Pelayanan Publik (Studi Kasus Kabupaten Sukoharjo). PhD thesis, Universitas Islam Indonesia; 2020.

29. Shrayner YS, Vladimir YV. Maturity models of quality management system in high-tech industry: A systematic literature review. In: Proceedings of The 2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering, ElConRus 2019 Saint Petersburg and Moscow, Russia: IEEE; 2019. p. 1478–1484. https://ieeexplore.ieee.org/document/8656993.

30. Gunawan F, Ruldeviyani Y. Improving Data Quality in Crowdsourced Data for Indonesian Election Monitor: A Case Study in KawalPilpres. Journal of Physics: Conference Series 2020;1566(1):0–6. https://iopscience.iop.org/article/10.1088/1742-6596/1566/1/012095.

31. Setiadi Y, Hidayanto AN, Rachmawati F, Yohannes AYL. Data Quality Management Maturity Model : A Case Study in Higher Education's Human Resource Department. In: Proceedings of The 7th International Conference on Computing, Engineering and Design Sukabumi, Indonesia: IEEE; 2021. p. 1–6. https://ieeexplore.ieee.org/document/9664881.

How to cite this article: Payani A.S., Mayanda A.M., Adiresta A., Ruldeviyani Y. (2022), Data Quality Management Strategy To Improve Remote Sensing Data Quality: A Case Study On National Remote Sensing Data Bank, *IPTEK The Journal of Technology and Science*, *33(3):162-174*.