The Maturity Model of Data Quality Management in Banking Industry: PT XYZ Core System Customer Data

Rahmad Mulyadi[1*], Yova Ruldeviyani[2], Noverina Alfiany[3], Achmad Nizar Hidayanto[4]
[1,2,4] Faculty of Computer Science, Universitas Indonesia
[3] Faculty of Mathematics and Natural Science, Universitas Andalas
*email: rahmad.edu@gmail.com

ABSTRACT

PT XYZ, engaged in the financial industry, aims to become a leading company in Southeast Asia. Its core system has supported it with more than 200 million customer data. This massive amount of data is expected to create business opportunities, build a risk-aware culture, and increase supremacy in the business strategy of PT XYZ. These things can be achieved if the data used is of good quality data. In fact, found anomalies in a large number of customer data. In order to get recommendations for improving the quality of customer data, it is necessary to assess the quality of customer data. The customer data quality assessment in this study uses the method introduced by Loshin (2011). Loshin's Data Quality Management Model (DQMM) adopts a capability maturity level model in building its criteria matrix. Maturity levels obtained are 3.6 (expectations), 3.6 (dimensions), 4.4 (policy), 3.8 (procedures), 4.2 (governance), 3.8 (standardization), 4, 2 (technology), and 3.8 (performance management). Regarding the expectation that senior management can achieve the highest level of data quality, nine strategic recommendations were produced nine strategy recommendations were submitted to PT XYZ as the result of mapping between criteria that have not been met with data quality management activity in Data Management Body of Knowledge (DMBOK) version 2.0. Measurement and monitoring of good data quality is the most influential recommendation for PT XYZ.

Keywords: *DQMM, Financial Industry, Core System, Loshin's Management Model, Customer Data Quality, DMBOK.*

INTRODUCTION

Data is a precious and critical thing in today's digital era for both individuals and organizations or companies. Data can be used to simplify search, observe, and gain something. Organizations in the form of government make data as an added value to support the government's strategy and mission [1]. In this case, data is used to simplify the process of getting the ideal strategy, observing the domestic and foreign strategies of other countries, and winning the ideal strategy that can be used. Business organizations or companies that use data-driven decisions (DDD) produce 5-6% yield more than expected [2].

Good-quality data can be used for many things. Good data quality will have a positive impact on an organization or company in business processes and various forms of business fields. Good quality data becomes the basis for decision-making and will save a lot of costs for the data recovery process [3]. To be able to determine whether the data includes good quality data or not by measuring using parameters. The parameters that can usually be used in data measurement are the dimensions of data quality. The dimension of data quality is considered as a form of the characteristics of the data itself that can be measured to represent various aspects of data, such as data quality [4]. Currently, all organizations, not limited to

companies, have used data, especially good quality data, to support their business processes, including PT XYZ.

PT XYZ is one of the largest financial companies engaged in Southeast Asia. The company has more than 200 million customers. Therefore, there are more than 200 million pieces of data that PT XYZ must process to support the business. With such a massive volume of customer data, we found a large amount of poor-quality data. Data quality problems that occur at PT XYZ, among other data redundancy and poor-quality data, are found. In the observations made, there are duplications of customer data. Observation is done by selecting customer data in the core system of PT XYZ. The selection is made using a filter query, where the condition 'where unique number identifier' has been duplicated.

Based on some of the factors mentioned above, it can be concluded that there is a data distribution anomaly. The anomaly is a crucial point to be fixed immediately so that the decision-making process regarding data at PT XYZ becomes better. In addition, there are regulations that PT XYZ must comply with as a financial service provider. These regulations address many issues, including sanctions that financial service providers may face in relation to their customers. The reference regulations must be complied with by PT XYZ are Indonesian central bank Regulations No. 22/20/PBI/2020 regarding Bank Indonesia's Consumer Protection and Financial Services Authority Regulations No. 12 /POJK.03/2018 regarding the Implementation of Digital Banking Services by Commercial Banks. Of course, this has the potential to have a negative impact on the company, such as errors in business operations, inaccurate marketing business analysis, and even to mistakes in determining the company's strategy [5][6].

As the most prominent financial company, PT XYZ is required to comply with regulatory requirements such as submitting reports related to its customers. The reporting must be correct dan reflects the current condition of PT XYZ. If any data in the report is of poor quality, then the regulator will offer a warning to PT XYZ. Regulators also conduct audits of customer data at PT XYZ. Based on information from the data management work unit at PT XYZ, it was stated that there had been audit findings that specifically mentioned the quality of PT XYZ's customer data. Such findings on data quality are made by audits of financial services regulators.

Thus, it can be concluded that there are problems with customer data in the core system of PT XYZ still contains data that is not of good quality and not up to date, and there are still data redundancies. This leads to two major impacts faced by PT XYZ related to its customer data problems, namely: PT XYZ cannot maximally comply with regulations from the Indonesian central bank and the financial services authority, and PT XYZ is not able to maximize customer data to support PT XYZ operations.

One of the solutions to solving problems that can be done based on analysis using the people, process, and technology approach is to conduct a data quality maturity assessment. From the process area, it is mentioned that there has been an assessment of data quality, but this has not been proven because there is no documentation of these activities. The stages and results expected to arise from the process cannot be applied to customer data in the core system of PT XYZ.

In conducting a data quality maturity assessment, the DQMM (Data Quality Maturity Model) method with Loshin's framework approach is a commonly used method [10][11][19].

To generate strategy recommendations for improving data quality, DQMM is considered easier to implement. Slightly different from previous studies, The DQMM used in this study adheres closely to the rules of the CMM (Capability Maturity Model), where a higher level of maturity will not be obtained if all criteria at the previous level have not been fully achieved. This research also raises themes that are rarely discussed in the banking sector in Indonesia, namely, related to data quality management.

The main objective of this research is to measure the quality level of customer data in the core system of PT XYZ. To be able to deliver more benefits, especially to PT XYZ, this research also aims to provide recommendations for data quality management strategies in an effort to improve customer data quality at PT XYZ. The high-level recommendations provided are based on best practices in DMBOK. The results of this study are expected to be positive recommendations for PT XYZ to improve customer data quality. This research is also expected to be an additional reference for PT XYZ in an effort to improve data quality.

This article consists of 5 main sections with the following structure: The background that is the core of this research is explained at the beginning or in section 1. Section 2 describes various theories that are closely related to this research. It also represents the comparison and application of the frameworks used. Section 3 describes the methodology used to complete this research. The methodology starts with the research design, stages, and instruments used in the research. Section 4 explains the results of the measurements done and the explanation of conditions of each indicator used for recommendations that might be implemented at PT XYZ. Finally, Section 5 presents the conclusions and limitations of this research and what might be improved in future research.

LITERATURE REVIEW

1. Data, information, and knowledge

Data is defined as a representation of a fact or numbers associated with an object. Data currently stored digitally can be displayed in text form, numbers, graphs, images, sounds, and videos [4][7]. Data is the initial foundation of the formation of information. Knowledge is formed from various information and integrated experiences and develops a particular pattern. The pattern formed from the integration can easily be translated and has added value for an organization or company. Data can be used in different timeframes, both short-term and can be implicated in medium-term or long-term [4][8][9].

Data can be referred to as a fact about a situation that can be audio or visual. Data currently stored digitally has become crucial for organizations or companies, especially in determining short, medium, and even long-term strategies. The preparation of this strategy must be based on knowledge composed of a lot of information, which is obtained from a set of valuable or quality data.

2. Data quality

DMBOK: DAMA International stated that the fitness for use or benefits to the goals and expectations of data users [7]. Good quality data is data that can benefit the goals and expectations of its users. The need for good quality data applies to all organizations or companies, especially to companies that are transforming and innovating digitally [12].

Data quality management research or activities have been a concern since the mid-20th century AD [13]. Data quality issues persist due to the popularity of the data itself [14]. Nikiforova mentioned that there had been an analysis of more than 70 solutions that have been done on data quality. The research results revealed that most of the solutions generated were based on definition, grouping data quality dimensions, and their application to data sets. Data quality can be stated as the fulfillment of the satisfaction level of data users, which represents all types of objects expected by data users. Companies that have good data quality will get a high reputation in front of customers and even competitors. Research on data quality will continue to be of interest because of the popularity of the data itself.

3. Data quality management

Data quality management is a series of efforts or activities that are carried out continuously to improve data quality [15]. The action starts with determining the parameters for measuring the level of data quality, analysis, identification of anomalies, and correction process. These activities are expected to ensure that the information produced can meet the expectations and needs of data users.

Data Quality Maturity or Data Quality Capability Model is a form of framework used to measure the maturity level of data quality. The framework consists of eight components that need to be identified: data quality expectation, dimension of data quality, policy, procedure, governance, standard, technology, and dan performance management [16]. Based on the Capability Maturity Model (CMM) designed by the Software Engineering Institute at Carnegie Mellon University, there are five levels of maturity in the data quality of an organization, namely initial, repeatable, defined, managed, and optimizing.

The initial stage that must be done to develop Data Quality Management recommendations for an organization or company is to assess or measure the maturity of the current data quality [13]. To be able to measure data quality, there are various methodologies that can be used. The two most common measurement methodologies are TDQM and DQMM. TDQM (Total Data Quality Management) was introduced by Wang., (1998) and can be used for structured data as well as semi-structured data. In practice, TQDM activities are carried out using the TQDM cycle implemented by the previous researcher [17].

Broadly speaking, TDQM can be expressed as a quantitative method used to improve product or service quality using data migration. TDQM also enhances the use of data standards and business rules in data quality improvement. In its implementation, TDQM is carried out using four main stages: define, measure, analyze, and improve [18].

DQMM (Data Quality Management Maturity or Data Quality Maturity Model) is a form of assessing data quality by forming a model in the form of a gap between expectations and existing reality. The gap is obtained by determining data quality using each of the criteria used [19]. The most commonly used criterion for determining the maturity level of data quality is Loshin's Data Quality Maturity Model.

The classification of the assessment indicator for the maturity level of data quality is divided into eight: expectation, dimensions, policy, procedure, governance, standardization, technology, and performance management. By using the capability maturity model, a mapping of the maturity level can be done based on existing criteria.

In Corporate Data Quality Management, Lucas., 2010 uses DQMM to assess the achievement and performance of data quality. The use of DQMM tends to represent stakeholders' understanding of data quality. [19]. DQMM, in particular, adopted Loshin's framework, using the approach of extracting information on stakeholders from several aspects.

Both TDQM and DQMM are the methods used in previous comparative studies. The use of DQMM, especially those that adopt Loshin's framework, is the most feasible approach. Other methods, such as TDQM, require longer processes and procedures. This is because there are several stages of the process that must go through the procedures that apply at PT XYZ. One of the procedures involved by PT XYZ in the use of detailed company data is related to licensing, which requires the approval of several stakeholders, including senior management. Furthermore, a specific non-disclosure agreement is also needed for each data to be observed. Licensing the use of detailed data that can be used in measuring each dimension in TDQM requires a lot of resources and time.

DQMM using Loshin's framework, also has a complete data quality maturity level assessment dimension. The focus of this method is on improving data quality management. This method makes it easier to determine recommendations for strategies to improve data quality. One of them is the existence of criteria that can be applied at each level. The recommendation set has been used in several previous studies.

METHOD

1. Research Design

This research is applied research, which is conducted as an answer to questions at a practical level and can be applied directly in the primary research sector. Substances found in the basic research will be used as the subject in an effort to provide practical solutions. More specifically, this research belongs to evaluation research. Evaluation research is suitable for answering questions about the need, design, development, and outcomes of various activities carried out [18]. The research category is a case study with the object of research being a financial company, namely, PT XYZ.

2. Research Stage

    a. Data collection is carried out to explore further the problems, current conditions, and expectations of data quality at PT XYZ. Data collection is carried out through personal interview activities, observation, and review of various documents, especially internal documents. Interviews were conducted with experts closely related to data quality at PT XYZ. These experts have different qualification criteria, but they must be recognized as independent specialists in at least one of the fields relevant to the research being conducted [21]. There are several criteria for the experts interviewed, they must have tenure at PT XYZ of at least eight years, have international certifications, and have held manager-level or higher positions. Eight respondents meet the criteria consisting of; 3 department heads, five people have tenure above ten years, and three others range between 8 and 9 years, and all of them have international certifications such as CDMP, ISO/IEC 9001:2015, and ISO/IEC 27001:2013. In the document review, at least eight

internal documents are used. The internal document regulates data management, including governance and security.

b. After recognizing the problem from the data collection process, a literature review is carried out to obtain references for solving problems in the form of various scientific literature. By utilizing scientific literature in the form of previous research relevant to this research, method 3C+2S, compare, contrast, criticize, summarize, and synthesize, is used [22]. In this stage, five previous studies reviewed similar research, as shown in Table 1.

Table 1. Previous Studies

Author	Case Study	Methodology
Setiadi et al., (2021)	University of Indonesia (educational industries)	DQMM
Andini et al., (2020)	PT Telekomunikasi Indonesia (telecommunication industries)	TDQM
Bowo et al., (2019)	PT JAS (Financial industries)	TDQM
Sabtiana et al, (2018)	BPS-Statistics of Kaur Regency (government organization)	DQMM
Wibisono et al., (2018)	BMKG (government organization)	DQMM

c. Furthermore, research preparation is carried out using a predetermined methodology, DQMM, with Loshin's framework approach. The practice or formulation of the research is done to obtain various parameters and their use in calculating the level of data quality. This stage also defined all the research tools used.

d. Then the data quality level is calculated to get data quality expectations and current conditions at PT XYZ. Both the current state and data quality expectations are obtained using all criteria in the data quality indicators contained in Loshin's framework. For the current state of data quality, each criterion at a level must be fully met in order to move to the next level. For example, in the data expectation indicator, three of the five criteria at level 4 are met, so the data expectation indicator has not been able to reach the managed level. Figure 1 illustrates the methodology process for calculating quality-level data. Meanwhile, in calculating the expected value of data quality, higher criteria are used for each data quality indicator that has been met in the current condition. The Quality Level Data Calculation Process is shown in Figure 1.
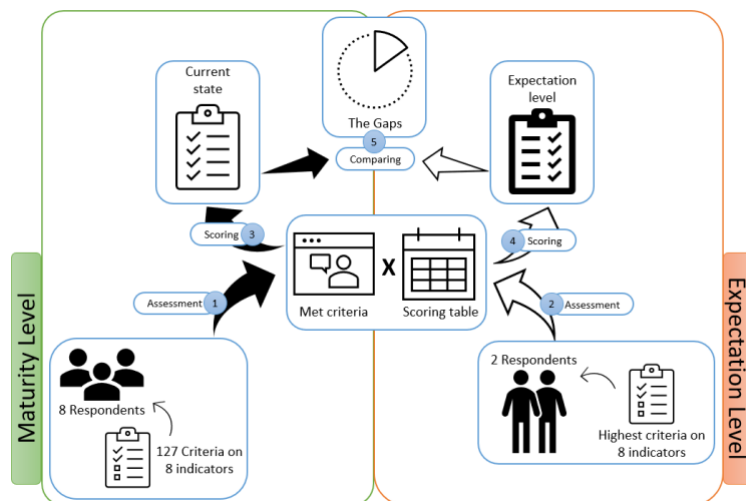

Figure 1. Overview of the Quality Level Data Calculation Process

e. Analysis of measurement results and preparation of recommendations is the next stage of the research conducted. The analysis is carried out by finding gaps between expectations and the current state of data quality at PT XYZ. The gap results will be mapped using Data Quality Management (DQM) Activity version 2.0 contained in the DMBOK – DAMA International 2017 to produce recommendations that best suit the needs of data quality at PT XYZ.

f. The last stage is the validation of the measurement results and recommendations to top-level management. Validation is done by presenting all measurement results and recommendations, then soliciting feedback and confirming whether all recommendations can be implemented. All feedback, including suggestions for improving the recommendations, is recorded until the final recommendations presented are recommendations that have been approved and refined by the top-level management of PT XYZ.

3. Research Instrument

a. Data quality maturity level criteria matrix based on the data quality management framework introduced by David Loshin. In this study, eight indicators with 127 criteria were used. The criteria matrix is compiled using the capability maturity model (CMM) level.

b. Based on the answers given by the respondents, a mapping of the predetermined maturity level is carried out. Each answer will contain the values x and 0. Score x is given for answers with positive connotations such as 'True,' 'Yes,' and 'Have,' and score 0 for answers with negative connotations, 'False,' 'No,' and 'Don't have.' In contrast to the scoring done by the previous researchers [12][19][23], the value of x is determined by dividing the maximum maturity score by the total number of criteria in each component in Loshin's Data Quality Management Maturity Model Matrix. Generally, the score is defined by

$$x = \frac{\max(Maturity\ score)}{\sum_{i=1}^{k} Criteria_i} \tag{1}$$

where k is the total number of criteria.

In data processing of maturity value calculations, the accumulated results of the calculation of each criterion are used decimal numbers with one number behind the comma. To determine these values, rounding up is also used for each value. In calculating the accumulation of scores that get more than 5 (five), natural numbers or positive integers with a value of 5 are used. Table 2 shows the scoring values for each criterion met.

c. Theory and practice found in DAMA-DMBOK (Data Management Body of Knowledge) or Data Quality Management (DQM) Activity version 2.0 is used as a tool to provide recommendations for improving data quality. The recommendations were developed using gap analysis. The gap will result in different components that must be fulfilled. This is done so that the recommendations produced become ideal. Table 2 shows the coding and DQM Activity version 2.0 used.

Table 2. Scoring for Positive Connotation Answer

Data Quality Indicator	*Criteria*	*Scoring*
Data Expectation	18	0,278
Data Dimensions	14	0,357
Data Policy	17	0,294
Data Procedure	20	0,250
Data Governance	18	0,278
Data Standardization	20	0,250
Data Technology	12	0,417
Data Performance Management	8	0,625

RESULTS AND DISCUSSION

For the current data quality maturity and the senior management expectation level value at PT XYZ has been found. Both values can be visualized, as shown in Table 3. Next, the difference between the two values is calculated to produce a gap. The following will discuss the gaps of each indicator which has also been described previously for each condition.

Using the test results as a form of validating the values measured for each data quality indicator used can be described as the form of data quality mapping that exists at PT XYZ. Figure 2 shows the visualization results of the current condition and expectations that will be addressed by data quality at PT XYZ for customer data based on Table 3.
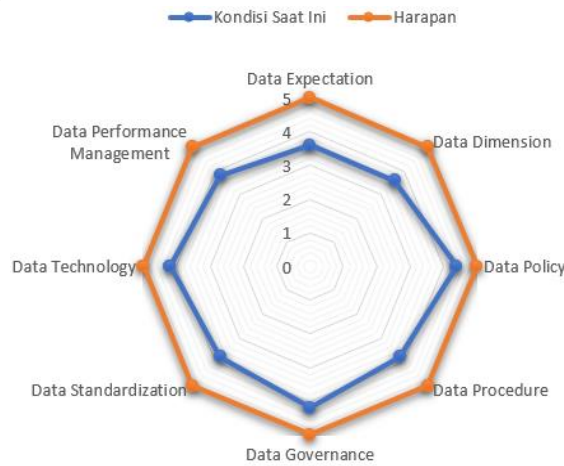

Figure 2. Data Quality Visualization Gap at PT XYZ

Table 3. Measurement Results & Expectations Indicator Data Quality

Data Quality Indicator	*Current Level*	*Expectation Level*
Data Expectation	3,6	5
Data Dimensions	3,6	5
Data Policy	4,4	5
Data Procedure	3,8	5
Data Governance	4,2	5
Data Standardization	3,8	5
Data Technology	4,2	5
Data Performance Management	3,8	5

1. Data Expectation

Data is defined as a representation of a fact or numbers associated with an object. Data currently stored digitally can be displayed in text form, numbers, graphs, images, sounds, and videos [4][7]. Data is the initial foundation of the formation of information. Knowledge is formed from various information and integrated experiences and forms a particular pattern. The pattern formed from the integration can easily be translated and has added value for an organization or company. Data can be used in different timeframes, both short-term and can be implicated in medium-term or long-term [4][8][9].

2. Data Dimension

PT XYZ already has data quality rules. The data quality rules refer to internal regulations, one of which is circular letter number 80-DIR/MDP/12/2021 on Standardization of Customer Data Quality. These data quality rules show that PT XYZ focuses on the quality of its customer data on the data dimension indicator. PT XYZ has been able to define the dimensions of customer data quality measurement used in data quality rules. Three dimensions are used: accuracy, completeness, and consistency. There is also a purpose for using each dimension used. PT XYZ has set and monitored the SLA, although its implementation still has limitations.

3. Data Policy

For the data policy indicator, it appears from the test results that PT XYZ has guidelines and rules that can be used to improve data quality. The existing regulations and guidelines are made using best practices that refer to the DMBOK guidelines. There is also a single source of data that can be used and privacy-related rules and documented mechanisms for resolving data quality issues. Regarding solving data problems at PT XYZ, the relevant work unit has received a service level agreement (SLA) that must be fulfilled.

So far, PT XYZ has made all guidelines by referring to the applicable regulations. This can be evidenced by the various reports that must be provided to the regulator at a particular time periodically. As a financial services business that has a stringent regulator, reports containing customer data must be provided in strict compliance with existing regulatory rules.

There are two criteria in the data policy indicator, which have now begun to be partially implemented at PT XYZ. For example, notification of inappropriate data can already be done but still manually.

4. Data Procedure

Handling data quality issues at PT XYZ has been coordinated according to existing criteria. Data quality management has been used by all organizations involved in addressing data issues. Resolving data quality issues is done from the main level to the operational work unit level. Defining the root cause of data problems at PT XYZ relates to the customer onboarding process.

5. Data Governance

The condition of PT XYZ is currently very concerned about data quality, which is one of the drivers of the high value of the data governance indicator in the test results. In terms of governance, PT XYZ has designed, created, and documented the organizational structure,

working principles, and operational standards and procedures related to data quality. It is used by all work units in PT XYZ. PT XYZ also has a governance committee.

The Company's Data Management Division has been working to make the current data quality management (DQM) practices more transparent. The current challenge is how implementation or practice can be adopted by all work units, both business and operational.

PT XYZ has tried to realize one of the criteria related to respect for its employees. Currently, PT XYZ has organized a data quality award to be able to provide a form of appreciation to work units that make the best efforts to improve the quality of customer data. In practice, the data quality award only covers the best operational work units and does not directly reward workers who make efforts to improve data quality.

6. Data Standardization

The data standards have been outlined in internal provisions in circular number 80-DIR/MDP/12/2021 on Standardization of Customer Data Quality. The provision regarding the standard customer data format at PT XYZ has explicitly been mentioned. PT XYZ has seriously adopted the concept of master data. This is evident in how PT XYZ created a separate organization responsible for master data management.

7. Data Technology

Regarding technology, PT XYZ is a financial services company that is very concerned about this. Digital transformation is one of the drivers of PT XYZ to continue to improve and utilize technology for various aspects of its business and operations. In an effort to enhance the quality of customer data at PT XYZ, the use of technology is massive. Data technology indicators have been widely applied at PT XYZ. The application of the technology has been carried out with a validation process and is starting to move towards automation. The data technology used has referred to existing business processes and, of course, is equipped with straightforward procedures. Some of PT XYZ's workers have also used the adopted technology. It is expected that with the technology, users other than the technical team can also define data quality rules according to business needs.

8. Data Performance Management

PT XYZ is currently able to identify impacts related to data quality, and some workers of PT XYZ have known it. PT XYZ, a financial service provider in general, has a customer profiling process used for many things. One of the uses of profiling data is for marketing activities. A company as big as PT XYZ needs good-quality data profiling to dominate the market. To be able to do many things with the data it has, PT XYZ adopts the most ideal and widely used framework in many industries, namely those found in DAMA-DMBOK.

To be able to provide recommendations that are in accordance with existing best practices, a mapping of the existing criteria gap at PT XYZ will be carried out with DAMA-DMBOK's data quality management activity. Based on the mapping results, ten data quality management activities can be used in the category to improve data quality at PT XYZ. Four activities can be used in more than two categories. The fourth activity is to prioritize actions based on business impact and develop preventive and corrective actions, data quality is adequately measured and monitored, and findings and the level of data quality can be reported. Details of the number of quality management data activities used in criteria that include data quality gaps at PT XYZ are shown in Table 4 .

Table 4. Summary of DQM Activities that Can Use

Data Quality Management Activity		Count
Code	**Description**	
DP1	Defining High-Quality Data	1
DP2	Defining the Data Quality Strategy	0
DP3	Defining Critical Data for Simple Measurements	0
DP4	Defining Existing Rules and Patterns in Initial Assessment	2
DP5	Prioritizing Issues in Data Quality Measurement	1
DP6	Developing Root Cause Analysis on Data Quality Measurement	1
DP7	Prioritizing Actions Based on Business Impact	3
DP8	Develop Preventive and Corrective Measures	5
DP9	Confirming the Action Plan	1
DP10	Develop Data Quality Operational Procedures	1
DP11	Performing Error Correction on Data Quality	0
DP12	Data Quality is Properly Measured and Monitored	6
DP13	Reportable Findings and Data Quality Level	4

In outline, the strategy recommendations PT XYZ can do to improve the quality of its customer data are as follows: Increased awareness and understanding of workers related to data management based on risk-based thinking. This can be supported by well-socialized governance and rules. Next is to make observations related to using and implementing existing data quality rules in all work units at PT XYZ. This is done to ascertain whether all work units have implemented the existing laws.

PT XYZ must also measure data quality management with systematic stages, such as determining goals to stages of improvement equipped with a structured timeline to improve its data quality. Each step or implementation of improvement recommendations must be organized and executed based on priorities developed with data stakeholders. Stakeholders do monitoring of the implementation. So far, data quality management monitoring has been done mainly by the 1st line. In the future, PT XYZ must involve the 2nd line, which is Risk Management, to monitor quality management data and efforts to improve it.

Then the report on measuring customer data is made more specific according to the criteria set out in DMBOK - DAMA International. Make a checklist of all report criteria so that the reports displayed are more effective and acceptable to stakeholders, especially senior management at PT XYZ. One possible form of report presentation is key risk indicators (KRI). Finally, PT XYZ can organize senior management meetings or utilize various forums in which senior management is present. Reports on the results of data quality measurements carried out must be displayed in senior management meetings to continue getting relevant stakeholders' attention and support. One forum that can be utilized is the risk management committee (RMC). RMC itself has routinely conducted once a month, and almost all discussions in the RMC are of concern to senior management, even up to the C level. The Unfulfilled Criteria and DQM Activities that PT XYZ can use for detail is shown in Table 5.

Table 5. Unfulfilled Criteria and DQM Activities that PT XYZ Can Use

No.	Unfulfilled Criteria	DQM Activity Code
Data Expectation		
1.	PT XYZ can formulate prioritization of fulfilling data quality expectations as a follow-up to the analysis of the impact of data problems on business	DP7
2.	Data quality measurement guidelines are clearly defined at PT XYZ.	DP4
3.	Compliance with data quality rules can be linked to employee performance targets such as the utilization of PT XYZ employee KPIs that affect data quality.	DP7
4.	Objective goals in data quality improvement have been set in accordance with or more than the capabilities of PT XYZ.	DP7
5.	Control over data validation can be integrated with existing business processes at PT XYZ.	DP6
Data Dimension		
6.	Data quality dimensions can be mapped with the impact on each business cluster (micro, retail, medium & wholesales) in PT XYZ.	DP4
7.	Service Level Agreement (SLA) for data quality including the repair of poor-quality data at PT XYZ has been met.	DP13
8.	The data quality SLA that has been set at PT XYZ can be monitored properly.	DP12
9.	Data quality can be part of the system development life cycle at PT XYZ.	DP8
Data Policy		
10.	There is an automatic notification if there is an anomaly in PT XYZ customer data.	DP12
11.	PT XYZ workers can automatically comply with the data quality policy set by the company (self-governing)	DP8
Data Procedure		
12.	Data control has been designed to be able to be merged into different business applications at PT XYZ.	DP7
13.	Data weaknesses are recognized early in the information flow e.g., in the development phase of SDLC or DDLC	DP5
14.	Data control has been implemented throughout the organization or work unit at PT XYZ.	DP7
15.	Each related work unit in PT XYZ issues data quality measurement results	DP12
16.	DQM practices in PT XYZ have been transparent	DP13
Data Governance		
17.	The reporting and remediation framework at PT XYZ has collaborated in implementing statistical process controls to maintain the quality of data to stay at the specified limits	DP13
18.	Data quality performance measures at PT XYZ are continuously reviewed so that opportunities for data quality improvement continue to increase.	DP12
19.	Staff or workers at PT XYZ have been rewarded if they achieve data governance performance goals.	DP8
Data Standardization		
20.	The Data Standards Oversight Board at PT XYZ has overseen the maintenance of internal standards and compliance in line with external standards used, e.g., regulatory standards.	DP12
21.	The concept of master data at PT XYZ has been managed in a master data environment.	DP8
22.	PT XYZ has defined and validated a taxonomy for customer data standards.	DP4
23.	Compliance with standards is integrated in the technical structure-oriented policies at PT XYZ.	DP10
24.	The data standardization process at PT XYZ has been done automatically	DP8
Data Technology		
25.	Impact analysis at PT XYZ has been supported by dashboard and reporting applications.	DP13
26.	Non-IT users such as business users at PT XYZ can dynamically define and modify data quality rules and data dimensions.	DP1
Data Performance Management		
27.	The audit conducted at PT XYZ has been based on compliance with the rules related to the existing data quality dimensions.	DP12
28.	The overall performance of the organization or working unit at PT XYZ can be improved by modifying policies through rules.	DP9

CONCLUSIONS

This research is conducted to measure the level of customer data quality in the core system of PT XYZ. The measurement is done using an approach commonly used academically, Loshin's Data Quality Management Framework. Within the framework, there are eight indicators. Based on the measurement results of customer data quality on the core system of PT XYZ, the value of each indicator is obtained; data expectation (3,8), data dimension (3,4), data policy (4,0), data procedure (3,5), data governance (3,9), data standardization (3,1), data technology (3,8), and data performance management (3,7).

After determining the level of data quality at PT XYZ, a gap measurement is done between the current level of data quality and senior management's expectations of customer data quality. This is done to be able to produce well-documented data quality management strategy recommendations in an effort to improve data quality, especially customer data at PT XYZ.

Based on the existing gaps, ten strategy recommendations were generated from the quality management activity version 2.0 data based on DMBOK - DAMA International, namely defining data with a high level of quality, defining existing rules and patterns in the initial assessment, developing root cause analysis on data quality measurement, prioritizing problems in data quality measurement, prioritizing actions based on business impact, developing preventive and corrective actions, confirm the action plan, develop data quality operational procedures, data quality is measured and appropriately monitored, findings and levels of data quality can be reported.

The results of this study are expected to be used as an initial parameter in measuring the quality of existing customer data at PT XYZ using an academic approach. PT XYZ, which has adopted the DMBOK framework, can immediately use this research because it aligns with what is currently implemented. On the other hand, PT XYZ can also take measurements with different methods to continue getting the best indicators that can be applied to improve customer data quality. Future research can use different approaches to select criteria more suitable for the company or organization of the research object.

In this research, there is no priority sorting of data quality indicators. Prioritizing data quality indicators allows future research to produce more specific recommendations for strategies to improve data quality in companies or organizations. Future research can focus more on recommendations for improving data quality by providing detailed information in the form of implementation time, the person in charge (PIC) of each recommendation given, and the results and outputs obtained from each recommendation given.

REFERENCES

[1] M. Fleckenstein, and L. Fellows, "Modern Data Strategy". Cham, Switzerland: Springer, 2018.

[2] E. Brynjolfsson, L. M. Hitt, and H. H. Kim, "Strength in numbers: How does data-driven decision making affect firm performance?" SSRN, 2011.

[3] G. Morbey, "Data Quality for Decision Makers (2nd ed.)". Wiesbaden: Springer Fachmedien Wiesbaden, Imprint: Springer Gabler, 2013.

[4] N. Laranjeiro, S. N. Soydemir and J. Bernardino, "A Survey on Data Quality: Classifying Poor Data," IEEE 21st Pacific Rim International Symposium on Dependable Computing (PRDC), 2015, pp. 179-188.

[5] D. Loshin, "Master Data Management and Data Quality" BT - The Practitioner's Guide to Data Quality Improvement. In MK Series on Business Intelligence, 2011.

[6] Wijayanti, W., Hidayanto, A. N., Wilantika, N., Adawati, I. R., & S. B Yudhoatmojo,. Data quality assessment on higher education: A case study of institute of statistics. 2018 International Seminar on Research of Information Technology and Intelligent Systems, ISRITI 2018, 231–236.

[7] International, D. (2017). Dama-Dmbok – data management body of knowledge (2nd ed.). Technics Publications Llc.

[8] Cichy, C., & Rass, S. (2019). An overview of data quality frameworks. IEEE Access, 7, 24634–24648.

[9] Ramasamy, A., Sisay, B., & Bahiru, A. (2021). A data science framework for data quality assessment and inconsistency detection. International Journal of Advanced Computer Science and Applications, 12(4).

[10] Setiadi, Y., Hidayanto, A. N., Rachmawati, F., & Yohannes, A. Y. (2021). Data Quality Management Maturity Model: A case study in higher education's Human Resource Department. 2021 IEEE 7th International Conference on Computing, Engineering and Design (ICCED).

[11] Sabtiana, R., Yudhoatmojo, S. B., & Hidayanto, A. N. (2018) "Data Quality Management Maturity Model: A Case Study in BPS-Statistics of Kaur Regency, Bengkulu Province, 2017," 6th International Conference on Cyber and IT Service Management (CITSM), pp. 1-4.

[12] Indriany, H. S., Hidayanto, A. N., Wantania, L. J., Santoso, B., Putri, W. U., & Pinuri, W. (2021). Data Quality Management Maturity: Case Study National Narcotics Board. 2021 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT).

[13] Lucas, A. 2010. Corporate data quality management: From theory to practice. 5th Iberian Conference on Information Systems and Technologies, pp. 1-7.

[14] Nikiforova, A. (2020). Definition and evaluation of data quality: User-oriented data object-driven approach to data quality assessment. Baltic Journal of Modern Computing, 8(3), 391–432.

[15] Mosley, M., Earley, S., & Brackett, M. (2009). The Dama Guide to the Data Management Body of Knowledge (1st ed.). Technics Pubns Llc.

[16] Loshin, D. (2011). The practitioner's guide to data quality improvement. Elsevier.

[17] Bowo, W. A., Suhanto, A., Naisuty, M., Ma'mun, S., Hidayanto, A. N., & Habsari, I. C. (2019). Data Quality Assessment: A case study of PT jas using TDQM framework. 2019 Fourth International Conference on Informatics and Computing (ICIC).

[18] Wang, R. Y. (1998). A product perspective on total data quality management. Communications of the ACM, 41(2), 58–65.

[19] Wibisono, S. B., Hidayanto, A. N., & Nugroho, W. S. (2018). Data Quality Management Maturity Measurement of government-owned property transaction in BMKG. CommIT (Communication and Information Technology) Journal, 12(2), 59–72.

[20] Rosen, A. M. (2019). Effective Research Methods for Any Project. Chantilly: The Teaching Company.

[21] Sekaran, U., & Bougie, R. (2016). Research Methods for Business: A Skill Building Approach 7th edition. Chichester: John Wiley & Sons Ltd.

[22] Li, S., & Wang, H. (2018). Traditional Literature Review and Research Synthesis. The Palgrave Handbook of Applied Linguistics Research Methodology, 123–144.

[23] Tri Putri, D., & Ruldeviyani, Y. (n.d.). Data Quality Assesment For Suspicious Report Using Loshin's Framework And Dmbok : A Case Study Of Financial Transaction Analysis Institution. Jurnal Penelitian Teknologi Informasi Dan Komunikasi, 1, 17–26.