Data and Information Quality Framework Development:
Proposed for Indonesia Higher Education

Halimah HUSAIN, Department of Chemistry, Universitas Negeri Makassar, Makassar, Indonesia, email: halimah.husain@unm.ac.id

Faisal SYAFAR, Computing and Information Systems Research Centre, Department of Electronics, Universitas Negeri Makassar, Makassar, Indonesia, email: faisal.syafar@unm.ac.id

Edy SABARA, Computing and Information Systems Research Centre, Department of Electronics, Universitas Negeri Makassar, Makassar, Indonesia, email: edysabara66@unm.ac.id

Sofyan SYAFAR, STIA-LAN Makassar, Department of Public Administration, Makassar, Indonesia, email: sf_syafar@yahoo.com

Abstract

The main objective of this research is to find solutions to improve the quality of data and information on higher education in Indonesia. The target is the availability of a structured and comprehensive Framework and strategy accompanied by manual book adoption and implementation, which is expected to encourage efforts to obtain quality data and information (accurate, complete, timely, and consistent) utilizing information technology in supporting governance good governance. In the long run, it is expected that all higher education institutions in Indonesia have an information system that is organized with valid and reliable data quality. This quality data and information is used for policy formulation, process clarity procedures, data cleansing approaches, adequate planning, and appropriate decision making, and management of higher education institution management at all levels and levels of management (institutions, faculties, departments and study programs) towards the realization of the national education industry in Indonesia. To achieve this goal, this study uses a mixed method approach, namely quantitative methods (National and national studies and surveys) and qualitative (case studies). Delphi studies were conducted over three rounds involving experts in data / information quality and higher education. National surveys are carried out using online questionnaires, while case studies are conducted with semi-structured interview techniques. Framework development is carried out by triangulation between Delpi findings data, studies and case studies. The research population is State and Private Universities (Universities, Institutions, Colleges, Academies) which are spread across all major islands in Indonesia (Sumatra, Java, Bali, Kalimantan, Sulawesi, Maluku and Papua).

Introduction

Data and information have become an increasingly important resource in supporting educational activities, therefore higher education seeks to always produce and collect more supporting data. However, it does not mean that the more data, the better the information becomes, or better decision-making is guaranteed. According to Gao and Koronios (2014), it is estimated that more than 70% of data produced by many institutions has never been used. Even Levitin and Redman (1998) concluded that most institutions have a lot of data, but they cannot use it well; while contrarily, the really needed data is actually not available.

The need for accurate data/information is increasing but it turns out that the education or academic information system is still unable to produce accurate, complete and timely data and information. Various problems are still faced in the implementation of academic information system, including the absence of the same perception among providers of higher education. The implementation of academic information systems has not been carried out efficiently, there has been "redundant" data, and duplication of activities, besides the quality of data collected is still low, there are even data that

are not in accordance with needs, timeliness of reports is also low, feedback system is not optimal, utilization data and information at the level of faculty/ department/study program for advocacy, program planning, monitoring and management are still low, resulting in inefficient use of resources.

The improved education Data and Information Quality (DIQ) Framework for Higher Education (PT) developed in this study consists of a comprehensive structure to address the DIQ problem with a strategic, organized approach, and proposes solutions in creating quality information-oriented data. In order to develop the framework, there are four research questions that need to be answered.

Research Questions

1. What is the current condition of the quality of data and information of higher education?
2. What are the current quality and data quality problems in higher education?
3. What factors influence the quality of data and information?
4. How can the quality of data and information in higher education be improved through an integrated framework development?

Research purposes

1. To identify the conditions and quality problems of data and information related to the management of higher education, along with the factors that influence them.
2. To develop an integrated framework on improving the quality of higher education data and information, which includes approaches and strategies:
   - Improving the quality of higher education data management including: collection, storage, processing, analysis and dissemination of data and information.
   - Improving the DIQ of education based on quality assurance and the control system.
   - Strengthening higher education information system resources which include the use of ICT, human resources (HR), financing, and facilities and infrastructure (physical assets)

Urgency of Research

There is very limited or even there has not been serious research and adequate literature on the quality of higher education data and information in Indonesia. It is, therefore, we considered very urgent and important to carry out this research to understand the DIQ issues of PT. This research was carried out using a holistic approach covering aspects of technology, organization, and personal (TOP).

Theoretical Foundation

Data quality

In the beginning, data quality (DQ) was only interpreted from the perspective or dimension of accuracy, but then many studies identified that DQ consisted of several dimensions besides accuracy. The four dimensions of DQ that must be possessed by a data set most often mentioned in various literatures are accuracy (accurate), completeness (complete), timeliness (on time) and consistency (consistent) (Liu & Chi, 2002; Naumann, 2002; Bouzeghoub & Peralta, 2004 and Batini et al., 2004). A data set might be able to fulfill three of the four dimensions but one other dimension is not fulfilled. To fulfill or improve one dimension can interfere with other dimensions. For example, it is possible to improve or meet the dimensions of timeliness, but perhaps the accuracy dimension is sacrificed (Neely & Pardo, 2002) or both can be achieved with additional costs (Ballou & Pazer, 1995). Meanwhile, every industry sometimes has different standard requirements or DQ dimension priorities (Giannoccaro et al., 1999).
Data whose quality is suitable for one particular institution (Syafar et al., 2014a; Syafar et al., 2014b), is not necessarily considered sufficient for other institutions or organizations (Ballou & Pazer, 1995).

Furthermore, the dimensions of DQ are considered appropriate for a particular decision making, not necessarily valid for the other nine decisions. That is why the definition of DQ which has been proposed by Wang and Strong (1996) which states that "quality data is data that is suitable for use by data users" was adopted in this study.

Maintaining data quality is often a problem, because only with quality data can be a reference for effective decision making. Examples of many factors that can hinder the quality of data in an institution, company or industry in Indonesia according to Buana, IBGM and Wirawati, NGP. (2018), are: inadequate management structures to ensure data completeness, timeliness and accuracy of data; inadequate rules, training, and procedural guidelines for those involved in data collection; and inconsistencies between services related to data collection.

Higher Education Data and Information Management

As an organization that has a strategic role in educating the nation, every higher education institution in Indonesia should have a Long Term Development Plan that describes the direction of the development of higher education institutions. Based on a report from the Secretary of the Directorate General of Higher Education, almost all State Universities in Indonesia already have a Long Term Development Plan, but in reality, the implementers of education institution development have never referred to the Long-term Development Plan that they have prepared. This shows that the Higher Education Institution's Long Term Development Plan is compiled not based on data and information in higher education institutions.

The need for information systems above can only be fulfilled if higher education institutions have used quality data as raw material for the preparation and processing of quality, effective and efficient information. Quality information is needed to meet the reporting needs and executive information that can be trusted as a reference for decision making by the leadership elements (such as Chancellor / Dean / Chair of Study Program / Head of School, etc).

There are various types of technical specifications, operational and administrative systems in managing higher education assets, which not only manage the operation of equipment, but also provide maintenance support throughout the asset life cycle, especially physical assets (Syafar et al., 2017; Syafar et al., 2015) . The use of information technology and the process of automation in asset management means improving the quality of asset management data and information to achieve quality output (Syafar et al., 2015).

TOP Approach

Mitroff & Linstone (1993) argues that every phenomenon, subsystem or system needs to be analyzed from what they call the multiple perspectives method (a combination of several perspectives) using different ways to see, find solutions to a problem. This method or method is shown in the TOP model (technology, organization, personal) from Linstone (1999) and Mitroff & Linstone (1993). The TOP model allows analysts to see the context of the problem both from the Technical (technological) or Organizational (company / industry) or Personal (individual) side.

Mitroff & Linstone (1993) shows that these three perspectives can be applied as "looking in three ways" to a problem that arises for, or in a particular phenomenon or system. Werhane (2002) further notes that using the TOP perspective is very important, because we can take into account the fact that each individual or group, organization / institution, or system, creates a frame through a series of models that are independent or incomplete. In other words, looking at a problem from a single perspective or approach is often imperfect compared to considering it with more than one approach.

Research Design

The research design consisted of four stages: literature review, Delphi study, national survey, and case study.

Stage of Research

The first stage, is a literature review (focused) that aims to explore the quality of data in general (generic) which will be used as the basis for developing the basic concept of the DIQ Framework. This conceptual DIQ framework will guide researchers at the next research stage. Identification and categories of DIQ problems and the factors that influence them are based on the TOP approach (Technology / technical, Organization / Institution, Personal / individual).

The second stage, Delphi study. The Delphi technique officially originated from the RAND Corporation in the early 1950s as a method for exploring and predicting future technology and potential political and policy issues surrounding military activity (Gordon 2003; Vernon 2009). The Delphi study has been using in area of mobile technology studies (Syafar et al., 2013). The goal is to:
- create research forums with DIQ experts and experts in higher education systems;
- get input on DIQ specific issues;
- uncover additional DIQ issues related to PT other than those which obtained from the literature;
- improving TOP factors and categories that have been obtained in stage I.
- The Delphi technique was chosen to produce consensus by asking opinions from expert groups relevant to the topic of the DIQ PT (1) DIQ Experts and (2) PT system experts (4 people each in the country and 4 foreigners = 16 respondents)
- Conducted in 3 (three rounds).

Delphi study data analysis

Responses from expert groups are summarized at the end of the rounds, then disseminated again in the next round to get a response. Through the process of convergence with the content analysis approach: identification of general trends, similarities / similarities. Finally a consensus can be reached at the end of the third round.

Third stage, National survey

The aim of the national survey is to:
answer research questions especially the first, second and third questions, namely:
- disclose the current state of the DIQ PT or in Indonesia;
- understand general perceptions of the quality of data and information related to PT resource management;
- provide an overview of the factors that have an impact on the DIQ, which later it will be verified by the findings obtained in the case study in the next research phase.

At this stage at least 1000 respondents will be sent a questionnaire (online) link and / or via e-mail attachment (email) with the target role / position of the respondent being the PT leadership element (Chancellor, Assistant / Deputy Chancellor; Dean, Assistant / Deputy Dean, Program Chairperson Study, Director / Head of ICT Center, Director / Head of Quality Assurance center), Data Collector, Operator, and Technician. The population of this survey are State and Private Higher Education Institutions (Universities, Institutions, Colleges, Academies) which are spread across all major islands in Indonesia (Sumatra, Java, Bali, Kalimantan, Sulawesi, Maluku and Papua).

National survey data analysis

The SPSS program will be used for survey data analysis including ANOVA (individual variables of each category of PT); MANOVA (a combination of the 6 basic DIQ problems: accuracy, completeness, redundancy, integrity, consistency, timeliness) followed by Multiple Regression; and t-test (to see if there are significant differences of opinion from respondents).
Fourth stage, Case Study. The purpose of the case study is to:

- get input on how important each of the DIQ factors is;
- verify the findings of the DIQ survey conducted in stages II and III;
- determine the rank order of the DIQ factors;
- identifying key (main) factors of DIQ that need to be a priority;
- test the relationship between the DIQ factors (if found there).

Semi-structured interview case studies are designed for ten (10) different PTs. Interviews will be conducted with PT leadership elements, IT personnel, data administrators, asset managers, operator technicians, field staff, maintenance technicians and (external) contractors / suppliers.

Data analysis of the case study

Within-case analysis (DIQ analysis between all target respondent groups in the same college) and Between-case analysis (DIQ analysis with comparison between the same group of respondents in different types of colleges). Analysis Ranking order data using SPSS is used to determine the sequence of DIQ problem factors for each group T, O and P, from the most important (main or key) to least important.

Development of DQ Framework

The results of the study in the case study were triangulated with the results of the Delphi study and survey to obtain a complete set of factors influencing the DIQ. According to Guion (2012) the phenomenon examined by using multi methods can be well understood so that high levels of truth can be obtained if approached from various points of view. Therefore triangulation is an effort made to ensure the correctness of data or information obtained by researchers from a variety of different perspectives by reducing as much as possible the differences that occur during data collection and analysis. The ranking order of all factors is sorted by importance. The factors and sub-sets of factors received with the majority of respondents' support can be identified as key (main) DIQ factors. According to Syafar and Husain (2017); Syafar and Gao (2013), the purpose of triangulation process is to determine the empirical validity of developing the DIQ Framework.

Conclusion

This research will be carried out in 3 stages and what will be done for 3 years. This research expected provides a comprehensive understanding of the philosophy, framework and methodology of adequate data and information governance so that higher education can have valid and reliable data and information through the formulation of data management policies and procedures, clarity of ownership, and adequate data cleansing and excellent planning approaches.

References

Buana, IBGM and Wirawati, NGP. (2018), 'Pengaruh Kualitas Sistem Informasi, Kualitas Informasi, dan Perceived Usefulness Pada Kepuasan Pengguna Sistem Informasi Akuntansi', *E-Jurnal Akuntansi, Universitas Udayana*, 22 (1), 683-713.

Ballou, DP. and Pazer, HL. (1995), 'Designing Information Systems to Optimize the Accuracy-timeliness Tradeoff', *Information Systems Research,* 6 (1), 51- 72.

Batini, C, Catarci, T and Scannapiceco, M. (2004), 'A survey of Data Quality Issues in Cooperative Information Systems', paper presented at the the 23rd International Conference on Conceptual Modeling, (ER2004), Shanghai, China.

Bouzeghoub, M and Peralta, V. (2004), 'A Framework for Analysis of Data Freshness', paper presented at the First ACM International Workshop on Information Quality in Information Systems (IQIS 2004), Paris, France.

Gao, J. and Koronios, A. (2014), 'Data Quality Issues and Challenges in RFID Adoption', paper presented at the 4th International Conference on Supply Chain Management and Information Systems (SCMIS 2006) Taichung, Taiwan, July 5-7, 2006.

Giannoccaro, A, Shanks, G. and Darke, P. (1999), 'Stakeholder Perceptions of Data Quality in a Data Warehouse Environment', *Australian Computer Journal,* 31 (4), 110-117.

Linstone, HA. (1999), Decision Making for Technology Executives: Using Multiple Perspectives to Improve Performance, Artech House Publisher.

Liu, L and Chi, L. (2002), 'Evolutionary Data Quality. A Theory-Specific View', paper presented at the 7th Intl. Conf. on Information Quality (IQ'02). Massachusetts Institute of Technology, Cambridge, MA, USA.

Mitroff, I and Linstone, H. (1993), The Unbounded Mind: Breaking the Chains of Traditional Business Thinking, OxFord University Press, New York.

Naumann, F, Freytag, JC and Leser, U. (2004), 'Completeness of Integrated Information Sources', *Information Systems,* 29 (7), 583-615.

Neely, MP. (2002), 'A framework and associated software tool for the analysis of source data for a data warehouse: Development and exploratory study', Information Science, State University of New York at Albany.

Syafar, F. and Husain, H. (2017), 'Development of an integrated framework for successful adoption and implementation of mobile collaboration technology in indonesian healthcare. Proceedings of the 30th IBIMA, 2017, paper 11,108-114.

Syafar, F., Husain, H., Harun, S., Ridwansyah, and Sokku, S. (2017). Key data and information quality requirements for asset management in higher education: a case study. Proceedings of the 30th IBIMA, 2017, paper 55, 1670-1677.

Syafar, F, Gao, J. and Du, T. (2015),' Mobile Collaboration Technology to Support Maintenance Enterprise Systems in Large Industry', The 21st American Conference on Information Systems.

Syafar, F., Gao, J. and Du, T. (2014), 'Refereed proceedings of the 18th Pacific Asia Conference on Information Systems (PACIS): Mobile-Enabled Collaborative Maintenance. Chengdu, China.

Syafar, F., Gao, J. and Du, JT (2014a), 'Current Mobile Technology Roles to support Collaborative. Proceedings of the 24th International Business Information Management Association Conference - Crafting Global Competitive Economies: 2020 Vision Strategic Planning and Smart Implementation.

Syafar, F., Gao, J. and Du, T. (2013), Refereed proceedings of the 17th Pacific Asia Conference on Information Systems (PACIS): Applying The International Delphi Technique In A Study Of Mobile Collaborative Maintenance Requirements. Jeju Island, Korea.

Syafar, F. and Gao, J (2013). Development of a framework for adoption and implementation of mobile collaboration technology in engineering asset management. Proceedings of the 20th International Business Information Management Association Conference.

Wang, RY. and Strong, DM. (1996), 'Beyond Accuracy: What Data Quality Means to Data Consumers', *Journal of Management Information Systems (JMIS),* 12 (4), pp. 5-34.

Werhane, PH. (2002), 'Moral Imagination and Systems Thinking', *Journal of Business Ethics,* 38, 33-42