ENHANCING QUALITY ASSURANCE MECHANISM THROUGH DATA ANALYSIS APPROACH

Byron Joseph A. Hallar
Doctor in Information Technology

ABSTRACT

Quality assurance is a holistic approach utilized in higher education institutions which covers all the elements of educational processes. It is aimed at providing better services to students and other stakeholders ensuring that these services meet certain quality standards. This study sought to develop an enhanced quality assurance mechanism intended to expedite the process of a state university's leveling by developing an information system that incorporated the useful functions of data consolidation and clustering and supported by data analytics technologies. The study utilized the descriptive mixed methods research design and included fifteen (15) participants who are involved in the operation of the leveling instrument tool for the State Universities and Colleges (SUCs). Relevant data were gathered using questionnaires and through observations and interviews. The developed SUCs Leveling tool has a high level of validity as assessed by the information technology experts. The overall results of the participants' evaluation, the enhanced SUC Leveling Tool for SUCs were found to be compliant to a very great extent with the quality standard characteristics defined by ISO/IEC 9126-1 in terms of functionality, reliability, usability, efficiency, maintainability and portability. The developed system has the ability to extract knowledge hidden in complex and varied sources of data and visualize them in human-readable form. The data mining and analytics tool is freely available for academic use and uses the data sets generated by the developed system as its input.

INTRODUCTION

State Universities and Colleges in the Philippines are ranked by the use of a "leveling" instrument jointly prescribed by the Philippine Government's Commission on Higher Education and the Department of Budget and Management. This is part of the country's Quality Assurance process mandated for these educational institutions. Performed annually, the "leveling" process determines the rank of the university or college by means of a five-point typology, with Level I as the lowest and Level V as the highest. The intent of the process is to reward institutional performance and to provide incentives for improvement. The rank or "level" of the institution determines the level of state support and the level of bonuses awarded to individual employees. It is also a matter of prestige among the schools to attain the highest possible rank.

The process of determining the ranking level of an educational institution requires the analysis of voluminous amounts of data. Each educational institution is required to assemble and collate the data and submit it to the national government for final processing. In actual practice, gathering the required data is labor-intensive and time consuming. There is no existing information system in use for rapid data entry, collation and analysis. This study titled "Enhancing Quality Assurance Mechanism through Data Analysis Approach" was intended to expedite the process of SUC Leveling. Supported by data analytics technologies, the developed information system incorporated useful functions such as data consolidation and clustering. Piloted at the Mariano Marcos State University in Northern Philippines, the researcher intended to eventually deploy the developed information system at a regional and nationwide basis.

Statement of the Problem

This study aimed to develop an enhanced quality assurance mechanism that utilizes the data analytics approach.

Specifically, it attempted to answer the following questions:

1. What is the profile of the participants in terms of:
    1.1 Unit of Work;
    1.2 Position;
    1.3 Academic Rank;
    1.4 Highest Educational Attainment; and
    1.5 Field of Specialization?
2. What is the performance level of the existing system as assessed by the participants in terms of:
    2.1 Functionality;
    2.2 Reliability;
    2.3 Usability;
    2.4 Efficiency;
    2.5 Maintainability; and
    2.6 Portability?
3. Is there a significant difference in the assessment of the participants when grouped according to the profile variables with respect to:
    3.1 Functionality;
    3.2 Reliability;
    3.3 Usability;
    3.4 Efficiency;
    3.5 Maintainability; and
    3.6 Portability?
4. To what extent does the developed enhanced leveling tool for SUCs comply with the quality standard characteristics of software as defined by ISO/IEC 9126-1 in terms of:
    4.1 Functionality;
    4.2 Reliability;
    4.3 Usability;
    4.4 Efficiency;
    4.5 Maintainability; and
    4.6 Portability?
5. What is the degree of validity of the developed enhanced leveling tool system as evaluated by the information technology experts?

6. Is there a significant difference between the features of the existing system and the developed system as assessed by the participants?

7. What is the source data generated by the developed enhanced leveling tool system that is needed for data modeling and analysis?

METHODOLOGY

Research Design

The Descriptive Qualitative-Method was used for this study. The method was used as it describes the nature of the situation as it exists at the time of the study. It involved gathering data that describe events and then organizes, tabulates, depicts, and describes the data. It also used description as a tool to organize data into patterns that emerge during analysis (Myers, 2012).

Moreover, it is concerned with conditions of relations of relationships that exist; practices that prevail; beliefs, processes that are going on; effects that are being felt, or trends that are developing. The process of descriptive research goes beyond mere gathering and tabulation of data. It involves an element or interpretation of meaning or significance of what is described. Thus, description is often combined with comparison and contrast involving measurements, classifications, interpretation and evaluation. This method was used since it can greatly assess, determine and pin-point the impact of the proposed system to its future users. Moreover, this method allows a flexible approach, thus, when important new issues and questions arise within the duration of the study, further investigation may be conducted.

Descriptive research, on the other hand, is a type of research that is mainly concerned with describing the nature or condition and the degree in detail of the present situation. This method is used to describe the nature of a situation, as it exists at the time of the study and to explore the causes of a particular phenomenon. The aim of descriptive research is to obtain an accurate profile of the people, events or situations. With this research type, it is essential that the researcher already has a clear view or picture of the phenomenon being investigated before the data collection procedure is carried out. The researcher used this method to obtain first hand data from the participants so as to formulate rational and sound conclusions and recommendations for the study. The descriptive approach is quick and practical in terms of the financial aspect.

In this study, the descriptive research method was implemented

specifically the summative evaluation type to address issues concerning the usage, implementation, and to track problems that may arise during the deployment of the technology. The Research and Development (R&D) Method by Penuel, et al. (2011) and Burkhardt, et al. (2003) was also employed as a guide for the activities conducted in the research particularly with the development of the software. With the end product being a software system for actual production use, the sequential nature of R&D is particularly suitable for the application of basic research, applied research and project development.

Software Development Methodology

The researcher selected the relatively new, widely accepted Scrum methodology / management framework for the software development portion of this study. The name "Scrum" is said to have originated from the "Scrum Formation" of the Rugby sport. Scrum is an Agile type of methodology that aims to develop a software product incrementally or iteratively by building a series of prototypes and constantly adjusting them to user requirements (Sommerville & Sawyer, 1997).

Participants of the Study

For the needs assessment, the researcher included participants involved in the operation of the leveling instrument tool for the SUCs to get the necessary information needed for the study.

The personnel included in the process of getting the baseline information as shown in Table 1 are the participants of the study.

Table 1
Study Participants

Participants	Frequency
CHED IT Coordinator	1
IT Experts	4
Administrative Staff	2
Support Staff	8
Total	**15**

Instrumentation

The instruments were based on PIECES Framework by Wetherbe (1984). Relevant data were gathered using questionnaires and through observation checklist and interview guide. These instruments were utilized to collect sufficient information needed for the study.

Observation. This is an approach of analyzing computer system requirements. This extensively helped the researcher in terms of step-by-step proceeding in developing the system. The researcher used this technique to supplement the facts that were adopted in this study. The researcher personally observed and experienced how the job would be handled, performed and accomplished in the workplace by the participants.

Interview. The researcher used this technique in gathering significant facts and information that were utilized in the study. Casual interviews were conducted with the ones who have direct involvement in the proposed system in order to understand the current system and determine their opinions towards the proposed system.

Questionnaire. This instrument was used to gather additional data and information needed for the study. The survey questionnaire was constructed based on ISO 9126-1, a model to standardize the quality factors of software. This tool was the primary source of data and information. The researcher prepared a questionnaire that was validated by a panel of experts and the adviser and distributed to the participants.

Document Analysis/ Records Review. Office documents were sources of the records containing performance rating data for offices involved in the SUC leveling process.

Data Gathering Procedure

The researcher sought permission from the presidents of involved SUCs, and the CHED Regional Director to gather data from their concerned agencies or offices. The concerned offices include the university registrar, university accountant, directors of the different

offices involved in the leveling tool (SSD, Alumni Affairs, HRMO, QAO, Research and Extension and the CMR). Face to face interviews with the focal persons were conducted and a survey was conducted via questionnaire to determine the needs and to evaluate the solution. Relevant documents and records from the offices were analyzed to gain additional insights needed in formulating a solution.

Data Analysis

After gathering data, the statistical tools used in the analysis of data were the Weighted Mean, Standard Deviation and, Analysis of Variance (ANOVA). The weighted mean was used to determine the average responses on each item of the five (5) options in the questionnaire namely, 5 (Very Great Extent /Accepted Unconditionally), 4 (Great Extent /Accepted with minor condition), 3 (Moderate Extent), 2 (Low Extent/Accepted with major condition) and 1 (Very Low Extent/ Reject).

The points that were used are:

5 points = Very Great Extent/Accepted Unconditionally
4 points = Great Extent/Accepted with minor condition
3 points = Moderate Extent
2 points = Low Extent/Accepted with major condition
1 point = Very Low Extent/Reject

Table 2
Range of Weighted Mean and Its Interpretation

Mean Range	Qualitative Interpretation
4.20 – 5.00	Very Great Extent (as a computer user) The Proposed System is Accepted unconditionally
3.40 – 4.19	Great Extent (as a computer user) The Proposed System is Accepted with minor condition
2.60 – 3.39	Moderate Extent
1.80 – 2.59	Low Extent (as a computer user) The Proposed System is Accepted with major condition
1.00 – 1.79	Very Low Extent (as a computer user) The Proposed System is Rejected

Table 2 shows the mean range and the qualitative interpretation. This was used as a guide to determine the acceptability of the newly developed system from the different users involved in the SUC leveling tool. Standard deviation was used to determine the degree of dispersion or disagreement among the participants. A low standard deviation indicates that the data points tend to be close to the mean (also called the expected value) of the set, while a high standard deviation indicates that the data points are spread out over a wider range of values.

T-test or Analysis of Variance. It was used to determine if there is a significant difference in the assessment of the participants when grouped according to the profile variables with respect to the system's functionality, reliability, usability, efficiency, maintainability and portability.

Also to determine if there is a significant difference between the features of the existing system and the proposed system as assessed by the participants, the level of significance was set at 0.05.

RESULTS AND DISCUSSION

Profile of the Participants

A substantial number of the participants or forty percent (40%) hold the rank of director who works in administrative offices. Twenty percent (20%) of the participants hold ranks of Professor I and Associate Professor V, respectively, while those who are Assistant Professor II, Associate Professors III and IV account for 13.33% each, respectively. The smallest number of participants hold ranks of Supervising Administrative Officer, Administrative Officer V, Associate Professor I, and Professor III account for 6.70% each, respectively. A sizable number of the participants at 46.70% are information technology specialists who hold master's degrees or have earned doctorate units in relation to the discipline. This is followed by specialists in the field of public administration account for 13.33% of the participants. All other specializations account for 6.70% each of the participants.

The findings for the profile of participants can be correlated to

the relevant sections of the MMSU Operations Manual (2nd Edition 2016). Section 2.0 "Organization and Governance" and Sub-section 2.5 "Officials and Administration" of the said manual enumerate and define the position titles as well as the associated duties and responsibilities that are consistent with the findings for the profile of the participants.

Performance Level of the Existing System as Assessed by the Participants

All of the participants (100%) agreed that the performance level of the existing system is compliant only to a moderate extent with the particular attributes of functionality, reliability, usability, efficiency, maintainability and portability.

Test for Significant Difference in the Assessment of the Participants When Grouped According to the Profile Variables

There is no significant difference in the features of the existing system as assessed by the participants.

The Extent of Compliance by the Enhanced SUC Leveling Tool with the Quality Standard Characteristics of Software as Defined by ISO/ IEC 9126-1

Based on the overall results of the participants' evaluation, the enhanced SUCs Leveling Tool was found to be compliant to a very great extent with the quality standard characteristics defined by ISO/ IEC 9126-1 in terms of functionality, reliability, usability, efficiency, maintainability and portability.

Degree of Validity of the Developed Enhanced Leveling Tool System as Evaluated by the Information Technology Experts

The developed SUC leveling system was found to have a high degree of validity based on the results of the validation by information technology experts.

Test for Significant Difference Between the Features of the Existing and the Developed System as Assessed by the Participants

Based on the obtained t-test value, the probability is less than 0.001 which is lower than the level of significance at 0.05. Thus, there is a significant difference between the features of the existing and developed system as assessed by the participants.

Data Analytics Method

The Data Analytics approach utilized the methods of Clustering, Decision Trees, and Cross Validation to discover knowledge from the data collated by the developed system.

CONCLUSION

Based on the findings, the following conclusions are drawn:

Most of the participants possess advanced educational qualifications and hold mid-level to high ranking positions in the University. Thus, the work assignments involved in the SUC Leveling process require highly qualified personnel. This is supported by the mandates set by the MMSU Operations Manual.

The participants considered the performance level of the existing system to be only moderately functional, reliable, usable, maintainable and portable. As a manual, paper-based scheme, the existing system has inherent inefficiencies and flaws that do not exist in modern information systems.

The participants' assessment on the different features of the existing system is the same. The individual capabilities of the users of the existing system are similar to each other. They share common experiences in using the existing system, thus, their similar assessments.

The high extent of compliance of the developed SUC Leveling Tool with the quality standard characteristics of software as defined by ISO/IEC 9126-1 shows that it is ready for production use. This high level of compliance with this internationally accepted standard means that

the developed tool satisfactorily met the needs of its users and of the host University.

The developed SUC Leveling tool has a high level of validity as revealed in the evaluation done by the information technology experts. This shows that it is technically sound from an information technology perspective and is corroborated by the evaluation of the tool by the participants.

There is a significant difference in the features of the existing and developed system, in favor of the latter. As assessed by the participants, the developed system has a substantial improvement compared to the existing one. This implies that the developed system solves the problems of the existing approach and introduces a better tool for performing SUC Leveling work.

The visualized knowledge generated by the Rapid Miner application was validated by the Planning and Quality Assurance Office. Derived from the data organized by the developed system, this knowledge was deemed to be highly useful for SUC leveling work and for supplying the information required by accreditors.

Generally, the participants are aware of some problems existing in the academic community, specifically, on SUC Leveling Tool. On the other hand, the participants expressed their willingness to adopt the new system to improve their workflow and solve the existing problems in the course of SUC leveling tool system.

Since the participants rated the system "strongly agree" in terms of all the six (6) attributes in the ISO 9126-1 characteristics of a good software product, they would be willing to adopt it.

On the data analytics approach, the participants specifically the Director of Planning and Quality Assurance Office (PQAO), appreciated the tool due to its ability to extract various knowledge hidden in the complex and varied sources of data and visualize them in human-readable form. The data mining and analytics tool is freely available for academic use and uses the data sets generated by the developed system as its input.

RECOMMENDATIONS

From the foregoing conclusions, the following are the researcher's recommendations:

1. The developed enhanced leveling tool for SUCs be deployed to eventually supplant the existing system in order to meet the University objectives as set in the MMSU Information Systems Strategic Plan.
2. The study could be presented to the Commission on Higher Education and the system may be eventually hosted by its web information systems for use of other SUCs in the whole country.
3. High-level personnel who are to use the software may be provided with trainings and knowledge enhancement in data mining and analytics.
4. More research could be conducted to adopt the system for nationwide usage as well as to enhance its capabilities.

References

Accrediting Agency of Chartered Colleges and Universities of the Philippines, Inc. (2003, January). J*ournal of Philippine Higher Education Quality Assurance, 1*(1).

Bagnall, A., Ratanamahatana, C., Keogh, E., Lonardi, S., & Janacek, G. (2006). A bit level representation for time series data mining with shape based similarity. *Data Mining and Knowledge Discovery, 13*, 11–40. DOI: 10.1007/s10618-005-0028-0

Bagnall, A., & Janacek, G. (2005). Clustering time series with clipped data. *Machine Learning, 58*(2-3), 151-178. https://doi.org/10.1007/s10994-005-5825-6

Burkhardt, H., & Schoenfeld, A. H. (2003). Improving educational research: Toward a more useful, more influential, and better-funded enterprise. *Educational Researcher, 32*(9), 3–14. https://doi.org/10.3102/0013189X032009003

Castano, M. C. N. & Cabanda, E. (2011). Sources of efficiency and productivity growth in the Philippine state universities and colleges: A non-parametric approach. *International Business & Economics Research Journal, 6*. DOI: 10.19030/iber.v6i6.3379.

Chang, W. H. T. & Lee, Y. H. (2000). Telecommunications data mining for target marketing. *Journal of Computers, 12*(4), 60-74.

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research (4th ed.)*. Upper Saddle River, NJ: Merrill Prentice Hall.

Cuenca, J. S. (2011, July). Efficiency of state universities and colleges in the Philippines: A data envelopment analysis. *Philippine Institute for Development Studies*. Retrieved from http://dirp3.pids.gov.ph/ris/dps/pidsdps1114.pdf

Han, J., Kamber, M., & Pei, J. (2006). *Data mining: Concepts and techniques (3rd ed.)*. Morgan Kaufmann. Retrieved from http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf

Hastie, T., Tibshirani, R., & Friedman, J. (2008). *The elements of statistical learning: Data mining, inference, and prediction*. New York: Springer. Retrieved from https://web.stanford.edu/~hastie/Papers/ESLII.pdf

Herschel, G. (2008). *Magic quadrant for customer data-mining applications*. Retrieved from http://www.spss.com.hk/pdfs/gartner_magic_quadrant.pdf

Hofmann, M., & Klinkenberg, R. (n.d.). Rapid miner: Data mining use cases and business analytics applications. *Chapman & Hall/CRC Data Mining and Knowledge Discovery Series*. CRC Press.

International Organization for Standardization. (2011). *Systems and software engineering - Systems and software quality requirements and evaluation (SQuaRE) - System and software quality models (ISO/IEC 25010:2011).*

Kadous, M.W., & Sammut, C. (2005). Classification of multivariate time series and structured data using constructive induction. *Machine Learning, 58*. https://doi.org/10.1007/s10994-005-5826-5

Klösgen, W., & Zytkow, J. M. (2002). *Handbook of data mining and knowledge discovery*. Oxford University Press, New York.

Laviña,C. G., Manabo, R. D., Hernandez, G. D., Habanida, F. L., Lacorte, A. M., & Ebron, J. G.(2016). *Outcomes-based practical guide to thesis and capstone project writing in computing*. Mindshapers Co., Inc.

Locke, L. F., Spirduso, W. W., & Silverman, S. J. (2013). *Proposals that work: A guide for planning dissertations and grant proposals (6th ed.)*. Thousand Oaks, CA: Sage.

Norman, D. A. (2002). *Emotion & Design: Attractive things work better. Interactions Magazine, 9*(4), 36–42.

O'brien, J. A. & Marakas, G. M. (2010). *Management information systems (10th ed.)*. New York: McGraw-Hill/Irwin.

Patil, P. S., Rao, S., & Patil, S. B. (2011). Optimization of data warehousing system: Simplification in reporting and analysis. *International Journal of Computer Applications*, 33-37. Foundation of Computer Science.

Paz, A. M., Gerardo, B. D., & Tanguilig, B. T. III. (2014). Development of college completion model based on K-means clustering algorithm. *International Journal of Computer and Communication Engineering, 3*(3). Retrieved from http://www.ijcce.org/papers/314-CS025.pdf

Singh, S., Firdaus, T. &. Sharma, A. K. (2015). Survey on big data using data mining. *International Journal of Engineering Development and Research, 3*(4). Retrieved from https://www.ijedr.org/papers/IJEDR1504022.pdf

Witten, I., Frank, E., & Hall, M. (2000). *Data mining: Practical machine learning tools and techniques (2nd ed.)*.