An Investigation into Data Management as a Strategic Information Tool and Its Importance at the Durban University of Technology

Ms Ramani Francis

Senior Information Analyst, Department of Management Information
Durban University of Technology, 41-43 ML Sultan Road, Durban, 4001, KwaZulu-Natal, South Africa.
Email: ramani@dut.ac.za

Professor D C Jinabhai

Director: Department of Human Resource Management, Faculty of Management Sciences
Durban University of Technology, 41-43 ML Sultan Road, Durban, 4001, KwaZulu-Natal, South Africa.
Email: Jinabhai@dut.ac.za

Professor N Dorasamy

Senior Lecturer: Department of Public Management & Economics, Faculty of Management Sciences
Durban University of Technology, 41-43 ML Sultan Road, Durban, 4001, KwaZulu-Natal, South Africa.
Email: nirmala@dut.ac.za

Abstract

In the early 1990's, various investigations were launched into higher education policies for post-apartheid South Africa. It was in these debates that the South African Post Secondary Education (SAPSE) funding formula was established. It was later established that it was a flawed document as suggested by Bunting (1994:141-149) and cited by Cloete, Maassen, Fehnel, Moja, Gibbon and Perold (2007:93). The SAPSE funding framework which was established in 1982 for the historically White universities applied throughout the late 1980s and early 1990s to all other universities and technikons, was eventually abolished. This was essentially an apartheid funding framework and it could not be used in the current transformed higher education system. The New Funding Framework (2003) gives advantage to comprehensive and traditional universities. It is also graduate output driven. The HEMIS transformed system is committed to equity and strong linkages with national development needs. It was argued by most stakeholder Institutions that with the abolition of the SAPSE funding framework, it would result in major benefits for some groups of institutions and equate to major costs for others. Thus, the significance of this study is to ensure an effective and efficient way for the data to be maintained in order to support the mission and vision of the Durban University of Technology. Bitzer (2009:354) states that South African universities have become highly unpredictable. Universities have moved from an isolated, divided system in the apartheid era toward a single coordinated system. This system has higher levels of participation and responsiveness. Yet it is within a developing economy and poor quality of primary and secondary education that South African higher education therefore faces a multitude of challenges. Bitzer (2009:354) further explains how conditions of unpredictability further worsens the complexities of planning, policy setting and decision making which then impacts negatively on corporate governance and management activities. The purpose was to investigate data management as a strategic information tool and its importance at the Durban University of Technology. The problem revolved around, inter alia, data management and accuracy as structured interventions impacting sound decision making. The research design adopted a quantitative methodological approach that used a precoded self administered questionnaire. The empirical component involved a survey method considering that it was an in-house investigation where the target population equated to 174 respondents. A significant 74% response rate was obtained using the personal method of data collection. Several hypotheses were formulated relating to data quality initiatives, data owners and their responsibility and frequency of data analysis in order to determine accuracy. These were tested using the Pearson chi-square test and data that was analyzed to determine frequencies and percentages of responses. The data was analyzed using the computerized Statistical Program for Social Sciences (SPSS) program. A highly significant finding was that 95.31% of the respondents strongly agreed that data management and integrity is of utmost importance at the Durban University of Technology. One recommendation suggests that for the Durban University of Technology to manage its data as an asset, a policy on data integrity and integration policy should be developed and implemented. Another recommendation highlighted that staff should strive to attain proper classification on the database, considering that this directly impacts on the accuracy of the HEMIS submissions to the Ministry of Education for the state allocated subsidy. The study concludes with directions for further research as well.

Keywords: Data, data integrity, data ownership, data management, data governance

1. Introduction

Data management refers to techniques used to organize, structure, and manage data including database management and data administration (Wang, 2006:569). This definition is comparatively broad and encompasses a number of topics namely; data governance, data integrity, data architecture, analysis and design, data security, data quality management. The focus of this study is based on data integrity as an essential component in relation to data management. According to Beyda (2000:117) data integrity is the prevention or correction of possible errors in data transmission.

Data integrity is a necessity at the Durban University of Technology (DUT), especially in relation to the mandatory data submissions for a specific reporting year to the Department of Higher Education and Training (DHET) from which the institution's subsidy is derived. The institution's subsidy is also known as Government appropriations. Bunting (1994:125) provides the definition as "the fund to which are assigned the amounts received from the central government under the universities' subsidy formula (for both current and fixed asset expenditure). The central government's share of the interest and principal due on long term loans e.g. for buildings or major capital equipment which is also assigned to this fund".

Ensor (2004:339) explains how the newly elected government was faced with the challenge of nurturing democracy from the apartheid regime. Thus, in the past 10 years there has been a significant change in the higher education funding environment. Consequently, the Department of Higher Education and Training (DHET) requires a high level of data integrity from a management information reporting perspective.

One of the most significant challenges faced at the Durban University of Technology is maintaining data integrity throughout the institution. According to Davenport (1994:122), "effective information management must begin by thinking about how people use information and not how people use machines". There is and must be only one main purpose towards improving data management which is to improve customer and stakeholder satisfaction by increasing the efficiency and effectiveness of business processes in an institution.

1.1 Objectives

The primary aim of the study is to investigate data management as a strategic information tool and its importance at the Durban University of Technology (DUT). The objectives of the study are:

- To improve awareness of data management within the Durban University of Technology and help stakeholders investigate the state of data management.
- To provide stakeholders with relevant information regarding data management and effectively apply this information in order to improve data integrity.
- To effectively utilize data management as a strategic information tool in order to provide all levels of management and stakeholders with management information to support strategic decision making, planning, policy development and quality processes.
- To explore the feasibility and effectiveness of data management using Higher Education Management Information System (HEMIS) to the DHET.
- To determine what processes exist for data verification.
- To determine what constitutes best practice in data management and data integrity at DUT.

2. Materials and Methods

2.1 Higher Education Management Information System (HEMIS)

The benefits of HEMIS as contained in The Department of Education's New Higher Education Statistical Reporting System for Students and Staff – A Concept Paper (1999:7) includes the following:

- It would reduce the institutional cost for assembling the data.
- Reduction in costs for the institution and the Department of Higher Education and Training relating to the checking and correction of data.
- The Department of Education will be able to store, access and analyse data more efficiently and effectively.
- There will be more timely provision of data to the Department of Higher Education and Training.
- When the system is fully operational, it will be possible for institutions to compare themselves to national data.
- The Department of Higher Education and Training will be able to undertake a more flexible analysis of data for strategic decision making.

The computer software provided by DHET, that is currently used by all HEI's for HEMIS reporting, is Valpac. This software contains a Help File that has specifications for the unit records of student, staff and building space collections. This help file is numbered accordingly for the specific version and it is subject to change whenever it is revised.

2.2   Strategic tools for data management

Evans (2008:293) shows that just how an artist or craftsperson cannot succeed in achieving high quality services and processes without a good set of tools, so too with organizations. Some practitioners have now adapted a variety of tools from other disciplines like statistics, operations research and creative problem solving to help design, improve and control processes. These tools provide a method by which problems and issues can be viewed objectively. Goetsch and Davies (2010:352) explain that strategic tools are tools for collecting and displaying information in ways to help the human brain grasp thoughts and ideas. According to Oja and Parsons (2011:624), the best data management tool depends on several factors and it is therefore important to consider the cost, versatility and ease of use. Downey (2007:4) specifies that a strategic analytical tool is essential to ensure consistency of data. According to Downey (2007:4) certain factors should be considered when using this tool:

- The tool should be able to fulfill the goals of the organizations.
- The benefit of the tool must be well defined and actionable.
- This tool must benefit from input as well as from collaboration with people and functions with the organization.

These types of tools are generally time consuming, so it is imperative that key stakeholders are well aware of the approximate time it would take for a full analysis.

2.3   Research methodology and design

Accurate and efficient data collection is central to any research. This research adopts a quantitative paradigm as close ended questions were asked in the research instrument. Quantitative research, as defined by Cooper and Schindler (2006:198), attempts precise measurement of something. Quantitative methodologies usually measure consumer behaviour, knowledge, opinions or attitudes. These answer questions related to how much, how often, how many, when and who. Quantitative data often consists of particular responses that are coded, categorized and reduced to numbers so that these data may be manipulated for statistical analysis Cooper and Schindler, (2006:198).

2.4   Secondary data

Secondary data is that which has been collected, collated and analysed by others as opposed to that which one would have collected personally. The secondary data used in this research was the following, as listed by Riley, Wood, Clark, Wilkie, Szivas (2000:107):

- Central and local government studies, reports, policies and rules.
- Academic journals.
- Textbooks.
- Other policy documents from the Department of Higher Education and Training.
- Other published material from Department of Higher Education and Training workshops.
- Internet sites and web pages.
- Audit Reports from DUT.

2.5   Primary data

According to Brewerton and Millward (2000:104), the proper construction of a questionnaire is critical to the accuracy of responses. Respondents would be contacted telephonically informing them of the purpose/reason for the completion of the questionnaire. The questionnaire was hand-delivered, e-mailed or faxed through to the target respondents. The respondent had approximately one week to complete the questionnaire. However, further time was permitted, if requested in order to encourage respondents to complete the questionnaire. A covering letter was attached and a summary of the findings would be e-mailed to the respondents on completion of the study.

2.6 Measuring instrument

Ghauri, Gronhaug and Kristianslund (1995:41) explain that the quality of information depends considerably on the measurement procedures used in the gathering of data. The research instrument to be used in the collection of data would be the survey questionnaire. In addition, a letter of informed consent was obtained from the Research Director prior to conducting the empirical investigation at the Durban University of Technology.

2.7 Target population

This study is an in-house investigation and thus the target population compromised of all Executive Deans, Heads of Departments, Information and Communication Technology staff, Management Information staff and all Faculty officers as well as the staff within the faculty office. The total number of target respondents was obtained from the institutional ITS database was 174.

3. Results

The data collected from the responses was analysed with Predictive Analytic Software (PASW) version 18.0 and SPSS (Statistical Package for the Social Sciences) version 16 for windows. The initial results were presented in the form of graphs and cross tabulations. A total of 174 questionnaires were distributed to the target respondents. The number of returned questionnaires totaled 128 which represented a high response rate of 74%. The data was captured by the researcher from the pre-coded questionnaires returned.

3.1 Reliability of the items in the measuring instrument

Muijs (2011:217) states that reliability refers to the property of a measurement instrument that gives similar results for similar inputs. Cronbach's alpha is a measure of reliability. Treiman (2009:245) explains how reliability is defined as the proportion of the variability in the responses to the survey that is the result of differences in the respondents. The computation of Cronbach's alpha was based on the number of items on the survey (k) and the ratio of the average inter-item covariance to the average item variance.

$$\alpha = \frac{k(cov/var)}{1 + (k-1)(cov/var)}$$

The item variances are all equal, this ratio simplifies to the average inter-item correlation and the result is known as the Standardized item alpha (or Spearman-Brown stepped-up reliability coefficient).

$$\alpha = \frac{kr}{1 + (k-1)r}$$

The coefficient of 0.921 reported for these items is an estimate of the true alpha, which in turn is a lower bound for the true reliability.

3.2 Descriptive statistics

Treiman (2009:114) states that presenting descriptive statistics are very informative as it allows the reader to understand the most basic aspects of the data being analysed. Salkind (2010:8) states that descriptive statistics is used to organize and describe the characteristics of a collection of data. McMillian and Schumacher (2008:150) concur that the use of descriptive statistics is a fundamental way to summarise data, and it is indispensable in the interpretation of the results of the analysis. This section presented the descriptive statistics based on the demographic information of the study. It is presented using tables, cross-tabulations and various types of statistical tests employed to yield a statistical value. Bryman and Cramer (2009:199) describe cross-tabulations as one of the simplest and most frequently used ways of demonstrating the presence or absence of a relationship. The analysis below indicates the scoring patterns of the respondents for the variables that constituted the different categories of the measuring instrument that was personally administered to the respondents.

3.3 Frequencies

Frequencies are used to determine how often a respondent made a certain selection to a dataset. McMillan and

Schumacher (2006:153) explain frequency to be the number of times the same score was attained, while Sekaran and Bougie (2010:313) refer to frequencies as the number of times various sub-categories of a certain phenomenon occur. The percentage can then be easily calculated from the frequencies.

Table 1: Summary of respondents by race and age (n = 128)

		Race					Total
		Black	**Coloured**	**Indian**	**White**	**Other**	
Under 25	Count	2	0	0	0	0	2
	% of Total	1.6%	0.0%	0.0%	0.0%	0.0%	1.6%
25 - 30	Count	6	0	5	0	0	11
	% of Total	4.7%	0.0%	3.9%	0.0%	0.0%	8.6%
31 - 35	Count	11	2	3	0	0	16
	% of Total	8.6%	1.6%	2.3%	0.0%	0.0%	12.5%
36 - 40	Count	4	1	15	2	0	22
	% of Total	3.1%	0.8%	11.7%	1.6%	0.0%	17.2%
41 - 45	Count	4	1	13	6	0	24
	% of Total	3.1%	0.8%	10.2%	4.7%	0.0%	18.8%
46 - 50	Count	2	0	11	2	0	15
	% of Total	1.6%	0.0%	8.6%	1.6%	0.0%	11.7%
51 - 55	Count	1	0	11	6	1	19
	% of Total	0.8%	0.0%	8.6%	4.7%	0.8%	14.8%
56 - 60	Count	0	1	4	8	0	13
	% of Total	0.0%	0.8%	3.1%	6.3%	0.0%	10.2%
> 60	Count	0	0	3	3	0	6
	% of Total	0.0%	0.0%	2.3%	2.3%	0.0%	4.7%
Total	Count	30	5	65	27	1	128
	% of Total	23.4%	3.9%	50.8%	21.1%	0.8%	100.0%

Table 1 indicates that almost half of the target respondents in the survey were (50.8%) Indian, with Black and White respondents constituting similar percentages. The lowest response was from the Coloured race group (3.9%). This population is in proportion of the current staff compliment of the selected departments at the Durban University of Technology. Table 4.2 also reveals that the majority of the respondents (93.7%) were between 25 and 60 years of age while the minor age category of respondents comprised of only 1.6%. Table 4.2 also illustrates that the majority of respondents fell between the 41 and 45 years of age category (18.8%).

Figure 1: Data Management and Integrity
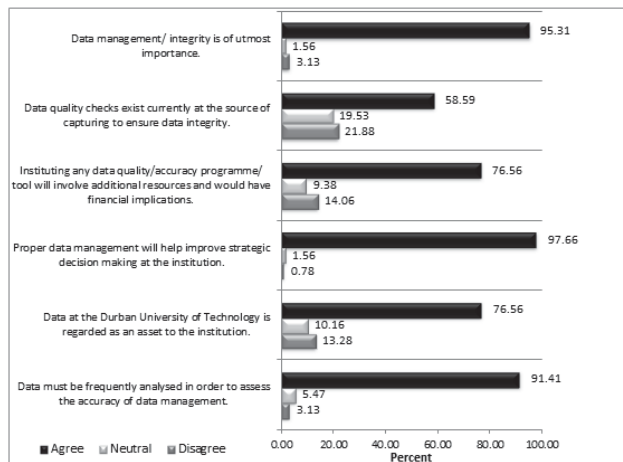
As illustrated in Figure 1, 95.31% of the respondents strongly agreed that data management and integrity is of utmost importance which is a highly significant result while a mere (3.13%) respondents did not agree. A neutral response was reported by (1.56%) of the respondents. It would appear that the majority of the respondents in this study agreed that data management and integrity is of utmost importance.

3.4 Hypotheses Testing

Wilson (2010:237) asserts that hypothesis testing is one of the main methods to test for significance using inferential statistics. It involves an analysis of some aspect of the statement or questions that generates a statistical value. This will determine whether the hypothesis can or cannot be rejected. The traditional approach to reporting a result requires a statement of statistical significance. For the statistical tests a confidence interval of 95% was used to statistically test the various hypotheses. A p-value is generated from a test statistic. A significant result is indicated if the p value is less that 0.05.

In this study the Pearson Chi square test was performed for the various hypotheses formulated. Karl Pearson as referenced by Treiman (2009:93) was the principal developer of the linear regression and correlation and in recognition of this the ordinary least square and correlation is known as Pearson's r. Muijs (2011:124) refers to Pearson's r as whether or not a high score on one variable is associated with a high score on the other.

H_1 There is a significant correlation between data owners and the responsibility for data management and integrity.

Table 2: Frequency of data owners and responsibility for accuracy (n = 128)

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	61.612[a]	40	0.016
Likelihood Ratio	57.949	40	0.033
Linear-by-Linear Association	.903	1	0.342
N of Valid Cases	128		

*Pearson Chi-square = 61.612, df = 40, Significance p<0.016

Table revealed that the Chi-square test result showed that the p value is 0.016, which is less than 0.05. This result indicates that there is a statistically significant correlation between the data owners and the responsibility for data management/integrity. Eckerson (2002:6) highlights how data defects easily creep into systems. Thus, maintaining data quality at acceptable levels takes considerable effort and coordination throughout the organization. Kleppner (2010:9) affirms that data owners (researchers) are ultimately responsible for the integrity of data. Friedman (2006:4) argues that data management is a major initiative for large organizations as they strive for greater integrity and consistency across the organization. Organizations should therefore understand the role of data management in relation to their overall business performance and decisive decision making.

4. Discussions

The main reason for this study was to investigate data management as a strategic information tool and its importance at the Durban University of Technology. What was revealing was that staff are generally aware of the significance of the data that they dealt with from the time as they received it until it gets captured and verified. Data management and data integrity may be distinct yet have overlapping fields. Each has its own focus and each discipline fulfills different purposes within an organization. However, it is becoming more evident that the coordination of data management and data integrity can increase an organization's efficiency and effectiveness. Arising from the empirical analysis of the data the following recommendations are suggested.

4.1 Recommendations

4.1.1 Knowledge Transfer

It is recommended that staff in the faculty office should be given on the job training by their Faculty Officer who have extensive knowledge. Similarly, Ellis (2005:29) asserts that the importance of developing employee knowledge beyond

basic understanding of job requirements is not required and is largely an activity that should be separate from work itself. Alavi and Leidner (2001:128) contend that the ease with which individuals are able to transfer their explicit components of their knowledge to their work situation is expected to transfer more than verbal or face-to-face communication.

4.1.2 Verification Process

It is critical that verification process should be done immediately after the data is captured in order to reduce the errors. Eckerson (2002:4) emphasizes that validation routines cannot catch typographical errors where the data represents a valid value. While Cascarino (2007:11) contends that when it comes to an on-line system, it permits remote entry of data and also allows for concurrent processing of data. It is imperative that there should be transaction authenticity as well as accuracy and completeness of the data management. Madanjit (2006:18) also argued that users at the Durban University of Technology sometime captured only fields that they were using and did not capture required fields by the institution. He also stated that data was sometime not verified against source documents. Hence, the verification of data should be expedited immediately for its accuracy and integrity.

4.1.3 Assigning Accountability For Data Management

Kavanagh and Suppert (2007:3) state that there must be accountability in order to achieve the results and realizing the value for the organization. Similarly, at the Durban University of Technology there should be accountability especially at the level of capturing. Madanjit (2006:16) contended that in order to ensure accuracy of data at the Durban University of Technology, Executive Deans, the Registrar, Directors and Heads of Departments should take ownership of the data by being accountable and responsible for the integrity of the data.

5. Acknowledgement

References

Alavi, M. and Leidner, D.E. (2001). Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *MIS Quarterly*. 25 (1): 107-136.
Beyda, W.J. (2000). *Data Communications from Basics to Broadband*. 3rd Edition. New Jersey: Prentice-Hall.
Bitzer, E. (2009). *Higher Education in South Africa – A Scholarly look behind the Scenes*. Cape Town: Sun Media Stellenbosch.
Brewerton, P. and Millward, L. (2006). *Organizational Research Methods.* London: Sage Publications.
Bryman, A. and Cramer, D. (2009). *Quantitative Data Analysis with SPSS 14, 15 and 16.* New York: Routledge.
Bunting, I. (1994). *A Legacy of Inequality Higher Education in South Africa*. Cape Town: UCT Press (Pty) Ltd.
Cascarino, R. (2007). *Auditor's Guide to Information Systems Auditing*. New Jersey: John Wiley & Sons Inc.
Cloete, N., Maassen, P., Fehnel, R., Moja, T., Gibbon, T. and Perold, H. (2007). *Transformation in Higher Education.* The Netherlands: Springer.
Cooper, D.R. and Schindler, P.S. (2006). *Business Research Methods.* 4th Edition. New York: McGraw Hill.
Davenport, T.H. (1994). Saving IT's Soul: Human-Centered Information Management. *Harvard Business Review*. 72 (2): 119-131.
Eckerson, W.W. (2002). Data Warehousing Special Report: Data Quality and the bottom line. *The Data Warehousing Institute.* (Online). Available at: http://www.estv.ipv.pt/PaginasPessoais/jloureiro/ESI_AID2007_2008/fichas/TP06_anexo1.pdf. Date accessed: 10 April 2012.
Ellis, S. (2005). *Knowledge-Based Working*. Oxford: Chandos Publishing.
Ensor, P. (2004). Contesting Discourses in Higher Education Curriculum Restructuring in South Africa. *Higher Education.* 48 (3): 339-359. (Online). Available at: http://www.jstor.org/stable/4151521. Date accessed: 1 December 2011.
Evans, J.R. (2008). *Quality and Performance Excellence*. 5th Edition. Ohio: Thomson South Western.
Friedman, T. (2006). *Key Issues for Data Management and Integration*. Connecticut: Gartner Inc.

Ghauri, P., Gronhaug, K. and Kristianslund, I. (1995). *Research Methods in Business Studies*. London: Prentice Hall.

Goetsch, D.L. and Davis, S.B. (2010). *Quality Management for Organizational excellence: Introduction to Total Quality.* New Jersey: Prentice Hall.

Kavanagh, S.C. and Suppert, M. (2007). *We're all in IT together: aligning technology with business through IT governance.* (Online). Available at: http://www.thefreelibrary-com. Date accessed: 8 September 2008.

Kleppner, D. (2010). *Ensuring the Integrity, Accessibility, and Stewardship of research Data in the Digital age.* (On-line). Available at: http://docs.lib.purdue.edu/iatul2010/conf/day1/10. Date accessed: 3 April 2012.

Madanjit, M. (2006). *Management Policy and the Decision Making process including Data Ownership and Validation.* Deputy Vice Chancellor Financial and Information Services. Presented at the ITS User Group, Botswana, 13 March 2006.

McMillan, J.H. and Schumacher, S. (2006). *Research in Education Evidence – based Inquiry*. 7th Edition. Boston: Pearson Education Inc.

Muijs, D. (2011). *Doing Quantitative Research in Education with SPSS.* 2nd Edition. California: SAGE Publications Inc.

Oja, D. and Parsons, J.J. (2011). *Computer Concepts.* Boston: Cengage Learning.

Republic of South Africa. (1999). *A New Higher Education Statistical Reporting System for Students and Staff – A Concept Paper.* Department of Education. Pretoria: Government Printer.

Riley, M., Wood, R.C., Clark, M.A., Wilkie, E. and Szivas, E. (2000). *Researching and Writing Dissertations in Business and Management.* London: Thomas Learning.

Salkind, N.J. (2010). *Statistics for People Who think they hate Statistics.* London: Sage Publications Ltd.

Sekaran, U. and Bougie, R. (2010). *Research Methods for Business.* 5th Edition. West Sussex: John Wiley & Sons Ltd.

Treiman, D.J. (2009). *Quantitative Data Analysis Doing Social Research to Test Ideas.* San Francisco: John Wiley & Sons Inc.

Wang, J. (2006). *Encyclopedia of Data Warehousing and Mining.* London: Idea Group Reference.

Wilson, J. (2010). *Essentials of Business Research*. California: Sage Publications Inc.