

Towards Developing An Instrument In Measuring The Need for InfoVis

Akanmu Semiu Ayobami¹, and Zulikha Jamaludin²

¹School of Computing, Universiti Utara Malaysia (UUM), Sintok, Malaysia, ayobami.sm@gmail.com

²School of Computing, Universiti Utara Malaysia (UUM), Sintok, Malaysia, zulie@uum.edu.my

ABSTRACT

The increasing trend of data volume and its superabundance have been endangering institutional data and specifically, higher education institutions (HEI) students' data. The most challenging is the need for HEI to make sense from their large datasets, through gaining insights and understanding the pattern and trends of events therein. Deductions from extant literatures strongly indicate the presence of information overload as the factor constraining HEI decision makers from making wealthy use of the institutional datasets. However, no study has empirically investigated the presence of information overload in HEI students' data management. To attend to this, this study aims at developing an instrument to be used in measuring the presence of information overload, and justifiably, the need for Information Visualization (InfoVis) –being an argued better tool for institutional data management. This study employs quantitative research method with administration of 9-item survey questionnaire. Thirty-two (32) respondents are purposively drawn among HEI decision makers. Descriptive statistics is used as the statistical technique to find the mean value of the computed variable based on the normal Likert 5-point scale. The result of the instrument reliability test gives a value of 0.712 as the Cronbach's Alpha value which suggests that the items designed are internally consistent, and a weighted mean value of 4.03 strongly supports the hypothesis that HEI experiences information overload.

Keywords: information visualization, information overload, higher education institutions (HEI), institutional data management

I INTRODUCTION

InfoVis is one of the two categories of Visualization(Spence, 2007; Shneiderman&Plaisant, 2010). Visualization, according to Ware (2000) is a graphical representation of concepts and data in a manner that supports decision making. The feature that stands InfoVis out is its presentation and representation of different forms of datasets in an

interpretative manner to all categories of users irrespective of their computer tech savviness. Also, InfoVis as a concept through its tools supports users' interactions in ways that enhance insights about the datasets, aid exploratory data analysis and decision making (Spence, 2007; Shneiderman&Plaisant, 2010; Ware, 2000). The ability to visualize: 'to form mental model or mental image of something' (Spence, 2007) is the core value of an InfoVis.

Notably, the problem-solving nature of InfoVis has been the driving force for researches in the field. InfoVis as a research domain is interested in data discovery and analysis through visual exploration [4]. As Meyer (2012) and Stasko(2008) explain, InfoVis centers on how to make sense of data. Making sense of data in institutional data management, just like other domains however depend on how understandably represented and visually presented is the data to the users.

In institutional data management, especially managing the students' data of HEI, literature review extensively suggests that problem of information overload exists in the domain. The constraint to wealthy use of the massive datasets and profound decision making process is argued to be caused by the limitations of currently used data management tools, inexistence of the domain-focused InfoVis for HEI and unavoidable data growth. This study is therefore necessitated to develop an instrument in measuring the need for InfoVis by empirically investigating and testing this hypothesis that: HEI experiences information overload. The acceptance of this hypothesis is posited to subsequently justify the need for InfoVis for managing HEI students' data.

II STUDIES ON HEI DATA MANAGEMENT TOOLS

Studies on HEI data management tools ranging from software systems, data mining technologies, data warehousing, are not many. Nevertheless, review of the available ones shows their limitations, and thus necessitates the need for this study.

Sarker et al. (2010) worked on HEI data warehousing. The outstanding contribution of the work is delivery

of a web-based interface that allows stakeholders in HEI to remotely access and share data from the data repository. However, the non-inclusion of exploratory data analysis and interactive features will limit the extent of insights to be derived from the data repository and minimize the user-control capacity. Also, students' data management was not involved in the scope of the tool.

Delavari et al. (2008) and Bresfelan et al. (2009) delivered data mining technologies, with the objective of improving the data analytical ability of the HEI data management tool. Though Bresfelan et al. (2009) was specifically to proffer better curricula and syllabi design, both lack visualization and interactive features, just as Pocius and Reklaitis(2000) that delivered a user interface for the internetworked HEI data repositories without the inclusion of students' data.

On another hand, El-Fattah (2012) and Pinto et al. (2012) worked on visual data model and InfoVis respectively. El-Fattah's(2012) deliverable is usable for distributing teaching resources, but it does not involve students' data, and also lacks exploratory data analysis and interaction. Pinto et al.'s(2012) work – being InfoVis tools – contain integrated features for visualization and interaction, but none of the four InfoVis tools centered on students' data.

Summarily, previous works on HEI data management have shown that much is still needed to be done towards its decision making support tool, especially as it bothers on students' data. The inherent limitations are the pointer to the cause of information overload. This study therefore develops a survey instrument to measure information overload, and thus justifies the need for InfoVis in managing the HEI students' data. Figure 1 presents a graphical representation of the severity of the limitation as observed from the reviewed literatures.

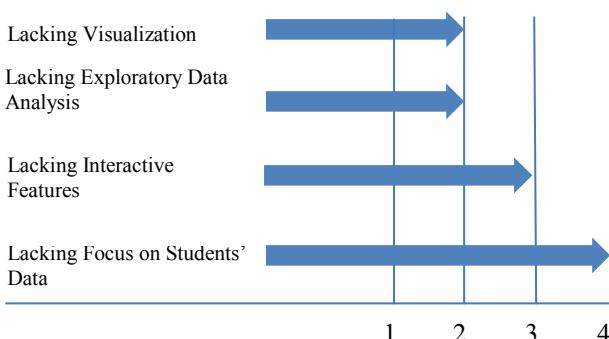


Figure 1: Graphical Representation of the Limitations

III SURVEY INSTRUMENT DEVELOPMENT AND DATA COLLECTION PROCESS

The 9-item survey questionnaire is developed through adaptation of items and conceptual explanation from past literatures. Table 1 presents the summary showing the supporting references and codes for the items.

Table 1: Summary of the supporting references and codes for the items

Code	Items	Sources
IO1	Sensitive data is not secure from unauthorized access	Keim et al. (2008)
IO2	We are experiencing growth in our volume of data	Haksever(2000)
IO3	The growth of our data volume needs better analytical tool	Whelan (2011); Yanosky(2009)
IO4	Our data analysis tool cannot be used by some of the administrative personnel	Yanosky(2009)
IO5	Employees do not understand their responsibilities in the use of data	Fekete and Plaisant(2002); Keim et al. (2008)
IO6	Employees cannot manipulate our datasets and transform it to their desired information	Fekete and Plaisant(2002); Keim et al. (2008)
IO7	We get maximum administrative decision support from our institutional data	Fekete, van Wijk, Stasko, and North (2008)
IO8	We get maximum academic (teaching and learning) value from our institutional data	Fekete, van Wijk, Stasko, and North (2008)
IO9	We get maximum business (students' enrolment, industries consulting) value from our institutional data	Fekete, van Wijk, Stasko, and North (2008)

The questionnaires were redistributed to thirty-two (32) respondents purposively selected among HEI decision makers. Data collected are analyzed using Statistical Package for the Social Sciences (SPSS).

IV DATA ANALYSIS

A. Items Recoding

As shown in table 1, the items designed to measure the experience of information overload in HEI were firstly coded before input into the analysis tool. It should be noted that items IO7, IO8 and IO9 are worded in an opposite direction as against the phenomenon that is under study. This is to check for bias in the responses gathered (Pallant, 2011). In order to avoid its negative influence on the statistical result, they are recoded and formed into another items

with code REOI7, REOI8 and REOI9 respectively using the SPSS.

B. Reliability Test

A reliability test was conducted to validate the internal consistency of these items and show that they duly measured the intended phenomenon. For a stronger Cronbach's Alpha value, item IO1 is recommended for deletion. The reliability test gives a Cronbach's Alpha value of 0.712. According to Pallant(2011), a value of 0.6 and above for Cronbach's Alpha represents an internal consistency of the items measuring the construct.

C. Descriptive Statistics

For the descriptive statistics, the mean value of the construct computed as NTOI, i.e. the summation of IO2, IO3, IO4, IO5, IO6, REOI7, REOI8 and REOI9 is given as 4.03. These items give weighted sum for the assessment of the phenomenon under study; Information Overload in HEI.

V INTERPRETATION AND FINDINGS

A Cronbach's Alpha value greater than 0.6 is acceptable, while that greater than 0.7 is preferable (Pallant, 2011). Therefore, Cronbach's Alpha value of 0.712 is preferable. This shows that items IO2, IO3, IO4, IO5, IO6, REOI7, REOI8 and REOI9 are suitable for the assessment of information overload.

The mean value of the descriptive statistics given as 4.03 on 5-point Likert scale 1 = Strongly Disagree, 2= Disagree, 3= Neutral, 4= Agree and 5= Strongly Agree, showed that the hypothesis: HEI data management experiences information overload is accepted.

VI LIMITATION OF THE STUDY

The limitation of the study is the size of its sample which could not afford the opportunity of conducting construct validity using factor analysis. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy result is given as 0.47 which is below the expected least value of 0.6 (Pallant, 2011). Therefore, future study is recommended to conduct the construct validity of the construct measured by the developed instrument using a larger sample size. However, items developed for

measuring information overload in HEIs can be adapted to all domains.

VII CONCLUSION

Having observed the limitations in the previously deployed tools for institutional data management, empirically assessed information overload, and established its experience in HEI data management, we suggest that deployment of InfoVis tool is the best approach to deal with the posed constraints in the HEI decision making process. InfoVis helps in exploratory data analysis, with dynamic presentation and appropriate representation. Its applicability has been domain specific in view of addressing the problem of information overload, i.e. getting lost in data which may be irrelevant to the task at hand, processed in an inappropriate way or presented in an inappropriate way (Keim et al., 2008). It is opined that with appropriate information representation, and presentation through a befitting integration of visualization, interaction and visual data mining techniques, an InfoVis tool that aligns with the Shneiderman's information seeking mantra: Overview first, zoom and filter and then detail-on-demand (Shneiderman, 1996) will solve the problem of information overload. Our future work encompasses designing the InfoVis conceptual framework and subsequently developing the InfoVis prototype.

REFERENCES

Bresfelean, V.P., Ghisoiu, N., Lacaurezeanu, R., & Sitar-Taut, D.A (2009). Towards the Development of Decision Support in Academic Environments, *Proceedings of International Conference on Information Technology Interfaces*, ITI, '09, pp. 343- 348

Delavari, N., Phon-Amnuaisuk, S., & Beikzadeh.M.R. (2008). *Data Mining Application in Higher Learning Institutions*, Institute of Mathematics and Informatics, Vilnus, 7(1), 31-54.

El-Fattah, M.A. (2012). Visual Model for Managing Educational Capacity Utilization in Egyptian Universities, in *Proceedings of International Conference on Information Technology Based Higher Education and Training*, pp. 1-4.

Fekete, J-D. &Plaisant, C. (2002). Interactive Information Visualization of a Million Items, *Proceedings of the IEEE Symposium on Information Visualization* (InfoVis '02).

Fekete, J-N., van Wijk, J.J., Stasko, J.T. & North, C. (2008). The Value of Information Visualization. *Information Visualization: Human-Centred Issues and Perspectives* (E.d.), pp. 1-18.

Haksever, M.A. (2000). A Model to Predict the Occurrence of Information Overload of Project Managers, *Construction Informatics Digital Library*, 7 (8), 4-15

Keim, D., Andrienko, G., Fekete, J-D., Gorg, C., Kohlhammer, J. & Melancon, G. (2008). Visual Analytics: Definition, Process, and Challenges. A Kerren et al. (Eds): *Information Visualization, LNCS* 4950, pp. 154-175. Springer-Verlag Berlin Heidelberg.

Meyer, M. (2012). *Designing Visualization for Biological Data*, A Keynote talk at 3rd Leonardo Satellite Symposium of Arts, Humanities and Complex Networks (NetSci, 2012)

Pallant, J. (2011). *SPSS Survival Manual: A Step-by-Step Guide to Data Analysis Using SPSS Program*, 4th Edition. Australia: Allen & Unwin.

Pinto, M., Raposo, R. & Ramos, F. (2012). Comparison of Emerging Information Visualization Tools for Higher Education, *Proceedings of 16th International Conference on Information Visualization*, pp. 100-105.

Pocius, K. & Reklaitis, V. (2000). Web-based Services for Technology Education Management and Planning, *Proceedings of International Workshop on Advanced Learning Technologies*, IWALT, 2000, pp. 284- 285.

Sarker, Farhana, Davis, Hugh and Tiropanis, & Thanasis (2010). The Role of Institutional Repositories in addressing Higher Education Challenges. The Second International Workshop on Semantic Web Application in Higher Education, University of Southampton, UK.

Shneiderman, B. &Plaisant, C. (2010). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 5th Edition U.S.A: Pearson Higher Education.

Shneiderman, B. (1996). The eyes have it: a task by data type taxonomy for information visualizations, presented at *IEEE Symposium on Visual Languages*, 1996, Boulder, CO, USA, pp. 336-343.

Spence, R. (2007). *Information Visualization: Design for Interaction*, 2nd Edition, England. Pearson Education Limited.

Stasko, J. (2008). Visualization for Information Exploration and Analysis, *Proceedings of IEEE Symposium on Visual Languages and Human-Centric Computing* (VL/HCC '08).

Ware, C. (2000). *Information Visualization: Perception for Design*. United States of America: Morgan Kaufmann Publishers.

Whelan, E. (2011). *Managing Information Overload: Examining The Role of the Human Filter*, Kemmy Business School, Ireland

Yanosky, R. (2009). *Institutional Data Management in Higher Education*. A Report from EduCause Centre for Applied Research. Retrieved on 6th June, 2013 from <http://net.educause.edu/library/pdf/EKF/EKF0908.pdf>