Big Data and a Smarter University: A Literature Review

Sameer Hayikader *, Mohd. Toriq Khan bin Mohamad Niyaz Khan **, Abdulrahman Dahlan ***

* sameer797.kader@gmail.com, ** mohdtoriqkhan@hotmail.com, *** arad@iium.edu.com

Department of Information Systems, Kulliyyah of Information and Communication Technology,
International Islamic University Malaysia

Abstract- Using big data in a university is 'the thing' today since almost all information today is digitalized. Big data ensures better data processing, analysis and sharing, thus increasing productivity and researchers would have a great time doing their researches. A smart university is a university which uses technologies which could increase the productivity of knowledge management. Changing into a smart university is definitely a smart move because it is important to be competitive in a world full of formidable competitors. Research has proven that smart universities such as Caltech University and Northwestern University which are among the top universities in the world has also endeavored into the big data arena.
.

Index Terms- Data warehouse, big data, smart university, sharing knowledge.

I. INTRODUCTION

A data warehouse is designed to integrate data from various sources to a single platform and structure. Data warehouse is different from other applications it has its own specification and characteristics which make it unique. It provides tools to meet the information needs of the users at all levels, not just for complex data queries, but as well as in term of retrieving a quick, accurate, and often insightful information. A data warehouse is designed so that its users can recognize the information they want and access that information using simple tools [1]. The data warehouse is not only designed to have a single platform, but it also to have more exploring objectives such as summarization and data mining [2]. Data integration means collecting and combining data from various source systems for usage by users. Data integration was limited to transactional systems and their applications. Big data processing needs to be implemented as a data-driven architecture that includes the analysis and design of data processing, which all the data within the enterprise are categorized according to the data type [3]. Data warehouse and big data integration must respond quickly to business needs, balance requirements against implementation options using their technical knowledge, and build the requisite structures and processes. University of Cambridge which ranks number two in the QS top universities ranking is using big data and they are still researching its potentials and sharing their knowledge in this what they called "growing field [20]." Another example among the top university that uses big data is university of Washington. University of Washington offers education and research in big data. They now have several projects regarding big data such as the AstroDB, Myria, CQMS, Data Eco$y$tem and SQLShare [21].

II. BACKGROUND

Universities that operate its system based on multiple systems of different data sources that include Oracle, MySQL, SQL server etc. will affect the reliability and the quality of data in term of incomplete data integration. A centralized system means that database and information are centralized and that will make the university work more efficiency [4]. Another problem is that staffs and top managers have to spend much time in business decisions based on limited data from different department. Data warehouse has several benefits once it's launch as it will enhance business intelligence, cut time consuming , generates ROI profit. And also organizations that have implemented data warehouses and complementary BI systems have generated more revenue and saved more money than organizations that haven't invested in BI systems and data warehouses [5].

Another advantage of having centralized database is that sharing all information and knowledge can be done much easier and accessed into it in a single location. By having a centralized database looking for the information in database can be faster because the search engine needs to check only a single location rather than multiple locations to get the results. Sharing knowledge and Information can be organized easier in a single location. Centralized database makes adding server to the database site not complicated so that the company will not have to balance the needs of a distributed database [6]. Combine spreadsheets from each department in order to generate a report manually cost time. Process that done manually in universities in order to get a report or approval from various departments or to some analysis from colleagues will take longer time than expected. By implementing a data warehouse will help centralize and gather data and make it available to all team members more effectively [7].

The first smart tertiary education centre in Malaysia is KDU University College. KDU focuses on learning from people and individual experiences. KDU is using a technology called the 'Grammarly@EDU' where it facilitates academic writing [18]. Example of an international smart university is University of Michigan, where they would perform full-time research in a faculty laboratory, have weekly luncheon presentations where faculty discuss their research, have short presentations of their summer accomplishments at the conclusion of the program, obtain clinical exposure opportunities in the form of physician

shadowing and students will be able to receive a stipend of $4400, plus housing [19]. By becoming a smart university, not only the university's ranking will rise, but also other aspects related to the university, such as increased profit, student enrollment, researching abilities and publishing and the university would also be able to share and receive knowledge from other universities [12][16][17].

III.   PROBLEM STATEMENTS

1) Malaysian universities are still left behind in terms of university world ranking due to several shortcomings including the lack of technology implementation which includes big data and offering big data courses.

2) Universities in Malaysia lack the initiatives to implement big data and becoming smart university.

IV.   LITERATURE REVIEW

Big data is in many ways an evolution of data warehousing and the business benefits of big data are potentially revolutionary [8]. According to [9], big data was included as a data warehouse modernization. Today we can see that big data has been descended to many fields such as the healthcare, government, e-commerce and universities [22][10]. However, according to [11], data warehousing and big data are actually two separate things; that data warehousing is only architecture to store data while big data is a technology. In the context of this research, as far as a university is concerned, in relation to data warehousing and big data, they can be used to store a huge archive of students', lecturers' and employees' data and that data can be used to be as a reference for future lecturers and employees to attract future students or stakeholders or to look into better ways to teach students which in return, would generate more income for the university. Just like what a research on the Bank of America had done towards its customers on what kind of loan was the most requested by its stakeholders using data mining which is the basis for data warehousing and big data and they have found out that home equity loan was the most requested loan and they acted upon that data and they successfully managed to increase their stakeholders' response rate from 0.7 percent to 7 percent [12]. All of these researches didn't provide the actual feelings of the recipients or stakeholders' feelings towards the implementation of big data into data warehouse into a university environment, which will be the aim of this research besides performance statistics. A smart university is actually defined first by the 'smart' term which according to [13], as a strategic approach to economic development through targeted support to Research and Innovation. While according to [14] & [15], the smart term is actually abbreviation for S, Specific, M is Measurable, A is Achievable, R is Relevant and T is Timely. Proving that big data can bring better achievements to a university can be seen from University of Florida. Its vice president and chief information officer, Elias Eldayrie said that "The processing and analysis of our faculty's research has been greatly increased, meaning their important work in health care, climatology, and other critical areas makes it into the global conversation faster. We have

faculty conducting brain injury research. We have faculty studying how oil spills impact the fisheries along the Gulf of Mexico, and so many more projects of significant state and worldwide impact. Their research in these fields of study can get to the marketplace quicker thanks to the computing abilities we now have in place at UF (University Florida)." Big data had reduced the time needed to process large data queries of 10 million to 100 million records from 27 hours to a mere 3 seconds, allowing the UF faculties to analyze their research more efficiently and share findings with top institutions, as well as the scientific organizations and government agencies sponsoring UF's expansive research portfolio [16]. [17], the world top university of 2014 had also launched a summer school on big data because Caltech University believes that students and postdocs need to master such skills in order to be effective researchers today. Clearly, big data and data warehousing are fields to be explored and mastered to secure a better future not only for business, but even for universities. A university's rank is one of the criteria for a student's choice to study in that particular university [24]. A university's rank is determined by these criteria which include academic reputation, employer reputation, student-to-faculty ratio, citations per faculty and international faculty ratio and international student ratio [23].

Based on all of these researches, there are a lot of benefits in becoming a smart university, such as increased education quality, research output that is by increasing the analysis and understanding of lots of information through the use of big data [22], ranking and performance. To become a smart university, a university should have the required technologies such as a good server, classrooms with technological devices equipped, a special technology that would help the universities functions and having the S.M.A.R.T objectives which includes efficient and reliable employees and efficient standard of procedures (SOP) or processes in order to complete tasks [25] or have a strategic approach to economic development through targeted support to Research and Innovation [13].

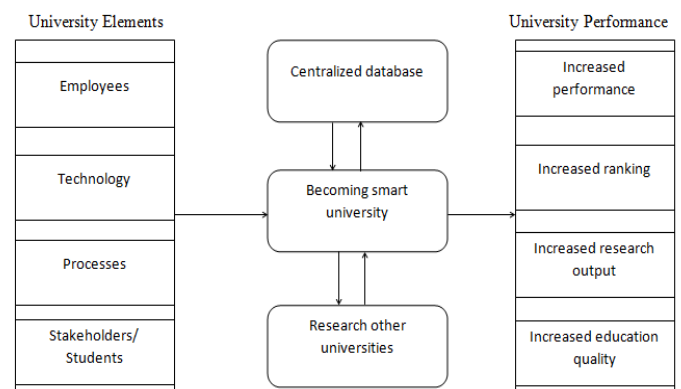### V.   INITIAL RESEARCH FRAMEWORK

Figure 1. initial research framework.

VI.   SHARING KNOWLEDGE IN A SMART UNIVERSITY USING BIG DATA

Knowledge sharing in software enterprise is basically about teams working together to achieve shared objectives. A smart

university can make the use of big Data technologies and social network to share knowledge and experience by making processes, artifacts, and knowledge more explicit and sharable.
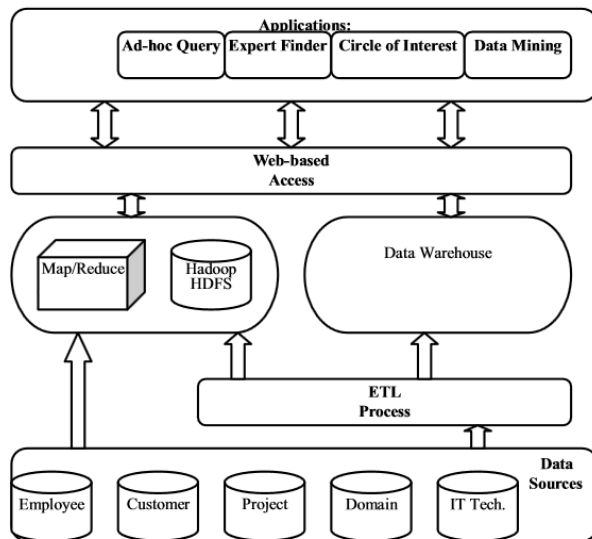

Figure2. A framework of knowledge sharing in software enterprise. [26]

There are three components included In this framework of knowledge sharing which are:

1) Data sources: information on employees' personal information, information on customer with industry characteristics, information on completed project in organization, domain information on software engineering, knowledge on IT Tech. These are operational data sources.

2) Hadoop, HDFS and Data warehouse: managing and storing structured knowledge and unstructured knowledge on people, projects, products, and processes.

3) Applications: including knowledge artifacts looking up and sharing and mining using Ad-hoc query, Expert finder, Circle of interests and data mining tools.

VII.   CONCLUSION AND FUTURE WORK

As a conclusion, the integration of big data into university for a smarter university is definitely something to be looked into as it would provide a better future for both students and the universities employees. Besides increasing profit and the ranking for the university, it would also increase the stakeholders' satisfaction towards the services of the university. This project will ensure continuous innovation to the university, which may open ways to other innovations for a better university since this integration of big data and data warehouse will provide the

capability of storing large amount of data, it will provide easy access to all data in a centralized database from all departments. As well as to address the importance of using Big data in sharing knowledge. For future work, surveys and analysis towards the implementation of big data and becoming smart university will be investigated to improve university performance.

REFERENCES

[1] A. Perkins, "Critical Success Factors for Data Warehouse Engineering," p. 1.

[2] J. Wang, Data Warehousing and Mining : Concepts, Methodologies, Tools, and Application, New York: Information Science reference, 2008.

[3] K. Krishnan, "Search Data Management," Tech Target, [Online]. Available: http://searchdatamanagement.techtarget.com/feature/Author-explains-big-data-integration-in-the-data-warehouse. [Accessed 27 2 2014].

[4] "Effective Database Management," [Online]. Available: http://effectivedatabase.com/resources/why-should-you-have-a-centralized-system/. [Accessed 25 2 2015].

[5] S. B. Team, "TIBCO spotfire's," 10 August 2011. [Online]. Available: http://spotfire.tibco.com/blog/?p=7597. [Accessed 25 Feb 2015].

[6] M. McMahon, "wise GEEK," 10 Feb 2015. [Online]. Available: www.wisegeek.com/what-is-a-centralized-database.htm. [Accessed 25 Feb 2015].

[7] Editor, "TechAdvisory," 2014. [Online]. Available: http://www.techadvisory.org/2014/08/does-your-company-need-a-data-warehouse/. [Accessed 25 Feb 2015].

[8] Oracle, "Oracle Database 12c for Data Warehousing and Big Data", Oracle White Paper, 2014.

[9] IBM, "Data Warehouse Modernization," 2015. [Online]. Available: http://www01.ibm.com/software/data/bigdata/use-cases/data-warehouse.html. [Accessed 24 Feb 2015].

[10] Chen, Hsinchun, Chiang, Roger H.L. & Storey, Veda C. "Business Intelligence and Analytics: from Big Data to Big Impact," MIS Quarterly, 2012.

[11] Inmon, Bill, "Big Data Implementation vs. Data Warehousing," 2013. [Online]. Available: http://www.b-eye-network.com/view/17017. [Accessed 24 Feb 2015].

[12] Berry, Michael J.A. & Linoff, Gordon S., " Data Mining Techniques for Marketing , Sales and Customer Relationship Management. A Case Study in Business Data Mining". (Second Edition, pp. 22-28), 2004. [Online] Available: Google Book. [Accessed 24 Feb 2015].

[13] Kempton, Louise, Goddard, John Edwards, Barbara Hegyi, Fatime & Perez, Susana-Elena, "Universities and Smart Specialisation. Background to 'Smart Specialisation' (pp. 4)," European Commision Joint Research Centre, 2013.

[14] UMass Dartmouth, "Creating S.M.A.R.T Goals," 2015. [Online]. Available: http://www.umassd.edu/fycm/goalsetting/resources/smartgoals/. [Accessed 24 Feb 2015].

[15] Wake Forest University, "SMART Goal Setting Instructions," 2012. [Online]. Office of Personal & Career Development. [Accessed 24 Feb 2015].

[16] University of Florida, "University of Florida receives $500,000 'Big Data' boost from IBM," 2014. [Online]. Availabe: http://news.ufl.edu/archive/2014/09/university-of-florida-receives-500000-big-data-boost-from-ibm.html. [Accessed 24 Feb 2015].

[17] Caltech University. "Big Data Summer School Is in Session—Virtually," 2014. [Online]. Availabe: http://www.caltech.edu/news/big-data-summer-school-session-virtually-43622. [Accessed 24 Feb 2015].

[18] KDU University College, "Student Entry", 2015, [Online]. Available: http://www.kdu.edu.my/teaching-learning/student-entry. [Accessed 10 March 2015].

[19] University of Michigan, "UM-SMART Undergrad Summer Program, Summer at Michigan for Undergraduate Research Training UM-SMART," Available: http://medicine.umich.edu/medschool/education/mdphd-program/um-smart-undergrad-summer-program. [Accessed 14 March 2015].

[20] University of Cambridge, "Cambridge Big Data," [Online]. Available: http://www.bigdata.cam.ac.uk/

[21] University of Washington, "Big Data Research and Education" [Online]. Available: http://www.cs.washington.edu/research/bigdata

[22] Shaw, Jonathan, "Why Big Data is a Big Deal," [Online]. Available: http://harvardmagazine.com/2014/03/why-big-data-is-a-big-deal

[23] QS Top Universities, "QS World Universities Rankings: Methodology," [Online]. Available: http://www.topuniversities.com/university-rankings-articles/world-university-rankings/qs-world-university-rankings-methodology

[24] National Union of Students (NUS), "NUS Student Experience Report," (pp. 3). Student Experience Report, 2008.

[25] Wayne State University, "S.M.A.R.T Objectives," [Online]. Available: http://hr.wayne.edu/leads/phase1/smart-objectives.php

[26] W. l. LIU, "Framework of sharing knowledge in software orgnization using big data," *IEEE xplore,* p. 271.

AUTHORS

First Author – Sameer Hayikader, Master of Information Technology, Department of Information Systems, Kulliyyah of Information and Communication Technology, International Islamic University Malaysia and sameer797.kader@gmail.com.

Second Author – Mohd Toriq Khan Bin Mohamad Nyiaz Khan, Master of Information Technology, Department of Information Systems, Kulliyyah of Information and Communication Technology, International Islamic University Malaysia and mohdtoriqkhan@hotmail.com.

Third Author – Abdulrahman Dahlan, Lecturer and Senior Academic Fellow, Department of Information Systems, Kulliyyah of Information and Communication Technology, International Islamic University Malaysia and arad@iium.edu.com.