TECHNOLOGY ENHANCED ANALYTICS (TEA) IN HIGHER EDUCATION

Ben Kei Daniel and Russell Butson
Educational Technology Group - Higher Education Development Centre - University of Otago - Dunedin, New Zealand

ABSTRACT

This paper examines the role of Big Data Analytics in addressing contemporary challenges associated with current changes in institutions of higher education. The paper first explores the potential of Big Data Analytics to support instructors, students and policy analysts to make better evidence based decisions. Secondly, the paper presents an institutional framework for exploring Big Data at the University of Otago in New Zealand. Thirdly, a series of use-case scenarios are presented to demonstrate the benefits of Big Data in Higher Education, and some of the challenges associated with implementation. Finally the paper concludes by outlining future directions relating to the institutional project on Big Data at the University of Otago.

KEYWORDS

Big Data, learning analytics, Higher Education

1. INTRODUCTION

Institutions of higher education are increasingly facing unprecedented challenges due to increasing and diverse student profiles and levels of literacy, a decline in government funding, dynamics in market conditions resulting in a reduction in the value of endowments coming from alumni and other stakeholders, declining support from business and private sectors, increasing operational costs, growing regulatory demands (government, regulatory bodies, and private sectors) for continuous monitoring of performance, transparency and accountability (Hazelkorn, 2007).

Additionally, higher education institutions are being called upon to expand the number of students, increase the proportion of students in certain disciplines and address the pervasive and long-standing underrepresentation of minorities. In response, many institutions are under pressure to compete globally in order to attract more international students and highly qualified academic staff, adding more operational challenges.

Further, corporate-academic partnerships are increasing. However, to attract and sustain partnerships, corporations require institutions of higher education to demonstrate a commitment to the utilization and development of advanced technologies that are likely to support applied research outputs, and with potentials for knowledge transfer and commercialization.

Also, within the institutions of higher education, new technologies continue to have a significant impact on academic careers as research and teaching become more reliant on these technologies (Economist, 2008). Likewise emerging social technologies are transforming the way students interact with others and their learning environments. As learning technologies continue to penetrate all facets of higher education, a plethora of useful ‘data traces’ are generated. However, leveraging these data traces has many challenges, both at a technical and policy level. While rudimentary data analytics has always had a place in universities, this new more pervasive movement has the potential to reveal a vast array of currently unknown data that is likely to transform our current conceptions and practices of higher education.

While there is a growing appreciation for the need for ‘rich’ evidence-based data extracted from analytics for effective decision-making (Oblinger 2012), the area is still evolving. More work is required in the areas of institutional data warehousing, aggregation, and analysis. This paper to outline a process by which conceptual ideas concerning analytics can be realized through the design and implementation of a framework for Technology Enhanced Analytics (TEA) within the Higher Education Sector.

What is significant about our approach is the identification and inclusion of various stakeholders that traditionally haven't been considered within Higher Education analytics. It is our belief that the identification and distribution of appropriate analytics to these stakeholders is best approached through a central warehouse model, governed by key institutional representatives charged with the development and deployment of policies aimed at leveraging the benefits of Big Data institutional analytics.

2. BIG DATA AND ANALYTICS

Using data for making decisions is not new; business organizations have been storing and analyzing large volumes of data since the advent of data warehouse systems since the early 1990s. For instance business have employed business intelligence (BI) techniques to various data warehouse systems to discern insights on consumers' behaviours, detecting useful patterns and creating models that can explain present customers' behaviours and predict future trends. Web analytics (WA), an early approach to BI, focuses on analysis of webpage page visits to understand and improve how people use the Web. Over the years, business has grown beyond WA developing more sophisticated techniques to track and trace social actions, such as bookmarking to social sites, posting to twitter or blogs, and commenting on stories to predict and recommend Web pages of interest.

As the rate of growth in data volumes continues to escalate, business organizations continue to seek for ways to capture, store and analyze greater levels of human and machine-generated data. In 2012 the term Big Data emerged as an approach for dealing with increasing volumes and the variability of massive data generated by users and technology environments (e.g. open source software and loud architecture).

Current literature suggests that Big Data refers to data which is fundamentally too big and moves too fast, exceeding the processing capacity of conventional database systems (Manyika, et., al. 2010). Generally Big Data has come to be identified by three fundamental characteristics:

- Volume—large amount of information is often challenging to store, process, and transfer, analyses and present.
- Velocity—relating to increasing rate at which information flows within an organization— (e.g. organizations dealing with financial information have ability to deal with this).
- Variety referring to data in diverse format both structured and unstructured.

Due to its complexity, Big Data requires exceptional technologies to efficiently process large quantities of varied data within tolerable time elapses. Current areas of research on Big Data tend to focus on both technical and applied aspects. The technical aspects of Big Data include distributed computing, algorithm development, integrated systems, network and database architecture, and storage. Applied areas of research tend to emphasis ways to examine the implications and applications of Big Data in education, health care, government, business and social services. More specifically, the application of Big Data in higher education is concerned with approaches and techniques aimed at efficiently collecting, aggregating, analyzing, and interpreting vast amounts of information stored in institutional systems.

3. LEARNING ANALYTICS AND BIG DATA IN HIGHER EDUCATION

Long and Siemens (2011) indicated that Big Data presents the most dramatic framework in efficiently utilizing the vast array of data and ultimately shaping the future of higher education. The application of Big Data in higher education was also echoed by Wagner and Ice (2012), who noted that technological developments have certainly served as catalysts for the move toward the growth of analytics in higher education. In the context of higher education, Big Data connotes the interpretation of a wide range of administrative and operational data gathered processes aimed at assessing institutional performance and progress in order to predict future performance, and identify potential issues related to academic programming, research, teaching and learning (Hrabowski III, Suess & Fritz, 2011; Picciano, 2012).

As an emerging field within education, a number of scholars have contended that learning analytics with the Big Data framework is well positioned to address some of the key challenges currently facing higher education (see for example Siemen, 2011; Dawson, 2013). At this early stage much of the work on data analytics within higher education is coming from interdisciplinary research spanning the fields of

Educational Technology, Statistics, Mathematics, Computer science and Information Science. A core element of the current work is centred on data mining. Luan (2002) describes the features of data mining techniques as clustering and prediction. The clustering aspect of data mining offers comprehensive analysis while the predicting functions estimate the likelihood for a variety of outcomes (Romero & Ventura, 2010).

While educational data mining tends to focus on developing new tools for discovering patterns in data, Big Data (learning analytics) for instance focuses on applying tools and techniques to analyze large sets of data. Analytics also provides researchers with opportunities to carry out real-time analysis of activities. By performing retrospective analysis of student data, predictive models can be created to examine students at risk and provide appropriate intervention, and hence, enabling instructors to adapt their teaching or initiate tutoring, tailored assignments, and continuous assessment (EDUCAUSE, 2011; US Department of Education, 2012).

Big Data in higher education also covers database systems that store large quantities of longitudinally data on students and down to very specific transactions and activities on learning and teaching. When students interact with learning technologies, they leave behind data trails which can reveal their sentiments, social connections, intentions and goals. Researchers can use such data to examine patterns of student performance over time—from one semester to another or from one year to another.

The added-value of Big Data is the ability to identify useful data and turn it into useable information by identifying patterns and deviations from patterns. Schleicher of OECD, 2013 reported that: “Big Data is the foundation on which education can reinvent its business model and build the coalition of governments, businesses, and social entrepreneurs that can bring together the evidence, innovation and resources to make lifelong learning a reality for all. So the next educational superpower might be the one that can combine the hierarchy of institutions with the power of collaborative information flows and social networks.”

Further, Big Data analytics could be applied to examine student entry on a course assessment, discussion board entries, blog entries, or wiki activity could be recorded, generating thousands of transactions per student per course. This data would be collected in real or near real time as it is transacted and then analyzed to suggest courses of action. As Siemens (2011) indicated that “[learning] analytics are a foundational tool for informed change in education” and provide evidence on which to form understanding and make informed (rather than instinctive) decisions.

Big Data can also address the challenges associated with finding information at the right time when data is dispersed across several unlinked different data systems in institutions. By identifying ways of aggregating data across systems, Big Data can help improve decision-making capability. Though Big Data is an emergent research area in higher education, there are higher education institutions that have implemented tools to capture, process and use Big Data. For instance, Arizona State University is using predictive analytics to increase graduation rates. Purdue University developed the Signals project in 2007, which gathers information from student information system, course management systems, and course gradebooks to generate a risk level for students, and those designated as at-risk are targeted for outreach.

Further, University of Wollongong in Australia implemented Social Networks Adapting Pedagogical Practice (SNAPP), a tool designed to expand on the basic information gathered within learning management systems, which included how often and for how long students interact with posted material. SNAPP enable visual analytics to display how students interact with discussion forum posts, giving significance to the socio-constructivist activities of students.

4. DATA ANALYTICS AT THE UNIVERSITY OF OTAGO

University of Otago is a research intensive University, and the oldest in the Southern Hemisphere. The University has an extraordinary record of accomplishment in research leadership and teaching. Over the years, the University has served as a wellspring of research and creative endeavor, and in providing public service. Like many another institutions, the University has its share of challenges.

Currently an institutional collaborative project titled Technology Enhanced Analytics (UO-TEA) consisting of an interdisciplinary team is being established to explore the potential of data analytics to address a number of these challenges. Over the next year the group aims to explore the implications of Big Data within the institution and ultimately develop platforms for data collection, aggregation, and build a data

ISBN: 978-972-8939-99-1 © 2013

warehouse that aligns with the needs of the various stakeholders: students, instructors, policy, and researchers. To do this a four element framework has been developed (see figure 1).

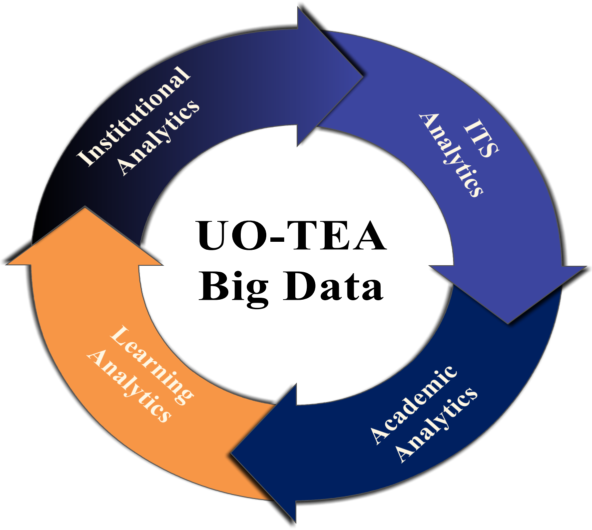

Figure 1. Figures Components of Big Data at Otago

4.1 Figures Components of Big Data at Otago

4.1.1 Institutional analytics

Institutional analytics refers to a variety of operational data that can be analyzed to help with effective decisions about making improvements at the institutional level. Institutional analytics include assessment policy analytics, instructional analytics, and structural analytics. Institutional analytics make use of reports, data warehouses and data dashboards that provide an institution with the capability to make timely data-driven decisions across all departments and divisions.

4.1.2 Information Technology Analytics

Information technology (IT) analytics covers usage and performance data which helps with monitoring required for developing or deploying technology, developing data standards, tools, processes, organizational synergies and policies. Information technology analytics aim at integrating data from a variety of systems—student information, learning management, and alumni systems, as well as systems managing learning experiences outside the classroom. Results of information technology analytics are used to develop rigorous data modeling and analysis to reveal the obstacles to student access and usability, and to evaluate any attempts at intervention. Freeman and Suess (2010) reported with analytics, IT systems can help by refining the associated business processes to collect critical data that might not have been collected institutionally, and by showing how data in separate systems can become very useful when captured and correlated.

4.1.3 Academic/Program Analytics

Academic analytics provides overall information about what is happening in a specific program and how to address performance challenges. Academic analytics combines large data sets with statistical techniques and predictive modelling to improve decision making. Academic analytics provide data that administrators can use to support the strategic decision-making process as well as a method for benchmarking in comparison to other institutions.

The goal of an academic analytics program is also to help those charged with strategic planning in a learning environment to measure, collect, interpret, report and share data in an effective manner so that operational activities related to academic programming and student strengths and weaknesses can be identified and appropriately rectified.

4.1.4 Learning Analytics

Learning analytics is concerned with the measurement, collection, and analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs (Siemens & Long, 2011). More broadly, learning analytics software and techniques are commonly used for improving processes and workflows, measuring academic and institutional data and generally improving organizational effectiveness (Jones, 2012). Although such usage is often referred to as learning analytics, it is more associated with 'academic analytics' (Goldstein and Katz, 2005). Learning analytics is undertaken more at the teaching and learning level of an institution and is largely concerned with improving learner success (Jones, 2012).

4.2 Data Analytical Framework at Otago

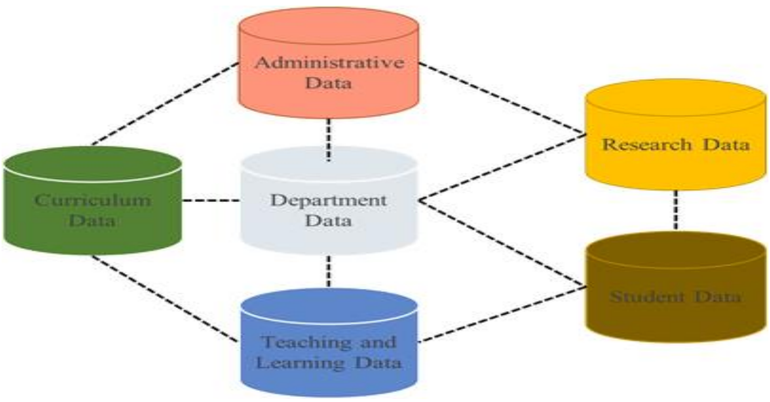

Figure 2. Big Data Analytical Framework at Otago

4.3 System Scenario

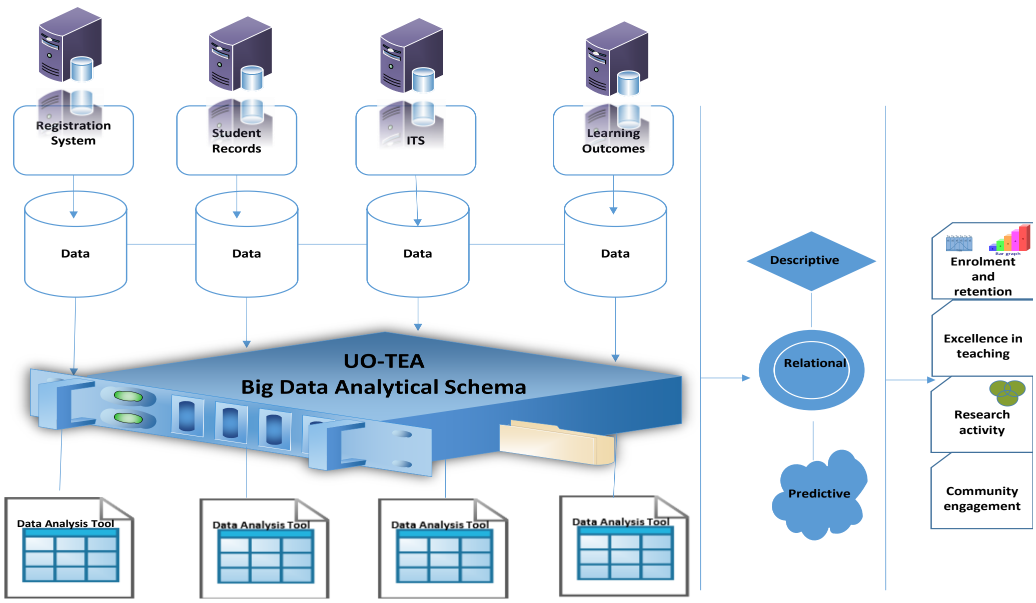

Figure 3. System Scenario

4.4 Simple Process Report Request Scenario

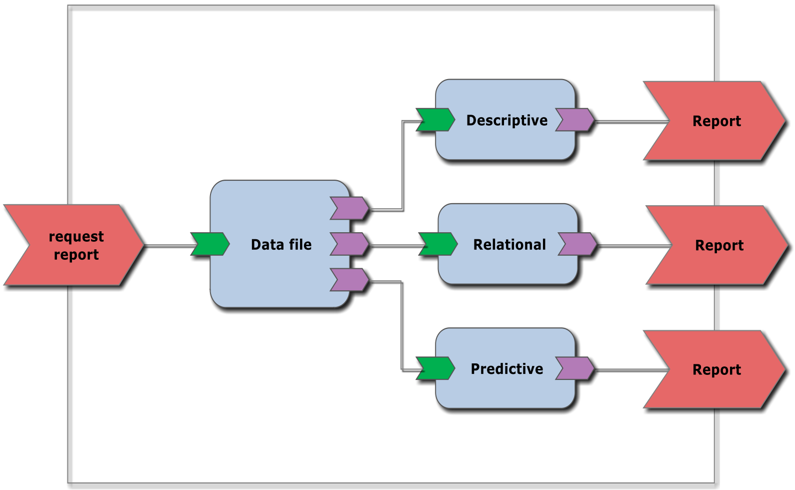

Figure 4. Simple Process Report Request Scenario

4.4.1 Possible Project Performance Outcomes

- Better understanding of institutional Big Data at University of Otago
- Better understanding of the requirements for effective data preparation for Big Data analytics
- A solid foundation for Big Data utilization
- Improved standardized and streamlined data processes
- Consistent ways to effectively leverage data analytics for improved accuracy, deeper knowledge and real time decision making
- Better data-driven decision making and practice
- Foundation for hypothesis testing, web experimenting, scenario modelling, simulation, sensibility and data mining

4.4.2 Possible Project Process Outcomes

- Better tools for collecting, processing, analysing and interpretation of data
- Better data system interoperability and system linking
- Enhanced data analytics and predictive modelling
- Better real-time rendering of analytics on students and instructors performances
- Reliable and comparable performance indicators and metrics within departments and divisions
- Better utilization of historical institutional data to make informed decisions
- Better ability to develop and utilize "what if" scenarios for exploring data to predict possible outcomes

5. CHALLENGES OF IMPLEMENTATION

We anticipate a number of challenges associated with the collecting and implementation of analytic techniques for analyzing Big Data in higher education. For instance, the costs associated with collecting, storing, and developing algorithms to mine data can be time consuming and complex. Furthermore, most of institutional data systems are not interoperable, so aggregating administrative data and classroom and online data can pose additional challenges. While combining data sets from across a variety of unconnected systems can be extremely difficult it offers better comprehensive insights that inevitable lead to improve capabilities of predictive modelling. Dringus (2012) suggested that one way of overcoming these problems, is to increase institutional transparency by clearly demonstrating the changes that analytics can help to achieve.

Big Data can be used to help carry out targeted decisions and faster decisions, for promotion purposes (marketing) or to protect our interests. Emerging evidence from research and practice communities suggests that learning analytics may enable learning experiences that are more personal, more convenient, and more engaging and may also have a direct positive impact on student retention. Analytics also has the potential to help learners and instructors recognize danger signs before threats to learning success materialize (Wagner & Ice, 2012). However, wide institutional acceptance of learning analytics requires a clear institutional strategy and the usability of analytics software packages. Further, as stated by Ali et al. (2013), perceived usefulness is one of the strongest drivers influencing users' intentions of adopting a software tool.

A report by the US Department of Education (2013) suggested that the successful implementation of Big Data in higher institution would depend on collaborative initiatives between various departments in a given institution. For instance, the involvement of information technology services departments in planning for data collection and use is deemed critical. This is consistent with views that the value of Big Data Analytics will be based on the ability to co-create governing structures and delivery of more progressive and better policies and strategies currently used (Schleicher, 2013). Wagner and Ice (2012) also pointed out that by increasing collaborative ventures on Big Data initiatives help all groups take ownership of the challenge involving student performance and persistence. Dringus (2012) suggested that the practice of learning analytics should be transparent and flexible to make it accessible to educators (Dringus, 2012; Dyckhoff et al., 2012).

In many instances, there is a divide between those who know how to extract data and what data is available, and those who know what data is required and how it would best be used. As Romero and Ventura (2010) note, analytics has traditionally been difficult for non-specialists to generate (and generate in meaningful context), to visualize in compelling ways, or to understand, limiting their observability and decreasing their impact (Macfadyen & Dawson, 2012).

The importance of communicating these ideas is also acknowledged by Macfadyen and Dawson (2012), who found analytics to have a negative or neutral impact on educational planning. They advocate delving into "the socio-technical sphere to ensure analytics data are presented to those involved in strategic positions in ways that have the power to motivate organizational adoption and cultural change."

Although the existence of an 'online learning environment' is often implied as necessary for the practice of analytics, most types of data are not specific to the web. Data can be generated from any interaction an instructor has with a student. It is the ability to obtain data in greater volumes and track students' activities with precision that has contributed to the development of Big Data as a research field in higher education. Becker (2013) believes that there are three interactive components to be studied when collecting data for analytics: location, population and timing. Location is defined by where and how students are accessing the learning space, while population refers to the characteristics of the group of learners participating in the learning space. Timing can be defined by any unit, from second or minute to semester or year.

Finally, Big Data raises the topic of the ethics of data collection in regard to quality of data, privacy, security and ownership. It also raises the question of an institutions responsibility for taking action based on the information available (Jones, 2012). Dringus (2012) suggests that bringing transparency to learning analytics as a practice could be used to help deter any potentially wrongful use of data. As the amount of data available for use is ever-increasing, the benefits will come from good learning management, reliable data warehousing and management, flexible and transparent data mining and extraction, and accurate and responsible reporting.

6. FUTURE DIRECTIONS

We are currently reviewing work on Big Data analytics in Higher Education and exploring data management and governance structures. This work will result to a detailed description of the current conceptual and theoretical underpinnings of Big Data analytics in higher education, as well as key performance indicators, metrics and methods for capturing, processing and visualizing data. We also intend to develop a set of diagnostic tools and an integrated technology enhanced data analytic framework and ultimately a Data warehouse for Big Data Analytics.

REFERENCES

Ali, L., Adasi, M., Gasevic, D., Jovanovic, J., and Hatala, M. (2013). Factors influencing beliefs for adoption of a learning analytics tool: An empirical Study. *Computers & Education*, 62, pp. 130–148.

Dringus, L. (2012). Learning Analytics Considered Harmful. *Journal of Asynchronous Learning Networks*, 16 (3).

Dyckhoff, A. L., Zielke, D., Bültmann, M., Chatti, M. A., and Schroeder, U. (2012). Design and Implementation of a Learning Analytics Toolkit for Teachers. *Educational Technology & Society*, 15 (3), pp. 58–76.

EDUCAUSE (2011) Learning Initiative, "7 Things You Should Know about First-Generation Learning Analytics," December 2011.Making the grade 2011. A study of the top 10 issues facing higher education institutions. Retrieved from http://www.deloitte.com/assets/DcomIreland/Local%20Assets/Documents/Public%20sector/IE_PS_making%20the%20grade_IRL_0411_WEB.pdf

Hrabowski III, F.A. & Suess, J. (2010). "Reclaiming the Lead: Higher Education's Future and Implications for Technology,"*EDUCAUSE Review*, vol. 45, no. 6 (November/December 2010). Retrieved from <http:// www.educause.edu/library/ERM1068>

Hazelkorn, E. (2007). The impact of league tables and ranking systems on higher education decision making. *Higher education management and policy*, 19 (2), pp.1-24

Hrabowski, III, F.A., Suess, J. & Fritz, J. (2011). Assessment and Analytics in Institutional Transformation. Assessment and Analytics in Institutional Transformation (EDUCAUSE Review). *EDUCAUSE Review*, vol. 46, no. 5 (September/October 2011). Retrieved August 8th 2013 from [http://www.educause.edu/ero/article/assessment-and-analytics-institutional-transformation]

Jones, S. (2012). Technology Review: the possibilities of learning analytics to improve learner-centered decision-making. *Community College Enterprise*, 18 (1), 89-92.

Luan, J. (2002). Data mining and its applications in higher education. In A. Serban & J. Luan (Eds.), *Knowledge management: Building a competitive advantage in higher education* (pp. 17-36). Josey-Bass: PA.

Oblinger, D.G. (2012). Let's talk analytics. *EDUCAUSE Review*, July/August, 10-13.

The future of higher education: *The Economist intelligence unit 2008 report. How technology will shape learning.* Retrieved on August 2 2013 from [http://graphics.eiu.com/upload/The%20Future%20of%20Universities.pdf]

Macfadyen, L. P., and Dawson, S. (2010). Mining LMS data to develop an ''early warning system" for educators: A proof of concept. *Computers & Education* 54, 588-599.

Macfadyen, L. P., and Dawson, S. (2012). Numbers Are Not Enough. Why e-Learning Analytics Failed to Inform an Institutional Strategic Plan. *Educational Technology & Society*, 15 (3), 149-163.

IBM What is Big Data? — Bringing Big Data to the enterprise. 01.ibm.com.

Manyika, J., M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers. (2011). *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute. Retrieved on August 8th 2013 from: http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation.

Picciano, A.G. (2012). The Evolution of Big Data and Learning Analytics in American Higher Education, 16 (3), pp.9-20. Journal of Asynchronous Learning Networks.

Romero, C. R. & Ventura.S (2010). "Educational Data Mining: A Review of the State of the Art." IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews 40 (6): 601–618.

Schleicher, A. (2013). Big Data and PISA. Retrieved on 4th August 2013 from http://oecdeducationtoday.blogspot.co.nz/2013/07/big-data-and-pisa.html?m=1

Siemens, G., and Long, P. (2011). Penetrating the Fog: Analytics in Learning and Education. EDUCAUSE Review, 46 (5), 30.

Siemens, G. (2011). How data and analytics can improve education, July 2011, Retrieved on August 8 from [http://radar.oreilly.com/2011/07/education-data-analytics-learning.html]

U.S. Department of Education, Office of Educational Technology, Enhancing Teaching and Learning through Educational Data Mining and Learning Analytics: *An Issue Brief, Washington*, D.C., 2012.

Wagner, E. & Ice, P. (2012). Data changes everything: delivering on the promise of learning analytics in higher education. *Educause Review* JULY/AU GUST 2012, pp. 33-42.