When Learning Analytics Meets E-Learning

Betul C. Czerkawski
University of Arizona- South Campus
bcozkan@email.arizona.edu

Abstract

While student data systems are nothing new and most educators have been dealing with student data for many years, learning analytics has emerged as a new concept to capture educational big data. Learning analytics is about better understanding of the learning and teaching process and interpreting student data to improve their success and learning experiences. This paper provides an overview to learning analytics in higher education and more specifically, in e-learning. It also explores some of the issues around learning analytics.

Learning Analytics

As the amount of educational data grows, it is becoming challenging for higher education institutions to organize and understand large complex data sets. More than any time in the history, learners are leaving digital traces in their choices when learning (UNESCO, 2012) and these traces or analytics provide invaluable opportunity for the learning institutions to increase their efficiency and effectiveness. "Learning analytics places a greater emphasis on the qualitative data that originate from learning behavior" (Becker, 2013, p. 63) while analyzing quantitative metrics. In the last few years, the Horizon Report included learning analytics (LA) as one of the emergent areas of research since 2011 and in the most recent edition it provided the following description for it:

Learning analytics research uses data analysis to inform decisions made on every tier of the education system, leveraging student data to deliver personalized learning, enable adaptive pedagogies and practices, and identify learning issues in time for them to be solved (Horizon Report, 2014, p.38).

In other words, learning analytics refer to educational 'big data'. Big data, uses a statistical analysis approach, and was initially developed for market researchers to understand consumer experience trends. In recent years, there has been a big push for learning institutions to interpret student experience using the data they produce, so instruction can be individualized for student needs and student performance can be predicted in future planning efforts (See Table 1 for current LA projects). In addition, retention and student support especially for at-risk students are some of the first places where learning analytics are applied.

Table 1: Higher Education Projects in Learning Analytics

Projects
Name: Gradecraft Institution: University of Michigan Website: https://gradecraft.com/

Purpose: Provide motivating assessments using digital badges
Name: STAR Institution: University of Hawaii Website: https://www.star.hawaii.edu:10012/studentinterface/ Purpose: Chart student's academic plan and send alerts when students veer off their path.
Name: Open Academic Analytics Initiative Institution: Marist College Website: http://tinyurl.com/nm7nf2w Purpose: Open source academic alert system that uses predictive modeling to increase student success
Name: Visualizing Collaborative Knowledge Work Institution: Ball State University Website: http://www.emergingmediainitiative.com/uncategorized/learning-analytics/ Purpose: Encourage continuous feedback, formative evaluation and meta-cognition among collaborators using human-computer interaction and interface design
Name: Grade Performance Status (GPS) Institution: University of Northern Arizona Website: http://nau.edu/University-College/GPS/Students/GPS-Grade/ Purpose: Generate feedback alerts for academic standing including attendance and academic performance.
Name: "Check my activity" Institution: University of Maryland, Baltimore County Website: http://www.educause.edu/ero/article/video-demo-umbcs-check-my-activity-tool-students Purpose: A self-service feedback tool integrated in Blackboard LMS for students so they can check their progress in online courses
Name: Signals Institution: Purdue University Website: http://www.itap.purdue.edu/studio/signals/ Purpose: Identify at-risk students using Blackboard Vista data mining techniques; predicting modeling for early detection and feedback
Name: Chico State Learning Analytics Research Project Institution: Chico State University Website: http://tinyurl.com/acz5f7o Purpose: Study relationship between academic achievement and LMS use.

Data mining techniques have existed in higher education for over a decade. Learner analytics is different from data mining, in that analytics support data-driven decision-making and provides information for educators, policy makers and administrators to improve learning process. Siemens (2013) adds that learning analytics is about sense-making and interpretation, whereas educational data mining is more about developing methods and models for existing data in educational settings so that bigger questions about learning could be answered. Moreover, as Swan (2012) suggests, previous data mining efforts focused on reporting or archiving student data and never emphasized teaching and learning. Data mining tries to organize and reduce educational data whereas data analytics views the entire data systems to better understand learner behaviors.

Learning analytics is not yet considered as an academic discipline with established methodological approaches but it is definitely a new research arena for scholars. Organizations such as the Society for Learning Analytics Research and the International Educational Data Mining Society, are trying to establish a research community on learning analytics (Siemens, 2013). Learning analytics uses techniques such as web analytics, social network analysis (SNA), predictive modeling, natural language processing (Crow, 2013) and cognitive modeling and has roots in artificial intelligence, statistical analysis, machine learning, and business intelligence (Siemens, 2013). Building on established techniques facilitates rapid advancement in the short term but learning analytics as a field of study still lacks a common terminology or coherent structure.

The purpose of this paper is to provide an overview of learning analytics, explore some of the issues and challenges around it and discuss the e-learning and learning analytics relationship. The author concludes with a brief discussion of major implications of learner analytics for e-learning.

Advantages and Concerns

With the maturation of web tracking tools, learning analytics enables education institutions to gather much needed data on students' learning experiences. This data could be used in areas such as personalized learning, adaptive technologies and tools, identification of learning problems, program measurement and evaluation, as well as improved learning and teaching experiences (Horizon Report, 2014). Among these benefits, the most striking one is the ability to create individualized environments for students, which may lead to 'flexible' educational frameworks of the kind that educators have been discussing for decades. Students do not learn at the same speed and level and their progression varies from student to student (Dietz-Uhler & Hurn, 2013).

Being able to customize instruction for individual needs and identify learning difficulties as they arise are very powerful ideas for flexible education. Finally, learning analytics could offer "increased accountability at all levels of education" (Dietz- Uhler & Hurn, 2013, p. 20).

While the benefits of learning analytics are exciting for most educators, there are also some concerns about who owns student data, how the data is used, how possible errors in data are handled and the validity of interpretations. The concern is that big data can give commercial businesses or large companies new abilities to manipulate user data and can result in unfair competitions (Bollier, 2010). In addition, violations of user privacy and freedom are major issues. In the context of e-learning, there are questions about leaning management systems' effectiveness at capturing learning experiences to produce useful data. For instance, Dringus (2012) posits that learning management systems (LMSs) do not present student data in a meaningful way to the instructor. Instructors do not have sufficient tools in the system to manage the data and draw conclusions. Therefore, while it is promising, learning analytics results in poor decision-making about student learning.

Kay, Kom and Oppenheim (2012) suggest that learning institutions 1) have to be clear and transparent about what data they are collecting and the purpose of data collection, 2) give learners a chance to opt out if they wish to and 3) provide a mechanism of complaint and take-down if unforeseen consequences arise from handling data. Slade and Prinsloo (2013) add to this that "data collected through learning analytics should have an agreed-upon lifespan and

expiration date, as well as mechanisms for students to request data deletion under agreed-upon criteria" (p.13). Learning analytics deal with student data. It is particularly important to ensure safety and clarity as well as consideration to all ethical issues around collecting and maintaining individual data.

Learning Analytics Techniques

Various authors (Siemens, 2012; Crow, 2013) describe different techniques that could be utilized in learning analytics. For instance, rather than merely offering broadly defined techniques, Bienkowski, Feng and Means (2012) offer the following specific techniques that could be used with online students: user knowledge modeling, user behavior modeling, user experience modeling, user profiling, domain modeling, learning component analysis and instructional principle analysis, trend analysis and adaptation and personalization. The author suggests the following techniques, especially for those who are new to learning analytics field: pattern analysis, domain analysis, social network analysis and trend analytics.
Learning analytics predominately uses "pattern analysis" (i.e. discovering patterns in existing data), since this is a well established technique in data mining literature. The pattern analysis technique helps researchers predict future student behaviors. In addition to prediction, pattern analysis also allows for "model building" for improving educational success. Model building could be done for groups of students, but the current tendency is to use individual 'user or learner models' and explain who the learner is, what he knows, his aspirations, motivations and satisfaction so his future online behaviors can be predicted to improve his learning experience.

In addition to pattern analysis, "domain analysis" focuses on content analysis: what topic needs to be presented, in what order and to what degree. As Bienkowski, Feng and Means (2012) suggest, domain analysis is a much more experimental technique compared to pattern analysis. However, it is these pedagogical experimentations such as applying learning theories and principles that create effective instruction.

Finally, learning analytics uses "social network analysis" and "trend analysis" extensively. In social network analysis learners' online interactions are recorded to gauge their participation and engagement level. UNESCO's 2012 report on learning analytics reports: "connections learners forge with each other and the resulting group structures correlate with more or less effective learning" (p. 6). Trend analysis, on the other hand, looks at the changes occurring in a period of time to identify areas of user interest. For trend analysis to be successful longitudinal learner records are gathered and analyzed for common occurrences.

Because of the difficulty of data interpretation, attention is currently being paid to "visual" data analytics where researchers "view" the data to make sense of it. However, visualization does not only facilitate interpretation but also brings a human touch to data analysis, which is lacking in data mining research. The strength of visual data analytics lies in its capability "to expose patterns, trends and exceptions in very large heterogonous and dynamic datasets collected from complex systems" (Bienkowski, Feng and Means, 2012, p.15).

Learning Analytics Tools

As shown in Table 1, most learning analytics tools are developed in-house by individual universities to meet the specific needs of their students. The majority of these projects could be

considered as 'academic analytics', which is a slightly different concept than learning analytics. Academic analytics are used at the institutional or regional level and benefit administrators, policy makers and government organizations. Learning analytics, on the other hand, are about course or departmental level analytics and benefit learners and faculty members (Siemens & Long, 2013). Regardless of the scope and use of findings, both academic and learning analytics use similar techniques and tools. Following is a list of free tools that could be used by online educators interested in analytics. They are summarized from simple LA tools to complex ones.

Google Analytics (http://www.google.com/analytics): While it is not designed with LA in mind, this simple tool (only basic version is free) could be beneficial to beginner scholars as it provides a good example of web analytics. Offered by Google, Google Analytics reports data about any website's traffic and the source of usage. It also provides statistics about the site's users, their social network preferences and use of search engines. One other advantage of Google Analytics is its availability for mobile devices. However, comparing to tools designed specifically for learning analytics (see SNAPP and Netlytics below) Google Analytics does not provide the full picture of student behaviors that online instructors seek.

SNAPP (http://www.snappvis.org): SNAPP is acronym for Social Networks Adapting Pedagogical Practice and is developed by Office for Learning and teaching in Australia. It is used for conducting real-time social network analysis using data visualization techniques. All LMSs, commercial or open-source, could integrate SNAPP to capture student interactions in online courses and especially in discussion forums. Online instructors could use this tool to diagnose non-participating students, gauge the level of interaction among students and provide timely feedback if group structures do not work. SNAPP is also a good tool for identifying behavioral issues such as isolation.

Netlytic (https://netlytic.org/home): Netlytics is another social network analysis tool but its scope goes beyond of what SNAPP can do at the course level. Netlytics is designed to analyze textual data as well as students' social media interaction in programs such as Twitter, YouTube, blogs, online forums and chats. Because this tool is cloud-based, online students and instructors do not have to install or download any software but use their open ID to log onto the site. Similar to SNAPP, this tool can identify level of student engagement but also is capable of gauging group relationship and interactions and inform the users if an 'online community' is present.

R Project (http://www.r-project.org/): R is an open source programming language that is widely used in academic programs. While this is a high level tool that requires expertise in programming, its capability in analyzing big chunks of data, statistical computing and graphical analysis makes it a very powerful tool for learning analytics.

Weka (http://www.cs.waikato.ac.nz/ml/weka): Developed at the University of Waikato, New Zealand in 1993, Weka is probably one of the oldest tools available. Originally designed for data mining, Weka uses machine learning. It analyzes both textual and multimedia data and produces visual reports in charts and graphics.

All of these tools could be integrated within a learning management system but there are also products that are specific to certain LMSs. For instance, almost all LMSs today come with the warehouse technologies needed for conducting learning analytics (e.g., Moodle SmartKlass, Canvas Analytics, D2L insights, Blackboard Analytics, Sakai Apereo Learning Analytics

Initiative).

E-Learning and Learning Analytics Considerations for Higher Education

Web tracking mechanisms that provide data for learning, work extremely well with online students. As a matter of fact, to use learning analytics transactions should be electronic or online rather than manual (Picciano, 2012). Siemens (2013) goes further and suggests that it was e-learning that contributed to the development of learning analytics. Siemens also states that it is students' online behaviors such as using an LMS, social software as well as their clicks, navigation patterns, time on task, social networks, information flow, and concept development through discussions that allow learning analytics to record and analyze educational data (p. 1384).

All learning management systems include reporting metrics such as access time, time spent on certain pages, number of posts, etc. but learning analytics refer to a higher level of sophistication in terms of recording students' online behaviors. Currently, most LMS companies and some start-up companies are offering 'learning analytics' packages that go with the LMS. Various dashboard software and data warehouses compile key learner metrics and aggregates data for instructors and administrators (West, 2012).

Recently, there are also some discussions about the relationship between learning management systems and learning analytics. Many students do not rely on LMSs for their learning and use free and open source programs available on the internet. In a study conducted by Pardo and Cloos (2011), the researchers developed virtual machines to record student behaviors outside of the LMS. The student-generated data was then sent to a central server where analysts interpreted it. Considering that LMSs are not interoperable, more studies are expected to be conducted in regard to virtual machines.

Today, e-learning courses are capable of analyzing a wide range of student behaviors. For instance, online instructors can review how much time students devote to reading teaching materials, what resources they use, where they get information and how quickly they master online content (West, 2012). Similar to business intelligence applications, dashboards are appearing in e-learning courses to provide learner analytics functionalities. Some of the advanced functionalities present data from other universities and enable complex reporting procedures that study relationships between learning behaviors (UNESCO, 2012). In addition to course instructors, students can also use some of the analytics tools, which allow them to compare their performance to others and help them gauge their own engagement level.

While tools and techniques have been developed quickly, not all online instructors are familiar with or experienced in learner analytics. As Swan (2012) suggests, most online instructors should inform themselves about learning analytics and understand their potential to improve student experiences since most of the developments in learning analytics show an upward development of the field. It appears that vendors and commercial companies are more active in providing training and webinars to online instructors than higher education administrators and institutional researchers.

For learning analytics to work in e-learning there are a few issues that need to be considered. The first issue is about the limitations of learning analytics software. Learning online in today's digital world is not exclusively for those who are taking fully online courses. All higher

education students use online platforms and information networks regardless of the mode of instructional delivery in their courses. The distributed nature of online student data (i.e. data interoperability) or erroneous or incomplete datasets (Reyes, 2015) makes data quality and collection a major issue for learning analytics. Therefore, e-learning instructors should understand that the learning analytics software they use may not capture all students' learning experiences. For instance, although LMSs are a major source of big data, mobile devices are used more and more by the students. In online education, learning analytics should be viewed beyond LMSs. Additionally, more research needs to be conducted to develop analytical tools that gauge student behaviors. Existing tools and learning management systems do not have sufficient capabilities.

Another challenge with learning analytics software is that before conducting any analytics, the best predictors to determine students' social, behavioral and linguistic responses should be identified. In other words, e-learning instructors should construct very clearly-defined measures and predictors of behaviors to be used, before performing any analysis via learning analytics. When analyzing learning analytics data, researchers should always remember that central to learning analytics are human and social processes, therefore they should not get bogged down with analytical methods and techniques. Online student data needs human interpretation. In terms of data analysis and interpretation, more standardization is also needed.

Data analytics is a very costly and time-intensive, requiring effort that may well be beyond and online instructors' workload. To ease up instructors' load, more case studies or examples should be available to online instructors so they can correctly analyze, interpret and draw conclusions from student data. In addition, classroom instructors should have appropriate tools available to collect student data.

Finally, learner analytics necessitates collaboration between classroom instructors, administrators, IT staff and institutional researchers. Currently, there are few trained individuals in higher education that use learning analytics, so this collaboration is mandatory. Such collaborations will also promote producing meaningful results and will help with synchronizing various efforts occurring in college campuses. In the long run, these collaborative efforts will result in developing a culture of using data to improve learner success as well.

Implications for E-Learning

Learning analytics uses intelligent online data and promises "to transform educational research into a data-driven science, and educational institutions into organizations that make evidence-based decisions (UNESCO, 2012, p.12). While many important issues exist in the full adoption of this idea there are a few important implications for e-learning. The first is that intelligent data means development of Semantic Web or Web 3.0 tools, which may revolutionize online learning and create more open and interactive learning experiences. Today many commercial companies such as Amazon and Netflix keep track of users' online activities and offer them personalized recommendations. Educational equivalences of such experiences are being developed at various universities around the country and it is expected that similar ideas will be prevalent in e-learning in a few years.

Second, with the advance of learning analytics online instructors will develop skills as analysts so they can collect and interpret student data and offer solutions. In a way, many instructors already do this but with learning analytics new skills in data mining will be necessary. Along the

same lines, more user-friendly tools will have to be introduced and tested by e-learning faculty.

Finally, within the learning analysis framework, assessment of learning shifts from final outcome assessment to process assessment. Becker (2013) suggests that this "in-process assessment draws its data from the daily learning activity of students within their social and informational networks" (p. 64) and it is the heart of educational assessment. As the data analytics field grows, "embedded assessment" techniques that were much emphasized in the U.S. Department of Education's National Education Technology Plan (NETP) (U.S. Department of Education, 2010) as well as formative assessments will be more prominent in higher education. According to Bienkowski, Feng and Means (2012) along with embedded assessment, "interconnected feedback" systems will shape online education in the future. The interconnected feedback is about gathering student data online, interpreting them and share the final data-driven decisions among various stakeholders of learning institutions.

When it comes to learning analytics, commercial companies have more developed tool sets than most learning institutions. The fast development cycles that companies use also help them develop advanced and functional tools. Perhaps this would also work for higher education e-learning endeavors but it is almost impossible in the learning analytics field to avoid collaborations with commercial data analytics companies. In the near future we can expect that institutional research offices, establish strong ties with the commercial companies to utilize some of the knowledge base.

Conclusions

In conclusion, it can be said that learning analytics empowers learners to understand the wealth of data related to learning (Clow, 2013). The idea of learning analytics "provides a new model for college and university leaders to improve teaching, learning, organizational efficiency, and decision making and as a consequence, serve as a foundation for systemic change" (Siemens & Long, 2011, p.32). Clow warns us, however, that the process is not simple or straightforward. The data analytics is not only about student data. In order to achieve change in learning institutions a broader and contextualized understanding should be developed along with attention to all ethical issues around student data. It is important to recognize that learning analytics hosts many opportunities and challenges for the society at large and our efforts should focus on converting the possibilities to a general public good (UNESCO, 2012).

References

Becker, B. (2013) Learning analytics: Insights into the natural learning behavior of our students. *Behavioral & Social Sciences Librarian, 32*(1), 63-67, doi:10.1080/01639269.2013.751804

Bienkowski, M., Feng, M., & Means, B. (2012, October). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. U.S. Department of Education. Office of Educational Technology. Retrieved from http://www.cra.org/ccc/files /docs/learning-analytics-ed.pdf

Bollier, D. (2010). The promise and peril of big data. The Aspen Institute, CO, USA, 2010. Retrieved from http://www.aspeninstitute.org/sites/default/files/content/docs/pubs

/The_Promise_and_Peril_of_Big_Data.pdf.

Clow, D. (2013) An overview of learning analytics. *Teaching in Higher Education, 18*(6), 683-695, doi: 10.1080/13562517.2013.827653.

Dietz-Uhler, B., & Hurn, J. E. (2013). Using learning analytics to predict (and improve) student success: A faculty perspective. *Journal of Interactive Online Learning. 12*(1). 17-26.

Dringus, L.P. (2012). Learning analytics considered harmful. *Journal of Asynchronous Learning Environments. 16*(3). 87-100.

Horizon Report (2014). 2014 Higher education edition. NMC. Accessed from http://www.nmc.org/pdf/2014-nmc-horizon-report-he-EN.pdf.

Kay, D., Kom, N., & Oppenheim. C. (2012). Legal, risk and ethical aspects of analytics in higher education. *JISC CETIS Analytics Series. 1*(6) http://publications.cetis.ac.uk/wp-content/uploads/2012/11/Legal-Risk-and-Ethical-Aspects-of-Analytics-in-Higher-Education-Vol1-No6.pdf.

Pardo, A., & Kloos, C. D. (2011). Stepping out of the box: Toward analytics outside the learning management system. In Proceedings of the *1st International Conference on Learning Analytics and Knowledge*. 163-167. doi:10.1145/2090116.2090142

Picciano, A. G. (2012). The evolution of big data and learning analytics in American higher education. *Journal of Asynchronous Learning Networks. 16*(3). 9-20.

Reyes, J. A. (2015, March/April). The skinny on big data in education. Learning analytics simplified. *TechTrends, 59*(2). 75-79.

Siemens, G. & Long, P. (2011, September/October). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review, 46*(5). 31-40.

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist, 57*(10), 1380–1400.

Slade, S. & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, *57*(10) 1509–1528. http://oro.open.ac.uk/36594/2/ECE12B6B.pdf.

Swan, K. (2012, June). Introduction to the special issue on learning analytics. *Journal of Asynchronous Learning Networks, 16*(3), 5-7.

UNESCO Institute for Information Technologies in Education. (2012, November). Learning analytics. Policy Brief. Retrieved from http://iite.unesco.org/files/policy_briefs/pdf/en/learning_analytics.pdf.

U. S. Department of Education. (2010). Transforming American education: Learning powered by technology. National Educational Technology Plan. Retrieved from http://www.ed.gov/sites/default/files/netp2010-execsumm.pdf.

West, D. M. (2012, September). Big data for education: Data mining, data analytics, and Web dashboards. Brookings Institution. Governance Studies. Retrieved from http://www.brookings.edu/~/media/research/files/papers/2012/9 /04%20education%20technology%20west/04%20education%20technology%20west.pdf.

Online Journal of Distance Learning Administration, Volume XVIII, Number 2, Spring 2015
University of West Georgia, Distance Education Center
Back to the Online Journal of Distance Learning Administration Contents