Learning Analytics Methods, Benefits, and Challenges in Higher Education: A Systematic Literature Review

John T. Avella, Mansureh Kebritchi, Sandra G. Nunn, Therese Kanai
University of Phoenix

Abstract

Higher education for the 21st century continues to promote discoveries in the field through learning analytics (LA). The problem is that the rapid embrace of of LA diverts educators' attention from clearly identifying requirements and implications of using LA in higher education. LA is a promising emerging field, yet higher education stakeholders need to become further familiar with issues related to the use of LA in higher education. Few studies have synthesized previous studies to provide an overview of LA issues in higher education. To address the problem, a systemic literature review was conducted to provide an overview of methods, benefits, and challenges of using LA in higher education. The literature review revealed that LA uses various methods including visual data analysis techniques, social network analysis, semantic, and educational data mining including prediction, clustering, relationship mining, discovery with models, and separation of data for human judgment to analyze data. The benefits include targeted course offerings, curriculum development, student learning outcomes, behavior and process, personalized learning, improved instructor performance, post-educational employment opportunities, and enhanced research in the field of education. Challenges include issues related to data tracking, collection, evaluation, analysis; lack of connection to learning sciences; optimizing learning environments, and ethical and privacy issues. Such a comprehensive overview provides an integrative report for faculty, course developers, and administrators about methods, benefits, and challenges of LA so that they may apply LA more effectively to improve teaching and learning in higher education.

Introduction

The advancement of technology has provided the opportunity to track and store students' learning activities as big data sets within online environments. Big data refers to the capability of storing large quantities of data over an extended period and down to particular transactions (Picciano, 2012). Users can take big data from different sources to include learning management systems (e.g., Blackboard), open

source platforms (e.g., Moodle), open social platforms (e.g., LinkedIn), and different web tools such Meerkat-Ed and Snapp (Reyes, 2015). Similar to decision making driven by data, analytics refers to the scientific process that examines data to formulate conclusions and to present paths to make decisions (Picciano, 2012). According to Brown (2012), the process of systematically collecting and analyzing large data sets from online sources for the purpose of improving learning processes is called learning analytics (LA). LA is an emerging field in education. Experts in online learning in American higher education predict that within the next few years learning analytics will be widely used in online education to identify students' pattern of behaviors and to improve students' learning and retention rates.

Learning analytics, educational data mining, and academic analytics are closely related concepts (Bienkowski, Feng, & Means, 2012; Elias, 2011). Educational data mining focuses on developing and implementing methods with a goal of promoting discoveries from data in educational settings. It examines patterns in a large data set related to students' actions. The methods may be utilized to form a better understanding of the educational settings and learners. Hung, Hsu, and Rice (2012) defined data mining as data analysis techniques which when applied extract hidden knowledge consisting of tasks consisting of *pattern discovery* as well as *predictive modeling*. Romero and Ventura (2010) provided a definition of educational data mining that uses data mining algorithms with the objective of solving educational issues. Academic analytics refers to an application of the principles and tools of business intelligence to academia with the goal of improving educational institutions' decision-making and performance (Campbell, De Blois, & Oblinger, 2007). Academic analytics combines "large data sets, statistical techniques, and predictive modeling" (Campbell et al., 2007, p. 42).

Learning analytics uses predictive models that provide actionable information. It is a multi-disciplinary approach based on data processing, technology-learning enhancement, educational data mining, and visualization (Scheffel, Drachsler, Stoyanov, & Specht, 2014). The purpose of LA is to tailor educational opportunities to the individual learner's need and ability through actions such as intervening with students at risk or providing feedback and instructional content. Conversely, educational data mining tries to generate systematic and automated responses to learners. While LA focuses on the application of known methods and models to address issues affecting student learning and the organizational learning system, educational data mining focuses on the development of new computational data analysis methods (Bienkowski et al., 2012).

There has been some criticism that higher education managers and the economic framing of education drive the process of big data mining (Clow, 2013); however, empirical studies indicated that LA can be useful for improving education. LA increases awareness of learners and educators in their current situations that can help them make constructive decisions and more effectively perform their tasks (Scheffel et al., 2014). One of the main applications of learning analytics is tracking and predicting learners' performance as well as identifying potential problematic issues and students at risk (EDUCAUSE, 2010; Johnson, Smith, Willis, Levine, & Haywood, 2011). Some universities have already used LA in various courses to improve learning. For example, Purdue University used predictive modeling based on data collected from the course management system to identify students at risk and provide intervention. The University of Alabama improved student retention by forming a predictive model for students at risk based on the large data set of learners' demographics. In another case, Northern Arizona University connected resource use, risk level, and students' achievement by forming a predicting model to identify which students would benefit from which resource (Campbell et al., 2007). These are some examples of pioneer higher education institutions that applied LA.

Although there have been studies related to using LA in higher education institutions within the last several years, LA is still an emerging field of education. Higher education stakeholders including leaders, administrators, instructors, and course developers need to become familiar with LA methods and application in higher education (Scheffel et al., 2014). The problem is few studies have synthesized the

previously conducted studies or provided a combined overview of issues concerning the use of LA in higher education. To address this literature gap and enhance application of LA in higher education, this study conducted a literature review. It provides an overview of methods, benefits, and challenges of using LA in higher education institutions for administrators, instructors, and course developers who are not expert in LA and need to develop a basic understanding about LA. As the use of LA is becoming increasingly popular and urgent in higher education, providing such an overview is critical to enhancing higher education stakeholders' understanding about LA.

Method

To address the research problem, researchers conducted a literature review using the procedure suggested by Cooper (1988) for synthesizing the literature. This systematic procedure helped to (a) formulate the problem, (b) collect data, (c) evaluate the appropriateness of the data, (d) analyze and interpret relevant data, and (e) organize and present the results. Then results were compared with current issues in a large higher education institution.

Formulating the problem. The problem is that embracing LA in evaluating data in higher education diverts educators' attention from clearly identifying methods, benefits, and challenges of using LA in higher education. These three key components need further clarification for higher education stakeholders to help them effectively apply learning analytics in higher education. Educators have to go through the daunting task of sifting through the literature to become familiar with LA methods, benefits, and challenges. To help solve the problem, the following questions guided this review:

1. What are the methods for conducting learning analytics in education?
2. What are the benefits of using learning analytics in education?
3. What are the challenges of using learning analytics in education?

Further identifying and describing LA methods, benefits, and challenges can better help educators in higher education to incorporate LA to improve students' learning.

Data collection. The purpose of data collection was to find empirical studies including quantitative, qualitative, mixed methods, and literature reviews published in peer-reviewed journals since 2000 to identify methods, challenges, and benefits of LA in higher education. The keywords that were used included *learning analytics and methods*, *learning analytics and benefits*, and *learning analytics and challenges*. Other keywords included *data mining and education*, *learning analytics and education*, and *learning analytics*. The databases used for literature research included Google Scholar, Educational Resources Information Center (ERIC), ProQuest, and EBSCO HOST.

Data evaluation and analysis. Based on the described procedure, 112 articles were found. Of these, 10 focused on issues related to learning analytics methods, 16 on benefits, and 18 focused on challenges. The remaining articles were excluded from this review because they could not be used to address the main three questions of the study. Only articles that were directly related to LA methods, benefits, and challenges and helped answer the three research questions were included in this review. The method described by Cooper (1988) was appropriate to guide a systematic review of the literature. The researchers exhausted the literature using the above-described procedure, keywords, and databases. Further, researchers limited the search of the literature to the specified keywords and databases. Therefore, this literature may not include sources not available via the searched criteria and databases. Table 1 provides the citations of sources included in the results section.

Table 1 *Sources Found Corresponding to the Research Questions*

Focus	Sources
Learning Analytics Methods and Analyses	Baker (2010)
	Baker and Yacef (2009)
	Bienkowski, Feng, and Means (2012)
	Campbell and Oblinger (2007)
	Clow (2012)
	Clow (2013)
	Dawson and Siemens (2014)
	Johnson, Levine, Smith, and Stone (2010)
	Reyes (2015)
	Romero and Ventura (2010)
Learning Analytics Benefits	Althubaiti and Alkhazim (2014)
	AlShammari, Aldhafiri, and Al-Shammari (2013)
	Arnold and Pistilli (2012
	Armayor and Leonard (2010)
	Bhardwaj and Pal (2011)
	DiCerbo (2014)
	Dietz-Uhler and Hurn (2013)
	Grummon (2009)
	Hsinchun, Chiang, and Storey (2012)
	Hung and Zhang (2012)
	Jantawan and Tsai (2013)
	Kostoglou, Vassilakopoulos, and Koilias (2013)
	Mardikyan and Badur (2011)
	Picciano (2012)
	Sharda, Adomako Asamoah, and Ponna (2013)
	Xu and Recker (2012)
Learning Analytics Challenges	
	Bottles, Begoli, and Worley (2014)
	Brown (2012)
	Buckingham Shum and Ferguson (2012)
	Dyckhoff, Zielke, Bültmann, Chatti, and Schroeder (2012)
	Johnson, Smith, Willis, Levine, and Haywood (2011)
	Kay, Korn, and Oppenheimer (2012)
	Ferguson (2012)
	Fournier, Kop, and Sitlia (2011)
	Lias and Elias (2011)
	McNeely and Hahm (2014)
	Pea (2014)
	Picciano (2014)
	Schroeder (2012)

Sclater (2014a)
Sclater (2014b)
Slade and Prinsloo (2013)
Vahdat et al. (2015)
West (2012)

Results

Based upon the literature review, the results obtained to answer the three research questions are provided in this section.

Learning Analytics Methods in Education

The following methods and analysis approaches inform faculty, educators, and administrators in higher education who are not experts in LA about the available methods reported in the literature. Such an overview provides an integrative report for educators and saves them from the daunting task of a literature search to become familiar with different LA methods.

Learning analytics process. With the current advent of both blended and online learning opportunities, big data and learning analytics are predicted to play a significant role in education in future years. When discussing learning analytics methods in education, it is important to provide a background regarding the flow of analytical information. The flow of analytical information can be traced from the students to the stakeholders within the framework of a hierarchy. When provided the opportunity to offer input and make recommendations, stakeholders can help enrich the learning experiences of students (Reyes, 2015). Researchers play a role as they validate and report their research results to inform stakeholders of best practices. Further, learning analytics also provides insight to instructors and students in the educational setting.

To streamline the flow of information and provide a structured process for collecting and analyzing data in learning analytics, researchers suggested a macro-level process for conducting learning analytics in educational settings. Campbell and Oblinger (2007) proposed five stages of capturing data, reporting the data pattern and trends, predicting a model based on the data by using statistical regression, acting by using an intervention based on the model to improve learning, and refining the developed model. Similarly, a learning analytics cycle was suggested by Clow (2012, 2013) in which researchers collect data from the learners, process the data into metrics, and use the results to perform an intervention that affects the students. The cycle continues as researchers collect additional data from the students for the next cycle of learning analytics.

Learning analytics analysis. Learning analytics focuses on data related to learners' interactions with course content, other students, and instructors. LA integrates and uses analysis techniques related to data mining, data visualization, machine learning, learning sciences, psychology, social network analysis, semantics, artificial intelligence, e-learning, and social aspects (Bienkowski et al., 2012; Dawson & Siemens, 2014). Social network analysis includes analysis of relationships between learners as well as between learners and instructors to identify disconnected students or influencers. Social analysis refers to the analysis of metadata to determine learners' types of engagement within educational settings (Bienkowski et al., 2012).

Data visualization tools and techniques. Visual data analysis includes highly advanced computational methods and graphics to expose patterns and trends in large, complex datasets (Johnson, Levine, Smith, & Stone, 2010). One of the standard techniques is visual interactive principal components

analysis; it can be used to reduce many variables into a few by finding elements within datasets. Table 2 shows websites that offer tools for data visualization. Gapminder uses the visual interactive approach to help analyze datasets. IBM Many Eyes has tools such as map-based clouds, charts, and graphs to create a visualization. FlowingData allows users to upload their data and create visualizations. A variety of additional visual analysis websites and tools are gathered by the National Visualization and Analytics Center and available at the Visualization Community website (Bienkowski et al., 2012).

Table 2 *Data Visualization Websites and Tools*

Website and Application	URL
Gapminder: reduces datasets into few	http://www.gapminder.org
IBM Many Eyes: creates data visualization	http://www.boostlabs.com/ibms-many-eyes-online-data-visualization-tool/
FlowingData: uploads the data and creates visualization	https://flowingdata.com/
Visualization Community: includes data visualization tools and websites	http://vacommunity.org/

Additionally, learning analytics uses educational data mining methods to analyze large datasets. Within educational data mining, researchers currently use a variety of popular methods. Classified into five categories, these methods consist of *prediction*, *clustering*, *relationship mining*, *discovery with models*, and *separation of data for use in the process of human judgment* (Baker, 2010; Baker & Yacef, 2009; Romero & Ventura, 2010). The final two categories are of significance within the field of education. This section discusses each of the five categories in detail below.

Predication. *Predication* involves developing a model that uses both a predicted variable and predicator variables. A predicted variable represents a particular component of the data, whereas predicator variables consist of a combination of other data elements. Researchers classify predication into three categories known as classification, regression, and density estimation. Baker (2010) described the three categories as classification methods with the use of decision trees, logistic regression, and support vector machine regression. Regression centers around a continuous variable as the predicted variable. Further, it uses linear regression, neural networks, and support vector machine regression. For density estimation, a probability density function is the predicted variable and the use of kernel functions.

Clustering. *Clustering* entails the discovery of a set of data points that form a logical group together. Therefore, observation reveals the resultant formation of some clusters from the full dataset. The use of clustering becomes most valuable when the categories within a group are unknown. How appropriate the set of clusters is may be evaluated by how well the set of clusters fits the data. Baker (2010) asserted that the goal of clustering involved the discovery of data points that formed a natural group together as well as the full dataset. By dividing a collection of data into logical clusters, researchers can assess how cluster sets explain the meaning of the data.

Relationship mining. The method of *relationship mining* focuses on the goal of discovering relationships between variables in a set comprised of a large number of variables. Forms of relationship mining may include learning which variables are related to a single variable or discovering what is the strongest relationship between two variables. Two criteria are necessary for relationship mining: statistical significance and interestingness (Baker, 2010).

Discovery with models. In the next method known as *discovery with models*, the goal is to develop a model using one of the following methods: predication, clustering, or knowledge engineering. Knowledge engineering uses human reasoning for model development. When using the discovery with models method, a prediction model influences a model's generalization across different contexts (Baker, 2010).

Separation of data for use in the process of human judgment. Researchers classify the *separation of data for use in the process of human judgment* method as a visualization method, in which educational data have a particular structure and meaning rooted within that structure. This method possesses two distinct goals identification and classification. Baker (2010) cited the importance of distilling data for identification when the display of data permits easy identification of well-known patterns which may be difficult to express formally. The learning curve represents an example of this concept. For example, the x-axis represents opportunities to practice a particular skill while the y-axis represents performance. This graphical representation can display Performance as the percentage correct or the amount of time that it takes to respond.

Learning Analytics Benefits in Education

Examination of the literature reveals how the use of big data is beneficial for higher education and includes various aspects from learning analytics that closely examine the educational process to improve learning. Another benefit includes the use of academic analytics that make alterations as a result of the application of algorithms to various points of data to improve learning. Through careful analysis of big data, researchers can determine useful information that can benefit educational institutions, students, instructors, and researchers in various ways. These stakeholder benefits include targeted course offerings, curriculum development, student learning outcomes and behavior, personalized learning, improved instructor performance, post-educational employment opportunities, and improved research in the field of education.

Identifying target courses. An initial benefit that evolves from using big data analysis in education is the ability of educational institutions to identify targeted courses that more closely align with student needs and preferences for their program of study. By examining trends in student enrollment and interests in various disciplines, institutions can focus educational and teaching resources in programs that maximize student enrollment in the most needed areas of study. Schools can better predict graduate numbers for long-term planning of enrollment (Althubaiti & Alkhazim, 2014).

Curriculum improvement. Using big data allows instructors to make changes and adjustments to improve curriculum development in the educational system, such as in the use of curricular mapping of data (Armayor & Leonard, 2010). Through the analysis of big data, educators can determine weaknesses in student learning and comprehension to determine whether or not improvements to the curriculum may prove necessary. Instructors can engage in educational strategic planning to ensure that the learning curriculum targets student needs to maximize learning potential.

Student learning outcome, behavior, and process. Another key benefit of big data and text mining focuses on the ability of schools and instructors to determine student learning outcomes in the

educational process as well as determine how to improve student performance (Bhardwaj & Pal, 2011). Researchers noted that the use of educational data mining contributed to positive results in the learning process (AlShammari, Aldhafiri, & Al-Shammari, 2013). Analysis of the data can help educators understand the student learning experience through learner interactions with technology tools such as e-learning and mobile learning (Hung & Zhang, 2012). Use of big data also reveals learning behavior, the impact on adaptive learning, and level of persistence (DiCerbo, 2014) in the learning process. By understanding the effects on learner outcomes, use of this data also reveals how to make improvements in student learning and performance in academic coursework. Therefore, LA allows instructors to evaluate forms of knowledge and adjust educational content accordingly.

Personalized learning. Arnold and Pistilli (2012) discussed an early intervention system that demonstrates the benefits and power of learning analytics. As an example, Course Signal provides students with real-time feedback. The components of students' grades, demographic characteristics, academic background, and demonstrated effort are all addressed. The system employs a personalized email and a stoplight, specific color method to indicate progress or lack thereof. Using learning analytics, the concept of personalized learning reveals student success. Dietz-Uhler and Hurn (2013) asserted that course designers do not account for students who do not begin specific coursework at the same learning stage and who do not proceed, learn, and master course competencies at the same pace. Learning analytics allows faculty to use data collected by the learning management system to observe the frequency of student login. Instructors can also see student interaction within the course, total engagement, pace, and grades. These components serve as predictors of students' potential success or failure. Learning analytics allows for real-time reception of the pertinent data, review as well as the incorporation of data, and real-time feedback for every student.

Improved instructor performance. Using this data also helps to assess instructor performance (Mardikyan & Badur, 2011). The use of data provides an opportunity to improve instructor development so that instructors are better prepared to work with students in a technological learning environment. Through the acquisition of data generated from instructor usage of technology and research tools in online libraries (Xu & Recker, 2012), analysts can determine online behaviors by educators. Therefore, use of this information can help identify areas in need of improvement by the instructor to facilitate enhanced instructor-student interactions in the educational environment.

Post-educational employment. Using big data allows educational institutions to identify post-education employment opportunities for graduates and help target education that more closely aligns with employment market needs. It can also predict graduate employment, unemployment, or undetermined situations about job opportunities (Jantawan & Tsai, 2013). Using big data can help stakeholders in the educational system better understand vocational prospects for students and better assess student learning programs for occupational compatibility (Kostoglou, Vassilakopoulos, & Koilias, 2013). In a global learning environment, this type of information not only can facilitate better educational and post-education vocational planning, but also may prove useful to organizations as they make hiring and budgeting decisions for college graduates in different disciplines.

Learning analytics practitioners and research community. The research community also benefits from the use of big data in education. Researchers can more easily share information and collaborate. They can identify gaps between industry and academia so that research can determine how to overcome problems. Also, useful data analysis represents an important component of the ability of scholars to generate knowledge as well as continue to progress in research disciplines (Sharda, Adomako, Asamoah, & Ponna, 2013). However, these benefits are also offset by the need for trained personnel who can use and apply analytics appropriately. Current researchers note a looming future gap in practitioners possessing requisite analytical skill sets in the area of business intelligence and analytics. Picciano (2012) noted a lack of sufficiently trained database administrators and designers to address present needs. This

issue has become an important focus for academic and business researchers seeking to overcome this problem through improved education in this area (Hsinchun, Chiang, & Storey, 2012).

Review of the benefits illustrates the usefulness of big data in education. Data analysis provides educational stakeholders a comprehensive overview of the performance of the institution, curriculum, instructors, students, and post-educational employment outlooks. It also provides scholars and researchers with needed information to identify gaps between education and industry so that educators and institutions can overcome these deficiencies in course offerings. More important, the ability of big data to provide these revelations can help the field of education make significant progress to improve learning processes. Further, LA promotes education as it strives to keep pace with the growth of information systems and new technologies. Therefore, the use of LA can help stakeholders understand how well the higher education process is working (Grummon, 2009). The use of learning analytics can prove useful in higher education performance evaluation at the university level. It can provide higher education institutions, instructors, and students improved metrics by which to measure the effectiveness of teaching methods, the engagement of learners in the educational environment, and the efficiency of the learning process using technology.

Learning Analytics Challenges in Education

The review of the literature revealed the LA challenges about data tracking, data collection, data analysis, a connection with learning sciences, learning environment optimization, emerging technology, and ethical concerns regarding legal and privacy issues.

Data tracking. The digital tracking of information is a technique used by analysts to determine how best to present new learning opportunities as the wave of education continues to move forward into the second decade of the 21st Century. The tracking of big data represents the monitoring system. Current trend tracking indicators regarding the delivery and dissemination of instruction depend on the learning management system used by the institution. Platforms such as Moodle, Canvas, EPIC, and Blackboard have the capability to track the number of times an individual logs into the course room. These platforms also provide significant documentation to determine how involved the student was upon their login. Such tracking provides those who plan and implement new educational programs with valuable information. The monitoring reveals how engaging the curriculum presented is, as well as identifying areas that cause confusion (Brown, 2012).

Data collection. The collection of data can be a challenge when looking at LA. Nonetheless, it represents an important component in planning for continued implementation of educational program growth (Bottles, Begoli, & Worley, 2014). Educators must consider several elements. They must consider the availability of resources at a venue. Next, instructors must establish a viable social platform as it directly relates to interactions between learners to synthesize the educational content. Finally, instructors must discriminate whether the learner population possesses the requisite suitability for this type of learning environment and knowledge acquisition. Besides these challenges, gaps exist because of the inability to share proprietary information gathered by the institution. Further, another problem emerges because the creation of the ideal framework to disseminate educational curriculum takes teamwork, especially among the organizations bidding against one another to capture the learner population who want to engage in this type of learning experience.

Evaluation process. An important consideration of data collection concerns how learning analytics has become a force in the evaluation process. As greater amounts of educational resources become available online, there is a subsequent increase in the total data available regarding learning interactions. For learning analytics to help instructor evaluation to function appropriately, data needs to be delivered in a timely and accurate manner (Picciano, 2014). Learning analytics can provide powerful

tools for developing meaning from interactions and actions within a higher education learning environment (Fournier, Kop, & Sitlia, 2011). With the unprecedented explosion of available data for online interactions, it is critical for the continued development of the evaluation process. LA can translate from other fields as interest in the data growth in education becomes more focused. Lias and Elias (2011) noted that statistical evaluation of rich data sources already exists within other professions and fields.

Data analysis. Technical challenges also exist from the assimilation of the data analysis because of the presentation format of the data. Erroneous data can skew the findings causing a misinterpretation of the overall population. Such scenarios are commonplace in the online learning environment. For example, an instructor may create a student profile to isolate an assignment that requires grading, test the ease of submission process, or to determine if there are any gaps in the presentation of the curriculum as it appears for students. Creation of a non-existent learner introduces redundant information that appears in the course without identification. This data does not represent student information but rather misinformation created by the instructor that flows into the big data pool of information (McNeely & Hahm, 2014). When manually conducting data analysis, this information can be easily identified from the population. However, working with data collection from the learning analysis vantage point adds a significant margin of error to the outcome of overall results.

Learning sciences connection. According to Pea (2014), personalized learning and learning opportunities demonstrate an inability to leverage learning analytics optimally; therefore, "the endgame is personalized Cyberlearning at scale for everyone on the planet for any knowledge domain" (p. 17). Ferguson (2012) asserted that to optimize and fully understand learning requires understanding how knowledge develops and how to support knowledge development. Further, researchers must understand the components of identity, reputation, and affect. Researchers must find ways to connect "cognition, metacognition, and pedagogy" (Vahdat et al., 2015, p. 299) to help improve learning processes. With a stronger connection to learning sciences, learning analytics can promote effective learning design.

Learning environment optimization. Ferguson (2012) noted that as learners expand the boundaries of the learning management system into open or blended learning settings, researchers must discover the problems faced by students and how to determine success from the learners' perspectives. This process will encumber a shift toward more challenging datasets that may include mobile, biometric, and mood data. Besides the individual learning aspect of learning analytics, researchers are seeking to address another component known as social learning analytics. In this context, social learning analytics focuses on the collaboration and interaction of learners in a socialized learning environment, not just on individual learning outcomes (Buckingham Shum & Ferguson, 2012).

Emerging technology. The full potential of learning analytics relating to learning requires continued and emerging technology that presently remains in the younger stages. This revelation presents a challenge as the technology continues to develop to stay constant with the growth of learning analytics. Further, to fully understand the method and practice of teaching, more research is needed. Research focusing on learning analytics and pedagogy is still in the beginning stages (Dyckhoff, Zielke, Bültmann, Chatti, & Schroeder, 2012).

Ethical and privacy issues. Another issue that emerges about learning analytics concerns the ethical, legal, and risk considerations (Kay, Korn, & Oppenheimer, 2012). Because of dynamic changes in technology as well as how users store data and applications in cloud-based systems, "the challenges of privacy and control continue to affect adoption and deployment" (Johnson et al., 2011, p. 3). Further, the ethical and legal complexities of learning analytics challenge institutions that seek to implement their usage (Sclater, 2014a). For example, these considerations can include obvious areas of privacy considerations such as consent, data accuracy, how to respect privacy, maintaining anonymity, opting out

of data gathering, and the potential effects to students. Additional concerns include data interpretation, data ownership, data preservation, sharing data with parties outside of the institution, and proper training of staff members regarding the handling of data (Sclater, 2014b). Further, the question becomes who owns this aggregate data, because having an infrastructure with the capacity to house large amounts of information becomes a daunting task (West, 2012). Because of these different issues, institutions must achieve a balanced approach to safeguard data while also assuring benefits to the educational process through the use of four guiding principles. These principles consist of clear communication, care, proper consent, and complaint (Kay et al., 2012). Institutions must demonstrate adherence to legal and ethical parameters to safeguard student privacy concerns while also achieving the educational goals for students and educators.

To accomplish privacy and educational goals, Slade and Prinsloo (2013) proposed the use of an ethical framework based on six principles that strives to maintain educational effectiveness for instructors and students. The first principle proposes that learning analytics should focus on understanding what works within the educational process and the moral necessity to use information. The second principle advances the notion that students must collaborate with the institution and willingly provide informed consent for institutions to collect, use, and store data. The third principle asserts that a student's identity and education performance is temporary. Therefore, this information should possess an expiration date as well as permit students to request the deletion of data under predetermined guidelines. Principle four notes that student achievement consists of various factors that can demonstrate a diverse range of considerations and complexities. For this reason, data may not always reflect student learning accurately. Principle five states that transparency should serve as an important guideline to ensure that data usage conforms to appropriate standards based on its purpose, who may access the data, and privacy protection. Finally, principle six suggests that institutions cannot afford to ignore the value of learning analytics and its value to learning outcomes. While learning analytics poses different ethical, legal, and risk challenges that stakeholders must understand, proper awareness and scrutiny can ensure that users adhere to ethical and legal parameters to balance privacy protections of students with institutional information needs.

Discussion

Analysis of the literature provided a comprehensive overview of the analytical learning methods, benefits, and challenges regarding the use of big data in education. Examination of these aspects showed positive contributions. However, the literature review also revealed negative considerations regarding the use of learning analytics. For these reasons, a closer examination of these aspects can reveal other insights.

This study revealed that learning analytics is an interdisciplinary field or *Jackdaw* field of inquiry as suggested by Clow (2013). As such, learning analytics selects and uses methods and analysis techniques from other disciplines as appropriate to achieve the goal of improving education. LA offers interventions based on the predicted models grounded in the large datasets in educational institutions. Learning analytics follows a macro-level process of data collection, data analysis, prediction model development, intervention, and refinement of the predicted model (Campbell & Oblinger, 2007; Clow, 2012, 2013). There is a hierarchy framework in which the flow of analytical information traces from students to the stakeholders. This research reveals that the process must include stakeholders to offer input and feedback. Use of an inclusive strategy will lead to an enhanced experience for the students. Learning analytics uses various approaches including visual data analysis techniques, social network analysis, semantic, and educational data mining to analyze the data (Bienkowski et al., 2012; Dawson & Siemens, 2014). In particular, this study described the two widely used approaches of data visualization and education data mining method of Baker (2010), which includes prediction, clustering, relationship mining, discovery with models, and separation of data for human judgment. Learning analytics allows for

stakeholders—such as instructors and learners—to receive insights about the educational setting. Further, researchers must validate and report research results as well as promote best practices for stakeholders.

From using these various methods, big data in education exhibits many benefits in the educational environment. Review of the literature revealed how data analytics can yield benefits to education by helping institutions and instructors to improve course offerings. LA can also help establish subject curriculums that more closely align with the needs of students as well as the needs of industry. Further, analytics can provide useful feedback regarding learner outcomes in the educational process as well as the effectiveness of instructors within the classroom. Moreover, analytics can provide useful information to the research community to help identify and resolve gaps between academia and different industries. Though LA appears to demonstrate a comprehensive representation of benefits within education, one must remember that the use of learning analytics is still in the early stages of development. Therefore, additional benefits may emerge as the use of analytics evolves and technology improves. Further, as people acquire greater skills in the field of analytics, more sophisticated uses of big data may also emerge to provide more useful insights in educational environments.

While learning analytics demonstrates benefits to educators and students through improved lectures and student learning outcomes (Scheffel et al., 2014), a close review of the ethical, legal, and risk concerns demonstrate the complexities of balancing challenges with educational concerns. Though current technological capabilities allow the acquisition, storage, and access of data, users must remain current on technology innovations that may provide improvements to these systems and how institutions can balance stakeholder needs with ethical considerations. To accomplish this, people who work with data must possess the requisite training in learning analytics to understand how to use the data productively to achieve meaningful results. Further, users must also understand the various ethical and legal aspects as well as the inherent risks of data usage. To protect privacy and maintain ethical standards, guidelines should inform appropriate data use. Further, accountability at all levels of data use will require an appropriate system of checks and balances. Additionally, data security should be well implemented and maintained to ensure privacy protections. In summary, mechanisms must provide appropriate transparency, data controls by students, information security, and accountability safeguards (Pardo & Siemens, 2014).

Regardless of the insights revealed by these components, the use of big data in education remains a relatively new area of focus in the field of education. Though the methods provide useful information within the current environment, analysts within the educational community must continue to develop and improve the methods needed. By properly refining the use of data to understand its meaning and implications, researchers can determine how best to use the information to improve the educational process at all levels and for all stakeholders.

Recommendations

It is evident that data acquisition using learning analytics demonstrates considerable implications for the future of the educational process. Not only does LA allow institutions to understand trends for planning purposes, but it also allows educators and learners to focus on needed skills for the 21st century. However, users must continue to grow with the acquisition and application of new technology that can provide enhanced capabilities to meet educational needs. To help improve educational outcomes in the learning environment, researchers will need to explore the use of different types of data sets from biometrics, learner moods, and mobile devices (Ferguson, 2012). The use of biometrics, such as fingerprints, facial recognition, iris recognition, and hand or finger geometry, would provide user authentication to verify user access as well as attendance, work, and testing in the educational process (Sayed & Jradi, 2014). Use of this technology would serve as an additional layer of protection for data

and would help to mitigate ethical, legal, and risk concerns. Because of the predicted rise in data acquisition and storage, researchers could integrate biometric services into a cloud-based system or different e-learning environments (Peer, Bule, Gros, & Štruc, 2013).

Another recommendation would be to identify ways to better streamline research priorities to ensure the use of the most pertinent information. For example, researchers could focus on refinement of learning analytics to help achieve better learning outcomes. Further, better standardization of assessments could improve student engagement and help establish learner competency levels. The expansion of learning analytics that includes different learning contexts could provide an improved understanding of learner-to-learner interaction within the educational environment (Pea, 2014). By addressing these recommendations, learning analytics will continue to be a valuable means of predicting future trends and needs in the field of education. As online education continues to flourish into the 21st century, educators need to continue to embrace technological changes in the educational sector. By doing so, these technological innovations will continue to serve as the driving force to make online teaching a successful process to achieve learning outcomes for all students.

Contributions, Implications of the Study, and Results

This study has revealed the different methods used by learning analytics not only to show how the use of big data can benefit education but also to reveal the challenges faced by stakeholders in the educational process. The field of learning analytics has been and will continue to expand greatly in part due to the ability to store increased amounts of data and a large diversity of research strands. As a result of greater availability and access to data, learning analytics will provide an increased understanding of the patterns of learner behavior, networks, and interactions. West (2012) cited the example of a total of sixteen states presently utilizing data mining techniques and learning analytics to identify students classified as at-risk. However, this study also revealed that learning analytics underscores the need for greater understanding of how to analyze data to optimize results and use the information to improve the educational process at all levels.

Going forward, schools must recognize the importance of implementing a data-driven approach to education. The use of performance systems allows for increased and more productive decision-making, the identification of trends and problematic areas, and the more efficient allocation of resources. Moreover, performance systems reveal how stakeholders within education must strive to continually improve methods, processes, and knowledge. Doing so ensures that the use of learning analytics continues to serve the needs of educators, students, institutions, and researchers throughout the educational process. Further, using learning analytics will ensure positive steps forward in the educational system as well as ensure greater outcomes for all stakeholders.

 Learning analytics is still an emerging field of education. Higher education stakeholders need to become further familiar with learning analytics applications in education (Scheffel et al., 2014). Nonetheless, there is a gap in literature to provide an integrative report about learning analytics in higher education. To bridge the gap, this review of literature provided an overview of the methods, benefits, and challenges regarding the use of learning analytics for leaders, administrators, instructors, and course developers in higher education. It is hoped that the overview enhances the application of learning analytics in higher education and supports the use of large datasets to improve the quality of teaching and learning.

References

AlShammari, I. A., Aldhafiri, M. D., & Al-Shammari, Z. (2013). A meta-analysis of educational data mining on improvements in learning outcomes. *College Student Journal, 47*(2), 326-333.

Althubaiti, A., & Alkhazim, M. (2014). Medical colleges in Saudi Arabia: Can we predict graduate numbers? *Higher Education Studies, 4*(3), 1-8.

Armayor, G.M., & Leonard, S. T. (2010). Graphic strategies for analyzing and interpreting curricular mapping data. *American Journal of Pharmaceutical Education, 74*(5), 1-10.

Arnold, K. E., & Pistilli, M. D. (2012, April 29). Course signals at Purdue: Using learning analytics to increase student success. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 267-270). New York, NY: ACM. doi: 10.1145/2330601.2330666

Baker, R. (2010). Data mining for education. *International Encyclopedia of Education*, *7*, 112-118.

Baker, R., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, *1*(1), 3–16.

Bhardwaj, B. K., & Pal, S. (2011). Data mining: A prediction for performance improvement using classification. *International Journal of Computer Science and Information Security, 9*(4), 136-140.

Bienkowski, M., Feng, M., & Means, B. (2012). *Enhancing teaching and learning through educational data mining and learning analytics: An issue brief*. U.S. Department of Education, Office of Educational Technology. Washington, D.C. Retrieved from http://www.ed.gov/technology.

Bottles, K., Begoli, E., & Worley, B. (2014). Understanding the pros and cons of big data analytics. *Physician Executive*, *40*(4), 6-12.

Brown, M. (2012). Learning analytics: Moving from concept to practice. *EDUCAUSE Learning Initiative*. Retrieved from http://net.educause.edu/ir/library/pdf/ELIB1203.pdf

Buckingham Shum, S., & Ferguson, R. (2012). Social learning analytics. *Educational Technology & Society*, *15*(3), 3-26.

Campbell, J. P., De Blois, P. B., & Oblinger, D. G. (2007). Academic analytics: A new tool for a new era. *Educause Review*, *42*(4), 40-57. Retrieved from http://www.educause.edu/ero/article/academic-analytics-new-tool-new-era

Campbell, J. P., & Oblinger, D. G. (2007). Academic analytics. *Educause*. Retrieved from http://net.educause.edu/ir/library/pdf/pub6101.pdf.

Clow, D. (2012). The learning analytics cycle: Closing the loop effectively. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 134-138). New York, NY: ACM. doi:10.1145/2330601.2330636

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

Cooper, H. (1988). The structure of knowledge synthesis: A taxonomy of literature reviews. *Knowledge in Society, 1*, 104-126.

Dawson, S., & Siemens, G. (2014, September). Analytics to literacies: The development of a learning analytics framework for multiliteracies assessment. *International Review of Research in Open and Distance Learning*, *15*(4), 284-305.

DiCerbo, K. E. (2014). Game-based assessment of persistence. *Journal of Educational Technology & Society, 17*(1), 17-28.

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

Dyckhoff, A. L., Zielke, D., Bültmann, M., Chatti, M. A., & Schroeder, U. (2012). Design and implementation of a learning analytics toolkit for teachers. *Journal of Educational Technology & Society*, *15*(3), 58-76.

EDUCAUSE. (2010). Next generation learning challenges: Learner analytics premises. *EDUCAUSE Publications*. Retrieved from http://www.educause.edu/Resources/NextGenerationLearningChalleng/215028

Elias, T. (2011). *Learning analytics: Definitions, processes and potential* (Report). Retrieved from http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf

Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, *4*(5/6), 304–317.

Fournier, H., Kop, R., & Sitlia, H. (2011). The value of learning analytics to networked learning on a personal learning environment. *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge* (pp. 246-250). New York, NY: ACM. doi:10.1145/2567574.2567613

Grummon, P. T. H. (2009). Trends in higher education. *Planning for Higher Education*, *37*(4), 48-57.

Hsinchun, C., Chiang, R. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, *36*(4), 1165-1188.

Hung, J.L., Hsu, Y.C., & Rice, K. (2012). Integrating data mining in program evaluation of k-12 online education. *Educational Technology & Society*, *15*(3), 27-41.

Hung, J., & Zhang, K. (2012). Examining mobile learning trends 2003-2008: A categorical meta-trend analysis using text mining techniques. *Journal of Computing in Higher Education, 24*(1), 1-17. doi:http://dx.doi.org/10.1007/s12528-011-9044-9

Jantawan, B., & Tsai, C. (2013). The application of data mining to build classification model for predicting graduate employment. *International Journal of Computer Science and Information Security, 11*(10), 1-7.

Johnson, L., Levine, A., Smith, R., & Stone, S. (2010). *The horizon report: 2010 edition* (Report). Retrieved from http://www.nmc.org/pdf/2010-Horizon-Report.pdf

Johnson, L., Smith, R., Willis, H., Levine, A., & Haywood, K., (2011). *The horizon report: 2011 edition* (Report). Retrieved from https://net.educause.edu/ir/library/pdf/HR2011.pdf

Kay, D., Korn, N., & Oppenheim, C. (2012). *Legal, risk and ethical aspects of analytics in higher education* (White Paper). Retrieved from http://publications.cetis.ac.uk/wp-content/uploads/2012/11/Legal-Risk-and-Ethical-Aspects-of-Analytics-in-Higher-Education-Vol1-No6.pdf

Kostoglou, V., Vassilakopoulos, M., & Koilias, C. (2013). Higher technological education specialties and graduates' vocational status and prospects. *Education & Training, 55*(6), 520-537. doi:http://dx.doi.org/10.1108/ET-03-2012-0026

Lias, T. E., & Elias, T. (2011). *Learning analytics: The definitions, the processes, and the potential* (Report). Retrieved from http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf

Mardikyan, S., & Badur, B. (2011). Analyzing teaching performance of instructors using data mining techniques. *Informatics in Education*, *10*(2), 245-257.

McNeely, C. L., & Hahm, J. (2014). The big (data) bang: Policy, prospects, and challenges. *Review of Policy* Research, *31*(4), 304-310.

Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, *45*(3), 438-450. doi:10.1111/bjet.12152

Pea, R. (2014). *The learning analytics workgroup: A report on building the field of learning analytics for personalized learning at scale* (Report). Retrieved from https://ed.stanford.edu/sites/default/files/law_report_complete_09-02-2014.pdf

Peer, P., Bule, J., Gros, J. Ž., & Štruc, V. (2013). Building cloud-based biometric services. *Informatica*, *37*(2), 115-122.

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.

Picciano, A. G. (2014). Big data and learning analytics in blended learning environments: Benefits and concerns. *International Journal of Artificial Intelligence and Interactive Multimedia*, *2*(7), 35-43.

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

Romero, C., & 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.

Sayed, M., & Jradi, F. (2014). Biometrics: Effectiveness and applications within the blended learning environment. *Computer Engineering and Intelligent Systems*, *5*(5), 1-9.

Scheffel, M., Drachsler, H., Stoyanov S., & Specht, M. (2014). Quality indicators for learning analytics. *Educational Technology & Society*, *17*(4), 117–132.

Sclater, N. (2014a, September 18). *Code of practice "essential" for learning analytics*. Retrieved from http://analytics.jiscinvolve.org/wp/2014/09/18/code-of-practice-essential-for-learning-analytics/

Sclater, N. (2014b, November). *Code of practice for learning analytics: A literature review of the ethical and legal issues*. Retrieved from http://repository.jisc.ac.uk/5661/1/Learning_Analytics_A-_Literature_Review.pdf

Sharda, R., Adomako Asamoah, D., & Ponna, N. (2013). Research and pedagogy in business analytics: Opportunities and illustrative examples. *Journal of Computing & Information Technology*, *21*(3), 171-183. doi:10.2498/cit.1002194

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, *57*(10), 1510-1529. doi: 10.1177/0002764213479366

Vahdat, M., Ghio, A., Oneto, L., Anguita, D., Funk, M., & Rauterberg, M. (2015). Advances in learning analytics and educational data mining. *Proceedings from 2015 European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, *Belgium*. Retrieved from http://www.idemployee.id.tue.nl/g.w.m.rauterberg/publications/ESANN2015paper1.pdf

West, D. M. (2012, September). Data mining, data analytics, and web dashboards. *Governance Studies at Brookings*, 1-10. Retrieved from Brookings.edu website at: http://www.brookings.edu/~/media/research/files/papers/2012/9/04%20education%20technology%20west/04%20education%20technology%20west.pdf

Xu, B., & Recker, M. (2012). Teaching analytics: A clustering and triangulation study of digital library user data. *Journal of Educational Technology & Society, 15*(3), 103-115.