Learning Analytics : definitions, applications and related fields

A study for future challenges

Entesar A. Almosallam, Henda Chorfi Ouertani

King Saud University, Information Technology Department
{ealmosallam, houertani}@ksu.edu.sa

Abstract. In the last few decades, the number of people connected online for educational purpose is increasing dramatically and consequently a huge quantity of data is being generated. This data is mainly "traces" or "digital breadcrumbs" that students leave as they interact with online learning environments. Confident that this data can teach us about learners' behaviors and help us enhancing learning experience, there has been a growing interest in the automatic analysis of such data. A research area referred to as Learning Analytics (LA) is identified. It is considered by many researchers as a strategic trend in education. Nevertheless, LA cannot be considered as a new field, it actually derives from different related fields such as Educational Data Mining, Academic Analytics, Action research, Personalized Adaptive Learning.

In this paper, we begin with an examination of the educational factors that have driven the need and the development of analytics in education. We study connections between LA and its most related fields (Educational Data Mining and Academic Analytics). We summarize this interconnection in a table showing for each field the objectives, the stakeholders, the methods and the initial trigger behind the analysis actions. After that we study and run through LA applications presented in the International Learning Analytics & Knowledge Conferences during the three last years. Finally, we conclude by identifying some challenges in the area of LA in relation to the driven factors related to Educational Data.

Keywords: Learning Analytics, Educational Data, Analytics Applications.

T. Herawan et al. (eds.), *Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013)*, Lecture Notes in Electrical Engineering 285, DOI: 10.1007/978-981-4585-18-7_81,

1 Introduction

In the last few decades, the number of people connected online for educational purpose is increasing dramatically and consequently a huge quantity of data surrounding these interactions is being generated. In higher education, most of the learning systems used by colleges and universities usually collect comprehensive log data associated with learners' behaviors and actions. However, the reports generated based on this data provide very general and limited information [1]. Recently, there has been a growing interest in the automatic analysis of educational data and how this data can be used to improve teaching and learning. This interest has seen increasingly rapid advances in the emerging field of Learning Analytics (LA).

LA is emerging as a way for educational institutions to use the data they usually collect and save for more than just reporting – to gain a better understanding of what they are doing, and so gain a strategic improvement. Generally, LA deals with the development of methods that harness educational data to support the learning process [2]. It is part of a growing academic trend towards big and open data: LA is identified in the 2013 Horizon Report **[7]** as a key future trend for education, and as having real potential to improve learner experience.

Although LA is still an emerging discipline; it is not new. It has roots in various fields, involving machine learning, artificial intelligence, information retrieval, statistics, visualization, and research models in general [2]. However, what is new is the rise of quantity and quality of captured data about the learning processes [4]. As a consequence, analytics have gained attention in education.

LA is a multi-disciplinary field in which several related areas of research converge. These include academic analytics, action research, educational data mining, recommender systems, and personalized adaptive learning [5]. LA borrows from these different related fields and applies several existing techniques.

The remainder of this paper is structured as follows: the next section provides some definitions of LA. How LA can change education is presented in Section 3. Section 4 identifies the related areas overlapped with LA. A through up LA applications during the last 3 years is discussed in section 5. Conclusions and challenges are presented in Section 6.

2 What is Learning Analytics?

Different definitions have been provided for the term 'Learning Analytics'. Learning Analytics is defined on LAK11 website [1] - the first conference on LA - as "the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs".

According to [6], LA "refers to the use of intelligent data, learner-produced data, and analysis models to discover information and social connections, and to predict and advise on learning". EDUCAUSE's Next Generation learning initiative [2] describes LA as: "the use of data and models to predict student progress and performance, and the ability to act on that information" [6].

[1] https://tekri.athabascau.ca/analytics/
[2] http://nextgenlearning.org/

Another definition that experts in this field provide is: "Learning Analytics refers to the interpretation of a wide range of data produced by and gathered on behalf of students in order to assess academic progress, predict future performance, and spot potential issues" [7]. Even these definitions are different; it is noticeable that they focus on changing educational data to insights and actions aiming to improve learning [6].

3 How can LA changes education?

The 2013 Horizon Report [7] includes LA in the 2-3 years of adoption. It is predicted that LA will be considered as a foundational tool for informed change in education within that time [8]. Evidence inferred from LA tools can support us to make better decisions. We can better understand learning behaviors and how the learning process outputs are produced. LA helps in identifying learners who are at risk and provides the guide to the decision makers to choose and apply the best actions that will help learners to success. Techniques used in LA tools are very compelling for adoption in education [10]. Classical methods of student assessment and evaluation can report and show students results collected generally. However, L.A. permits to track student performance by integrating their data that is collected through a variety of sources. It gives a clear idea of what students are doing and their gained learning in the whole learning process. Moreover, these evaluations are usually done after the completion of the course letting impossible the interventions while the course is in progress. Also, it is difficult to find out some important information about how students use the educational content, the time consumed, and interaction with other learners. However, all these benefits and more could be accomplished using LA.

LA could offer a holistic view of the learning experience for both learners and teachers to discover the data, depending on the needs of their educational context [9]. LA can provide dramatic change in education, by altering how educational institutions develop curriculum, present it, evaluate student learning, provide learning feedback, and even allocate resources [8].

However, a combination of technological and educational factors drives the need and the development of analytics in education. Some of these factors are related to the *educational data*, the *educational task*, or the *stakeholders* [11].

3.1 Educational data

LA is data-driven approach [14], as it is based on the collection, process, and analysis of collected educational data.

In educational environments, there are many different types of *educational data* available for analyzing. These data are specific to the educational area, and therefore have essential meaning, and relationships with other data. For example, it is more difficult when dealing with *distributed* data from multiple sources with different formats rather than dealing with data that is *central* in one source [12]. Also, the *scope* and the *size* of the data to analyze is important factor. As most of the educational institutions move to online learning, this means that they deal with increasingly large sets of data [11]. Furthermore, it is also necessary to take privacy aspects of the learners and the educational institutions into consideration. The capabilities of LA tools will be limited when the data is *private* in contrast with *public* and *open* data [13].

3.2 Educational Tasks

In terms of educational tasks, there are many factors to the development of LA depending on the objective of applying LA in each application area. Although the objectives and stakeholders of all application areas overlap, they require analytics using different metrics and work on different scales. On one hand, there are many applications or tasks in educational environments that have been resolved through LA.

Most of the possible objectives of LA were stated in table 2 such as: monitoring, analysis, prediction, intervention, tutoring/mentoring, assessment, feedback, adaptation, personalization, recommendation, and reflection. On the other hand, the values that LA applications extract can be very helpful for different stakeholders mainly: learners, teachers and administrators [9, 10]. In the table 1, we present the most common educational tasks that have employed LA techniques grouped by stakeholders.

Table 1. educational tasks that employed LA techniques grouped by stakeholders

Learners	Teachers	Administrators
• Self-reflection • Assessment and feedback • Adaptation and Personalization	• Curriculum development • Interventions and Recommendations	• Monitoring • Prediction • Making decisions

3.3 Stakeholders

Learners. LA can be a valuable tool to promote self-reflection. Reflective learning offers the chance of learning by evaluating past work in order to improve future experiences and promote continuous learning. Students can benefit from data compared within the same course, across classes, to draw conclusions and self-reflect on the progress of their learning practice. [11]

In addition, by using LA, the process of evaluation of learners can be enhanced to produce a ***real-time assessment*** rather than the assessment at the end of a course [8]. LA supports the assessment to improve the efficiency of the learning process by providing ***feedback*** to different stakeholders. Feedback provides interesting information generated based on data about the user's interests and the learning context [5].

In terms of ***Adaptation and personalization***, LA would be able to adapt and provide personalized improvements that adjust to the needs of each learner based on the previous academic performance [8]. The objective of LA here is to help learners decide what to do next by organizing learning resources and activities according to the preferences of each learner in adaptive way. This adaptation and personalization of the content of learning process, ensuring that each learner receives resources and teaching that reflect their current knowledge state [9].

Teachers. Teachers can benefit from LA to determine the appropriate curricula for students that adapt with their interests and preferences. When using LA, teaching components could be sharable and reusable between teachers easily. In addition, some teachers employ LA to identify specific parts in a course that cause students failures. Then, teachers can adjust curriculum or modify learning activities to enhance learning process. It helps teachers to know which academic practices need to be curbed and which need to be encouraged [9].

Indeed, the most common use of LA is to identify students who may need additional support based on their past and current activities and accomplishments and to provide ***early interventions*** to help them improve their performance. The effective analysis and ***prediction*** of the learner performance can help the teacher in intervention by suggesting appropriate actions that should be taken to assist learners achieve better outcomes [9].

In terms of ***recommendations***, LA can provide valuable insight into the factors that influence learners' success. It can provide early alerts or indications of which students are at risk in their learning behavior. By recognizing these students and offering early interventions, educational institutions can translate that data into actionable insights to reduce related problems dramatically [8].

Administrators. Today, LA is most often used in higher education for ***administrative decisions*** [9]. There is an increasing emphasis on the use of metrics and analytics for higher ***education*** in order to improve learning outcomes [10]. As a result, the majority of colleges and universities are now placing a heavy emphasis on the use of LA. This will encourage a more rapid adoption of LA as a way of meeting those needs.

LA has the ability to track student activities and generate reports in order to support decision-making by the educational institution. LA monitor and evaluate the learning process with the purpose of continuously improving the learning environment. Examining how students use a learning system and analyzing student accomplishments can help administrators find patterns and make decision on the future design of the learning process.

4 Classification of Learning Analytics and their related areas

In this section, we studied connections between LA and its different related fields (Educational Data Mining, Academic Analytics). To do that, we provide a brief definition for each field then we summarize them in table 2 showing for each field the objectives, the stakeholders, the methods and the initial trigger behind the analysis actions. While LA focused on the educational issues and how can we optimize opportunities for improve learning [11], Academic Analytics (AA) focused on the administrative concerns and how to help administrative users in higher education to make better learning opportunities and decisions [5]. It combines data with statistical techniques and predictive modeling to help determine which students may face academic difficulty, allowing interventions to help them succeed. The examples in the AA literature refer mostly to the problem of detecting "at risk students", that is, those students that might drop out of a course or abandon their studies [5]. AA usually focused on enrollment management and the prediction of student academic success. AA were restricted to statistical software. On the other hand, Educational Data Mining (EDM) focused on the technical challenges and how can we extract value from the large volumes of learning-related data [13]. From a technical perspective, EDM is the application of data mining techniques to educational data, and so, its objective is to support teachers and students in analyzing the learning process [27].

The objectives, analysis domain, data, and process in LA, AA and EDM are quite similar. All of them focus on the educational domain, work with data originating from educational environments, and convert this data into relevant information with the aim of improving the learning process. However, the techniques used for LA can be different from those used in other fields. In addition to data mining techniques, LA further includes other methods, such as statistical and visualization tools or social network analysis (SNA) techniques, and puts them into practice for studying their actual effectiveness on the improvement of teaching and learning [11].

Table 2. Classification of LAs and other related research areas

Field	Stakeholders	Objectives	Methods	Data
AA	Educational institutions	Enrollment management, prediction Support decision making	Statistical methods	Educational environments data
EDM	Teachers-Students	Convert data into relevant information to improve learning process	Data mining techniques (clustering, classification, association rules)	Educational environments data
LA	Learners - Teachers - Educational institutions	Enrollment, prediction, Reflection, Adaptation, Personalization, recommendation	Quantitative methods Data mining techniques (clustering, classification, association rules)	Educational environments data

5 Learning Analytics Applications

The number of publications around LA and its related areas has grown rapidly in the last few years. In this section, we review the most relevant studies in this field from 2011 to 2013. We mainly reviewed studies from the LAK11, LAK12, LAK13 conferences, which discussed concrete LA applications. To note that LA publications from 2011 to 2013 are not limited solely to the aforementioned ones. Several other conferences have included special tracks addressing LA and its related topics. However, the publications at LAK conferences presented a compelling picture of the potential of LA. Therefore, we restricted our literature review to these three conferences as they represent the rich diversity of research and project directions currently underway in the field of LA. We summarize just a few of the papers to try to capture a sense of this. An overview of these papers and related learning application is presented in table 3. We summarize characteristics of 10 LA applications: their educational environments, tasks and features, for whom they are intended, what data they track and how they have been facilitated with LA in practice.

Table 3. An overview of LA applications

Tool Name	Edu. Environment	Edu. Tasks	Edu. Data	Stakeholder	LA Metric	Visualization
Video-games [15] – 2012	Educational video games	**Task:** Assessment	In-game situations and interaction traces	Learners	Tracking techniques, rule-based system	Human-readable reports
AAT [16] – 2011	Within the Moodle Analytics project	**Task:** access and analyze student behavior data in learning systems	Data from one or a set of courses hosted in one or several databases	Learning designers teachers.	Statistics, customized queries and activity reports	Graphical user interface (GUI)
CAFe [17] – 2011	eTwinning-an online community for European teachers (digital mediated learning network)	**Task:** Self-monitoring tool for teachers • monitor their positions in the eTwinning network, • knowing their achievements • Recognizing their weaknesses and shortages.	Data dumps gathered from relational database in the eTwinning project - anonymous data sets	Teachers coordinators researchers	Implicit assessment methods of some indicators stored in XML files	Graphs, time series charts, and bar charts
SNAPP [18] – 2011	learning management systems	**Task**: to visualize the network of interactions resulting from discussion forum posts and replies	Student behavioral patterns and learning activities	Teachers	Thematic Analysis	Social network visualization
Course Signals [19] - 2012	Learning Management Systems	**Task**: Prediction - real-time feedback to students	Edu. data collected by instructional tools	Learners	Statistical & mining techniques - predictive student success algorithms	Traffic signal indicators
GLASS [20] – 2012	Web-based visualization platform	Visualization	Datasets stored using the schema that allows to capture events occurring during the use of various computer applications.	Learners and teachers		Time lines and charts of activity events
Learn-B [21] - 2012	Self-regulated workplace learning in organizational context	**Tasks:** Reflective learning & collaboration **Features:** Makes use of ontologies for data linking and annotation	Learners progress data	Learners	Statistics & Social Waves	A set of various visualization charts
E^{2}Coach [22] - 2012	Adaptive- recommender system for physics education	**Task:** Adaptive recommendations and interventions **Features:** Provides each student with an individualized coach.	Real time, content specific student performance data	Learners	Actionable intelligence	Dashboards

HOU2LEARN [23] - 2013	Personal Learning Environment	**Tasks:** To encourage social networking and collaboration **Features:** Open platform	E-portfolios of learners	University professors and high school teachers	Assessment method based on rubrics	Dashboards
StepUp! [24] - 2013	Visualization tool	Awareness – reflection - track student activity	Traces of learning activities collected automatically by software trackers	Learners	Statistics - tracking techniques	Dashboards

6 Conclusion and challenges

Learning Analytics promises to harness the power of advances in data analysis techniques to improve understandings of teaching and learning, and to tailor education to individual learners more effectively. However, LA requires key researchers to address a number of challenges, including questions about privacy, heterogeneity and integration, appropriate visualization, data structure, meaningful indicators and Costs issues. **Table 4** recapitulates some of the challenges and educational issues that need to be addressed during the LA development.

Table 4. Educational challenges and issues surrounding LA development

Challenges	Description
Privacy	Challenges to the ownership and use of the data - Who gets access to the data?
Heterogeneity	Different sources and formats
Intended stakeholders	The kinds of data and analysis employed depend on the intended audience and stakeholder.
Visualization	Appropriate and understandable information visualization for the different stackeholders
Data structure	The data can be structured (logged data) or unstructured (interaction data)
Lack of unique identifiers and meaningful indicators	Different stakeholders use different technologies in different ways. Stakeholders have different point of views.
Costs issues	Costs to store big data and producing LA tools

References

1. Elias, T. 2011. "Learning Analytics: Definitions, Processes and Potential".
2. George Siemens, "Structure and Logic of the Learning Analytics Field", 2013, Retrieved June 3, 2013, from http://www.learninganalytics.net/?p=185
3. Maurizio De Rose, " Future trends on data analytics to support your learners" , 2013, Retrieved May 15, 2013, from http://www.jisc.ac.uk/inform/inform35/FutureTrendsDataAnalytics.html
4. George Siemens , "Sensemaking: Beyond Analytics as a Technical Activity", April 11, 2012
5. Chatti, Mohamed Amine, et al. "A reference model for learning analytics."International Journal of Technology Enhanced Learning 4.5 (2012): 318-331.
6. Siemens, G. (2010). What are Learning Analytics? Retrieved May 14, 2013, from http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/
7. Johnson, L.,AdamsBecker, S.,Cummins,M., Estrada,V., Freeman,A., and Ludgate, H. (2013). NMC Horizon Report: 2013 Higher Education Edition. Austin, Texas: The New Media Consortium
8. How data and analytics can improve education, Retrieved May 14, 2013, from http://strata.oreilly.com/2011/07/education-data-analytics-learning.html
9. Jie Wu, "Revolutionize Education Using Learning Analytics", Retrieved May 11, 2013, from http://www.aboutjiewu.com/Resources/projects/INFO203/LearningAnalytics.pdf
10. Malcolm Brown, "Learning Analytics: The Coming Third Wave", April 2011, Retrieved May 12, 2013, from http://www.educause.edu/library/resources/learning-analytics-coming-third-wave
11. Ferguson, Rebecca. "Learning analytics: drivers, developments and challenges." International Journal of Technology Enhanced Learning 4.5 (2012): 304-317.
12. Suthers, Daniel, and Devan Rosen. "A unified framework for multi-level analysis of distributed learning." Proceedings of the 1st International Conference on Learning Analytics and Knowledge. ACM, 2011.
13. Romero, Cristóbal, and Sebastián Ventura. "Educational data mining: a review of the state of the art." Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 40.6 (2010): 601-618.
14. Bader-Natal, Ari, and Thomas Lotze. "Evolving a learning analytics platform."Proceedings of the 1st International Conference on Learning Analytics and Knowledge. ACM, 2011.
15. Serrano-Laguna, Ángel, et al."Tracing a Little for Big Improvements: Application of Learning Analytics and Videogames for Student Assessment."Procedia Computer Science 15 (2012).
16. Graf, Sabine, et al. "AAT: a tool for accessing and analysing students' behaviour data in learning systems." Proceedings of the 1st International Conference on Learning Analytics and Knowledge. ACM, 2011.
17. Song, Ergang, et al. "Learning analytics at large: the lifelong learning network of 160,000 European teachers." Towards Ubiquitous Learning. Springer Berlin Heidelberg, 2011. 398-411.
18. Bakharia, Aneesha, and Shane Dawson. "SNAPP: a bird's-eye view of temporal participant interaction." Proceedings of the 1st International Conference on Learning Analytics and Knowledge. ACM, 2011.
19. Arnold, Kimberly E., and Matthew D. Pistilli. "Course Signals at Purdue: Using learning analytics to increase student success." Proceedings of the 2nd International Conference on Learning Analytics and Knowledge. ACM, 2012.
20. Leony, Derick, et al. "GLASS: a learning analytics visualization tool."Proceedings of the 2nd International Conference on Learning Analytics and Knowledge. ACM, 2012.
21. Siadaty, Melody, et al. "Learn-B: a social analytics-enabled tool for self-regulated workplace learning." Proceedings of the 2nd International Conference on Learning Analytics and Knowledge. ACM, 2012.

22. McKay, Tim, Kate Miller, and Jared Tritz. "What to do with actionable intelligence: E 2 Coach as an intervention engine." Proceedings of the 2nd International Conference on Learning Analytics and Knowledge. ACM, 2012.
23. Koulocheri, Eleni, and Michalis Xenos. "Considering formal assessment in learning analytics within a PLE: the HOU2LEARN case." Proceedings of the Third International Conference on Learning Analytics and Knowledge. ACM, 2013.
24. Santos, Jose Luis, et al. "Addressing learner issues with StepUp!: an Evaluation." Proceedings of the Third International Conference on Learning Analytics and Knowledge. ACM, 2013.
25. Bauer, Florian, and Martin Kaltenböck. "Linked Open Data: The Essentials."Edition mono/monochrom, Vienna (2011).
26. The Open Data Handbook, Retrieved June 19, 2013, from http://opendatahandbook.org/
27. Siemens, George, and Ryan SJ D. Baker. "Learning analytics and educational data mining: towards communication and collaboration." Proceedings of the 2nd International Conference on Learning Analytics and Knowledge. ACM, 2012.