INTRODUCTION TO LEARNING ANALYTICS ADOPTION IN HIGHER EDUCATION INSTITUTIONS

Abstract. Educational data mining and Learning analytics represent a pair of research disciplines, which cover the majority of these data mining techniques, methods, applications as well as data mining tools in the area of education. The main aim of the paper is to summarize the main characteristics of Learning Analytics and focus on the approaches and frameworks used for its successful adoption and implementation of the environment of higher educational institutions.

Keywords: Learning analytics; data mining; educational technology; decision support

Introduction. Many researchers try to apply data mining methods to the wide range of education related problems. However, the educational data has rarely been mined intelligently with the goal of improving learning and informing teaching practice, although evidence from other sectors such as marketing, retail, health or technology suggests that the effective use of data can offer the education sector the potential to enhance its systems, approaches, and outcomes (Manyika et al., 2011).

The review of the current literature indicates, that educational data could be sucessfuly used in areas such user behavioral pattern modelling, user knowledge and experience modelling, user profiling, personalization and adaptive personalized learning, adaptive technologies and tools, identification of learning problems, study program measurement and evaluation, as well as improvement of learning and teaching experiences (Johnson et al., 2013).

The analysis of data collected from the interaction of users with educational and information technology has attracted much attention as a promising approach for advancing our understanding the learning process. This promise motivated the emergence of the new research field, learning analytics (LAK), and its closely related discipline, educational data mining (EDM). These research fields represent a pair of research disciplines, which cover the majority of data mining techniques, methods, techniques, and applications as well as data mining tools in the area of education.

Learning Analytics, which creates the main topic of this paper, is an emerging field of research that aspires to use data analysis to support decisions made at every level of the
educational institutions (Johnson et al., 2013). In other words, the main objective of LAK is a support of decision-making processes at the educational institutions. LAK deals with a gathering, measuring, and analysis of available data about stakeholders for understanding and optimizing of the learning process and the whole environment, where the learning process is realized (Baker and Siemens, 2014).

LAK offers the capacity to investigate the rising amount of learner data with the goal of understanding the activities and behaviours associated with effective learning and to leverage this knowledge in optimizing our educational systems (Bienkowski, Feng and Means, 2012).

According to the Gartner predictions for the digitalization of education, LAK has a large potential in the area of higher education (Tsai and Gasevic, 2016). Measuring student outcomes, growing personalized learning and adaptive learning technologies, and investing in effective learning analytics belong to the key areas of LAK implementation. Moreover, there are a lot of other examples of the internal and external pressure to use analytics to assess student success and resource management at the HEIs. However, the challenge of successful institutional change for learning analytics implementation calls for new adaptive forms of leadership, collaboration, policy development and strategic planning.

Therefore, the main aim of the paper is to summarize the main characteristics of LAK and focus on the approaches and frameworks used for its successful adoption or implementation to the higher educational institutions (HEI) environment.

The rest of the paper is structured as follows. The next section provides all necessary information about the current trends in the learning analytics. The third section brings a short comparison of the LAK with the closely related field Educational data mining. The provision of LAK for different kinds of stakeholders is shortly introduced in the next section. The fifth section summarizes the information about existing systematic models of LAK implementation. The last section discusses different aspects of LAK implementation at the higher educational institution and summarizes challenges and trends for better adoption of the LAK at the HEI.

**Learning Analytics.** The most widely used definition of the LAK describes LAK as a research discipline, which deals with a gathering, measuring, and analysis of available data about stakeholders for understanding and optimizing of the learning process and the whole environment, where the learning process is realized (Baker and Siemens, 2014).

Learning Analytics represents a new research discipline, which

– applies modern statistical techniques and methods on the system level of the learning process,
– targets the institutional level, and
– tries to support the decision-making processes of the educational institutions.

There are also known some less formal definitions of LAK, for example “learning analytics is about collecting traces that learners leave behind and using those traces to improve learning” or learning analytics is “the process of developing actionable insights through problem definition and the application of statistical models and analysis against existing and/or simulated future data” (Duval, 2012).

The original study about LAK defined analytics in a singular sense (Arroway et al., 2016; Gašević 2014). The current version of the study deepens understanding of LAK by distinguishing institutional analytics from learning analytics. While institutional analytics is intended to improve services or business practices, learning analytics is intended to enhance or improve student success. Although both share many characteristics surrounding interest, investment, and implementation, institutional analytics currently dominates. LAK remains somewhat less evolved on most campuses and often is rather known as educational data mining. The term learning analytics is used throughout the paper for its large acceptance in
the community. The differences between LAK and educational data mining are summarized in the next section.

LAK offers the possibility of implementing real-time assessment and feedback systems and processes at a scale that are focused on improvement of learning, development of self-regulated learning skills, and student success (Macfadyen et al., 2014).

The authors of LAK publications try to combine the research approaches and methods, which are typical for the different research fields such as computer science, psychological science, and sociology with the aim to improve the learning process on the institutional level considering the technical, pedagogical and social aspects of education.

The LAK initiatives include topics that are (Wong, 2016):

- learner-related, attempting to model the behaviour of the learners
- action-related, analysing learner’s logs and predicting outcomes
- assessment-related, communicating requirements and achievements, and personalising feedback
- content-related, identifying knowledge gaps, competencies, and misconceptions
- social-related, promoting learning in a social environment
- curriculum-related, assisting in mappings of objectives and competencies to graduate attributes
- adoption-related, discussing challenges and solutions around the adoption of learning analytics
- wellness-related, developing tools and interventions that support the wellness of students
- support-related, developing tools that support key stakeholders in teaching and learning
- context-related, contextualising learning analytics solutions based on the "analytics tone" of the HEI.

In other words, LAK deals with measuring, collecting, analysing and reporting of the data, which were created during the HEI stakeholders’ activity, for the purpose of the understanding, optimization or improvement of the learning process. At the same time, LAK utilizes the same activities in order to optimization of the learning environment (Siemens and Baker, 2012).

LAK focuses on the interpretation of the wide range of data with the aim of evaluation of the learning progress, future student’s performance, or potential study problems identification. The LAK outcomes can be applied not only to the evaluation of the student’s performance but also more widely. The LAK outcomes can help to review the study programs, syllabus, even the educational institutions (Johnson et al., 2011).

**Comparison of LAK and EDM.** According to Romero et al. (Romero and Ventura, 2010), Educational Data Mining (EDM) is an emerging interdisciplinary research area that deals with the application of data mining techniques to educational data. EDM deals with the development of new techniques and methods for discovering stakeholders’ behavioural patterns from unstructured, semi-structured or structured data, which comes from the interaction of stakeholders within the virtual learning environments, adaptive or intelligent tutoring systems or other educational software.

A typical EDM process converts raw data coming from this educational data sources into useful information that could potentially have a great impact on the educational research and practice (Romero et al., 2014). EDM tries to solve problems, which arise during different phases of a learning process. Romero et al. (2010) presented a comprehensive, up-to-date overview of the current state of data mining in education. The comprehensive state-of-the-art review of EDM research discipline was edited by Peña-Ayala (Peña-Ayala, 2014b, Peña-

Considering the current state of educational technology research field, it is necessary to compare EDM and LAK for completeness sake. LAK and EDM differ in their origins, techniques, fields of emphasis and types of discovery (Baker and Siemens, 2014).

However, LAK and EDM constitute an ecosystem of methods and techniques that successively gather, process, report and act on machine-readable data on an ongoing basis to advance the educational environment and reflect on learning processes. In general, these techniques and methods initially emphasize on measurement and data collection and preparation for processing during the learning activities. Consequently, they focus on further analysis, reporting of data and interpretation of results, targeting to inform and empower learners, instructors and management of the institution about performance and goal achievements, and facilitate decision making accordingly (Papamitsiou and Economides, 2014).

The common application domain of the aforementioned research fields is user modelling, modelling the students' knowledge, behaviour or experience during the learning process, user profiling, domain modelling, effectiveness measurement, trend analysis, recommendation and improvement of the learning process.

Both disciplines try to design models, tasks, methods, and algorithms for data exploration from the educational settings, to analyze educational data, to find out patterns and to make predictions that characterize learners, to observe the progress of the students, to allow an identification of the critical points of study and to define actions for improvement (Peña-Ayala, 2014a, Drlik et al., 2014).

A wide range of EDM and LAK methods has emerged over the last several years. Some are roughly similar to those seen in the use of data mining in other domains, whereas others are unique to educational data mining (Baker and Siemens, 2014). Predictive methods, structure discovery, relation mining, a distillation of data for human judgment and discovery with models belong to the most frequent categories of the advanced EDM and LAK methods (Baker and Siemens, 2014, Rodriguez, 2011).

While the EDM and LAK have similar application domains and objectives, use the same data sources, suppose the same knowledge of research methods, the technical, ideological and methodological background is different (Siemens and Baker, 2012) Table 1 summarizes the main differences between LAK and EDM.

<table>
<thead>
<tr>
<th></th>
<th>LAK</th>
<th>EDM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery</td>
<td>Leveraging human judgement is a key; automated discovery is a tool to accomplish this goal</td>
<td>Automated discovery is a key; leveraging human judgment is a tool to accomplish this goal</td>
</tr>
<tr>
<td>Reduction and Holism</td>
<td>Stronger emphasis on understanding systems as wholes, in their full complexity</td>
<td>Stronger emphasis on reducing to components and analyzing individual components and relationships between them</td>
</tr>
<tr>
<td>Origins</td>
<td>LAK has stronger origins in semantic web, &quot;intelligent curriculum,&quot; outcome prediction, and systemic interventions</td>
<td>EDM has strong origins in educational software and student modeling, with a significant community in predicting course outcomes</td>
</tr>
<tr>
<td>Adaptation and Personalization</td>
<td>Greater focus on informing and empowering instructors and learners</td>
<td>Greater focus on automated adaptation (e.g. by the computer with no human in the loop)</td>
</tr>
<tr>
<td>Techniques</td>
<td>Social network analysis, sentiment</td>
<td>Classification, clustering, Bayesian</td>
</tr>
</tbody>
</table>
The preferred approach of knowledge discovering represents the first difference between EDM and LAK research fields. EDM prefers automation of knowledge discovering steps, i.e. the straightforward integration of EDM methods into the learning environment. On the other hand, LAK prefers and supports the decision-making processes managed by the people.

The different understanding of adaptation and personalization can be considered the second important difference between the EDM and LA. EDM models are often used as the basis for automatized adaptation, adaptivity or user model creating. By contrast, the LAK models are developed mainly for the purpose of the better awareness of teachers and students about the results or issues of the learning process.

The researchers in the EDM prefer the top-down approach. They usually divide the observed problem into several smaller parts, which consequently they analyse in detail. The researchers in the LAK provide the holistic approach, they try to see the whole complexity of the observed phenomenon.

As was mentioned earlier, the origin of EDM and LAK is different. EDM is closely related to the educational software and VLEs development. It focuses on the students modelling and forecasting of their performance and successfulness in learning goals achievement. LAK originates in the semantic web. It tries to design and develop intelligently managed study programs.

Finally, the EDM and LAK use different methods and techniques. According to Romero et al. (2014), EDM often uses classification, clustering, Bayesian networks, visualization, discovering by models. LAK prefers social nets analysis, sentiment analysis or influence analysis.

**LAK Provision for Different Kind of Stakeholders.** Successful adoption of the LAK at the HEI should bring provision for all kinds of stakeholders of HEI. In other words, there are three common internal drivers for the adoption of learning analytics (Tsai and Gasevic, 2017), which motivate HEIs to adopt or implement some aspects of LAK in their processes:

- Learner-driver: to encourage students taking responsibility for their own studies by providing data-based information or guidance.
- Teaching-driver: to identify learning problems, improve teaching delivery, and allow timely, evidence-based support.
- Institution-driver: to inform strategic plans, manage resources, and improve institutional performances, such as retention rate and student satisfaction.

The provision for students is considered the most accepted and required results of successful LAK adoption at the HEI. LAK can add distinct value to teaching and learning practice by providing greater insight into the student learning process to identify the impact of curriculum and learning strategies, while at the same time facilitating individual learner progress.

The ways the LAK can enhance students’ learning experience through different stakeholders, is shown in figure 1 (Tynan and Buckingham Shum, 2013). Only suitable and sensitive combination of the responsibilities of all participating kinds of stakeholders can contribute to the positive acceptance of the LAK potential at the HEI.
Fig. 1. Ways that learning analytics enhances student learning experience through different key stakeholders (Tynan and Buckingham Shum, 2013).

The students’ activity represents the typical data source of the LAK methods. It includes, for example, assignment uploading, posts sent to discussions, test and questionnaires attempts, but also the relationship between the students (Bienkowski, Feng, and Means, 2012). After pre-processing of these data, the models are obtained, which teachers and managers can use for improvement of student’s learning path or study program optimization.

Students’ success is influenced by several factors, which can be divided into four categories:

– Entry characteristics (age, previous education, gender),
– academic compatibility (the subject of study, preparation for the study, quality of study materials, module assessment strategy),
– social and academic integration, and
– other external factors.

The impact of these indicators on the student’s success is different from the LAK point of view and needs further research (Buckingham Shum, 2013).

According to Buckingham Shum (2013), a clear vision should be developed to galvanise effort across the HEI on the focused use of LAK to drive students’ success. This vision assumes to use and apply information strategically using specified indicators and measures to retain students and progress them to complete their study goals (Fig. 2). Moreover, this vision assumes using sophisticated GUI in the form of dashboards, reports and other tools, which are especially prepared for different kind of stakeholders and reduce the amount of the information with which they normally work.

The LAK–supported approaches to assessment of learning assumes a technological layer that is capable of capturing, storing, managing, visualizing and processing big educational data – the millions of events occurring in diverse learning scenarios and platforms. Transformation of assessment practices to embrace and integrate learning analytics tools and strategies in support of teaching and learning, therefore, demands effective institutional technology infrastructures (Tynan and Buckingham Shum, 2013).
Fig. 2. Evaluation cycle of the OU Strategic Analytics Model (Tynan and Buckingham Shum, 2013).

This vision should be fulfilled if these needs are achieved at a macro level as well as micro level. The macro level has to aggregate information about the students learning experience at the institutional level to inform the management about the strategic priorities that will improve students’ retention and progression. On the other hand, a micro level assumes to use analytics to drive short, medium and long-term interventions (Buckingham Shum, 2013).

It is necessary to emphasize that the relationship between LAK and other pedagogical approaches has not been described in enough detail yet. The LAK methods are useful for data analysis, which has been obtained during the application of a particular pedagogical or didactical approach or teaching strategy. It is evident that such data depends on the chosen approach. Even though the correct interpretation of the results can be useful, and can help the teacher to find out many ways, how she can intervene to the learning process and how she can guide it appropriately. Subsequently, the final pedagogical intervention of the teacher influences the behaviour of the students. It results in the source data modifications, which serve as input to the LAK methods. The whole lifecycle of the LAK methods using in the learning process is closed (Drachsler and Greller, 2012).

**Systematic Approach to LAK Adoption in HEI.** The LAK adoption and implementation require a lot of procedural changes at different levels of HEI environment. Firstly, it is of great importance that institutions adopt learning analytics under clear guidelines that are grounded in cultural, social, economic and political contexts specific for each institution and are based on existing best practices for learning analytics and learning theories. The systemic approach to LAK which takes into account these assumptions is covered by the EU project “Supporting Higher Education to Integrate Learning Analytics (SHEILA).” It has been developed to assist European universities to become more mature users and custodians of digital data about their students as they learn online (Gashevic et al. 2014).

The main deliverable of the SHEILA project is a policy development framework that supports higher education institution in LAK adoption and implementation. This project uses
a modified framework RAPID (Research and Policy in Development) Outcome Mapping Approach (ROMA), which has the following six steps (Fig. 3) (Macfadyen et al., 2014):

1. map political context,
2. identify key stakeholders,
3. identify desired behaviour change,
4. develop engagement strategy,
5. analyse internal capacity to effect change, and
6. establish monitoring and learning frameworks.

Fig. 3. The RAPID Outcome Mapping Approach (ROMA) (Macfadyen et al., 2014).

ROMA was originally designed for policy engagement and influence in international development (Young and Mendizabal, 2009) ROMA serves as a policy and planning heuristic for learning analytics implementation. The six steps of ROMA model is focused on evidence-based policy change. It is designed to be used iteratively and to allow refinement and adaptation of policy goals and the resulting strategic plans over time and as contexts change, emphasizing the provisional nature of any solutions.

Importantly, the ROMA process begins with a systematic effort at mapping institutional context, the people, political structures, policies, institutions and processes that may help or hinder change. This critical activity allows institutions to identify the key factors specific to their own context that may influence (positively or negatively) the implementation process, and therefore also has the potential to illuminate points of intervention and shape strategic planning.

Other great contributions and relevant sources of up-to-date information related to the LAK adoption at HEI are available under the initiatives of EDUCASE, LACE, SOLAR as well as JISC. Their outcomes are partially included in this paper.

If the HEI would like to adopt LAK into its processes, it should consider the outcomes of the framework like SHEILA, as well as follow the next six key themes that should be included in an institutional policy (Tsai and Gasevic, 2017):

- Privacy and transparency
- Roles and responsibilities
- Objectives of learning analytics
– Risks and challenges
– Data management
– Research and data analysis

Moreover, according to Drachsler and Greller, the meaningful utilization of the LAK at HEIs assumes the cooperation of six essential dimensions (Drachsler and Greller, 2012): stakeholders, internal limitations, external constraints, objectives, data, and instruments (Fig. 4).

The modification of the individual parameter or representative, which characterize given dimension, has an influence on the other dimensions. This modification can cause changes in the observed outcomes. It means that only the balance between the various dimensions leads to the optimal utilization of the LAK methods in HEI.

Fig. 4. Dimensions of Learning analytics in the context of the learning process (Greller and Drachsler, 2012).

Nowadays, it is evident, that an adoption of LAK is a longlasting and iterative process. Although there is a growing interest in LAK among HEIs, the maturity levels of higher education institutions in terms of being “student data informed” are only at early stages (Wong, 2017). Most institutions are still in a preparatory or early stage of adoption, i.e. showing awareness of analytics and using some basic reports (Fig. 5). They are predominantly in the stages marked as Aware or Experimentation according to the general sophistication model of LAK (Fig. 5).
The issues related to the successful LAK adoption and implementation are very complex. Therefore, there are currently described only several case studies, which can be very useful for all other potential LAK implementators.

The Open University can be considered an institution, which is a leader in adopting LAK into its processes. Its research team proposed ten key areas that build the underpinning strengths required for the effective deployment of LAK at the university (Fig. 6).

It is evident based on figure 6 that the LAK adoption is an iterative process, which does not fit to all (Gašević et al., 2016). Considering the previously mentioned approaches, it is, therefore, necessary to effectively estimate the rate of the fulfilment of each key area from the perspective of resources, processes, people, responsibilities and their impact on the expected results.

In this context, HEI faces to the six challenges of LAK regarding the adoption of learning analytics (Tsai and Gasevic, 2017):

- Challenge 1: There is a shortage of leadership capabilities to ensure that implementation of learning analytics is strategically planned and monitored.
- Challenge 2: There are infrequent institutional examples of equal engagement with different stakeholders at various levels.
- Challenge 3: There is a shortage of pedagogy-based approaches to removing learning barriers that have been identified by analytics.
- Challenge 4: There are insufficient training opportunities to equip end users with the ability to employ learning analytics.
- Challenge 5: There are a limited number of studies empirically validating the impact of analytics-triggered interventions.
- Challenge 6: There is limited availability of policies that are tailored for learning analytics-specific practice to address issues of privacy and ethics as well as challenges identified above.
The challenges of implementing institution-wide learning analytics include both procedural concerns and practical applications. These challenges highlight the need to develop a comprehensive policy that meets the requirements of LAK and considers multiple dimensions including an institution’s context, stakeholders therein, pedagogical applications, institutional capacities, success evaluation, legal and ethical considerations, and a strategy that aligns with the institution’s missions (Tsai and Gasevic, 2016).

While the vision of improving student learning and assessment through the implementation of effective learning analytics tools and approaches holds promise, the further real challenges of implementation are significant: the challenge of influencing stakeholder understanding of assessment in education and the challenge of developing the necessary institutional, technological infrastructure to support the undertaking. Meanwhile, any such changes must coexist with the institution’s business as usual obligations (Head and Alford, 2013).

The difficulties in resolving these challenges are reflected by the fact that most of the HEIs are still in the “Experimentation”/”Organization” stage in the learning analytics sophistication model (Fig. 4). Many institutions have not defined clear strategies for learning analytics, whereas those that have implemented centrally-supported projects or planned to do so often initiated learning analytics under wider digitization strategies or teaching and learning strategies. Addressing the procedural concerns and practical applications, as well as establishing strategic leadership and monitoring are crucial for HEIs to move towards the stage of “Organizational Transformation” (Tsai and Gasevic, 2016).

Discussion and Conclusions. Pineda (2016) outlined some of the key trends surrounding organizational and technical patterns obtained through Readiness Assessments conducted as part of JISC’s learning analytics project (JISC, 2016). The Readiness Assessment process is designed to be collaborative and conducted onsite with a variety of key stakeholders across the institution. Pineda (2016) identified organizational and technical trends, which relate to the successful adoption and implementation at the HEI.

As she noted, from an organizational perspective, there is a high level of support for learning analytics. While there are concerns, as expressed above, the learning and teaching, academic staff, and student groups all feel strongly that learning analytics would be of benefit to the institutions. However, only if it will be implemented correctly and with the full participation of all required groups of stakeholders. On the other hand, the technical trends center on the availability, access, and use of data, as well as the staffing needs to deploy and maintain a learning analytics environment.

There are a lot of unanswered questions, which relate to the technical, ethical and legal issues of using the LAK methods. Many HEI will not have own capacity to implement a wide range of methods or approaches based on the LAK. Therefore, some predictions can be made, which suppose, the LAK will be delivered in the form of services, APIs, and cloud solutions. These predictions expect (Buckingham, 2011):

- The commoditization of analytics services and tools,
- Embedding of institutional analytics and diffusion of lessons learnt from the robust patterns,
- Emergence of analytics and recommendations engines grounded in theories of learning and sensemaking,
- Just in time intervention tools using analysis of online engagement,
- Larger personalization using demographic and previous study outcomes history,
- Multimedia indexing and logging for detecting their use by the students,
- Merging data from cloud applications,
- Analysing social networks used in the educational process.
However, as was mentioned before, to realize these predictions, the necessary shifts in the culture, technological infrastructure, and teaching practices of higher education, from assessment–for–accountability to assessment–for–learning should be achieved through piecemeal implementation of new LAK tools and services (Tsai and Gasevic, 2017).

Acknowledgement. The research leading to these results has received, within the framework of the IRNet project, funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/ under REA grant agreement No: PIRSES-GA-2013-612536.

REFERENCES


Buckingham Shum, S. (2013). Designing Systemic Learning Analytics at the Open University. August 2017,

https://www.slideshare.net/sbs/designing-systemic-learning-analytics-at-the-open-university


Duval, E. (2012). Learning Analytics and Educational Data Mining, Erik Duval’s Weblog, 30 January 2012, July 2017,

https://erikduval.wordpress.com/2012/01/30/learning-analytics-and-educational-data-mining/


JISC, Effective Learning Analytics, 5th UK Learning Analytics Network Event (2016), August 2017,

https://analytics.jiscinvolve.org/wp/2016/02/16/notes-and-presentations-from-5th-uk-learning-analytics-network-event-in-london/


Tynan, B., Buckingham Shum, S. (2013). Designing Systemic Learning Analytics at the Open University. Webinar in the SoLAR online open course Strategy & Policy for Systemic Learning Analytics.