Papamitsiou, Z., & Economides, A. (2014). Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence. *Educational Technology & Society*, *17* (4), 49–64.

Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence

Zacharoula Papamitsiou* and Anastasios A. Economides
Interdepartmental Progr. of Postgraduate Studies in Information Systems, University of Macedonia, 156 Egnatia Avenue, Thessaloniki, 54636, Greece // papamits@uom.edu.gr // economid@uom.gr
*Corresponding author

ABSTRACT

This paper aims to provide the reader with a comprehensive background for understanding current knowledge on Learning Analytics (LA) and Educational Data Mining (EDM) and its impact on adaptive learning. It constitutes an overview of empirical evidence behind key objectives of the potential adoption of LA/EDM in generic educational strategic planning. We examined the literature on experimental case studies conducted in the domain during the past six years (2008-2013). Search terms identified 209 mature pieces of research work, but inclusion criteria limited the key studies to 40. We analyzed the research questions, methodology and findings of these published papers and categorized them accordingly. We used non-statistical methods to evaluate and interpret findings of the collected studies. The results have highlighted four distinct major directions of the LA/EDM empirical research. We discuss on the emerged added value of LA/EDM research and highlight the significance of further implications. Finally, we set our thoughts on possible uncharted key questions to investigate both from pedagogical and technical considerations.

Keywords

Adaptive learning, Educational data mining, Empirical evidence, Learning analytics, Systematic review

Introduction

The information overload, originating from the growing quantity of "Big Data" during the past decade, requires the introduction and integration of new processing approaches into everyday objects and activities ("ubiquitous and pervasive computing") (Cook & Das, 2012; Kwon & Sim, 2013). Handling large amounts of data manually is prohibitive. Several computational methods have been proposed in the literature to do this analysis.

In commercial fields, business and organizations are deploying sophisticated analytic techniques to evaluate rich data sources, identify patterns within the data and exploit these patterns in decision making (Chaudhuri, Dayal & Narasayya, 2011). These techniques combine strategic planning procedures with informational technology instruments, summarized under the term "Business Intelligence" (Eckerson, 2006; Jourdan, Rainer & Marshall, 2008). They constitute a well-established process that allows for synthesizing "vast amount of data into powerful decision making capabilities" (Baker, 2007, p. 2).

Recently researchers and developers from the educational community started exploring the potential adoption of analogous techniques for gaining insight into online learners' activities. Two areas under development oriented towards the inclusion and exploration of big data capabilities in education are Educational Data Mining (EDM) and Learning Analytics (LA) and their respective communities.

EDM is concerned with "developing, researching, and applying computerized methods to detect patterns in large collections of educational data that would otherwise be hard or impossible to analyze due to the enormous volume of data within which they exist" (Romero & Ventura, 2013, p. 12). Respectively, LA is an area of research related to business intelligence, web analytics, academic analytics, action analytics and predictive analytics. According to the definitions introduced during the 1st International Conference on Learning Analytics and Knowledge (LAK), LA is "the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and environments in which it occurs" (https://tekri.athabascau.ca/analytics/).

Explaining the previous definitions, LA and EDM constitute an ecosystem of methods and techniques (in general procedures) that successively gather, process, report and act on machine-readable data on an ongoing basis in order to advance the educational environment and reflect on learning processes. In general, these procedures initially emphasize on measurement and data collection and preparation for processing during the learning activities. Next, they focus on further analysis, reporting of data and interpretation of results, targeting to inform and empower learners, instructors and organization about performance and goal achievement, and facilitate decision making accordingly.

Both communities share similar goals and focus where learning science and data-driven analytics intersect. However, they differ in their origins, techniques, fields of emphasis and types of discovery (Chatti et al., 2012; Romero & Ventura, 2013; Siemens & Baker, 2012). Romero and Ventura (2013) presented an up-to-date comprehensive overview of the current state in data mining in education. In their overview, the authors do not present research results as empirical evidence. Their focus targets on the objectives, methods, knowledge discovery processes and tools adopted in EDM research. Analogous attempts were presented by Ferguson (2012) and Bienkowski et al. (2012) in the state of LA in 2012 and in an issue brief respectively.

All of these previous studies claim that as far as it concerns the approach to gaining insights into learning processes, LA adopts a holistic framework, seeking to understand systems in their full complexity. On the other hand, EDM adopts a reductionistic viewpoint by analyzing individual components, seeking for new patterns in data and modifying respective algorithms. In other words, these two research areas are complementary and in order to capture the whole picture, someone should follow their traces alongside each other.

Motivation and rationale of the study

The motivation for this review derived from the fact that empirical evidence is required for theoretical frameworks to gain acceptance in the scientific community. A search in relevant literature did not reveal any review of empirical evidence of the added value of research in both domains. Consequently, there was a need to supply the audience with an accredited overview. This paper aims to fill that gap.

The value of any single study is derived from how it fits with and expands previous work, as well as from the study's intrinsic properties. Thus, putting together all the unbiased, credible results from previous research would be a step towards understanding the whole picture and construct a map of our knowledge on the domain. In a sense, the rationale of our study was to manage the overwhelming amount of publications through a critical exploration, evaluation and synthesis of the previous empirical results that worth reflection.

This paper's goal is to carry out a systematic review of empirical evidence in order to contribute towards:

- a complete documentation of the applied research approaches so far,
- a feasibility study that captures the strengths and weaknesses of research in the domain, and
- the identification of possible threats, and thus motivate the research community redefine or refine related questions or hypotheses for further research (opportunities).

The research questions

The following research questions need to be addressed, and are distinguished into primary (generalized: set to fulfill the goals of the review) and secondary (sub-objectives/specific: refine the primary - explanatory):

RQ1 (Primary)-*Research Objectives*: *Which are the basic research objectives of LA/EDM so far (in terms of measurable metrics), and which methods do researchers follow to achieve these goals?*

- **RQ1.1 (Secondary) - *Efficacy of implementation*:** *What are the significant results from previous research that constitute empirical evidence regarding the impact of LA/EDM implementation?*

- **RQ1.2 (Secondary) - *Interpretation of the results*:** *What do these results indicate regarding the added value of this technology?*

RQ2 (Primary) - *Future challenges*: *Which other emerging research technologies should be explored through the LA/EDM viewpoint?*

Research methodology

The followed methodology qualifies this article as a systematic qualitative review of empirical research results concerning LA/EDM (Okoli & Schabram, 2010).

In order to conduct the literature review we defined a review protocol, consisting of four discrete stages: a) searching the literature – data collection, b) reviewing and assessing the search results – selection of primary studies, c) analyzing, coding and synthesizing the results, and d) reporting the review.

During the first stage, our goal was to collect the appropriate studies. For that reason, we determined and accessed the article pool and declared the key search terminology. We extensively and iteratively searched international databases of authoritative academic resources and publishers, including Scopus, ERIC, Google Scholar, Science Direct, DBLP and ACM Digital Library. We also scanned International Journals and selected Conference Proceedings. The search terms included *learning analytics, learning analytics tools, learning analytics case studies, educational data mining, knowledge discovery in education*. The search process spanned from March 2013 to August 2013. The time frame of the search was bound within the last six years (2008-2013), in which emergence and adoption of LA/EDM has grown.

Due to the orientation of our work towards the practical implementation and exploitation of LA/EDM, at the end of the data collection stage, we explicitly determined the article inclusion/exclusion criteria (Table 1).

Table 1. Inclusion/exclusion criteria

Include	Exclude
• Articles published in Journals with Impact Factor • Full-length articles published in International Conference/Workshop Proceedings • Present quantitative results • Date from 2008 to 2013	• Articles that do not present empirical data (e.g. theoretical and conceptual articles, essays, tool demonstration, etc.) • Short papers from conferences/workshops • Book chapters

The search procedure, and after deleting the duplicate records, yielded 209 results. 40 of them are published in International Journals and 169 articles were presented at International Conferences. Then we assessed the quality of the collected literature according to the following rigorous quantitative/qualitative rules:

- Number of citations
- Degree of underlying innovation (e.g., significant changes in techniques/equipment or software, as proposed by UNESCO: http://www.uis.unesco.org/ScienceTechnology/Pages/st-data-collection-innovation.aspx)
- In depth illustration of the followed methodology (e.g., clear settings, fully-explained experimental procedure, etc.)
- Sufficient presentation of the findings (e.g., analytical discussion of findings and interpretation of results, use of figures and tables when needed, etc.)

Throughout the assessment procedure we identified that among the 209 retrieved articles, 40 of them were considered more central to our review (*key studies*), based on the combination of the above rules.

Next, we proceeded on with an article classification according to the adopted research strategy (category), research discipline (topic), learning settings, research objectives (goals), data gathering (sources and data-types) and analysis technique (method) and results. Finally, we used non-statistical methods to evaluate and interpret findings of the collected studies, and conduct the synthesis of this review.

Limitations

Although there are papers concerning EDM that are published before 2008, this year was a landmark for the independent growth of this research domain (1st Conference in EDM - http://www.educationaldatamining.org/EDM2008/). For that reason we decided to examine the literature on experimental case studies conducted in the domain from 2008 to 2013. Furthermore, indicative reviews on former work can be found in Romero and Ventura (2007) and Romero and Ventura (2010).

We should note that, despite the fact that appreciated research works have been published in respectable conferences like Intelligent Tutoring Systems (ITS), Artificial Intelligence in EDucation (AIED) and User Modeling, Adaptation and Personalization (UMAP), in this review we included published papers only from the EDM and LAK conferences. We acknowledge the leadenness of the previously mentioned conferences, but, in this work, we wanted to isolate the LA/EDM research and focus on its strengths, weaknesses, opportunities and threats.

We should also mention that the papers presented at the 6th International Conference on Educational Data Mining were excluded from the review process, since at the time of the search process these articles had had no citations. However, we decided to include recently published Journal articles (published in 2013) with lower number of citations, as indicative of current trends in the domain.

Results

In this section, we present our findings based on the analysis of the published case studies. We used non-statistical methods to evaluate and interpret findings of the collected studies.

According to the followed *research strategy*, most of the published case studies are exploratory or experimental studies. Some of them are evaluation studies, while others are empirical studies or surveys. Furthermore, the *research topics* differ from study to study, but most of them focus on science, technology, engineering, and mathematics (STEM).

Based on the *learning settings* of the studies (illustrated in Table 2), most studies are conducted within Virtual Learning Environments (VLEs) and/or Learning Management Systems (LMSs). Other popular learning settings are Cognitive Tutors (CTs), computer-based and web-based environments, mobile settings, and more recently, Massive Open Online Courses (MOOCs) and social learning platforms.

Table 2. Classification of case studies according to the learning settings

Learning setting[*]	Authors & Year (Paper Ref.)
VLEs / LMSs[a]	Lin, Hsieh & Chuang, 2009; Lykourentzou et al., 2009a; Lykourentzou et al., 2009b; Macfadyen & Dawson, 2010; Merceron & Yacef, 2008; Romero et al., 2008; Romero-Zaldivar et al., 2012; Tanes et al., 2011
MOOC/social learning[b]	Clow & Makriyiannis, 2011; Fournier et al., 2011; Kizilcec et al., 2013
Web-based education[c]	Abdous, He & Yen, 2012; Giesbers et al., 2013; He, 2013; Khribi et al., 2009; Li et al., 2011; Romero et al., 2009
Cognitive tutors[d]	Baker et al., 2008; Moridis & Economides, 2009; Pardos et al., 2013; Shih, Koedinger & Scheines, 2008
Computer-based education[e]	Ali et al., 2012; Barla et al., 2010; Blikstein, 2011; Jeong & Biswas, 2008; Levy & Wilensky, 2011; Santos et al., 2012; Thai-Nghe et al., 2011
Multimodality[f]	Worsley & Blikstein, 2013
Mobility[g]	Chen & Chen, 2009; Leong et al., 2012

Note. [a]VLEs/LMSs: *controlled environment (VLE/LMS), used for gathering learner and activity data,* [b]MOOC/social learning: *informal, social learning setting,* [c]Web-based education: *web-based e-learning environments except from VLEs, LMSs and MOOCs,* [d]Cognitive tutors: *special software, utilized for the needs of the study,* [e]Computer-based education: *other environments that include some type of computer technology (e.g. desktop applications, etc.) except from those belonging to one of the other categories,* [f]Multimodality: *learner data in different modalities,* [g]Mobility: *mobile devices used as the primary learning mediator.*

The authors *gathered data* from different *data sources*, including log files from the goal-oriented implemented systems, questionnaires, interviews, Google analytics, open datasets from the dataTEL Challenge (http://www.teleurope.eu/pg/groups/9405/datatel/), virtual machines, and many more. In particular, researchers tracked different *types of data* in order to measure students' participation and login frequency, number of chat messages between participants and questions submitted to the instructors, response times on answering questions and solving tasks, resources accessed, previous grades, final grades in courses, detailed profiles, preferences from LMSs, forum and discussion posts, affect observations (e.g. bored, frustrated, confused, happy, etc.) and many more.

Another important parameter is the *data mining method* adopted by authors *to analyze* the gathered data. In the field of LA/EDM, the most popular method is classification, followed by clustering, regression (logistic/multiple) and more recently, discovery with models. In addition, *algorithmic criteria* computed for comparison of methods include precision, accuracy, sensitivity, coherence, fitness measures (e.g. cosine, confidence, lift, etc.), similarity weights, etc. Table 3 displays the classification of the key studies according to the data mining method they adopt.

Table 3. Classification of case studies according to the analysis method

Data analysis method	Authors & Year (Paper Ref.)
Classification	Baker et al., 2008; Barla et al., 2010; Chen & Chen, 2009; Dejaeger et al., 2012; Dekker et al., 2009; Jeong & Biswas, 2008; Guruler et al. 2010; Guo, 2010; Huang & Fang, 2013; Khribi et al., 2009; Kizilcec et al., 2013; Klašnja-Milićević et al., 2011; Li et al., 2011; Lin et al., 2013; Lykourentzou et al., 2009a; Lykourentzou et al., 2009b; Moridis & Economides, 2009; Pardos et al., 2013; Romero et al., 2008; Thai-Nghe et al., 2011
Clustering	Abdous, He & Yen, 2012; Chen & Chen, 2009; Khribi et al., 2009; Kizilcec et al., 2013; Klašnja-Milićević et al., 2011; Lykourentzou et al., 2009b; Romero et al., 2009
Regression	Abdous, He & Yen, 2012; Macfadyen & Dawson, 2010; Romero-Zaldivar et al., 2012
Text mining	He, 2013; Leong et al., 2012; Lin, Hsieh & Chuang, 2009
Association rule mining	Merceron & Yacef, 2008; Romero et al., 2009
Social Network Analysis	Fournier et al., 2011; Macfadyen & Dawson, 2010
Discovery with models	Ali et al., 2012; Pardos et al., 2013; Shih, Koedinger & Scheines, 2008
Visualization	Clow & Makriyiannis, 2011; Fournier et al., 2011; Santos et al., 2012
Statistics	Giesbers et al., 2013; Guo, 2010

The article classification according to *the research objectives (goals)* is illustrated in Table 4. As seen in this table, the majority of studies investigate issues related to student/student behavior modeling and prediction of performance, followed by increase of students' and teachers' reflection and awareness and improvement of provided feedback and assessment services.

Table 4. Classification of case studies according to the research objectives

Research objectives (goals)	Authors & Year (Paper Ref.)
Student/Student behavior modeling	Abdous, He & Yen, 2012; Baker et al., 2008; Blikstein, 2011; Fournier et al., 2011; He, 2013; Jeong & Biswas, 2008; Kizilcec et al., 2013; Levy & Wilensky, 2011; Li et al., 2011; Pardos et al., 2013; Romero et al., 2008; Shih, Koedinger & Scheines, 2008
Prediction of performance	Abdous, He & Yen, 2012; Huang & Fang, 2013; Lykourentzou et al., 2009b; Macfadyen & Dawson, 2010; Moridis & Economides, 2009; Pardos et al., 2013; Romero et al., 2008; Romero-Zaldivar et al., 2012; Shih, Koedinger & Scheines, 2008; Thai-Nghe et al., 2011
Increase (self-) reflection & (self-) awareness	Ali et al., 2012; Clow and Makriyiannis, 2011; Fournier et al., 2011; Macfadyen & Dawson, 2010; Santos et al., 2012
Prediction of dropout & retention	Dejaeger et al., 2012; Dekker et al., 2009; Giesbers et al., 2013; Guo, 2010; Guruler et al. 2010; Kizilcec et al., 2013; Lykourentzou et al., 2009a
Improve assessment & feedback services	Ali et al., 2012; Barla et al., 2010; Chen & Chen, 2009; Leong et al., 2012; Tanes et al., 2011; Worsley & Blikstein, 2013; Wilson et al., 2011
Recommendation of resources	Khribi et al., 2009; Klašnja-Milićević et al., 2011; Romero et al., 2009; Thai-Nghe et al., 2011; Verbert et al., 2011

Key studies analysis

In this section we present the findings of the review process and answer on the initially set research questions RQ1 and RQ1.1. The rest of the research questions (mostly the results of the case studies and their comparative evaluation, as well as current and future trends, possible gaps and new research directions) are discussed in next section.

RQ1: *Which are the basic research objectives of LA/EDM so far (in terms of measurable metrics), and which methods do researchers follow to achieve these goals?*

Student/student behavior modeling

As seen from table 4, detection, identification and modeling of students' learning behavior is a primary research objective. More specifically, the authors seek to identify learning strategies and when they occur, and model affective and metacognitive states (Abdous, He & Yen, 2012; Baker et al., 2008; Blikstein, 2011; Jeong & Biswas, 2008; Levy & Wilensky, 2011; Shih, Koedinger & Scheines, 2008). For example, Abdous, He and Yen (2012) and He (2013) tried to correlate interactions within a Live Video Streaming (LVS) environment to students' final grades in order to predict their performance, discover behavior patterns in LVSs that lead to increased performance, and understand the ways students are engaged into online activities. In another case study, Blikstein (2011) logged automatically-generated data during programming activity in order to understand students' trajectories and detect programming strategies within Open-Ended Learning Environments (OELEs). Furthermore, Shih, Koedinger and Scheines, (2008) used worked examples and logged response times to model the students' time-spent in terms of "thinking about a hint" and "reflecting on a hint" for capturing behaviors that are related to reasoning and self-explanation during requesting hints within a CT environment. In another self-reasoning example, Jeong and Biswas (2008) tried to analyze students' behavior based on the sequence of actions, and to infer learning strategies within a teachable agent environment.

Another orientation is the discovery and modeling of the respective behaviors within MOOCs (Fournier et al., 2011; Kizilcec et al., 2013). The authors tried to identify meaningful, high-level patterns of participation, engagement and disengagement in learning activities in this recently introduced learning setting.

Updated and extended review on student modeling approaches can be found in Chrysafiadi and Virvou (2013) and in Pena-Ayala (2014).

Prediction of performance

Authors also explore, identify and evaluate various factors as indicators of performance for prediction purposes. Among these factors, demographic characteristics, grades (in pre-requisite courses, during assessment quizzes and their final scores), students' portfolios, multimodal skills, students' participation, enrollment and engagement in activity and students' mood and affective states are acknowledged as the most common ones (Abdous, He & Yen, 2012; Huang & Fang, 2013; Lykourentzou et al., 2009b; Macfadyen & Dawson, 2010; Moridis & Economides, 2009; Pardos et al., 2013; Romero-Zaldivar et al., 2012). For example, Macfadyen and Dawson (2010) examined the effect of variables tracked within an LMS-supported course (e.g., total number of discussion messages posted, total time online, number of web links visited, etc.) on students' final grade. In another example, Lykourentzou et al., (2009b) used neural networks to accurately cluster students at early stages of a multiple choice quiz activity.

Moreover, researchers investigated affective factors that influence learning outcomes (within an ITS) (Pardos et al., 2013) and used simulation environments (virtual appliances that appear to learners as regular desktop applications) to monitor students and predict their performance (Romero-Zaldivar et al., 2012). Pardos et al. (2013) employed discovery with models, post-hoc analysis of tutor logged data and sensor-free detectors of affect (based on classification algorithms), while Romero-Zaldivar et al. (2012) tracked events (such as work-time, commands, compile, etc.) and analyzed the gathered data with multiple regression for the estimation of the variance of performance.

Increase (self-)reflection and (self-)awareness

Another crucial issue in EDM/LA research that authors attempt to address is how to increase the instructors' awareness, identify "disconnected" students and evaluate visualizations regarding their capabilities on informing students about their progress and compared to peers. In order to provide instructors with pedagogically meaningful information and to help them extract such information on their own, the researchers embedded multiple representations of feedback types (Ali et al., 2012) and multiple widget technology for personalization of learning environments (Santos et al., 2012). Alternatively, content analysis for threaded discussion forums was explored regarding its monitoring capabilities (Lin et al., 2013). In particular, the authors aimed to facilitate the automated coding process within a repository of postings in an online course, in order less monitoring of the discussion to be needed by the instructor. Furthermore, Merceron and Yacef (2008) employed association rule mining to extract meaningful association rules to inform teachers about usage of extra learning material. The authors investigated students' usage of learning resources and self-evaluation exercises and its possible impact on final grades.

In the context of social/open learning, the researchers explored the usefulness and motivation capabilities of dashboard-like applications regarding their self-reflection and self-awareness opportunities (Clow & Makriyiannis, 2011). In particular, the request was related to the effect of "expert users" presence on participants' awareness of their own contribution and participation in an online reputation system with positive feedback only. Furthermore, Fournier et al. (2011) searched for crucial moments of learning based on interactions in MOOCs. In this case, the authors examined the impact of visualized provision of useful information to social learners regarding their participation and social interactions.

Prediction of dropout and retention

Prediction of dropout and retention are also key issues for LA/EDM research. In order to predict students' dropout at early stages, Lykourentzou et al. (2009a) applied a combination of three machine learning techniques on detailed students' profiles from an LMS environment. The authors compared the accuracy, sensitivity and precision measures of the proposed method to others in literature. From a similar point of view, Dekker et al. (2009) tried to predict students' dropout and identify factors of success based on the use of different classification algorithms. The authors compared the accuracy and performance of these algorithms to make a selection between classifiers. In particular, they used classifiers for prediction of dropout based on simple "early" data (from first year enrollment) and boosted accuracy with cost-sensitive learning.

More recently, Kizilcec et al. (2013) classified learners according to their interactions (video lectures and assessment) with course content in learning activities in MOOCs. Next, they clustered engagement patterns, and finally, they compared clusters based on learners' characteristics and behavior.

The issue of motivating engagement in learning activities and consequently increasing students' satisfaction and retention was also explored (Dejaeger et al., 2012; Giesbers et al., 2013; Guo, 2010; Guruler et al., 2010). Demographics and factors like achievement rates and final performance were associated to students' motivation to remain engaged and actively enrolled in courses. Identification of success factors urged Giesbers et al. (2013) to investigate the relationship between observed student behavior (i.e., actual usage of synchronous tools), motivation, and performance on a final exam. The researchers explored whether actual usage of synchronous tools increases the motivation to participate in online courses that support these tools. Similarly, Guo (2010) used statistical measures and neural network techniques for prediction of students' retention. The researcher examined the number of students enrolled in each course and the distinction rate in final grades. Furthermore, Dejaeger et al. (2012) explored measures of students' satisfaction for retaining student population. The authors investigated a number of constructs of satisfaction (e.g., perceived usefulness of training, perceived training efficiency, etc.) along with class related variables.

Improve feedback and assessment services

Many researchers have explored the use of LA/EDM in producing meaningful feedback. Feedback is strongly related to reflection and awareness and could be informative regarding students' dropout intentions. For that reason, provision of appropriate forms/types of feedback was a major issue for Ali et al. (2012), Clow and Makriyiannis (2011) and Macfadyen and Dawson (2010), formerly presented. Visualization of feedback was also crucial for Tanes et al. (2011). The authors explored instructors' perceptions of feedback types in relation to students' success.

Complementary to that, in the mobile learning context, Leong et al. (2012) explored the impact and usefulness of SMS free-text feedback to teacher regarding the feelings of students, after a lecture. Their goal was to visualize positive and negative aspects of the lecture by taking advantage of the limited SMS length and the use of emoticons in order to provide free-text feedback to teacher.

In addition to these studies, an extensive area of LA/EDM research deals with issues related to using LA/EDM for adaptive assessment of goal achievement during activities. The landscape in this domain is quite distributed and diverse. Selection of the most appropriate next task during adaptive testing, students' satisfaction level during mobile formative assessment, as well as construction of sophisticated measures of assessment (Barla et al., 2010; Chen and Chen, 2009; Wilson et al., 2011; Worsley & Blikstein, 2013) have emerged. Barla et al. (2010) focused on assessment capabilities of EDM methods and combined three different classification methods for selection of the most appropriate next task during adaptive testing. In a different context, Chen and Chen (2009) developed a tool that uses six computational intelligence theories according to the web-based learning portfolios of an individual learner, in order to measure students' satisfaction during mobile formative assessment. Furthermore, Worsley and Blikstein (2013) aimed to detect metrics that could be used primarily as formative assessment tools of sophisticated learning skills acquisition in process-oriented assessment. A combination of speech recognition with knowledge tracing was proposed by the authors as method for multimodal assessment.

Recommendation of resources

Another major issue in dataset-driven research concerns data resources and their management. Research in this domain focuses on a technical aspect. The approaches include similarity calculation mechanisms deployment, comparison of the performance of different mining algorithms, aggregation of different datasets in the context of dataset-driven research, suggestion of infrastructures for storing and forwarding learning-resources metadata (Romero et al., 2009; Thai-Nghe et al., 2011; Verbert et al., 2011) for resource recommendation in larger scale and across different contexts. Examples of algorithmic approaches also include recommendations according to the affective state of the learner (Santos & Boticario, 2012), implementation of collaborative filtering to sequence learning activities, hybrid recommendations based on learner and content modeling (Khribi et al., 2009; Klašnja-Milićević et al., 2011) and more.

Verbert et al. (2011) presented an analysis of publicly available datasets for TEL that can be used for LA in order to support recommendations (of resources or activities) for learning. In addition, the authors evaluated the performance of user-based and item-based collaborative filtering algorithms and measured their accuracy and coverage through metrics implementation. Moreover, Romero et al. (2009) explored user profile information and web-usage mining for recommendation of resources (here, hyperlinks). The authors compared the performance of three different mining algorithms.

A comprehensive review on recommender systems in the TEL context can be found in Manouselis et al. (2013).

RQ1.1: *What are the significant results from previous research that constitute empirical evidence regarding the impact of LA/EDM implementation?*

According to the research objectives explored by the authors, Table 5a displays a categorization of the algorithmic-oriented findings from the collected studies.

Table 5a. Classification of the results of LA/EDM case studies (algorithmic)

Objective	Results
Student/ student behavior modeling	• Quantitative analysis could be applied for reporting on participants' activity, while qualitative analysis could be more effective on revealing deeper concepts related to learning (Fournier et al., 2011). • Comprehensibility of the results strongly depends on human judgment - produced models are not equally interpretable by the teachers (Fournier et al., 2011; Romero et al., 2008).
Prediction of performance	• Adding more predictor variables does not help improve the average prediction accuracy of the mathematical models explored by Huang and Fang (2013) for prediction of performance. However, neural networks method leads to better prediction results compared to those of the regression analysis method (Lykourentzou et al., 2009b).

	• Reducing the size of the training set by removing very high and very low probabilities of obtaining a correct answer without knowing the skill and obtaining an incorrect answer even though the student knows the skill, and forecasting techniques that embed sequential information (temporality) into the factorization process may improve the predictive model of students' performance (Baker et al., 2008; Thai-Nghe et al., 2011).
Increase (self-) reflection & (self-) awareness	• Genre classification methods can automate the coding process in a forum and handle issues like imbalanced distribution of discussion postings (Lin, Hsieh & Chuang, 2009) • LMS log data are not data mining "friendly" (i.e., not stored the same way, data consolidation requires complex manipulations, etc.) (Merceron & Yacef, 2008). • Comparison of measures of interestingness of association rules did not significantly improved decision making for discarding a rule (Merceron & Yacef, 2008).
Prediction of dropout & retention	• Combination of machine learning techniques afforded more reliable results, which depend on the level of detail of available students' data (Lykourentzou et al., 2009a). • Simple classifiers had higher accuracy than sophisticated ones and cost-sensitive learning helps to bias classification errors (Dekker et al., 2009). • While investigating disengagement in MOOCs, the cross-cluster comparison can help understanding the reasons why learners remain to a cluster (Kizilcec et al., 2013).
Recommendation of resources	• A combination of students' clustering and sequential pattern mining is suitable for the discovery of personalized recommendations (Romero et al., 2009), while content based filtering and collaborative filtering approaches are valid recommendation strategies (Khribi et al., 2009), but further research should be conducted (Verbert et al., 2011)

Table 5b displays a categorization of the pedagogy-oriented findings. The learning context of the studies has been taken under consideration, as well. That is because we wanted to maintain the targeted applicability of the results.

Table 6b. Classification of the results of LA/EDM case studies (pedagogical)

Objective	Results	
	Formal Learning	Non-Formal Learning
Student/ student behavior modeling	• Students' detected critical moments during programming reflect students' behavior and their perceived learning benefits, both in Secondary and Higher Education (Blikstein, 2011; Levy & Wilensky, 2011). • In secondary education, learning by teaching provides better opportunities for retaining metacognitive learning strategies (Jeong & Biswas, 2008), while worked examples are effective indicators of self-explanation and learning gain (Shih, Koedinger & Scheines, 2008). • Specifying the moments that teacher should intervene requires to better distinguish between students who use worked examples, how they use them and their response times (Shih, Koedinger & Scheines, 2008).	• The presence of "experts" (that is users or organizations with advanced expertise or reputation on the field of study) has a significant impact on the highly unequal distribution of activities within a functioning social network (Clow & Makriyiannis, 2011). • At a level of interactivity among learners or between learners and teachers, questions posed to instructors and chat messages posted among students (both in number and their content) are correlated (Abdous, He & Yen, 2012). • The discovery of four trajectories (auditing, completing, disengaging, sampling learners) (Kizilcec et al., 2013) roughly describes engagement that makes sense in MOOCs.
Prediction of performance	• Both in Secondary and Higher Education, the number of quizzes passed is the main determinant of performance (i.e., the final grade), while others, such as number of posts, frequencies of the events and time-spent could identify activities that are related to higher or lower marks (Romero et al., 2008; Romero-Zaldivar et al., 2012; Shih, Koedinger & Scheines, 2008).	• Abdous, He and Yen (2012) couldn't predict performance based on students' participation and online interactions. • Giesbers et al. (2013) and Macfadyen and Dawson (2010) found a significant positive relationship between participation and grades.

	• In Secondary Education, engaged concentration and frustration are correlated with positive learning outcomes, while boredom and confusion are negatively correlated with performance (Pardos et al., 2013).	
Increase (self-) reflection and (self-) awareness	• In Secondary Education, and from the instructors' perspective, coding discussion posts in a forum can assist the teacher to automatically monitor the forum and maintain its quality (Lin, Hsieh & Chuang, 2009). • In Higher Education, meaningful rules increases teachers' awareness regarding the students' usage of additional material within LMSs (Merceron & Yacef, 2008). • SMS text increases instructor's awareness on students' affective states in order to modify the lecture (Leong et al., 2012).	• Dashboard-like applications and multiple feedback representations could increase (self-)awareness and perceived value of provided feedback (Ali et al., 2012; Macfadyen & Dawson, 2010; Santos et al., 2012). • From the learners' point of view, students want to be aware of what their peers are doing, but they don't like to be tracked outside a course environment due to privacy concerns (Santos et al., 2012). • Identification of disconnected students based on their networking activity ended up to clusters of students with similar participatory behavior (Macfadyen & Dawson, 2010).
Prediction of dropout and retention	• Monitoring students' activity with virtual machines and applying data-driven machine learning methods on students' profiles and log files (mostly grades and assessment quiz scores) from LMS databases allow detecting students at-risk at an early stage (Lykourentzou et al., 2009a; Romero-Zaldivar et al., 2012). • Students want to feel that they belong to the course in order to engage and enroll (Guo, 2010). Improving students' course satisfaction can be used to reduce students' dropout (Dejaeger et al., 2012; Guo, 2010).	• In MOOCs, the most common detected reasons for disengagement were personal commitments, work conflict and course overload (Kizilcec et al., 2013). • In web-videoconference settings there was not found a relation between motivation to participate and dropout (Giesbers et al., 2013). • Types of registration to the university as well as the family income seem to affect more the students' retention (Guruler et al., 2010).
Improve feedback and assessment services	• In Higher Education, adaptive selection of the most appropriate next task improved testing outcomes mostly for below-average students (Barla et al., 2010). • In Elementary Education, web-based learning portfolios of an individual learner during mobile formative assessment granted similar results to those of summative assessment (Chen & Chen, 2009).	
Recommendation of resources	• Additional learner attributes (e.g., experience level indicators, learning interests, learning styles, learning goals and competences and background information), student's expected performance on tasks, his recent navigation history (within a number or resources) or learner's affective traits should be taken under consideration in recommendation processes (Khribi et al., 2009; Klašnja-Milićević et al., 2011; Santos & Boticario, 2012; Thai-Nghe et al., 2011; Verbert et al., 2011).	

Discussion and future research

From the former analysis it becomes apparent that recently, the educational research community has started applying sophisticated algorithmic methods on gathered (mostly raw) data for understanding learning mechanisms through an in-depth exploration of their relations and meaning. As seen in Tables 5a and 5b, the landscape of the LA/EDM research combines diverse and often conflicting aspects and results related to gaining insight into learning processes. However, the above results have highlighted four distinct major axis of the LA/EDM empirical research including:

- *Pedagogy-oriented issues* (e.g., student modeling, prediction of performance, assessment and feedback, reflection and awareness): several studies focus on pedagogically meaningful analysis on collected students' data in order to shed light to the whole picture from students/students' behavior modeling to self-regulated learning.

- *Contextualization of learning* (e.g., multimodality, mobility, etc.): a number of studies gathered data from the learning context itself and focus on positioning learning within specific conditions and attributes.

- *Networked learning* (e.g., MOOCs, social learning platforms, etc.)*:* some case studies try to identify patterns within the social aspect of learning and the MOOCs, where the number of participants rapidly increases and the interactions between learners and the learners and the content are text/video-based.

- *Educational resources handling:* fewer, but not neglected studies raise the issue of organizing and recommending educational resources from data pools, and selecting the most appropriate algorithmic method for making suggestions.

However, these four axis are not completely autonomous, since significant overlaps may occur. For example, student modeling (i.e. a pedagogy-oriented issue) can still be explored in MOOCs (i.e., a form of Networked learning). However, this statement could only constitute a limitation which does not deduce the added value of the findings.

RQ1.2: *What these results indicate regarding the added value of this technology?*

One of the most important goals of the systematic review was to reveal the added value of the field explored. From the above analysis of findings derives that analysis of user interactions in order to "control" the information generated through technology has always been a request. LA/EDM research results indicate that data integration from multiple sources can improve the accuracy of a learner profile and subsequent adaptation and personalization of content. Exploration of students' behavior within educational contexts that support multimodality and mobility could lead to shaping a holistic picture of how, when and where learning occurs.

Researchers set the educational context within limits in which previously it was almost impossible to infer behavior patterns, due to their high levels of granularity. In such advanced learning contexts, LA/EDM research community determines simple and/or sophisticated factors as predictors of performance and explores their predictive value and capabilities by tracking actual data and changes on behavioral data. The goal is to identify the most significant factors in order to develop better systems. These systems will allow students to monitor their own progress and will help them evaluate and adjust their learning strategies to improve their performance in terms of learning outcomes.

Moreover, the social dimension of learning and the opportunity of selectively participating in MOOCs are also explored with encouraging results. Consequently, the research community could gain insight into the learning mechanisms that previously were a "black box."

RQ2: *Which other emerging research technologies should be explored through the LA/EDM viewpoint?*

Complementary, the literature overview has revealed a number of unexplored issues in this rapidly grown domain, including (but not limited to) the following:

Suggested incorporation of other emerging research technologies with LA/EDM

Game-based learning (GBL) has been acknowledged for its positive impact on learners. According to Collony et al. (2012, p.1), "playing computer games is linked to a variety of perceptual, cognitive, behavioral, affective and motivational impacts and outcomes." One interesting research question is if and how LA/EDM methods could be applied to report and visualize learning processes during GBL. In other words, how can LA/EDM be applied on GBL to detect patterns and construct measures that are transferable to other OELEs, in order to assess advanced skills development.

Another field evolving in a rapid pace is mobile and ubiquitous learning. Mobile learning has been acknowledged for the unique opportunity of offering authentic learning experiences anytime and anywhere (Tatar et al., 2003). Although two of the selected studies were conducted in a mobile context (Chen and Chen, 2009; Leong et al., 2012), none of them associated or explored the effect of the context on the attained results. LA/EDM research could investigate the appropriateness of the popular methods in the above context in order to provide sophisticated, personalized learning services through mobile applications.

Furthermore, according to Piaget's theory of cognitive development, sensorimotor learning is the first stage of human learning (Piaget, 1952). Sensorimotor learning refers to improvement, through practice, in the performance of sensory-guided motor behavior (Krakauer and Mazzoni, 2011). Due to its high relevance to the brain anatomy and functionality, sensorimotor learning has recently been under the lenses of neuroscience research (e.g., Catmur, 2013). LA/EDM has not been previously examined for sensorimotor learning or combined to neuroscience research. It would be interesting to study transformation of learning experience into strategy development (knowledge transfer) by exploring big neuroscience data.

Technology acceptance is also a well addressed issue in educational research. Regarding learning analytics acceptance, Ali et al. (2012) proposed a model that considers only two parameters – ease of use and perceived usefulness. However, more parameters should be explored in order to create a reliable learning analytics acceptance model. An appreciated model for computer based assessment acceptance was proposed by Terzis and Economides (2011). Researchers from the LA/EDM domain could also examine respective models that are suitable for the purposes of LA tools.

Finally, the review process didn't yield any article related to learning "meta"-analytics (i.e., feeding machine readable results from the LA/EDM procedures to another data-driven system for diving decision making without the mediation of the human judgment parameter). It would be interesting to take advantage of the plethora of results from LA/EDM research towards introducing innovative intelligent tutoring systems or fully automated educational recommender systems.

Conclusions

Previous literature reviews on LA/EDM research provided significant insight into the conceptual basis of this rapidly growing domain. However, these studies did not conduct an analysis of actual research results. The current paper presents a systematic review of empirical evidence of LA/EDM research. We searched the literature and gathered representative, mature and highly-cited articles of real case studies with actual data, both from LA and EDM domains. The analysis of selected case studies and their results shed light on the approaches followed by the respective research communities and revealed the potential of this emerging field of educational research. Along with the arising opportunities, we discovered a number of gaps that require the researchers' attention. Table 6 illustrates our findings regarding the strengths, weaknesses, opportunities and threats (SWOT) of LA/EDM research.

Table 7. SWOT of LA/EDM research

Strengths	*Weaknesses*
• Large volumes of available educational data →increased accuracy of experimental results. • Use of pre-existing powerful and valid algorithmic methods. • Interpretable multiple visualizations to support learners/teachers. • More precise user models for guiding adaptation and personalization of systems. • Reveal critical moments and patterns of learning. • Gain insight to learning strategies and behaviors.	• Misinterpretation of results due to human judgment factors - focus on reporting, not decision. • Heterogeneous data sources: not yet a unified data descriptive vocabulary – data representation issues. • Mostly quantitative research results. Qualitative methods have not yet provided significant results. • Information overload – complex systems. • Uncertainty: "are we ready yet?" So far, only skilled teachers/instructors could interpret the results correctly.
Opportunities	*Threats*
• Use of Open Linked Data for data standardization and compatibility among different tools and applications → generalized platform development. • Multimodal and affective learning opportunities	• Ethical issues – data privacy. • Over-analysis: the depth of analysis becomes profound and the results lack generality. The "over-granularity" approaches so far might threaten the holistic picture

based on sophisticated metrics. • Self-reflection/ self-awareness/ self-learning in intelligent, autonomous and massive systems. • Feed machine readable results from the LA/EDM procedures to other data-driven systems for diving decision making. • Acceptance Model: e.g., perceived usefulness, goal expectancy, perceived playfulness, trust, etc.	being explored; look at the tree and miss the forest. • Possibility of pattern misclassification. • Trust: contradictory findings during implementations.

Beyond learning perceptions and attitudes collected through questionnaires, every “click” within an electronic learning environment may be valuable actual information that can be tracked and analyzed. Every simple or more complex action within such environments can be isolated, identified and classified through computational methods into meaningful patterns. Every type of interaction can be coded into behavioral schemes and decoded into interpretable guidance for decision making. This is the point where learning science, psychology, pedagogy and computer science intersect. The issue of understanding the deeper learning processes by deconstructing them into more simple, distinct mechanisms remains in the middle of this cross-path.

We believe that this active research area will continue contributing with valuable pieces of work towards the development of powerful and mostly accurate learning services both to learners and teachers.

References

Abdous, M., He, W., & Yen, C.-J. (2012). Using data mining for predicting relationships between online question theme and final grade. *Educational Technology & Society, 15*(3), 77–88.

Ali, L., Hatala, M., Gašević, D., & Jovanović, J. (2012). A qualitative evaluation of evolution of a learning analytics tool. *Computers & Education, 58*(1), 470–489.

Baker, B. (2007). *A conceptual framework for making knowledge actionable through capital formation.* (Doctoral dissertation). University of Maryland University College, Maryland. Retrieved from ABI/INFORM Global. (Publication No. AAT 3254328).

Baker, R.S.J.d., Corbett, A.T., & Aleven, V. (2008). Improving contextual models of guessing and slipping with a truncated training set. In R. de Baker, T. Barnes, & J. Beck (Eds.), *Proceedings of the 1st International Conference on Educational Data Mining* (pp. 67–76). Retrieved from http://repository.cmu.edu/cgi/viewcontent.cgi?article=1008&context=hcii

Barla, M., Bieliková, M., Ezzeddinne, A. B., Kramár, T., Šimko, M., & Vozár, O. (2010). On the impact of adaptive test question selection for learning efficiency. *Computers & Education*, 55(2), 846–857.

Bienkowski, M., Feng, M., & Means, B. (2012). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. *US Department of Education, Office of Educational Technology, 1-57.*

Blikstein, P. (2011). Using learning analytics to assess students' behavior in open-ended programming tasks. In P. Long, G. Siemens, G. Conole, & D. Gasevic (Eds.), *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (pp. 110-116). New York, NY: ACM.

Catmur, C. (2013). Sensorimotor learning and the ontogeny of the mirror neuron system. *Neuroscience Letters, 540*, 21–27.

Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning, 4*(5), 318–331.

Chaudhuri, S., Dayal, U., & Narasayya, V. (2011). An overview of business intelligence technology. *Communications of the ACM*, 54(8), 88–98.

Chen, C.-M., & Chen, M.-C. (2009). Mobile formative assessment tool based on data mining techniques for supporting web-based learning. *Computers & Education, 52*(1), 256–273.

Chrysafiadi, K., & Virvou, M. (2013). Student modeling approaches: A literature review for the last decade. *Expert Systems with Applications*, 40(11), 4715–4729.

Clow, D., & Makriyannis, E. (2011). iSpot analysed: participatory learning and reputation. In P. Long, G. Siemens, G. Conole, & D. Gasevic (Eds.), *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (pp. 34–43). New York, NY: ACM.

Connolly, T .M., Boyle, E. A, MacArthur, E., & Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education, 59*(2), 661–686.

Cook, D. J., & Das, S. K. (2012). Pervasive computing at scale: Transforming the state of the art. *Pervasive & Mobile Computing,* 8(1), 22–35.

Dejaeger, K., Goethals, F., Giangreco, A., Mola, L., & Baesens, B. (2012). Gaining insight into student satisfaction using comprehensible data mining techniques. *European Journal of Operational Research, 218*(2), 548–562.

Dekker, G. W., Pechenizkiy, M., & Vleeshouwers, J. M. (2009). Predicting students drop out: A case study. In T. Barnes, M. Desmarais, C. Romero, & S. Ventura. (Eds), *Proceedings of the 2nd International Conference on Educational Data Mining* (pp. 41–50). Retrieved from http://www.educationaldatamining.org/EDM2009/uploads/proceedings/dekker.pdf

Eckerson, W. W. (2006). *Performance dashboards: Measuring, monitoring, and managing your business*. Hoboken, NJ: John Wiley & Sons.

Ferguson, R. (2012). *The state of learning analytics in 2012: A review and future challenges*. (Technical Report KMI-2012). Retrieved from http://kmi.open.ac.uk/publications/techreport/kmi-12-01

Fournier, H., Kop, R., & Sitlia, H. (2011). The value of learning analytics to networked learning on a personal learning environment. In P. Long, G. Siemens, G. Conole, & D. Gasevic (Eds.), *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (pp. 104–109). New York, NY: ACM.

Giesbers, B., Rienties, B., Tempelaar, D., & Gijselaers, W. (2013). Investigating the relations between motivation, tool use, participation, and performance in an e-learning course using web-videoconferencing. *Computers in Human Behavior,* 29(1), 285–292.

Guruler, H., Istanbullu, A., & Karahasan, M. (2010). A new student performance analysing system using knowledge discovery in higher educational databases. *Computers & Education, 55*(1), 247–254.

Guo, W. W. (2010). Incorporating statistical and neural network approaches for student course satisfaction analysis and prediction. *Expert Systems with Applications, 37*(4), 3358–3365.

He, W. (2013). Examining students' online interaction in a live video streaming environment using data mining and text mining. *Computers in Human Behavior*, *29*(1), 90–102.

Huang, S., & Fang, N. (2013). Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models. *Computers & Education, 61*, 133–145.

Jeong, H., & Biswas, G. (2008). Mining student behavior models in Learning-by-teaching environments. In R. de Baker, T. Barnes, & J. Beck (Eds), *Proceedings of the 1st International Conference on Educational Data Mining* (pp. 127–136). Retrieved from http://www.teachableagents.org/papers/2008/jeongEDM.pdf

Jourdan, Z., Rainer, R. K., & Marshall, T. E. (2008). Business intelligence: An analysis of the literature. *Information Systems Management, 25*(2), 121–131.

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In D. Suthers, K. Verbert, E. Duval, & X. Ochoa (Eds.), *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge* (pp. 170–179), New York, NY: ACM.

Khribi, M. K., Jemni, M., & Nasraoui, O. (2009). Automatic recommendations for e-learning personalization based on web usage mining techniques and information retrieval. *Educational Technology & Society, 12*(4), 30–42.

Klašnja-Milićević, A., Vesin, B., Ivanović, M., & Budimac, Z. (2011). E-learning personalization based on hybrid recommendation strategy and learning style identification. *Computers & Education, 56*(3), 885–899.

Krakauer, J. W., & Mazzoni, P. (2011). Human sensorimotor learning: adaptation, skill, and beyond. *Current Opinion in Neurobiology, 21*(4), 636–644.

Kwon, O., & Sim, J. M. (2013). Effects of data set features on the performances of classification algorithms. *Expert Systems with Applications*, *40*(5), 1847–1857.

Leong, C-K., Lee, Y-H., & Mak, W-K. (2012). Mining sentiments in SMS texts for teaching evaluation. *Expert Systems with Applications*, 39(3), 2584–2589.

Levy, S. T., & Wilensky, U. (2011). Mining students' inquiry actions for understanding of complex systems. *Computers & Education, 56*(3), 556-573

Li, N., Cohen, W. W., Koedinger, K. R., & Matsuda, N. (2011). A machine learning approach for automatic student model discovery. In M. Pechenizkiy, T. Calders, C. Conati, S. Ventura, C. Romero, & J. Stamper (Eds.), *Proceedings of the 4th International Conference on Educational Data Mining* (pp. 31–40). Retrieved from http://pact.cs.cmu.edu/pubs/Li,%20Cohen,%20Koedinger,%20Matsuda-edm-2011.pdf

Lin, F.-R., Hsieh, L.-S., & Chuang, F.-T. (2009). Discovering genres of online discussion threads via text mining. *Computers & Education, 52*(2), 481–495.

Lin, C.-F., Yeh, Y., Hung, Y. H., & Chang, R. (2013). Data mining for providing a personalized learning path in creativity: An application of decision trees. *Computers & Education, 68*, 199–210.

Lykourentzou, I., Giannoukos, I., Nikolopoulos, V., Mpardis, G., & Loumos, V. (2009a). Dropout prediction in e-learning courses through the combination of machine learning techniques. *Computers & Education, 53*(3), 950–965.

Lykourentzou, I., Giannoukos, I., Mpardis, G., Nikolopoulos, V., & Loumos, V. (2009b). Early and dynamic student achievement prediction in e-learning courses using neural networks. *Journal of the American Society for Information Science and Technology,* 60(2), 372–380.

Manouselis, N., Drachsler, H., Verbert, K., & Duval, E. (2013). *Recommender systems for learning.* New York, NY: Springer.

Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & Education, 54*(2), 588–599.

Merceron, A., & Yacef, K. (2008). Interestingness measures for association rules in educational data. In R. de Baker, T. Barnes, & J. Beck (Eds), *Proceedings of the 1st International Conference on Educational Data Mining* (pp. 57–66). Retrieved from http://www.educationaldatamining.org/EDM2008/uploads/proc/6_Yacef_18.pdf

Moridis, C. N., & Economides, A. A. (2009). Prediction of student's mood during an online test using formula-based and neural network-based method. *Computers & Education, 53*(3), 644–652.

Okoli, C., & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research. *Sprouts: Working Papers on Information Systems*, 10(26). Retrieved from http://sprouts.aisnet.org/10-26

Pardos, Z. A., Baker, R. S. J. d., San Pedro, M. O. C. Z., Gowda, S. M., & Gowda, S. M. (2013). Affective states and state tests: investigating how affect throughout the school year predicts end of year learning outcomes. In D. Suthers, K. Verbert, E. Duval, & X. Ochoa (Eds.) *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge* (pp. 117–124). New York, NY: ACM.

Peña-Ayala, A. (2014). Educational data mining: A survey and a data mining-based analysis of recent works. *Expert Systems with Applications, 41*(4), 1432–1462.

Piaget, J. (1952). *The origins of intelligence in children*. New York, NY: International Universities Press.

Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, *1*, 135–146.

Romero, C., & Ventura, S. (2010). 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), 601–618.

Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 3*(1), 12–27.

Romero, C., Ventura, S., Espejo, P. G., & Hervás, C. (2008). Data mining algorithms to classify students. In R. de Baker, T. Barnes, & J. Beck (Eds), *Proceedings of the 1st International Conference on Educational Data Mining* (pp. 8–17). Retrieved from http://sci2s.ugr.es/keel/pdf/keel/congreso/Data%20Mining%20Algorithms%20to%20Classify%20Students.pdf

Romero, C., Ventura, S., Zafra, A., & de Bra, P. (2009). Applying web usage mining for personalizing hyperlinks in web-based adaptive educational systems. *Computers & Education, 53*(3), 828–840.

Romero-Zaldivar, V-A., Pardo, A., Burgos, D., & Kloos, C. D. (2012). Monitoring student progress using virtual appliances: a case study. *Computers & Education, 58*(4), 1058–1067.

Santos, O. C., & Boticario, J. G. (2012, September). Affective issues in semantic educational recommender systems. In N. Manouselis, H. Draschler, K. Verber, & O. C. Santos (Eds.), *Proceedings of the 2nd Workshop on Recommender Systems for Technology Enhanced Learning* (Vol. 896, pp. 71-82). Retrieved from http://ceur-ws.org/Vol-896/paper6.pdf

Santos, J. L., Govaerts, S., Verbert, K., & Duval, E. (2012). Goal-oriented visualizations of activity tracking: A case study with engineering students. In S. Buckingham Shum, D. Gasevic, & R. Ferguson (Eds.), *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 143–152), New York, NY: ACM

Shih, B., Koedinger, K. R., & Scheines, R. (2008). A response time model for bottom-out hints as worked examples. In R. de Baker, T. Barnes, & J. Beck (Eds), *Proceedings of the 1st International Conference on Educational Data Mining* (pp. 117–126). Retrieved from http://www.cmu.edu/pier/pdfs/12_Shih_35%20copy.pdf

Siemens, G., & Baker, R. S. J. D. (2012). Learning analytics and educational data mining: Towards communication and collaboration. In S. Buckingham Shum, D. Gasevic, & R. Ferguson (Eds.), *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 252–254). New York, NY: ACM.

Thai-Nghe, N., Horváth, T., & Schmidt-Thieme, L. (2011). Factorization Models for forecasting student performance. In M. Pechenizkiy, T. Calders, C. Conati, S. Ventura, C. Romero, & J. Stamper (Eds.), *Proceedings of the 4th International Conference on Educational Data Mining, EDM 2011, July 6-8,* Eindhoven, the Netherlands (pp. 11–20). Retrieved from http://www.ismll.uni-hildesheim.de/pub/pdfs/Nguyen_et_al_EDM_2011.pdf

Tanes, Z., Arnold, K. E, King, A. S., & Remnet, M. A. (2011). Using Signals for appropriate feedback: Perceptions and practices. *Computers & Education, 57*(4), 2414–2422.

Tatar, D., Roschelle, J., Vahey, P., & Penuel, W. R. (2003). Handhelds go to school: Lessons Learned. *IEEE Computer, 36*(9), 30–37.

Terzis, V., & Economides, A. A. (2011). The acceptance and use of computer based assessment. *Computers & Education, 56*(4), 1032–1044.

Verbert, K., Drachsler, H., Manouselis, N., Wolpers, M., Vuorikari, R., & Duval, E. (2011). Dataset-driven research for improving recommender systems for learning. In P. Long, G. Siemens, G. Conole, & D. Gasevic (Eds.), *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (pp. 44–53). New York, NY: ACM.

Wilson, K., Boyd, C., Chen, L., & Jamal, S. (2011). Improving student performance in a first-year geography course: Examining the importance of computer-assisted formative assessment. *Computers & Education, 57*(2), 1493–1500.

Worsley, M., & Blikstein, P. (2013). Towards the development of multimodal action based assessment. In D. Suthers, K. Verbert, E. Duval, & X. Ochoa (Eds.), *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge* (pp. 94–101), New York, NY: ACM.