Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 159 (2019) 1957–1965

Procedia Computer Science

www.elsevier.com/locate/procedia

23rd International Conference on Knowledge-Based and Intelligent Information & Engineering Systems

Knowledge-Aware Learning Analytics for Smart Learning

Weiqin Chen[a,b]*

[a]SLATE, Universtiy of Bergen, POB. 7807, N-5020 Bergen, Norway
[b]Oslo Metropolitan University, POB. 4, St. Olavs plass, 0130 Oslo, Norway

Abstract

With the increasing development and adoption of digital technologies for education, more data gathered from educational contexts are being analyzed to give actionable insights to stakeholders. As a data-driven approach for better understanding and optimizing learning and the learning environment, learning analytics has the potential to contribute to smart learning. However, current learning analytics lacks knowledge awareness, an important component in smart learning. This paper draws upon research in the domain of smart learning, reflects on current research on methods and processes in learning analytics, and proposes a framework for knowledge-aware learning analytics for smart learning.

Peer-review under responsibility of KES International.

Keywords: Knowledge; smart learning; learning analytics

1. Introduction

The last decade has seen an increasing number of publications analyzing educational data and providing actionable insights into learning and the learning environment. Learning analytics (LA) is gaining attention as an emerging research field. LA is “the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs” [1].

LA analyzes a large amount of data in education and presents the results to stakeholders (learners, teachers, faculty, etc.) for evidence-based decision-making. For example, LA provides an overview of learning data in a

* Corresponding author. Tel.: +47-67238671; fax: +47 67235000.
E-mail address: weiqin.chen@oslomet.no

Peer-review under responsibility of KES International.
10.1016/j.procs.2019.09.368

dashboard for learners to reflect on their own activities, engagement, and progress [2] and for teachers to reflect on their teaching practice and make decisions about necessary interventions [3]. LA can also perform predicative analysis, and the results can be used to encourage success and prevent failure or drop-out [4]. Some LA applications also recommend learning resources based on the analysis of learner activities [5]. In contrast to educational data mining (EDM), which is focused on developing methods to explore the unique types of data generated in educational settings, LA is more concerned with sense-making and action [6].

Smart learning is another emerging area that is gaining momentum. It is defined as learning in interactive, intelligent, and tailored environments, supported by advanced digital technologies and services (e.g., context-awareness, augmented reality, cloud computing, social networking services) [7]. Hwang [8] identified three criteria for a smart learning environment. Besides context-awareness, a smart learning environment should be able to offer instant and adaptive support to learners and should be able to adapt the user interface and the subject contents to meet the personal factors and learning status of individual learners. Hwang [8] further stated that adaptive support should include learning guidance, feedback, hints, and learning tools based on learners' needs. To provide adaptive support and an adaptive user interface, understanding learners, their preferences, and performance, as well as online and real-world status and context, is necessary.

LA can contribute to this understanding by providing insights into the above-mentioned features based on learners' behavior data in online and real-world contexts. Such insights can be considered a model of learners' behavior. However, profiling learners without considering the knowledge aspect can only offer an incomplete view of the learning experience. To achieve adaptation, learners' knowledge must also be understood, and a mechanism to reason and provide adaptive learning support based on the learner model and pedagogical knowledge is necessary.

This paper aims to draw upon research in the field of smart learning, reflect on current research on methods and processes in LA, and propose a framework for knowledge-aware LA for smart learning.

2. Learning Analytics and Its Relevance for Smart Learning

In LA, learners' interaction data are gathered, analyzed, and presented as interpretable guidance for stakeholders. Papamitsiou and Economides [9] identified six research objectives for data-driven analytics in education in their systematic literature review. These objectives are as follows: student/student behavior modeling, prediction of performance, increase of students' and teachers' (self-) reflection and (self-) awareness, prediction of dropout and retention, improvement of provided feedback and assessment services, and recommendation of resources. These objectives can be directly mapped to the design of a smart learning environment. According to Spector [10], a learning environment may be considered smart when it "makes use of adaptive technologies or when it is designed to include innovative features and capabilities that improve understanding and performance." Smart learning not only refers to the idea of improving learning but also emphasizes the need for adaptation and personalization, accounting for the places where learning occurs [11].

An LA dashboard is widely used for visualizing learning traces for learners and teachers. It provides graphical representation of the current and historical state of the learner and learning process to enable flexible decision-making. In a previous study, the authors compared 15 dashboard applications and summarized the learning dashboard characteristics [12]. The comparison showed that outputs from LA visualized in a learning dashboard included activity data of the learners such as time spent, social interaction, document and tool use, artifacts produced, and exercise results/quizzes. These visualizations were presented either to teachers or students or both to support awareness, reflection, and sense-making.

Besides showing matrices and indicators to teachers and learners, some LA applications present results from content analysis such as essay content summaries [13] and provide contextually relevant feedback and follow-up instructions based on the content of students' short answers [14]. For knowledge-building communities, LA applications analyze and visualize the relationships among topics or groups of topics. This can provide teachers with insights into students' mental models and misconceptions [15]. Such outcomes from content or discourse analytics represent an extended learner model with a knowledge status that goes one step further in the direction of providing content-based adaptation and support.

The concept of smart LA was introduced in 2016 and is considered a subset of LA that focuses on supporting the features and processes of smart learning [16]. Knowledge structure is one of the main elements of smart LA. For example, the PeT analytics system makes use of a task-related knowledge structure and hierarchical task model with an unlimited number of layers [17]. Another smart LA system takes a pre-defined curriculum as a starting point. Based on the output of the LA, the system provides a personalized learning path and a degree of guidance [18]. In this system, the pedagogical model is expected to be reusable for other domains, which implies that the pedagogical knowledge is somewhat explicitly represented. LA has been found to help acquire new knowledge structures from the analyzed data and derive better knowledge structures that benefit the dynamic adoption of learners' needs [19].

3. Knowledge in Smart Learning Environments

Lister [20] illustrated how digital knowledge content plays a pivotal role in learning design and learner interactions occurring in smart learning, both for the content of learning and as part of the learning process. In a smart learning environment, three types of knowledge are used to provide adaptive learning support: domain knowledge, pedagogical knowledge, and learner knowledge. These types of knowledge are also represented in intelligent tutoring systems (ITSs) [21].

3.1. Domain knowledge

Domain knowledge, also called expert knowledge, represents the facts and rules of a particular domain in which learners learn. In ITSs, great effort has been made to discover and codify the domain knowledge and make it more explicit in the form of ontology or a concept map. The explicitly represented domain knowledge serves two purposes: as the source of knowledge to be presented to the learners and as standard knowledge to assess the learners' overall progress. In most current digital learning systems, including e-learning, m-learning, and massive open online courses (MOOCs), the domain knowledge is more implicit. It is hidden in the modules and steps and is often called knowledge content. For example, in a typical MOOC, the domain knowledge—what the participants are supposed to learn—is organized into modules and steps. Each module can have several steps, and each step can present a topic or a unit of knowledge, described by videos or texts. Quizzes and discussions are often added at the end of each step or each module. The learner's overall progress is often assessed by the results of quizzes and what the learner has clicked on through the steps and modules.

3.2. Pedagogical knowledge

Pedagogical knowledge, also called teaching strategy or teaching knowledge, uses knowledge about learners and domain knowledge to decide which pedagogical activities will be performed, such as giving hints to overcome impasses in performance, advice, support, explanations, different practice tasks, etc. [22]. ITSs also aim to represent pedagogical knowledge explicitly, providing the possibility to adapt and improve teaching strategies over time and to reuse the strategies for other domains. In most current digital learning systems, including e-learning, m-learning, and MOOCs, the pedagogical knowledge lies in the minds of the course design and teaching team. Based on the learner's progress, as visualized in dashboards in the system, the course team makes decisions on pedagogical interventions.

3.3. Learner knowledge

Learner knowledge is related to the domain knowledge the learner is supposed to learn. It represents the learners' emerging knowledge from their behaviors and learning. In ITSs, learner knowledge is part of the learner model, which represents an understanding of the learner and is, in most cases, explicit. With this model, the ITSs can provide adaptive interventions to learners. In most current digital learning systems, including e-learning, m-learning, and MOOCs, the learner model is represented in the dashboard with behavior information based on clickstreams, such as how many steps learners have gone through, how many quizzes they have tried, and how much time they have spent.

Knowledge can also be generated during the learning process. For example, in a knowledge-building process, groups generate new knowledge. This learner-generated knowledge can be part of the learner model used for adaptive support. It can also become part of the domain knowledge to be learned by other learners. Cognitive modeling approaches have been used to trace how learners develop knowledge in ITSs [23], and these approaches remain relevant for LA research [6, 24].

Hwang [8] suggested that a smart learning environment should "not only enable learners to access digital resources and interact with learning systems in any place and at any time, but also should actively provide the necessary learning guidance, hints, supportive tools, or learning suggestions to them in the right place, at the right time, and in the right form". To achieve such adaptation, LA must consider the different types of knowledge.

4. Related Research and Historical Roots

Bienkowski et al. [26] identified the following application areas of LA and EDM: modeling user knowledge, behavior, and experience; creating profiles of users; modeling knowledge domains; trend analysis; and personalization and adaptation. These show the importance of learner knowledge, domain knowledge and pedagogical knowledge in LA.

According to an early definition by George Siemens [27], LA is the use of intelligent data, learner-produced data, and analysis models to discover information and social connections for predicting and advising people's learning. Based on this definition, Siemens presented an integrated knowledge and LA model (Fig. 1) in which intelligent curriculum covers connected knowledge, semantic data, and linked data. He considered that the knowledge, attitudes, and skills required in any domain can be rendered as a network of relations, which can be presented by semantic web and linked data. He further stated that knowledge domains can be mapped and that learner activity can be evaluated in relation to these maps [28]. Although Siemens later criticized this definition and claimed that "learning analytics—at an advanced and integrated implementation—can do away with pre-fab curriculum models," research in learner modeling and the three levels of knowledge (domain knowledge, learner knowledge, and pedagogical knowledge) in ITSs are still relevant to LA research [6]. For example, concept networks derived from

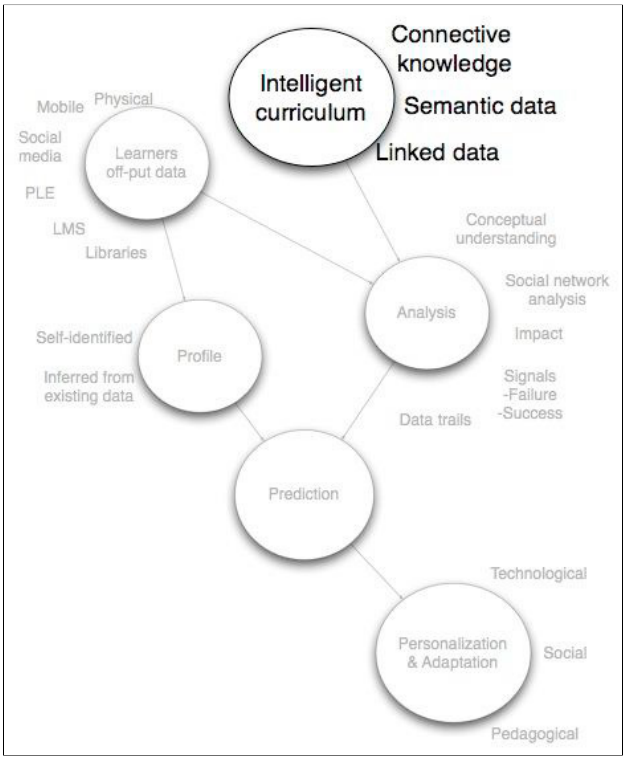

Fig. 1. Integrated knowledge and learning analytics model with intelligent curriculum highlighted [25].

learner-generated texts analysis can reveal learners' mental models and misconceptions. This can be a basis for enriching domain taxonomies and for curriculum revision [29] as well as for generating pedagogical interventions to address misconceptions.

The integrated learning analytics platform proposed by Siemens, et al. [30] includes the development of four specific tools and resources (Fig. 2):

1. Learning analytics engine
2. Adaptive content engine
3. Intervention engine: recommendations, automated support
4. Dashboard, reporting, and visualization tools

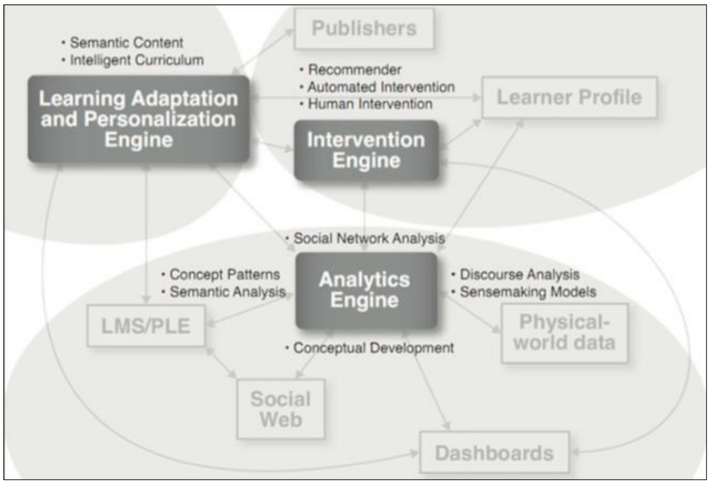

Fig. 2. An integrated learning analytics platform [30].

The learning analytics engine gathers and processes data using different analytical methods. The adaptive content engine provides personalized learning pathway and learning content to learners based on their profiles. In order to achieve personalization, learning materials should be designed to reflect the knowledge architecture of a domain. Once knowledge domains have been articulated or mapped, learner data, profile information, and curricular data can be brought together and analyzed to determine learner knowledge in relation to the knowledge structure of a domain [6]. The intervention engine tracks learner progress and provides automated interventions or enables human intervention by providing recommendations to teachers and using prediction models developed in the learning analytics engine. System-generated interventions can for example range from a simple alert about a leaner's likeliness to succeed to requiring an at-risk learner to take a specific action tailored to this particular learner in order to address the concerns. The dashboard presents visualized data to create self-awareness and sense-making and assist individuals in making decisions about teaching and learning.

Hwang [8] proposed a framework for smart learning environment. This framework includes six modules and six databases (Fig. 3).

1. The learning status detecting module detects learner's real-world status and context. This information is stored in the learning portfolio database.
2. The learning performance evaluation module assesses learners' performance by conducting tests online or in the real world. The tests given to the learner are related to the learning goals and selected from the test bank database.
3. The adaptive learning task module provides learning tasks to learners based on their progress, performance, personal factors and their learning objectives, which are stored in the portfolio database. The learning tasks are selected from learning sheets and materials.

4. The adaptive learning content module recommends and organizes learning materials, and adapts the user interface based on the learner information in the portfolio database. The tailored learning materials are presented in different learning tools.
5. The personal learning support module takes into consideration of the current task and learning materials and learner profile, and provides just-in-time guidance, hints and recommendations.
6. The inference engine makes decisions by reasoning with the learner status and context based on the pedagogical knowledge stored in the knowledge base. The reasoning results in dynamic values of learning tasks, materials, strategies and tools for each learner.

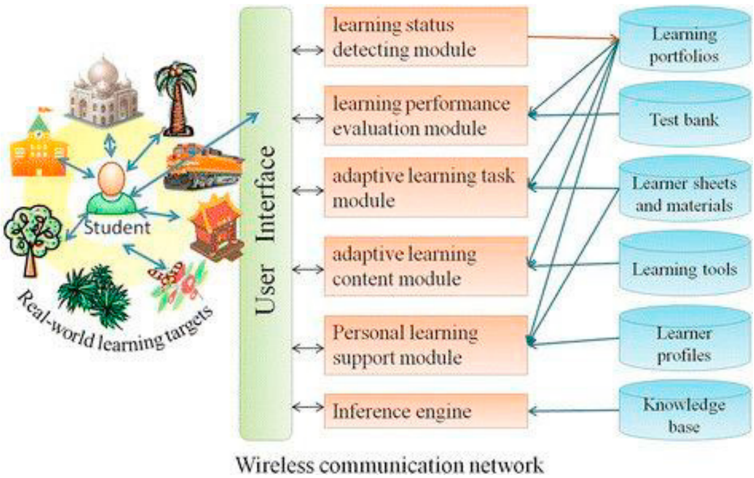

Fig. 3. Framework of a smart learning environment [8].

In this framework, LA covers the learner data collection and analysis and contributes to the learning portfolio and learner profiles. The outputs of learning analytics can then be used by the inference engine to provide adaptive learning experience. In this framework, the pedagogical knowledge seems to be explicitly represented, however, the domain knowledge is represented in the form of learning sheets, materials and the test bank. The learner knowledge is included in the learning portfolios.

5. Framework of a Knowledge-Aware LA for Smart Learning

In this section, a scenario is used to illustrate the proposed framework of a knowledge-aware LA for smart learning (Fig. 4). In the scenario, Anna is a visiting student at the University of Kyoto who uses a location-based mobile app to learn about the history of Kyoto. She has registered as a user in this app and has created a profile with her personal learning preferences about locations, eras in history, and subjects (such as architecture and culture). She has also determined how and when she prefers to receive notifications and guidance.

The app has a pre-defined domain knowledge base, including a concept map for each of the subjects and a list of tasks, learning resources, and quizzes associated with each concept. The knowledge base is used by the inference engine, together with the pedagogical knowledge and learner model, to create adaptive tasks, content, and guidance. The learner model is generated by the analysis module based on the data gathered though the app. The learner model includes learner preferences, activities, context, performance, and a concept map of the learner’s knowledge.

Every day Anna chooses one task from the list of tasks in the app and follows the required route and answers quizzes along the way. The recommended tasks are initially generated based on her profile. Once she starts carrying out the tasks, the app gathers more data about her activities, context and locations (1). These data are transferred in real time through the mobile data network to a cloud service (2). The analysis module processes the data about Anna in the cloud and updates her learner model (3). The analysis processes not only interaction data, but also data bout the context, the history domain, as well as other data about Anna that are already in the cloud.

The visualization in the app shows her statistics on her completed tasks and quizzes, correct and incorrect answers, hint used, learning resource access, time spent on different activities, and her current knowledge status mapped to the domain concept map (4). This can help her with reflection on her activities and performances. What she finds most helpful is that the system provides her with tailored tasks, quizzes, learning resources and support based on her current task or quiz in the app, her physical location in the city. Such adaptation is generated by the interface engine with input from pedagogical and domain knowledge base as well as her learner model (5).

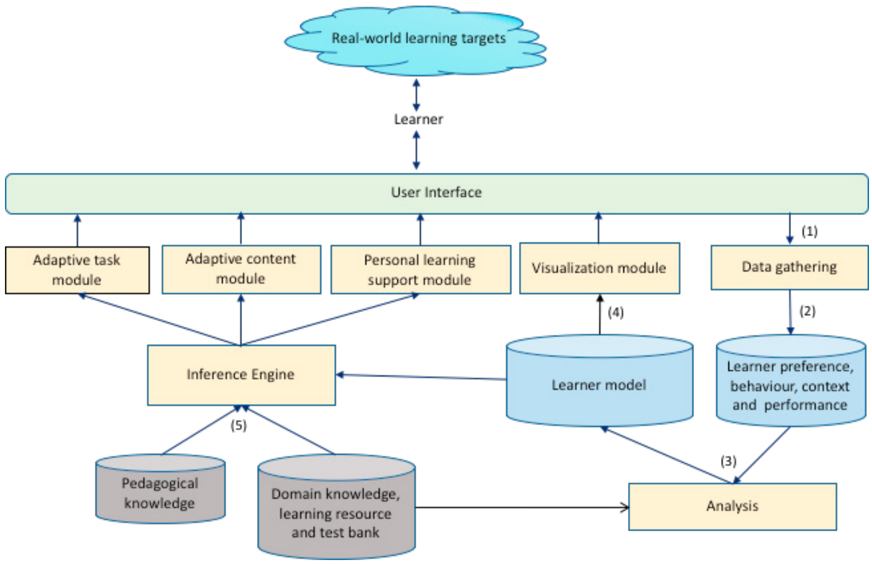

Fig. 4. A framework of a knowledge-aware LA for smart learning.

The LA in this framework contributes to building and updating the learner model. A learner model based on LA without domain knowledge does not represent a complete picture of the learner and limits the ability of the inference engine to provide adaptive content and support. In a smart learning environment where context awareness and adaptive support are two of the most important elements, knowledge-aware LA can provide a richer learner model, which can support better adaptation.

6. Conclusion and Future Work

The last few years have seen an increasing number of publications related to knowledge gaps of learners and visualization of such gaps. Some have used domain ontology and knowledge maps [31]. Kinshuk and Kumar [32] reviewed case studies in smart learning analytics and stated that smart learning analytics and smart learning environments are indispensable and the smartness depends on how adaptive the system is. We argue that in order to achieve true adaptivity, it is necessary to consider domain knowledge and pedagogy knowledge, and more research is necessary in terms of using these knowledges for learning analytics in smart learning environments. We hope that the framework proposed in this paper could contribute to knowledge-aware learning analytics. By integrating knowledge into today's data-driven approaches, we may provide stakeholders a deeper understanding of learning, opportunities for better decision making and a better understanding of the impacts of the decisions.

Future work should focus on the validation of this framework, identify the drawbacks and challenges in adopting it for integrating knowledge in learning analytics, and demonstrate the power and potential of blended knowledge and data analytics approach in education.

References

[1] Siemens, George. (2011a) *Learning and Academic Analytics*. Available at: https://www.learninganalytics.net/?p=131.
[2] Clow, Doug and Elpida Makriyannis. (2011) iSpot analysed: participatory learning and reputation. In: Phillip Long, George Siemens, Gráinne Conole and Dragan Gašević (eds) *Proceedings of the 1st International Conference on Learning Analytics and Knowledge.* New York, NY: ACM, 34-43.
[3] Lin, Chun Fu, Yu-Chu Yeh, Yu Hsin Hung and Ray I Chang. (2013) Data mining for providing a personalized learning path in creativity: An application of decision trees. *Computers & Education* 68: 199–210.
[4] Dekker, Gerben W. , Mykola Pechenizkiy and Jan M. Vleeshouwers. (2009) Predicting students drop out: A case study. In: Tiffany Barnes, Michel Desmarais, Cristóbal Romero and Sebastian Ventura (eds) *Proceedings of the 2nd International Conference on Educational Data Mining*. International Educational Data Mining Society, 41–50.
[5] Verbert, Katrien, Hendrik Drachsler, Nikos Manouselis, Martin Wolpers, Riina Vuorikari and Erik Duval. (2011) Dataset-driven research for improving recommender systems for learning. In: Phillip Long, George Siemens, Gráinne Conole and Dragan Gašević (eds) *Proceedings of the 1st International Conference on Learning Analytics and Knowledge.* New York, NY: ACM, 44–53.
[6] Siemens, George. (2013) Learning Analytics: The Emergence of a Discipline. *American Behavioral Scientist* 57: 1380–1400.
[7] Lee, Junghwan, Hangjung Zo and Hwansoo Lee. (2014) Smart learning adoption in employees and HRD managers. *British Journal of Educational Technology* 45: 1082–1096.
[8] Hwang, Gwo-Jen. (2014) Definition, framework and research issues of smart learning environments – a context-aware ubiquitous learning perspective. *Smart Learning Environments Open Journal* 1: 1-14.
[9] Papamitsiou, Zacharoula and Anastasios A. Economides. (2014) Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence *Journal of Educational Technology & Society* 17: 49-64.
[10] Spector, Jonathan Michael. (2014) Conceptualizing the emerging field of smart learning environments. *Smart Learning Environments* 1: 1-10.
[11] Gros, Begoña. (2016) The design of smart educational environments. *Smart Learning Environments* 3.
[12] Verbert, Katrien, Erik Duval, Joris Klerkx, Sten Govaerts and José Luis Santos. (2013) Learning Analytics Dashboard Applications. *American Behavioral Scientist* 57: 1500–1509.
[13] Whitelock, Denise, Alison Twiner, John T. E. Richardson, Debora Field and Stephen Pulman. (2015) OpenEssayist: A Supply and demand learning analytics tool for drafting academic essays. In: Josh Baron, Grace Lynch, Nicole Maziarz, Paulo Blikstein, Agathe Merceron and George Siemens (eds) *Proceedings of the 5th International Conference on Learning Analytics and Knowledge (LAK '15).* New York: NY: ACM, 208–212.
[14] Dzikovska, Myroslava, Natalie Steinhauser, Elaine Farrow, Johanna Moore and Gwendolyn Campbell. (2014) BEETLE II: Deep natural language understanding and automatic feedback generation for intelligent tutoring in basic electricity and electronics. *International Journal of Artificial Intelligence in Education* 24: 284–332.
[15] Hoppe, Heinz Ulrich, Michael Erkens, Gill Clough, Oliver Daems and Anne Adams. (2013) Using Network Text Analysis to characterise teachers' and students' conceptualisations in science domains. In: Ravi Vatrapu, Peter Reimann, Wolfgang Halb and Susan Bull (eds) *Proceedings of the 2nd International Workshop on Teaching Analytics (IWTA 2013).*
[16] Giannakos, Michail N., Demetrios G. Sampson and Łukasz Kidziński. (2016) Introduction to smart learning analytics: foundations and developments in video-based learning. *Smart Learning Environments* 3(12).
[17] Boulanger, David, Jérémie Seanosky, Michael Baddeley, Vivekanandan Kumar and Kinshuk. (2014) Learning analytics in the energy industry: measuring competences in emergency procedures. *IEEE 6th International Conference on Technology for Education (T4E).* Washington, DC: IEEE Computer Society.
[18] Bacca, Jorge, Silvia Baldiris, Ramon Fabregat, Kinshuk and Sabine Graf. (2015) Mobile augmented reality in vocational education and training. *Procedia Computer Science* 75: 49-58.
[19] Clemens, Clayton, Vivekanandan Kumar, David Boulanger, Jérémie Seanosky and Kinshuk. (2017) Learning traces, competence, and causal inference for English composition. In: Jonathan Michael Spector, Vivekanandan Kumar, Alfred Essa, Yueh-Min Huang, Rob Koper, Richard A. W. Tortorella, Ting-Wen Chang, Yanyan Li and Zhizhen Zhang (eds) *Frontiers of Cyberlearning – Emerging Technologies for Teaching and Learning.* Singapore: Springer, 49-67.
[20] Lister, Penelope J. (2018) A smarter knowledge commons for smart learning. *Smart Learning Environments* 5(8).
[21] Nwana, Hyacinth S. (1990) Intelligent tutoring systems: an overview. *Artificial Intelligence Review* 4: 252–277.
[22] Self, John A. (1988) Student models: what use are they? In: Paolo Ercoli and R. Lewis (eds) *Artificial Intelligence Tools in Education.* New York, NY: Elsevier Science Inc., 73–86.
[23] Burns, Hugh L. and Charles G. Capps. (1988) Foundations of intelligent tutoring systems: An introduction. In: Martha C. Polson and J. Jeffrey Richardson (eds) *Foundations of Intelligent Tutoring Systems.* Hillsdale, N.J.: Lawrence Erlbaum Associates, 1-19.
[24] Koedinger, Kenneth. (2001) Cognitive Tutors as Modeling Tool and Instructional Model. In: Kenneth D. Forbus and Paul J. Feltovich (eds) *Smart Machines in Education: The Coming Revolution in Educational Technology.* Menlo Park, CA: AAAI/MIT Press, 145–168.
[25] Siemens, George. (2011b) *Learning Analytics: a foundation for informed change in Higher education*. Presentation to EDUCAUSE ELI. Available at: https://events.educause.edu/eli/webinars/2011/eli-web-seminar-january-10.
[26] Bienkowski, Marie, Mingyu Feng and Barbara Means. (2012) Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. Washington, D.C.: U.S. Department of Education, Office of Educational Technology.
[27] Siemens, George. (2010) *What are Learning Analytics?* Available at: http://www.elearnspace.org/blog/2010/08/25/what-are-learning-

analytics/.
[28] Long, Phillip and George Siemens. (2011) Penetrating the fog: analytics in learning and education. *EDUCAUSE Review* 48: 31–40.
[29] Hoppe, Heinz Ulrich. (2017) Computational Methods for the Analysis of Learning and Knowledge Building Communities. In: George Siemens Charles Lang, Alyssa Wise, Dragan Gašević (ed) *Handbook of Learning Analytics*. Society for Learning Analytics Research, 23-33
[30] Siemens, George, Dragan Gasevic, Caroline Haythornthwaite, Shane Dawson, Simon Buckingham Shum, Rebecca Ferguson, Erik Duval, Katrien Verbert and Ryan S. J. D. Baker. (2011) *Open Learning Analytics: an integrated & modularized platform*. Available at: http://www.solaresearch.org/OpenLearningAnalytics.pdf.
[31] Flanagan, Brendan, Rwitajit Majumdar, Gökhan Akçapınar, Jingyun Wang and Hiroaki Ogata. (2019) Knowledge Map Creation for Modeling Learning Behaviors in Digital Learning Environments. In: Jim Cunningham, Nicole Hoover, Sharon Hsiao, Grace Lynch, Katie Mccarthy, Christopher Brooks, Rebecca Ferguson and Heinz Ulrich Hoppe (eds) *Companion Proceedings 9th International Conference on Learning Analytics & Knowledge (LAK19)*. Society for Learning Analytics Research, 428-436.
[32] Kinshuk and Vivekanandan Kumar. (2018) Advancing learning through smart learning analytics: a review of case studies. *Asian Association of Open Universities Journal* 13: 1-12.