# Introduction into Sparks of the Learning Analytics Future

**Mykola Pechenizkiy**

Department of Computer Science

Eindhoven University of Technology (TU/e)

The Netherlands

m.pechenizkiy@tue.nl


**Dragan Gašević**

Schools of Education and Informatics

University of Edinburgh

United Kingdom

dgasevic@acm.org

**ABSTRACT:** This section offers a compilation of 16 extended abstracts summarizing research of the doctoral students who participated in the Second Learning Analytics Summer Institute (LASI 2014) held at Harvard University in July 2014. The abstracts highlight the motivation, main goals and expected contributions to the field from the ongoing learning analytics doctoral research around the globe. These works cover several major topics in learning analytics including novel methods for automated annotations, longitudinal analytic studies, networking analytics, multi-modal analytics, dashboards, and data-driven feedback and personalization. The assumed settings include the traditional classroom, online and mobile learning, blended learning, and massive open online course education models.

**KEYWORDS:** learning analytics, doctoral students, future of learning analytics, Learning Analytics Summer Institute, LASI 2014, SoLAR

## 1  SECOND LEARNING ANALYTICS SUMMER INSTITUTE (LASI 2014)

Learning analytics is an emerging multidisciplinary field with connections to educational data mining and other neighboring fields such as the learning sciences and learning at scale. The goal of learning analytics is to support learning assessment to improve understanding of teaching and learning through the collection and analysis of data about learners – their behavior and performance. The field is rapidly growing in terms of the development of innovative methodologies and applications. As such, there is parallel interest developing amongst students pursuing a PhD in learning analytics and closely related areas that will bring further advancements to the field.

The Society for Learning Analytics Research (SoLAR) in collaboration with Harvard Graduate School of Education and with assistance from the International Educational Data Mining Society (IEDMS) organized the 2nd Learning Analytics Summer Institute (LASI 2014) held at Harvard University,

Cambridge, MA, June 30 – July 2, 2014. The institute attracted over 130 participants including university leaders, academics, emergent researchers pursuing their PhDs in the area of learning analytics, and representatives of educational technology vendors. The summer institute had two major goals: i) to build the field of learning analytics through cross disciplinary interactions, identification of research and teaching needs, advancement of learning analytics methods, and connecting isolated researchers and research communities; and ii) to develop the skills and knowledge of doctoral students and academics, equipping them to engage actively in learning analytics research and teaching.

The development of early career researchers, in general, and doctoral students, in particular, is seen as critical for the future sustainability of the learning analytics field. Thirty one doctoral students participated in LASI 2014. To increase their exposure during LASI 2014 and for other participants to understand the current and future trends in learning analytics, a special "lightening session" was held on the morning of the first day of LASI 2014. The session was a great success and attracted much attention. During LASI 2014 an idea emerged [George Siemens is held responsible ☺!] that a special section in the Journal of Learning Analytics should be arranged that would feature these sparks of the learning analytics future. And, here we are only a few months later!

## 2  PAPERS IN THE SPECIAL SECTION

Although all 31 doctoral students were invited to submit their extended abstract, 16 submitted their extended abstracts (3-4 pages) for inclusion in this special section on the *sparks of the learning analytics future*. The students were asked to be either sole authors of the abstracts or to co-author them with their doctoral supervisors only. All the extended abstracts were reviewed to meet the quality expectations of the journal. The abstracts highlight the motivation, main goals and expected contributions to the field from the ongoing learning analytics PhD research around the globe. These works cover several major topics in learning analytics including novel methods for automated annotations, longitudinal analytic studies, networking analytics, multi-modal analytics, dashboards, and data-driven feedback and personalization. The assumed settings include the traditional classroom, online and mobile learning, blended learning, and MOOC education models. Despite the vast diversity of topics and settings studied, it is possible to see the many interesting connections that develop across investigations – some of which are highlighted below.

Tailored learning and assessment has been one of the central topics in online education. There are numerous challenging problems that need to be addressed to establish tailored learning and assessment as mainstream practices. Some of these challenges are addressed in the work selected for this special section. For instance, Yuan and Recker (2014) study how to identify and characterize the differences between the behaviours of novices and elders based on the usage logs from web-based tools used in online educational communities. While, Lang (2014) investigates how to advance the state-of-the-art Bayesian Knowledge Tracing techniques with Inverse Bayes for modeling student performance, and in particular predicting the next action by a student in a cognitive tutor. Ali, Hatala, and Gašević (2014) examine how the learning strategies and achievement goal orientations of students relate to their

academic behaviours and performance. In noting that assignments targeting the upper limits of student ability could aid students in their learning, Ivančević (2014) studies how learning analytics can help the construction of such tests in programming. Papamitsiou and Economides (2014) focus on the analysis of temporal features (e.g. time spent on test questions) to facilitate adaptive assessment and feedback personalization. Ye and Biswas (2014) study how a suitable representation for more effective MOOC student dropout and performance prediction can be constructed by means of feature selection and feature extraction. Studying effects of using a technology on learning performance is another related issue. For example, Demmans Epp (2014) explores the use of an adaptive mobile tool for language learning. Learning analytics can help in understanding the relationship between application usage and learning, and in this particular case, between specific mobile learning activities and English vocabulary learning.

In general, interpretation of raw data and extraction of meaningful features for further analysis is a grand challenge by itself, especially when it comes to multi-modal analytics. Raca and Dillenbourg (2014) study how to extract and explore visual cues from video-recordings of an observational system for measuring classroom activity to provide teachers with additional information about students' level of attention as well as about their own energy and attention distribution in the classroom. Kaczorowski and Raimondi (2014) focuses on how to facilitate flexible learning optimized for diverse elementary-aged students learning mathematics, including students with learning disabilities. The authors argue that although state-of-the-art video-coding software is useful for studying engagement, independence, accuracy, and specific affordances and barriers to learning with mobile technology, the many-sided analysis of interviews, field notes, and video codes was necessary to come up with conclusions.

Indeed, learning analytics often call for interdisciplinary approaches for modeling students' interactions and performance. In particular, Dowell and Graesser (2014) advocate the adoption of psychological theories of discourse comprehension for applying computational linguistics methodologies to identify pedagogically valuable discourse features from text to model and monitor students' latent cognitive, affective, and social processes. Thus, facilitating timely and adaptive support for learners. Gilmore (2014) advocates for a combination of learning analytics and conventional research methods in learning sciences. Goffman's (1959) perspective on front stage (e.g. logged online learning) and backstage (e.g. what student does at home, offline) behaviors is adopted to come up with a more holistic modeling of student learning.

Investigating various student–student and student–information interactions occurring within a networked learning environment and their effect on student learning is also becoming possible. Different tools are requested to facilitate many-sides analysis of relevant data reflecting these interaction processes. Hence, Joksimović, Gašević, and Hatala (2014) propose a theoretical model of networked learning that allows for a comprehensive and scalable analysis of how and why students interact. Kovanovic, Gašević, and Hatala (2014) develop a learning analytics framework for inquiry-based digital learning, focusing on i) facilitation of automated coding by means of text classification and ii) investigation of how learning processes unfold over time through student interactions with information and peers.

Wherever possible, learning analytics attempts to steer away from black-box solutions for generating hints, recommendations, or optimizing the presentation of learning materials and their sequencing. Instead, the goal is to make students and teachers aware of the types of analytics and data collection processes undertaken to help teachers obtain insights into the learning process and progress of their students. For this purpose, different explanatory models, visual analytics techniques, and dashboards have been proposed. Thus, Charleer, Klerkx, and Duval (2014) study learning dashboards, i.e. how information visualization techniques can help teachers and students deal with the abundance of educational data in interactive and collaborative ways with the use of tabletops and whiteboards. Another interesting problem is how to help students with their progress and career advancement. Wang, Paquette, and Baker (2014) conduct a longitudinal study of MOOCs learners integrating course-related data with longitudinal data. Kobayashi, Mol, and Kismihók (2014) attempt to help students to transfer from education to employment by exposing them to relevant labour market information.

## 3  CONCLUDING REMARKS

We hope you will find the contributions presented in this section informative for understanding what the future of learning analytics may bring. Although being still in early days, the field is certainly poised for a great future with the talented early career researchers who already demonstrate high creative potential and strong methodological rigor in their research.  We are very much grateful to all the doctoral students who participated in LASI 2014 and especially to those who contributed to this section. We look forward to learning more about their future research and following their career progression. They are definitely worthwhile watching as many among them are the rising stars of the learning analytics field!

## REFERENCES

Ali, L., Hatala, M. Gašević, & Winne, P. (2014). Leveraging MSLQ Data for Predicting Student Achievement Goal Orientations. *Journal of Learning Analytics, 1*(3), 157-160.

Charleer, S., Klerkx, J., & Duval, E. (2014). Learning Dashboards. *Journal of Learning Analytics, 1*(3), 199-202.

Demmans Epp, C. (2014). Mobile Adaptive Communication Support for Vocabulary Acquisition. *Journal of Learning Analytics, 1*(3), 173-175.

Dowell, N. M. M. & Graesser, A. C. (2014). Modelling Learners' Cognitive, Affective, and Social Processes through Language and Discourse. *Journal of Learning Analytics, 1*(3) 183-186.

Gilmore, D. (2014). Goffman's Front Stage and Backstage Behaviours in Online Education. *Journal of Learning Analytics, 1*(3), 187-190.

Goffman, E. (1959). *Presentation of self in everyday life*. Garden City, NY: Double Day Anchor Books.

Ivančević, V. (2014). Constructing Programming Tests from an Item Pool: Pushing the Limits of Student Knowledge using Assessment and Learning Analytics. *Journal of Learning Analytics, 1*(3) 161-164.

Joksimović, S., Gašević, D., Hatala, M. (2014). Learning Analytics for Networked Learning Models. *Journal of Learning Analytics, 1*(3), 191-194.

Kaczorowski, T. & Raimondi, S. (2014). eWorkbooks for Mathematics: Mapping the Independent Learning Experiences of Elementary Students with Learning Disabilities . *Journal of Learning Analytics, 1*(3), 179-182.

Kobayashi, V., Mol, S. T., & Kismihók, G. (2014). Labour Market Driven Learning Analytics. *Journal of Learning Analytics, 1*(3), 207-210.

Kovanovic, V., Gašević, D., & Hatala, M. (2014). Learning Analytics for Communities of Inquiry. *Journal of Learning Analytics, 1*(3), 195-198.

Lang, C.  (2014). An Adaptive Model of Student Performance Using Inverse Bayes. *Journal of Learning Analytics, 1*(3), 154-156.

Papamitsiou, Z. & Economides, A. A. (2014). Temporal Learning Analytics for Adaptive Assessment. *Journal of Learning Analytics, 1*(3), 165-168.

Raca, M. & Dillenbourg, P. (2014). Classroom Social Signal Analysis. *Journal of Learning Analytics, 1*(3), 176-178.

Wang, Y., Paquette, L., & Baker, R. (2014). A Longitudinal Study on Learner Career Advancement in MOOCs. *Journal of Learning Analytics, 1*(3), 203-206.

Ye, B. & Biswas, G. (2014). Early Prediction of Student Dropout and Performance in MOOCs using Higher Granularity Temporal Information. *Journal of Learning Analytics, 1*(3), 169-172.

Yuan, M. & Recker, M. (2014). Characterizing User Behaviours and Products in an Online Educational Community: A Comparison between Novices and Elders. *Journal of Learning Analytics, 1*(3), 150-153.