

Learning Analytics Summer Institute Spain 2019: Learning Analytics in Higher Education

Vigo, Spain, June 27-28, 2019

Editors:

Manuel Caeiro-Rodríguez
Universidad de Vigo, España

Ángel Hernández-García
Universidad Politécnica de Madrid, España

Pedro J. Muñoz-Merino
Universidad Carlos III de Madrid, España

ISBN: 978-84-16829-40-8

Learning Analytics Summer Institute Spain 2019: Learning Analytics in Higher Education

Manuel Caeiro-Rodríguez¹ [0000-0002-2784-6060], Ángel Hernández-García² [0000-0002-6549-9549] and Pedro J. Muñoz-Merino³ [0000-0002-2552-4674]

¹ Department of Telematics Engineering, Universidade de Vigo, Campus Lagoas-Marcosende, 36310 Vigo, Spain
mcaeiro@det.uvigo.es

² Departamento de Ingeniería de Organización, Administración de Empresas y Estadística, ETSI de Telecomunicación, Universidad Politécnica de Madrid, Av. Complutense 30, 28040 Madrid, Spain
angel.hernandezj@upm.es

³ Department of Telematics Engineering, Universidad Carlos III de Madrid, Av. Universidad, 30, 28911 Leganés, Madrid, Spain
pedmume@it.uc3m.es

Preface to the Conference Proceedings

The seventh¹ edition of the Learning Analytics Summer Institute Spain, LASI Spain 19² was held in Vigo on June 27th and 28th, 2019. Under the main theme of “Learning Analytics in Higher Education”, the conference was organized by Universidade de Vigo, in collaboration with the SNOLA (Spanish Network of Learning Analytics) research network and TELGalicia. LASI Spain 19 is conceived as a platform to catalyze educators, technologists, researchers, enterprise and policymakers around shaping the next generation of learning infrastructures to truly serve the needs now facing the education sector. LASI Spain 19 is part of the Learning Analytics Summer Institute locals; LASI worldwide events, sponsored by SoLAR (Society for Learning Analytics Research), are strategic events that bring the right mix of people together for an intensive ‘summer camp’ that serves as an intellectual and social springboard to accelerate the maturation of learning analytics.

This year’s edition of LASI Spain focused on a specific context of application of learning analytics: Higher Education. Learning Analytics has been called to improve

¹ The previous editions of LASI Spain as official LASI-local event include the following:

- LASI Spain 2013 in Madrid: <http://www.emadridnet.org/index.php/es/eventos2/312-seminario-emadrid-learning-analytics-summer-institutue>
- LASI Spain 2014 in Madrid: <https://canal.uned.es/serial/index/id/1303>
- LASI Spain 2015 in Bilbao: <https://blogs.deusto.es/lasi2015Bilbao>
- LASI Spain 2016 in Bilbao: <http://lasi16.snola.es>
- LASI Spain 2017 in Madrid: <http://lasi17.snola.es>
- LASI Spain 2018 in León: <http://lasi18.snola.es>

² <https://lasi19.snola.es>

learning practice by transforming the ways we develop and carry out learning and teaching activities and processes. In the particular context of Higher Education, the pervasive integration of digital technology is influencing both teaching and learning practices, allowing access to new resources, functionality and data. Nowadays, existing online learning environments are used to communicate with students, distribute educational resources and perform learning activities. These technologies have already been adopted and their use is common practice in education. Now it is time to move forward and to put the focus on increasing “quality”. This is a central idea shared by the scientific contributions and keynotes presented in LASI Spain 19: that learning analytics can play a key role in this new landscape and contribute to a significant change.

The programme of LASI Spain 19 comprised a wide range of activities that brought together representatives of academia, practitioners and policymakers, including keynotes by international experts on learning analytics in Europe, presentation sessions of scientific studies on learning analytics in Higher Education, discussion panels and workshops. The different activities gave attendants the opportunity to have a comprehensive view of the state of affairs in learning analytics in Higher Education, share experiences in the design and application of learning analytics techniques and showcase innovative pieces of research.

The keynotes of LASI Spain 19 provided a thorough overview of learning analytics in Higher Education and gave attendees insight about new and promising analysis techniques to improve the design and implementation of learning processes in IT-mediated education. More precisely, the keynotes of LASI Spain 19 were as follows:

- In “*Can Learning Analytics Transform Higher Education?*”, Mar Pérez-Sanagustín (Université Paul Sabatier Toulouse III) set out to review the past, present and future of learning analytics in Higher Education. From the perspective of the era of ‘big data’, Dr. Pérez-Sanagustín explained how ‘big data’ were introduced in Higher Education through the emerging discipline of learning analytics. Owing to the introduction of information technologies in Higher Education systems and learning processes, we are living a transition from a time of data scarcity, with students’ grades being the central data, to a time of abundance. Digital environments that collect the students’ digital “fingerprints” in different contexts, generate massive datasets that offer great opportunities for research and to support students’, teachers’ and managers’ tasks, and only by analyzing and understanding the student’s learning process it is possible to provide adequate support and improve it. The presentation emphasized that learning analytics facilitates the provision of aggregated information to every agent involved in the learning process, in order to act at the most appropriate time, define learning strategies and plan and execute interventions. Dr. Pérez-Sanagustín went through the different stages of learning analytics —descriptive, diagnostic, predictive and prescriptive— and highlighted the need of a ‘learning analytics culture’ within Higher Education institutions for an effective change that allows Universities to benefit from the advantages made possible by learning analytics. It was argued that such change is, however, context-dependent, and

requires both bottom-up and top-down approaches to foster diffusion of effective learning analytics practices. The presentation also discussed different challenges faced by the learning analytics community as a whole. Questions such as ‘In what context and with what objective are Higher Education institutions incorporating learning analytics into their processes?’, ‘What is the impact of the use of learning analytics in our institutions for students, teachers and managers?’ need to be addressed. Finally, the presentation of examples of application of learning analytics from research projects in Europe and Latin America invited reflection on the potential of learning analytics as the catalyst for the transformation of Higher Education.

- In “*Process mining in Education: Current state and opportunities*”, Manuel Lama (University of Santiago de Compostela) focused on process mining techniques in Higher Education. Process mining aims to understand what is really happening in a process from the data generated over time. In the last decade these techniques have been successfully applied to several application domains, such as industry, public administrations or finance. Nonetheless, the use of process mining in education has been relatively low due, among other reasons, to the need to adapt and to make flexible the educational processes to the profile and behavior of the students. The presentation focused on presenting the main opportunities that process mining techniques offer to facilitate decision making by teachers and managers in Education.
- In “*Big data-based technology implementation at UNED: Ethical considerations*”, José Luis Aznarte (The National Distance Education University, UNED) brought forward some of the most up-to-date topics in learning analytics: privacy, ethics of data collection, handling and analysis, and limits and good practices in learning analytics. Dr. Aznarte shared his experience at UNED and described a roadmap, strategies and an evidence-based framework for the definition and implementation of ethically responsible learning analytics in Higher Education. Using the ongoing ED³ project at UNED as an example, Dr. Aznarte advocated for a participative process involving all learning agents to design an ethical framework of data use, collection and curation, analysis, intervention and predictive modelling.

The keynote by Dr. Aznarte served as starting point for discussion of pending issues and the future of learning analytics in a roundtable under the title “*Learning Analytics in Higher Education: Opportunities, threats, strengths and weaknesses*”, sponsored by the IEEE Spain Section and promoted by the Spanish Chapter of the IEEE Education Society. The discussion panel included key representatives of European Higher Education institutions, bringing together academic research and policymaking: Manuel Caeiro-Rodríguez (Universidade de Vigo), José Luis Aznarte (The National Distance Education University, UNED), Óscar Rubiños, (Universidade de Vigo), Pedro Muñoz-Merino (Universidad Carlos III de Madrid), Ángel Hernández-García (Universidad Politécnica de Madrid) and Mar Pérez-Sanagustín (Université Paul Sabatier Toulouse III). Beginning with an overview of ethical principles for the application of learning analytics in the discussants’ institutions, the talk then shifted to the differentiation between ethical and legal frameworks and the need for definition of transparent ethical

frameworks in learning analytics. At this point, the discussants debated about the difference between educational *data* (given) and *capta* (captured or collected), and the necessity to address this difference and only use *data* in learning analytics. The discussion then moved to the potential of (but also difficulty in) data integration across the institutions to perform smart and responsible use of educational data, and the need to adapt any valid framework to the specific context of the Higher Education institution.

LASI Spain 19 also welcomed the celebration of a workshop directed by Pedro J. Muñoz-Merino and Mar Pérez-Sanagustín under the theme “LALA Project: Connecting Europe and Latin America for Learning Analytics”. The workshop presented the results and ongoing research within the LALA project³ and included the presentation of the LALA framework, which aims to guide the design, implementation and use of learning analytical tools in Higher Education institutions in Latin America. Participants had the chance to have a hands-on practice of the application of the LALA framework to the specific context of their institutions.

Finally, the academic community attending LASI Spain 19 had the opportunity to describe and discuss recent scholar developments in the field in two sessions that included selected research studies from the open call for papers for the conference, three of which were presented the first day of the conference with the remaining five being presented on the second day. These proceedings include the 8 selected contributions after double-blinded peer review. The remainder of this preface summarizes the studies presented at LASI Spain 2019 in order of presentation, and shows the diversity of approaches to learning analytics in Higher Education:

“Application of Learning Analytics techniques on blended learning environments for university students” (Sheila Lucero Sánchez-López, Rebeca P. Díaz-Redondo and Ana Fernández-Vilas) presents an exploratory analysis of student activity in different Moodle modules. The data uses log-data of three cohorts of university students in a programming course, and differentiates between actions related to content or class notes and actions related to interpersonal activities. The analysis consists on a category-based classification that differentiates between ‘code’, ‘content’ and ‘course administration’, and applies content analysis using two corpora —code and content. The authors conclude that the analysis of messages can provide insightful feedback to instructors and help identify topics of interest or those that are not being completely learnt, information that may be used for remediation practices.

“Using Simva to evaluate serious games and collect game learning analytics data” (Cristina Alonso-Fernández, Iván José Perez-Colado, Antonio Calvo-Morata, Manuel Freire-Morán, Ivan Martínez-Ortiz and Baltasar Fernández-Manjón) describes *Simva*, a tool that may be employed to validate serious games using pre-post experiments. The study describes the application of *Simva* in three different serious games, after presenting the architecture of the tool. The study proposes a pragmatical approach to run controlled experiments, orchestrating conditions and data collection instruments embedded in serious games. The authors argue about the potential of *Simva* to simplify the validation of serious games and student assessment, but they also identify key issues when

³ <https://www.lalapproject.org>

conducting experiments in real settings to validate serious games: ensuring users privacy, the heterogeneity of data sources and the difficulty in transitioning from pre-post experiments to Game Learning Analytics.

“Extending a dashboard meta-model to account for users’ characteristics and goals for enhancing personalization” (Andrea Vázquez-Ingelmo, Francisco José García-Peñalvo, Roberto Therón and Miguel Ángel Conde) presents the extension of a meta-model for dashboard personalization. The meta-model extends a generic dashboard to account for users’ characteristics (preferences, disabilities, knowledge about different domains, visualization literacy and bias, including action, perceptual or social bias) and goals, which can be broken down into individual and more specific tasks. The conceptual study discusses the pivotal role of characteristics and goals in user-personalization of dashboards for learning analytics in order to provide users with a complete view of the data necessary for decision making and to foster self-regulated learning and improve academic achievement.

“Predicting student performance over time. A case study for a blended-learning engineering course” (Juan Antonio Martínez, Joaquim Campuzano and Teresa Sancho-Vinuesa) proposes a comparison of prediction models to test their accuracy as estimators of at-risk students. The study uses data from out-of-school activities of a first-year engineering course that follows a flipped-classroom methodology. The results show that the longer the period of analysis, the more accurate the models are, but also suggest that early periods lack accuracy and would not be optimal for interventions.

“Analyzing Students’ Persistence using an Event-Based Model” (Pedro Manuel Moreno-Marcos, Pedro J. Muñoz-Merino, Carlos Alario-Hoyos and Carlos Delgado-Kloos) proposes a way to measure students’ persistence when solving specific exercises. The study considers that students show persistence when they do not give up after failing an exercise. The analysis uses data from different courses and divides students into two groups: high-persistence and mid-persistence students. The study finds a positive correlation between this type of persistence and average grades, but concludes that there is no relationship between persistence and dropout or video visualizations in the educational scenarios under analysis.

“A Data Value Chain to Support the Processing of Multimodal Evidence in Authentic Learning Scenarios” (Shashi Kant Shankar, Adolfo Ruiz-Calleja, Sergio Serrano-Iglesias, Alejandro Ortega-Arranz, Paraskevi Topali and Alejandra Martínez-Monés) presents and examines four different multimodal learning analytics (MMLA) scenarios under the lens of the data value chain (DVC). The four scenarios provide a wide view of the complexity and heterogeneity of applying MMLA in different learning environments. The proposal of the data value chain identifies a total of seven multimodal data processing activities, divided into three groups: (i) data discovery related to the collection, annotation, curation, structuring and transformation of heterogeneous datasets (collect & annotate, prepare, and organize); (ii) data fusion, focused on integration of different datasets to generate a coherent view of multimodal evidence; and (iii) data exploitation, which includes analysis-related activities (analysis, visualization and decision-making).

“Predictors and Early Warning Systems in Higher Education - A Systematic Literature Review” (Martín Liz-Domínguez, Manuel Caeiro-Rodríguez, Martín Llamas-

Nistal and Fernando Mikic-Fonte) introduced a review of the literature of the use of learning analytics techniques applied to early-warning systems. Liz-Domínguez et al. conclude that most of the existing research focuses on predictive algorithms and tools to detect and analyze at-risk students; e.g. a student failing or dropping out a course. The study reviews the defining characteristics of existing predictive models considering input data, prediction goal and key aspects related to their use in practice, and shows that this is a hot topic in the current learning analytics landscape.

“Predicting early dropout student is a matter of checking completed quizzes: the case of an online statistics module” (Josep Figuerola-Cañas and Teresa Sancho-Vinuesa) deals with the prediction of student dropout. Predictive learning is one of the most popular application of learning analytics, while dropout is one of the most studied problems in learning and instruction in Higher Education. This study proposes easy-to-use classifiers based on decision trees to detect students at risk of dropout. The predictor variables include quiz results and forum activity, but only quiz results —particularly, completion of quizzes along the course period— were significant for the prediction.

Acknowledgements

The authors want to thank the financial support provided by TELGalicía (under project ED431D 2017/12) and the IEEE Spain section. We also thank the members of the Organization Committee and Scientific Programme Committee for their dedication and knowledge, as well as all the authors who submitted their valuable contributions to LASI Spain 19.

LASI Spain 19 Committees

General Chair

Manuel Caeiro-Rodríguez
Universidade de Vigo, Spain

Programme Committee Chairs

Pedro J. Muñoz-Merino
Universidad Carlos III de Madrid, Spain
Ángel Hernández-García
Universidad Politécnica de Madrid, Spain

Organization Chairs

Martín Llamas Nistal
Universidade de Vigo, Spain
Fernando Mikic-Fonte
Universidade de Vigo, Spain
Martín Liz Domínguez
Universidade de Vigo, Spain
Andrea Vázquez-Ingelmo
Universidad de Salamanca, Spain
Juan Manuel Santos-Gago
Universidade de Vigo, Spain

Scientific Committee

Ainhoa Álvarez
Euskal Herriko Unibertsitatea, Spain
Miguel L. Bote-Lorenzo
Universidad de Valladolid, Spain
Ruth Cobos
Universidad Autónoma de Madrid, Spain
Miguel Á. Conde
Universidad de León, Spain
Juan Cruz-Benito
Universidad de Salamanca, Spain
Davinia Hernández-Leo
Universitat Pompeu Fabra, Spain
Mikel Larrañaga
Euskal Herriko Unibertsitatea, Spain

Alejandra Martínez-Monés

Universidad de Valladolid, Spain

Rafael Pastor-Vargas

Universidad Nacional de Educación a Distancia, Spain

Cristóbal Romero

Universidad de Córdoba, Spain

Salvador Ros

Universidad Nacional de Educación a Distancia, Spain

Teresa Sancho-Vinuesa

Universitat Oberta de Catalunya, Spain

Andrea Vázquez-Ingelmo

Universidad de Salamanca, Spain

Mikel Villamañe

Euskal Herriko Unibertsitatea, Spain