Achievements and challenges in learning analytics in Spain: The view of SNOLA

(Logros y retos en analítica del aprendizaje en España: La perspectiva de SNOLA)

Alejandra Martínez-Monés
Yannis Dimitriadis
Emiliano Acquila-Natale
Ainhoa Álvarez
Manuel Caeiro-Rodríguez
Ruth Cobos
Miguel Ángel Conde-González
Francisco José García-Peñalvo
Davinia Hernández-Leo
Iratxe Menchaca Sierra
Pedro J. Muñoz-Merino
Salvador Ros
Teresa Sancho-Vinuesa

DOI: http://dx.doi.org/10.5944/ried.23.2.26541

How to reference this article:
Abstract

As in other research fields, the development of learning analytics is influenced by the networks of researchers that contribute to it. This paper describes one of such networks: the Spanish Network of Learning Analytics (SNOLA). The paper presents the research lines of the members of SNOLA, as well as the main challenges that learning analytics has to address in the next few years as perceived by these researchers. This analysis is based on SNOLA’s archival data and on a survey carried out to the current members of the network. Although this approach does not cover all the activity related to learning analytics in Spain, the results provide a representative overview of the current state of research related to learning analytics in this context. The paper describes these trends and the main challenges, among which we can point out the need to adopt an ethical commitment with data, to develop systems that respond to the requirements of the end users, and to reach a wider institutional impact.

Keywords: groups and organizations; data processing; trend.

Resumen

Tal y como ocurre en otros campos de investigación, el desarrollo de la analítica del aprendizaje está influído por las redes de investigadores que contribuyen al mismo. Este artículo describe una de estas redes: la Red Española de Analítica de Aprendizaje (SNOLA). El artículo presenta las líneas de investigación de los miembros de SNOLA, así como los principales retos que la analítica del aprendizaje tiene que afrontar en los próximos años desde la visión de estos investigadores. Este análisis está basado en datos de archivo de SNOLA y en una encuesta realizada a los actuales miembros de la red. Aunque esta aproximación no cubre toda la actividad relacionada con analítica del aprendizaje en España, los resultados proporcionan una visión general representativa del estado de la investigación relacionada con analítica del aprendizaje en dicho contexto. El artículo muestra cuáles son estas tendencias y los principales retos, entre los que se encuentran la necesidad de adoptar un compromiso ético con los datos, desarrollar sistemas que respondan a las necesidades de los usuarios y alcanzar mayor impacto institucional.

Palabras clave: grupos y organizaciones; procesamiento de datos; tendencia.
people, whereas in some cases they were created on the basis of a founded project. One example of these officially recognized networks at the European level is the Learning Analytics Community Exchange (LACE) (LACE, 2020), which contributed with some outstanding work between the years 2015 and 2017 and is still active in connection with SOLAR.

In Spain a group of researchers working in the LA field formed the Spanish Network of Learning Analytics (SNOLA) (SNOLA, 2020) in 2013. This network was officially recognized in 2015 by the Spanish Government. In 2020 SNOLA has started a new period of activity as an officially recognized thematic network, with new groups that add more variety and quality to the network. Although SNOLA does not represent all the work being done in LA in Spain, it comprises a wide representation of groups. The description of the work done by its members provides a good picture of what is being done in LA research in Spain, in line with recent works providing a more general overview of the field (Papamitsiou, Giannakos, & Ochoa, 2020).

Scientific networks on specific areas can be developed at different levels, e.g. global/regional/national, or coming from other existing networks and societies. The analysis and study of the networks at different levels, origins or focuses may be relevant in order to understand the trends and challenges that are in the center of attention of these networks. This is especially relevant for the field of LA that emerged in the last ten years as a consequence of the interest of a number of researchers that were working in related research areas (i.e., Artificial Intelligence in Education, Computer Supported Collaborative Learning, Educational Data Mining, or Technology Enhanced Learning), each one of them with its own community behind. Although some studies have been undertaken in order to provide an overview of the global landscape in LA (Ferguson et al., 2016), it is still important to further analyze and understand the evolution, trends and challenges that are detected in an active network at a finer grain level (national).

This paper presents SNOLA: its motivation, goals, activities, and expected outcomes in the following years. It provides an overview of the work carried out by its members, giving a partial but meaningful panorama of current research and challenges of LA in Spain. This overview is based on archival information about the activity of the network in the past years and its working plan for the future. As a second contribution, the paper presents an outline of the challenges that should be faced by the LA research field in the next few years. The description of the activity of the groups and their vision of the challenges was collected through an open-ended survey that was filled out by the representative of each group in the network. Using this survey, we have collected the main research lines of each group related to LA, as well as the participants’ view of the main challenges of the LA research field in the near future.

The structure of the paper is as follows. First, we describe SNOLA, its motivation and main goals, the activities it has carried out in the past and what impact is expected from the network in the future. Then, the research groups belonging to
SNOLA are introduced. Third, we describe the work done by its members in the last 5 years, followed by a summary and discussion of the main challenges for LA described by its members.

**SNOLA – MOTIVATION AND ACTIVITIES**

Education is a key element in our society. Institutional reports by the European Commission (EC, 2016) have pointed out the potential of LA for the improvement of the quality of the European education system, but also the fact that many research questions and technical issues need to be solved to reach this goal. The need to face these issues has led the European Union to make a series of recommendations for the field of LA (Ferguson et al., 2016). One of these recommendations is the need for structured organizations to support the use of LA; the urgency of bringing together all relevant stakeholders to collect and analyze the evidence about the present and expected benefits of LA; and to identify the needs of the implied actors, especially students and teachers.

SNOLA appeared with the main aim to face these demands. It was established informally in 2013 by Spanish researchers interested in the interchange of knowledge and expertise in the area of LA. Its mission was to create synergies and explore possibilities of collaboration among them, as well as becoming a reference in the field of LA in Spain. SNOLA also aimed to promote the connection between Spanish researchers with other international networks in different regions of the world, including Latin America, Australia, US and Europe. The work of the network was reinforced by its official recognition as a Thematic Network of Excellence by the Spanish Government in 2015 and in 2020.

One of the main activities of the network since its creation has been the organization of scientific events that have served to disseminate the work of the network and to extend the connections to other researchers, both in Spain and at an international level. Since 2013, SNOLA has organized numerous events: the series of Learning Analytics Summer Schools, (LASI Spain) which have taken place annually since its first edition in 2013, in Madrid, up to its 2019 edition in Vigo (Caeiro-Rodríguez, Hernández-García, & Muñoz-Merino, 2019).

These networking activities have helped to extend the network to other forums, such as the events organized by the e-Madrid network, the organization of the special track on LA in the TEEM conference from 2013 (Conde-González & Hernández-García, 2013) to 2019 (Conde-González & Hernández-García, 2019), and the organization of the workshops WLA-CISTI and WAPLA. This intense activity has been possible due to the existence of the network itself. The communication and coordination channels set up by the SNOLA have facilitated to build up working teams to organize these events in a smooth and efficient fashion, which would have been not possible without the existence of the network.
The network has maintained contact with relevant researchers in the field of LA at the international level, who have offered webinars organized by the network, or have participated as keynote speakers at different editions of LASI and the special tracks on LA organized by SNOLA. These connections have materialized as well in the participation of some of the members of the network in international projects. For example, in Europe, the project SHEILA (Tsai et al., 2018) has worked in the deployment of LA policies at an institutional level. This project has had a continuation in Latin America with the project of Learning Analytics in Latin America (LALA, 2020).

Finally, it is interesting to note that the number of organizations (companies, public administrations, universities, educational institutions in general) and persons interested in the network goes beyond the list of the official members of SNOLA. The SNOLA webpage (SNOLA, 2020) has become a hub for information about LA in Spain, with news, resources, software, and publications that provide a good overview of the current state of research in the area in our context.

**SNOLA COMPONENTS**

Besides the networking and dissemination activities, the value and richness of the network is derived from the activity of its members. This section describes the groups that belong to SNOLA and the main research lines in which they are involved.

SNOLA, as an officially recognized thematic network, comprises 12 research groups, all of them belonging to different universities across Spain. These groups are summarized in Table 1, which includes a link to their URL. As it can be observed, most of its members come from a technological background, from pure Information Technology (IT)-related areas like Telematics Engineering, Computer Languages and Systems, Artificial Intelligence, and Computer Architecture; or applied areas, such as Didactics of Mathematics or Business Organization. Therefore, SNOLA as a whole has been traditionally more focused on technology-oriented contributions rather than other aspects which are also important (e.g. pedagogical, philosophical, etc.). This has been also the case with the wider LA field at an international level. However, the community is fully aware of the need to consider the ethical, social, and pedagogical issues raised by the domain in which these technological contributions are presented. The increasing weight of these issues in current discussions in the field points out to the need of a more interdisciplinary approach in the composition of the groups and of the network. A challenge of SNOLA is to attract researchers that come from social and pedagogical fields to collaborate closely with its members, in order to face the more global challenges posed nowadays in LA. These challenges are discussed more deeply in the following section.

Table 1 evidences a considerable experience in LA of the members of SNOLA, given the range of years in the field (from 5 to 20 years). Some of these groups were already working in research problems that are now considered as part of LA.
years before the field was recognized as a separate research area with this name. This is explained by the fact that most of the groups that belong to SNOLA joined the specific LA field from a broader research field related to the use of computers in education, like e-learning, Technology Enhanced Learning (TEL), Computer Supported Collaborative Learning (CSCL), Educational Data Mining (EDM) or Artificial Intelligence in Education (AIED).

Table 1. Description of the members of SNOLA.

<table>
<thead>
<tr>
<th>GROUP</th>
<th>URL</th>
<th>Knowledge Area</th>
<th>University</th>
<th>Years in the area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deusto LearningLab</td>
<td><a href="http://dtlearning.deusto.es">http://dtlearning.deusto.es</a></td>
<td>New technologies applied to Education</td>
<td>UDeusto</td>
<td>7</td>
</tr>
<tr>
<td>Group for Adaptive Teaching-Learning Environments (GaLan)</td>
<td><a href="https://galan.ehu.eus/Galan">https://galan.ehu.eus/Galan</a></td>
<td>Computer Languages and Systems</td>
<td>UPV/EHU</td>
<td>8</td>
</tr>
<tr>
<td>Telematics Systems and Engineering (GIST)</td>
<td><a href="https://bidi.uvigo.es/es/grupo/gist-grupo-de-ingenieria-de-sistemas-telematicos">https://bidi.uvigo.es/es/grupo/gist-grupo-de-ingenieria-de-sistemas-telematicos</a></td>
<td>Telematics Engineering</td>
<td>UVigo</td>
<td>7</td>
</tr>
<tr>
<td>Gradient Lab – Group of Telematic Applications and Services (GAST)</td>
<td><a href="http://gradient.gast.it.uc3m.es">http://gradient.gast.it.uc3m.es</a></td>
<td>Telematics Engineering</td>
<td>UC3M</td>
<td>17</td>
</tr>
<tr>
<td>GRoup of Interaction And e-Learning (GRIAL)</td>
<td><a href="http://grial.usal.es">http://grial.usal.es</a></td>
<td>Computer Science and Artificial Intelligence, Computer Languages and Systems, Research Methods and Assessment in Education</td>
<td>USal</td>
<td>11</td>
</tr>
<tr>
<td>GROUP</td>
<td>URL</td>
<td>Knowledge Area</td>
<td>University</td>
<td>Years in the area</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Intelligent and Collaborative Systems Group (GSIC)</td>
<td><a href="http://www.gsic.uva.es">http://www.gsic.uva.es</a></td>
<td>Computer Languages and Systems, Telematics Engineering, Didactics and School Organization</td>
<td>UVa</td>
<td>20</td>
</tr>
<tr>
<td>Learning Analytics for Innovation and Knowledge Application in Higher Education (LAiKA)</td>
<td><a href="http://laika.blogs.uoc.edu">http://laika.blogs.uoc.edu</a></td>
<td>Online mathematics education, Computer Science</td>
<td>UOC</td>
<td>6</td>
</tr>
<tr>
<td>Research Group on Interactive and Distributed Technologies for Education (TIDE)</td>
<td><a href="https://www.upf.edu/web/tide">https://www.upf.edu/web/tide</a></td>
<td>Telematics Engineering, Computer Science, Cognitive Sciences</td>
<td>UPF</td>
<td>6</td>
</tr>
</tbody>
</table>

*For each member, the name of the Group, its URL, the knowledge area of its members, the university to which they belong, and the number of years working in LA and related areas previous to 2011*

**SNOLA LINES OF RESEARCH**

In order to provide an overview of the work done by the network, the groups reported their research lines in an internal survey, which was inspired by the classification of LA systems presented in Omedes (2018). A total of 34 distinct research lines were identified. For each research line, the following data has been collected: brief description; who is the main user of the system/functionality; educational level to which it has been applied/thought of; data sources and analysis techniques; the type of system (whether it is descriptive or prescriptive); and the
main expected benefits of the system. Regarding this feature, we started with the four possible values suggested by Omedes (2018), and added three additional ones that emerged from the data provided by the partners. This set of categories are shown in Table 2, with a description of their meaning, the code we will use to refer to them in this section, and the number of research lines that were reported to fall under each research goal.

Table 2. Categories and codes*

<table>
<thead>
<tr>
<th>Code</th>
<th>Expected benefit</th>
<th>Description</th>
<th>Num of lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILRP</td>
<td>Increase learner retention and performance</td>
<td>Reduce dropout rates and increase students’ performance. Having the right insights allow for performing proactive tutoring and intervention.</td>
<td>26</td>
</tr>
<tr>
<td>IQCCLE</td>
<td>Improve the quality of the content, the course or the learning environment</td>
<td>Discover content consumption patterns, understand content quality issues, and provide personalized learning experiences (adaptive learning).</td>
<td>17</td>
</tr>
<tr>
<td>PDS</td>
<td>Proactively drive success</td>
<td>Identify and promote success factors as well as understand students’ pathways leading to graduation (curriculum design).</td>
<td>11</td>
</tr>
<tr>
<td>ACE</td>
<td>Allocate costs effectively</td>
<td>Help in discovering which resources work and which don’t. Selective investment strategies may be better designed based on our analytics.</td>
<td>3</td>
</tr>
<tr>
<td>IIL</td>
<td>Identify indicators for learning</td>
<td>Define proxies for learning based on data.</td>
<td>7</td>
</tr>
<tr>
<td>CBLM</td>
<td>Create better learner models</td>
<td>Identify elements of the learner model.</td>
<td>4</td>
</tr>
<tr>
<td>PRFA</td>
<td>Promote regulation, formative assessment of self-reflection about the learning process</td>
<td>Show elements of learning to the learner (or teacher) and enable their reflection about learning.</td>
<td>3</td>
</tr>
</tbody>
</table>

*Categories (and codes) used to describe the different research goals of a learning analytics system / functionality and how many times these goals were reported with a research line. The first four categories are the ones suggested by Omedes (2018). The categories in the last three rows emerged from the responses to the survey.

The data reported by SNOLA members has been classified according to their similarities in a number of categories. This classification is not orthogonal: some of the research lines might belong to different categories, but we present them in only one of them, to make the description more clear and simpler to follow.
The first category is composed of predictive LA systems. Prediction of learning variables, such as students’ performance or students at risk is one of the most well-known functionalities of LA applications (Peña-Ayala, 2018). It is therefore not surprising that it is the most frequently reported research line, with six lines (see Table 3). Some of these lines focus on the prediction of learning results and dropout in MOOCs (Moreno-Marcos et al., 2020; Cobos & Olmos, 2018), or on-line and blended courses (Martínez, Campuzano, Sancho-Vinuesa, & Valderrama, 2019; Agudo-Peregrina, Iglesias-Pradas, Conde-González, & Hernández-García, 2014). In general, these systems aim at deriving high-level indicators from the low-level data provided by the system logs (Alexandron, Ruipérez-Valiente, Chen, Muñoz-Merino, & Pritchard, 2017). One of such indicators is academic engagement (Bote-Lorenzo & Gómez-Sánchez, 2018). The users of these systems are students and teachers, who are expected to benefit from actionable information provided by the predictive systems. In some of the cases they also address managers in charge of the institutional learning environment, or researchers that aim at identifying indicators for learning (IIL) based on the low-level data provided by the system. In general, works related to prediction need a certain amount of data. For this reason, they are frequent in virtual learning contexts at scale, enacted on LMSs (like Moodle) or MOOCs, although other environments like version control systems are considered as well (Guerrero-Higuera, DeCastro-García, Rodríguez-Lera, Matellán, & Conde, 2019). As it can be expected, these research lines use different versions of predictive analysis techniques from artificial intelligence and machine learning. The main expected benefits from these research lines are to improve learner retention and performance (ILRP) as well as to improve the quality of the course content and the learning experience (IQCCLE).

Table 3. Research lines in SNOLA

<table>
<thead>
<tr>
<th>Research line</th>
<th>Publication</th>
<th>User(s)</th>
<th>Data sources</th>
<th>Analysis techniques</th>
<th>Type</th>
<th>Expected Benefit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prediction of learning results and dropout</td>
<td>(Moreno-Marcos et al., 2020)</td>
<td>S / T / M</td>
<td>Students’ actions on the system (MOOC)</td>
<td>Random Forest, Regression, Neural Networks, Decision Trees</td>
<td>P</td>
<td>ILRP, IQCCLE</td>
</tr>
<tr>
<td>Research line</td>
<td>Publication</td>
<td>User(s)</td>
<td>Data sources</td>
<td>Analysis techniques</td>
<td>Type</td>
<td>Expected Benefit</td>
</tr>
<tr>
<td>------------------------------------------------------------------------------</td>
<td>-------------------------------------------</td>
<td>---------</td>
<td>--------------------------------------------------</td>
<td>----------------------------------------------------------</td>
<td>------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>Prediction of learning results and dropout</td>
<td>(Cobos &amp; Olmos, 2018)</td>
<td>T / M</td>
<td>Students actions on the system (MOOC)</td>
<td>Predictive analytics, Machine Learning, Statistical analysis</td>
<td>P</td>
<td>ILRP, IIL, CBLM, PRFA</td>
</tr>
<tr>
<td>Identification of engineering students at risk</td>
<td>(Martínez et al. 2019)</td>
<td>S / T</td>
<td>Students actions on the system (Moodle and Virtual Campus)</td>
<td>Predictive analysis</td>
<td>P</td>
<td>ILRP, PDS</td>
</tr>
<tr>
<td>Actionable information based on prediction of academic engagement in MOOCs</td>
<td>(Bote-Lorenzo &amp; Gómez-Sánchez, 2018)</td>
<td>S / T</td>
<td>Students’ actions on the system (MOOC)</td>
<td>Feature selection, Machine Learning</td>
<td>P</td>
<td>ILRP, I QCCCLE</td>
</tr>
<tr>
<td>Analysis and classification of student interaction data with prediction purposes (Interactions)</td>
<td>(Agudo-Peregrina et al., 2014)</td>
<td>T / M / R</td>
<td>Student Activity (Moodle log data)</td>
<td>Log data classification, Regression</td>
<td>P</td>
<td>ILRP, I QCCCLE, IIL</td>
</tr>
<tr>
<td>Educational data mining</td>
<td>(Guerrero-Higuera et al., 2019)</td>
<td>S / T</td>
<td>Students actions on the system (version management system)</td>
<td>ML</td>
<td>P</td>
<td>ILRP, I QCCCLE</td>
</tr>
<tr>
<td>Definition of high-level actionable indicators based on low level data</td>
<td>(Alexandron et al., 2017)</td>
<td>S / T</td>
<td>Students’ actions on the system (MOOC)</td>
<td>ML, Artificial Intelligence Techniques, Semantic modelling, Heuristics</td>
<td>P</td>
<td>ILRP, I QCCCLE</td>
</tr>
</tbody>
</table>

*Research lines in SNOLA related to the prediction of dropout and students’ performance. Codes used: User(s): T/S/M/R/ID (Teachers/Students/Manager/Researchers/Instructional Designers); Type: D/P (Descriptive/Predictive); Expected Benefit: See Table 2*
Visual Analytics

A second category of research lines frequent in SNOLA is visual analytics (see Table 4). The focus on data visualization is also reported by Peña-Ayala (2018) as a main functionality of LA systems. In SNOLA, the groups that work in this research line aim at providing dashboards oriented to teachers (Gómez-Aguilar, García-Peñalvo, & Therón, 2014), students (Tobarra et al., 2014) or both (Ruipérez-Valiente, Muñoz-Merino, Leony & Delgado Kloos 2015). These systems are generally aimed at providing dashboards with indicators derived from the log data, but some of them are oriented to more specific purposes, like the one by Chaparro-Peláez, Iglesias-Pradas, Rodríguez-Sedano and Acquila-Natale (2019), which aims at supporting peer- and self-assessment processes. Most of these systems have a descriptive nature, although some others combine descriptive and prescriptive objectives (Tobarra et al., 2014). Sometimes these systems are also designed with the aim of helping managers allocate costs effectively (Cobos, Gil, Lareo, & Vargas, 2016), besides other outcomes, such as helping teachers identify indicators for learning. These systems use techniques specific to visual analytics, sometimes supported by heuristics or descriptive statistics. Some of the reported research lines are currently focused on frameworks to facilitate the automatic generation of visualizations, like the system by Vázquez-Ingelmo, García-Peñalvo, and Therón (2019), to generate dashboards adapted to the needs of different types of users, or the one by Hernández-García and Suárez-Navas, (2017) oriented to the creation of graph data based on social network analysis.

Table 4. Research lines related to Visual Analytics (See codes are the same as the ones described in Table 3)

<table>
<thead>
<tr>
<th>Research line</th>
<th>Publication</th>
<th>User(s)</th>
<th>Data sources</th>
<th>Analysis techniques</th>
<th>Type</th>
<th>Expected benefit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual analytics of eLearning systems (VeLA)</td>
<td>(Gómez-Aguilar et al., 2014)</td>
<td>T</td>
<td>Students’ actions on the VLE, Grades</td>
<td>Visual analytics</td>
<td>D</td>
<td>ILRP</td>
</tr>
<tr>
<td>LA Dashboards for virtual labs</td>
<td>(Tobarra et al., 2014)</td>
<td>S</td>
<td>Platforms logs</td>
<td>Heuristics</td>
<td>D/P</td>
<td>ILRP, PDS</td>
</tr>
<tr>
<td>Visual Analytics of students’ actions</td>
<td>(Ruipérez-Valiente, et al., 2015)</td>
<td>S / T</td>
<td>Students’ actions on the system (MOOC)</td>
<td>Visual analytics</td>
<td>D</td>
<td>ILRP, IQCCLE</td>
</tr>
<tr>
<td>Research line</td>
<td>Publication</td>
<td>User(s)</td>
<td>Data sources</td>
<td>Analysis techniques</td>
<td>Type</td>
<td>Expected benefit</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------</td>
<td>---------</td>
<td>--------------</td>
<td>---------------------</td>
<td>------</td>
<td>------------------</td>
</tr>
<tr>
<td>LA Dashboard for MOOCs</td>
<td>(Cobos et al., 2016)</td>
<td>T / M</td>
<td>Students’ actions on the system (MOOC), grades, demographics, self-reported data</td>
<td>Descriptive Statistics</td>
<td>D</td>
<td>PDS, ACE, IIL</td>
</tr>
<tr>
<td>Visualization of peer and self-assessment data in Moodle (MWDEX)</td>
<td>(Chaparro-Peláez, et al., 2019)</td>
<td>T</td>
<td>Peer-assessment grades (Moodle Workshops)</td>
<td>Visual Analytics</td>
<td>D</td>
<td>ILRP, IQCCLE, IIL</td>
</tr>
<tr>
<td>Automatic generation of adapted dashboards</td>
<td>(Vázquez-Ingelmo et al., 2019)</td>
<td>S / T / M / R</td>
<td>Multi-Dimensional Analysis (MDA), Machine Learning (ML)</td>
<td></td>
<td>D</td>
<td>ILRP, IQCCLE, PDS, ACE</td>
</tr>
<tr>
<td>Graph generation of educational data in online learning for social network analytics (GraphFES)</td>
<td>(Hernández-García &amp; Suárez-Navas, 2017)</td>
<td>T / M</td>
<td>Student activity (Moodle log data-Forums)</td>
<td>Social Network Analysis, Data visualization</td>
<td>D</td>
<td>ILRP, IQCCLE, IIL</td>
</tr>
</tbody>
</table>

**Support to active learning strategies**

The systems in the previous categories are based on distance learning and/or MOOCs, but LA is also used to promote other types of learning. Table 5 describes the lines of research related to this goal. These types of learning include collaborative learning (Amarasinghe, Hernández-Leo, & Jonsson, 2019); adaptive learning (Muñoz-Merino, Novillo & Delgado Kloos, 2018); peer feedback (Er, Dimitriadis, & Gaseviç, 2019); social learning (Claros, Cobos, & Collazos, 2015), and flipped classrooms (Rubio-Fernández, Muñoz-Merino, & Delgado Kloos, 2019). All these
proposals share their orientation for both teachers and students. In general, they are expected to improve learner retention and performance (ILRP); to proactively drive success (PDS), by analyzing the paths that make the students perform better and using this knowledge to make recommendations to the users; and at improving the quality of the course content and the learning experience (IQCCCLE). In many cases, the results of the LA system are used to improve the learning design (LD) of the courses. This link between LA and LD is gaining momentum in the field of LA and may have higher impact in the upcoming years.

A slightly different approach related to this category is the one taken by Manso-Vázquez, Caeiro-Rodríguez and Llamas-Nistal, (2018) that aims at defining design criteria for self-regulated learning, including the definition of the traces that should be generated by these systems according to the xAPI specification.

Table 5. Research lines related to the use of LA to support different types of online learning (see Table 3 for the description of the codes)

<table>
<thead>
<tr>
<th>Research line</th>
<th>Publication</th>
<th>User(s)</th>
<th>Data sources</th>
<th>Analysis techniques</th>
<th>Type</th>
<th>Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptive learning based on user models</td>
<td>(Muñoz-Merino et al., 2018)</td>
<td>S / T</td>
<td>Students’ actions on the system (Intelligent Tutoring Systems)</td>
<td>Bayesian networks, rules, Item Response Theory.</td>
<td>P</td>
<td>ILRP, IQCCCLE</td>
</tr>
<tr>
<td>Support to dialogic peer feedback (Synergy)</td>
<td>(Er et al., 2019)</td>
<td>S / T</td>
<td>Students actions on the system, content of the feedback</td>
<td>Descriptive statistics</td>
<td>D</td>
<td>ILRP</td>
</tr>
<tr>
<td>Social learning supported by learning analytics</td>
<td>(Claros et al., 2015)</td>
<td>S / T</td>
<td>Students actions on the system (content and social)</td>
<td>SNA, CSCL</td>
<td>D</td>
<td>IQCCCLE, PRFA</td>
</tr>
</tbody>
</table>
Learning analytics for Learning Design

The fifth category recognizes the importance of the relation between learning design and LA in a different direction: using LA to support LD processes. This is the common goal of the two research lines reported in Table 6. The system by Michos, Hernández-Leo and Albó (2018) uses the actions on a social network for teachers to support learning design processes, while the line reported by Wiley, Dimitriadis, Bradford and Linn (2020) aims at developing a framework for developing and evaluating learning analytics for learning design.

Table 6. Research lines related to the use of LA to support learning design processes (see Table 3 for the description of the codes)
Assessment support

Assessment is a central aspect of learning. In response to this fact, some of the groups in SNOLA have devoted their work to this topic. The work by Villamañe, Larrañaga and Álvarez (2017) aims at supporting these processes by providing indicators to assist evaluators in adjusting their grades. Others have focused on supporting the assessment of 21st century skills (Menchaca Sierra, Guenaga, & Solabarrieta, 2018), some of them with a focus on workgroup assessment (Tobarra et al., 2017; Conde, Colomo-Palacios, García-Peñalvo, & Larrucea, 2018; Hernández-García, Acquila-Natale, Chaparro-Peláez, & Conde, 2018).

Table 7. Research lines related to assessment (see Table 3 for the description of the codes)

<table>
<thead>
<tr>
<th>Research line</th>
<th>Publication</th>
<th>User(s)</th>
<th>Data sources</th>
<th>Analysis techniques</th>
<th>Type</th>
<th>Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition and adjustment of assessment processes</td>
<td>(Villamañe et al., 2017)</td>
<td>S / T / ID</td>
<td>Students’ answers, grades</td>
<td>Statistics, Regression, NNLS, Data visualization</td>
<td>P</td>
<td>IQCCLE, PDS, ACE</td>
</tr>
<tr>
<td>Learning analytics for the assessment of 21st-century skills</td>
<td>(Menchaca et al., 2018)</td>
<td>S / T</td>
<td>Grades</td>
<td>Heuristics</td>
<td>P</td>
<td>PDS</td>
</tr>
<tr>
<td>Analysis of Moodle logs for decision making and workgroup assessment</td>
<td>(Tobarra et al., 2017)</td>
<td>S / T</td>
<td>MOOC platform logs</td>
<td>Heuristic</td>
<td>D</td>
<td>ILRP, IIL, CBLM</td>
</tr>
</tbody>
</table>
Multimodal and contextual data

One of the challenges in LA is to improve the quality of the data used to derive indicators. In this order, one category of the research works reported in this study is devoted to the definition of new data sources for LA (see Table 7). These works explore the use of human-generated data like grades (Villamañe, Alvarez, & Larrañaga, 2020) or information provided by teachers in the learning design (Rodríguez-Triana, Martínez-Monés, Asensio-Pérez, & Dimitriadis, 2015); biometric signals captured with sensors (de Arriba-Pérez, Caeiro-Rodríguez, & Santos-Gago, 2018); or multimodal data including both self-reported and sensor-based data (Vujovic & Hernández-Leo, 2019). Some of these works are exploratory, and are aimed at identifying indicators for learning based on new data sources or at defining better learner models.

Table 8. Research lines related to the use of multimodal data and/or context-aware data (see Table 3 for the description of the codes)
Emotion and sentiment analysis

One of the ways to design improved learner models is the capability to model emotions. A group of research lines is devoted to this goal (see Table 9). For example, some of the works reported perform sentiment analysis on data taken from the posts of students in forums (Ros et al., 2017) or the actions of the students in MOOCs and MOOC contents (Cobos, Jurado, & Blázquez-Herranz, 2019). Another approach is the one taken by Ruiz, Urretavizcaya, Rodríguez and Fernández-Castro (2018) who modeled emotions based on self-reported data.
Table 9. Research lines related to emotion modelling and sentiment analysis data (see Table 3 for the description of the codes)

<table>
<thead>
<tr>
<th>Research line</th>
<th>Publication</th>
<th>User(s)</th>
<th>Data sources</th>
<th>Analysis techniques</th>
<th>Type</th>
<th>Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social and sentiment analysis</td>
<td>(Ros et al., 2017)</td>
<td>S / T</td>
<td>Forum messages</td>
<td>Heuristics</td>
<td>D</td>
<td>ILRP</td>
</tr>
<tr>
<td>Academic success prediction based on emotion modelling (PresenceClick)</td>
<td>(Ruiz et al., 2018)</td>
<td>S / T</td>
<td>Sensors, self-reported emotions</td>
<td>Transition matrix, Decision trees, Data visualization</td>
<td>P</td>
<td>ILRP, PDS</td>
</tr>
<tr>
<td>Sentiment Analysis</td>
<td>(Cobos et al., 2019)</td>
<td>T / M</td>
<td>Student actions on the system (MOOC), MOOC contents</td>
<td>Descriptive analytics, Natural Language Processing (NLP), Sentiment Analysis</td>
<td>D</td>
<td>ILRP, IQCCLPDS</td>
</tr>
</tbody>
</table>

A final research line is related to the definition of frameworks for the adoption of LA, enacted by the participation of one of the groups in the projects SHEILA (Tsai et al., 2018) and LALA. Both projects deal with one of the challenges posed to LA at a national and international level: institutional adoption of LA. The following section deals with this and other challenges for LA in the context of SNOLA.

MAIN CHALLENGES TO LA IN SPAIN FROM THE PERSPECTIVE OF SNOLA

One of the missions of SNOLA is to identify the research challenges that should drive the attention of the research and development in the field for the next few years. In the survey used for this study, the members of the network were asked to identify the main challenges faced by the LA field. A total of 36 challenges were identified by the respondents. The answers were analyzed and clustered around main topics. This section provides a brief summary of their responses.
Table 10. Main categories of challenges identified from the responses to the survey, and the frequency of the responses that could be attributed to that category

<table>
<thead>
<tr>
<th>Type of challenge</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethical, privacy and security issues</td>
<td>7</td>
</tr>
<tr>
<td>Increase adoption by end users</td>
<td>8</td>
</tr>
<tr>
<td>Quality of the process and the results</td>
<td>6</td>
</tr>
<tr>
<td>Personalization / Adaptation / Interoperability</td>
<td>5</td>
</tr>
<tr>
<td>Pedagogical challenges</td>
<td>5</td>
</tr>
<tr>
<td>Apply LA at an institutional level</td>
<td>5</td>
</tr>
</tbody>
</table>

- **Ethical, privacy and security issues were among the issues most frequently mentioned.** Overall, the survey shows that there is a strong concern about the need for an ethical commitment with the data by the researchers working in LA. There is a common understanding that our work has to comply with the new regulations about data protection—e.g. the General Data Protection Regulation (EU, 2016)—, and that all stakeholders need more training and tools to address this need (Drarschler & Greller, 2016). This is a global challenge in LA at an international level (Slade & Prinsloo, 2013), but it has local implications, related to national-level regulations and the ways that local users (individuals, schools, universities) in our context are (or are not) developing procedures to deal with ethical and data privacy and security issues.

- Partially related to the ethical commitment with data is the need to **increase the quality of the analytical processes and the results.** Some of the respondents pointed out that LA needs to provide valid and reliable results, so that users have the confidence to apply them in their practice. This challenge involves working in improved indicators, which provide more accurate proxies to learning, developing tools to check the validity and reliability of the LA-results, or collecting high-quality data. Some of the responses pointed out to the need of considering contextual issues to enrich the analysis and increase the validity of the results.

- Besides defining better learner models, another group of challenges addressed pedagogical issues, focusing on **improving real learning.** This goal is at the core of the definition of LA (Gašević, Dawson and Siemens, 2015). However, its achievement has proven to be especially difficult because—as pointed out by one of the participants—LA has to deal with complex sociotechnical systems. While the general idea is to keep on doing research about learning and building improved learner models, one alternative approach to deal with the overall difficulty to address learning is suggested by another participant “to focus on providing better tools for teachers, which will lead to better learning”.

- **Increase adoption by end users.** LA solutions need to address the demands of their users while considering their constraints. This goal could be achieved...
by means of co-design approaches that consider the users’ perspective from the outset, or by reinforcing teachers’ training, thus helping to develop a culture of data-based analysis. This focus on users is shared by Ferguson et al. (2016), who pointed out that the field of LA is nowadays more represented by the “supply side” (researchers offering tools and systems) than by the “demand side” (end users), and that this should be addressed in order for LA to provide more actionable applications.

- **Increase personalization / interoperability of data and tools.** A complementary approach to increase adoption is to focus on providing tools able to adapt / to be personalized to the needs of different types of users and contexts. In this regard, one particular issue that has to be solved is to design interoperable solutions that may enable the integration of the tools in existing educational systems and may be perceived as easier to adopt by end-users.

- **Take measures to apply LA at an institutional level.** Finally, several participants pointed out the need for organizational structures to support the use of LA and help educational leaders implement these challenges. Some specific actions in this line would be to set up an institution, such as a center for data driven studies, or to promote initiatives around LA at institutional levels, like CRUE (Conference of the Spanish University Principals), which is the main link between the Universities and the Spanish Government. This problem of encouraging a wider institutional adoption can be addressed by considering the recommendations given by the SHEILA project (Tsai et al., 2018), in which a SNOLA member participated as a partner.

**CONCLUSIONS**

SNOLA has played a main role in the promotion of LA in Spain since its foundation in 2013, through the activity carried out by the network, including the organization of conferences and LA-related events, as well as the contributions of each member group to the field. The review to the research lines reported by the members of SNOLA shows a multifaceted and intense activity, covering a wide range of topics and approaches and including connections with other LA research groups and organizations at an international level.

Most of the members of the network have a technical background. As a consequence, their research lines are mostly focused on technological contributions, as described in this paper. In spite of this bias, the members of SNOLA are aware of the global challenges in the field, are not restricted to technological issues. Among these challenges, the participants highlight the need for an ethical commitment with data; the design of systems that are able to adapt to the needs and demands of their users; and the establishment of a wider institutional framework to support and foster advances in the field. In order to face these challenges, the network might benefit from establishing connections with networks, groups, or individuals coming...
from complementary fields, such as pedagogy, psychology and/or philosophy. These problems are aligned with the ones pointed out by Pelánek (2020) in a recent review of the challenges of the field at an international level.

This work is just a first step towards the analysis of the impact of SNOLA in the development of LA in our context, which may be complemented with new analyses (i.e., a study on the structure of dynamics of the network), or with a deeper discussion considering research in LA at an international level.

ACKNOWLEDGEMENTS

This research has been co-funded by the National Research Agency of the Spanish Ministry of Science, Innovation and Universities and the Structural Funds (FSE and FEDER) under project grants RED2018-102725-T, TIN2017-85179-C3-1-R, TIN2017-85179-C3-2-R, TIN2017-85179-C3-3-R and TIN2016-80172-R; by FEDER/Castille and Leon Regional Government grant VA257P18; by the Basque Government under grant number IT980-16 and by the Catalan Government under grant number 2017SGR1619. This work has been co-funded by the Madrid Regional Government, through the project e-Madrid-CM (S2018/TCS-4307), the e-Madrid-CM project is also co-financed by the Structural Funds (FSE and FEDER). D. Hernández-Leo acknowledges the support by ICREA under the ICREA Academia programme.

REFERENCES


are about learning. *TechTrends, 59*(1), 64-71.
according to a proposed taxonomy. WIREs Data Mining and Knowledge Discovery, 8(3), e1243. https://doi.org/10.1002/widm.1243


Sound in Collaborative Learning, 891-892.

ACADEMIC AND PROFESSIONAL PROFILE OF THE AUTHORS

Alejandra Martínez Monés is Associate Professor at the Department of Computer Science (Universidad de Valladolid). She is a member of the GSIC research group. http://orcid.org/0000-0003-3201-0345
E-mail: amartine@infor.uva.es

Yannis Dimitriadis is a Full Professor at the Department of Signal Theory, Communications and Telematics Engineering, Universidad de Valladolid, and coordinator of the GSIC research group. http://orcid.org/0000-0001-7275-2242
E-mail: vannis@tel.uva.es

Dr. Emiliano Acquila-Natale is Assistant Professor at the Department of Organization Engineering, Business Administration and Statistics (School of Telecommunication Engineering, Universidad Politécnica de Madrid). http://orcid.org/0000-0003-2164-8386
E-mail: emiliano.acquila@upm.es

Ainhoa Álvarez is a faculty member of the Computer Languages and Systems’ Department at the University of the Basque Country UPV/EHU. http://orcid.org/0000-0003-0735-5958
E-mail: ainhoa.alvarez@ehu.eus

Manuel Caeiro Rodríguez is faculty member of the Department of Telematic Engineering at the University of Vigo. He is member of the research group of Telematic Engineering Systems. http://orcid.org/0000-0002-2784-6060
E-mail: mcaeiror@gmail.com

Ruth Cobos is Associate Professor at the Department of Computer Science Engineering of Universidad Autónoma de Madrid. She is member of the GHIA research group. http://orcid.org/0000-0002-3411-3009
E-mail: ruth.cobos@uam.es
Miguel Ángel Conde-González is Associate professor at the University of León. He is a member of the Robotics research group of the University of León and GRIAL research group of the University of Salamanca. http://orcid.org/0000-0001-5881-7775
E-mail: mcong@unileon.es

Francisco José García-Peñalvo is a Full Professor at the Department of Computer Science, University of Salamanca, member of the Research Institute for Educational Sciences and head of the GRIAL research group. http://orcid.org/0000-0001-9987-5584
E-mail: fgarcia@usal.es

Davinia Hernández-Leo is Full Professor and Icrea Academia Fellow at the Department of Information and Communication Technologies, Universitat Pompeu Fabra, and the head of the TIDE research group. http://orcid.org/0000-0003-9548-7455
E-mail: davinia.hernandez-leo@upf.edu

Iratxe Menchaca Sierra is Associate Researcher at Faculty of Engineering of University of Deusto of Bilbao. She is member of the Deusto LearningLab group. http://orcid.org/0000-0002-6594-6716
E-mail: iratxe.mentxaka@deusto.es

Pedro J. Muñoz-Merino in Associate Professor at the Department of Telematics Engineering at Universidad Carlos III de Madrid. http://orcid.org/0000-0002-2552-4674
E-mail: pedmume@it.uc3m.es

Salvador Ros is Associate Professor at UNED (Spanish Distance University) at the School of Computer Science. http://orcid.org/0000-0001-6330-4958
E-mail: sros@scc.uned.es

Teresa Sancho-Vinuesa is Full Professor at the School of Computer Science, Multimedia and Communication. She is the Director of the Degree on Applied Data Science. http://orcid.org/0000-0002-0642-2912
E-mail: tsancho@uoc.edu

Date of receipt: 24/01/2020
Date of acceptance: 27/02/2020
Date of layout: 11/04/2020