Online-Journals.org

iJET International Journal of Emerging Technologies in Learning

iJET | eISSN: 1863-0383 | Vol. 19 No. 7 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijet.v19i07.50313

PAPER

# A Systematic Review of Software for Learning Analytics in Higher Education

**Muhammad Afzaal**(✉), **Jalal Nouri**

Department of Computer and Systems Sciences, Stockholm University, Stockholm, Sweden

muhammad.afzaal@dsv.su.se

**ABSTRACT**

Learning analytics (LA) is an important area of study in technology-enhanced learning that has emerged during the last decade. In earlier years, several systematic reviews have been conducted that focused on the theories behind LA or on empirical studies that utilized LA-based methods to improve learning and teaching processes in higher education. However, to date, there has been no systematic review of papers that have adopted a software perspective to report on the many forms of learning analytics software (LAS) that have been developed, despite these being used more frequently than before in higher education to support learning and teaching processes. To fill this gap, this paper presents a systematic review of LAS with the aim of critically scrutinizing the ways in which the use of interactive software in real-world settings may both support students in improving their academic performance and assist teachers in various pedagogical practices. A thematic analysis of 75 articles was conducted, resulting in the identification of three categories of LAS: at-risk student identification (ARSI) software; self-regulation software; and collaborative learning software. For each of these categories, we analyzed (i) the embedded functionality; (ii) the stakeholder (teacher and student) for which the functionality is intended; (iii) the analytical and visualization approaches implemented; and (iv) the limitations of the software that require future attention. Based on the findings of our review, we propose future directions for the development of learning analytics software.

**KEYWORDS**

learning analytics (LA), learning analytics software (LAS), systematic review, identification of at-risk students, computer-supported collaborative learning (CSCL), self-regulated learning

## 1 INTRODUCTION

Learning analytics (LA) is a thriving domain within the field of technology-enhanced learning due to its potential to enhance both learning and teaching methods. Several systematic reviews [1–6] have been conducted that focus on the implementation of LA methods to improve learning and teaching processes.

Afzaal, M., Nouri, J. (2024). A Systematic Review of Software for Learning Analytics in Higher Education. *International Journal of Emerging Technologies in Learning (iJET)*, 19(7), pp. 17–43. https://doi.org/10.3991/ijet.v19i07.50313

Article submitted 2024-05-29. Revision uploaded 2024-07-11. Final acceptance 2024-07-11.

One salient implication is the development of learning analytics software (LAS), which is a software, tool, or dashboard that processes students' data using different approaches such as machine learning (ML), data mining, and analytics to improve the learning and teaching experiences and provide students and teachers with a snapshot of how they are progressing in their courses. In recent years, interest in LAS development has remained high, and higher education institutions serve as an ideal context for LAS due to the complexity of learning environments, resource optimization needs, the drive for continuous improvement, and compliance and accountability requirements [7]. Moreover, LAS in higher education empowers educators, administrators, and institutions with data-driven insights to support student success, enhance teaching and learning strategies, and make informed decisions. By leveraging the power of data, higher education institutions can improve student outcomes and foster a more personalized and effective learning environment.

The growth in the use of LAS in higher education has both provided an opportunity for studied and institutions to capture the digital footprints of students from learning management systems (LMSs) and enabled the identification of patterns within data to predict future events and make informed recommendations that can improve decision-making and learning outcomes [8]. Some types of LAS aim to trigger early performance alerts for at-risk students, whereas others aim to act as metacognitive tools that can make students aware of their learning performance and support self-reflection, and still others aim to leverage student interactions in a collaborative work setting. Previously, educational data mining (EDM) and artificial intelligence in education (AIED) were used to develop such software, although in recent times LA has become as prevalent as EDM and AIED. More than half of the existing LAS is used in higher education to scaffold students and assist teachers in offering adaptive or personalized guidance based on insights from LA [9]. These systems are expected to provide institutions with opportunities to boost student retention rates, support student progress, and, most importantly, provide personalized learning on a large scale in the near future. The developers of LAS are therefore particularly interested in understanding and optimizing learning and teaching processes by focusing on indicators of knowledge construction, creativity, self-directed learning, and self-regulation, in addition to monitoring academic progress [10, 11].

Recently, some studies [9, 12, 13] investigated the available forms of LA-based software, although the authors conducted only a superficial analysis of the theoretical view, design, and impact of such software. Despite continuous progress in LA study leading to the development of multiple LAS with various purposes, the existing literature lacks a comprehensive analysis of these LAS in which there is a focus on the implemented functionalities, the analytical approaches employed, and the limitations of these approaches. Our systematic review of real-world implementations of LAS aimed to both explore the techniques of LA used for data collection, analysis, and visualization and discover possible aspects that are not covered by existing software. Compared to other studies, the significance of the present work lies in the fact that the software reviewed here was implemented and evaluated in real educational environments, which constituted a criterion for the inclusion of the articles.

In this systematic review paper, we analyze several types of LAS that were developed to support both students and teachers and published in academic study repositories (ARRs). However, there may be many successful forms of LAS out there that are not mentioned in ARRs. Hence, the focus of this paper was on reviewing study-evaluated software that also reflects the different practices used by contemporary LA practitioners. The following three study questions (RQs) guided our review of learning analytics software:

**RQ1:** What types of LAS have been developed, and for what purposes?
**RQ2:** What functions are embedded in existing LAS?
**RQ3:** What analytics approaches are employed in LAS, and what are their limitations?

## 2 RELATED WORK

This section summarizes the findings of systematic reviews in which LA-based tools and applications were used to assist students and teachers in higher educational environments. Recently, several systematic reviews [1–6, 14] were published on LA in higher education. The purpose of these reviews was to examine the data presentation, the LA approaches used, the LA factors operationalized, the impact of learning interventions, and the data-driven learning design decisions made to improve student retention rates, learning outcomes, students' study success, and the learning process. Viberg et al. [6] also investigated the effectiveness of LA interventions for student underperformance and highlighted the limitations found based on the available evidence relating to LA interventions. These systematic reviews provided a broad view of current scientific knowledge about LA approaches; however, these approaches were not implemented in the form of software or tools.

Romero and Ventura [13] reviewed LA and educational software based on EDM techniques for three different modes of learning: traditional, computer-based, and blended learning. The purposes of their review were to both classify existing software into these three different modes and provide a short description of the datasets and LA approaches used in each mode. Matcha et al. [9] and Pérez-Álvarez et al. [12] conducted a systematic review of self-regulated learning (SRL) software and tools in which LA approaches were implemented using student data. Their work had a specific focus on SRL tools and software and, in particular, on different dimensions such as theory, design, feedback, impact, and quality. The most significant limitations of all these studies are a lack of exploration of the analytical approaches used for data analysis and the fact that none focused on the evaluation of software used in real-world settings.

Although the increasing amount of study on LA has led to the emergence of multiple types of LAS, to the best of our knowledge, there are no general and comprehensive reviews of LAS that have outlined the purposes and aims of the software applications used in university practices, the specific functions implemented, the analytical approaches employed, and their limitations.

## 3 METHODOLOGY

The primary focus of our systematic review was of explorations of LAS that had been published at the time our search was carried out (20 October 2021). The definition of LAS is a software, tool, or dashboard that processes students' data using different approaches such as ML, data mining, and analytics to improve the learning and teaching experiences and provide students and teachers with a snapshot of how they are progressing in their courses. The guidelines proposed by Kitchenham and Brereton [15] were followed, and the procedure used to conduct this systematic review is illustrated in Figure 1. It consisted of three steps: a literature search, a selection strategy, and an extraction process. Three independent studied performed the search process twice, first in January 2021 and again on October 20, 2021.

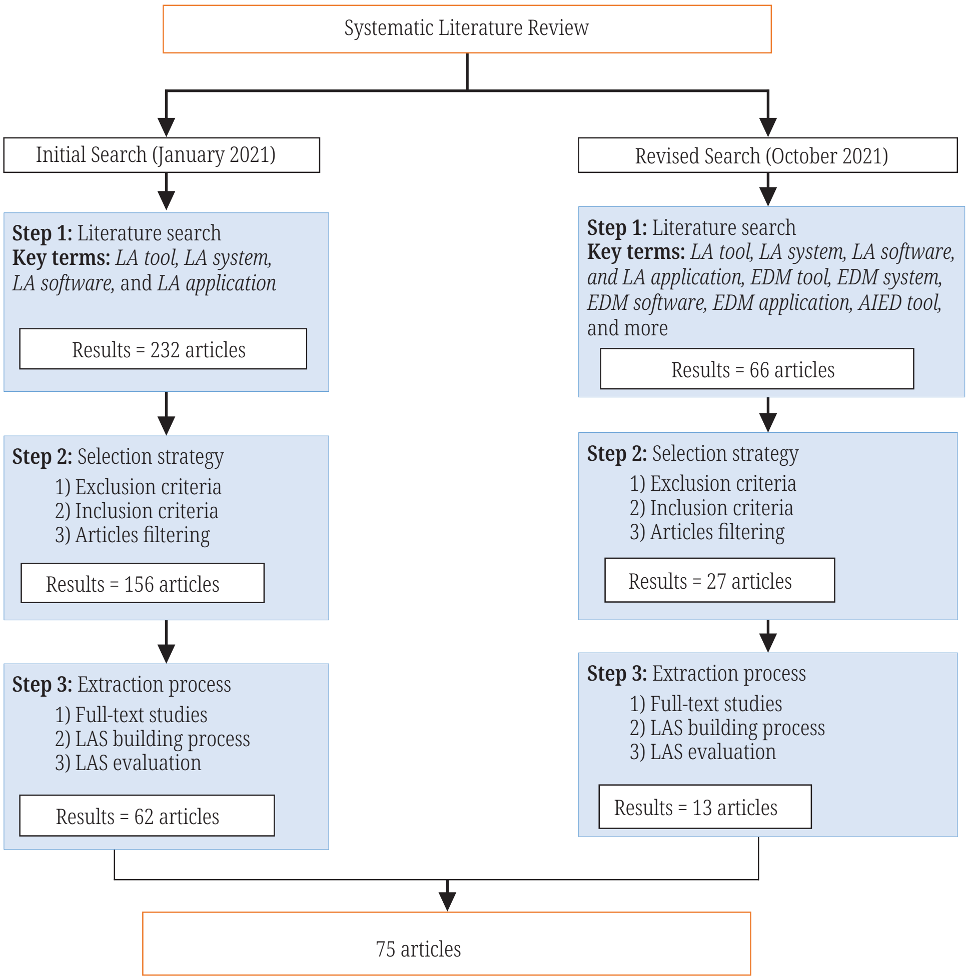


**Fig. 1.** Process of the systematic review

## 3.1 Literature search

In the first step, a pool of articles was retrieved via a systematic search of well-known and well-established databases in the fields of education and computing. This search was performed automatically and manually, using search strings derived from keywords defined based on the concepts of our study questions, as shown in Table 1. Using Boolean logical operators (OR, AND), each combination of keywords generated a set of search strings, as presented in Table 2. For the automatic search, we applied the proposed search strings to eight major bibliographic databases (IEEE Explorer, ACM Digital Library, ScienceDirect, Google Scholar, Wiley, Springer, Scopus, and Web of Science) with the intention of collecting evidence from reliable and high-quality sources. In the manual search, any related work that was identified while studying another paper was also added to our list of papers. The search strings or queries were tailored using a set of chosen keywords ("tool," "application," "software," "system," and "dashboard") along with logical conjunctions with emerging techniques, including LA, EDM, and artificial intelligence in education.

**Table 1.** List of keywords used for searches

| Type | Search Keywords |
|---|---|
| Research subject | software; application; system; tool; dashboard |
| Domains | learning analytics; LA; educational data mining; EDM; artificial intelligence in education; AI in education; AIED |

In the initial search conducted in January 2021, a total of 232 relevant articles were obtained by running the query: ("learning analytics" OR "LA") AND ("tool" OR "application" OR "software" OR "system" OR "dashboard"). For each query, after the top 30 or 40 records, the databases started to provide articles from other domains, such as medicine and engineering. A revised search was conducted in October 2021 in which two more queries were run, in addition to the previous ones: ("educational data mining" OR "EDM") AND ("tool" OR "application" OR "software" OR "system" OR "dashboard"), and ("artificial intelligence in education" OR "AI in education" OR "AIED") AND ("tool" OR "application" OR "software" OR "system" OR "dashboard"). The reason for adding these queries was the overlap between available papers on LA and those on EDM and AI in education. The revised search yielded 66 more articles.

**Table 2.** Search strings (queries)

| Search Strings |
|---|
| ("learning analytics" OR "LA") AND ("software" OR "system" OR "tool" OR "application" OR "dashboard") |
| ("educational data mining" OR "EDM") AND ("software" OR "system" OR "tool" OR "application" OR "dashboard") |
| ("artificial intelligence in education" OR "AI in education" OR "AIED") AND ("software" OR "system" OR "tool" OR "application" OR "dashboard") |

## 3.2 Selection strategy

The articles were selected based on exclusion and inclusion criteria, as shown in Tables 3 and 4. A three-step selection strategy was applied to identify potentially relevant articles. First, exclusion criteria were applied to remove all duplicate, non-English, incomplete (where only the abstract was available), short (poster), and gray literature, while emphasizing the selection of 232 full-text scientific journal articles and conference papers. Second, inclusion criteria were applied, in which the scope and study context of each article (higher education) were noted from the title, keywords, and abstract. Third, after applying these exclusion and inclusion criteria, the articles were filtered based on the introduction, headings, graphic information, tabular data, and conclusion. As a result, 156 articles were selected by the initial search in January 2021, and an additional 27 articles were selected from the revised search in October 2021.

**Table 3.** Exclusion criteria for literature selection

| Exclusion Criteria | Description |
|---|---|
| Duplicate articles | Same article retrieved from different databases |
| Non-English articles | Articles not written in English |
| Incomplete paper | Papers for which only the abstracts were available rather than the full text |
| Short paper | Posters or small conference papers with less than five pages |
| Grey literature | Other study materials such as reports and pre-prints |

### 3.3 Extraction process

In the final step of the search, an extraction process was applied to the filtered articles to remove model proposals, conceptual frameworks, and preliminary versions of articles (i.e., conference articles that were later extended to journal articles). By reading empirical analyses, experimentation details, and case studies, we extracted only articles containing a full-text study with detailed study of an LAS building process, implementation, and evaluation in real-world scenarios. As a result of this extraction process, a total of 75 articles were extracted, of which 62 were identified during the January 2021 search and 13 were identified during the October 2021 search.

**Table 4.** Inclusion criteria for literature selection

| Inclusion Criteria | Description |
|---|---|
| Quality papers | Peer-reviewed journal and conference articles |
| Scope | LAS in the domain of higher education |
| Research context | Higher education |

### 3.4 Analysis

The final 75 papers were read and analyzed in order to analyze the articles. This analysis focused on the name of the software, the category, the functions implemented, the analytical approaches used, and their limitations. In light of our study questions, two studied independently coded each article; we then calculated the percentage agreement between the coders and the value of Cohen's Kappa, a statistical measure [16]. For the 75 articles, the agreement between coders was 89%, and Cohen's Kappa was 0.67. Coding conflicts were first discussed between both coders and then reconciled for use in the final results. Based on this analysis, articles were thematized into three LAS categories (refer Table 5): (i) LAS for identifying at-risk students through early warning; (ii) LAS for computer-supported collaborative learning; and (iii) LAS for self-regulated learning. Overall, 91% of the assessed software explicitly mentioned the software category in the paper's purpose. The remaining 9% of the assessed software did not explicitly mention the category of the software; however, they did address one of the tasks of the mentioned categories; for example, Santos et al. [17] developed software that provides goal-oriented self-evaluation support to students; however, while in the text the authors did mention the software category, the tasks of goal setting and self-evaluation were tabulated under the self-regulated learning category of software.

These software categories we used are conventional and have been reported in the literature [18–20]. This analysis found that one function that provides reflection to students about their learning overlaps in three categories. For example, software for collaborative learning and self-regulated learning provides opportunities for reflection for students on their tasks. However, the utilization of data and the implementation of analytical approaches were contrasting.

**Table 5.** Learning analytics software and their respective categories

| LA Software | Reference | LA Software | Reference |
|---|---|---|---|
| **Category: At-Risk Student Identification** | | | |
| Course Signal | [21] | Advisory Dialogue | [33] |
| Student Success System | [22] | Student Engagement Viewer | [34] |
| Key Splitting Milestones | [23] | Adaptive Learning System | [35] |
| OU Analyse | [24] | CA-LAD | [19] |
| Student Activity Meter | [25] | GradeCraft | [36] |
| Early Warning Software | [26] | HEFCE | [37] |
| Student Explorer | [27] | LARA | [38] |
| IEWS | [28] | LAPLE | [39] |
| ALAS-KA | [29] | EMODA | [40] |
| LOCO-Analyst | [30] | I-LAD | [41] |
| MEWS | [31] | ITLAT | [42] |
| GAR-Based EWS | [32] | | |
| **Category: Computer-Supported Collaborative Learning** | | | |
| SNAPP | [43] | Dyad Body Posture | [20] |
| CanvasNet | [44] | BLINC | [53] |
| PyramidApp | [45] | FCA | [54] |
| MTClassroom | [46] | PIKMO System | [55] |
| ViLLE | [47] | Scripted-CL Sessions | [56] |
| Team Formation | [48] | SRES | [57] |
| Starburst | [49] | Metafora | [58] |
| COLLECE | [50] | SocialLearn | [59] |
| VCRI | [51] | Facilitators First | [60] |
| SST | [52] | SF-SLA | [61] |
| **Category: Self-Regulated Learning** | | | |
| OpenEssayist | [62] | Instrumentation Tools | [78] |
| Ask-Elle | [63] | C-LAD | [79] |
| Mastery Grids | [64] | TF-LAD | [80] |

*(Continued)*

**Table 5.** Learning analytics software and their respective categories *(Continued)*

| LA Software | Reference | LA Software | Reference |
|---|---|---|---|
| **Category: Self-Regulated Learning** | | | |
| AutoTeach | [65] | SR-LAD | [17] |
| LearnTracker | [66] | Learning Pulse | [81] |
| U-Behavior | [67] | ELAT | [82] |
| WIDE | [68] | RIDT | [83] |
| ESM | [69] | nStudy | [84] |
| AMBA | [18] | AcaWriter | [85] |
| Compod | [70] | OnTask | [86] |
| The Lifelong Learning Hub | [71] | CR-LAD | [87] |
| MyLA | [72] | LISSA | [88] |
| Mobile LAP | [73] | Learning skills dashboard | [89] |
| NoteMyProgress | [74, 75] | F-LAD | [90] |
| Next-TELL | [76] | SlimStampen | [91] |
| RiPPLE | [77] | DIFIAR | [92] |

# 4 RESULTS

In this section, we review the LAS-outlined embedded functions and LA approaches implemented in software, and we highlight the limitations of these approaches.

## 4.1 Research question 1: What types of learning analytics software have been developed, and for what purposes?

One of the main objectives of this systematic review was to identify the types of LAS that have been developed and implemented. Our analysis led to the identification of three categories: ARSI, CSCL, and SRL. Figure 2 depicts the number of LAS in each type or category and demonstrates that SRL software outnumbered ARSI and CSCL software. The purpose of each category of software is different; for example, ARSI software provides information to teachers about students who may fail, receive low grades, or drop out of a course, whereas CSCL software facilitates collaboration among students and teachers and supports teachers in the identification of isolated students and the creation of learning groups, while SRL software provides information that can help students' reflection and guide them to make necessary changes that lead to successful SRL. On the other hand, Figure 3 presents the year-wise summary of the reviewed LAS and shows a trend that demonstrates that the interest in developing new forms of LAS has increased over time, especially in recent years (e.g., 2020 and 2021).

We also performed an analysis to identify the maturity level and development settings of the reviewed LAS. Regarding the maturity level, Figure 4 shows that around 71% ($n$ = 53) of LAS were standalone versions, meaning they could operate independently, and 16% ($n$ = 12) of LAS were extension or additional plugins to an existing LMS. In contrast, the remaining LAS were still in the prototype stage or were subkits of different existing small-level tools that teachers and students commonly use in courses. Concerning deployment settings, 76% ($n$ = 57) of the analyzed LAS were deployed on students in different educational settings (shown in Figure 5). For example, the most common setting was undergraduate courses where students use a LAS throughout the course, while a few LAS were also deployed in PhD-level courses. Conversely, the remaining 24% of LAS were deployed with teachers and experts (e.g., studied or developers). These deployments had two main goals: (1) since 13% of LAS was developed for teachers only, it was obviously to help teachers or domain experts perform LAS evaluations; and (2) to understand the design and practical issues of these LAS before deploying them in a real educational setting.

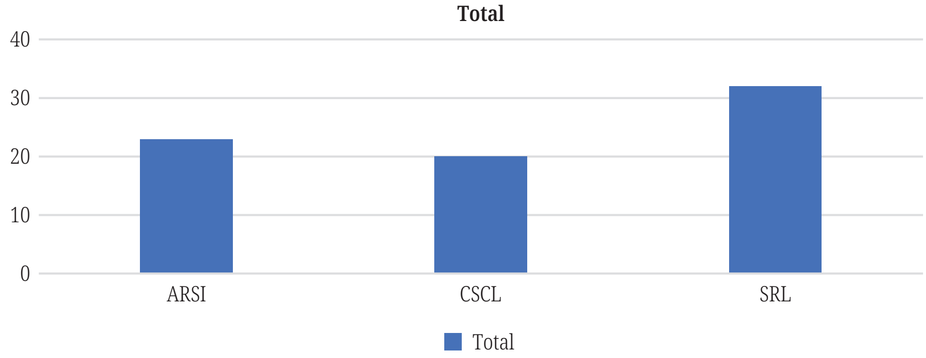


**Fig. 2.** Reviewed learning analytics software categories

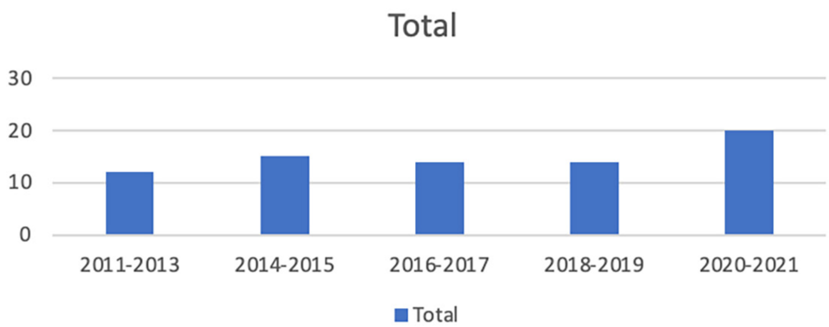


**Fig. 3.** Year-wise summary of reviewed learning analytics software

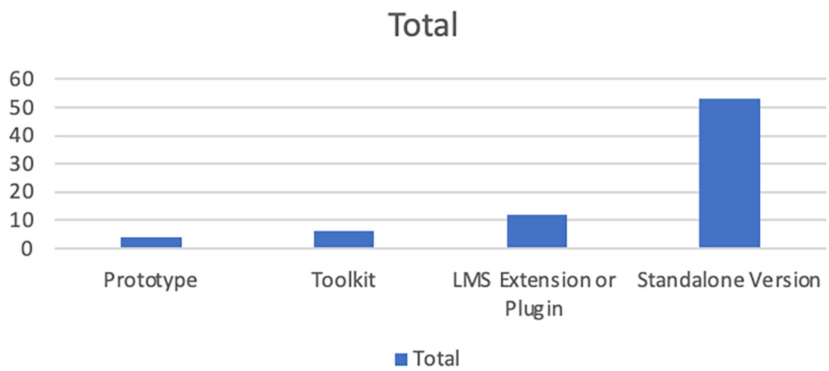


**Fig. 4.** Level of maturity of reviewed learning analytics software

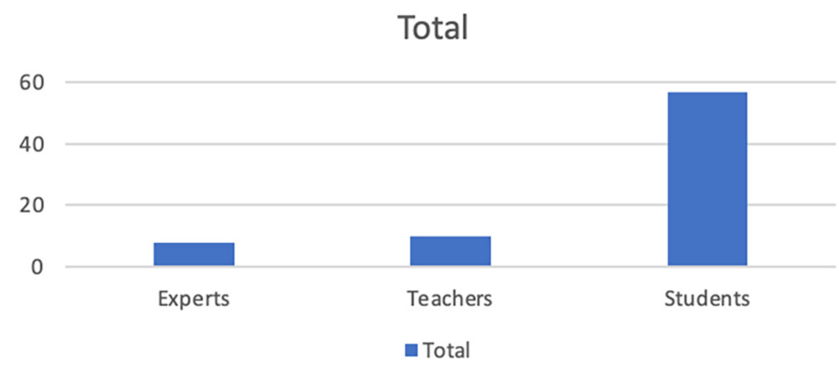


**Fig. 5.** Deployment settings of reviewed learning analytic software

## 4.2 Research question 2: What functions are embedded in existing learning analytics software?

At-risk student identification software offers several functions, assigning a risk status (e.g., at-risk, possibly at-risk, not at-risk) to each student based on learning activities—the primary one [31], de Quincey et al. [37]. The risk status function enables teachers to design an intervention that might help retain a high-risk student. Therefore, ARSI software offers functionality allowing teachers to create interventions for students based on their risk level, for example, in the form of a personalized email or a private message (Cavanagh et al. [35]). Other forms of ARSI software also allow teachers to make improvements by authoring study content, assessments, feedback, and study plans adapted to the identified students' needs (Weng et al. [34]). Moreover, ARSI software provides reflection functions for students and teachers based on students' activities in order to raise awareness and help students understand why they have been assigned a particular status [36, 42].

On the other hand, CSCL software with LA support offers a range of functions, such as identifying isolated or disengaged students [43]. Some software of this type also provides a monitoring function for students and teachers that both helps teachers monitor student collaboration on a task and supports students in comparing their engagement with their peers [46, 60]. Some software in this category can also support teachers in forming collaborative groups based on student feedback on their collaborative learning experiences (Manathunga and Hernández-Leo [45]). Another essential function is improving student interaction by enhancing their awareness using a diverse range of data, including discussion forums, LMS behavior, body postures, and voice recordings [58, 61].

Self-regulated learning software supports diverse functions to help students regulate their learning. It frequently starts with goal setting and planning; this kind of software allows students to set learning goals, plan their activities (e.g., assignments), and meet deadlines [79]. SRL software also provides performance monitoring functions for both students and teachers. On the one hand, it allows students to visualize their performance in activities and compare it with their peers via the analysis of various statistics (e.g., reading time and the status of specific goals) (Kia et al. [72], [75]). On the other hand, it both supports teachers in monitoring students' progress and helps regulate students' needs [80]. Beyond those, the most common functions in SRL are self-evaluation and reflection functions that provide *intelligent feedback* (e.g., statistical, visual, and textual) about students' activities related to their learning goals that helps them reflect and make timely decisions [17, 77].

### 4.3 Research question 3: What analytics approaches are employed in learning analytics software, and what are their limitations?

In this review, we also sought to identify the analytical approaches that were applied to create the LAS functions and the limitations of these approaches.

**Learning analytics software for identifying at-risk students.** The prime feature of ARSI software is to predict student risk status; therefore, several approaches have been implemented for this purpose. For example, three studies [21, 26, 37] implemented different ML algorithms (e.g., random forest, logistic regression, C4.5, decision, and regression trees) on students' LMS data to classify them into different risk statuses (e.g., pass or fail), and they achieved 95% accuracy with a regression tree approach. On the other hand, Baneres et al. [93] performed gradual at-risk modelling, where sub-models were built for each assessment activity to determine the probability of passing or failing a course. The created models were evaluated on two different courses, and the best accuracy of 95.78% was achieved using kNN. Likewise, Cano and Leonard [31] implemented an incremental multi-view genetic algorithm to identify dropout students. However, these approaches cannot predict student risk status at the assignment or task level, and it can be helpful for a student to know the prediction of the following upcoming assignment. Therefore, four studies [22–24, 34–41] implemented ML and visualization approaches at the assignment level to predict success status. For instance, Hlosta et al. [23] implemented a splitting milestone technique consisting of three steps. First, the students' activities were divided into a time-sliced format; then, to identify the essential milestones, the algorithm continuously examined the differences between successful and unsuccessful students. Finally, the best splitting values were generated from the identified milestone. Herodotou et al. [24] employed an ensemble approach using the three ML algorithms of Naïve Bayes, kNN, and CART, and they developed four predictive models. Combining the results of these four models improved the overall accuracy compared to using a single model.

The formulation of interventions is an important feature and is considered essential for the success of at-risk students. Therefore, De Laet et al. [33] introduced three modules, where the first module visualized the entire learning path on a single screen using historical academic records; the second module simulated the workload; and the third module performed a comparison with peers. These modules help teachers formulate an effective intervention for students. On the other hand, Fu et al. [39] worked on programming log data to discover students' weaknesses and classify them as "outstanding," "active," or "struggling" in real-time to support teachers in intervention formation. Improving course content is also considered essential to the success of students. Therefore, Weng et al. [34] processed log data from the BookRoll tool attached to Moodle and assisted teachers by improving learning content through the provision of a statistical demonstration based on actions taken by students while reading e-books. Hwang et al. [19] conducted a cognitive analysis of students by implementing techniques using a "fuzzy" membership function and fuzzy rules, which enabled the provision of suitable learning material to each student. In the same context, Cavanagh et al. [35] proposed several strategies, including manual preparation of assessments, a learning path, and the provision of suitable alternative content. Moreover, pedagogical teaching practices were also proposed using domain-based, concept-based, and student-based learning analytics.

Learning analytics approaches also offer interactive visualizations to raise awareness among teachers and students about their learning and teaching activities. For instance, Holman et al. [36] developed several visualizations such as performance summaries, class standing, and badge analytics, to reflect on. Similarly, Govaerts et al. [25] tracked and visualized the changes in students' activities by dividing student tracking data into 14 parts, which could help students detect learning patterns. Conversely, Ruipérez-Valiente et al. [29] identified 21 indicators (e.g., video efficiency, exercise efficiency, level of emotions, etc.) for teachers to guide them concerning individual students and the overall class status. In a similar vein, Ali et al. [30, 42] presented statistical information about students' interactions with course content. Likewise, Ruiz et al. [38] captured students' emotions and provided reflection to both teacher and student, then measured the effects on student academic performance.

**Limitation:** Predictive models for identifying at-risk students have been developed for specific courses and fed only with LMS data and are likely to have poor generalisability beyond the original context [37]. Moreover, visualization is an effective method of forecasting risk level, yet the use of complex visualizations may not be easy to understand, giving rise to the risk of incorrect decision-making and severe repercussions on students' learning. Regarding intervention formation, the approaches analyzed in this paper could not monitor the impact of interventions on students' learning, making it difficult for teachers to adapt their interventions to meet students' needs. Moreover, no system was found to be universal for all types of students, as different students come from different academic and socioeconomic backgrounds, and the learning difficulties faced by each student are also difficult to identify. Recognition of the impact of awareness methods was missing, although the use of data from previous students could form the basis for intelligent recommendations that might help current students find successful learning trajectories.

**Approaches for computer-supported collaborative learning.** Computer-supported learning software is primarily focused on visualization approaches to identify isolated students, monitor their participation, and provide support and awareness about their interactions. Bakharia and Dawson [43] implemented an ego-network based on student interaction data to identify isolated students. This approach extracted student discussion data to generate a student learning network. For example, a network would be "loose" if a participant submitted a post to which no other participants responded. On the other hand, to identify students who seldom participated in collaborative discussions, Radu et al. [20] used Kinetic sensors and Hololens headsets to collect data on the participants' joint coordinates and their gaze information. They then analyzed the data using a K-means posture clustering technique based on an elbow method.

Regarding student participation monitoring, Martinez-Maldonado et al. [46] implemented a proportion visualization technique using multi-touch tabletops data to represent the active participation of each student in a given collaborative activity. In the same context, Han et al. [54] developed an adaptive argument support system in a face-to-face environment by applying Pearson's correlation analyzes to group discussion data. The desired student learning status was visualized by applying color-coding to collaborative scripts using a rule-based algorithm. Since adaptive argumentation may raise the issues of students' perceived choice and lack of motivation, Challco et al. [56] introduced an ontology-based gamification scheme for collaborative session scripts in which game design exercises encoded student knowledge. Different view panels on the dashboard were provided based on the student role, using role models for visualization. Using a different approach, Yoo and Jin [55] applied interaction analysis to determine student participation and employed a comparability analysis implementation to compare student engagement with peers.

Forming collaborative groups is another helpful feature; therefore, Manathunga and Hernández-Leo [45] developed behavioral rules (e.g., flow, control, awareness, and support) to construct active working groups. However, rules were not efficient in forming heterogeneous and homogeneous groups. Duque et al. [50] therefore applied the concept of data depth to student data in terms of both homogeneous analysis indicators (HOAIs), such as communication skills and work speed, and heterogeneous indicators (HEAIs), such as the quality of presentations and documentation and success in problem-solving, and they divided students into groups based on a minimum score (the difference between the HOAIs and HEAIs). Similarly, Alberola et al. [48] applied Bayesian learning to students' feedback on their peers to identify and use eight behavioral patterns for future group formation.

In terms of social interaction and awareness, several studies [44, 53, 55, 61] implemented text analysis techniques on interactions to create networks and representations for students that provide guidance and insights about their collaboration. For example, Ouyang et al. [61] implemented multi-method analytics (i.e., social network analysis, text mining technique, and content analysis) to produce three network representations, namely the social network, the topic network, and the cognitive network, so that students can monitor their own and peers' social interactions. Worsley et al. [53] enhanced the collaborative experience of students by first transcribing real-world audio and video discussion and then applying sentiment analysis to provide guidance. On the other hand, Wise et al. [49] implemented a hyperbolic tree structure to both encourage students to interact more with their peers and practice a visualization based on a night sky metaphor to promote student participation.

**Limitations:** Isolated student identification approaches did not consider *passive participation*, whereby participants read or browse messages but do not respond. Moreover, these approaches are one-dimensional in that they are exclusively based on the number of student interactions in discussion forums. Supporting group formation based on student feedback may overlook important factors, such as the interpersonal relationships between students and their past performance and skills. Furthermore, these approaches do not provide an early success status update and cannot predict whether or not a collaborative group will complete an assigned task within the allotted time. Lastly, visualization approaches were found to improve interaction with peers; however, they could become difficult to interpret, such as visualizations used in vast networks, such as massive courses.

**Approaches for self-regulated learning.** Self-regulated learning software starts with goal setting, planning, and time management, which are considered essential steps in regulating students' learning. Muslim et al. [83] achieved this by enabling students to define their goals, pose questions, and self-define their indicators based on personalized and goal-oriented LA. To help students track their learning activity goals, Gaftandzhieva et al. [73] developed a hierarchical system of measurable indicators based on LMS data that could help students meet their activity goals and improve course success. However, achievement levels and learning indicators do not provide meaningful information to help students reach their desired goals. Hence, Kia et al. [72] proposed a progress-towards-the-final-grade (goal) approach based on assignment dates for an entire course and their calculated influence on the student's final grade, which was able to help students spend time on only those assignments that had more impact on their final grade. On the other hand, Jivet et al. [79] employed a qualitative analysis approach to coding each goal into three binary variables, describing the presence of a learning component and a performance or time-frame component in the goal text. Concerning managing time, Sense et al. [91] proposed a learner-specific rate approach to the forgetting of scheduling items using each student's "forget rate," which inspires students to use their study time more effectively. Similarly, Tabuenca et al. [66] utilized mobile time logs from students' SRL activities to calculate their achievement levels. This approach uses the time spent variable and the scores obtained in SRL activities to calculate the probability of achieving a specific goal based on their relative frequency.

Performance comparison and monitoring are also considered essential components of SRL software. Therefore, three studies [18, 72, 88] implemented ML and visualization approaches that support students in comparing their performance with their peers. For example, Aljohani et al. [18] implemented ML methods on student tracking data to analyze students' online activities and presented peer comparisons using a graphical representation. Similarly, Charleer et al. [88] processed students' historical data (grades) to make learning paths that help students monitor and compare their performance with peers. Three studies [64, 89, 90] provided skill level comparison by employing a mastery progress grid and skill meter to help students compare their level of understanding of different course concepts and contents. On the other hand, Bull et al. [76] developed a *skills meter* that both helps teachers compare each student's understanding of the course concepts at the activity level and allows teachers to alter the weightings of activities in the model. Similarly, Dourado et al. [80] implemented a discovery approach that processes LMS event data by removing irrelevant events and extracting relevant events. After that, identical sequences were grouped to build a tree with the most common learning paths.

Self-evaluation and reflection are also considered essential features of SRL software, according to Whitelock et al. [62]. In four studies by [67, 72, 75, 81], who implemented visualization approaches to provide evaluation and reflection, the authors extracted features from textual assignments (e.g., common words, central ideas, and grammatical structures) and generated a tag cloud to provide an evaluation of the written material. Similarly, Knight et al. [85] implemented NLP techniques on the rhetorical utterances in students' texts to identify sentences that communicate a specific rhetorical function. The system includes a computational parser that processes each sentence in the text and identifies rhetorical moves. Likewise, Kia et al. [72] and Khosravi et al. [77] proposed a resource selection strategy that tabulates access to the course materials, categorizes the results based on each resource, and then generates a graphical representation that helps students determine ("recommend") which resource is vital to read or watch versus which has already been read or watched. With a focus on programming tasks, two studies [63, 65] applied code tracking and a hint generator based on a set of rules derived from teacher-specified annotated solutions. These generated rules were used to identify code errors and provide hints. On the other hand, Bodily et al. [87] developed a content recommender system that uses correct-answer probability to give feedback on improving mastery of each concept. This real-time feedback is presented at the unit level, which allows students to study more strategically as they can easily see their knowledge gaps. Sending feedback based on students' needs is also challenging; Tabuenca et al. [66] implemented a notification-based feedback approach that sent customized feedback notifications rather than static daily notifications based on the student's tracking data (TD). Conversely, Pardo et al. [86] implemented a personalized messaging system that provides feedback based on students' needs data to suggest actions.

**Limitations:** Goal-setting and planning approaches were found to be helpful, but they could not identify the most appropriate time for performing a specific task or the most effective time slot. Performance comparison approaches did not provide information that explains why some students perform better than others. Without such information, students are left without guidance on how to self-regulate. On the other hand, approaches for programming task evaluation involving a comparative analysis with expert solutions could have adverse outcomes, as students might develop unique solutions not in line with expert solutions. In this case, hints based on expert solutions may hinder a student's self-learning process. Finally, recommendation approaches signal to students which content or resource is essential to improving their academic performance based on the utilization of specific content within their class. This information might not be helpful until the impact of these resources on the student's performance has been calculated.

# 5 DISCUSSION

## 5.1 Findings

This study systematically reviewed the available forms of LAS for higher education. Starting with the first study question, we categorized the reviewed LAS based on purpose into the three categories of ARSI, CSCL, and SRL. Furthermore, we revealed that most of the LAS were standalone versions, meaning they could operate independently, and the remainder were plugins for LMSs or toolkits. The analyzed LAS were employed by students and teachers in diverse higher education courses. However, a few software packages were also assessed by experts to understand the design and practical issues of LAS before deploying them. Moreover, a year-wise analysis showed that interest in developing LAS has increased recently because of the adoption of digital tools, the availability of vast amounts of student data, and the opportunity to understand and optimize the learning experience.

For our second study question, we aimed to identify the embedded functions in the existing LAS. The results demonstrate that LAS delivers several essential functions to both teachers and students. As shown in Figure 6, teachers were offered three sets of functions related to detection, examination, and formation. The first set of detection functions allows teachers to identify at-risk and isolated students and discover the areas in which they are struggling. The second set of functions was linked to examining student behavior, active participation, and risk level, while the last function was related to intervention and feedback support for teachers. The functions offered to students can be categorized into four types. The first involves reflection and evaluation, whereby students can acquire understanding about why they are failing a course, and they can obtain automated assessment and feedback on their tasks and performance. The second type of function provides monitoring based on active participation in collaborative tasks, and it delivers comparisons and analyses of success rates compared to peers. Several interactive functions were added to the existing social network framework to improve social interaction, with a view to understanding and identifying highly relevant discussion topics and rating peers' options (questions and answers) to gain more information. Last, to motivate and remind students about their goals, which were determined before they started the course, the assessed LAS provided functions that allow students to self-regulate based on these goals. According to the LAS evaluations presented in the reviewed articles, the embedded functions were found to provide significant support in terms of helping both students and teachers reach their goals.

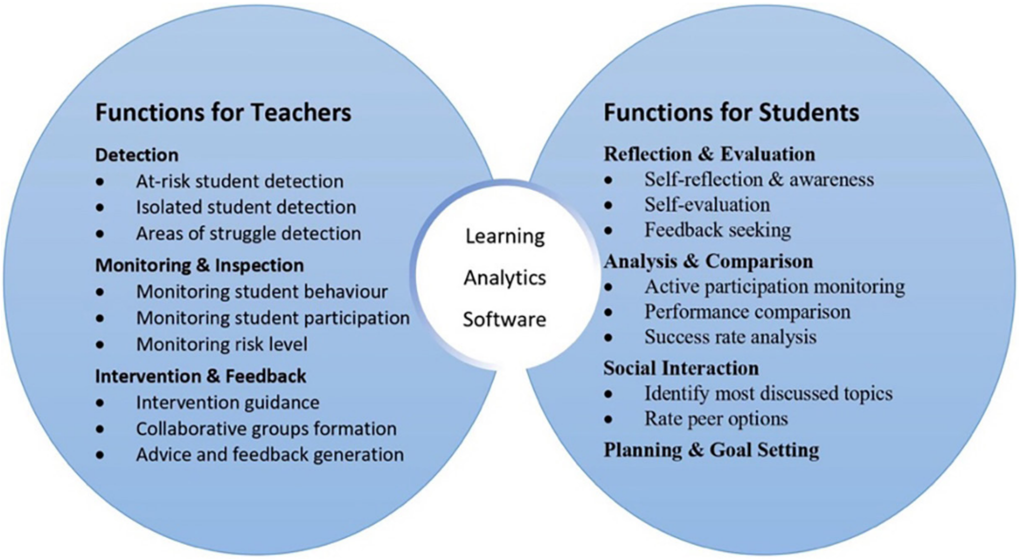


**Fig. 6.** Overview of functions offered by learning analytics software

To answer our third study question, we provided an overview of the approaches implemented by LAS and their corresponding limitations. As seen in Figure 7, an overall 42% of the LAS-implemented visualization approaches mainly showed students analytics in a way that could be beneficial. Data mining approaches were implemented by 25% of the assessed LAS to identify student behavior patterns. In contrast, 18% of the assessed LAS employed ML approaches for prediction or detection tasks. NLP and statistical analysis were the least implemented approaches; the use of NLP approaches was to process and understand textual data, while statistical analysis was used to find correlations between students or activities. ARSI software implemented ML algorithms (e.g., Naïve Bayes, kNN, regression, decision tree, CART, logistic regression, and random forest) to identify at-risk students and project risk levels. Data mining approaches improved the identification process by identifying the weaknesses in students' behavior, which helped teachers formulate student interventions. Additionally, visualiszation approaches provide awareness among students about their class standing, progress, and skill level. Although these analytical approaches provided satisfactory results, the transferability of ML approaches to other courses was questioned [37]. Furthermore, approaches were mainly fed with LMS data, overlooking learning that occurs in a blended manner.

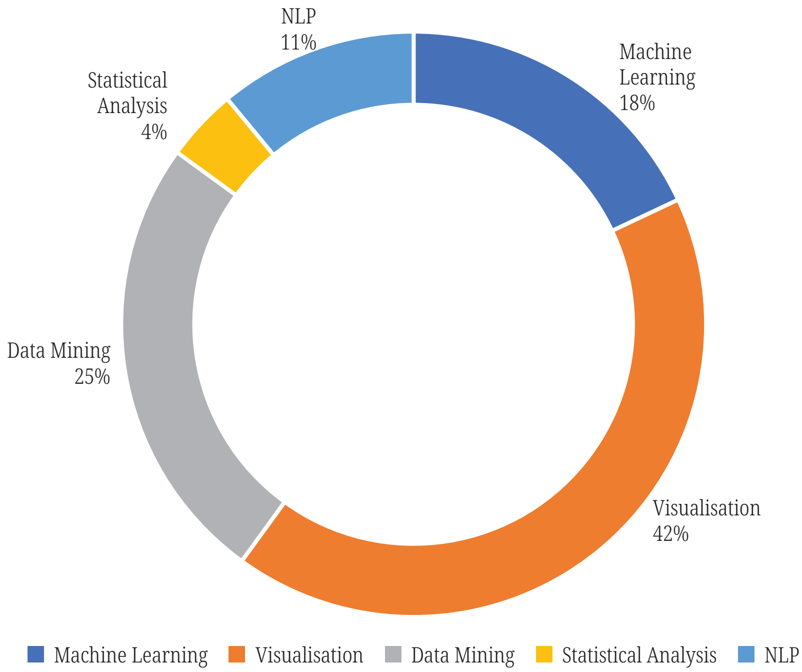


**Fig. 7.** Overview of approaches offered by learning analytics software

Computer-supported collaborative learning software predominantly employed data mining approaches such as Bayesian learning, algorithmic rules, and data depth analysis to create optimized learning groups and maintain their working flow and social interactions. However, these approaches did not include passive participation, which occurs within forums when participants read or browse messages but do not respond [94]. Conversely, text analysis approaches were employed to develop social, topical, and cognitive networks to enhance students' understanding of their and their peers' social interactions. However, networks were one-dimensional and based only on the number of interactions (quantity) and therefore could not capture the complex nature of collaborative learning [95]. SRL systems focus on visualization approaches to provide feedback and compare students' learning skills. In contrast, data mining and ML techniques helped to make these graphical representations more informative. For example, the *tag cloud* was one of the visualizations to assess textual assignments generated by performing syntactic and semantic analysis on textual data. On the other hand, graph-oriented visualizations provided comparison among students and feedback about their daily activities, which were produced using ML algorithms and statistical analysis of students' data. Visual notification was another approach based on analytical information (e.g., time-devoted or achievable grades) sent through emails to provide self-reflection and goal achievement status. However, the interpretation of system-generated feedback is a drawback of visualization approaches due to their complex structuring, whereas lacking recommendations based on behavior analysis was the crucial constraint of the proposed approaches.

### 5.2 Implications

We consider the findings of this study interesting from three perspectives. From a student's perspective, at-risk status identification, monitoring, reflection, and awareness functions help students better understand the learning process and improve their learning outcomes [78]. As a result, students can complete their courses, change their learning behavior, achieve their learning goals, and effectively collaborate on tasks. Another practical implication of LAS is that recommendations and timely feedback can motivate students to perform more actions (e.g., complete their learning tasks), leading to improved assessment performance [96]. As students start regulating themselves, this indirectly supports teachers, who are responsible (especially in massive and distance courses) for continually intervening and guiding students through course activities, for instance, by emphasizing the importance of a task, the time needed to complete it, and the required score to pass. Since LAS handles these responsibilities and gives students information regularly, the teacher's job becomes more manageable. However, LAS can impact students negatively because LAS are overloaded with information, and many students are using these kinds of LAS for the first time [60]. Overall, students appeared to feel that most LAS were helpful. Still, a few functions (e.g., the interaction dashboard) may have been unfamiliar or relatively less dependable, and students may require skills to interpret that information [55]. Therefore, LAS that are built according to students' capacities can be more effective.

From the teacher's perspective, LAS can be practically helpful in informing teachers about at-risk students, disengaged students in collaborative tasks, and students who require special attention or help to proceed with their learning, especially in distance learning scenarios. Such insights could lead to proactive teaching practice, where teachers approach students flagged as at-risk, discuss their progress and performance, identify possible learning difficulties, and provide real-time, tailored support to accommodate a student's learning needs [24]. From a cultural perspective, the effect of learning analytics is not determined solely by the features of the technology itself; more importantly, *how* the technologies are used is critical due to the complex interplay between stakeholders, the learning contexts, and the affordances of the new LAS. Since the LAS feeds the information back to the students themselves, the visualizations must be tailored to the students' learning needs, goals, and levels of experience [61]. Therefore, merely introducing the LAS to students is not enough; engaging students in the design and development processes is necessary to understand their learning needs, contexts, and goals. In this way, the software design can better align with students' learning practices and foster better learning. Overall, a human-centered, participatory design is one possible means to: build collaboration between multiple stakeholders; improve the design and implementation of LAS; and evolve subsequent practices of online teaching and learning.

### 5.3 Identified gaps and future research recommendations

Despite the contributions made by LAS, several study gaps and areas for future investigation were identified. These gaps provide opportunities for further exploration and advancement in the field. The following sections outline the key study gaps and suggest potential future study directions.

**Customizable learning analytics dashboards.** Learning analytics software provides various dashboards that allow students to monitor their interactions, academic performance, and learning behavior throughout their courses. However, the analytics on the dashboards are general for all students and specified by teachers, academics, and institutions [79]. There is a need for a student-centered approach to empower students to make their own decisions about what analytics they wish to see on their dashboards. A future study might address this gap by considering customizable dashboards on which students can choose the analytics they want to monitor towards achieving their desired academic performance. Customisable dashboards could also be helpful in observing students' decision-making concerning what part of the analytics is essential for them and what analytics they want to monitor on a learning platform. Moreover, how information selection decisions relate to their learning performance during courses might also be measured.

**Explainability of prediction and detection.** The prediction of at-risk students is a common feature of LAS that is used in feedback and intervention generation. Such feedback is insightful to some extent; however, users are not provided with the reasons behind such a status or prediction and what appropriate steps are needed to avoid possible failure. Future studied might address this gap by using explainable ML that can understand data and explain the predictions made by ML models [96]. Explainable ML would clarify why a student might fail the course and how this could be avoided (e.g., suggest remedial actions); furthermore, it could also support teachers in generating more insightful interventions for students at a personalized level.

**Analytics interpretability and informed actions.** Students appeared to agree that the provided analytics and visualizations are challenging to interpret; as a result, it becomes difficult for students to take suitable action [60]. Designers should not presume that offering more information will automatically lead to better effects. Because providing information automatically increases students' cognitive overload, corresponding action support is needed. Therefore, improvements in analytics interpretability are required, mainly for students with inadequate data literacy; it is difficult for them to quickly understand the analytical results, identify valuable information from the results, and take informed action without external support [84]. Furthermore, in the CSCL context, future iterations might also remind students to take concrete steps, such as interacting with their less-engaged peers, noting others' contrasting perspectives, or elaborating their ideas in certain directions [61].

**Long-term implementation and scalability.** While many studies have examined the use of LAS in controlled settings, there is a gap in the study addressing its long-term implementation and scalability in real-world educational contexts. Future studied should investigate the challenges and strategies for the successful integration and sustained use of LAS across diverse educational institutions. Exploring factors that influence the adoption of learning analytics initiatives, institutional support, and scalability can provide valuable insights for practitioners and policymakers.

**Individual differences and personalization.** Learning analytics software has the potential to support personalized learning experiences by tailoring instruction to individual students' needs. However, there is a study gap in understanding how to effectively leverage learning analytics to accommodate and address individual differences, such as learning styles, prior knowledge, and cognitive abilities [3]. A future study might explore how LAS can adapt to different students and provide personalized recommendations and interventions. Investigating the optimal balance between automation and human involvement in the personalization process is also essential.

**Data privacy and ethics.** As LAS collects and analyze large amounts of student data, there are concerns about data privacy and ethical implications [97]. Future studied should focus on developing privacy-preserving techniques and ensuring the responsible use of student data. Additionally, investigating the perceptions and concerns of students, teachers, and parents regarding data privacy could help shape ethical guidelines for the implementation of learning analytics software.

### 5.4 Limitations and future directions of this systematic review

The limitations of this study lie in the search method used and the analysis of LAS. In terms of the search method, it might be possible that our keywords and the corresponding set of queries were insufficient to search the entire literature. Furthermore, our search of only a select handful of databases may mean that some papers on LAS were overlooked, and the exclusion of books, dissertations, and non-English and short papers in our literature selection strategy may have led to the omission of important information on LAS. Also, the thematization of LAS introduced a limitation of vague boundaries among the three categories of ARSI, CSCL, and SRL software. In our analysis of software functions, we did not explore the scalability, computational speed, and resource requirements for each form of LAS. Furthermore, the structure and size of the datasets analyzed by the LAS were not investigated. In the future, we plan to study existing forms of LAS from the perspectives of generalisability, scalability, and resource efficiency in real-time environments. To improve our search method, we will both extend the set of search queries by adding more relevant keywords and widen the search to conduct a more comprehensive, systematic review. Grey literature, such as books and dissertations, will also be incorporated into future studies. Moreover, the analysis of the LAS is not deep or comprehensive enough, and more in-depth analysis could be possible through concrete studies (e.g, of tools to help with self-regulation).

## 6 CONCLUSION

The use of LA to build LAS has provided an opportunity to deliver insightful recommendations to students and teachers based on large-scale, heterogeneous data drawn from learning environments. Several types of commercial and open-source LAS have been developed with diverse objectives, such as identifying at-risk students, enhancing collaborative learning, advancing SRL, and monitoring students' behaviors. In this paper, we present a systematic review of 46 types of LAS that were implemented and evaluated in the real-world setting of higher education. The goals of this review were to explore the implementation details of embedded functions of LAS and to highlight the limitations that might be overcome in future study. This study has achieved the following: (i) we presented the main functions of the software in terms of supporting students and teachers to improve their learning performance and teaching practice; (ii) we highlighted the LA techniques implemented in the software functions; and (iii) we identified limitations and opportunities for future study and the creation of software that could help advance these fields. However, it should be noted that many of the more well-known types of LAS have not been described in academic studies and hence were not included in this study. Our focus was on reviewing a particular range of study-driven software in order to obtain a deeper knowledge of the embedded functions and implemented LA approaches.

## 7 ACKNOWLEDGEMENT

I extend my heartfelt gratitude to Aayesha Aayesha, Panagiotis Papapetrou, and Uno Fors for their invaluable support and guidance throughout the development of this study article. Their insightful feedback and suggestions were instrumental in improving the structure and quality of this work. Thank you for your unwavering assistance and encouragement.

## 8 REFERENCES

[1] D. Ifenthaler and J. Y.-K. Yau, "Utilising learning analytics to support study success in higher education: A systematic review," *Educational Technology Research and Development*, vol. 68, pp. 1961–1990, 2020. https://doi.org/10.1007/s11423-020-09788-z

[2] J. Knobbout and E. Van Der Stappen, "Where is the learning in learning analytics? A systematic literature review on the operationalization of learning-related constructs in the evaluation of learning analytics interventions," *IEEE Transactions on Learning Technologies*, vol. 13, no. 3, pp. 631–645, 2020. https://doi.org/10.1109/TLT.2020.2999970

[3] A. Larrabee Sønderlund, E. Hughes, and J. Smith, "The efficacy of learning analytics interventions in higher education: A systematic review," *British Journal of Educational Technology (BERA)*, vol. 50, no. 5, pp. 2594–2618, 2019. https://doi.org/10.1111/bjet.12720

[4] K. Mangaroska and M. Giannakos, "Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning," *IEEE Transactions on Learning Technologies*, vol. 12, no. 4, pp. 516–534, 2018. https://doi.org/10.1109/TLT.2018.2868673

[5] B. A. Schwendimann *et al.*, "Perceiving learning at a glance: A systematic literature review of learning dashboard research," *IEEE Transactions on Learning Technologies*, vol. 10, no. 1, pp. 30–41, 2016. https://doi.org/10.1109/TLT.2016.2599522

[6] O. Viberg, M. Hatakka, O. Bälter, and A. Mavroudi, "The current landscape of learning analytics in higher education," *Computers in Human Behavior*, vol. 89, pp. 98–110, 2018. https://doi.org/10.1016/j.chb.2018.07.027

[7] P. Leitner, M. Khalil, and M. Ebner, "Learning analytics in higher education—A literature review," in *Learning Analytics: Fundaments, Applications, and Trends: A View of the Current State of the Art to Enhance E-learning*, A. Peña-Ayala, Eds., Springer, Cham., vol. 94, pp. 1–23, 2017. https://doi.org/10.1007/978-3-319-52977-6_1

[8] R. Ferguson, "Learning analytics: Drivers, developments and challenges," *International Journal of Technology Enhanced Learning*, vol. 4, nos. 5–6, pp. 304–317, 2013. https://doi.org/10.1504/IJTEL.2012.051816

[9] W. Matcha, D. Gašević, and A. Pardo, "A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective," *IEEE Transactions on Learning Technologies*, vol. 13, no. 2, pp. 226–245, 2020. https://doi.org/10.1109/TLT.2019.2916802

[10] Y. Tabaa and A. Medouri, "LASyM: A learning analytics system for MOOCs," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 4, no. 5, 2013. https://doi.org/10.14569/IJACSA.2013.040516

[11] R. Nusche, "Student assessment: Putting the learner at the centre," in *Synergies for Better Learning: An International Perspective on Evaluation*, OECD Publishing, Paris, 2013, pp. 139–269. https://doi.org/10.1787/9789264190658-7-en

[12] R. Pérez-Álvarez, J. Maldonado-Mahauad, and M. Pérez-Sanagustín, "Tools to support self-regulated learning in online environments: Literature review," in *Lifelong Technology-Enhanced Learning, EC-TEL 2018*, in Lecture Notes in Computer Science, V. Pammer-Schindler, M. Pérez-Sanagustín, H. Drachsler, R. Elferink, and M. Scheffel, Eds., Springer, Cham., vol. 11082, 2018, pp. 16–30. https://doi.org/10.1007/978-3-319-98572-5_2

[13] C. Romero and S. Ventura, "Educational data mining and learning analytics: An updated survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 10, no. 3, p. e1355, 2020. https://doi.org/10.1002/widm.1355

[14] N. Bergdahl, J. Nouri, T. Karunaratne, M. Afzaal, and M. Saqr, "Learning analytics for blended learning: A systematic review of theory, methodology, and ethical considerations," *International Journal of Learning Analytics and Artificial Intelligence for Education (iJAI)*, vol. 2, no. 2, pp. 46–79, 2020. https://doi.org/10.3991/ijai.v2i2.17887

[15] B. Kitchenham and P. Brereton, "A systematic review of systematic review process research in software engineering," *Information and Software Technology*, vol. 55, no. 12, pp. 2049–2075, 2013. https://doi.org/10.1016/j.infsof.2013.07.010

[16] N. J. M. Blackman and J. J. Koval, "Interval estimation for Cohen's kappa as a measure of agreement," *Statistics in Medicine*, vol. 19, no. 5, pp. 723–741, 2000. https://doi.org/10.1002/(SICI)1097-0258(20000315)19:5<723::AID-SIM379>3.0.CO;2-A

[17] J. L. Santos, S. Govaerts, K. Verbert, and E. Duval, "Goal-oriented visualizations of activity tracking: A case study with engineering students," in *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 2012, pp. 143–152. https://doi.org/10.1145/2330601.2330639

[18] N. R. Aljohani, A. Daud, R. A. Abbasi, J. S. Alowibdi, M. Basheri, and M. A. Aslam, "An integrated framework for course adapted student learning analytics dashboard," *Computers in Human Behavior*, vol. 92, pp. 679–690, 2019. https://doi.org/10.1016/j.chb.2018.03.035

[19] G.-J. Hwang, H.-Y. Sung, S.-C. Chang, and X.-C. Huang, "A fuzzy expert system-based adaptive learning approach to improving students' learning performances by considering affective and cognitive factors," *Computers and Education: Artificial Intelligence*, vol. 1, p. 100003, 2020. https://doi.org/10.1016/j.caeai.2020.100003

[20] I. Radu, E. Tu, and B. Schneider, "Relationships between body postures and collaborative learning states in an augmented reality study," in *Artificial Intelligence in Education. AIED 2020*, in Lecture Notes in Computer Science, I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, and E. Millán, Eds., Springer, Cham., vol. 12164, 2020, pp. 257–262. https://doi.org/10.1007/978-3-030-52240-7_47

[21] K. E. Arnold and M. D. Pistilli, "Course signals at Purdue: Using learning analytics to increase student success," in *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 2012, pp. 267–270. https://doi.org/10.1145/2330601.2330666

[22] A. Essa and H. Ayad, "Student success system: Risk analytics and data visualization using ensembles of predictive models," in *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 2012, pp. 158–161. https://doi.org/10.1145/2330601.2330641

[23] M. Hlosta, J. Kocvara, D. Beran, and Z. Zdrahal, "Visualisation of key splitting milestones to support interventions," 2019.

[24] C. Herodotou, B. Rienties, A. Boroowa, Z. Zdrahal, and M. Hlosta, "A large-scale implementation of predictive learning analytics in higher education: The teachers' role and perspective," *Educational Technology Research and Development*, vol. 67, pp. 1273–1306, 2019. https://doi.org/10.1007/s11423-019-09685-0

[25] S. Govaerts, K. Verbert, E. Duval, and A. Pardo, "The student activity meter for awareness and self-reflection," in *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, 2012, pp. 869–884. https://doi.org/10.1145/2212776.2212860

[26] Y.-H. Hu, C.-L. Lo, and S.-P. Shih, "Developing early warning systems to predict students' online learning performance," *Computers in Human Behavior*, vol. 36, pp. 469–478, 2014. https://doi.org/10.1016/j.chb.2014.04.002

[27] A. E. Krumm, R. J. Waddington, S. D. Teasley, and S. Lonn, "A learning management system-based early warning system for academic advising in undergraduate engineering," in *Learning Analytics*, J. Larusson and B. White, Eds., Springer, New York, NY, 2014, pp. 103–119. https://doi.org/10.1007/978-1-4614-3305-7_6

[28] H. Goker and H. I. Bulbul, "Improving an early warning system to prediction of student examination achievement," in *2014 13th International Conference on Machine Learning and Applications*, 2014, pp. 568–573. https://doi.org/10.1109/ICMLA.2014.114

[29] J. A. Ruipérez-Valiente, P. J. Muñoz-Merino, D. Leony, and C. D. Kloos, "ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform," *Computers in Human Behavior*, vol. 47, pp. 139–148, 2015. https://doi.org/10.1016/j.chb.2014.07.002

[30] L. Ali, M. Hatala, D. Gašević, and J. Jovanović, "A qualitative evaluation of evolution of a learning analytics tool," *Computers and Education*, vol. 58, no. 1, pp. 470–489, 2012. https://doi.org/10.1016/j.compedu.2011.08.030

[31] A. Cano and J. D. Leonard, "Interpretable multiview early warning system adapted to underrepresented student populations," *IEEE Transactions on Learning Technologies*, vol. 12, no. 2, pp. 198–211, 2019. https://doi.org/10.1109/TLT.2019.2911079

[32] D. Bañeres, M. E. Rodríguez, A. E. Guerrero-Roldán, and A. Karadeniz, "An early warning system to detect at-risk students in online higher education," *Applied Sciences*, vol. 10, no. 13, p. 4427, 2020. https://doi.org/10.3390/app10134427

[33] T. De Laet, M. Millecamp, M. Ortiz-Rojas, A. Jimenez, R. Maya, and K. Verbert, "Adoption and impact of a learning analytics dashboard supporting the advisor—Student dialogue in a higher education institute in Latin America," *British Journal of Educational Technology (BERT)*, vol. 51, no. 4, pp. 1002–1018, 2020. https://doi.org/10.1111/bjet.12962

[34] J.-X. Weng, A. Y. Huang, O. H. Lu, I. Y. Chen, and S. J. Yang, "The implementation of precision education for learning analytics," in *2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, 2020, pp. 327–332. https://doi.org/10.1109/TALE48869.2020.9368432

[35] T. Cavanagh, B. Chen, R. A. M. Lahcen, and J. R. Paradiso, "Constructing a design framework and pedagogical approach for adaptive learning in higher education: A practitioner's perspective," *International Review of Research in Open and Distributed Learning*, vol. 21, no. 1, pp. 173–197, 2020. https://doi.org/10.19173/irrodl.v21i1.4557

[36] C. Holman, S. Aguilar, and B. Fishman, "GradeCraft: What can we learn from a game-inspired learning management system?" in *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, 2013, pp. 260–264. https://doi.org/10.1145/2460296.2460350

[37] E. de Quincey, C. Briggs, T. Kyriacou, and R. Waller, "Student centred design of a learning analytics system," in *Proceedings of the 9th International Conference on Learning Analytics and Knowledge*, 2019, pp. 353–362. https://doi.org/10.1145/3303772.3303793

[38] S. Ruiz, S. Charleer, M. Urretavizcaya, J. Klerkx, I. Fernández-Castro, and E. Duval, "Supporting learning by considering emotions: Tracking and visualization a case study," in *Proceedings of the Sixth International Conference on Learning Analytics and Knowledge*, 2016, pp. 254–263. https://doi.org/10.1145/2883851.2883888

[39] X. Fu, A. Shimada, H. Ogata, Y. Taniguchi, and D. Suehiro, "Real-time learning analytics for C programming language courses," in *Proceedings of the Seventh International Learning Analytics and Knowledge Conference*, 2017, pp. 280–288. https://doi.org/10.1145/3027385.3027407

[40] M. Ez-Zaouia and E. Lavoué, "EMODA: A tutor oriented multimodal and contextual emotional dashboard," in *Proceedings of the Seventh International Learning Analytics and Knowledge Conference*, 2017, pp. 429–438. https://doi.org/10.1145/3027385.3027434

[41] N. Diana, M. Eagle, J. Stamper, S. Grover, M. Bienkowski, and S. Basu, "An instructor dashboard for real-time analytics in interactive programming assignments," in *Proceedings of the Seventh International Learning Analytics and Knowledge Conference*, 2017, pp. 272–279. https://doi.org/10.1145/3027385.3027441

[42] A. F. Wise and Y. Jung, "Teaching with analytics: Towards a situated model of instructional decision-making," *Journal of Learning Analytics*, vol. 6, no. 2, pp. 53–69, 2019. https://doi.org/10.18608/jla.2019.62.4

[43] A. Bakharia and S. Dawson, "SNAPP: A bird's-eye view of temporal participant interaction," in *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, 2011, pp. 168–173. https://doi.org/10.1145/2090116.2090144

[44] B. Chen, Y.-H. Chang, F. Ouyang, and W. Zhou, "Fostering student engagement in online discussion through social learning analytics," *The Internet and Higher Education*, vol. 37, pp. 21–30, 2018. https://doi.org/10.1016/j.iheduc.2017.12.002

[45] K. Manathunga and D. Hernández-Leo, "PyramidApp: Scalable method enabling collaboration in the classroom," in *European Conference on Technology Enhanced Learning*, 2016, pp. 422–427. https://doi.org/10.1007/978-3-319-45153-4_37

[46] R. Martinez-Maldonado, A. Clayphan, and J. Kay, "Deploying and visualising Teacher's scripts of small group activities in a multi-surface classroom ecology: A study in-the-wild," *Computer Supported Cooperative Work (CSCW)*, vol. 24, pp. 177–221, 2015. https://doi.org/10.1007/s10606-015-9217-6

[47] M.-J. Laakso, E. Kaila, and T. Rajala, "ViLLE–collaborative education tool: Designing and utilizing an exercise-based learning environment," *Education and Information Technologies*, vol. 23, pp. 1655–1676, 2018. https://doi.org/10.1007/s10639-017-9659-1

[48] J. M. Alberola, E. Del Val, V. Sanchez-Anguix, A. Palomares, and M. D. Teruel, "An artificial intelligence tool for heterogeneous team formation in the classroom," *Knowledge-Based Systems*, vol. 101, pp. 1–14, 2016. https://doi.org/10.1016/j.knosys.2016.02.010

[49] A. F. Wise, Y. Zhao, and S. N. Hausknecht, "Learning analytics for online discussions: Embedded and extracted approaches," *Journal of Learning Analytics*, vol. 1, no. 2, pp. 48–71, 2014. https://doi.org/10.18608/jla.2014.12.4

[50] R. Duque, D. Gómez-Pérez, A. Nieto-Reyes, and C. Bravo, "Analyzing collaboration and interaction in learning environments to form learner groups," *Computers in Human Behavior*, vol. 47, pp. 42–49, 2015. https://doi.org/10.1016/j.chb.2014.07.012

[51] A. van Leeuwen, "Learning analytics to support teachers during synchronous CSCL: Balancing between overview and overload," *Journal of Learning Analytics*, vol. 2, no. 2, pp. 138–162, 2015. https://doi.org/10.18608/jla.2015.22.11

[52] D. A. Gómez-Aguilar, Á. Hernández-García, F. J. García-Peñalvo, and R. Therón, "Tap into visual analysis of customization of grouping of activities in eLearning," *Computers in Human Behavior*, vol. 47, pp. 60–67, 2015. https://doi.org/10.1016/j.chb.2014.11.001

[53] M. Worsley, K. Anderson, N. Melo, and J. Jang, "Designing analytics for collaboration literacy and student empowerment," *Journal of Learning Analytics*, vol. 8, no. 1, pp. 30–48, 2021. https://doi.org/10.18608/jla.2021.7242

[54] J. Han, K. H. Kim, W. Rhee, and Y. H. Cho, "Learning analytics dashboards for adaptive support in face-to-face collaborative argumentation," *Computers and Education*, vol. 163, 2021. https://doi.org/10.1016/j.compedu.2020.104041

[55] M. Yoo and S.-H. Jin, "Development and evaluation of learning analytics dashboards to support online discussion activities," *Educational Technology and Society*, vol. 23, no. 2, pp. 1–18, 2020.

[56] G. C. Challco, I. I. Bittencourt, and S. Isotani, "Can ontologies support the gamification of scripted collaborative learning sessions?" in *International Conference on Artificial Intelligence in Education*, 2020, pp. 79–91. https://doi.org/10.1007/978-3-030-52237-7_7

[57] J. McDonald *et al.*, "Cross-institutional collaboration to support student engagement: SRES version 2," in *33rd ASCILITE Conference, Adelaide*, 2016. https://doi.org/10.14742/apubs.2016.808

[58] T. Dragon *et al.*, "Metafora: A web-based platform for learning to learn together in science and mathematics," *IEEE Transactions on Learning Technologies*, vol. 6, no. 3, pp. 197–207, 2013. https://doi.org/10.1109/TLT.2013.4

[59] V. McCall, G. Mooney, P. Rutherford, and A. Gilmour, "The potential of online academic communities for teaching staff: Findings from a pilot study of the SocialLearn platform," *International Journal of Web Based Communities*, vol. 10, no. 4, pp. 404–425, 2014. https://doi.org/10.1504/IJWBC.2014.065392

[60] L. V. Guillain and B. Schneider, "Facilitators first: Building a tool with facilitators to foster a more collaborative makerspace community through movement traces," in *LAK21: 11th International Learning Analytics and Knowledge Conference,* 2021, pp. 533–539. https://doi.org/10.1145/3448139.3448194

[61] F. Ouyang, X. Li, P. Jiao, X. Peng, and W. Chen, "The three-iterative design and implementation of a student-facing social learning analytics tool," *International Journal of Distance Education Technologies (IJDET)*, vol. 19, no. 4, pp. 1–22, 2021. https://doi.org/10.4018/IJDET.286738

[62] D. Whitelock, D. Field, S. Pulman, J. T. Richardson, and N. Van Labeke, "OpenEssayist: An automated feedback system that supports university students as they write summative essays," 2013.

[63] A. Gerdes, B. Heeren, J. Jeuring, and L. T. van Binsbergen, "Ask-Elle: An adaptable programming tutor for haskell giving automated feedback," *International Journal of Artificial Intelligence in Education*, vol. 27, pp. 65–100, 2017. https://doi.org/10.1007/s40593-015-0080-x

[64] J. Guerra, R. Hosseini, S. Somyurek, and P. Brusilovsky, "An intelligent interface for learning content: Combining an open learner model and social comparison to support self-regulated learning and engagement," in *Proceedings of the 21st International Conference on Intelligent User Interfaces*, 2016, pp. 152–163. https://doi.org/10.1145/2856767.2856784

[65] P. Antonucci, C. Estler, D. Nikolić, M. Piccioni, and B. Meyer, "An incremental hint system for automated programming assignments," in *Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education*, 2015, pp. 320–325. https://doi.org/10.1145/2729094.2742607

[66] B. Tabuenca, M. Kalz, H. Drachsler, and M. Specht, "Time will tell: The role of mobile learning analytics in self-regulated learning," *Computers and Education*, vol. 89, pp. 53–74, 2015. https://doi.org/10.1016/j.compedu.2015.08.004

[67] K. McKenna, B. Pouska, M. C. Moraes, and J. E. Folkestad, "Visual-form learning analytics: A tool for critical reflection and feedback," *Contemporary Educational Technology*, vol. 10, no. 3, pp. 214–228, 2019. https://doi.org/10.30935/cet.589989

[68] S. Mohammed and A. M. Abdelazziz, "WIDE an interactive web integrated development environment to practice C programming in distance education," in *2013 1st International Conference of the Portuguese Society for Engineering Education (CISPEE)*, 2013, pp. 1–6. https://doi.org/10.1109/CISPEE.2013.6701964

[69] B. Tabuenca, M. Kalz, D. Börner, S. Ternier, and M. Specht, "Where is my time? Identifying productive time of lifelong learners for effective feedback services," in *International Computer Assisted Assessment Conference*, vol. 439, 2014, pp. 149–161. https://doi.org/10.1007/978-3-319-08657-6_15

[70] A. Nussbaumer, E.-C. Hillemann, C. Gütl, and D. Albert, "A competence-based service for supporting self-regulated learning in virtual environments," *Journal of Learning Analytics*, vol. 2, no. 1, pp. 101–133, 2015. https://doi.org/10.18608/jla.2015.21.6

[71] B. Tabuenca, M. Kalz, and M. Specht, "Lifelong learning hub: A seamless tracking tool for mobile learning," in *European Conference on Technology Enhanced Learning*, vol. 8719, 2014, pp. 534–537. https://doi.org/10.1007/978-3-319-11200-8_59

[72] F. S. Kia, S. D. Teasley, M. Hatala, S. A. Karabenick, and M. Kay, "How patterns of students dashboard use are related to their achievement and self-regulatory engagement," in *Proceedings of the Tenth International Conference on Learning Analytics and Knowledge*, 2020, pp. 340–349. https://doi.org/10.1145/3375462.3375472

[73] S. Gaftandzhieva, R. Doneva, S. Petrov, and G. Totkov, "Mobile learning analytics application: Using Students' big data to improve student success," *International Journal on Information Technologies and Security*, vol. 10, no. 3, pp. 53–64, 2018.

[74] R. Pérez-Álvarez, J. J. Maldonado-Mahauad, D. Sapunar-Opazo, and M. Pérez-Sanagustín, "NoteMyProgress: A Tool to Support Learners' Self-Regulated Learning Strategies in MOOC Environments," in *European Conference on Technology Enhanced Learning*, 2017, pp. 460–466. https://doi.org/10.1007/978-3-319-66610-5_43

[75] R. Pérez-Álvarez, J. Maldonado-Mahauad, and M. Pérez-Sanagustín, "Design of a tool to support self-regulated learning strategies in MOOCs," *J. Univers. Comput. Sci.*, vol. 24, no. 8, pp. 1090–1109, 2018.

[76] S. Bull, M. D. Johnson, M. Alotaibi, W. Byrne, and G. Cierniak, "Visualising multiple data sources in an lndependent open learner model," in *International Conference on Artificial Intelligence in Education*, 2013, pp. 199–208. https://doi.org/10.1007/978-3-642-39112-5_21

[77] H. Khosravi, K. Kitto, and J. J. Williams, "RiPPLE: A crowdsourced adaptive platform for recommendation of learning activities," *Journal of Learning Analytics*, vol. 6, no. 3, pp. 91–105, 2019. https://doi.org/10.18608/jla.2019.63.12

[78] J. Van Der Graaf *et al.*, "Do instrumentation tools capture self-regulated learning?" in *LAK21: 11th International Learning Analytics and Knowledge Conference*, 2021, pp. 438–448. https://doi.org/10.1145/3448139.3448181

[79] I. Jivet, J. Wong, M. Scheffel, M. Valle Torre, M. Specht, and H. Drachsler, "Quantum of choice: How learners' feedback monitoring decisions, goals and self-regulated learning skills are related," in *LAK21: 11th International Learning Analytics and Knowledge Conference*, 2021, pp. 416–427. https://doi.org/10.1145/3448139.3448179

[80] R. A. Dourado, R. L. Rodrigues, N. Ferreira, R. F. Mello, A. S. Gomes, and K. Verbert, "A teacher-facing learning analytics dashboard for process-oriented feedback in online learning," in *LAK21: 11th International Learning Analytics and Knowledge Conference*, 2021, pp. 482–489. https://doi.org/10.1145/3448139.3448187

[81] D. Di Mitri, M. Scheffel, H. Drachsler, D. Börner, S. Ternier, and M. Specht, "Learning pulse: A machine learning approach for predicting performance in self-regulated learning using multimodal data," in *Proceedings of the Seventh International Learning Analytics and Knowledge Conference*, 2017, pp. 188–197. https://doi.org/10.1145/3027385.3027447

[82] M. V. Torre, E. Tan, and C. Hauff, "edX log data analysis made easy: Introducing ELAT: An open-source, privacy-aware and browser-based edX log data analysis tool," in *Proceedings of the Tenth International Conference on Learning Analytics and Knowledge*, 2020, pp. 502–511. https://doi.org/10.1145/3375462.3375510

[83] A. Muslim, M. A. Chatti, T. Mahapatra, and U. Schroeder, "A rule-based indicator definition tool for personalized learning analytics," in *Proceedings of the Sixth International Conference on Learning Analytics and Knowledge*, 2016, pp. 264–273. https://doi.org/10.1145/2883851.2883921

[84] P. H. Winne and A. F. Hadwin, "nStudy: Tracing and supporting self-regulated learning in the internet," *International Handbook of Metacognition and Learning Technologies*, vol. 28, pp. 293–308, 2013. https://doi.org/10.1007/978-1-4419-5546-3_20

[85] S. Knight, A. Shibani, S. Abel, A. Gibson, and P. Ryan, "AcaWriter: A learning analytics tool for formative feedback on academic writing," *Journal of Writing Research*, vol. 12, no. 1, pp. 141–186, 2020. https://doi.org/10.17239/jowr-2020.12.01.06

[86] A. Pardo *et al.*, "OnTask: Delivering data-informed, personalized learning support actions," *Journal of Learning Analytics*, vol. 5, no. 3, pp. 235–249, 2018. https://doi.org/10.18608/jla.2018.53.15

[87] R. Bodily, T. K. Ikahihifo, B. Mackley, and C. R. Graham, "The design, development, and implementation of student-facing learning analytics dashboards," *Journal of Computing in Higher Education*, vol. 30, pp. 572–598, 2018. https://doi.org/10.1007/s12528-018-9186-0

[88] S. Charleer, A. V. Moere, J. Klerkx, K. Verbert, and T. De Laet, "Learning analytics dashboards to support adviser-student dialogue," *IEEE Transactions on Learning Technologies*, vol. 11, no. 3, pp. 389–399, 2017. https://doi.org/10.1109/TLT.2017.2720670

[89] T. Broos, L. Peeters, K. Verbert, C. Van Soom, G. Langie, and T. De Laet, "Dashboard for actionable feedback on learning skills: Scalability and usefulness," in *Learning and Collaboration Technologies. Technology in Education. LCT 2017,* in Lecture Notes in Computer Science, P. Zaphiris and A. Ioannou, Eds., Springer, Cham., vol. 10296, 2017, pp. 229–241. https://doi.org/10.1007/978-3-319-58515-4_18

[90] T. Broos, K. Verbert, G. Langie, C. Van Soom, and T. De Laet, "Multi-institutional positioning test feedback dashboard for aspiring students: Lessons learnt from a case study in flanders," in *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, 2018, pp. 51–55. https://doi.org/10.1145/3170358.3170419

[91] F. Sense, M. van der Velde, and H. van Rijn, "Predicting university students' exam performance using a model-based adaptive fact-learning system," *Journal of Learning Analytics*, vol. 8, no. 3, pp. 155–169, 2021. https://doi.org/10.18608/jla.2021.6590

[92] M. Afzaal *et al.*, "Explainable AI for data-driven feedback and intelligent action recommendations to support students' self-regulation," *Front. in Artifi. Intell.*, vol. 4, p. 723447, 2021. https://doi.org/10.3389/frai.2021.723447

[93] D. Baneres, M. E. Rodríguez-Gonzalez, and M. Serra, "An early feedback prediction system for learners at-risk within a first-year higher education course," *IEEE Transactions on Learning Technologies*, vol. 12, no. 2, pp. 249–263, 2019. https://doi.org/10.1109/TLT.2019.2912167

[94] T. Staubitz and C. Meinel, "Collaborative learning in moocs approaches and experiments," in *2018 IEEE Frontiers in Education Conference (FIE)*, 2018, pp. 1–9. https://doi.org/10.1109/FIE.2018.8659340

[95] B. Schneider and R. Pea, "Real-time mutual gaze perception enhances collaborative learning and collaboration quality," in *Educational Media and Technology Yearbook*, M. Orey and R. Branch, Eds., Springer, Cham., vol. 40, 2017, pp. 99–125. https://doi.org/10.1007/978-3-319-45001-8_7

[96] M. Afzaal *et al.*, "Explainable AI for data-driven feedback and intelligent action recommendations to support students self-regulation," *Front. in Artif. Intell.*, vol. 4, 2021. https://doi.org/10.3389/frai.2021.723447

[97] L. D. Bennion *et al.*, "Early identification of struggling learners: Using prematriculation and early academic performance data," *Perspectives on Medical Education*, vol. 8, no. 5, pp. 298–304, 2019. https://doi.org/10.1007/S40037-019-00539-2

## 9 AUTHORS

**Muhammad Afzaal** is with the Department of Computer and Systems Sciences, Stockholm University, Stockholm, Sweden (E-mail: muhammad.afzaal@dsv.su.se).

**Jalal Nouri** is with the Department of Computer and Systems Sciences, Stockholm University, Stockholm, Sweden.