

## Design and Implementation of a Learning Analytics Toolkit for Teachers

**Anna Lea Dyckhoff\***, **Dennis Zielke**, **Mareike Bültmann**, **Mohamed Amine Chatti** and **Ulrik Schroeder**

Computer Supported Learning Group, RWTH Aachen University, D-52056 Aachen, Germany //  
dyckhoff@informatik.rwth-aachen.de // chatti@informatik.rwth-aachen.de // schroeder@informatik.rwth-aachen.de  
// dennis.zielke@rwth-aachen.de // mareike.bueltmann@rwth-aachen.de

\*Correspondence author

### ABSTRACT

Learning Analytics can provide powerful tools for teachers in order to support them in the iterative process of improving the effectiveness of their courses and to collaterally enhance their students' performance. In this paper, we present the theoretical background, design, implementation, and evaluation details of eLAT, a Learning Analytics Toolkit, which enables teachers to explore and correlate learning object usage, user properties, user behavior, as well as assessment results based on graphical indicators. The primary aim of the development of eLAT is to process large data sets in microseconds with regard to individual data analysis interests of teachers and data privacy issues, in order to help them to self-reflect on their technology-enhanced teaching and learning scenarios and to identify opportunities for interventions and improvements.

### Keywords

Learning analytics, EDM, Improving educational software, Teacher support

### Introduction

Learning Management Systems (LMS) or Virtual Learning Environments (VLE) are widely used and have become part of the common toolkits of educators (Schroeder, 2009). One of the main goals of the integration of traditional teaching methods with technology enhancements is the improvement of teaching and learning quality in large university courses with many students. But does utilizing a VLE automatically improve teaching and learning? In our experience, many teachers just upload existing files, like lecture slides, handouts and exercises, when starting to use a VLE. Thereby availability of learning resources is improved. For improving teaching and learning it could be helpful to create more motivating, challenging, and engaging learning materials and e.g., collaborative scenarios to improve learning among large groups of students. Teachers could e.g., use audio and video recordings of their lectures or provide interactive, demonstrative multimedia examples and quizzes. If they put effort in the design of such online learning activities, they need tools that help them observe the consequences of their actions and evaluate their teaching interventions. They need to have appropriate access to data to assess changing behaviors and performances of their students to estimate the level of improvement that has been achieved in the learning environment.

With the establishment of TEL, a new research field, called Learning Analytics, is emerging (Elias, 2011). This research field borrows and synthesizes techniques from different related fields, such as Educational Data Mining (EDM), Academic Analytics, Social Network Analysis or Business Intelligence (BI), to harness them for converting educational data into useful information and thereon to motivate actions, like self-reflecting ones previous teaching or learning activities, to foster improved teaching and learning. The main goal of BI is to turn enterprise data into useful information for management decision support. However, Learning Analytics, Academic Analytics, as well as EDM more specifically focus on tools and methods for exploring data coming from educational contexts. While Academic Analytics take a university-wide perspective, including also e.g., organizational and financial issues (Campbell & Oblinger, 2007), Learning Analytics as well as EDM focus specifically on data about teaching and learning.

Siemens (2010) defines Learning Analytics as “the use of intelligent data, learner-produced data, and analysis models to discover information and social connections, and to predict and advise on learning.” It can support teachers and students to take action based on the evaluation of educational data. However, the technology to deliver this potential is still very young and research on understanding the pedagogical usefulness of Learning Analytics is still in its infancy (Johnson et al., 2011b; Johnson et al., 2012).

It is a current goal at RWTH Aachen University to enhance its VLE—the learning and teaching portal L<sup>2</sup>P (Gebhardt et al., 2007)—with user-friendly tools for Learning Analytics, in order to equip their teachers and tutors with means to evaluate the effectiveness of TEL within their instructional design and courses offered. These teachers still face difficulties, deterring them from integrating cyclical reflective research activities, comparable to Action Research, into everyday practice. Action Research is characterized by a continuing effort to closely interlink, relate and confront action and reflection, to reflect upon one’s conscious and unconscious doings in order to develop one’s actions, and to act reflectively in order to develop one’s knowledge.“ (Altrichter et al., 2005, p. 6). A pre-eminent barrier is the additional workload, originating from tasks of collecting, integrating, and analyzing raw data from log files of their VLE (Altenbernd-Giani et al., 2009). To tackle these issues, we have developed the “exploratory Learning Analytics Toolkit” (eLAT). The main aim of eLAT is to support reflection on and improvement of online teaching methods based on personal interests and observations. To help teachers reflect on their teaching according to their own interests, the desired Learning Analytics tool is required to provide a clear, simple, easily interpretable and usable interface while, at the same time, being powerful and flexible enough for data and information exploration purposes. Therefore, eLAT was designed to enable teachers to explore and correlate content usage, user properties, user behavior, as well as assessment results based on individually selected graphical indicators. Dyckhoff et al. (2011) gave a short overview of the toolkit. In the remainder of this paper, we present the theoretical background, design, implementation, and evaluation results of eLAT in more detail.

In section 2 (theoretical background), we provide information on the theoretical background and briefly describe the results of a requirements analysis and its implications for Learning Analytics tools. In section 3 (eLAT: exploratory Learning Analytics Toolkit), we discuss the design, implementation, and evaluation of eLAT. In section 4 (Related Work), we compare our approach to state-of-the-art solutions. Even though there are some approaches to support teachers in their ongoing evaluation and improvement activities (e.g., Johnson et al., 2011a; García-Saiz & Zorilla PantaLeón, 2011; Mazza & Dimitrova, 2007; Pedraza-Perez et al., 2011), many challenges remain (Chatti et al., 2012). Examples include integration with other VLEs and integration of diverse data sources, minimizing the time delay between the capture and use of data, consideration of data privacy issues, protection of students’ identities and prevention of data misuse, enabling data exploration and visualization manipulation based on individual data analysis interests, providing the right information to the right people right away, and investigating which captured variables may be pedagogically meaningful (Elias, 2011; Johnson et al, 2011b). By developing eLAT, we have tried to tackle these challenges. The final section 5 (Conclusion and Outlook) gives a summary of the main findings of this study according to the challenges mentioned above and outlines perspectives for future work.

## **Theoretical background**

In TEL, masses of data can be collected from different kinds of student actions, such as solving assignments, taking exams, online social interaction, participating in discussion forums, and extracurricular activities. This data can be used for Learning Analytics to extract valuable information, which might be helpful for teachers to reflect on their instructional design and management of their courses.

Usable EDM and Learning Analytics tools for teachers that support cyclical research activities are still missing in most current VLE or are far from satisfactory (Hijon & Carlos, 2006). Romero et al. state “[...] data mining tools are normally designed more for power and flexibility than for simplicity. Most of the current data mining tools are too complex for educators to use and their features go well beyond the scope of what an educator might require” (Romero et al., 2007, p. 369). If tracking data is provided in a VLE, it is often incomprehensible, poorly organized, and difficult to follow, because of its tabular format. As a result, only skilled and technically savvy users can utilize it (Mazza & Dimitrova, 2007). But even for them it might be too time consuming. Moreover, unnecessary personal information of students can be observed by teachers or even fellow students, i.e., data privacy issues are ignored in the design of most VLE (Loser & Herrmann, 2009). Legal issues have to be taken into account to prevent malpractice. In Germany, for example, teachers are not allowed to have access to everything a student does in their online courses. They are only supposed to have access to data that is relevant for teaching in a form that is transparent to students (Directive 95/46/EC, 1995; Federal Data Protection Act, 1990). However, many research questions of teachers are concerned with general learning processes of a whole group of students in contrast to gaining more knowledge about a single student. Therefore, to ensure data privacy, student data could be pseudonymized in a preprocessing step by using, e.g., a hash instead of a student-ID, and by presenting summarized results in form of visualizations that rather show group processes and do not allow to focus on one student. Further

deficiencies of reporting tools are related to usability and clarity as well as completeness of the delivered results, such as the lack of possibilities to integrate results of online questionnaires with data from logs.

Many teachers are motivated to evaluate their courses and they already have research questions related to their teaching in mind. For example, a teacher who offers weekly online exercises has the intention to help her students to prepare for an exam. But she is not sure if the currently available exercises are helpful enough for this purpose. Therefore, the teachers would like to know if those students who practice with her online exercises on a continually basis are better in the final exam than students who do not use them. A Learning Analytics toolkit could help her to do research on this hypothesis by automatically collecting, analyzing, and visualizing the right data in an appropriate way. Yet, most monitoring and reporting tools found in current VLEs are designed to collect, analyze, and visualize data in a static tabular form that was predefined by system developers. Teachers face the difficulty that appropriate and usable Learning Analytics tools that help them answer their individual questions continuously and efficiently are missing in prevalent VLEs, since most of the work in the area of Learning Analytics is conceptual (Johnson et al., 2011b).

Teachers should have access to Learning Analytics tools (e.g., provided via dashboards) that can be integrated into a VLE or other learning environments. These tools should allow for interactive configuration in such a way that its users could easily analyze and interpret available data based on individual interests. Results of Learning Analytics and EDM should be presented in a clear and understandable format, e.g., information visualizations that are understandable without data mining expert knowledge. Card et al. (1999) define the term visualization as “the use of computer-supported, interactive, visual representations of data to amplify cognition.” It “promises to help us speed our understanding and action in a world of increasing information volumes” (Card, 2003, p. 542). It has been widely argued in the EDM literature that teachers can grasp the information more easily and quickly when presented through comprehensible information visualizations. Mazza and Dimitrova (2007, p. 138), for instance, write: “...the effectiveness of [Course Management Systems] can be improved by integrating [Information Visualization] techniques to generate appropriate graphical representations ....” Also, results of a recent study by Ali et al. (2012) showed that “visualization can be an effective mean to deal with larger amounts of data in order to sustain the cognitive load of educators at an acceptable level” (p. 486) and “multiple ways of visualizing data increase the perceived value of different feedback types” (p. 488). Still, it should be noted that in some analytical cases visualizations could be ineffective with respect to a textual or a tabular interface, e.g., if details about many items are very important.

In our approach, we only indicate certain facts about the usage and properties of the learning environment and try to visualize them appropriately. Therefore, we revert to the concept of *indicators*, which can be described as specific calculators with corresponding visualizations, tied to a specific question. For example, if the teacher’s question is “Are those students who practice with online exercise on a continually basis are better in the final exam than students who do not use them,” the corresponding indicator could show a chart that quickly facilitates a visual data comparison. Indicator concepts have been used before. Glahn (2009) for example, introduced the concept of *smart indicators*, which he defined as “a context aware indicator system, which dynamically aligns data sources, data aggregation, and data presentation to the current context of a learner” (Glahn, 2009, p. 19). However, in our case the target group differs. The eLAT indicators are collecting and visualizing data of students to present them to teachers.

A typical Learning Analytics process is depicted in figure 1. The process starts with the data-gathering step. In this step, data is collected from different learners’ activities when they interact with learning elements within a VLE, LMS or a personal learning environment (PLE). Examples of these activities include participation in collaborative exercises, writing a forum post or reading a document. In the data collection step, it is crucial to address data privacy issues. Often the output of the data extraction and preprocessing step is transferred into a separate database. The second step of the Learning Analytics process is the mining of the preprocessed data, based on different mining techniques, such as clustering, classification, association rule mining, and social network analysis. Thereafter, the results of the mining process can be presented as a widget, which might be integrated into a VLE, a dashboard, or a PLE. Based on appropriate graphical visualizations of the analyzed data, teachers are supposed to be able to more quickly interpret the visualized information, reflect on the impact of their teaching method on the learning behavior and performance of their students, and draw first conclusions about the effectiveness of their teaching, i.e., consider if their goals have been reached. Furthermore, unexpected findings should motivate them to iteratively improve their teaching interventions. However, having a graphical visualization does not guarantee that teachers will be able to

interpret the information represented correctly. Indicators must be designed and evaluated carefully. Also, the system should provide instructions for interpretation.

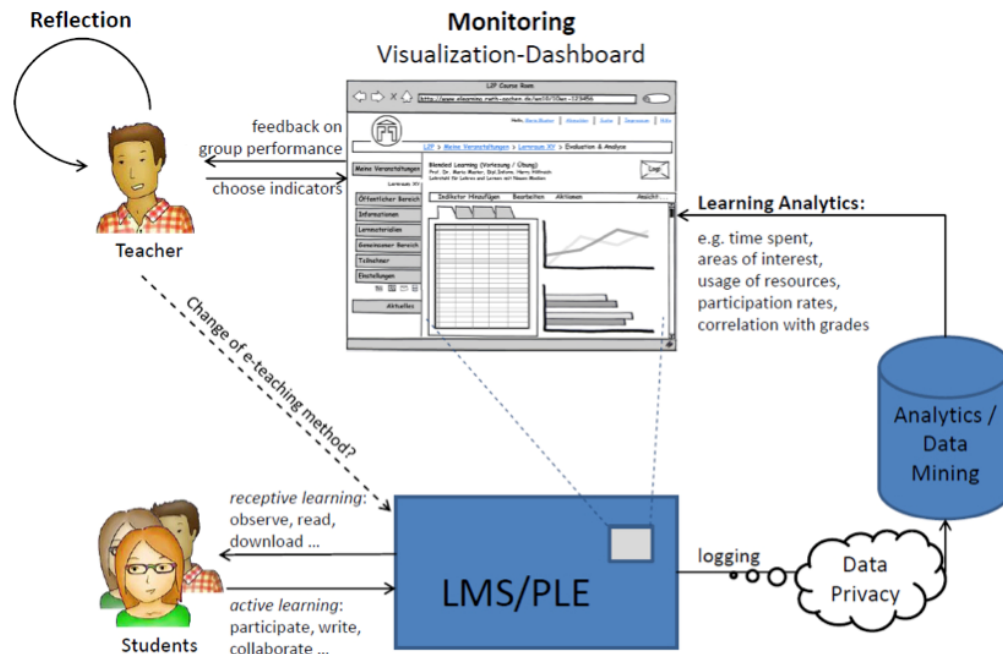


Figure 1. The Learning Analytics Process

Requirements for developing such dedicated systems have been collected in a former study by analyzing interests and needs of the target group in more detail (Dyckhoff, 2010). Results of this study showed that teachers already have various questions about their instructional design and the utilization of learning materials, the students' learning behaviors and correlations between objects of teaching and learning as well as outcomes. Their intentions can be e.g., to find out how well the overall instructional design is appreciated, to learn more about the needs of all or a specific group of students, or to better understand learning processes in general.

The conclusion from the study mentioned above was that Learning Analytics tools should support teachers by collecting, integrating, and analyzing data of different sources as well as by providing a step-by-step guidance including semi-automated processes, instead of just presenting large tables of data. "It is undisputable that statistics in isolation represent only one aspect of any real-world situation. To make more meaningful interpretations, [educators] often need to look at the two or more statistics together" (Ali et al., 2012, p. 484). Hence, teachers should be able to choose from a flexible and extendable set of indicators. The system should guide the user throughout the research process, help him or her form research questions, recommend and provide appropriate methods for data collection, integrate data from different sources, and support its collaboratively organized analysis. Such a Learning Analytics tool could e.g., provide extendable lists of supported research questions (indicators) and suitable qualitative as well as quantitative methods for data collection, visualization and analysis. Furthermore, it should be possible to use and integrate the tool with any kind of VLE and learning software.

### eLAT: exploratory Learning Analytics Toolkit

In the following sections, we introduce eLAT by giving an overview about results of the requirement analysis, development stages, design decisions, and evaluation phases, concluding with a discussion of our basic findings.

### Requirements

Requirements for eLAT have been collected through literature analysis (Dyckhoff, 2010), as well as by informally talking to teachers, eLearning experts and system administrators at RWTH Aachen. The requirements analysis concluded the following main design goals:

- *Usability*: prepare an understandable user interface (UI), appropriate methods for data visualization, and guide the user through the analytics process.
- *Usefulness*: provide relevant, meaningful indicators that help teachers to gain insight in the learning behavior of their students and support them in reflecting on their teaching.
- *Interoperability*: ensure compatibility for any kind of VLE by allowing for integration of different data sources.
- *Extensibility*: allow for incremental extension of analytics functionality after the system has been deployed without rewriting code.
- *Reusability*: target for a building-block approach to make sure that re-using simpler ones can implement more complex functions.
- *Real-time operation*: make sure that the toolkit can return answers within microseconds to allow for an exploratory user experience.
- *Data Privacy*: preserve confidential user information and protect the identities of the users at all times.

*Usability and usefulness*: Every course is different, depending on the teachers and students who are involved in it. There are different teaching strategies, different learning goals, etc. Among the teachers there are some who have not used a Learning Analytics tool before, as well as advanced users. A Learning Analytics tool should be easy to use and understandable for all users. It must be usable for both: the beginner, who just looks at it for the first time, as well as for the expert, who already has a specific question and wants to perform deeper analysis. For beginners, a Learning Analytics tool should enable a direct entry and it should motivate to occupy themselves more with the underlying data, e.g., through a dashboards solution. Experts should find ways to explore and do further analysis to keep them well on the ball. Varying learning scenarios will also demand differing sets of indicators. An important future research task is to find out which indicators are useful for whom in what situation.

*Interoperability, extensibility, and reusability*: Most existing Learning Analytics tools cannot be easily adapted for a different VLE. In addition, new e-learning systems are being developed that may contain useful data for Learning Analytics. Also, learning may take place on informal learning platforms. Therefore, an interoperable Learning Analytics tool that integrates with other systems, and can collect and analyze data from different platforms is required.

*Real-time operation*: New issues on a course may arise at any time during a semester and should then usually be answered directly, so that timely improvements can be made. Also, new questions that are worth to be examined more closely, may arise during the answering process of ongoing questions. Therefore, a Learning Analytics tool should provide current data and comprehensive data analysis capabilities and be available at all times, not only at the end of the semester. Also, interactive analysis and visualization features, like filtering options for exploring the data in more detail, should deliver results and changing visualizations within microseconds.

*Data privacy*: Personal data should be protected at all times to prevent abuse. Data privacy acts ensure such protection (e.g., Directive 95/46/EC, 1995; Federal Data Protection Act, 1990). However, an exception is made for teaching and research projects, under the condition that the data will be handled transparently and purposefully (Federal Data Protection Act, 1990). In addition, students or a data protection officer could be asked to consent to the collection and analysis of student data. Many questions regarding teaching, however, do not aim to examine records of individual students. Rather, data of the totality of students or subgroups with specific characteristics are interesting for drawing conclusions on learning processes. Data could be stored and processed pseudonomized to protect the users. As further protection, the tool could ensure that certain kinds of analyses cannot be executed in certain situations, where they would lead to the identification of individual students.

## **Development stages and evaluation methods**

eLAT was iteratively and incremental developed within two main stages that partially overlapped to meet the requirements described above: (stage 1) the implementation and testing of a backend framework as well as (stage 2) the design and evaluation of a UI (frontend).

In the first stage eLAT was designed as a prototype to evaluate different software architectural approaches for Learning Analytics using different VLE platforms. During winter term 2010/2011, we selected four courses that were using the learning and teaching portal L<sup>2</sup>P of RWTH Aachen University. The courses differed in course sizes

(1370, 338, 220, and 38 registered students), learning technologies and teaching styles to ensure realistic usage scenarios. We logged the students' activity, interaction and (in one case the) assessment data over the duration of three months. By using the data of real courses, it was possible to learn more about meaningfulness of already implemented indicators and to let the teachers of the courses participate in the development process of eLAT. In this way, we could get immediate feedback and comments on prototype stages that already processed analytics based on the real data.

The design and evaluation of a UI (stage 2) started parallel to the first development stage described above. It was iteratively conducted as well, whereat each of overall three iterations had a specific objective. The first iteration dealt with the collection and definition of the content. Since eLAT was designed to enable teachers to explore educational data of their students and courses based on graphical indicators, it involved the collection of indicators as well as assigning priorities to them. Thus, semi-structured interviews were conducted to evaluate a set of graphical indicators and to get to know further user requirements. Semi-structured interviews are used to collect facts, opinion and attitudes of the interview partners (Naderer, 2007). The interviews are guided through prepared questions, but it is also possible to ask questions spontaneously to investigate interesting details (Lindlof and Tylor, 2002). The second iteration focused on the layout and data presentation of the UI. The evaluations of the first and second iteration were performed with the help of paper prototypes, while in the third iteration a functional UI, which was implemented based on previous evaluation results, was used to investigate interactivity and usability aspects.

Layout and data presentation were designed and evaluated in terms of heuristic evaluation, cognitive walkthrough and pluralistic walkthrough. A heuristic evaluation uses approved usability principles or guidelines to investigate the usability of a UI. Thus, problems can be discovered with less effort in an early development step (Nielsen, 1992). A cognitive walkthrough is more formal than heuristic evaluation. It needs a specification of the UI and tasks to evaluate the usability. With the help of the tasks the UI can be processed step by step to discover usability problems (Polson et al., 1992; Dix et al., 2004). Both methods were chosen to evaluate the prototypes of the UI in an early stage of development. The pluralistic walkthrough is similar to the cognitive walkthrough. It is a meeting of experts of different domains, such as users, designers, and usability specialist. They discuss elements of the interface prototype according to the view of the users (Bias, 1994). We used a variant of the pluralistic walkthrough where a domain expert, a usability experts and the designer discussed the interface from the users' perspective. Main results of these studies are presented in the section "User interface".

The third iteration, which was mainly concerned with interaction, included a qualitative think-aloud study. Here users were asked to perform tasks with a software prototype, whereat they were talking about what they were doing and thinking. During the tasks the evaluator observed them. This method was chosen to identify areas of interactions where users can make mistakes (Dix et al., 2004).

In the following sections, we present the resulting UI, use cases, design and implementation details of eLAT as well as overall evaluations results.

## **User interface**

The structure and layout of the eLAT user interface (UI) are the result of an iterative approach and were derived from our user studies, which have been elaborately discussed in Bültmann (2011). The UI is designed as a launch pad, which is similar to a dashboard but provides more comprehensive analysis options, additional to an initial overview (Few, 2006).

A monitoring view helps to observe several indicators at once (figure 2). Furthermore, analysis views provide a deeper insight into the data of chosen indicators by making it possible to drill down into details by changing parameters of an indicator. Additionally, a mouse over effect shows details about the currently regarded information (figure 3).

In the *monitoring view*, the content of the launch pad is grouped into four widgets. The widgets are containers for indicators related to the categories "document usage," "assessment/ performance," "user activity" and "communication." Each indicator has its own tab in the widget. This hierarchical layout is supposed to help users to get a better overview about the current learning situation. By using widgets and tabs it is possible to put all indicators

on one single screen. This concept also helps in terms of personalization, because widgets can be arranged flexible by the users.

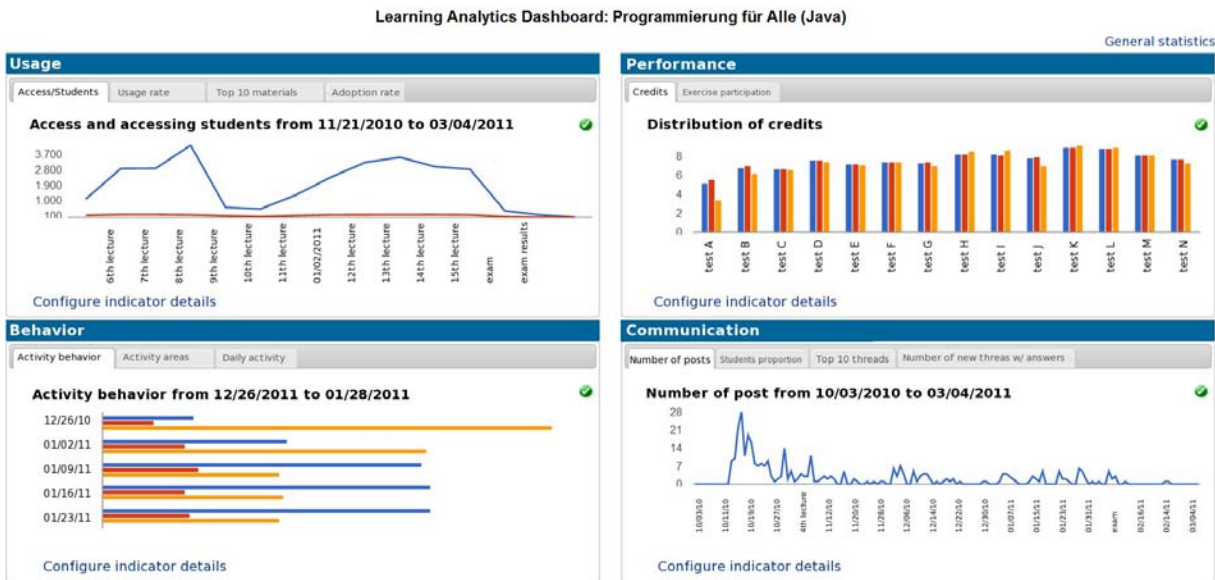


Figure 2. Monitoring view of the eLAT user interface

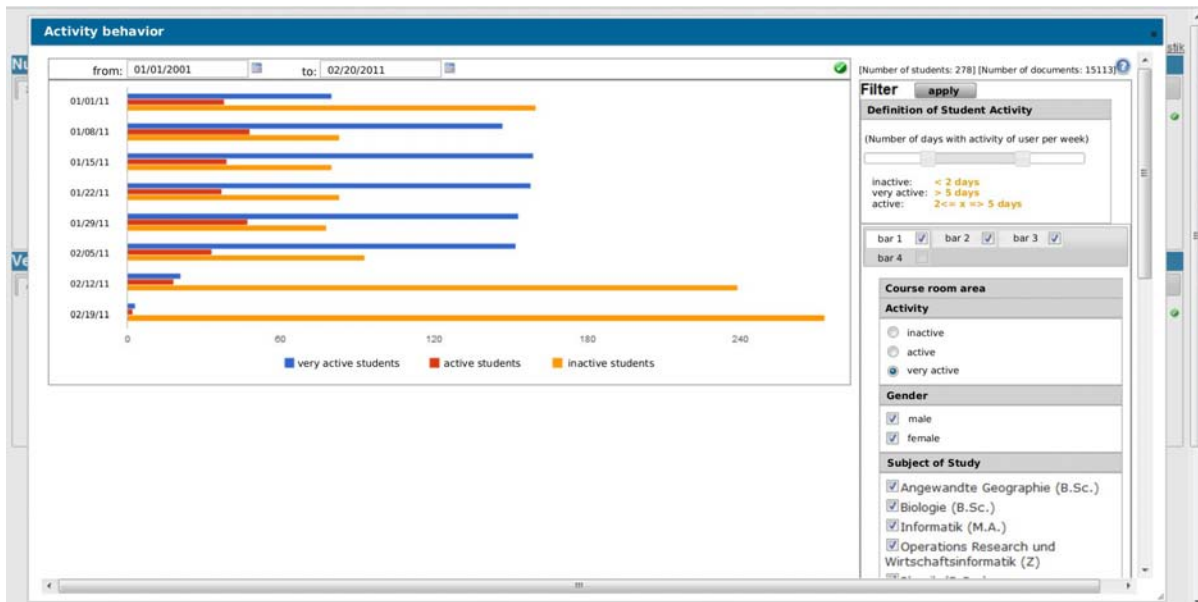


Figure 3. Analysis view of the indicator "Activity behavior"

The *analysis view* of each indicator is consistently accessible in the corresponding tab by clicking "Configure indicator details" (figure 2). This detailed view of the indicator is shown as an overlay on top of the monitoring view (figure 3). Layout and functionality have been designed in a consistent way to gain a better usability (Few, 2006). On the right side of the analysis view is a filtering menu. The filtering of the presented data is context-dependent according to the currently selected indicator. For each tab in the filtering menu of any indicator, the user can determine which information the indicator should present. Hence, there are many options for data exploration, such as, comparing the activity of male and female users or the activity of students of different study programs. But not all filters can be used for each context. Because of data privacy regulations, we cannot allow the use of any kind of user properties like gender or study course, when there are less than a certain number of students, e.g. five users with that

certain property in a course. The following paragraphs give an overview about six implemented indicators, which were rated to be interesting as result of the evaluations.

Figure 3 shows the analysis view of the indicator “Activity behavior”. Student data is divided into the three groups “very active students” (blue bars), “active students” (red bars), and “inactive students” (yellow bars), and shows their weekly distribution over a time span. An “active student” is determined by calculating the average number of days per week, at which a student was active, i.e., logged into the system. In the current configuration of the indicator “activity behavior”, shown in figure 3, a student is defined to be “active” if he or she logs in at least once a week. A student is defined to be “very active” if he or she logs in on more than five days a week. The user can change the time span and the definition of an “active student.”

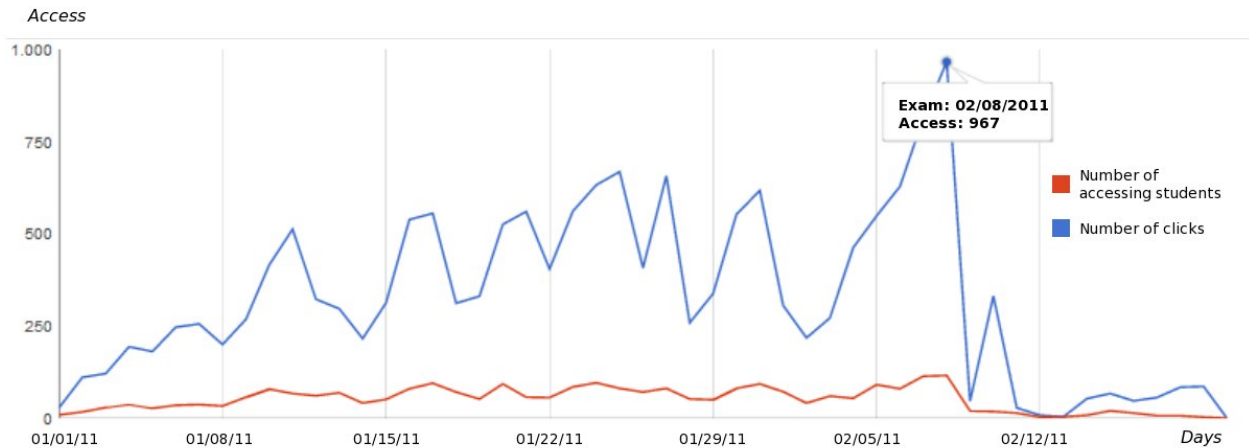


Figure 4. Indicator “Access and accessing students”

The data in figures 3–7 is based on a programming course, which was finished with a final exam at February 8<sup>th</sup> 2010. As expected, the indicator above shows an increase of “inactive students” after the exam date. The indicator “Activity behavior” (figure 3) indicates whether continues learning is taking place. A participant of our semi-structured interviews considered continues learning as a main factor for good exam results. As a sign of continues learning, the teacher might e.g., expect his students to log in at least twice a week to download new materials and stay up-to-date related to course information. The indicator “Activity behavior” can show tendencies of increasing or decreasing numbers of such active (groups of) students. High numbers of inactive students during the semester could bring the teacher to motivate his students to learn more regularly, e.g. through creating weekly exercises, or initialize further investigations on the reasons of low activity.

The indicator “access and accessing students” (figure 4) supports the teachers in monitoring the overall online activity of their course. It shows the number of accesses/clicks (blue line) over the number of unique students (red line) who accessed the virtual learning environment during a time span defined by the user. The blue line represents the sum of every single click on any resources in the learning environment per day or week. It is important for a teacher to observe if e.g., a small group of students clicks many times or many students click once on a resource. The data in figure 4 shows that almost every day about a third of 278 registered students accessed several resources. The peak at the end of the timeline demonstrates a strong increase in accesses before the final exam, but only a small increase in accessing students. Lines converge after the date of the exam. Probably, the students only come back to the virtual course room to check the exam results (one click per student). The indicator “Access and accessing students (figure 4) can show outliers of usual access behavior/frequency. Teachers can relate high or low usage e.g., to teaching events or holidays. They can quickly observe if changes of learning materials or didactics lead to changes in overall usage behaviors. This might motivate them to experiment with didactics to improve the overall access to the learning environment.



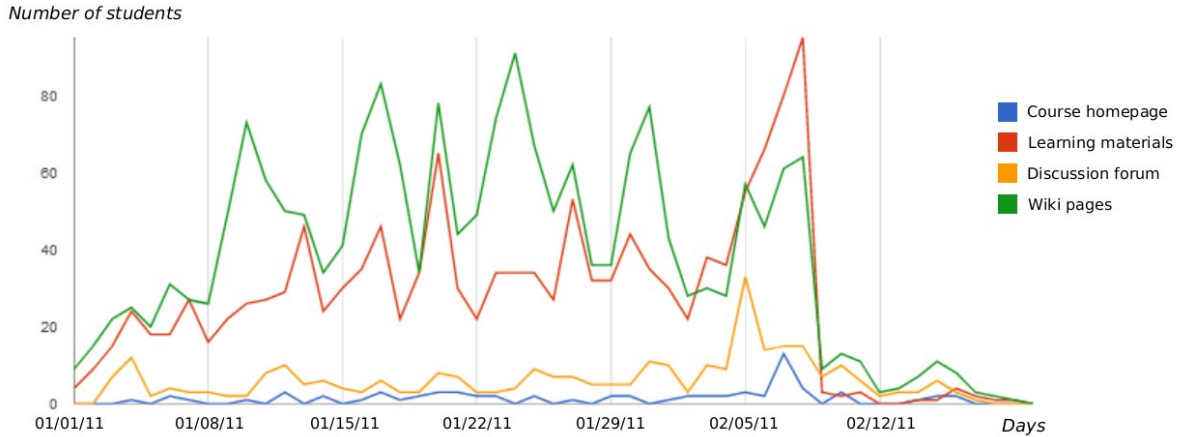


Figure 5. Indicator “Activity areas”

With help of the “Activity areas” indicator, shown in figure 5, teachers are supposed to identify whether and when students are accessing which parts/areas of a virtual course room per week in a defined time span. Hence, access rates between functions, like wiki pages or discussion forum, can be compared and related to teaching events as well. The x-axis of the indicator shows the days or weeks. If metadata on dates of course events, like the occurrence of specific lectures or the exam, are provided, these events can also be written on the x-axis. The y-axis records the number of students, who were active, i.e., clicked on resources, during that day or week in a specific part of the virtual course room. The red line in figure 5 e.g., shows the number of students accessing a document library with learning materials, such as lecture scripts and exercises. The red line and the yellow line, which represents the number of students, who accessed the discussion forum, peak out 1–3 day before the exam. Students seem to become more active in reading and discussing during that time, so that the case could be made that they are learning more intensively.

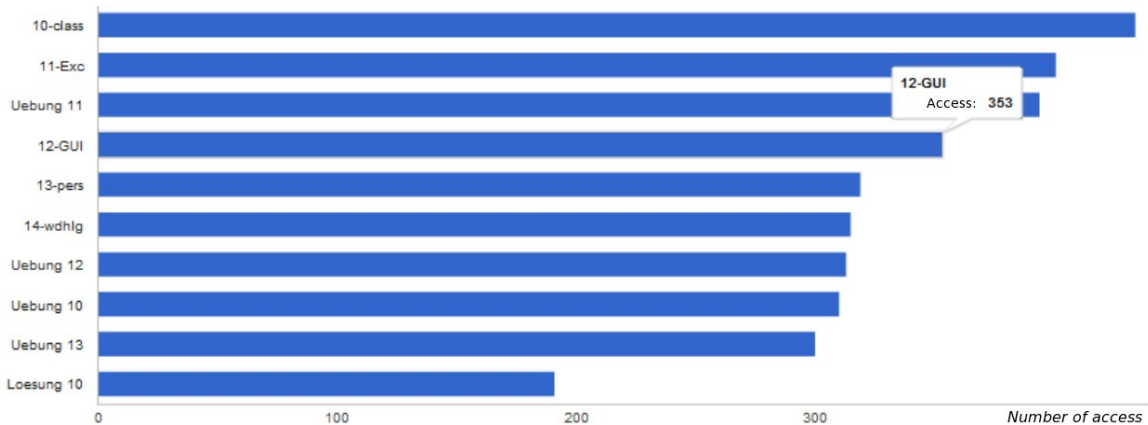


Figure 6. Indicator “Top 10 resources”

The “Top 10 resources” indicator (figure 6) gives an overview about the most accessed materials. It can help to identify active documents/items that have been accessed more than others. Such a popularity indicator could have differing reasons. A document that shows up in the “Top 10 resources” indicator e.g., could be useful for solving an exercise or might be difficult to understand. The learning materials, presented in figure 6, are exercises (“Uebung 10–13”), code examples (“10-class”), lecture scripts (“12-GUI” and “11-Exc”), a lecture summary (“14-wdhlg”) and an example solution (“Loesung 10”). This could indicate that students mainly have been learning by solving exercises. Based on this indicator, a teacher could start to explore the meaning of the high access of specific learning materials. In case of a difficulty of understanding, learning materials could be improved.

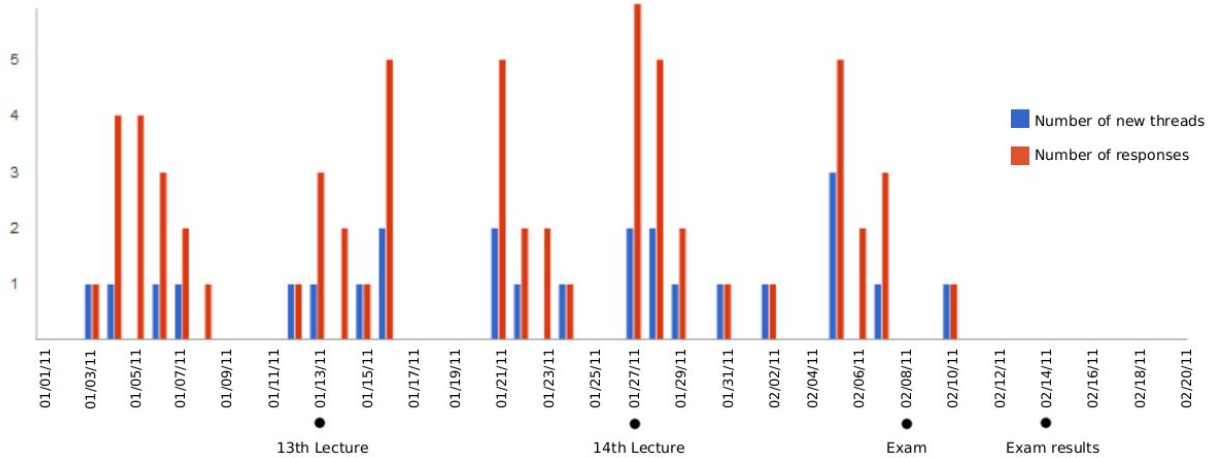


Figure 7. Indicator “Forum usage”

The “Forum usage” indicator (figure 7) represents the number of new threads with corresponding answers to these threads (x-axis) per day (y-axis). Teacher can more easily identify increasing discussions and, thus, determine, if collaboration among students is taking place, and whether there might be problems or not. Furthermore, by looking at a thread title of the observed communication activity, problems in understanding or in preparation of learning materials could be identified. Although this indicator does not show the answers per thread it can be an activity measure for forum usage.

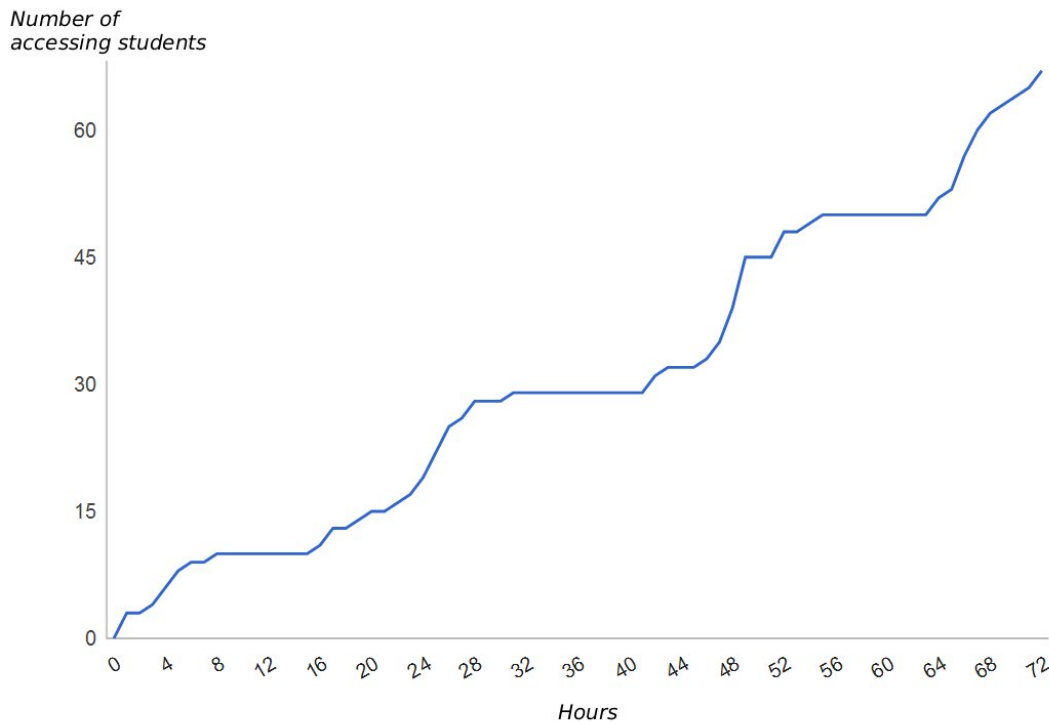


Figure 8. Indicator “Adoption rate”

The “Adoption rate” indicator (figure 8) deals with the time span from uploading a selected learning material to the time of access by students. With the help of the “adoption rate”, it can be identified how fast how many students access new materials. It also shows the number of students who have utilized a certain material and helps teachers to find out after which time the item has achieved a sufficient distribution amongst his or her students. In some courses this is helpful e.g., because the students might have a reading assignment. Thus, teachers are enabled to estimate,

how many students have at least accessed a document that they were supposed to read. If the adoption rate is lower than expected, this could be an explanation for low homework discussion participation during class.

Comprehensive and well-fitting indicators are crucial for an effective and successful application of a Learning Analytics tool. Therefore, we conducted semi-structured interviews with eight teachers (6 male, 2 female). During the interviews, example indicators were presented, shortly explained and discussed with the teachers to collect facts, opinion and new ideas. The evaluation goals named by all participants were quality and activity oriented. It is important for seven interview participants to monitor activity, like the frequency of logins or the date and occurrences of access. The principal point is to monitor, if continuous learning is taking place and if this leads to better learning results (mentioned by all participants). Therefore, it is important to relate the usage and access of learning material with performance. These evaluation goals are similar to the results found in our former study (Dyckhoff, 2010). Indicators concerning activity or achievements are valued most.

Seven participants rated the indicator “Access and accessing students” (figure 4) as helpful for a general overview, especially in the beginning of using a VLE. But the posed question, if it can always deliver interesting information, remained unanswered. Therefore, the indicator should be used in combination with other data. Six participants stated that “Adoption rate” (figure 8) is important. It can indicate student behavior. Teachers can relate it to teaching events, such as making an announcement, and observe changes in student behaviors. The “Top 10 resources” indicator (figure 6) was identified as a valuable measure (7 called it important and one somehow important). It helps to identify those resources that are somehow relevant for students. A “Top 20” might be better suited for courses with a large amount of resources. By monitoring, e.g., “Activity areas” (figure 5), it is possible to identify patterns, such as at what times students do their exercises, or whether they access learning materials before or after lectures. Activity measures in general were pointed out to be well suited for first impressions during analysis, i.e., for getting a quick overview about what is going on.

Regarding the performance, besides correlations with the activity of students, teachers also wished to take a deeper look at correlations with properties, like the program of study, the duration of study, and the mother languages of students. Five of eight participants stated that such indicators are important and one thought it is not important. This indicator could be important for adjusting teaching methods for specific groups of students.

Teachers had divergent opinions on examining active participation, e.g. the usage of forums or wiki pages (figure 7). Four participants rated it as an important measure, because communication, discussion and participation are represented in collaborative features. The other half rated it less important, because of a low participation rate in their courses or students using other collaboration tools in the cloud. The assessed usefulness of this indicator heavily depends on the participation of the students in an online course and the relevance a teacher ascribes to it. The reason can be found in the underlying structure of the hybrid courses. Communication often takes place outside the VLE because of blended learning settings at RWTH Aachen University. Students still talk and learn together outside the online learning environment, and thus, communication and its relation to learning cannot be measured adequately.

Additional to the indicators described above, we evaluated several other indicators. The final prototype does not implement all indicators evaluated in the previously conducted iterations due to the fact that not all the data needed was available at the time of implementation; e.g., it lacks of data corresponding to session information and durations or more detailed metadata about students. Those indicators, not described in this paper, have led to differing opinions. Their usefulness rating was low or very much dependent on the underlying didactical scenarios. Hence, we draw the overall conclusion that teachers should be enabled to explore data individually, e.g. by arranging their own sets of indicators per course, to facilitate improvement of teaching.

## **Framework design and implementation**

The architecture of eLAT is presented in an abstract manner in figure 9. The functional requirements for the toolkit demanded a very flexible coupling of VLE infrastructure, the implementation of an indicator evaluation process and a visualization system. eLAT encompasses three main components, namely an indicator framework, a mining database, and a visualizer application. At the heart of eLAT lies the indicator framework, which negotiates and provides report evaluation services, i.e., executions of indicator calculations, between the website and the mining

database, while the visualizer component provides an abstraction layer for visualizing different reports in an appropriate manner.

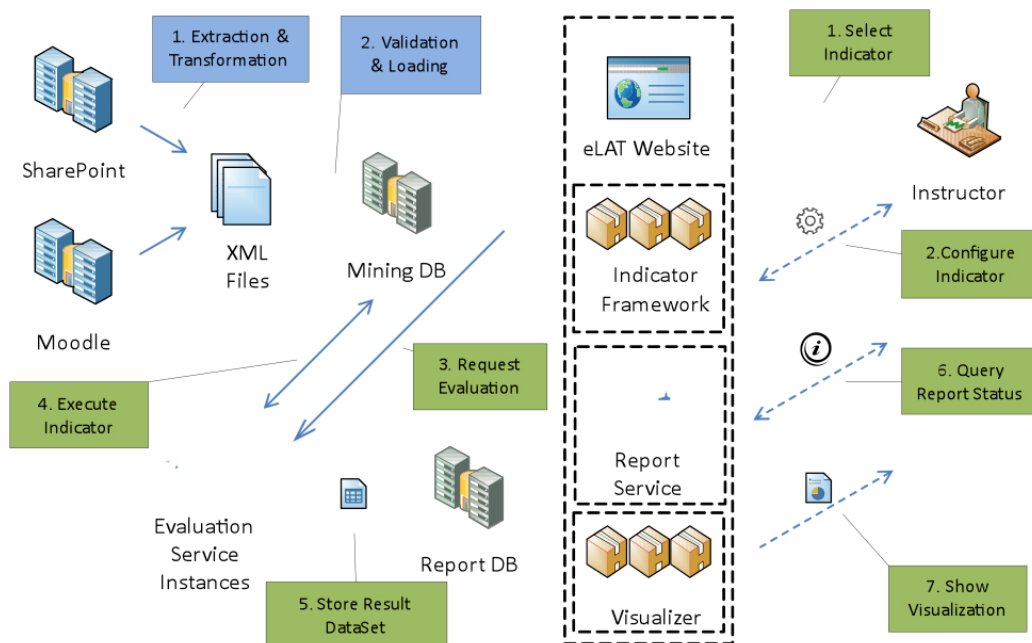


Figure 9. Indicator execution process

The implemented course of action for an indicator evaluation is illustrated in the green boxes in figure 9. Triggered by the instructor who is visiting the website and selecting an appropriate indicator (step 1), a Controller Factory in the indicator framework will then dynamically instantiate a controller for this indicator. This will create a view, containing user interfaces for all the available configuration properties (step 2). After the user has finished adjusting all the properties and the configuration has been validated, the framework will generate a report evaluation request, store it in the report database, and send an evaluation request to the evaluation service instances. In the meantime the user will be redirected to a waiting page that displays the current evaluation status and updates automatically (step 3+4). Once the evaluation has completed the report, consisting of the initial configuration and a dataset of raw data tables, it will be stored permanently in the report database (step 5). While querying the report status the client side scripts will eventually learn of the successful evaluation and load the dedicated JavaScripts that will then generate and show the appropriate visualization based on the raw data set obtained by the report service (step 6+7).

A very important part of the system architecture of eLAT deals with the extensibility and reusability of the existing code base, so that the scope of operations can grow. Therefore, a single indicator implementation makes use of smaller parts in the form of expressions that are performance-optimized database queries to retrieve specific result sets that can be useful for other indicators as well. The same practice is applied to the dynamic user interface generation for indicator parameters and the visualizers, which operate on standardized datasets and are therefore generic to the data inside the report. This leads to a relatively small effort for implementing new indicators.

Since we want to keep the eLAT implementation independent from a particular VLE, we have developed a neutral data model that supports all the major data types as well as an extension model to fit in special types. As illustrated in figure 9, we have modeled a learning environment as a general virtual learning space dedicated to a specific course, which could be implemented in any VLE. From the VLE instance we extract the information needed for further investigation and load them into the mining database:

- *User data:* This primarily stores the pseudonymous user identifier and the role of the user, e.g. “instructor”, “tutor” or “student”. It also holds custom properties like gender and study course, which might not be available for all users.
- *Content data:* In most VLE, there are multiple sources of content data, structured by using content areas and content lists or libraries, which can have predefined recurring types and certain properties for specific objectives. Additionally, there are content type extensions dedicated to extend content items with properties that are not

always available, like for example a forum post, which has a property for the depth inside the discussion thread – in any other case that post can be regarded like any other content item, with properties like title, creation date and the user responsible.

- *Assessment data*: We wanted to store assessment instances separately from the user submission to allow for different handling and support of various assessment types. This also supports an extension model to fit in properties like completion time or group work submissions, which are not common, but sometimes available.
- *Activity data*: This contains specific information about any content interaction along with timestamp and user information occurring in the learning environment.
- *Event data*: As typically every course has a calendar and special events like course start, assessment and exam dates, it is useful to use this information in the parameter selection and visualization process. We are interested in narrowing down the activity to certain time spans such as the time between two exams or visualizing the usage statistics of the forum activities only in the time between two evaluations of weekly assignments.

Most of the server-side code is written in .NET and uses Windows Communication Foundation (WCF) services for providing data and communication interfaces between the website, the evaluation service and the client site visualizations. The indicator website itself does not generate much load on the CPU or the database (answers below half a second), relative to the load the evaluation processes. This is why we decided to execute evaluation processes by a dedicated service that is running in multiple instances, ideally on a different physical machine, to approach our real-time operation requirements.

As a result of the semi-structured interviews, the possibility to filter data was identified as a substantial requirement. However, when implementing the filtering options, we did not yet entirely meet our goal of providing real-time operations, which are necessary for a smooth user experience. One reason for this lies in the structure of the database, which was created in the first development stage and, in its current form, has been designed according to a pattern used for operational databases rather than data warehouses. Hence, we implemented on-the-fly generation of lookup tables, containing pre-calculated intermediate results, to increase the efficiency of filtering. However, a thorough redesign of the database scheme with data warehouse principles in mind is required to address this issue in greater detail.

To allow for upgrading and enhancing of analysis methodology while the system is deployed, the implementation makes heavy use of the Managed Extensibility Framework, which allows for dynamic composition and object instantiation of software capabilities and is part of the latest version of the .NET Framework.

For our software prototype evaluations we extracted data for our four courses from L<sup>2</sup>P, which is currently run on SharePoint 2007, and one a Moodle instance, which was used in one course to host assessments. We employed custom code for each VLE to extract and transform data to our data model, as shown in the blue boxes in figure 8. From there the simplest way to load data into the mining database is by using the xml import feature. We have specified an xml schema, which can be used to create and validate xml files from any VLE data output.

To ensure data privacy, we had to make modifications to the data extraction process. Since we have no interest in the identity of each user, but still need to be able to uniquely distinguish their data, we chose to create a hash from the username with a learning environment specific salt, so that the transformation is not easily reversible. The algorithm used is MD5 with a hash size of 128 bits (Rivest, 1992).

In future versions, we still expect the need for custom data exports of indicator reports. Some advanced users have already expressed their interest in using the results in Excel or other analytics tools, like Weka (Hall et al., 2009) or Keel (Alcalá-Fdez et al., 2009), to start their own studies. Although we would like to provide as much functionality as possible inside the eLAT toolkit, we have designed interfaces for exporting reports.

## **Evaluation and discussion**

This section presents our basic evaluation findings and answers the two questions: Does eLAT meet the requirements “usability,” “usefulness,” “interoperability,” “extensibility,” “reusability,” “real-time operation” and “data privacy” sufficiently and what did we learn in relation to those challenges.

To evaluate if all requirements are met by our concept of a usable exploratory Learning Analytics toolkit, we have implemented an instantiation of our concept, namely the prototype eLAT, tested its technical functionality with four courses at RWTH Aachen University and conducted several usability evaluation iterations. Results of the first two iterations have been summarized in the section “User interface” by describing the UI and giving insights into indicator usefulness ratings.

The final usability study with the software prototype was conducted as a qualitative think aloud study. Four teachers (3 male, 1 female), participants of the former iterations, were asked to perform tasks, which aimed at keeping an eye on the learning progress. At first, the prototype was shown to the participants without giving a detailed explanation. The tasks were designed with increasing difficulty. All participants quickly understood the structure and were able to navigate the user interface. The monitoring and analysis views were well understood. One out of four participants tried to click and investigate the whole interface. During the evaluation, some problems were discovered.

One of the participants criticized the font size as being too small, but put into perspective that the font in the web browser could be enlarged by him and thus individually adopted. Furthermore, the mouse over effect, which shows details on the currently regarded information, helped to overcome the drawback of a small font. Two participants did not immediately understand the wording of the indicator “Adoption rate” (figure 8). Regarding the filtering, three participants requested more analysis options than possible. One wished the possibility to get into details by simply clicking on a line of a chart. As an improvement for filtering options, one participant mentioned the possibility of providing templates for common time spans during the lecture period, such as “examination phase” or “examination preparation phase”. Two participants stated that the tool should be extended regarding the correlation of activity and performance. Also the filtering possibilities should be extended and some filters also could be applied in other indicator contexts. Furthermore, filters according to the activity behavior of students could be defined. In addition, advanced users still need more personalization and more analysis functionality, e.g. the possibility to create new indicator templates.

However, although the evaluation showed that eLAT is usable, comprehensive field-tests with more courses from different disciplines still have to be carried out to gain more reliable data on pedagogical usefulness of the toolkit. We investigated the usability of several indicators. Yet, we do not have enough data to answer the question which captured variables may be pedagogically meaningful sufficiently. Indicators implemented in eLAT are rather general and focused on simplicity and understandability. For enabling data exploration and visualization manipulation based on individual data analysis interests, more significant indicators that are still easy to use, have to be designed, implemented, and evaluated. Reviewers suggested e.g., considering outliers of student behavior by using a box plot for visualization. May et al. (2010) suggested potentially meaningful indicators, such as, indicators for visualizing different user activities as unique colored spheres on a horizontal bar, or a radar graph for visualizing different aspects of the user’s level of interaction in a communication activity. Zorilla et al. (2010) chose a radar graph as well for presenting student and session profiles. Yet, the authors of both research projects mention evaluated difficulties in the understandability caused by the visualizations (May et al., 2010; Zorilla et al., 2010). These evaluation results conform to findings from our indicator evaluation: visualization types should be well known and simple to interpret. Fancy graphics might cause problems by requiring too much interpretation time, but they have the potential to deliver meaningful information. The most impact has a visualization that may present outliers of data or rather unexpected data about students. Therefore, the design of indicators should be carefully accompanied by user studies and supplemented by adequate interaction options and help facilities as needed.

For providing the right information to the right people, the development of more advanced functionality, like the systematic comparison of selected data of different courses should be pushed on. As a further advancement, the tool could analyze indicator usage and offer a sophisticated rating mechanism to recommend indicators depending on teachers’ data analysis goals. In addition, eLAT could expand its target group to learners for the purpose of self-monitoring.

Current Learning Analytics tools should be interoperable with different learning environments and systems. eLAT uses a neutral data model and has been tested with data of three different learning environments, namely the learning and teaching portal L2P, Moodle and Dynexite, an e-learning exercise tool. We are also working on a mapping between our data model and Contextualized Attention Metadata (CAM) (Schmitz et al., 2009) that will be supported in the next versions of eLAT. In the future, Learning Analytics tools may be linked more often to more open, networked, and lifelong learning environments, driven by new learning theories, such as Connectivism (Siemens,

2005) and LaaN (Chatti, 2010). Hence, integratability is an important factor for the sustainability of Learning Analytics tools. Furthermore, integration of diverse data sources may lead to more pedagogically meaningful indicators because a more holistic picture of the learner could be drawn. To support the implementation of new indicators, in terms of extensibility and reusability, indicator implementations make use of expressions that are performance-optimized database queries. Furthermore, visualizers operate on standardized datasets and are therefore generic to the data inside the report.

As a result of the semi-structured interview, the possibility to filter data was identified as a substantial requirement. However, when implementing the filtering options, we did not entirely meet our goal of providing real-time operations, which are necessary for a smooth user experience. The reason can be found in the data tables, which are not normalized in order to improve performance. To avoid too long calculations, we implemented on-the-fly generation of lookup tables containing pre-calculated intermediate results, thus, increasing the efficiency of filtering. However, a thorough redesign of the database scheme with data warehouse principles in mind is required to minimize the time delay between the capture and use of data.

Regarding data privacy, we chose to pseudonymize student data to protect students' identities and prevent data misuse. According to prior agreements with the universities data protection officer, creating a hash from the username pseudonymized the collected data for clearly assigning and saving personal data. We used a learning environment specific salt with the username, which is a value that makes it difficult to guess the original value, for security reasons. Data was exported from the VLE to eLAT's backend on a weekly basis during the semester. For transparency reasons, students of each course were informed verbally and in writing about the data collection and analysis goals. Of course, the more data is collected, the higher is the potential to recognize individual students during the data analysis. Therefore, we prohibit certain parameter selection options, if less than five students with that particular property are registered for a course. This leads to a trade-off of "data privacy" versus "pedagogical useful indicators". Teachers of small courses with few registered students will not have as much indicator options as large courses with many students. Also, courses with unequally distributed student properties, e.g. with very few female students, cannot be analyzed according to that property, i.e. questions on gender differences cannot not be investigated.

## **Related Work**

Several researchers are trying to solve some of the EDM and Learning Analytics challenges similar to the ones mentioned above.

Krüger et al. (2010) observed that VLEs are usually not designed for data analysis and mining. Therefore they also developed a data model to allow for efficient data mining in a VLE. A prior aim was to automate and alleviate the pre-processing, which is needed to explore, analyze, and mine VLE data. The data model and an implementation are oriented towards the structure of Moodle and have not yet been tested with other VLEs. Our own data model design differs from this work in terms of more detailed logging data concerning content interactions, a property bag mechanism for storing custom data type extensions, the introduction of events as its own data type and an assessment data model. Reasons for these differences are due to our requirements that demand the possibility for integrating custom content and user properties that diverge with different course type instances as well as the need for adding versatile information. An extension model enables instructors to extend the information set of his or her students by using surveys inside the course room, while preserving data privacy, and using the results and user submissions for cross-comparison in certain indicators. Furthermore, our data model was designed to fit with different kinds of VLEs and has been evaluated with L<sup>2</sup>P as well as Moodle.

Pedraza-Perez et al. (2011) have presented a first prototype of a java desktop tool addressed to non-data-mining experts, which supports simple execution of common data mining steps. A simple wizard-based interface helps the user to create, pre-process, visualize, and mine Moodle data files by clicking buttons that are named accordingly, e.g. "create data file" or "pre-process data file", whereby they choose between pre-defined and recommended data mining options. Even though its usage is simple, the interface is still oriented towards users with background knowledge about the data mining process. Teachers, who have no data mining experience, will first have to learn about the process and understand the wizard's vocabulary, such as "classification," "regression," or "association."

Lately, several research projects similar to eLAT are emerging. These projects show that information visualizations are gaining importance for increasing usability of Learning Analytics tools. Mazza and Dimitrova (2007) studied graphical user interfaces for EDM tools. They presented CourseVis, which has been built as an extension of WebCT. Its design is based on the results of a survey, which revealed that instructors need information on social, cognitive, and behavioral aspects about their students when running distant education courses with a VLE. CourseVis uses web log data similar to eLAT and also renders it graphically. An evaluation has shown that the graphical representations of CourseVis helped instructors to quickly and more accurately grasp information of students (Mazza, 2006). As a follow-up, the successful visualization principles of CourseVis have been implemented with a graphical interactive plug-in for student monitoring in Moodle, called GISMO (Mazza and Botturi, 2007). Another example is EDM Vis, an information visualization tool for exploring students' data logs that uses a tree structure to provide an overview of class performance, and interface elements to allow easy navigation and exploration of student behavior (Johnson et al., 2011a). The main differences of CourseVis, GISMO, and EDM Vis compared to eLAT are that the three former mentioned provide visualization in a more static way defined by the system developers. In our approach, teachers are able to dynamically choose by themselves which visual indicators are most helpful to answer their individual data analysis questions. They can change parameters of existing visualization and, hence, create sets of relevant indicators composed on dashboards that fit a specific teaching intervention.

Based on motivations and goals very similar to those of eLAT, two other tools are taking individual user perspectives into account: Graf et al. (2011) recently introduced the Academic Analytics tool (AAT) and García-Saiz and Zorilla Pantaleón (2011) presented the data mining application, called E-learning Web Miner (EIWM). AAT has been primarily developed for learning designers, but can also be used by teachers. The application allows its users to access and analyze students' behavior in Moodle online courses by customizing and performing analytical data base queries. This data analysis may be used as additional evidence for the formative evaluation of courses. Combined with students' evaluations and professor and tutor recommendations for changes, the analytics results are supposed to inform the work of learning designers, who can then adapt, revise and extend resources. Like eLAT, AAT is applicable for different VLEs and aims at being easily extendable, with respect to adding analysis techniques/indicators (Graf et al., 2011). EIWM is supposed to help teachers explore distance students' behavior in a VLE, e.g., Moodle or Blackboard. It offers a set of templates, which can be compared with eLAT's indicator approach. Currently these templates visualize data graphically related to three common questions of teachers, concerning resource usage, students' sessions, and students' profiles in a selected course. The templates contain all needed input attributes for the execution of the related data mining algorithms, which have been implemented on the basis of the Weka toolkit (García-Saiz & Zorilla Pantaleón, 2011). A challenge for Learning Analytics tools like EIWM or eLAT is not only to provide meaningful templates/indicators, but also to decrease the necessity to input parameters by users. As distinguished from EIWM currently eLAT only implements indicators based on simple statistical methods. In the future, we are going to implement further indicators based on data mining methods and evaluate their usefulness for facilitating analysis and improvement of teaching.

It is a complex task for teachers to input data mining parameters. Therefore, parameter-free data mining is a promising approach. Zorilla et al. (2011) compared *yacaree*, a parameter-free association miner with three well-known association rule miners. In their study *yacaree* was well-suited and superior in terms of usefulness to a teacher involved in the evaluation. Research like this will support the further development of usable Learning Analytics and EDM tools that support explorative analytics usage.

## **Conclusion and outlook**

In this paper, we presented the theoretical background, requirements, design, implementation, and evaluation details of eLAT; an exploratory Learning Analytics Toolkit that enables teachers to monitor and analyze their teaching activities. The main goal of eLAT is the improvement of teacher support with graphical analytics, which are useful because they allow extending the audience to "normal" instructors without prior knowledge in data mining techniques. With the help of eLAT, teachers are enabled to explore, reflect and evaluate teaching interventions based on their interests. Key EDM and Learning Analytics requirements, such as usability, interoperability, extensibility, reusability, and data privacy, have been tackled with the development of eLAT.

Currently, eLAT has been primarily developed with the intention to support teachers in their ongoing reflection, evaluation and improvement of their instructional design. In the future, we plan to enhance eLAT in ways that



students can use it as well. Furthermore, eLAT has been successfully tested with data collected from four courses. Future work will include the integration of eLAT in L<sup>2</sup>P and its field-testing with more courses from different disciplines, based on new indicators. We also plan to enhance eLAT with an intelligent recommendation component and evaluate its usefulness.

## Acknowledgements

This project was partly funded by the Excellence Initiative of the Federal and State Governments.

## References

- Alcalá-Fdez, J., Sánchez, L., García, S., del Jesus, M. J., Ventura, S., Garrell, J. M., ... Herrera, F. (2009). KEEL: A software tool to assess evolutionary algorithms to data mining problems. *Soft Computing*, 13(3), 307–318.
- Ali, L., Hatala, M., Gasevic, D., & Jovanovic, J. (2012). A qualitative evaluation of evolution of a learning analytics tool. *Computers & Education*, 58(1), 470–489.
- Altenbernd-Giani, E., Schroeder, U., & Akbari, M. (2009). Programmierungsveranstaltung unter der Lupe. [Getting granular on a programming course]. In A. Schwill & N. Apostolopoulo (Eds.), *Lecture Notes in Informatics: DeLFI 2009 - 7. e-Learning Fachtagung der Gesellschaft für Informatik* (pp. 55–66). Berlin, Germany: Köllen Druck & Verlag.
- Altrichter, H., Posch, P., & Somekh, B. (2005). *Teachers investigate their work – An introduction to the methods of action research*. London: Routledge.
- Bias, R. G. (1994). The pluralistic usability walkthrough: Coordinated empathies. In J. Nielsen, & R. Mack (Eds.), *Usability Inspection Methods* (pp. 63–76). New York, NY: Jolin Wiley & Sons, Inc.
- Federal Data Protection Act, §§ 1-11 (2003). Retrieved May 1, 2012, from [http://www.gesetze-im-internet.de/bdsg\\_1990/](http://www.gesetze-im-internet.de/bdsg_1990/)
- Bültmann, M. (2011). *Design und usability-studie eines learning analytics-werkzeugs*. [Design and usability study of a learning analytics tool]. (Unpublished master thesis). Aachen University, Aachen, Germany.
- Campbell, J. P., & Oblinger, D. (2007). *Academic analytics*. Retrieved from Educause website: <http://connect.educause.edu/library/abstract/AcademicAnalytics/45275>
- Card, S., Mackinlay, J., & Shneiderman, B. (1999). *Readings in information visualization: Using vision to think*. San Diego, CA: Morgan Kaufmann Publishers.
- Card, S. (2003). Information visualization. In A. Sears, & J. A. Jacko (Eds.), *The human-computer interaction handbook: Fundamentals, evolving technologies, and emerging applications*. New York, NY: Lawrence Erlbaum Assoc Inc.
- Chatti, M.A. (2010). *Personalization in technology enhanced learning: A social software perspective*. ( Unpublished doctoral thesis.) Aachen University, Aachen, Germany.
- Chatti, M.A., Dyckhoff, A.L., Schroeder, U., & Thüs, H. (2012). Forschungsfeld learning analytics. *icom – Zeitschrift für interaktive und kooperative Medien*, 1(1), 22–25.
- Directive 95/46/EC (1995). *European parliament and council directive 95/46/EC of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data*. Retrieved May 1, 2012, from [http://europa.eu/legislation\\_summaries/information\\_society/data\\_protection/](http://europa.eu/legislation_summaries/information_society/data_protection/)
- Dix, A. J., Finlay, J., Abowd, G. D., & Beale, R. (2004). *Human computer interaction* (3rd ed.). Harlow, UK: Pearson Education Limited.
- Dyckhoff, A. L. (2010). Towards tools for educational action research: A meta-analysis on applied research questions. In M. B. Nunes, & M. McPherson (Eds.), *Proceedings of the IADIS International Conference E-Learning 2010* (pp. 235–242). Freiburg, Germany: IADIS.
- Dyckhoff, A. L., Zielke, D., Chatti, M. A., & Schroeder, U. (2011). eLAT: An exploratory learning analytics tool for reflection and iterative improvement of technology enhanced learning. In M. Pechenizkiy et al. (Eds.), *Proceedings of the 3rd Conference on Educational Data Mining 2011* (pp. 355–356). Eindhoven, The Netherlands: International EDM Society.
- Elias, T. (2011). *Learning analytics: Definitions, processes and potential*. Retrieved July 29, 2011, from <http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf>
- Few, S. (2006). *Information dashboard design: The effective visual communication of data*. Sebastopol, CA: O'Reilly Media, Inc.

- Fritz, J. (2011). Classroom walls that talk: Using online course activity data of successful students to raise self-awareness of underperforming peers. *The Internet and Higher Education*, 14(2), 89–97.
- García-Saiz, D. & Zorilla Pantaleón, M.E. (2011). E-learning web miner: A data mining application to help instructors Involved in virtual courses. In M. Pechenizkiy et al. (Eds.), *Proceedings of the 3rd Conference on Educational Data Mining 2011* (pp. 323–324). Eindhoven, The Netherlands:
- Gebhardt, M., Rohde, Ph., & Schroeder, U. (2007). Effizientes und nachhaltiges eLearning an der RWTH Aachen durch das integrierte Lehr- und Lernportal L<sup>2</sup>P und das CAMPUS-Informationssystem. [Efficient and sustainable e-learning at RWTH Aachen through the integrated teaching and learning portal L<sup>2</sup>P and the Campus Information System]. In M. Gaedke, R. Borgeest (Eds.), *Proceedings of the Workshop IIM 2007* (pp. 129-144). Karlsruhe, Germany: Universitätsverlag Karlsruhe.
- Glahn, C. (2009). *Contextual support of social engagement and reflection on the Web*. (Unpublished doctoral thesis) Open University of the Netherlands, Heerlen, The Netherlands.
- Graf, S., Ives, C., Rahman, N., & Ferri, A. (2011). AAT - A tool for accessing and analysing students' behaviour data in learning systems. In P. Long et al. (Eds.), *Proceedings of the 1st International Conference on Learning Analytics and Knowledge 2011*. Banff, Alberta, Canada:
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: An update. *ACM SIGKDD Explorations Newsletters*, 11(1), 10–18.
- Hijon, R., & Carlos, R. (2006). E-learning platforms analysis and development of students tracking functionality. In E. Pearson & P. Bohman (Eds.), *Proceedings of the 18th World Conference on Educational Multimedia Hypermedia & Telecommunications* (pp. 2823–2828).
- Johnson, M.W., Eagle, M.J., Joseph, L., & Barnes, T. (2011a). The EDM vis tool. In M. Pechenizkiy et al. (Eds.), *Proceedings of the 3rd Conference on Educational Data Mining 2011* (pp. 349–350). Eindhoven, The Netherlands: International EDM Society.
- Johnson, L., Smith, R., Willis, H., Levine, A., & Haywood, K., (2011b). *The 2011 horizon report*. Austin, Texas: The New Media Consortium, Retrieved July 29, 2011, from <http://net.educause.edu/ir/library/pdf/HR2011.pdf>
- Johnson, L., Adams, S., & Cummins, M. (2012). *The NMC horizon report: 2012 higher education edition*. Austin, Texas: The New Media Consortium, Retrieved May 9, 2012, from <http://net.educause.edu/ir/library/pdf/HR2012.pdf>
- Krüger, A., Merceron, A., & Wolf, B. (2010). A data model to ease analysis and mining of educational data. In M. Pechenizkiy et al. (Eds.), *Proceedings of the 3rd International Conference on Educational Data Mining 2010* (pp. 131–140). Eindhoven, The Netherlands: International EDM Society.
- Lindlof, T. R., & Taylor, B. C. (2002). *Qualitative communication research methods* (2nd ed.). London, UK: Sage Publications.
- Loser, K.-U., & Herrmann, T. (2009). Ansätze zur Entwicklung datenschutzkonformer E-Learning-Plattformen. [Approaches to the development of privacy-compliant e-learning platforms.] *Lecture Notes in Informatics*, 153, 79–90. Retrieved from <http://subs.emis.de/LNI/Proceedings/Proceedings153/P-153.pdf>
- May, M., George, S., & Prévôt, P. (2010). Travis to enhance online tutoring and learning activities: Real time visualization of students tracking data. In M. B. Nunes & M. McPherson (Eds.), *Proceedings of the IADIS International Conference E- Learning 2010* (pp. 57–64). Freiburg, Germany: IADIS.
- Mazza, R. (2006). Evaluating information visualization applications with focus groups: The CourseVis experience. *Proceedings of the 2006 AVI workshop on BEyong time and errors: Novel evaluation methods for information visualization* (pp. 1-6). New York, NY: ACM Press.
- Mazza, R. & Botturi, L. (2007). Monitoring an online course with the GISMO tool: A case study. *Journal of Interactive Learning Research*, 18(2), 251–265.
- Mazza, R., & Dimitrova, V. (2007). CourseVis: A graphical student monitoring tool for supporting instructors in Web-based distance courses. *International Journal of Human-Computer Studies*, 65(2), 125–139.
- Naderer, G. (2007). *Qualitative marktforschung in theorie und praxis: Grundlagen, methoden und anwendungen*. [Qualitative market research in theory and practice: Principles, methods and applications]. Gabler Verlag: Gabler Verlag.
- Nielsen, J. (1992). Finding usability problems through heuristic evaluation. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 373–380). New York, NY: ACM.
- Pedraza-Perez, R., Romero, C., & Ventura, S. (2011). A Java desktop tool for mining Moodle data. In M. Pechenizkiy et al. (Eds.), *Proceedings of the 3rd Conference on Educational Data Mining 2011* (pp. 319–320). Eindhoven, The Netherlands: International EDM Society.
- Polson, P. G., Lewis, C., Rieman, J., & Wharton, C. (1992). Cognitive walkthroughs: A method for theory-based evaluation of user interfaces. *International Journal of Man-Machine Studies*, 36(5), 741–773.

- Rivest, R. (1992). *The MD5 message-digest algorithm*. Retrieved Mai 5, 2012, from <http://tools.ietf.org/html/rfc1321>
- Romero, C., Ventura, S., & García, E. (2007). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, 51(1), 368–384.
- Schmitz, H. C., Scheffel, M., Friedrich, M., Jahn, M., Niemann, K., & Wolpers, M. (2009). CAMera for PLE. *Lecture Notes in Computer Science*, 5794, 507–520.
- Schroeder, U. (2009). Web-based learning - Yes we can! Proceedings of the 8th International Conference on Advances In Web-based Learning 2009. *Lecture Notes in Computer Science*, 5686, 25–33.
- Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*, 2(1). Retrieved July 29, 2011, from [http://www.itdl.org/Journal/Jan\\_05/article01.htm](http://www.itdl.org/Journal/Jan_05/article01.htm)
- Siemens, G. (2010). *What are Learning Analytics?* Retrieved July 29, 2011, from <http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/>
- Zorilla, M. E., García, D., & Álvarez, E. (2010). A decision support system to improve e-learning environments. *Proceedings of the 2010 EDBT/ICDT workshops*. doi:10.1145/1754239.1754252
- Zorilla, M.E., García-Saiz, D., & Balcázar, J.L. (2011). Towards parameter-free data mining: Mining educational data with yacaree. In M. Pechenizkiy et al. (Eds.), *Proceedings of the 3rd Conference on Educational Data Mining 2011* (pp. 363–364). Eindhoven, The Netherlands: International EDM Society.