PANDA
A Platform for Open Learning Analytics

Marcus Hamann[1], Christian Saul[2] and Heinz-Dietrich Wuttke[1]
[1]*Ilmenau University of Technology, Integrated Communication Systems Group, Ilmenau, Germany*
[2]*Fraunhofer Institute for Digital Media Technology,*
Technology-enhanced Learning and Training Group, Ilmenau, Germany

Keywords: Open Learning Analytics, PANDA, xAPI, LRS, Personalization.

Abstract: Learning Analytics (LA) has emerged as a significant area of research in the field of technology-enhanced learning. It automatically analyzes educational data in order to enhance students learning experience and to foster their learning. Open learning analytics (OLA) extends this field in that it integrates data from distributed and heterogeneous sources, serves different stakeholders with very diverse interests and needs, and leverages a variety of statistical, visual and computational tools, methods and methodologies. This paper presents an OLA platform called PANDA that is currently being developed as part of a German research and development project. The platform allows different learning systems to publish data about learners and their contexts, and applies different methods and techniques for information visualization and discovery to analyze the collected data and to detect interesting patterns within.

1 INTRODUCTION

Over the last decade, Learning Analytics (LA) has emerged as a significant area of research in the field of technology-enhanced learning. It has been considered as one of the fastest growing areas of research related to education and technology (Broadfoot et al., 2012). According to the 1st International Conference on Learning Analytics and Knowledge (LAK'11), LA is defined as the "measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs". It was recognized very early that the field of LA offers promising possibilities for education and assessment. LA provides a variety of information that can be used to support monitoring and analysis, prediction and intervention, tutoring and mentoring, adaptation, personalization and recommendation as well as awareness and reflection (Chatti et al., 2014).

Even though LA has been cited in the latest NMC Horizon Report (2014 Higher Education Edition) as an important development with a time-to-adoption horizon of one year and less, a great deal of research is still required in this area. A particularly rich area for future research is Open Learning Analytics (OLA)[1]. OLA extends the field of LA in that it integrates data from distributed and heterogeneous sources, serves different stakeholders with very diverse interests and needs, and supports a variety of statistical, visual and computational tools, methods and methodologies. Following up on this, this paper proposes PANDA, an open learning analytics platform that is currently being developed as part of a German research and development project with the same name.

The remainder of the paper is organized as follows: The second chapter gives an insight into the project and its goals. The third chapter describes our proposed solution. Following up on this, chapter four gives an insight into the collecting and utilization of information about learning experience during using the PANDA environment. Finally, concluding remarks and references complete the paper.

[1]http://solaresearch.org/initiatives/ola/

2 THE PANDA PROJECT

PANDA is a two-year research and development project funded by the German government (Federal Ministry for Economic Affairs and Energy). Partners involved in the project are Ilmenau University of Technology, the Fraunhofer Institute for Digital Media Technology (IDMT) and the company Magh &

Hamann M., Saul C. and Wuttke H..
PANDA - A Platform for Open Learning Analytics.
DOI: 10.5220/0005489804670473
In *Proceedings of the 7th International Conference on Computer Supported Education* (CSEDU-2015), pages 467-473
ISBN: 978-989-758-107-6

Boppert. Each of the partners brings in an own learning system (Moodle, askMe! and Avendoo) with own data and different stakeholders with diverse interests and needs.

The main goal of the project is to develop a centralized platform that allows different learning systems to publish data about learners and their contexts anonymously with the benefit of getting useful information to foster learning in return. The PANDA platform applies different LA methods and techniques to analyze the collected data and to detect interesting patterns within. This includes on one hand methods and techniques that provide statistics in forms of reports and tables, visual representation in forms of charts, maps, etc., but on the other hand methods and techniques that extract information from a data set using methods from artificial intelligence, machine learning, data mining, etc.

The information is then made available to the different systems via well-defined interfaces (APIs). They are not only used to support tutoring and mentoring, but also allows the systems involved to address the individual user more specifically (personalization).

3 PROPOSED SOLUTION

The proposed PANDA platform is depicted in Figure 1. Basically, it consists of two components: a component that stores students' learning experiences and a second component that analyzes this information and applies LA methods and techniques. The (learning) data storage in the proposed platform is done using a Learning Record Store (LRS). It allows storing and retrieving learning experiences using the xAPI[2]. Beside the actual platform, there are different systems connected to it namely Avendoo[3], Moodle[4], askMe![5] and RemoteLab. From students point of view, they form a unified environment for learning (Avendoo, Moodle) and knowledge testing (askMe!, Remotelab) by providing course materials and interactive content objects (ICOs) like simulations and remote experiments.

The systems benefit from the PANDA platform in that it serves as connection point for the communication of learning experience. The modular design of the platform together with the usage of well-defined interfaces (xAPI) increases the compatibility with other systems and supports its prompt and widespread adoption. The following subsections describe the LRS and the LA component in more detail.

[2]http://www.adlnet.gov/tla/experience-api/

[3]http://www.avendoo.de/

[4]https://moodle.org/

[5]http://www.idmt.fraunhofer.de/askme

3.1 Learning Record Store

The Learning Record Store plays a significant role in the PANDA platform because it stores information about students' learning activities performed in different systems. It is a type of data repository designed to store learning experiences in so called statements generated by xAPI-compliant systems or reporting tools. The interface to the LRS is realized by the xAPI specification. xAPI is a communication mechanism based on activity streams that facilitates and integrates all types of learning and training and makes it possible to collect data about the wide range of experiences a person has (online and offline). The experiences are expressed as statements. Statements are the core of xAPI. In general, a statement has the following structure:

<actor (student)><verb><object>with <result>in <context>

The actor is represented by the student identified only by an ID. Verbs can be information about the usage of the system like "viewed", "passed" or "failed". The object, finally, addresses the element which is handled by the actor e.g. "Moodle chapter X", "question Y" or "ICO Z". These statements are human readable, meaningful if printed on an interface and also machine readable. By means of this structure, we can easily classify users by different verbs or objects in order to provide a high degree of adaptivity and personalization for our learning management system.

In the context of the PANDA project, the LRS stores the whole amount of information about the activities of the user such as the number of successfully finished tests or the amount of viewed learning material. By splitting the information into small information units these data are hidden for unauthorized persons because of their missing context.

3.2 Learning Analytics

The LA component plays a significant role in the PANDA platform because it analyzes data stored in the LRS and tries to discover patterns within. Then, these findings can be used by other systems to support tutoring, mentoring, adaptation, personalization, etc. Mainly four groups of techniques have received particular attention in the LA literature in the last couple of years (Chatti et al., 2014):

1. *Statistics*: This group includes all methods and

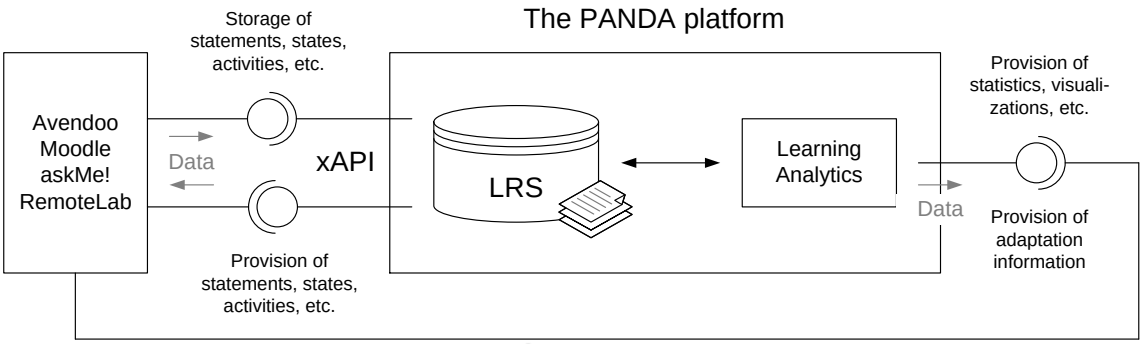

Figure 1: Overview of the PANDA platform.

techniques that provide statistics in forms of reports and tables.

2. *Information visualization*: This group includes all methods and techniques that provide visual representation in forms of charts, scatter-plots, maps, etc.
3. *Data mining*: This group includes all methods and techniques that discover useful patterns or knowledge from data sources.
4. *Social network analysis*: This group includes all methods and techniques that analyze the relationships between individuals or organizations.

The methods and techniques, the LA component applies, can mainly be assigned to the first three groups. For statistics and information visualization, the LA component uses the data stored in the LRS and provides them in a clear and understandable format. These visualizations (see Figure 2) can then be integrated on the own website or learning platform using e.g. a communication technique like JSONP[6]. Statistics and visualizations can be both *descriptive* and *predictive*. Descriptive means that they are in accordance with the facts and without any interpretation or assessment. Exemplary: Student A reached 5 out of 10 points. In contrast, predictive methods try to make statements about facts or the occurrence of events with a given probability. Exemplary: It is very likely that student B is not going to pass the final test successfully. In order to predict these future events, prediction models (mainly classification methods (Ming and Ming, 2012)) are going to be used. In addition, clustering algorithms are intended to be used to group students and student actions as well as sequential pattern mining to identify whether an event (or any observed construct) was the cause of another event (or observed construct) (Baker and Siemens, 2011).

[6]http://www.json-p.org/

The application of these methods using the data from the LRS will be used to improve both the learning and teaching methods. On the one hand we can use the information to support the user (e.g., give immediate feedback to wrong answers, links to relevant lectures, etc.), while on the other hand the teacher profits by the learning process that each user has passed. In that way we can accompany the user in the learning process in order to recognize user weaknesses and to motivate the user. The benefit for the teacher is mainly feedback of users weaknesses and learning material.

3.3 Data Privacy

A big concern of learning analytics is data privacy. However, it might not be in the interest of the student to disclose collected and stored personally identifiable information or other sensitive information. For that reason, the PANDA platform uses different ap-

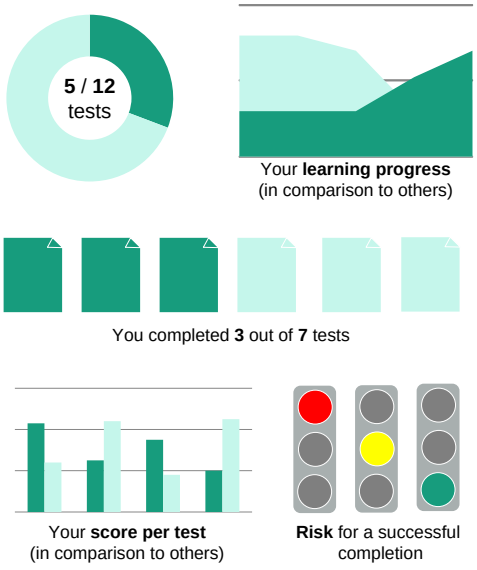

Figure 2: PANDA visualizations.

proaches to achieve a high level of anonymity. The most important precaution for protection of data privacy is the LRS. With the help of the LRS, we collect data anonymously and centralized outside of the learning environments (Avendoo, Moodle, askMe!, RemoteLab). In that way, each user is represented by an ID on an external server that provides the LRS.

Moreover, the learning environments are not able to exchange data about the user. Rather, each system collects special information about each user and transfer it to the LRS independently.

In addition to the above mentioned aspects the central data storage offers another advantage. From the anonymous data we can build user groups which can be evaluated separately. This user group specific evaluation, for instance according to skill level of each user, can be used to reach a high level of adaptivity and personalization of the whole system. The big advantage is that we do not have to poll each subsystem to receive all relevant information about one user or user group but that we can obtain all pertinent data from the LRS.

4 LEARNING EXPERIENCE - COLLECTING AND UTILIZATION DATA

In order to make the PANDA platform to work, in particular the LA component, information about learners interactions (i.e., learning experiences) need to be gathered from different systems. In the context of this project, the systems that mainly provide information about learning experiences are the Moodle learning platform, the remote laboratory hosted at the Ilmenau University of Technology and the Avendoo learning platform. In contrast, the system that take advantage of the PANDA platform (and its data) are mainly the askMe! system. In the following, these systems are described in more detail.

4.1 Moodle

Moodle is an open source learning platform originally developed by Martin Dougiamas. Today, the Moodle project is supported by a big community and is growing continuously. It combines a lot of modules such as ”books”, ”forums”, ”quizzes” and many more in order to support learning activities. In addition, it is highly adaptable in terms of including own plugins.

These facts make Moodle to a valuable learning platform and an important part of the PANDA project. In the context of the PANDA project, Moodle

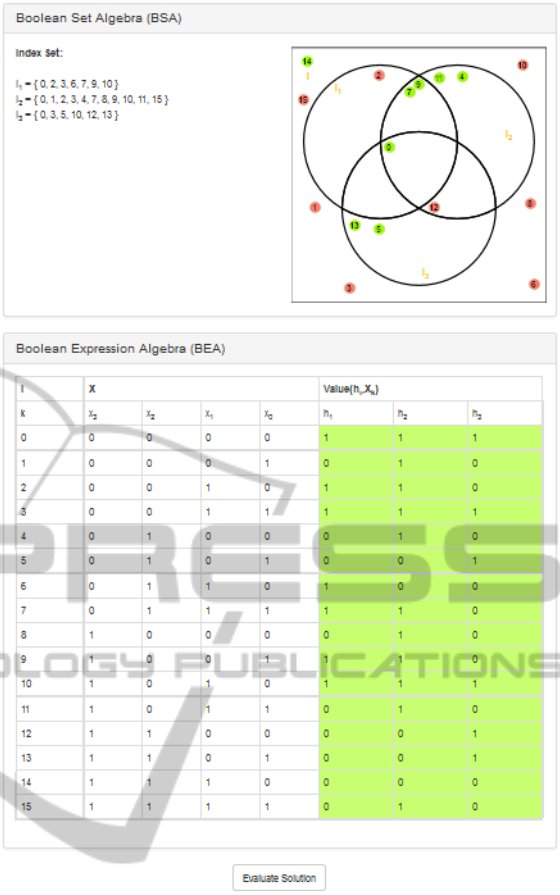

Figure 3: Moodle system.

is mainly used to provide learning content for the user. It allows structuring lectures and further information very generic. We are able to include external links as well as ICOs and the user can decide for himself which content is relevant.

Most advantageous for the PANDA project is the customizability of Moodle. Therefore, we have established an interface to the LRS. Thus, we can record the learning process of each user. After recording of data we can analyze users in order to provide adaptation information which in turn can be used by the Moodle system to improve the learning process. Thus, we can e.g. hide all other parts of the learning path to assist the user with a specific problem.

In summary, that implies, if a user acquired knowledge about one part of a lecture with the help of Moodle content and ICOs and passed a test about the knowledge, this information is recorded in the LRS without deeper knowledge about the user.

One excerpt of a course material inside the Moodle system, more precisely, an ICO is depicted in Figure 3. With the help of this ICO the student should

Figure 4: Remote Lab of Ilmenau University of Technology.

get deeper knowledge about Boolean set algebra. For that a random Venn diagram, representing sets of input combinations, will be presented to the student and he or she has to find a correct truth table matching the Venn diagram. During solving this problem the student can compare his or her solution with the correct one. This information will be sent to the LRS in order to support the student with the help of Learning Analytics.

4.2 Remote Laboratory

In general, a remote laboratory uses telecommunications to remotely conduct real experiments at a physical location whilst the scientist is utilizing this technology from a separate location. With the Remote Lab, called GOLDi (Grid of Online Lab Devices Ilmenau), the University of Technology wants to offer to students a working environment that is as close as possible to a real world laboratory.

The remote laboratory of Ilmenau University of Technology (Henke et al., 2012; Henke et al., 2013; Henke et al., 2014) allows students to design, verify and implement digital circuits and control systems. The lab consists of different programmable control units (embedded systems or FPGAs) and physical systems (e.g., an elevator, a conveyor or a high rack warehouse) that provide real-time experiments with real hardware equipment or simulations to students. All of these experiments empower students to solve complex design tasks, makes the learning environment very powerful and allows a very effective learning process.

Under real laboratory conditions disturbances can appear and lead to failures of the control algorithm that cannot be detected under virtual lab conditions.

While the student is interacting with the Remote Lab, it is able to generate fault messages in case of faulty inputs to the LRS which can be used by methods of Learning Analytics in order to provide improvements of the learning process. Figure 4 shows the graphical user interface of the Remote Lab in more detail.

4.3 askMe!

askMe! is a web-based e-assessment system being developed at the Fraunhofer IDMT that covers the whole life-cycle of e-assessments starting from creating questions, presenting them to the students up to preparing the results and presenting them to teachers, tutors, etc. The questions and tests can consider individual aspects so that e-assessments and their feedback can perfectly be tailored to students or groups of students (Saul and Wuttke, 2013). Moreover, the author of the adaptive tests is not limited to traditional question types such as multiple-choice, but can use Interactive Content Objects (ICOs) to create sophisticated (interactive) e-assessments (Saul and Wuttke, 2012). The latter aspect takes into account the as-

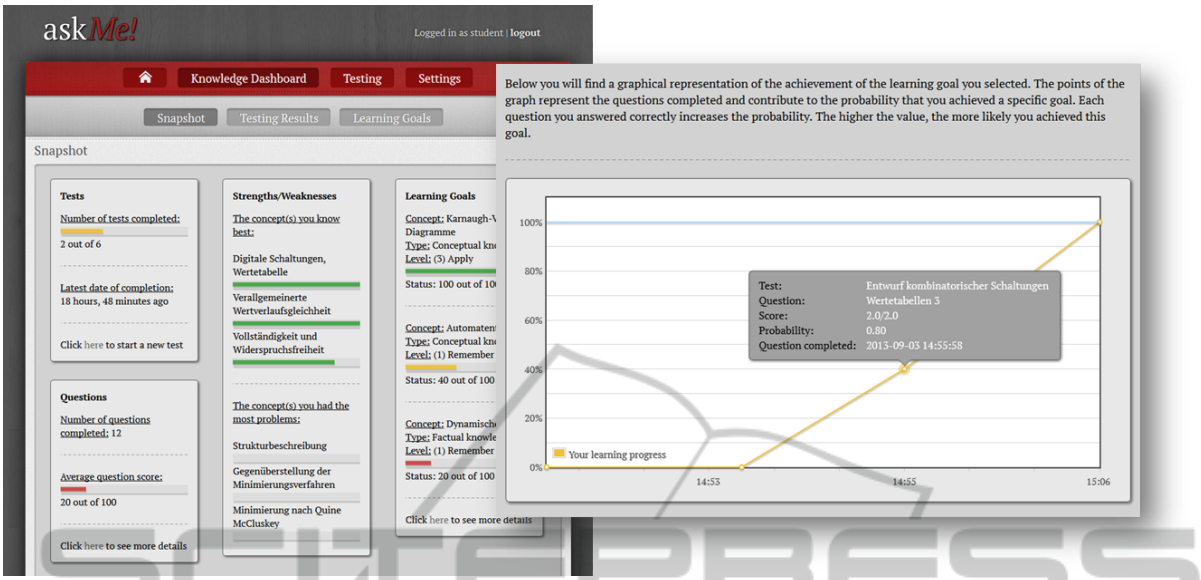

Figure 5: Knowledge dashboard of the askMe! system.

sumption that learning is the result of interaction and more specifically, the result of engagement with the subject matter. In order to deal with the different ICOs located elsewhere in the Web, a communication mechanism based on the xAPI specification has been developed.

In the context of the PANDA project, the LA aspect of the askMe! system is being further developed. Currently, the system focuses on information visualizations that are provided to both students and authors (Saul and Wuttke, 2014). When a student has completed a test, he/she will not be confronted with an abstract score, but will get detailed feedback on his/her strengths and weaknesses, which allows him/her to efficiently address specific deficits afterward. This information is presented in his/her knowledge dashboard (cf. Figure 5). This component not only provides students with a tabular and graphical overview of his/her testing results, but also with a detailed overview about his/her knowledge level according to the topics addressed by the respective test. In this way, the askMe! system is able to increase students self-awareness significantly. However, the presentation of statistics in askMe! is not limited to support students, but is also provided to authors. The information for this user group is presented in user and test statistics. This component presents an overview of students results in (adaptive or non-adaptive) tests as well as their individual learning progress.

These information visualizations in forms of charts and tables are now being expanded and supported with the help of PANDA. This includes the use of prediction methods in order to predict, which students are struggling with the content. Tutors and teachers can use this information (e.g., visualized as traffic light) to intervene either online or face-to-face. Other methods including heat maps, bar charts and line graphs show students' activities and testing results over time and support teachers in their decision-making. The basis of all these methods are students' interactions with the askMe! system, which are communicated to the LRS and processed by LA.

4.4 Avendoo

Avendoo is a learning platform developed by the company Magh & Boppert. In the frame of the PANDA project, Avendoo will not only provide information about students' learning activities, but will also use this information (together with information obtained from other systems) to adapt the learning courses (adaptive navigation support, (Brusilovsky, 2001)) and the system accordingly. For this, the system will make use of the LA component and the information it provides to implement adaptive link hiding (Brusilovsky, 2004).

5 CONCLUSIONS AND FUTURE WORK

This paper has proposed an OLA platform called PANDA. It supports both learning experience tracing and learning experience utilization. The learning experience tracing is realized by the Moodle learn-

ing platform and the Remote Laboratory. The objective here is to gather information about learners interactions such as learning experience. This part of the PANDA platform serves as input for the following learning experience utilization in order to process all relevant data for information visualization and discovery to find patterns in the learning process. The learning experience utilization is mainly realized by the askMe! system and the Avendoo learning platform.

Future work includes the (full) implementation of the LA component. This allows us to obtain valuable information from the PANDA platform and to use them for adaptation, personalization or individual support of the learner (e.g., to give better hints). Moreover, in order to handle the huge amount of data due to the number of users, an authoring tool is currently being developed. In addition, to cover also the higher order thinking skills while learning we intend to integrate the Remote Laboratory into the learning environment. Furthermore, a major challenge is to test our system with other learning content as well as other target groups.

Finally, it can be stated that the PANDA platform reflects the fact that learning takes place everywhere: in traditional education settings (e.g., LMS) as well as in more open-ended and less formal learning settings (e.g., PLEs, MOOCs). It has the potential to deal with the challenges in increasingly complex and fast-changing learning environments.

ACKNOWLEDGEMENTS

The research presented in this paper has been financially supported by the German Federal Ministry for Economic Affairs and Energy (BMWi) within the project ”Entwicklung einer Software für die Personalisierung von Lernprozessen durch Adaptivität, Nutzermodellierung und Datenanalyse der Lerner-Aktionen (PANDA)” under contract no. KF2329704KM3 and KF2250116KM3.

REFERENCES

Baker, R. and Siemens, G. (2011). Educational Data Mining and Learning Analytics. In *Cambridge Handbook of the Learning Sciences*, page 19. Cambridge University Press, Cambridge.

Broadfoot, P., Timmis, S., Payton, S., Oldfield, A., and Sutherland, R. (2012). Rethinking Assessment - Learning analytics and technology enhanced assessment (TEA). Technical report, University of Bristol.

Brusilovsky, P. (2001). Adaptive Hypermedia. *User Modeling and User-Adapted Interaction*, 11(1):87–110.

Brusilovsky, P. (2004). Adaptive Navigation Support: From Adaptive Hypermedia to the Adaptive Web and Beyond. *PsychNology Journal*, 2(1):7–23.

Chatti, A., Lukarov, V., Thüs, H., Muslim, A., Yousef, F., Wahid, U., Greven, C., Chakrabarti, A., and Schroeder, U. (2014). Learning Analytics: Challenges and Future Research Directions. *eleed*, 10(1).

Henke, K., Ostendorff, S., Vogel, S., and Wuttke, H.-D. (2012). A Grid Concept for Reliable, Flexible and Robust Remote Engineering Laboratories. *International Journal of Online Engineering (iJOE)*, 8:42–49.

Henke, K., Ostendorff, S., Wuttke, H.-D., Vietzke, T., and Lutze, C. (2013). Fields of Applications for Hybrid Online Labs. *International Journal of Online Engineering (iJOE)*, 9:20–30.

Henke, K., Wuttke, H.-D., Vietzke, T., and Ostendorff, S. (2014). Using Interactive Hybrid Online Labs for Rapid Prototyping of Digital Systems. *11th International Conference on Remote Engineering and Virtual Instrumentation (REV)*, 11:57–63.

Ming, N. and Ming, V. (2012). Predicting student outcomes from unstructured data. *CEUR Workshop Proceedings*, 872:11–16.

Saul, C. and Wuttke, H.-D. (2012). Towards a high-level Integration of interactive Tools with E-Assessments. In *Proceedings of the 12th IEEE International Conference on Advanced Learning Technologies (ICALT 2012)*, pages 597–598, Rome.

Saul, C. and Wuttke, H.-D. (2013). E-Assessment meets Personalization. In *Proceedings of the IEEE Global Engineering Education Conference (EDUCON 2013)*, pages 200–206, Berlin.

Saul, C. and Wuttke, H.-D. (2014). Turning Learners into effective better Learners: The Use of the askMe! System for Learning Analytics. In *Extended Proceedings of the 22th International Conference on User Modeling, Adaptation and Personalization (UMAP 2014)*, pages 57–60, Aalborg. CEUR-WS.