LEARNING ANALYTICS FOR HIGHER EDUCATION — PERSPECTIVES AND CHALLENGES

Introduction. The Horizon Report 2011 [1] identifies learning analytics as an emerging technology, which most likely will have an impact on universities within the next five years. The authors provide a first definition of learning analytics as the interpretation of student data. Data stem from student activities, which can be recorded by a learning management system, but also from extracurricular activities or online social interaction. They define the goal of learning analytics as enabling teachers "to tailor educational opportunities to each student’s level of need and ability".

In a wider sense, learning analytics may also assess curricula or programs. "It might also be used
by students themselves, creating opportunities for holistic synthesis across both formal and informal learning activities”.

A shorter definition of learning analytics can be found in the call for papers of the 1st International Conference on Learning Analytics and Knowledge [2]: “Learning analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs.”

In [3], three factors for the development of learning analytics are listed: big data, online learning and political concerns. Big data denote huge sets of, generally unstructured, data, which typically originate from social networks, together with analytic methods (“social media analysis”). With the success of learning management systems like moodle, universities deal with increasing sets of data about learning activities. Finally, in many countries, especially in the USA, there is a strong demand for learning analytics, to improve learning opportunities and educational results.

1. Promises and perspectives. At universities, there are three groups of people interested in learning analytics (stakeholders). The first group is interested in questions like: who are the students, how many students are presently active (drop out), how is their learning behaviour, how efficient are special programs in terms of student performance, what is the utilization of learning resources, what is the correlation between formal learning, informal learning, and student performance? With this “administrative view” on learning analytics, programs or curricula could be compared, and allocation of resources could be improved.

The second group are teachers: for them, learning analytics is a method to monitor ongoing learning, and to evaluate didactic concepts. Learning analytics supports personalization of teaching: groups of students with specific learning needs can be identified, problems can be identified, and learning offerings become more suitable for the specific needs of a student.

And finally, students themselves can get feedback from a learning analytics system.

2. Limitations and challenges. A major limitation for learning analytics is the lack of (good) data. As many courses still take place in the classroom, there is no (formal) online learning. If online learning is not mandatory, it is regarded as a supplement for students with learning problems; good and/or busy students don't use it. Therefore, data obtained from a learning management platform is incomplete and not very informative.

Another problem is a missing didactic concept for blended learning combining face-to-face teaching and online learning. Without a concept, including different forms of interaction and different forms of assessment, data in a learning management system consist basically of downloading learning resources (clicks) and of reading news. This data has low entropy, good results from learning analytics cannot be expected.

Traditionally, personalized interactions and user modeling have significant implications on data privacy [4]. Personal information about a user are collected and analyzed, which might not be in the interest of the user. Recently, collection of user data in social networks, and possible violation of data-privacy legislation, have been frequently published.

Many universities have strict data-privacy regulations, which ban the use of personal data, and thus limit learning analytics. This could be a policy for learning management systems to delete logging data containing personal data after a specified period of time, or the requirement to anonymize all student data before performing learning analytics. From an administrative view, this might be acceptable, but it deprives teachers from personalizing teaching and learning, and students cannot get feedback from learning analytics.

3. The learning-analytics tool LeMo. ”LeMo: Monitoring the Learning Process” is the name of a joint research project by three universities of applied sciences in Berlin, namely Beuth-Hochschule für Technik (BHT), Hochschule für Technik und Wirtschaft (HTW) and Hochschule für Wirtschaft und Recht (HWR). The tool developed in this project aims at supporting teachers, content providers and researchers in learning analytics. The LeMo tool extracts activity data from learning management systems like Moodle or Clix, as well as from learning platforms like the free encyclopedia
ChemgaPedia. LeMo offers a variety of filter options and allows for dynamic visualization, fostering learning analytics and visual analytics.

There are various approaches to data privacy in learning-analytics tools, for an example see [5]. To facilitate learning analytics at universities with strict data-privacy regulations, the LeMo prototype filters out almost all personal information, which could be obtained from a learning management system, e.g. name or student identification. The only exception is gender information, which is needed to do research on gender-specific issues. Filtering occurs during the ETL (extraction translation load) process, so names of teachers or students are not stored in the LeMo database. Persons are uniquely represented in LeMo by a hash value based on identification data obtained from the learning management system. Replacing names and IDs by a unique hash value can be seen as a process of de-identification, or pseudonymization. To achieve more anonymity, the k-anonymity method can be applied [6]. This implies, in the case of LeMo, the exclusion of courses from analysis, if persons could be re-identified applying appropriate filters.

Using pseudonyms for teachers (derived from login name) and students (derived from student identification) has various benefits. If the LeMo application is configured to use the same authentication scheme as the corresponding learning management system, LeMo can obtain a teacher's pseudonym from the login name. Thus, a teacher can be granted access to his/her own courses. In the case of pseudonyms derived from student identification, a teacher may "personalize" the outcomes of learning analytics, leading to individual learning offerings.

The LeMo project started out with a long requirement list (80 questions) from partner institutions and companies. From that list, 27 indicators were derived, which provide data analysis to help answering some of the questions. Presently, the LeMo prototype implements the analyses "Activities/Time", "Activities/Learning-Objects", "Frequent Paths", "Activity-Graph" and "Accumulated Activities/Weekday".

Through an activity graph, navigation of students within a course is visualized. Learning objects are represented as nodes, navigation steps as directed edges. Different colors indicate different types of learning objects. The width and color of an edge gives hints on the number of navigation steps between two learning objects. Activity graphs are interactive visualizations – a user can drag a learning object to see more details, and the graph automatically adapts to the new position of the node.

**Conclusions.** Learning analytics will be a valuable, and indispensable, tool for quality teaching and learning at universities, obtaining best results on courses with a well-defined blended-learning concept. Thus, learning analytics will be part of any quality-teaching-and-learning project at universities, aiming at better student performance, more and better learning opportunities, and a more effective.

Privacy and ethics issues are crucial for the acceptance, and success, of learning analytics. Personalized data provide insight in the learning state of a student, and lead to individual support and learning offerings. But personal data can also be misused, which further limits the application of learning-analytics tools.

**References**


Reviewer Ph. D., Prof. Odessa nat. polytechnic univ. Brovkov V.G.

Received December 24, 2012.