

HOW TO IMPLEMENT DATA-DRIVEN ANALYSIS OF EDUCATION: OPPORTUNITIES AND CHALLENGES

Olga Viberg, oviberg@kth.se

The topic of digitalisation of education attracts interest of the research community worldwide owing to unprecedented capabilities provided by technology to capture digital traces of today's students who, being 'digital natives', are active in technology-rich learning environments. The aim of this report is to present solutions on the topic that can be applicable in a Swedish context and raise an awareness of potential barriers and challenges. The solutions emerge as best practices, or as examples from recent research literature.

Status of research

Learning analytics is defined as "*the measurement, collections, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs*" (Long & Siemens, 2011, p.34).

While there are a number of software applications available, they are all at early stages of development and there might not be any "off-the-shelf" solution for applying learning analytics to a given context. This is for some fundamental reasons:

1. Most factors of interest regarding learning and education are not possible to define strictly.
2. Because pedagogical, technical and organisational aspects of learning are complex, they must be carefully interpreted within the used context.
3. The pedagogical, technical and organisational aspects of learning are more difficult to assess when examining teachers compared to giving teachers better tools for following the work of students.

This means that anybody who wants to apply learning analytics must clearly define *what* they want to achieve. Universal solutions that can be applied 'as-is' in every socio-cultural context do not exist. Nevertheless, it is possible to provide a map of the field in its current state and point to what the presumptive user needs to do in order to understand the output from systems, and to use the output adequately and ethically.

Best practices use cases

We have identified three areas of use of learning analytics where the research shows good results.

Identifying at-risk students

Predicting students' success in their studies is one of the dominating research focuses (predominantly in higher education). A considerable number of early warning systems aimed at teachers and students have been designed and implemented across institutions and countries (Arnold & Pistilli, 2012; Jayaprakash et al, 2014). The predictive models include measures such as students' demographic characteristics, their previous academic history, interaction with learning management systems, and course grades (performance on the course to date). One of the challenges to be considered when developing such a system relates to the institutions' capacity to extract data from different systems efficiently and timely. The overall advantage of early warning systems is to make early interventions that would lead to a higher academic success among students.

Learning strategies

Supporting the development of students' learning strategies with analytics has been studied. The literature suggests that such strategies and skills (e.g. time management and effort regulation) are highly valuable and may predict academic performance, especially in online learning environments (Broadbent, 2017). The developed tools have been aimed at both students and teachers and include mobile applications, with or without gamifications elements. They usually intended to support one type of strategy at a time. Typical measures include time-log records during the course, final grades, as well as clickstream data related to the performed learning activities. Overall, the use of such tools has shown improvements in the targeted strategies (Kwong et al, 2017; Tabuenca et al, 2015).

Learning design

To understand whether there is a mismatch between how students study in reality and how instructors designed a course, researchers examined the interplay between learning analytics and learning design. One frequent focus is on student engagement with a chosen learning management system, i.e., how engagement varies across different levels of performance. Typical measures comprise students' timing of engagement in relation to the learning design patterns (e.g. catching up patterns). By comparing and contrasting the assumptions in the learning design made by instructor with actual student behaviors, learning analytics can act as a reflective source and provide actionable feedback for how a course should be redesigned (Rodriguez-Triana et al, 2015).

A to-do list

For those who aim at performing learning analytics, we now present a set of guidelines to help structuring the work.

1. *Start defining a problem to be solved. Consider the following.*

- Do you want to support the learners, the teacher, or some external or internal evaluator, or someone else?
- What learning activities do you address?
- What is the motivation for the analysis?
- Are there any important ethical considerations (e.g., privacy)?

2. *Investigate what type of data you need and arrange for collecting them.*

One of the key challenges that learning analytics researchers and practitioners are facing to relates to the data collection procedures. Frequently used learning management systems are not open source and the data gathered is owned by the system provider. The system might present only selected amounts of data on the individual or group. In other words, it is not in a format or of completeness that can be used for further analysis. Presently, researchers conducted learning analytics studies use interaction data from:

- Learning management systems
- Massive open online courses
- Specifically developed mobile applications.

3. *Choose the type of output*

Different types of outputs (e.g., visualisations, tables) provide different possibilities for the analysis and relevant interpretations of the results, as well a different degree of accuracy.

- *Visualisation and learning analytics dashboard*

A learning dashboard aggregates and visualizes different indicators about learners, learning processes, and learning contexts (Schwendimann et al, 2016). Typical measures include learning resources used, time-on-task estimation and test results. Dashboards mirror the students' behavior back to students and teachers; they hence promote a dialogue between a student and an instructor, and increase awareness among teachers about learners' progress. A dashboard may also increase the awareness among students of their own progress in relation to effort spent and their time management.

- *Data-supported feedback*

The provision of timely and relevant feedback is closely associated with learning analytics. The relevant examples include gamified learning applications to provide formative feedback to the students. It also offers an overview for teachers about student achievements, and tools that support learners' peer- and self-assessments. Timely feedback provides teachers with simple, real-time insights of how students learn. Typical measures contain students' scores/grades and levels of competence from the summative assessment in relation to the continuously provided (formative) feedback. Gamification elements in the offered tools were found to increase motivation, especially among boys and men.

4. *Do a critical analysis*

When analysing the collected data, consider the following questions:

- What sources of error are there in the collected data?
- What significance might the errors in the data have on the results of the analysis?
- What is the uncertainty in the results? Are the data sufficient for trusting them?

One common source of error relates to the fact that only parts of a student's learning activities are performed in a chosen system, while the rest is done manually or in some other system from which the user data cannot be collected easily. This can give a skewed picture of the students' learning.

Summary

There are best practices of learning analytics where research shows that the results can lead to improvement in teaching and student learning. We encourage educational providers from schools, high schools, adult education and higher education to consider the use of learning analytics for improving their educational processes. To assist in this work, we provided a short to-do list as a guide.

References

- Arnold, K. E., & Pistilli, M. D. (2012, April). Course signals at Purdue: Using learning analytics to increase student success. In Proceedings of *the Second International Conference on Learning Analytics and Knowledge* (pp. 267-270). ACM.
- Broadbent, J. (2017). Comparing online and blended learner's self-regulated learning strategies and academic performance. *Internet and Higher Education*, 33, 24-32.
- Jayaprakash, S., Moody, E., Lauria, E., Regan, J., & Baron, J. (2014). Early alert of academic at-risk students: An Open Source Analytics Initiative. *Journal of Learning Analytics*, 1(1), 6-47.
- Kwong, T., Wong, E., & Yue, K. (2017). Bringing abstract academic integrity and ethical concepts into real-life situations. *Technology, Knowledge and Learning*, 22(3), 353-368.
- Long, P., & Siemens, G. (2011). Penetrating the fog: Analytics in learning and education. *Educause Review*, 46(5), 31-40.
- Rodríguez-Triana, M., Martínez-Monés, A., Asensio-Pérez, J., & Dimitriadis, Y. (2015). Scripting and monitoring meet each other: Aligning learning analytics and learning design to support teachers in orchestrating CSCL situations. *British Journal of Educational Technology*, 46(2), 330-343.
- Schwendimann, B., Rodríguez-Triana, M., Vozniuk, A., Prieto, L., Boroujeni, M., Holzer, A., Gillet, D., & Dillenbourg, P. (2016). Understanding learning at a glance: An overview of learning dashboard studies. In *Proceeding of the Six Conference on Learning Analytics & Knowledge Management*, Edinburgh, United Kingdom, April 25-26, 2016. DOI:[10.1145/2883851.2883930](https://doi.org/10.1145/2883851.2883930)
- Tabuenca, B., Kalz, M., Drachsler, H., & Specht, M. (2015). Time will tell: The role of mobile learning analytics in self-regulated learning. *Computers & Education*, 89, 53-74.