

Measurement of selected aspects of student behavior in online courses

Michal Černý

Abstract

The paper analyzes the possibilities of using learning analytics outside of LSM, for example in educational environments that are not associated with common large systems such as Moodle or Blackboard, but which can only be analyzed through web analytics tools. This topic has been marginalized in the discourse by the educational public for a long time. My research and methodological attention show that the data that can be obtained from these environments about the learning and information behaviour of students can be of interest to the design of the whole educational environment and can help with significant changes in the design of courses, measuring the attractiveness or clarity of individual pages etc. In this paper, we focus primarily on the combination of Google Analytics and Smartlook or Hotjar – tools for creating heatmaps.

Keywords: learning analytics, Google Analytics, Hotjar, heatmaps, web analytics, data driven education, online courses, LMS, learning environment.

Měření vybraných aspektů chování studentů v online kurzech

Abstrakt

Článek analyzuje možnosti využití analýzy učení mimo LSM, tedy v edukačních prostředích, která nejsou spojená s běžným robustními systémy jako je Moodle nebo Blackboard, ale které je možné analyzovat jen skrze nástroje webové analytiky. Toto téma bylo dlouhou dobu v diskursu pedagogickou veřejností marginalizováno. Naše výzkumná a metodologická pozornost ale ukazuje, že data, která lze z těchto prostředí získat o studijním a informačním chování studentů mohou být pro design celého edukačního prostředí zajímavá a mohou pomoci s významnými změnami v oblasti návrhu kurzů, měření atraktivnosti či srozumitelnosti jednotlivých stránek atp. V tomto článku se soustředíme především na kombinaci nástrojů Google Analytics a Smartlook respektive Hotjar pro tvorbu heatmap.

Klíčová slova: learning analytics, Google Analytics, Hotjar, heatmapy, webová analytika, daty řízené učení, online kurzy, LMS, vzdělávací prostředí.

DOI: 10.5507/epd.2019.015

Online courses represent a specific form of distance learning, that play a significant role in the tertiary education. It is a form associated with many myths, which on one hand prevents similar educational activities to gain significant role in the curriculum, but also ties up the creator with certain patterns of thoughts with which to cope. This situation, on-line courses, is much more intense than, for example, classical lectures or exercises where the level of difficulty and quality of content is tightened. This certainly does not mean that – especially regarding how the budgets of the faculty workplaces (i.e. with a distinctive person-credit parameter) are designed – that all online courses would be an ideal educational tool.

We would like to focus on some of the myths or mind schemes that appear in relation to online courses and which are also related to the possibilities of measuring students' behaviour. While in normal contact education, the ability to measure students' behaviour is very limited (and generally less used), in the case of online learning the tools are much broader.

The first myth is the idea that students do not really study online courses, or that they use them basically randomly. They just click through the organizational instructions, then move on to the test or task, and the study time ends. It is an idea of students like „minimax“ (according to Ivo Možný), who will achieve the maximum effect with minimal effort.

The second myth may be inadequate content quality – students will not learn what they would learn in a lecture or book in an online course. Here, it is true that the design of the online environment does not want to transfer the script to the online environment in 1 : 1 format, but it certainly does not mean lower quality. Quality is something that should be measurable, perhaps the easiest way is to compare the results of pre-tests and post-tests. The question is to what extent both approaches are effectively comparable and to what extent the form is associated with content so much that it is not a valid measurement.

Thirdly, there is a question of the social dimension of online education – learning is a social process and digital technology can lead to psychological and social isolation. This argument may be valid but at the same time we can say that a great number of online courses are based on intense interaction between individual students – from team projects, through peer reviews, webinars, collaborations in discussion forums, etc. Didactics of online courses is didactics of social interactions that can be quantified and measured again (Ferguson, 2012; Ferguson & Shum, 2012; Siemens & Baker, 2012).

If we are talking about myths, we do not necessarily mean misconceptions, but rather beliefs not based on rational arguments and quality analytical procedures. In this article, we cannot systematically discuss individual myths and their variations because of their limited scope, but we will at least attempt to indicate how web analytics tools in online courses (Long & Siemens, 2011; Siemens, 2013) can be used as one of the sources of information on how students learn online. The knowledge of this behaviour is then a prerequisite for designing new educational activities, interventions, or choosing the appropriate educational policy of the institution.

The article focuses primarily on higher education, i.e. a specific group of students. Data obtained through web analytics as well as other possible approaches can be used as one of the interesting tools for understanding learning behaviour (Morris 2005), as well as subject surveys or other forms of feedback, whether qualitative or quantitative. We would like to emphasize, in terms of methodology, that the use of the possibilities of web analytics is not a decisive or valid source of information on how students work in an online environment, but it can serve as a basis for further research or assist in the interpretation and understanding of other data.

What can be learned about student behaviour from Google Analytics?

The first analytical tool that can be used to understand the information and learning behaviour of users is Google Analytics (Clifton 2012; Cutroni 2010), which is by far the most common basic web analytics tool. It does not provide a view of individual student behaviour, but it allows to understand some specific activities or habits that may be

associated with a certain behaviour. (Romanowski & Konak 2016; Luo, Rocco & Schaad, 2015; Mc Guckin & Crowley, 2012)

For example, one can measure how much time the student spends on the task page on average, and how much time he spends on the study materials, etc. This is a basic quantifier that allows you to identify (possibly with other parameters such as the number of unique views) a page or part of the course that is interesting and beneficial to students and those that are not. The available data, though, are relatively limited, which implies an interesting thing – the time that students spend on assignments or general studies is more sensitive than the content to form and course design. Therefore, if the data in this area indicates that the focus of the students' interest is only the formal criteria for passing the course, it is primarily the problem of the design, rather than the educational setting.

Another interesting piece of information for a teacher may be, for example, the environment in which the student works and the device the student uses. If we focus on classroom climate in a classroom, it is also possible to work with this parameter in the online environment. The data from the Information Literacy Course of 2017 and 2016 show that students strongly prefer desktops (or laptops) to mobile devices. Their share is about one-fifth compared to usual usage. Thus, the study is for the user of a specific environment and activity. This also accounts for the fact that, for example, mobile phone users on the web generally spend less time than students using a tablet that is the most time-based learning platform.

Somewhat unclear is the fact that students in a university network study for a statistically shorter period than when they study at home (or generally non-university) environment. Here are two possible interpretations, the verification of which would be possible by another method (whether questionnaires or interviews) than web analytics. Students can be more dishevelled at home, so they must study longer, or the school environment is inappropriate or not friendly to study.

Interestingly, the Internet browser also affects visit times and depth of study. The most time spent on the course web is by students using Edge (or Internet Explorer), the shortest visit is seen by Safari users followed by Chrome users. Web browser speeds can have some influence, but the second factor may be that users with Edge (which is the default browser by Microsoft) can generally expect slower web traffic due to lower computer literacy.

However, with Google Analytics, you can also monitor other parameters that may be of interest to the educational institution. For example, differences in user behaviour by city size. Also interesting is the analysis of one of the more important positive features of e-learning – students can study at any time and from anywhere. The two courses we observed confirmed that. Students – due to their position towards Brno – were concentrated mainly in Brno, Moravia and Prague. When it comes to the time students go to courses, it is possible to identify two different traffic peaks, the date of submission

of the assignment or the sending of information by e-mail, as well as the time between 2 and 4 am when the course is accessed by a minimum of students. This is the ideal time to maintain the site.

Certainly, curious are students from abroad. Within the courses we are attending, there are always several students from Erasmus programmes who will register to the course because it is possible to study it distantly. Almost always they are students who devote extra time to the course. Most of them, based on the test data, also achieve above-average results.

Google Analytics also allows you to monitor your visitor flow. This model shows how students are moving around the web. It allows you to analyse the relationships between the web pages students visit so that they return to what they are doing in the course. These are data that are primarily interesting for innovation or an analysis of a course. Our two courses of the same course (one in Umbraco and the other in WordPress 2016) show a significantly different student behaviour that is probably not caused by anything other than course design. How the course (or generally the study material) appears has a great impact not only on the learning outcomes but on the whole information and learning behaviour. From the research point of view, it is certainly gratifying that A / B testing can also be done here.

Social Network Analysis

If we want to measure student behaviour in online courses, we cannot ignore the dimension of social interaction. It is possible to observe one of the basic pedagogical concepts that learning is a social issue that can be directly monitored. More options are available, the first is to work with LMS, when we assume that the whole learning process takes place in one authentication environment. In this case, it is possible to use different methods of monitoring the level of interaction – from reactions in discussion forums, up to common times of joining the system, or following similar educational paths and approaches (Slavin, 2008; Means, Toyama, Murphy, Bakia, M., & Jones, 2009; Clark, Nguyen & Sweller, 2011).

In tertiary education it is relatively not a problem to work with social networks in Czech environments, typically Facebook or Twitter. Both networks make it very interesting to draw a picture of how students do study. The easiest option is a qualitative content analysis that will track the behaviour of students in the group or on the page, or by a hashtag. In this way, along with the tracking of interactions (likes, asterisk, sharing, commentary), one can construct the basic idea of the structure of social interactions and of sentiment and emotional environment of the course. The measurement of sentiment by various automated tools is very important in terms of the breadth and plasticity of possible views of the student through the eyes of available data.

For teachers, it may be interesting to know what pages or users are among those students watch the most. Using analytical tools, for example, you can create a map of influential personalities, track pages that are important to the users, etc. This information may have a research, an intervention or a marketing dimension. From these data, it is also possible to conclude, for example, about what students are interested in and choose examples usable in study materials or try to set up a more natural communication framework.

Hotjar: heatmaps, records

In terms of understanding in what way students study in the online environment, it's important to be able to track them directly on the website. It is very interesting to analyse the direct interaction of an individual user, but also the ability to work with data of multiple visits. Perhaps the most famous service (though not the only one) for making such records and heatmaps is Hotjar, yet for example in Czech environment, it is also possible to see Smartlook.

Hotjar (and its alternatives) has two basic functions (Bergs 2017; Ribeiro 2017). The first ones are heat places that allow you to capture the places that users click (as Google Analytics can do), but especially the areas where the cursor moves. Thanks to the obtained data, it's relatively easy to see how extensive the pages are (where the students are willing to scroll) or identify on which site the individual student spends the most time. This analysis is particularly suited for identifying places that can be complex or interesting for various reasons and allow teachers to work with them in a more systematic way. Typical places that emerged during our research were those that contained difficult-to-understand formulations that students at the same time assessed as important or essential for understanding. It is clear students do not study online in a linear fashion, but that they focus on headlines and highlighted points or enumerations and then return to clarification or context.

This understanding enables the second function of these tools, namely the movement of the cursor on the page. Heatmaps themselves do not allow you to see where the movement was performed from. Rather, it provides a quick overview and identification of interesting places and provides an understanding of how learners in the online environment study and offers a careful analysis of cursor movement from the records.

For example, Hotjar allows to merge all data together into a single report, which then allows the researcher to analyse the data and easily interpret it in an appropriate manner. Certainly, not every mouse move must mean a specific point on the page, but because it is typically easy to work with hundreds or thousands of visits, it is possible to use standard statistical data processing and interpretations.

Learning analytics in LMS

The most comprehensive source of information on students' behaviour and learning habits is provided by the tools associated with LMS. Some LMSs offer advanced analytical modules and tools that allow you to study both individual student behaviour in the course as well as overall characteristics within individual learning objects or throughout the course. Such an approach then gives a comprehensive understanding of both the design and functioning of the educational environment as well as the work of the individual students.

Thanks to these data, for example, it is possible to identify the behaviour of the student, which will lead to an increased probability of learning failure or models of study behaviour that will be successful. Cognitive psychology shows that intellectual ability to define specific procedures for working with information and learning is essential for the efficiency and the teacher can then fit models work learn students, recommending them or work with students, such behaviour shows differently. (Siemens, 2013) This is a view that cannot be obtained through a heatmap or Google Analytics, and for both qualitative and quantitative analysis these are very good data sources. However, some issues or limitations need to be highlighted here. LMS is not a normal site, so it allows you to monitor your study habits within a very specific environment, not on a regular site. The second issue the personal data protection highlighted by the GDPR and the question of how far it will be possible to carry out analyses of this kind in the future. For the data to be stored, it is necessary to clearly describe their necessity for effective education, the student's consent for the data to be obtained without coercion and the processing and storage of the data to be performed within a clearly defined time frame is needed. However, other complications with GDPR will be even more present (Wachter, Mittelstadt & Floridi, 2017).

At the same time, it is important to say that the topic of learning failure, or its prediction based on different behavioural patterns, is one of the key areas of learning analytics. The paternal model of data access, which assumes that the core value is the effectiveness of education rather than the student's privacy, is precisely that assumption. It is also a topic that, unlike much of the current research, has a great economic potential, and much of the LMS is trying to implement some forms of it.

It is also necessary to emphasize that a lot of data do not need to be obtained directly from LMS – even if the web course contains authentication and the ability to perform activities such as discussion or task assignment or peer review, it is possible to obtain a large amount of data through various web analytics methods. In any case, these data from Google Analytics or other similar tools cannot be read which has a great potential in interpreting more general data or patterns of student behaviour.

Conclusion

In this article, we have attempted to indicate how it is possible to work with data on students' behaviour in online courses while maintaining reasonable research breadth. At the same time, we believe that it is a subject that will be further enhanced, whether regarding the development of blended learning or issues related to lifelong learning. We believe that there are many prejudices about online courses – at least in the academic world – that unnecessarily reduce the potential for effective development of these tools at universities. Their more careful exploration, based on data, as we have attempted to imply in this article, could not only help to ensure that such forms of education are incorporated meaningfully and purposefully, but also with their appropriate design and possible rational criticism.

References

- de Almeida Ribeiro, J. P. (2016). The use of web analytics on a small data set in an online media company: shifter's case study. Lisboa: NOVA IMS.
- Bergs, H. (2017). *Gaining insights into website traffic and online consumer behavior with Google Analytics*. (Master thesis). Reykjavik: Reykjavik University.
- Clark, R. C., Nguyen, F., & Sweller, J. (2011). *Efficiency in Learning: Evidence-Based Guidelines to Manage Cognitive Load*. New York, NY: John Wiley & Sons.
- Clifton, B. (2012). *Advanced web metrics with Google Analytics*. Indianapolis: John Wiley & Sons.
- Cutroni, J. (2010). *Google Analytics: Understanding Visitor Behavior*. Sebastopol, CA: O'Reilly Media, Inc.
- Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5–6), 304–317.
- Ferguson, R., & Shum, S. B. (2012, April). Social learning analytics: five approaches. In Dawson, S., & Association for Computing Machinery. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. New York, NY: ACM.
- Luo, H., Rocco, S., Schaad, C., & 2015 International Conference of Educational Innovation through Technology (EITT). (October 01, 2015). Using Google Analytics to Understand Online Learning: A Case Study of a Graduate-Level Online Course. In Hao Yang, H. et al. *Proceedings of EITT 2015 The Fourth International Conference of Educational Innovation through Technology*. Los Alamitos: IEEE Computer Society. DOI: 10.1109/EITT.2015.62.
- McGuckin, C., & Crowley, N. (November 01, 2012). Using Google Analytics to evaluate the impact of the CyberTraining project. *Cyberpsychology, Behavior, and Social Networking*, 15(11), 625–629.
- Means, B., Toyama, Y., Murphy, R., Bakia, M., Jones, K., SRI International., & United States. (2010). *Evaluation of Evidence-Based practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies*. Washington D.C.: U.S Department of Education. Retrieved from <https://www2.ed.gov/rschstat/eval/tech/evidence-based-practices/finalreport.pdf>.
- Morris, L. V., Finnegan, C., & Wu, S. S. (2005). Tracking student behavior, persistence, and achievement in online courses. *The Internet and Higher Education*, 8(3), 221–231.
- Romanowski, B., & Konak, A. (2016). Using Google Analytics to Improve the Course Website of a Database Course. In *ASEE Mid-Atlantic Regional Conference Papers. October 21–22, 2016*.

- Hempstead, NY: Hofstra University. Retrieved from <https://www.hofstra.edu/pdf/academics/colleges/seas/asee-fall-2016/asee-midatlantic-f2016-konak.pdf>.
- Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, 57(10), 1380–1400.
- Siemens, G., & Baker, R. S. (2012, April). Learning analytics and educational data mining: towards communication and collaboration. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. New York, NY: ACM.
- Siemens, G., & Long, Ph. (2011). Penetrating the fog: analytics in learning and education. *EDUCAUSE review* 46(5). 31–40.
- Slavin, R. E. (2008). Cooperative learning, success for all, and evidence-based reform in education. *Éducation et Didactique*, 2(2), 149–157.
- Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the general data protection regulation. *International Data Privacy Law*, 7(2), 76–99.

Contact:

RNDr. Michal Černý
Kabinet informačních studií a knihovnictví
Filozofická fakulta Masarykovy univerzity
Arna Nováka 1, Brno 602 00
Czech Republic
E-mail: mcerny@phil.muni.cz

RNDr. Michal Černý graduated from teaching physics and computer science. He has been involved in technology in education, online courses and digital competencies. He specializes in philosophy of information, philosophy of education and advanced methods of research of learning environment. He is the author of more than ten monographs and hundreds of articles on pedagogical, philosophical and didactic topics. He works at the Department of Information Studies and Librarianship at the Faculty of Arts and at the Department of Social Pedagogy at the Faculty of Education, Masaryk University in Brno, Czech Republic.