

Ethical Considerations in Learning Analytics: Ideas and Discussion

Extended abstract

Kaila Erkki¹, Kurvinen Einar²^[0000-0002-7872-1743], Apiola Mikko²

¹University of Helsinki

²University of Turku

erkki.kaila@helsinki.fi, emakur@utu.fi, mikko.apiola@utu.fi

Designing and developing the digital world crucially requires ethical perspectives. Common threats of today are algorithm-driven attempts to guide behavior, large-scale data-intensive analysis of technology usage and movements, personally tailored fake news, and new social issues generated by job-loss to automatisations. Learning analytics, the analysis of data generated by digital learning systems, is not without its own special set of ethical issues. In this short paper we wish to contribute and take part in this discourse and lay out an agenda for future discussions and activities.

Learning analytics is often considered as an essential part of the future of learning and teaching (Johnson et al. 2016). Learning analytics is often associated tightly with big data, artificial intelligence and machine learning (Elias 2011, Ferguson et al. 2012, Siemens & Baker 2012). The subjects of analysis in learning analytics can be anyone; kids, adolescents, adults or the elderly. This poses a new set of sensitive concerns in privacy and pedagogy for data analytics. Research scenarios in learning analytics often involve complex ethical considerations. An equal chance for learning should not be jeopardized. For a meaningful analysis, it is often necessary to combine data from learning activities with various background factors. However, such data is often sensitive and not easily available to analysis. Richer data means higher quality learning analytics. On the other hand, it means more ethical concerns, one of which is the pedagogically wise use of the results. If learning analytics is able to detect, eg. learning deficits or special talents, how should this information be used in a pedagogically wise way?

According to Slade & Prinsloo (2013), the ethical issues of learning analytics can be divided into three categories:

1. issues with location and interpretation of data,
2. issues with consent, privacy and anonymization, and
3. issues with data management.

Many of the issues in learning analytics are typical to data analytics in general. Still, some generally accepted rules in privacy might prove to be problematic in educational contexts. The General Data Protection Regulation (GDPR 2019) of European Union

Copyright © 2019 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

naturally affects educational technology as well. For example, the “right to be forgotten”, i.e. the right to have all data about you erased from a system, may be reasonable for a social media platform that you are no longer using, but may be problematic if your data in an educational system is used as a basis for your evaluation and grade.

Another example is automatic profiling, which according to Pönkä (2019) means “automatic classification that is based on an individual’s properties, interests and probabilities” or “automatic evaluations, recommendations and interpretations”. Again, this may sound reasonable to forbid if we are considering advertisement profilers. However, it might also mean that some scenarios in learning analytics become impossible. Such scenarios might include, eg. automatically identifying students at risk of failing a course or likely to repeat classes. In some systems, profiling is not done automatically, but the systems provide analytics data to support the decision-making of a teacher.

Slater (2015) has presented a taxonomy of ethical, legal and logistical issues in learning analytics, with a number of questions to consider. For example, what potential negative consequences for a student, such as isolation, bullying, could potentially result from opting out of voluntary data collection such as school photography. On the other hand, what is the impact of such incomplete or missing data for learning analytics? Other valid questions concern anonymity are: should students be able to discuss in course forums anonymously, and should students be able to decide what kind of data is collected and how it is used in analytics? And if so, what kind of outcomes this would have for example to plagiarism detection systems, if a number of students decide that their papers and theses should not be included in the systems’ archives? Moreover, how can teachers use continuous assessment instead of exams, if students can decide which data can be used by the teacher? Lot of similar questions have been raised (LAC 2019) by the Learning Analytics Committee appointed by the Finnish Ministry of Education and Culture.

The ethical issues in learning analytics are both similar and different to issues in data analytics in general. This means that some of the general rules and regulations may be difficult to apply as-is in educational contexts. In Finland, educational institutes already have more liberties than other institutes. For example, the copyright law is different for them. Educational institutes (or municipalities) can also decide, if they wish to use a digital tool in primary education. Such decisions may hinder students’ and their guardians’ freedom to opt-out. After all, municipalities are responsible and required by law to organize primary education. As seen in these few examples, educational institutes do have some mitigations to general rules. But how much do the rules bend?

AI-driven data analytics is often associated with threats such as algorithm-driven attempts to guide our behavior, personally tailored fake-news and massive data-intensive analysis of our technology usage and movement patterns. Since widespread utilization of learning analytics is yet to be conducted, it is still possible to do it responsibly by following strict ethical guidelines as long as we have an agreement of what those guidelines are. However, different stakeholders have different goals in learning analytics, and agreeing on the common guidelines has already proven to be a complex task.

References

1. Elias, T. (2011). Learning analytics. *Learning*, 1-22.
2. Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5/6), 304-317.
3. GDPR. (2016). General Data Protection Regulation. Retrieved from <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679> (Aug 13, 2019).
4. Johnson, L., Becker, S. A., Cummins, M., Estrada, V., Freeman, A., & Hall, C. (2016). *NMC horizon report: 2016 higher education edition* (pp. 1-50). The New Media Consortium.
5. LAC (2019). Learning Analytics Consortium. Retrieved from <https://wiki.eduuni.fi/pages/viewpage.action?pageId=61642250> (Aug. 13th, 2019).
6. Pönkä (2019). Tietosuoja opetuksessa. Retrieved at Aug. 13th, 2019. <https://www.slideshare.net/hponka/tietosuoja-ja-tietoturva-opetuksessa>
7. Siemens, G., & Baker, R. S. (2012). Learning analytics and educational data mining: towards communication and collaboration. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 252-254). ACM.
8. Sclater, N. (2015). A taxonomy of ethical, legal and logistical issues of learning analytics v1. 0.
9. Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, 57(10), 1510-1529.