

## Ethical Considerations on using Learning Analytics in Finnish Higher Education

Jussi Okkonen<sup>1\*</sup>, Tanja Helle<sup>1</sup>, Hanna Lindsten<sup>1</sup>

<sup>1</sup> Tampere University, Faculty of Information technologies and Communications Sciences,  
FIN-33014 Tampere University, Finland  
{jussi.okkonen, tanja.helle, hanna.lindsten}@tuni.fi,

**Abstract.** Utilization of learning analytics in universities is complex setting and there are several issues on motives, expectations and insights on ethics. There is lack of critical literature as the whole topic on learning analytics is still mostly addressed from the perspective of analytics per se. Current literature draws idealistic picture of well-functioning apparatus that supports studying process, enhances learning out-comes and eases the managerial burden and contradictions, even conflicts on utilizing learning analytics are only marginally discussed. The discussion on ethics of using learning analytics is spawned from the different practical and ideological issues regarding learning analytics. Taking the perspectives of privacy, ownership of data, and how studies are managed there are differences on expectations among the students and personnel. The aim of this paper is to examine and discuss the dissonance between expectations and hopes on ethical conduct of learning analytics in Finnish universities. The analysis is based on open ended replies of a survey collected among Finnish university students and staff in spring 2019.

**Keywords:** Learning Analytics · Higher Education · Ethics · User Expectations

### 1 Introduction

Digitalization of the most tuition related processes in universities and also several studying activities are put on digital platforms has opened new vistas for extensive utilization of analytics in higher education. Extensive implementation of data driven management has been executed during the last decade. Okkonen, Helle and Lindsten [1] point out the dichotomy, i.e. difference on managerial perspective, in expectations as students and teachers seek to manage daily activities on micro-level, but those in managerial position seek extensive leverage effect on meso- and macro-levels. These are not contradictory, yet they might not be served by the same development schemes and policies. Very often the development and implementation are regarding the administration, not single users. As single users teaching staff and students have aim to utilize digital tools for personal process management, thus maintain personal relationships and control. The administration perspective [over] emphasizes openness of the data, analytical approach, and seeking gains on mass as well as putting effort on what

brings reward to the institution. Those levels will be described more specifically later in data description.

The literature among learning analytics seem to have two trends, i.e. positivistic and critical. Positivistic tradition seems to concentrate on technical aspects, presenting verified or expected benefits, and having somewhat straightforward managerial perspective on learning analytics. This, to some extent, naïve approach diminishes users and forgets e.g. students as feeling subjects of learning analytics endeavor. The critical trend focuses on ethics of using data as well as discussion of privacy and agency. Ifenthaler and Schumacher [2] have approached student sentiments on learning analytics from privacy perspective. Okkonen et al [1] addresses the same issue by the expectation differences between students and other users. There seems to be certain privacy principles that are universal. The students seem to have need to have certain amount of privacy even their data is stored, analyzed, distributed and used in learning analytics systems and processes. Privacy refers to remaining anonymous to unknown as stated in [1] and [2]. Ifenthaler and Schumacher [2] also underline difference between private and public in sense the data about personal matters is private and data related to studying is public for official use. This also resonates the issue of transparency as people are comfortable being handled anonymous in large bodies of data still maintaining their privacy. Arnold and Sclater [3] underline the purpose for using the data as determining how student accept different learning analytics activities. The students seem to have strong opinions, yet in practice they seem to have little or no power if they are in institution. On the other hand, as pointed out by Okkonen et al [2] staff appreciates possibility to conduct various analytics related activities as well as little limitations on access to data.

There are also other attributes that affect on attitudes towards learning analytics. As stated in Whitelock-Wainwright et al [4] the student expectations on learning analytics questionnaire brought about there were ethical conduct and service as two major expectations. These seem to reflect their general expectations toward any service yet draws attention to role of students. In general learning analytics takes students into account as loyal subjects, not active participants. Corrin et al [5] similarly state that data management and data protection as well as privacy and access are the most important ethical issues. Whitelock-Wainwright et al [4] state students have no clear conception of the contents of learning analytics activities.

According to Schumacher and Ifenthaler [6] the main expectations of the student users of learning analytics are support for planning and management of studies, support for execution by organizing the learning and studying processes, giving feedback for self-assessment and reflection, suggesting adaptive recommendations for choosing courses, and providing personalized analysis on completed credits. Okkonen et al [2] point out that Finnish students are reluctant to be compared with their peers, especially without their consent. This seems to reflect the ethics of power as discussed in Slade et al [7] by the three ethical perspectives. The first one is location and analysis or interpretation of the data. The second is consent, privacy and anonymization of the data. The third is data management, i.e. possession of the data, data classification and storing the data.

Okkonen et al [2] states the considerations on ethics are due to use of learning analytics as the means and platform of teaching, tutoring and supervision as the expectations on gathering data, analyzing the data, using the data and the distribution of the

data within course, degree program or institution vary significantly between different user groups (cf. [8, 9, 10]). The students have privacy issues with their data, especially if they are expected to grant access to any random staff member. There is ambivalence on the topics as they are reluctant or indifferent as they seem to have no clear opinion on the issue. Management, administration and individual teachers point out that learning analytics calls for accessible data for certain processes or even to ad hoc purposes. As an analogy to discussion on data privacy where service providers justify expanding, even escalating, need for more thorough and detailed data on individuals by the more sophisticated services, more accurate information and wide array of data and information related products and services. This is almost opposite to how people with high awareness of privacy seem to address the topic. Especially when analytics is implemented in full extension the privacy issue, or how my data is used, takes new form. As presented in Okkonen et al [1] roughly 50 % of the students were willing to receive benchmarking or other algorithm powered data on their performance. This reflects to finding that university students considered themselves as part of academic community, not raw material, refined for the needs of the society.

The policies of utilizing learning analytics are established as there are several directives to restrict the implementation and use of analytics. Also, institutions have internal operation procedures communicated to staff. Therefore, the actual practice of learning analytics is externally moderated. As stated in Okkonen et al [1] it is not that divergent expectations exist, but mostly it is about on what staff, and especially teaching staff, is work on as the resources are limited and mostly learning analytics related chores are considered odd jobs and no specific resources are not allocated to execute those. The issue brought about in Okkonen et al [1] is that teaching staff has low expectations towards learning analytics as they need mostly up to date contacts to students and course related information. The realistic expectations of teaching staff seem to dilute the high hopes of data driven tutoring and counselling of students. It should be noted that the Finnish university system has so far used very little analytical information in student guidance as implementation of learning analytics or even operation management systems are in early stage and full utilization is yet to come. Okkonen et al [1] bring about that the students tend to be more positive about a system that they already use and whose logic they are familiar with. Some of the critical attitudes of the respondents to the information produced by the hypothetical system may be explained by its unfamiliarity as the case was for some respondents in Okkonen et al. [1].

This paper aims to elaborate the analysis of ethical considerations on using learning analytics. The ethics is manifested by the expectations and use of learning analytics. The data utilized for this paper consist of expectations on using learning analytics and therefore some interpretation is needed. Based on the open-ended responses extracted from more extensive dataset this paper approaches following research questions: What are the expected uses and ethical considerations on using registry data and How these expectations differ between students and staff. The aim is therefore to elaborate the insights on how using learning analytics is and can be justified in higher education.

## 2 Data and Analysis

The data of this study has been collected from six Finnish universities during the spring of 2019 and it continues the quantitative analysis presented in Okkonen et al [1]. The aim for this sequential analysis is to elaborate the results regarding the expectations on learning analytics to perspectives on ethics of learning analytics. The universities are Tampere University, Aalto University, LUT University, University of Turku, University of Eastern Finland and University of Oulu. The inquiry was disseminated through news bulletins of university intranet sites and targeted mailing lists. The purpose was to reach the widest possible range of respondents. The respondents were divided according to the user groups of the analytical data, into groups of students, teachers, teacher tutors, study coordinators and those responsible for education. Each of these groups had their own form to answer, questions targeted to take into account their possible, special needs on utilizing learning analytics. Heads of study affairs, heads on degree programs, deans and vice deans responsible for education were instructed to respond to the survey of those responsible for education. In one of the six universities (University of Oulu), the inquiry was only distributed to a teacher tutor. The way in which a poll was distributed in each university varied, so it is very difficult to estimate the number of respondents to the inquiry and thus the response rate. It is known that there is a total of 77 430 Bachelor's or Master's degree students in six universities. We assume that the different ways in which universities distributed the questionnaire has also had a strong impact on the number of respondents. In the analysis and interpretation of the material, it should be noted that the respondents have had experience with the registry systems used at their own university and in their responses, they reflect the experience of that particular system. Total amount of responses were 183 students and 170 staff members. It is also important to notice, that students responded to questionnaire, do not represent the average Finnish university student. According to their answers, 70 % of them are about to graduate within 5 years, which is the case only in 20% of university students in Finland.

The purpose of the five user surveys was primarily to find out what the different user groups consider to be important goals in utilizing analytical information. The survey was not collected primarily for research purpose, it was mainly to be used in the development of applications and information systems utilizing analytical information. This places limitations on the analysis of the material, for example the questions asked to different user groups are not always comparable.

All the questionnaires consisted of both multiple choice and open-ended answer options. Multiple choice answers were directly coded into SPSS which was used for statistical analysis. Open-ended answers were processed by outlining the key topics and creating a variety of categories which would describe responses as well as possible. It is also important to notice that a certain response could include in severe codes. Open-ended questions were coded because they gave a lot of knowledge to understand more deeply the reasons for different answers. Open-ended answers brought about additional information on the in ethical considerations and utilization of the analytical data in different setting and by different agents. Those two questions were presented exactly in the same way in all the forms so that made the comparison possible. In addition, there were also some questions which appeared on some forms. There were 86 responses from the students for question which explored the ethical issues

and 81 responses for a question related to how the university should utilize student data. That indicates that more than a third of student respondents answered to these two pure open-ended questions. That indicates that the students had clearly something to say about these issues.

As discussed also in Okkonen et al [1] it was found at the same time that the work of coordinators and teacher tutors, the guidance of studies, often lies at the interface between micro and meso levels, the coordination of teaching and study requirements and the student's personal curriculum. It can be said that the levels of utilization of analytical data are not so much divisible by the actors, but rather the processes that can be supported by analytical data. Those levels and processes concerning the analytical information is presented table below.

**Table 1:** Different agents' processes concerning the analytical information and levels (cf. Okkonen et al [1])

Agent	Process, the analytical information	concerning	Level
Student	Studying and planning studies		Micro
Coordinator/ Tutor teacher	Counselling, guidance		Micro
Teacher	Planning education		Meso

These findings were taken into account in the preparation of the questionnaires. The student's form focused on particular on ways of utilizing analytical knowledge that could support the student's planning, guidance and follow-up. The study coordinator and teacher tutor surveys focused on particular on the needs of study guidance and in the teachers' questionnaire on the information to be used in planning teaching of individual courses. In particular, the survey for those who are responsible for the studies focused on information management issues. To collect unpredictable information the question how the university should use the information collected in the systems was asked in all forms.

As stated in Okkonen et al [1] ethical concerns, especially regarding the private information about the individual student, were strongly raised in the students' interview. Those interviews were utilized in questionnaire forming. As a result, surveys attempted to identify students' readiness to allow different user groups to utilize student personal information as well as the need for other user groups to utilize such an information. In addition, an open-ended question asked the respondents about the possible ethical problems associated with the use of analytical information. In the survey the topic was approached from the perspective of "data in university repositories" not using term "learning analytics". This was conscious decision not to use abstract term that some might understand well, and some might have no reflection on their experience.

### 3 Results

In this study the ethical concerns found are augmenting those presented in the research setting justification. As discussed already in Okkonen et al [1] findings are roughly divided here in three categories. Ethical issues can be seen 1) in individual behavior, 2) in policy of institution using learning analytics or 3) in validity and relia-

bility of the data. The ethical considerations are reflected by the results of the study as those that are the main justification for the use of learning analytics, successfully implement learning analytics, and executing learning analytics policy in any institution the data is about. The ethical considerations will be presented through the content analysis of data. The ethical considerations were first divided into 11 categories before they were shaped into those three categories, which is a more abstract description of the data.

The utilization of any data, i.e. data in university registries, studying platforms, managerial systems etc., was asked in an open-ended question, "How the university should utilize the registry data?". There were 81 student responses and 87 staff responses. Table 2 summarizes the responses divided to different groups.

There were no significant differences between the students and staff. Most important purpose for using the data was designing and developing the education. The respondents also thought the data should be used for the benefit of the student. It was somewhat surprising that 12 % of staff respondents did not know how to utilize the data. However, the classification of the replies emphasizes the underlying idea of learning analytics activities serving the studying related processes more than serving just managerial good.

Different perspectives to utilizing the data were taken into account by asking about ethical considerations, "What kind of ethical considerations are related to utilizing of the registry data?", there were 86 student responses and 65 staff responses. The responses on ethical considerations varied more significantly and there were several issues brought up by both students and staff. Table 3 summarizes the findings.

**Table 2:** The purpose of the usage of the registry data

	Respondent group	
	Student	Staff
	%	%
For planning and developing the education	57 %	52 %
For the benefit of the student	35 %	43 %
For operational predicting and development the university	15 %	16 %
I don't know	3%	12 %
For the financial indicators	0 %	3 %

**Table 3:** Students' and staff's ethical considerations related to utilising the registry data

	Respondent group	
	Student	Staff
	N %	N %

	Respondent group	
	Student	Staff
	N %	N %
Transparency in collecting and using data	31 %	14 %
Misusage of data	30 %	12 %
Using sensitive and personal data	28 %	19 %
The access control	22 %	8 %
Data storage and leakage	11 %	15 %
System-level privacy	6 %	2 %
No ethical considerations	4 %	14 %
I don't know	4 %	7 %
Usage only for the counselling purposes	2 %	11 %
Only anonymous statistic data should be used	1 %	6 %
The ethics of individuals is sufficient	0 %	6 %

*Transparency* in collecting and using the data was emphasized by the students and this refers both the role of using the data for the benefit of the students and using the data without hidden agenda. Transparency in this sense in explicitly articulated motivation for gathering, processing and distributing data and information in the organization. The staff acknowledged the issue, yet they did not put as much emphasis on the topic.

Case was the same in concerns of *misusing the data*. Students saw it more possible, that there may be some staff members who do not use the data responsibly. *Using sensitive and personal data*, i.e. personal data protection and privacy, is the key issue in discussion on ethical and sustainable learning analytics. Respondents pointed out that all private information and especially information connected to individuals should not be distributed or used without their permission. Using sensitive data is also issue of explicit consent as information related to health or social issue are not subject to normal collection of data. Student seem to have no clear recognition on what data is available to whom. In the questionnaire it was also asked, if it was acceptable to pool information from different sources, like previous studies and suchlike. This possibility was highly rejected by students, at the same time as staff members saw it more useful in counselling purposes.

One of the critical notions on ethics is access to personal information. The students are more willing to restrict the access, and staff also acknowledge there is possibility to misuse of data. In some universities teaching staff has unrestricted access to all student data even it is not necessary in their work and ethical conduct is only ensured by disclaimer that data is allowed for official use only. The distinction between stu-

dents and staff is evident, since students see the issue from personal perspective but staff from the work-flow perspective. It is also noted, that students do not always know, what information is registered.

The issue related to sensitive and personal data is *data storage and possible leakage*. On the university level data is stored securely and, in most cases, it is used also as intended. However, there are some examples how staff members have leaked data or used it unethically e.g. by browsing personal data without acceptable reason or publishing information on individuals. In this study the issue is acknowledged more often by staff members.

Also, *system-level privacy* is an issue as data is considered to be pseudonymized and without personal credentials to general audience. However, one privacy related issue is that data should be used only for the counselling purposes. This refers to notion that data is used only for explicitly described purposes by explicitly nominated staff members, e.g. those who have connection to respective student. To extreme some respondents state that data should be used only for statistical purpose and individuals should not be studied at all as individuals.

The ethics of learning analytics seem to be uncharted domain and some respondents were totally unaware of ethic by stating “I do not know” what ethical considerations are related to learning analytics, or even that there are *no ethical issues* to consider. Especially the staff members should have put more attention to ethical use of data, but the topic is multifaceted and, in most universities, still without local agenda. On the other hand, the ethics of individuals is considered sufficient, yet obviously not without risks of misconduct.

There is a quotation of a student which summarizes many of the students’ concerns: “Students need to know when and for what purposes information is collected about them. Although data is processed anonymously, it should be possible to deny data collection. There must also be a clear definition of who is allowed to process the data and how. For example, teachers in individual courses or university management, should under no circumstances be given free access to individual student information”

## 4 Discussion

The findings related to ethical considerations seem to underline the issue of private and public. The students seem to be quite sensitive on how personal and private are defined in sense of data collected by universities. However, there seems to be distinction on how they define private and personal in sense on their online presence. On the other hand, the students are quite privacy aware in sense of data that is gathered during their studies. While, even it is was not directly addressed in the survey most of the users of e.g. social media, do not assumingly put much effort on restrictions and privacy as they seem to bother on their data. In studying issues, the distinction to personal is drawn in having a personal relationship with someone, e.g. tutor or counselor. At the same time, this seems to be bound to Finnish university system where students have high rivalry on access, no tuition fees and high degree of freedom. In Finland there is no tradition of using standardized tests to assess students or use those in universities either. Therefore, they do not need to be compared to each other during their studies. This seems to differ much of what is found in the literature.



The general ethics of learning analytics seem to be universal. The first ethical principle is to maintain control of the data on individuals as they should have at least nominal power on their data and they should be informed how it is used. The general sentiment is that they are happy if data is utilized for their benefit, yet not on controlling their actions. This is a two-sided topic as any data could be used in controlling subjects. The second ethical principle is privacy as students seem to have a clear understanding of whom they are willing to release their information to. The managerial, i.e. administrative, perspective is almost opposite but there should be local consensus on ethical conduct. The third principle is data protection as all information should be stored and distributed securely without risk of being leaked or used for other than designated purposes.

The qualitative results extracted from the survey augment the operationalized questionnaire and open new vistas on ethical perspectives of learning analytics. There are several limitations to be considered. The purpose of the initial survey was to gather user expectations of various user groups and each set of questions was designed to provide input for development of learning analytics, not to exhaustively and uniformly research ethics of learning analytics. The data presented above serves as a critical starting point to researching learning analytics in practice. These results open new questions when user experience, especially topic-specific experience, is assessed. The functionalities of learning analytics services provided by universities have an impact on planning studies and studying as information provided has a designated effect on both students and administration. Future research on ethics should pay attention to how certain features affect both student and staff users.

## References

1. Okkonen, J., Helle, T., Lindsten, H. (2020) *Expectation differences between students and staff of using learning analytics in Finnish universities*. In Proceedings of ICITS2020
2. Ifenthaler, Dirk & Schumacher, Clara. (2016) *Student perceptions of privacy principles for learning analytics*. Educational Technology Research and Development
3. Arnold, K. & Sclater N., (2017): Student perceptions of their privacy in learning analytics applications. LAK'17
4. Whitelock-Wainwright et al. 2019. The student expectations on learning analytics questionnaire
5. Corrin, L., Kennedy, G., French, S., Buckingham Shum S., Kitto, K., Pardo, A., West, D., Mirriahi, N., & Colvin, C. (2019). The Ethics of Learning Analytics in Australian Higher Education. A Discussion Paper.
6. Schumacher, C. & Ifenthaler, D. (2017). Features Students Really Expect from Learning Analytics. *Computers in Human Behavior*
7. Slade, S & Prinsloo, P [2013] *Learning Analytics: Ethical Issues and Dilemmas*. *American Behavioral Scientist* 57(10), 1510–1529.
8. Howell, J.A., Roberts, L.D., Seaman, K., Gibson D.C (2018) Are We on Our Way to Becoming a ‘‘Helicopter University’’? Academics’ Views on Learning Analytics Tech Know Learn ( 23: 1. <https://doi.org/10.1007/s10758-017-9329-9>

9. Roberts, L. D., Howell, J. A., Seaman, K., & Gibson, D. C. (2016b). Student attitudes toward learning analytics in higher education: “the fitbit version of the learning world”. *Frontiers in Psychology, 7*, 1959. doi: 10.3389/fpsyg.2016.01959.
10. Roberts, L., Chang, V., & Gibson, D. (2016a). Ethical considerations in adopting a university- and system-wide approach to data and learning analytics. In B. Kei Daniel (Ed.), *Big data and learning analytics in higher education* (pp. 89–108). Switzerland: Springer.