HOW DATA CAN IMPROVE THE QUALITY OF EDUCATION

CONTENTS

INTRODUCTION

In recent years, the Dutch higher education sector has acquired a wealth of insights into the potential use of data in education. Some of these insights are direct results of experiments with learning analytics, others derive from research into learning analytics.

Online education brings a number of major changes. One of these is lecturers having less insight into what students are doing. In a classroom, lecturers can see how students are getting on. It is obvious whether students are present or not, whether they are participating in class, and whether they have completed their assignments. In an online education environment, this overview is generally absent. This complicates things, as all lecturers want to support their students in the best possible way. The use of data in education could offer a solution for this problem. The overarching term for data use in education is 'learning analytics'. Learning analytics consists of the collection, analysis, and reporting of data about students and their environment. They help to understand and improve the learning processes and environment of teaching sessions.

Learning analytics gives lecturers better insights into online study behaviour, so they can support students and prevent them dropping out. Learning analytics also helps improve online materials. Data on how often, how long, when, and how materials are used, provide an indication of the extent in which materials are fit for purpose. Learning analytics is a logical consequence of online education. It is a way of obtaining insights into what is happening in an online learning environment.

Linking existing data

The use of data in higher education can be applied broader. Higher education institutions have access to a wide range of different data sources about students' prior education, how they landed at the institution, age, where they grew up, academic results, etc. Institutions also track data in physical environments. For example, additional Wi-Fi trackers may have been installed within the institution or heat sensors may be present in the lecture hall. Linking a number of these systems together provides insights into students' learning behaviour, the quality of the education, and the effectiveness of the institution. At an institutional level, learning analytics offers new opportunities for analysing trends, reducing the number of drop-outs, and working more efficiently. Information already known can be linked to new insights, as this information is now available digitally. This allows you to "nudge" students who are at risk of dropping out or underperforming, for example. Depending on the insights the data provides, you may need to take action.

Searching for links in a targeted way

Just a few years ago, people thought the ability of big data to collect as much data as possible, would result in ever more insights. Irrespective of whether or not this is correct, endless data collection and analysis is unthinkable in the post-GDPR era. It is not the way higher education institutions want to treat their students' and staff data. Learning analytics has the potential to deliver better education for students because it generates more insights. These insights, however, are always the result of specific questions. The first question one should ask is therefore: What do you want to achieve through the use of data? What exactly do you want to find and why?

[1] *This definition comes from the Learning Analytics Community Exchange (LACE). See also: http://www.laceproject.eu/faqs/learning-analytics/*

The follow-up question is: What data are essential to find an answer to your question? Prevent endless combing through data sets and search in a targeted way. Analysis in a minimal dataset is what learning analytics will be all about in 2019. As the practical example of VU in this issue shows, you can generate new insights with a carefully targeted process.

Another development that generated a lot of speculation several years ago, was the fact institutions could use learning analytics to predict which first-year students had little chance of completing their degrees. As things stand, there is little enthusiasm in Europe for selecting students on the basis of computergenerated predictions. The Dutch and Belgian higher education sectors believe strongly in the importance of treating the data of students and staff in an ethical, transparent way. However, statistics can be used to give students a realistic picture of their academic progress, as is the case at KU Leuven.

Document structure

What is the potential of learning analytics for the Dutch higher education sector? How do students and lecturers actually benefit? This is the subject of Chapter 1. Taking a number of practical examples, we describe what Dutch and Belgian higher education institutions regard as the key added value, having experimented with learning analytics for a number of years. Chapter 2 looks at the didactics, legal, ethical, and technical challenges and the potential impact on education. This chapter explores the insights derived from a previous learning analytics experiment SURF conducted with the institutions. We also demonstrate how these insights lead to a change in direction, with more opportunities for the ethical use of data. As well as tips on getting started with the use of data in education, this white paper contains tips for institutions that have successfully completed the experimental phase and are looking to scale up learning analytics.

Different types of analytics

The Learning Analytics Community Exchange (LACE) defines the *use of data in education* in a broad way: collection, analysis, and reporting of data about students and their environment in order to understand and improve the learning processes and environment in which the teaching sessions take place. But this does not say anything about the types of data used. Are you looking to predict the future, understand the past, or are you looking at real-time situations?

In this white paper, we use the term learning analytics consistently for all types of analytics. Some believe learning analytics, however, relates only to what happens during the learning process. Learning analytics provides insights into the study process, with the idea of being able to help or guide the student better. Under somewhat broader definitions, learning analytics may also include multimodal data, e.g. data from fitbits, Wi-Fi trackers, or detection doors within the institution.

Student analytics uses existing data from various sources. The data are made available to lecturers and academic advisers to enable them, for example, to assess potential at-risk students or students who could be challenged more.

Academic analytics involves using data to analyse trends and developments at an institutional or school level. These can be success rates but also, for example, predictions as to the likelihood of various groups of students completing a programme.

Machine learning is a development that largely takes place outside the education sector. It is rare for higher education institutions to run large quantities of data through an algorithm, hoping for the algorithm to provide new insights.

The precise definition of these terms is open to debate. The most important thing is that there are various ways of approaching data in education. It is paramount to consider how you wish to use learning analytics: Do you want to know what happens during students' learning process? Or do you want to analyse trends using existing sources? The definition of the form of data use is up to you.

THE POTENTIAL OF LEARNING ANALYTICS FOR THE DUTCH HIGHER EDUCATION SECTOR

LACE gathers evidence globally on the effectiveness of deploying learning analytics in education. The LACE Evidence Hub contains a wealth of examples of institutions that have started working with learning analytics. What is the added value for the Dutch higher education sector? Could the use of learning analytics mean interventions and tailored education? Or should the focus in the short-term be on improving learning materials?

Justian Knobbout is researching the impact of learning analytics at HU University of Applied Sciences Utrecht. He sees a significant need for low level analytics. Students and lecturers often benefit from a simple dashboard giving insight into what their students are doing in the digital learning environment. “Use learning analytics to evaluate the use of learning materials,” he says. “Ideally, this will give users far more of an insight into the learning process than is currently the case. It will enable lecturers to see exactly how they need to guide different students, and enables students to see what the next step in their learning process might be.”

But one course is not the other: for example, one lecturer may need a summary of the most discussed topics on the online forum, whereas another will primarily want to know how their video materials are being viewed. “A learning analytics system needs to be flexible and able to meet people’s various requirements,” says Knobbout. At the moment, this is not always the case.

Whilst, in terms of its breadth, data usage in education is constantly increasing, in terms of its depth, the potential of data usage is still not being fully exploited, Knobbout points out. He sees, for example, when implementing a new digital learning environment (DLO), institutions still regard learning analytics as a ‘nice to have’. “It’s far better to implement it from the start, than wondering in retrospect how you want to extract data from various applications that work together in a single environment. After all, that is almost impossible to do. You need to have a data storage infrastructure from the outset, which can be temporarily disconnected until you need it.”

The Amsterdam Center for Learning Analytics (ACLA), a research group based at VU set up in 2016, conducts research into the effectiveness of deploying learning analytics in education. Its director is Chris van Klaveren, also Associate Professor of Education Sciences & Economics of Education at VU University Amsterdam. He set up ACLA with Ilja Cornelisz and Martijn Meeter as he realised learning analytics requires close collaboration between different disciplines. The research group combines insights from the academic disciplines of economics, computer sciences, artificial intelligence, psychology, and educational science. He is enthusiastic about the insights derived from this collaboration. “We don’t have the knowledge to solve everything. By engaging in dialogue, you acknowledge each other’s gaps in knowledge and you are able to address them.”

Learning analytics is work that primarily involves people. “I’ve never seen data answer a question and algorithms don’t just happen, they are thought up and programmed by people”, says Van Klaveren. “The added value of learning analytics is no different to other tools that individualise or personalise education; it adds value if it improves something, be it the learning experience, be it the learning environment, or the learning outcome.” He believes the real-time nature of learning analytics has huge potential. “You monitor the right information and the learner can access this information at any time. This can potentially have a positive impact.”

According to Knobbout’s research, students who are already involved in their own learning process, in particular benefit from learning analytics. “They think about their learning behaviour and a dashboard with their personal data is really useful for this. But first-year students, who have not yet gotten into the swing of things, are more interested in other things than their learning process.”

PRACTICAL EXAMPLE

ACTIONABLE FEEDBACK: DATA THAT STUDENTS CAN USE

KU LEUVEN

The transition from secondary school to higher education is often a difficult one. That is why, in 2015, KU Leuven decided to carry out research into how existing data could be used to give students feedback during this transition process. First-year students now receive a dashboard four times a year that gives them insight into their learning and studyskills and academic results. The dashboards also give students specific tips regarding any improvements that may be required.

The data come from two data trails. The first of these is students' grades, which are derived from assessments, exams, and the benchmarking test, the test that is taken before the start of the academic year. The second data trail is self-reported data. All first-year students complete a validated questionnaire about their learning and studying skills, the so-called LASSI (Learning and Study Strategies Inventory) questionnaire. Based on this questionnaire, they receive the first dashboard at the start of the academic year. The dashboards they receive during the year indicate, amongst other things, their position in relation to their peers. They can also see how students who achieved a similar score in the past managed to do so. Tinne De Laet, Head of Tutorial Services in the Faculty of Engineering Science reveals: "We give students information they did not have before, data we can access easily from the databases of a higher education institution. It is not a complex algorithm and it does not involve advanced technology. Everyone can do so easily. The real challenge is in the design and in integrating it into student support." In other words, what do you actually want to show students? And how do you encourage them to take action when and where required?

De Laet: "We try to give students advice on things that are actionable. In other words, things that they have influence on. For example, we do not include gender or social economic status in the model, as these are factors students cannot change. We link the students' analytics to tips and references they can use to make improvements. KU Leuven offers a wide range of training courses on learning and studying skills."

The approach appears to be successful. The learning and studying skills dashboard encourages students to talk to their academic counsellors at an early stage in the semester. Students also indicate they find the learning analytics and tips useful.

The main ethical challenge is to not discourage students with statistics. With this in mind, a number of guidelines have been drawn up. For example, in the statistics, students were divided into three groups with each group including some students who are successful in their programme. De Laet: "It may happen that a student manages to do something that no one has ever done before, so we will never say that this might not be possible. The dashboard may say something that a student does not want to hear but it always puts things into perspective." The dashboard is transparent about the uncertainty of the prediction. And there is never any talk of probabilities. "It is nonsense to say someone has an 80% chance of obtaining a bachelor's degree in three years", stresses De Laet. "What we say is: of the students with the same grades as you, 80% obtained their bachelor's degree in three years. But this is based on what we know about you and we do not have a complete picture of who you are."

In general, students are enthusiastic about the project. This is illustrated by the students' request to have the dashboards rolled out across all bachelor programmes. One way or another, learning analytics is steadily taking root within KU Leuven. The project has been rolled out by the Faculty of Engineering Science to the Science and Technology tutorial group. In the last academic year, it covered 26 study programmes within KU Leuven. This reached 12,000 students. The next step is an institutionwide project.

PRACTICAL EXAMPLE

UNEXPECTED BENEFIT AT THE INSTITUTIONAL LEVEL

VRIJE UNIVERSITEIT

As part of a major project to digitalise the entire educational records system, Theo Bakker, at the time working for Deloitte, launched a sub-project, aimed at giving students greater insights into their academic progress through the use of combined data. Together with Deloitte, VU Amsterdam developed a model that predicted the probability of first-year students passing their first year of study. The model was used in a pilot designed to give study advisers and tutors an extra tool to help them in tutoring. VU Amsterdam developed a comprehensive Code of Practice for this for students. The pilot was not an outright success but the dataset still provides interesting insights on a weekly basis.

In total, VU Amsterdam has a dataset of more than 1,500 characteristics per student, derived from 80 data sources. Bakker is now strategic policy advisor and project leader for Student Analytics at VU Amsterdam, He is responsible for the 'Secure and reliable use of learning data' zone of the Acceleration plan education innovation through ICT. He explains: "We look at transactional data. In other words, data that are already contained in records. What schools do students come from? When did they apply for their study programme? Did they attend open days or "welcome days" in which they participate in courses? Did they take part in matching exercises? We also use data from the language test that all students take. And we use additional characteristics, such as prior education, age, and gender. Then we look at the results they achieve at VU Amsterdam."

Depending on the source, Bakker can provide new data on a weekly basis. He says this speed is one of the key benefits of learning analytics.

"Study advisers take action when a student comes to see them. These kinds of tools allow you to be proactive and make adjustments where required." This did not prove to be the case for the pilot: in the study advisers' opinion, the data rarely added to what the student had already told them. And they were worried the predictions would demotivate students. So, for the time being, supporting individual students based on predictions is no longer on the agenda. But there are plans for implementing a new project at a student level in 2020 in close consultation with students, study advisers, and tutors.

The pilot provided data that was useful in another respect, however. Large quantities of the data provided insights into the institution. Bakker: "Suddenly, we discovered that many of our first-year students had already followed a course at university and had not come straight out of high school. This gave us reason and proof to flag with the government that our funding is considerably lower than of other universities." These insights triggered the interest of the Executive Board. Now, policies are often revised on the basis of data. "We make new discoveries nearly every week," says Bakker. "I was surprised how little VU actually knew about itself. Three years ago, a lot of policy was based on ideas. Now it is evidence-based, wherever possible."

In 2019, VU Analytics is becoming an official team. VU Analytics will focus on new policy areas, such as recruitment and the job market. The team also wants to be able to answer more ad hoc questions from lecturers. Bakker: "I hope that more and more research groups will analyse these data".

PRACTICAL EXAMPLE

SURF LEARNING ANALYTICS EXPERIMENTS

EXPERIMENTS GIVE LECTURERS HANDS-ON EXPERIENCE

Together with a number of educational institutions, SURF developed an experimental environment where a number of institutions have taken their first steps in the area of learning analytics. The first experiments took place in 2016 with lecturers from VU Amsterdam and HU University of Applied Sciences Utrecht. In 2017, VU Amsterdam, Windesheim University of Applied Sciences, Eindhoven University of Technology, Zuyd University of Applied Sciences, and the vocational education and training college ROC Noorderpoort took part in the project. In total, there were twelve lecturers from six institutions. The institutions' experiences are described in the lessons learned.

VU

Ilja Cornelisz, co-founder of the Amsterdam Center for Learning Analytics (ACLA) and Statistics lecturer at VU, was involved in the development of SURF's experimental environment for learning analytics. As a lecturer, he participated in the experiment twice, in order to obtain better insights into the learning process of students and align the education offered more closely with their needs. He stresses learning analytics raises issues, but does not immediately provide a solution. It is a tool for engaging in dialogue with students and lecturers.

HU

The HU University of Applied Sciences Utrecht is carrying out research into how learning analytics can contribute to improving the teaching process. For three lecturers, the experiment was a way of putting learning analytics into practice in an easy way. For them, it is a great way of monitoring teaching.

Windesheim University of Applied Sciences

Yvonne van Vooren, a lecturer at Windesheim University of Applied Sciences, used learning analytics to gain better insight into her students' learning behaviour and skills. Based on the learning analytics, she discussed with students, ways for them to improve and ways to make changes to her lectures. She concluded that learning analytics is a good tool for improving lectures

TU/e

Eindhoven University of Technology had worked on their own policy framework for learning analytics, which made it interesting for them to participate in the experiment. By collecting as much data as possible, the participating lecturers hoped to gain more insights into the use of online teaching materials within the subject of Calculus. This method proved to be labour-intensive and not very fruitful.

Zuyd

The aim of Zuyd University of Applied Sciences' experiment was to give lecturers exposure to learning analytics. The institution concluded that this approach was somewhat too non-committal and, as a result, recommends learning analytics to be demand-driven.

Noorderpoort

ROC Noorderpoort is the first senior secondary vocational education and training institution (MBO) to gain learning analytics' experience in the SURF experimental environment. Using learning analytics for Digital Skills yielded a wealth of information for lecturer Jacob Poortstra and he is hoping for a possible follow-up.

THE CHALLENGES OF LEARNING ANALYTICS

Didactic challenges

The application of learning analytics starts with a good research question. This is hardly new information. 'What is the best education I want to offer, using the resources available to me?' is the recurring question. Because resources are becoming more and more advanced, you have to consider how you want to use them. Start doing this long before you begin your lecture or teaching process, otherwise it may turn out that data are not available, or that it is impossible to interpret them.

After deciding on what resources to use and how, a number of preconditions are necessary to achieve didactic added value. The Scotsman Niall Sclater, a consultant for Jisc, SURF's sister organisation in the UK, sums it up: "Do you have people in-house who understand the data, who know the data sources and can interpret them? People who can make predictive models or conduct statistical analyses? And more importantly, once all that is behind you, who can carry out interventions among the students? You can produce statistics until the cows come home, but if you cannot carry out interventions, you might as well not start at all."

The lecturers do not need to possess all these skills themselves however, in addition to thinking skills and the ability to interpret analyses, a certain amount of ICT skills is a must. Didactic challenges and technological challenges are linked together. Starting with learning analytics means, first of all, for the digital learning environment to function as required and for lecturers to have sufficient ICT skills. Many institutions are still lagging behind.

Legal challenges

The arrival of the GDPR has made many institutions wary about using learning analytics. Strictly speaking, however, not much has changed in terms of data usage possibilities. SURF is attempting to dispel the institutions' doubts in various ways. For example, by carrying out research into the appropriate measures for applying learning analytics. A publication about this was published in June 2019 (only in Dutch).

Compared to the Dutch Personal Data Protection Act, institutions using learning analytics under the GDPR must be more accountable internally and externally, says Niels Westerlaken, IT and Privacy Lawyer at Project Moore Advocaten, who is conducting this research. "The other steps have remained almost the same, but this particular step frightens project leaders, because you actually have to show what you are doing." The question about the subsequent appropriate measures cannot be answered clearly, because the impact of research using learning analytics can vary considerably. Westerlaken: "There are some general principles you must always comply: provide enough information, give a lot of control over the data, and specify what you want to do. For projects with an impact on individuals, where you might also merge data to make profiles, an appropriate measure could be to request consent from the people whose data you are using."

An institution might reason that data usage always benefits education, and therefore complies with the legal obligation. But that is making short shrift of this requirement. "Research carried out by institutions will serve a good purpose, but that is not a licence to apply learning analytics directly. Namely, there must be a balance of interests between the institution's objectives and the individual's privacy."

For lawyers, an initial court ruling in this area of education will be very interesting, as it will provide more clarity on how the courts interpret the law. Whether a measure is 'appropriate' is not set in stone. "Those involved have expectations, and we will have to see whether that matches what will happen in the future. explains Westerlaken. "In 5 or 10 years' time, students will perhaps expect learning analytics to be applied. A different scale of measures will be used compared to what we do now, in the phase it is reasonably new. Now you will really have to do your best for the benefit of the use to be clear. Not to cover yourself, but mainly to gain the trust of those whose data you are processing."

The institution is obliged to account for the data collected and its purpose. At present, students' expectations do not always correspond to reality, particularly when it comes to collecting data. Many students expect education may be improved, for example, on the basis of pilot tests in the digital learning environment. But it is very often less clear to them that, for example, therefore mistakes made at the individual level are monitored.

Westerlaken advises institutions to document experiments with learning analytics well. For internal and external accountability, but also to share the results outside the institution itself. How do you apply the principles? Which objectives are defined in advance? "If you can share that with other institutions, you're working towards a standard. That is educational for everyone." Consider, for example, the outcomes of Data Protection Impact Assessments, the Jisc Code of Conduct and the VU Code of Practice.

Privacy and other ethical challenges

Learning analytics will become a standard component of the higher education institution learning environment. That is what Hendrik Drachsler, professor at the Welten Institute of the Open University of The Netherlands and board member of the Society of Learning Analytics Research (SoLAR), expects. In 2015, he said: "I think that, in 5 or 10 years' time, we will not be discussing learning analytics any more, we will simply be using it. The lecturer will fetch a report from the system to use for the course, without even thinking about the fact he is using learning analytics." In 2018, he still believes, "it is just taking a bit longer than I expected due to the new privacy legislation, but we are seeing more and more learning analytics being used on a daily basis."

One of Drachsler's biggest nightmares is that one of his students will go to court because they object to their data being used for learning analytics. At the same time, this student would probably get good grades, because Drachsler sees the application of learning analytics as an excellent example of teaching students to think critically about data usage in general. He calls this data literacy. "In German, it would be called Daten-Mündigkeit. That goes somewhat further than literacy. Someone who is data-mündig has sought information on the use of their data, is opposed to it, and has every right to be. I regard trusted learning analytics, as I like to call it, as a possibility for training young people in this area."

According to Drachsler, higher education institutions wishing to apply learning analytics will, first of all, have to draw up a vision document and create a policy. Transparency is a priority, in this regard. Which data are being collected? Who has access to the data? What possibilities does the owner have? Drachsler: "A vision requires an innovation programme to make people aware of the direction they are considering of taking and the tools they are wanting to use to achieve their goals. Change management is often necessary. Lecturers still require a lot of training. Students also need more knowledge. Learning analytics requires a lengthy organisational process."

There are several resources available to support institutions in handling data transparently. Such as the DELICATE checklist, designed by Drachsler and others within LACE. Jisc identifies no less than 86 ethical and legal issues concerning learning analytics. Niall Sclater, who compiled the list, says: "The greatest challenge are cultural institutions: are the people ready to take decisions based on data?"

Technological challenges

The biggest technical challenge of learning analytics is dealing with large volumes of data from different sources. Data can originate from a student information system (grades and learning pathways), the digital learning environment (learning activities). But attendance data can also be a data source. Data are collected in a so-called Learning Record Store. In this store, data are structurally stored so they can be used for analysis and visualisation. A protocol is necessary to allow the various source systems the data come from to communicate with the Learning Record Store. Two protocols are contending to become the market standard: xAPI and IMS Caliper Analytics.

The technology for learning analytics is reasonably well-defined, but two developments have raised new issues. On the one hand, the target group of learning analytics is expanding: in addition to students, lecturers, and academic advisers, researchers want to use educational data. Interest is also increasing at an institutional level. All target groups want to see their own questions answered using learning analytics. The quantity of data sources will only increase, too. Students do not live in a digital learning environment but mainly outside it. By connecting data from the digital learning environment to additional data sources, such as fitness trackers and other existing data, a more complete picture is formed and it is possible to provide a better answer to a research question.

At the same time, the GDPR obligations must be met. This requires efficiency, among others: you collect data only for a pre-established purpose. The problem is, however, that too many data are being stored too quickly using the current technology.

Changes to SURF's learning analytics' architecture

These developments make it necessary to change SURF's learning analytics' architecture. Three considerations are currently leading to changes:

1. **Ownership.** The data are collected about the student, for the student. The student is in control of their personal data and can share their personal data with a lecturer or academic adviser. Data that has been rendered anonymous, that cannot be traced back to an individual student, may be used by others in the institution to answer their learning analytics' questions.

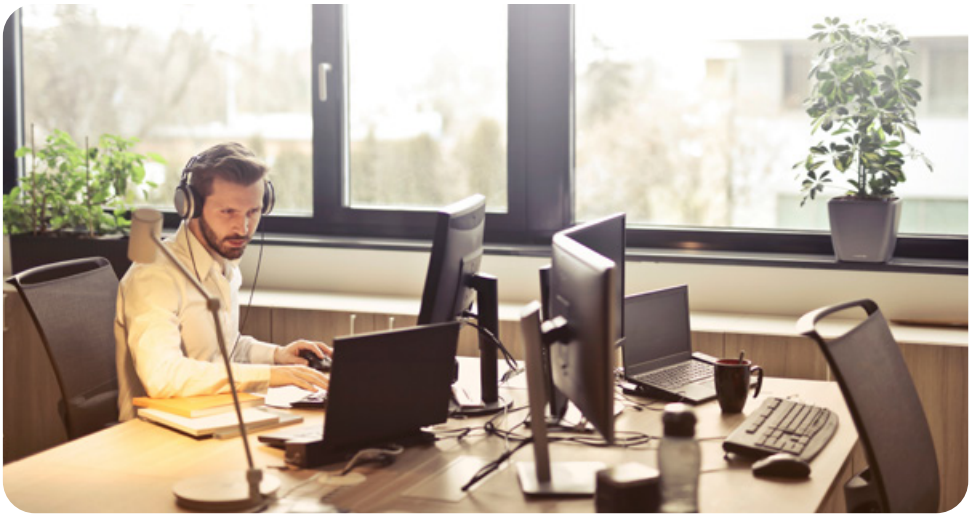
2. **Data minimisation.** Data are collected for a specific question. Suppose a lecturer wishes to know how often their video is viewed. Is it then relevant for them also to see that a student views it at 3 am and at 8 am? Not really. Therefore, only the frequency should be counted; timestamps may not be recorded. The challenge is to collect exactly enough data to be able to answer the research question adequately. This is the purpose of limitation , which makes it possible to perform learning analytics within the GDPR. Data minimisation is one of the most important privacy-by-design principles SURF wants to apply. Other privacy-by-design principles SURF uses include separating data, rendering data anonymous as much as possible, and abstracting data, as in the example above.

3. **Trust and transparency.** By giving students control over their data and collecting only data that are necessary to answer a specific question, the foundation is laid for a new platform based on trust. In addition, transparency about the questions used is necessary and the algorithms that provide the visualisations for the question (open algorithms). Additionally, SURF wants to be completely transparent about the operation of the new platform (open architecture). In this case, SURF expects learning analytics may get a new boost. Experts in the field will be involved in developing the research questions and open algorithms so it is clear to every stakeholder where the question and answer in the visualisation come from.

There will be a pilot with a platform where researchers and developers can formulate research questions and connect to data sources, with or without support from the institution's technical team. The platform offers various services, specifically focused on the questions from the various stakeholders. The services may be shared between institutions, whereby each institution benefits from new developments.

These considerations have a big impact on the architecture. The most important change is a new security layer for monitored data management. This layer ensures constant compliance with the GDPR. If a stakeholder formulates a research question and makes use of a service on the platform, this layer checks if the use matches the purpose. The system will automatically consent questions only requiring data that has been rendered anonymous and certain personalised questions in which the institution has a legitimate interest. For other research questions with a specific purpose or using a specific data type, a consent screen may be necessary for students. General consent will not be asked for data use, but for specific research questions when they arise. The first pilots with the new architecture will begin in 2020.

As shown in the practical example, the Open University in The Netherlands is busy working on the same considerations as SURF. The development of a trusted learning analytics system is showing large similarities to the architecture SURF is building.



[2] *The GDPR states that the processing of personal data must be linked to specific collection purposes. This is called purpose limitation. For more information, see the Handleiding Algemene veroordeling gegevensbescherming (General Data Protection Regulation handbook) and the Uitvoeringswet Algemene verordening gegevensbescherming (Dutch Act implementing the General Data Protection Regulation) (both in Dutch): https://autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/handleidingalgemeneverordeninggegevensbescherming.pdf*

PRACTICAL EXAMPLE

EN ROUTE TO TRUSTED LEARNING ANALYTICS

OPEN UNIVERSITY OF THE NETHERLANDS

The wide-ranging deployment and adoption of learning analytics in European higher education institutions is stagnating due to concerns about privacy and ethics relating to personal data and the GDPR. This uncertainty is leading to innovation managers having doubts about implementing learning analytics in their institution. In order to make the achievements of 10 years' research into learning analytics securely available to lecturers and students, the Open University in The Netherlands (OU), together with the DIPF - Leibniz Institute for Research and Information in Education in Germany - is working on a so-called Trusted Learning Analytics (TLA) infrastructure. This infrastructure tackles the current lack of trust and transparency in learning analytics. Hendrik Drachsler, one of the TLA initiators, says: "The TLA consortium is attempting to renew the 'contract' between students and education providers, not only by generating a high level of trust, but also by increasing the data knowledge and involvement of students. The design of the TLA infrastructure and tools is based on design processes that reward the same importance to ethical and privacy matters, as to functional requirements. The result is a system that achieves objectives technically, ethically, and humanely."

Trusted learning analytics must be developed from a user's perspective. Users should have total control over their data and not feel any distrust in relation to the system. In order to achieve this level of trust, it is crucial to transform as much 'black box' data into 'white box' data as possible. The consortium is attempting to be as open as possible about the algorithms applied and transparent about the methodologies used. An important objective of a TLA system is providing feedback, rather than it being used to make decisions. The TLA system should provide full access to the 'data subjects' - those whose data is analysed - and allow them to correct, comment on, and delete data. They must also be able to decide who has access to their personal data. In current learning analytics' infrastructures, this is often not possible. Drachsler: "For the OU, the TLA approach provides the basis for a new way of handling educational data, and for providing services and activities to stakeholders."

A TLA system requires the data subject, i.e. the student or the lecturer, to have a high level of skills, including critical thinking, information analysis, and self-regulation. The open-source TLA system attempts to create an environment in which they can train these crucial skills.

[3] *Drachsler, H. (2018). Trusted Learning Analytics. Synergie, 06, Fachmagazine für Digitalisierung der Lehre. (in English)* *https://synergie.blogs.uni-hamburg.de/ausgabe-06-beitrag-drachsler/*

Impact on education

Data only answers questions when the research has been designed well. It also does not offer ready-made solutions. Giving back information about behaviour does not mean that people are going to improve their behaviour. Chris van Klaveren from the Amsterdam Center for Learning Analytics (ACLA) says: “In its current form, learning analytics looks good and you can see many ‘footsteps’ again, but I suspect that this alone will not achieve any results. If we want to personalise education effectively, it is important that we do not just do research into what the best solution is on average. In this respect, the Bayesian Bandit models are interesting because they are based on the idea that we want to know how your dynamics works. In this way, we can also learn which interventions do not work and why they do not work.” He stresses the need for this. “Time is running out. The technology is getting better and better, with VR, dashboards, and eye-tracking but, truthfully, it will not get better at obtaining improved learning outcomes. Not as long as the diagnostic element is lacking.”

Hendrik Drachsler from the Open University in The Netherlands agrees with it to a certain extent. “Research shows that most dashboards are insufficiently evaluated and do not fulfil educational processes. There is no evidence that a dashboard has an actual impact on the learning process because that is precisely what is not being researched.” He still sees too many institutions that think providing a dashboard is sufficient. “It is not enough to present students or lecturers with a dashboard, it is about a permanent support process. This is the only way to improve the quality of education.” The SHEILA project (Supporting Higher Education to Integrate Learning Analytics) questioned 4,000 European students, lecturers, and experts about their expectations of learning analytics. This was captured in a framework that supports the implementation process.

Certain matters are also simply still unknown. During the first few years, research into learning analytics was mainly focused on technology. Since then, there is more attention on interventions. Justian Knobbout from the University of Applied Sciences Utrecht is curious about what the next phase will yield. He says: “The learning analytics process is cyclical. There are still too few examples of research into performing the cycle again. What happens when we continue to follow students and use old analytics to measure again, so we can carry out targeted interventions?” Research into data usage in education is still new. Much will have to be seen in the coming years.

[4] *This blog provides a brief explanation on the Bayesian Bandit: https://medium.com/@haydar_ai/learning-data-science-day-20-bayesian-bandit-problem-f7e87acfc2b2*

STARTING WITH LEARNING ANALYTICS

Starting with learning analytics is primarily a matter of doing. Here are 10 tips for institutions wishing to experiment with learning analytics:

1. Top-down or bottom-up? There is something to be said for both. A number of experts advise starting off small, with free software and small research questions. Others say, however, that a large project has more impact and offers more possibilities for learning analytics to become accepted within the organisation.
2. Start with a question. What problem would you like to solve? Or what would you like to know about education? Make the process interactive. Ask lecturers and students which questions they would like to see solved using learning analytics. Tinne De Laet from KU Leuven says: “We asked the academic advisers what we could make. Based on that, we came up with something they didn’t expect, but are really comfortable with.”
3. Involve the IT department from the start of the project.
4. Sort out privacy and ethical matters in advance. Consult legal counsels or specialists for advice.
5. Avoid jargon. Students do not know that they are engaging in “learning analytics”, or that they are making use of “learning dashboards”.
6. Be clear towards students about what data you are analysing and why. Show them they own their data. Ensure there is trust in the system.
7. Appoint a supervisor on a central level. Leadership is of great importance. Institutions that experiment successfully often have a strong leader, who is in a position to introduce learning analytics within the entire institution.
8. Announce you are busy working with learning analytics. Often, there are various initiatives taking place within institutions, of which people may not always be aware, and opportunities for collaboration are missed as a result.
9. Take the time. Connecting the various databases, sorting out privacy matters properly, and finding technical solutions all take time.
10. A major challenge is the cultural shift. Are people ready to take decisions on the basis of data? Give employees time to get used to the idea. Meanwhile, support early adopters.

UPSCALING LEARNING ANALYTICS

We asked experts for their advice to make learning analytics a well-rooted component of the organisation. Here are 15 tips for upscaling learning analytics.

1. Ensure that the highest management level of the institutions is on board.
2. Take participants along: the student council and works council.
3. Have the project managed by a steering group with a broad representation of stakeholders in an institution: scientists, students, study advisers, IT suppliers, and policy officers.
4. Make an annual calendar of strategic projects within the organisation so you can focus on current policy themes in time.
5. Do not just provide insights on policy questions, but actively participate in working groups on advice stemming from policy.
6. Have members of the team think about what they want to work on and in what order, including the institution's priorities.
7. In a Code of Practice, determine the purposes for which your analyses will or will not be used, the basis on which this happens, who has access to the data, who isultimately responsible, what data you have collected, for which policy advice it has has been used, for how long the data is stored, what rights students have, and how personal data is handled. Link this to the institution's ethical principles.
8. In the initial design, bear in mind that it must be scalable.
9. Start with the data you have. These are intrinsically very scalable.
10. Record which datasets are used, what data they contain and, who is responsible for the data.
11. Make use of a style guide that specifies exactly what variables in datasets are called and how code is to be documented.
12. Automate operational work employees do not enjoy doing as much as possible, like importing files.
13. Test each stage of the data processing system using automatic test scripts. This maintains the quality of your data, as well as the reliability of analyses.
14. Keep stakeholders involved in the project. The academic advisers of KU Leuven can determine the content of the dashboards themselves. In practice, they do not change the content much, but it gives them a feeling of ownership of the project. This in turn ensures acceptance.
15. Invest in communication. All stakeholders must be provided with the results.

CONCLUSION

SURF sees great potential in learning analytics as a contribution to improving the quality of education. There is added value to be gained, not only by the organisation, but especially by the students and lecturers. If they are convinced of the benefits of learning analytics, you will maximise the use of data.

In order to take substantial steps with data usage in Dutch higher education, the zone 'Secure and reliable use of study data' of the Acceleration plan for educational innovation with ICT was established. SIG Learning Analytics also plays a key role in the sharing of knowledge between institutions. In the future, SURF sees a possible role for itself as the provider of a basic system for learning analytics that complies with the GDPR.

COLOFON

Author
Jocelyn Manderveld
Marjolein van Trigt

Project Management
Jocelyn Manderveld

Layout
Vrije Stijl, Utrecht

Photography
Christina Morillo / pexels.com

Januari 2019

SURF

088 – 787 30 00
www.surf.nl