

Education is an important aspect of society and assessing the progress made by learners is key to appreciating, validating and if need be improving the educational system in place. Determining how to best assess learning outcomes continues to be a challenge for teachers and policy makers alike. From collaborative problem solving to using interactive programs, printed books and supplementary materials to independently finding information online, learning can take many forms as can assessment.



The goal of the [KNOWeSCAPE](#) workshop “Making sense of education indicators” [held on November 18-20th in Amsterdam](#) was to

gather a diverse group of experts interested in data and education to discuss how to adjust current data-based indicators to track more accurately a varied range of learning approaches. The discussion centered around data, privacy, indicators, and visualisations for the first two days of the event. During the final day the workshop participants joined the [Bee Collective Festival](#) on November 20th to crowdsource additional feedback on evaluating progress made by young learners.

During the introductory round on different education systems and software introduced by individuals in the room (Sugar, Eclass, Ustad, teaching.codes), it was outlined that some software designed for certain pedagogical approaches (e.g. Sugar) only needs to log basic information about learner activity. Student and teacher attendance itself can be considered to be an indicator; in low-income economies it is common to lose 10-24% of the education budget to teacher absenteeism. Ustad Mobile demonstrated software that can capture paper attendance sheets and process them into a database.



Privacy issues associated to using the full logs of online tools such as Eclass (or Moodle which it is based upon) were discussed along with the risks associated to letting learners teach themselves programming with potentially harmful code being created (intentionally or not) and run. This brings with it another challenge: tracking and assessing non-linear progress paths where learners are allowed to go back, sometimes several steps back, and try again or jump ahead to skip working on things they already know. Those systems are taken

as the basis for discussing the kind of data that can be used for monitoring education.

What kind of data ?

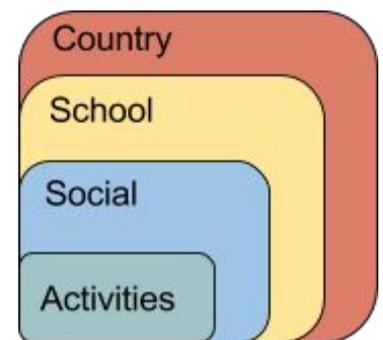


Participants were split into groups to consider the types of data being generated by different systems. The question to answer was “What kind of data point is, or could be, used to generate an indicator?”. The outcome of this exercise was four distinct groups of data points and indicators, which demonstrates how intertwined these two topics are. Indicators can hardly be designed without having knowledge of the data available, and it is somewhat pointless

to just generate as much data as possible just in case some of it could be of use someday. The outcomes were as follows:

- **School Level Data:** Knowing the type of pedagogy practised by the teacher (ratio between referential versus creative work, educational approaches, classroom methods) as well as the actual resources available to the teacher are important to understand the context of the data points.
- **Country Level Data:** This concerns the wider environment the school is located within (federal, provincial and municipal). It can explain and contextualise the school-level metadata. This includes but is not limited to: the accountability chain for the education performed, the physical location of the school, the general goal of the education structure, trends spanning several schools, the number of schools providing data as compared to the total number of schools (data coverage).
- **Activity Data:** This relates to the content produced and experienced by the learners, what when and how. The following data points have been lined up: number of errors made during an activity, level of complexity of the artifact produced (e.g. number of words used, numbers of programming blocks or number of colors used in paint), mime type of the content produced, time of the day, activity in front of the screen (keystrokes, mouse movements), timestamps for the start/end/pauses of/in an activity, name given to the content produced, user ID, machine ID. It was noted that some data, such as the wording or the color palette used for producing content could also be used as indicators of social behaviours and span over this group of data too.
- **Social Data:** Lastly there are data points that characterise the learner herself/himself and her/his position within the social structure of the school. Gender and age are two obvious data points. Student motivation for learning and her/his general attitude to it e.g. self-confidence of the student is another data point. The social network and the type of relations among the pupils is also to be considered. It is useful to know the preferred communication style/medium (text/voice), the directionality of the interactions (teaching/listening), the frequency of the interactions and the degree of self-disclosure (to which extents pupils are proud of showing their work to others).

It can be argued that there is in fact an implicit hierarchy among this data both in the way one group helps make sense of one another but also in the way this data is consumed. The activities performed by the learners can be influenced by their social network and behaviour, which in turn can be influenced by the practices in place at the school, those being influenced by the national policy at the country level. Those looking at country level data do not need to know about the details of the activities performed or the social behaviour of the learners. This influence chain also relates to levels of abstraction and aggregation of the data.



How to secure that data ?



Abstraction and aggregation were at the core of the discussion that followed. Creating data sets always come with a risk (see, for instance, the book [“Shooting your hard-drive into space”](#) arguing this point) which needs to be understood and mitigated. A post-it session was organised to reflect on data security. As a result of this exercise two groups were established: one concerned with different practices for securing data and another concerned with the actual threats.

- **Practices:** Obfuscating users and machine identifiers using a one-way hash is a first step and a must prior to releasing data. Another sensible thing to do is to encrypt the data. Aggregating the

data blurs the details and thus prevents identification of individuals and/ or any malicious party deanonymizing the data. The addition of a query API is also a good way to secure the data as it provides controlled, and eventually monitored, access to part of it rather than giving away complete dumps. Finally deleting the data after some time can be a good solution to ensuring it does not survive “too long”. This is what Eclass does by deleting all student data that is over one year old. This is efficient but there is a challenge in setting how long is too long. In that specific case one year is a bit too short as it does not enable students to look back at their past results throughout the course of their degree.

- **Risks:** It was noted that “If you want to beat a thief you have to think like one”, in other words assessing risk can entail data abuse, e.g. by using triangulation to figure out the identifiers from their hashed fingerprint. Even if the hash can not be reverted a sufficient amount of data can lead to finding out the identifier it is associated to. Informing all stakeholders of the risks is important. The risks posed by inherited data sets (e.g. an organisation can create a secured data set for an approved cause and hand it over to another organisation that will potentially misuse it) were also addressed.

The bottom line of the discussion was that hashing identifiers, controlling access, aggregating and limiting the lifespan of data set was a good set of practices for enabling data to flow to those who need it whilst mitigating risks and protecting the privacy of individuals.

How to make use of the data ?

One of the aims of this workshop was to discuss the type of indicators that can be computed using the data identified as being relevant to collect. Three particularly important points were raised:

- **Pick a model:** having a model for data processing is important when discussing with decision and policy makers which need a certain level of abstraction from the data. We discussed three of those models: the [Ackoff model of Data->Information->Knowledge->Wisdom](#), the [content/context/process dimensions by Pettigrew](#) and the [Learning Analytics Reference Model](#). All those have different characteristics. Pettigrew's model is the only one of the three to clearly feature the strategy behind the processing of the data. It is designed more for strategic change than for the exchange of information. This is clearly visible in the work of the [PISA working group](#) who used this model to produce several comparative indicators for OECD member and partner countries education systems. The Learning Analytics Reference Model is the only conceptual framework to clearly position indicators. In comparison, those indicators could be placed between ‘Information’ and ‘Knowledge’ in Ackoff’s model but this decision would have to be argued for.
- **Be aware of the context:** an indicator in itself is obscure. You need the entire context of the choices made for the analysis in order to fully appreciate it. There is also a potential cultural bias to account for along with personal experience. For example, it can be observed that having a bank account, a credit card or both is correlated with having better test results on financial literacy. The prominence of local family businesses requiring the presence of kids at home will impact their school attendance. Conversely the opportunity for a free school meal will motivate parents from low-income families to send their kids to school whatever the quality of education being received there. All those factors have to be taken into account when picking a set of identifiers and looking at their outcome.
- **Aim right and sound:** there are several types of indicators, all of them being a mathematical computation leading to an outcome. 80 different types of statistical inequality indicators can be



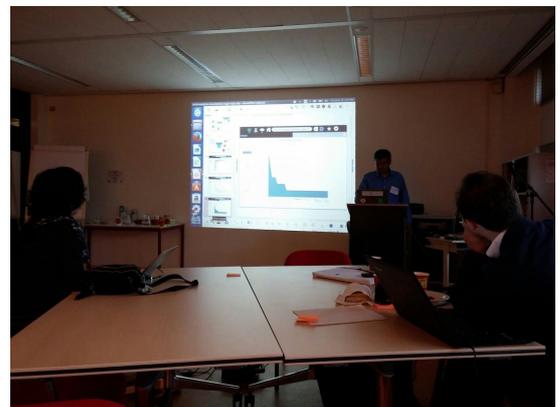
distinguished. This adds up to some 100 non-statistical indicators. These indicators provide background information; qualitative context is needed to make sense of them. This discussion can be guided by a goal e.g. increasing equality, helping struggling students, spot the specific skills of the best students to send them to specialised tracks, etc. The time relevance and actionability of indicators also has to be considered. Cars feature indicators meaning “go see a specialist”, others meaning “stop now!” and another set being informative of the current system status (RPM, speed, ...). A similar system could also be applied to education and other domains. Lastly, it is important not to [confuse correlation with causality](#) and to ensure that the statistical indicators are statistically significant when used.

This can be summarised as picking a set of indicators relevant for the goals at hand and ensuring those are correctly used and acted upon as necessary.

How to get from raw data to visual indicators ?

This last part of the discussion was about turning the output of indicators into

a visually attractive representation that is simple to understand. We looked at [XOVis](#), one of the current initiatives related to the learning environment Sugar. Education Data Mining (EDM) is not in focus currently for Sugar so there is no data being logged in the system with the direct goal of producing indicators. XOVis uses the data from the journal which logs the usage made of applications (“activities”). A qualitative approach with in-class



observations and interviews is needed to get a complete picture to supplement the data from the logs. From a technical point of view the data is first logged in the laptop running Sugar, then it is pushed to a school server running XSCE which acts as a micro-cloud. This micro-cloud finally pushes the data to a [central server](#) when connectivity permits it. The synchronisation is left up to CouchDB and Cloudant is used for the central instance. The setup is working but some key side-issues have to be kept in mind: 1) teachers had to be trained how to understand the graphs, what they meant and how they could read the results; 2) the time data about the usage date of activities is sometimes inaccurate because the clock is frequently set incorrectly on the laptops; and 3) some teachers asked to get data on a per-child basis which is tricky privacy-wise and contrary to aggregating the data. Although the development of XOVis was driven by the data available rather than the need from the teachers/ministry it proved to be useful so far to show that the laptops are being used. Further work is needed to exhibit more indicators.

Beyond the [common graphs](#) used in XOVis it was shown that indicators could be pictured as ‘traffic light’ style status indicators with directional arrows for trends. When relevant indicators are chosen that align with policy makers requirements attractive and accessible tools can be created.

Education is a wicked problem



Under the guidance of the team from the festival Bee Collective we applied the model of Decentralised Collaborative Organisation (DCO) to discuss the [paradigm shift occurring in education](#) and its assessment. Participants were asked to discuss which parts of the education system should be confined to the past, what we should be aiming for in the future and what elements are currently in transition.

The main outcomes were as follows:

- **Past:** rote learning, (excessive) direct instruction, (excessive) constructivism, excessive student:teacher ratios, physical punishment, standardised testing, testing as an end result, batch education.
- **Future:** more focus on developing strong metacognitive skills amongst learners, teachers as facilitators, teachers as learners, project-oriented teaching and learning, peer to peer learning, blended learning, physical and environmental awareness, critical thinking, creativity, active involvement of all parents in the school community.

Participants were encouraged to envision how education stakeholders (teachers, students, parents and others) could transition to a new paradigm using the [Bridges' Transition Model](#). Potential solutions were considered from the point of view of each stakeholder. The most promising solutions were voted on by the participants (e.g. the "Bees"). The top six solutions in order of popularity were as follows:

1. More frequent feedback (school to parents e.g. progress monitoring & reporting on the performance of both the child & the school in question) (9 votes)
2. Interdisciplinary learning (4 votes)
2. Meaningful, real life examples (4 votes)
5. Creating incentives for teachers (2 votes)
5. Putting the expectation on every child to train and qualify for a high skilled job vs. society's need for low skilled jobs. (2 votes)
6. Real-time progress visible for parents (1 vote)

Take away messages

Some take away messages from those three days:

- It could be interesting to develop a query API with support for access control lists (ACL) on top of the already well established [Tin-can protocol, also known as Experience API](#).
- There is a need for a framework to help stakeholders choose identifiers.
- Education is moving from the capacity of remembering undisputed facts to the capacity of solving problems and critically looking at information. This will call for adjusting the identifiers used to appreciate the quality of education.
- Education should not be seen as a step by step process with clear checklists. It is rather a continuous process allowing for different paths. This calls for indicators based on a continuous form of assessment.
- Education indicators can be statistically incoherent, used outside of their context or based on inaccurate data. This are important risks that need to be kept in mind at all times and mitigated as possible.
- A working paper should be produced to develop the ideas discussed during the workshop. This should include the further elaboration of the "[Learning Analytics Reference Model](#)".
- There is a difference between indexes and indicators. The exchange rate (for example \$:€) is considered a financial indicator. Unlike the Dow Jones, this is not an index value because its value is determined but it can be argued that the Dow Jones is also a financial indicator. It can then be concluded that both a determination and a calculation can be used to construct an indicator. What might be useful is to realize that you can [manipulate the value of a determination](#). It might be much

more difficult to manipulate an index value which is calculated. In that sense, the index value might have a larger internal safety and trustworthy position.

Written by Christophe Guéret and Benita Rowe. Photos by Christophe Guéret and Anna Bon. Many thanks to Sameer Verma, Roeland de Kok, Mike Dawson, Jahna Otterbacher, Giulia Rotundo, Senka Anastasova, Haluk Osman Bingol, Anna Bon, Bert Bredeweg, Cheah Waishiang, Applonia Mmbone and Nikolay Vitanov for their contributions.