1. Using learning analytics systems for educational policies

1.1 Focus of the use case

This use case considers the opportunity that a more extended use of learning analytics represents for an improved policymaking in the field of education.

There is no universally agreed definition of learning analytics, but an often-cited definition is:

> *"Learning analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs."*[1]

Learning analytics focuses on the individual student and his/her educator and is used predominantly in the institutional environment to inform on students' behaviour, to support formative assessments and most importantly, to improve institutional performance in relation to, for example, drop-out rates. A key characteristic of learning analytics data sources is therefore their level of granularity.

Learning analytics have the potential to eventually play a role in addressing many different policy issues concerning education, training, mobility and employment. The focus of this use case is on the potential benefit of learning analytics data to understand **transversal skill attainment**. Transversal skills, notably ICT and entrepreneurial skills, have become in high demand in the labour market, resulting also in more attention for these skills in education policies. This use case therefore considers big data can potentially enhance the design of educational strategies aimed at digital and entrepreneurial skills by drawing upon the insights from formative and summative student assessments.

Currently the attainment of transversal skills is measured through data from **international surveys**. The Global Entrepreneurship Monitor[2], for example, measures people's perception of their entrepreneurship capabilities and their aspirations towards starting a business; the survey of adult skills (PIAAC)[3], instead, measures the perception of having ICT skills. These types of surveys are a good source for information on the current level of skills; however, they give little insight in the process for the attainment of skills through learning (which in turn may predict future capabilities).

Forecasts related to skill attainment (and skills needs in the labour markets) are currently collected on a regular and systematic basis at national, international and European level. The European Centre for the Development of Vocational Training (CEDEFOP) provides bi-annual skill supply and demand forecasts[4], with insights into EU-wide skills needs and supply capacities, allowing for detailed comparisons across occupations, sectors and countries. Several Member States including the UK, Germany and Italy also produce their own forecasts.[5] The detailed projections and methodologies used are not all publicly available, though.

This use case focuses on the opportunity that the collection of micro-data on learning processes and the use of learning analytics can constitute for policymakers at national and European levels. From a design perspective, this use case can be considered to

---

[1] Society for Learning Analytics Research (SoLAR), 1st International Conference on Learning Analytics and Knowledge, Banff, Alberta, February 27–March 1, 2011, https://tekri.athabascau.ca/analytics/

[2] http://www.gemconsortium.org

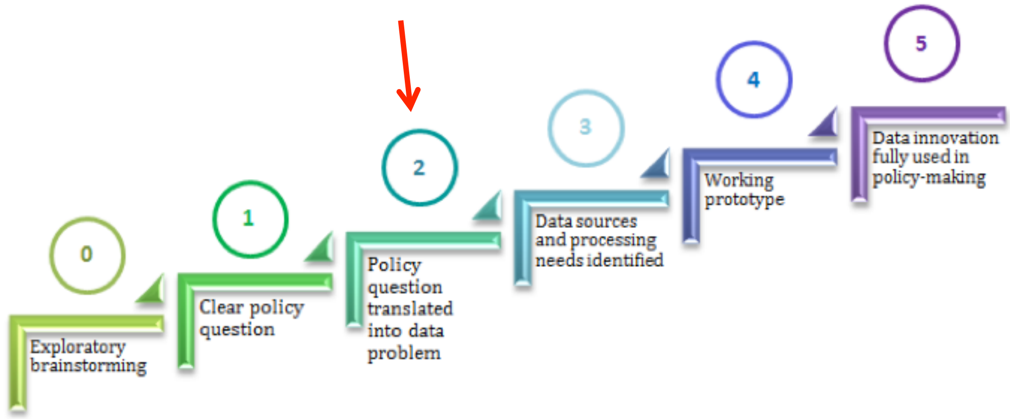
[3] http://www.oecd.org/site/piaac/

[4] http://www.cedefop.europa.eu/en/events-and-projects/projects/forecasting-skill-demand-and-supply

[5] CEDEFOP (2012) *Building on Skills Forecasts – Comparing Methods and Applications, Conference Proceedings*, European Centre for the Development of Vocational Training, Research Paper 18, Luxembourg: Publications Office of the European Union. Available at http://www.cedefop.europa.eu/en/publications-and-resources/publications/5518

have reached step 2: the policy question is translated into a data problem, but the data sources and processing needs and opportunities need some further consideration (Figure 5).

Figure 3 Use case readiness level



1.2 The rationale

Education systems are complex systems where performance depends on the interplay of multiple factors and actors. Analyses of educational processes build upon a deep understanding of the structure and interrelationships between the actors involved, in primis the students and their educators. Especially in the case of education, the national context, culture and history are important elements to take into account for policymaking. It implies the need for adaptive interventions and no one-size-fits-all models: national (and even regional) educational systems are diverse and specific.

International surveys are currently used to understand skills levels, but little is known about actual educational activities in these areas. This is mainly due to the nature of these skills; they are seldom attained through a degree completion (e.g. in computer science), but more often through distinct features of educational modules (e.g. ICT modules or ICT elements in other modules) and non-formal learning (e.g. online tutorials). Mapping transversal skill attainment in learning activities will provide insight into current levels and trends in skill attainment. By needing participants to provide the information to collect the information at individual level, surveys are often costly and are therefore conducted at most annually.

In this context, the expectation is that micro-data related to formative and summative assessments (at the student/institutional level) can constitute a precious source for the investigation of the **processes** allowing for the attainment of digital and entrepreneurial skills, and when analysed comparatively, be of relevance for a higher policy level analysis. The data should be collected at institutional rather than individual level, reducing the cost and burden on participants and can be updated more regularly than annually.

The monitoring of skill attainment based on learning analytics provides a number of advantages. The use of the data collected about the micro-level (students/institutions) can complement the survey data collected at the meso- or macro-level (i.e. country or European levels) by providing insight on **processes and enabling factors**, which data at the meso- and macro-level can rarely provide.

The data can be used for the monitoring of **skills levels** and the identification of **gaps**. The increasing demand for and development of data sources on skills gaps by policymakers demonstrate that there is a concrete need for this data. This could feed into multiple stages of the **policy cycle**. Ex ante, it could identify problems or generate foresight and agenda setting. Ex-post, the data could generate useful indicators for monitoring and evaluation. Most important, it could increase the effectiveness of policymaking by generating more insight in actual skill attainment than currently possible, which would enhance policy advice to Member States. Data demonstrating gaps in skills linked to disadvantaged groups could be used to develop or adjust interventions targeting specific groups (by awareness raising, supplying additional teachers, funding more courses, adapting courses, etc.). Furthermore regional data can be linked to labour market opportunities and be used to stimulate mobility between regions.

1.3 The policy context

Digital, entrepreneurial, transferable and multi-disciplinary skills and education programmes are topics that are high on the EU policy agenda. This is to be set against the context of, on the one hand, the shortage of ICT professionals on the EU job market and concerns about the number of entrepreneurs, and on the other hand, youth unemployment rates resulting from the economic crisis, global competition and automation.

The Treaty on the Functioning of the European Union, Articles 165 and 166, states that the EU contributes to the development of quality education and vocational training by encouraging cooperation between Member States and, if necessary, supporting and supplementing their action, while respecting the responsibility of the Member States for the content of teaching and the organisation of education systems and their cultural and linguistic diversity. The main working method the European Commission uses to support this goal is the Open Method of Coordination.

The European Strategy for Smart, Sustainable and Inclusive Growth (Europe 2020) sets outs clear targets for education and training: reducing the share of early school leavers to less than 10% and increasing the higher education attainment rates to at least 40% among the 30-34 year olds. The Europe 2020 strategy addresses these targets through two flagship initiatives: *Youth on the Move* that recently ended and *An Agenda for New Skills and New Jobs*, and through targeted initiatives such as *The grand Coalition for Digital Jobs*.

The Europe Education and Training 2020 Strategic Framework (ET 2020) guides European policy in education and training in more detail. Its main aims are:

- Making lifelong learning and mobility a reality;
- Improving the quality and efficiency of education and training;
- Promoting equity, social cohesion and active citizenship; and
- Enhancing creativity and innovation, including entrepreneurship, at all levels of education and training.

In a policy paper *Rethinking Education* published in 2012, the Commission stresses the importance of transversal skills such as entrepreneurial skills and skills related to science, technology, engineering and mathematics (STEM). The more recent Juncker Strategy also stresses the importance of skill attainment. Moreover, the strategy indicates that mobility should be encouraged in fields with persistent vacancies and skills mismatches. In summary, increasing the attainment of transversal skills through formal, informal and non-formal means throughout the whole educational trajectory – from elementary school to life long learning – is a high strategic priority for the European Union.

1.4 The data process: from data collection to analysis and visualisation

Data sources

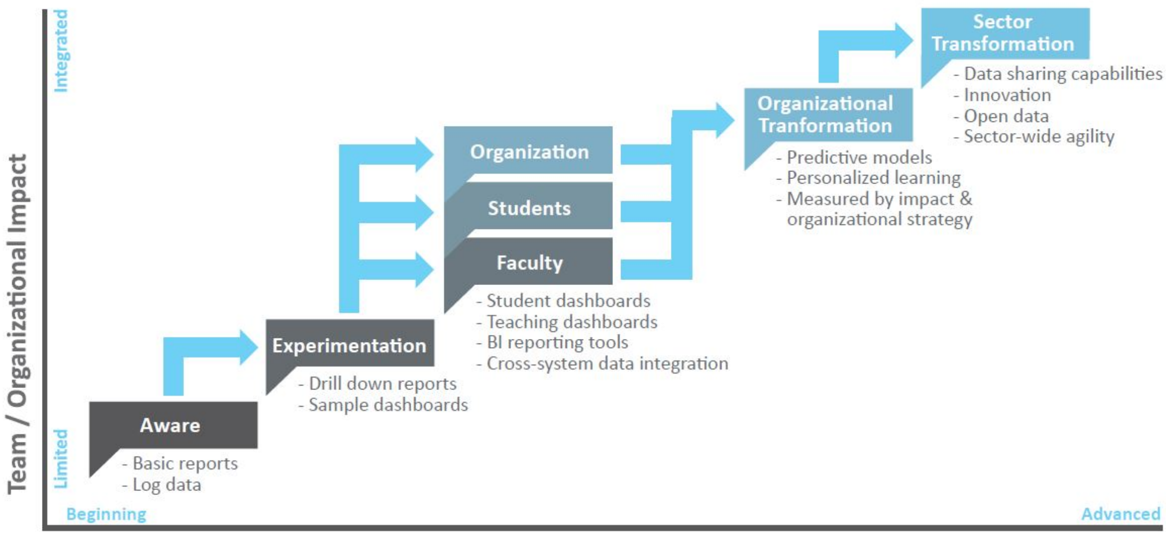
The main source for this use case would be **institutional learning analytics datasets**. In addition, social media data, LinkedIn data or data on job markets for ICT positions can be of use (e.g. based on the MOVIP pilot launched by the EC's Directorate-General CONNECT).

As these datasets regard micro-data, their use at an aggregated level (national or European) depends on the linking of datasets held at the institutional levels by means of **common identifiers**. Standardisation of the variables against which data are collected and indicators to use is therefore a critical condition in this context. In other words, coordination is required among and between the different levels in the system (institutional, national, European).

It should be noted, however, that the use of learning analytics in higher education institutions is still **relatively new and underdeveloped**. Bichsel[6] (2012) states, "Interest in analytics is high, but many institutions have yet to make progress beyond basic reporting." While progress has been made in these last years, consensus in the literature is that the majority of institutions are still at the awareness or experimentation stage in the Learning Analytics sophistication model (see the Figure below).[7]

The question is therefore whether learning analytics at the level of *organisations* will have matured sufficiently to be of use, in the near future, for *policymakers* at national and European level.

Figure 4 Maturity of Learning Analytics Deployment



Source: Siemens, G., Dawson, S., & Lynch, G. (2014). Improving the Quality and Productivity of the Higher Education Sector: Policy and Strategy for Systems-Level Deployment of Learning Analytics. Canberra, Australia: Office of Learning and Teaching, Australian Government. Retrieved from http://solaresearch.org/Policy_Strategy_Analytics.pdf

---

[6] Bichsel, J. (2012). Analytics in higher education: Benefits, barriers, progress, and recommendations. EDUCAUSE Center for Applied Research.

[7] An illustration is that in the UK, one of the leading countries in learning analytics, progress is steady but slow. http://repository.jisc.ac.uk/5657/1/Learning_analytics_report.pdf

Taking a broader perspective to the topic of investigation, i.e. micro-data on skill attainment, one should consider that individuals may attain skills through formal, non-formal and informal learning.

There are various data sources that can and are used to map skill attainment at the micro-level. These include: curriculum & learning outcomes of modules, online learning course completion and all non-formal learning outcomes. These data sources are discussed below.

Entrepreneurial and ICT skills can be part of **formal education** in certain courses of modules. Sometimes a whole programme is dedicated to these skills, such as computer science of entrepreneurship education. The graduation rates in these programmes are often used to assess skill attainment in these areas. However, many other programmes incorporate entrepreneurial and ICT skills in their modules. Moreover, students may complete certain modules, attain certain skills but not receive a diploma. Thus, to understand the level of transversal skills in the EU, also the mapping of the skills attained in these modules may be useful, in addition to the mapping of traditional programmes.

Developing a proxy for **skill attainment in modules** is not a straightforward exercise. The curriculum of modules could be an indicator. The challenge is to define what curriculum components contribute to a certain skill, and then measure this across a very differentiated set of modules. Thus coordination on the availability of data on modules and on the semantics of curriculum components is essential. Ideally this mapping of modules would be in line with the current developments in learning outcomes. In the last decades the European Commission has made considerable steps in the development of qualification frameworks based on learning outcomes (e.g. the European Qualification Framework and the European Credit System for Vocational Education & Training). These developments provide opportunities to use learning outcomes frameworks as a basis to map skill attainment. As this is still work in progress, it also poses some limitations. At the same time, it provides the opportunity to incorporate learning analytics into the discussion of learning outcomes. The next step would be to not only map the number of modules with certain learning outcomes completed, but also the degree to which learning outcomes are attained (e.g. by mapping grade data).

Another new data source that maps the attainment of transversal skills is **formal online learning.** There is an increasing supply of online courses and training material (e.g. MOOCs and online coding platforms). A limitation of this type of learning is the lack of accreditation of learning outcomes. However, initiatives exist that certify online courses or offer online badges for skill attainment.[8] These initiatives provide opportunities to map course completion. It is important to note that these online activities might not indicate an increase of skill attainment but a switch from offline to online learning resources. However, data generated by these online platforms is valuable to understand what activities are being undertaken and where, instead of the actual increase in volume

Finally, also all forms of transversal skill attainment in **non-formal learning** could be taken into account. Non-formal education is any type of deliberate learning activity (e.g. watching an online video, following a workshop or tutorial) that does not lead to accreditation in any form. It is thus hard to measure these activities. However, the rise of non-formal learning, using online platforms such as YouTube, has created opportunities for gathering data on the volume and type of learning activity.

[8] See for example the EU co-funded initiative: http://www.learningbadges.eu

Data analytics and visualisation

The data generated from learning analytics can be analysed in multiple ways. First, **descriptive statistics** on topics, learning outcomes, number of modules will give insight into the volume and type of skills attained. To map learning outcomes efficiently, **text mining** of curriculum information could be applied. When data is generated over time, **trend analysis** and **pattern detection** can be used to map the dynamics over time. However, attention needs to be paid to the attribution of trends in the data to actual trends in skill attainment. Finally, **predictive analysis** may be used to gain insight into future trends. For example, data on skill attainment in early higher education might be indicative of future study choices.

There are multiple **data visualisation** methods that can map skill attainment levels and trends, e.g. trend figures, geographical maps, dashboards etc. It is important to note, though, that visualisation is not a neutral representation of the data. A visualised image always conveys a certain story, especially when it features trends. As using learning analytics for international policymaking is a novel activity and the data sources are granular and inconsistent, one should be careful not to misinterpret the data. Therefore, it is important to first understand the context of the data behind the visualisation and to dedicate the needed attention to including context characteristics among the variables for visualisation.

1.5 Reflections on challenges and next steps

There are a number of considerable challenges for the implementation of this use case; one may even consider that seeing the current state of play, this use case is more hypothetical than realistically feasible in the next following years. The major challenges are as follows:

- A first challenge is in the **availability of learning analytics data sets**. This challenge is due to the relatively new and under-developed culture of learning analytics and data collection beyond traditional assessment in the national higher education systems.

- Secondly, **privacy concerns for using micro-data** remains a major challenge despite current initiatives to overcome refusals for sharing or aggregation by higher education institutions. Privacy issues in individual data related to education are not unlike privacy discussions in the field of health care. Individuals may not be willing to consent to sharing their data if they fear that it is accessible to organisations that may impact their personal lives, including their careers.

- Thirdly, a challenge that this use case faces - and which is common to all initiatives building on an aggregation of micro-data - is the **coherence** between data sets in terms of scope and definitions. Micro-data sets such as the ones resulting from learning analytics are built to satisfy the needs at an institutional level; this does not always coincide with the needs and intended use of these data at national or European level. Closely related to this topic is the challenge of **comparability** of the data depending on standardisation, their **consistency** over time, and their overall **reliability** and **quality**.

- Another issue that should be addressed is the **inclusion** of countries and regions with a weak educational data system or lower levels of online learning. Learning analytics relies on strong educational data systems, the right skills to extract data and the actual use of these systems as well as online educational platforms. These prerequisites will differ between countries and between regions. Consequently the data gathered might not reflect reality due to these differences.

Notwithstanding these challenges, learning analytics has large untapped potential for policymakers and gaining insight into the processes that allow for the attainment of transversal skills, beyond the formal learning modules, can indeed be of high importance for improved policymaking on the matter.

In case the EC wishes to investigate further the potential for this use case, we recommend the launch of a small pilot, focused on a limited number of educational institutes on the forefront of learning analytics. Taking into consideration that the data collection process would be a novel activity, requiring strong coordination, and a steep learning curve, a large-scale pilot would not be feasible.

Further reading

Related projects: LACE and application-centric projects LEA'S BOX, Pelars, WatchMe.

P. MacFadyen, S. Dawson, A. Pardo, D. Gasevic, 'Embracing Big Data in Complex Educational Systems: The Learning Analytics Imperative and the Policy Change' in *Research & Practice in Assessment,* Vol. 9, Winter 2014

Evidence hub *for learning analytics*

Learning analytics are more than measurement by Prof. D. Gasevic