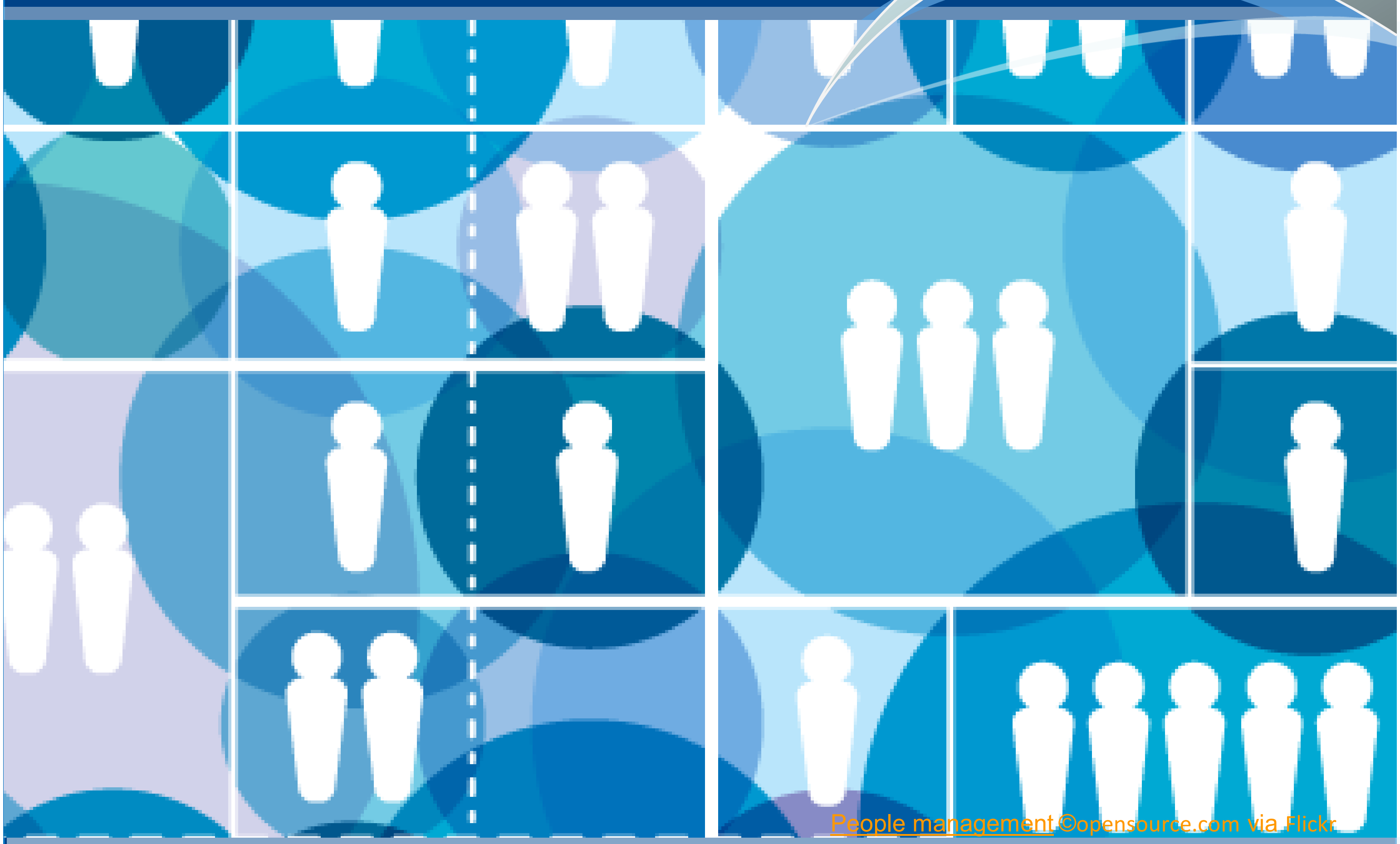
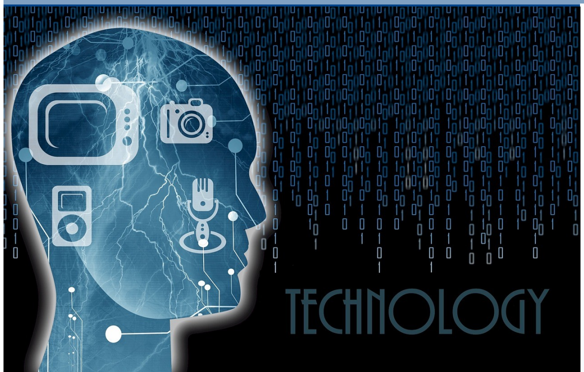
INTERNATIONAL COUNCIL FOR OPEN AND DISTANCE EDUCATION



People management ©opensource.com via Flickr

Higher Education and the Revolution of Learning Analytics



Anne Boyer
and
Geoffray Bonnin

"Learn from the masses, and then teach them."
Mao

ISBN: 978-82-93172-38-3

Digital tools everywhere, even in education...

The always broader usage of digital tools in everyday life as well as in all educational sectors leads to large or even very large sets of digital user-generated traces, collected via online pedagogical platforms, academic enrolment, libraries information systems, online assessment, social networks, etc. Each click on a webpage, each social interaction, each page accessed online, each use of an app on any mobile device leaves information that can be collected and stored to produce what is being called digital footprints.

Simultaneously, "big data" and Analytics technologies allow the exploitation and mining of increasingly large datasets. As in many other domains, such as business, e-commerce or e-health, the sector of higher education is discovering the impact that Analytics can have on its development, and how it can help institutions to manage their transition towards a University of the Future. Indeed, learning Analytics aims at providing educational leaders with data-founded models, helping them to improve the teaching and learning efficiency and quality. In that sense, Learning Analytics can be viewed as a potential foundation for a systemic change.

According to ECAR's[1] 2012 report[2], learning Analytics has become a hot topic in higher education, as there is a variety of examples of its successful use in many institutions. It is not only an operational set of tools to implement in an institution to achieve better performances, it is also a data driven support to educational strategies for all the stakeholders, from the managers to the students, and finally a very active and promising research field.

1 Educause Center for Applied Research.

2.https://library.educause.edu/resources/2012/6/2012-ecar-study-of-analytics-in-higher-education

...lead to learning analytics...

Learning Analytics is the emerging field defined as “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” (Long & Siemens, 2011)[3]. Learning Analytics is also defined by (CETIS[4], 2012)[5] as “the process of developing actionable insights through problem definition and the application of statistical models and analysis against existing and/or simulated future data, [allowing] institutions to experiment with data to gain insight, to improve the student learning experience and student outcomes, and identify improvements in efficiencies and effectiveness of provision.” This second definition insists on some major characteristics of the research field, such as the suggestion of actionable insight, allowing higher education institutions not only to better know what happens and understand what has happened, but also to predict with a high level of confidence what will happen or to recommend actions on how make something happen. All these aspects, which are transversal to Analytics in general, are well synthetized by the Gartner group, as shown in Figure 1.

Learning Analytics can be viewed as Analytics applied to educational data, four forms of learning Analytics can be of interest for a higher education institution:

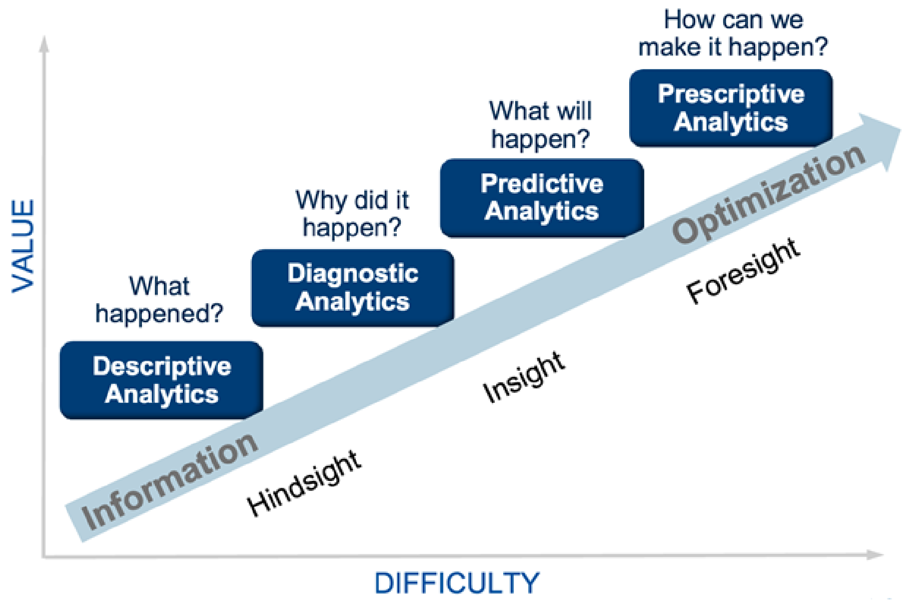


Figure 1 – Analytics (from Gartner)

- *Descriptive Analytics*, answering the question “what happened?”. Descriptive Analytics is the examination of data or digital content, characterized by basic computational or statistical techniques and visualizations such as pie charts, tables, bar charts or line graphs. Most of the time, they are aggregated within a dashboard with pre-defined metrics and indicators. This kind of Analytics can be used at different levels within the university, for example by the student to score his positioning in terms of performances, work load or engagement relatively to the others; or by the teacher, to have an idea of the impact of his pedagogy via assessment results, or time allocated by students to achieve a specific activity; by the institution, to have a realistic picture of the dropouts.

- *Diagnostic Analytics*, answering the question “why did it happen?”. Diagnostic Analytics examines data or content to better understand what events or features can explain the current situation. It is most of the time characterized by techniques such as data discovery, pattern mining or statistical correlations. Diagnostic Analytics allows understanding which event or combination of events lead to the current situation. Such Analytics can be used by students to discover what could explain why they succeed or fail; by teachers, to exhibit real learning paths and compare them with the *a priori* scheme they had in mind; or by the institution to benchmark impacts of specific actions on students achievement, such as extending opening hours of libraries, developing blended learning...

3. "Penetrating the Fog: Analytics in Learning and Education", by Phil Long and George Siemens September/October 2011 EDUCAUSE review, available at the address https://net.educause.edu/ir/library/pdf/erm1151.pdf

4 Centre for Educational Technology, Interoperability and Standards.

5 Cetis Analytics Series: "Case Study, Acting on Assessment Analytics" by Sheila MacNeill (Cetis) and Dr Cath Ellis (University of Huddersfield). April 2013, Cetis Analytics Se-

- *Predictive Analytics*, answering the question “what will happen?”. Predictive Analytics aims at anticipating a near future based on the exploitation of past events. It has to put the emphasis on the relevance of the resulting insights (confidence measures) and has to be easy to use by anyone, student, educator or decision maker. There is a clear need for designing tools accessible to a great range of users, from experts to occasional users. Predictive Analytics is a promising field, for all the actors of the learning process. It allows the student to know if he is working in a right direction to achieve his learning goal; the teacher to determine which students are at-risk, and then to intervene when it has still an impact on their success; the institution, to determine which courses to open next year.
- *Prescriptive Analytics*, answering the question “how can we make it happen?”. Prescriptive Analytics is another form of advanced Analytics, which examines data or digital content to determine realistic and efficient actions to plan to achieve a given goal. It is characterized by tools able to process big graph analysis, text and data mining, simulation or recommendation determination. It consists in the elicitation of future actions that should be performed to achieve a specific goal. This kind of Analytics can be provided to students to draw a personalized learning path; to a teacher to suggest relevant activities to his students; to the institution to identify trends of dropouts and to act before they occur.

Learning Analytics differs from academics Analytics, defined as “the process for providing higher education institutions with the data necessary to support operational and financial decision making.[6]” Analytics in general are powerful processes of data analysis and/or data mining to support informed decision-making. Learning Analytics is related specifically to data analysis and/or mining with the objective of addressing teaching and learning goals, while academic Analytics focus on data analysis and/or mining with the objective of making better operational and financial decisions.

Learning Analytics is a dynamic interdisciplinary field that gathers researchers from various disciplines, e.g., artificial intelligence, psychology, or educational sciences, with various stakeholders, e.g., practitioners, experts in legal issues, HEIs leaders, computer scientists, teachers and students.

6 van Barneveld, Arnold, and Campbell, "Analytics in Higher Education" quoted by Leveraging Analytics in Community Colleges by Treca Stark, Published Monday, September 14, 2015 on the Educause website.

...for a better learning efficiency...

Learning Analytics plays an important role in the improvement of the global quality and efficiency of learning and teaching, via a better understanding of the educational process, a more comprehensive student assessment and the personalization of education. As shown by a study conducted by Hanover Research[7], and mentioned in the Educause Horizon report 2016[8], "students have a desire for immediate and continual feedback as they learn. The findings indicate that almost two-thirds of the participating students believe the impact of Analytics reports on their academic performance is 'very positive'."

Extrapolating from some case studies made internationally, JISC[9] anticipates that learning Analytics could make significant contributions:

- To improve quality: educators could improve their own practice based on the information provided by Analytics. Learning Analytics can provide feedback to the teaching staff about the quality of their pedagogical content, the impact of the activities they provide, and their assessment process, to enable a continuous enhancement. Analytics can also be used by educators to monitor in real time the performance of their students and to adapt their teaching if, for example, they identify that students are struggling with some particular topic.

  Giving better information to students about how they are progressing and what they need to do to meet their educational goals is another important application. Learning Analytics is the opportunity for students to take control of their own learning, as it informs them about their current performance and helps them to make informed choices about what to study.
- To boost the retention rate: Analytics allows institutions to identify at-risk students at an early stage and to intervene with advices, additional materials and alternative activities, and to support them as long as it is still possible for them to recover.
- To enable the development of adaptive learning: Adaptive learning is emerging to help students to develop skills and knowledge in a more personalized and self-paced way. Learning Analytics enables the educational content to be tailored to the adequate level of understanding as they progress through it.

...as many successful projects tell us...

The ECAR[10] 2012 report mentions for example the Western Governors University and its assessment-based coaching reports to develop customized study plans for students, or the Paul Smith College and its early-alert program.

Many institutions around the world develop and implement learning Analytics projects. The literature reports many experiments that show how imaginative and innovative the higher education institutions are when they cope with challenges such as learning outcomes improvement, students' retention or pedagogical efficiency .

7 https://library.educause.edu/~/media/files/library/2016/2/hr2016.pdf

8 Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., and Hall, C. (2016). NMC Horizon Report: 2016 Higher Education Edition. Austin, Texas: The New Media Consortium

9 https://www.jisc.ac.uk/reports/learning-analytics-in-higher-education

10 Educause Center for Applied Research

An increasing number of learning Analytics projects are starting to reveal promising results. Such projects develop and integrate dashboards, visual representations of data in the institutional LMS, to personalize the learning experience. For instance, as can be found in the Educause Horizon report 2016[11] , The A4 learning project[12] (Universidad Internacional de La Rioja) combines data techniques with information visualization, providing each student with ongoing information that enables them to think critically about their learning and their goals. This 100% online institution also develops the iLIME project[13], whose purpose is to elaborate and implement an automatized itinerary-recommendation system aiming at helping teachers to make personalized recommendations to their students. The iLIME eLearning model has been designed for personalized learning, with a special focus on the combination of formal and informal settings.

Rio Salado College, Harvard University, and Austin Peay State University are three examples among others given in the Educause website[14] to illustrate how the use of learning Analytics enables to support student success and institutional effectiveness.

Rio Salado College[15] (part of Maricopa Community Colleges[16]) implements learning Analytics tools to predict at-risk behavior from activity factors, such as log-ins or site engagement, and to provide them with intervening responses. It operates a model that includes an in-house-developed course management system, some accelerated educational programs, and the ability for students to begin a course nearly any week of the year. Students have access to a dashboard, RioCompass, to monitor their progress toward degree completion, and instructors have access to a dashboard to regularly monitor available student Analytics. According to the Educause review[17], it allows the instructors to address student needs more expeditiously and aids in promoting retention.

Harvard University[18] mines classroom data using a system called Learning Catalytics[19], with the originality that it supports peer instruction. A basic pattern of learning/teaching (depending if you are a student or a teacher) is the following: "students log in to their interactive classroom, the professor produces a problem generated on the students' screen for them to answer; the system analyzes the answers and decides how to pair students as study partners; messaging appears on the students' screens indicating who they will work with; and the professor receives a map of information displaying how the students did so the professor can decide who to assist"[20] . Such a system allows the constitution of pertinent and data-founded students workgroups, thus enhancing efficient collaboration between learners, and enables educators to adjust instructional practices to provide learners with customized support.

The 2016 JISC report[21] summarizes in two informative maps (Figure 2 and Figure 3) some learning Analytics projects in the United Kingdom, Australia and the United States. All the projects referenced in these two maps show the variety and diversity of the implementation of learning Analytics, working towards a university vision and supporting its strategy.

11 Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., and Hall, C. (2016). *NMC Horizon Report: 8672 Higher Education Edition*. Austin, Texas: The New Media Consortium

12 go.nmc.org/iLime

13 http://research.unir.net/blog/ilime-operational-implementation-of-a-recommendation-model-for-informal-and-formal-learning/

14 van Barneveld, Arnold, and Campbell, "Analytics in Higher Education" quoted by Leveraging Analytics in Community Colleges by Treca Stark, Published Monday, September 14, 2015 on the Educause website.

15 http://www.riosalado.edu/

16 https://www2.maricopa.edu/

17 Leveraging Analytics in Community Colleges by Treca Stark, Published Monday, September 14, 2015 on the Educause website

18 http://www.harvard.edu/

19 http://atg.fas.harvard.edu/learning-catalytics

20 Leveraging Analytics in Community Colleges by Treca Stark, Published Monday, September 14, 2015 on the Educause website

21 https://www.jisc.ac.uk/reports/learning-analytics-in-higher-education

… with some key aspects for a safe and ethical use of learning analytics to keep in mind

Learning Analytics works with user-generated data and could be seen as a way of tracking students. A strong key issue is related to safety, privacy and security. For example, questions such as:

- Who owns the student data?
- What are the rights and responsibilities of the institutions, researchers and teams that use these data?
- What is an ethical practice and what is not?

It also highlights the question of the meaning of privacy nowadays, and the need to determine what is affordable or not. Other critical perspectives are related to student profiling, which could lead to biased behaviors and expectations from teachers. For instance, could a learning Analytics tool bias the behavior of a teacher when interacting with a student identified as "at-risk"?

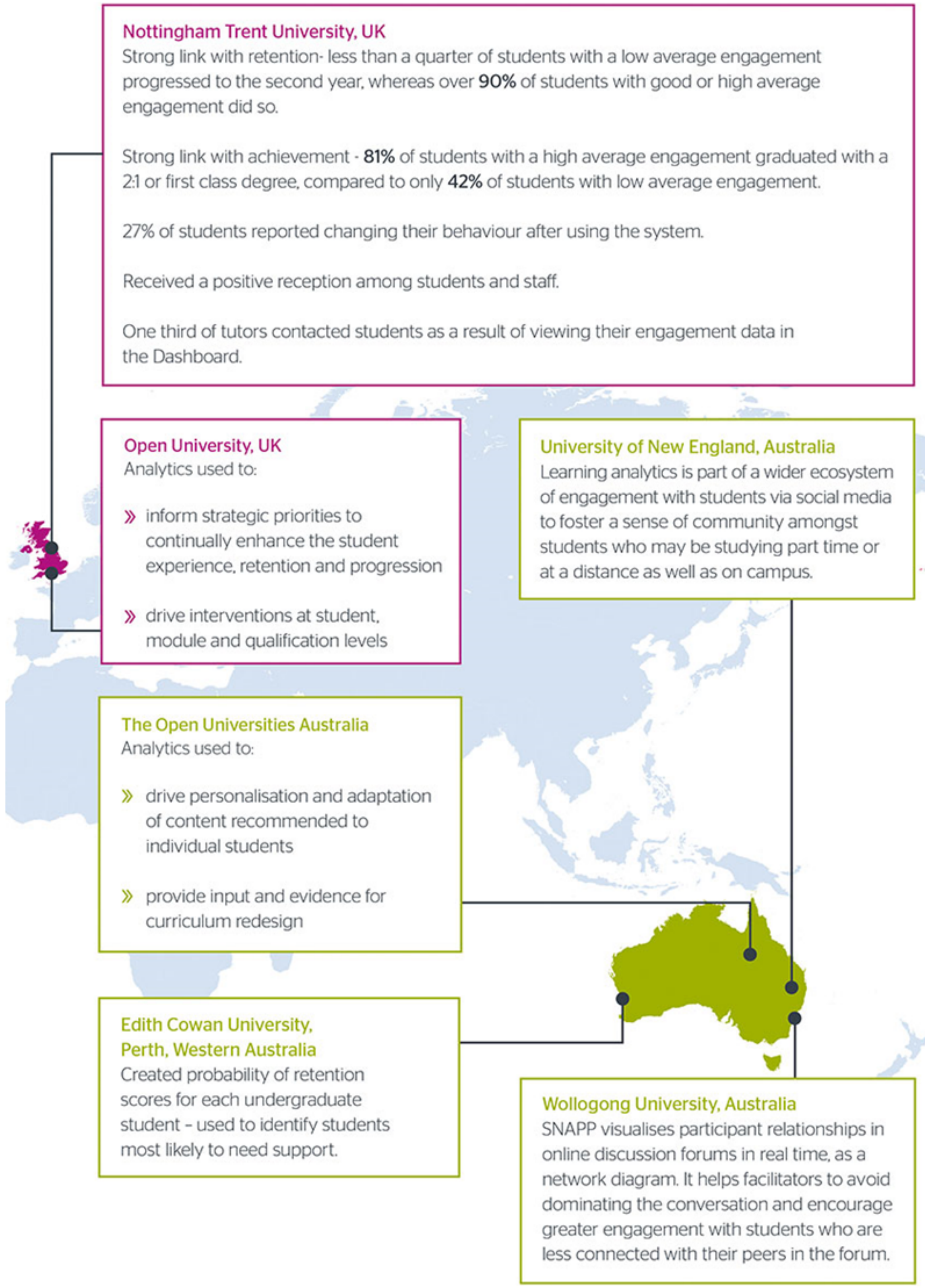


Figure 2 – Some projects in UK and Australia, from the JISC report

Arguments against learning Analytics include that counting clicks could lead to standardization of higher education. Transparency and communication are central aspects to consider when implementing a learning Analytics initiative. But using Analytics requires that we think carefully about what we need to know and what data is most likely to tell us that.

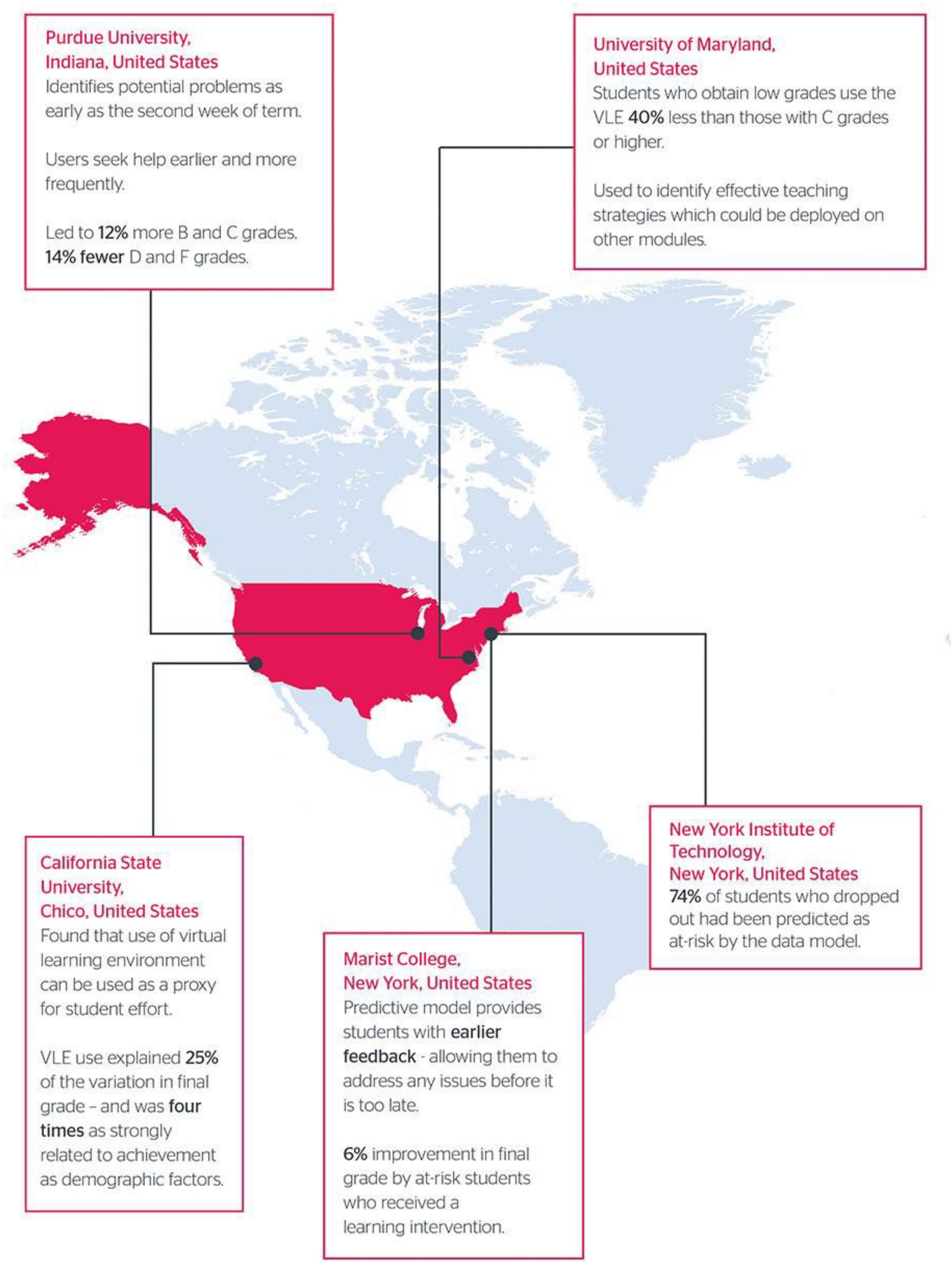


Figure 3 – Some projects in the USA, from the JISC report

… a cultural Revolution.

A large number of learning Analytics projects are currently being launched in many institutions all around the world; and the results are very promising, as it becomes obvious when considering both Figure 2 and Figure 3 from the JISC report.

Learning Analytics is not only beneficial for the learners and the teachers. Some research of the domain indicates that institutions can also take advantage of using its techniques.

For instance, Harrison et al. (2016)[22] measured the financial benefit for an institution when some notification system for the risk of failure is provided to the students. Their case study is a mostly distance university in Australia of about 20,000 students. They found that on average, each time a student dropped out, the university loses about $4,500, and that, on average, students that were identified as being at risk and to whom a notification was sent during the first year of study induced about $4000 in revenue, during the second year $2,675 and during the third year $3,479.

Another example is the work of Ménez et al. (2014)[23], which presents several learning Analytics techniques that can be used to help program coordinators to elaborate the curricula of their institution. They do so by exploiting academic history detailed data of an Ecuadorian university, which spans from 1978 to 2012. A first set of techniques are presented that allow to estimate the difficulty of some courses relatively to others, dependencies between courses, and the curriculum overall coherence. Then another technique is presented to determine which courses lead to student drop-out most often. Last, the performance of the students is analysed relatively to the number of courses they took, and their respective difficulty. Using these techniques, the authors discovered that some problematic courses could clearly be identified as well as some workloads that were clearly too high for most of the students. Interestingly, they also found that the lack of good results of students in the most difficult prerequisite courses did not have a very strong influence on the subsequent student results.

The paper written by Slade et. al.[24] explains how Learning Analytics enables educational institutions to gain knowledge on the learning behavior of their students, use it to improve the student retention rate and perform in-time student success.

While so many innovative projects are being launched, the effective deployment of learning Analytics seems to remain a long March in many universities. Most of the time, Analytics are compartmented to specific areas such as student retention, detection of at-risk students, and improvement of learner experience or learner-teacher interactions. As said by Mike Sharkey[25], VP of Analytics, “a successful Analytics implementation is a cultural challenge, not a technological one.” Creating a data driven culture within Higher Education institutions is not so easy.

22 Harrison, Scott, et al. "Measuring financial implications of an early alert system." Proceedings of the Sixth International Conference on Learning Analytics & Knowledge. ACM, 2016.

23 Méndez, Gonzalo, Xavier Ochoa, and Katherine Chiluiza. "Techniques for data-driven curriculum analysis." Proceedings of the Fourth International Conference on Learning Analytics And Knowledge. ACM, 2014.

24 Sharon Slade, Prinsloo Paul, Learning Analytics: ethical issues and dilemmas, American Behavioral Scientist, Oct 2013, Volume 57, No. 10, pp 1510-1529, DOI: 10.1177/0002764213479366.

25 quoted in https://Analytics.jiscinvolve.org/wp/files/2016/07/Jisc-LAN-Newman-Kevin-Mayles.pdf

1. How to dive into learning Analytics: little red book for decision makers

1.1. Ready for learning Analytics?

Learning Analytics can be considered from different points of view or levels depending on whether one is interested in a specific course, an education department, a university or a region. (Buckingham Shum, 2012)[26] organizes them into micro, meso, and macro Analytics. Each level has access to different sets in quantity and diversity of available data and contexts. They respond to different questions that provide specific views on the organizational level considered. (Chatti et al., 2014)[27] offers a reference model for learning Analytics based on four dimensions, which can be summarized by four questions: what (which types of data have to be collected, managed and used, who (which actors are concerned, who will receive the results), why (in which objectives do you analyze the data collected), and how (what methods are used to analyze the collected data). Educational institutions are interested in learning Analytics for different types of stakeholders, leading to a variety of objectives, such as:

- For the learner, question its achievements and its mode of behavior against other students;
- For the teachers, know at the earliest students calling for greater support or special monitoring;
- For teaching teams, improve the quality of education provided in the course and relevance of the developed materials;
- For training managers, decide which further actions are adapted to newly arrived audiences and adapt curricula.

Figure 4, from (Chatti, 2014) [28], summarizes the four dimensions of the model.

Thus, if we consider a specific course, it may be appropriate to focus on the level of student engagement in the material or activity level, while at the level of a department we may be more concerned with the detection of at-risk students, interventions and supports to implement. As pointed out in (Powell & MacNeil, 2012)[29], the various stakeholders may have different objectives in terms also of timescale (immediate need for the learner to medium term for the institution) and may require data, approaches and various tools. Just as an organizational level change results in a change of learning Analytics objectives, it also led to the use of different tools, which can be the analysis of social media for a class or trend analysis for a department. However, if the contributions of learning Analytics vary depending on the stakeholders (students, teachers, education officials, training administration), the learners are the first beneficiaries. In particular, it allows new uses, allowing them to follow highly customized training devices, tailored to their needs in real time, and get a "tailor-made" large scale to learn with a device capable of detecting the commitment deficits, and to take action to put the learner in a learning situation. These benefits are designed to have the same goal for the learner: optimize the acquisition and retention of knowledge.

26 Buckingham Shum S., « Learning Analytics ». UNESCO Policy brief.

Available at http://iite.unesco/pics/publications/en/files/3214711.pdf, 2012.

27 Chatti, M. A., Lukarov, V., Thüs, H., Muslim, A., Yousef, A. M. F., Wahid, U., Greven, C., Chakrabarti, A., Schroeder, U., « Learning Analytics: Challenges and Future Research Directions », eleed Iss. 10. Available at http://eleed.campussource.de/archive/10/4035, 2014.

28 Chatti, M. A., Lukarov, V., Thüs, H., Muslim, A., Yousef, A. M. F., Wahid, U., Greven, C., Chakrabarti, A., Schroeder, U., « Learning Analytics: Challenges and Future Research Directions », eleed Iss. 10. Available at http://eleed.campussource.de/archive/10/4035, 2014.

29 Powell S., MacNeil S., « Institutional Readiness for Analytics A Briefing Paper », CETIS Analytics Series, JISC CETIS, available at http://publications.cetis.ac.uk/wp-content/

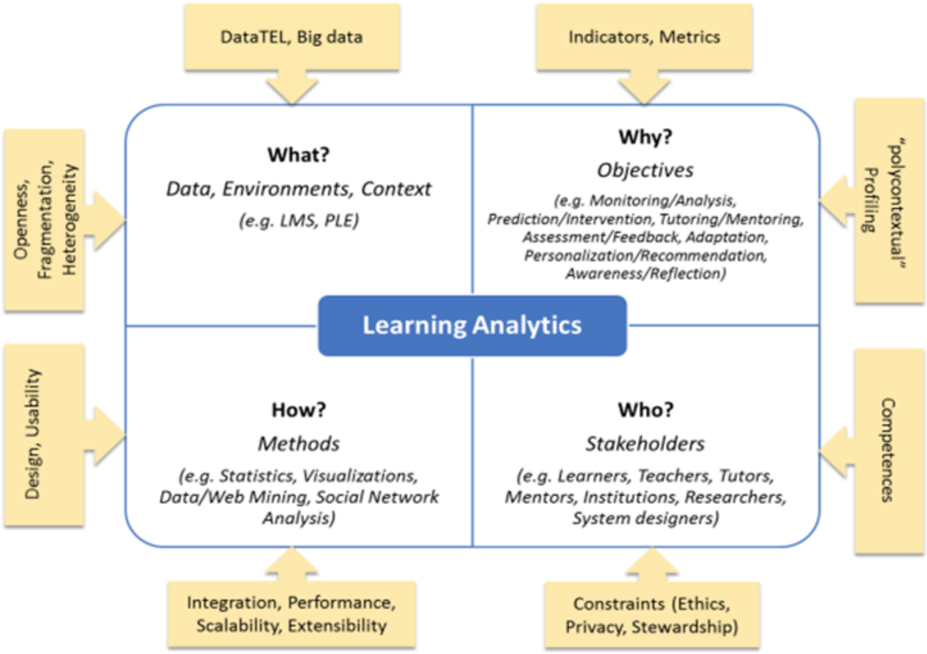


Figure 4 – Learning Analytics dimensions (from Chatti 2014)

1.2. Diving into a learning Analytics project

This sub-section aims at describing some thoughts or experiences gathered in the literature of the field about concrete implementations learning Analytics projects in higher education institutions. The main point of focus is the ability of institutions to carry out such projects while taking into account ethical considerations. This raises the question of the dynamics of adoption, and the inclusion of learning Analytics in the piloting of an institution.

When an institution decides to engage in the implementation of some learning Analytics tool, a first step may be to determine if the institution is ready: the implementation of learning Analytics involves financial and human investments, which induces the necessity to review the key elements that may contribute to the success of the initiative. A maturity indicator, along with a tool to evaluate it, is available at the following address: http://www.educause.edu/ecar/research-publications/ecar-analytics-maturity-index-higher-education

It was developed based on factor analysis of responses to the ECAR survey "Analytics in Higher Education." It consists in focus groups and discussions with academic leaders in information technology and institutional researchers, and domain experts. The ECAR maturity index for Analytics in higher education uses six dimensions (see Figure 5). The level of the institution is evaluated for each dimension by a score ranging from 1 to 5. A composite score representing the average of the six values is also provided. ECAR indicates that anonymity is ensured.

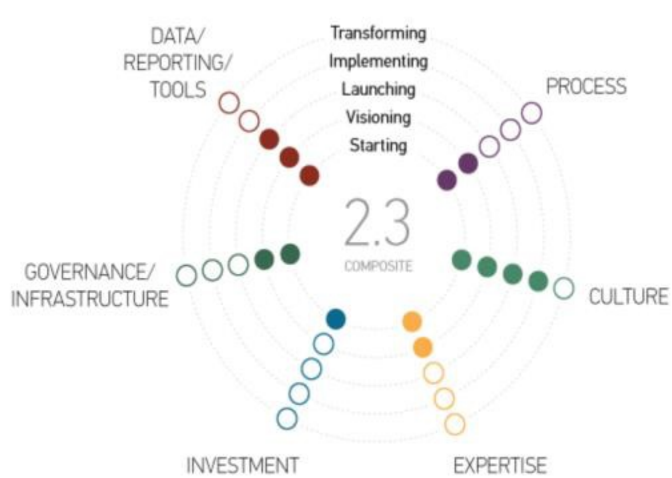


Figure 5 – The ECAR maturity index

(Arnold et al. 2014)[30] propose to revisit this indicator and introduce a new indicator called LARI (Learning Analytics Readiness Instrument). This composite indicator (see Figure 6) is made of 90 items grouped in five factors: (1) ability (expertise), (2) data, (3) governance and infrastructure, (4) culture and processes and (5) overall perception of maturity. The previous dimension of investment is removed, "culture" and "process" are merged in the same dimension and a new dimension is included (perception of maturity).

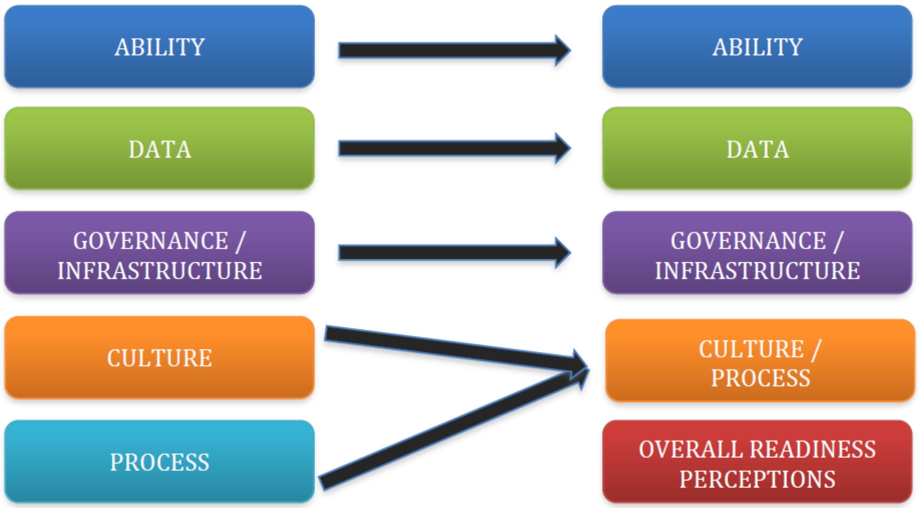


Figure 6 – The LARI tool

Once it has been determined that the institution is ready for implementing some learning Analytics tool, the development phase can begin, and different additional considerations can be taken into account. Figure 7 from the CETIS report[31] illustrates three key institutional considerations when developing a project of learning Analytics in a Higher Education institution that is "ready" for learning Analytics, namely:

- Provision of data, from various sources and of various levels of quality. This step may require specific developments to organize the collection and storage of data.
- Interpretation and visualization of data: use of the available data for reporting the relevance of the activity and representation of the information to make it accessible, precise and adapted to the user needs.
- Actionable information: what can lead to the implementation of concrete actions?

These three considerations correspond to three phases, each of which corresponds to different actors. The technical team is in charge of the first phase, analysts are in charge of the second and teachers and students are primarily in charge of the last. These phases are interconnected, and the concerned teams can work independently of each other.

30 Arnold K., Lonn S., Pistilli M., «An exercise in institutional reflection: the Learning Analytics Readiness Instrument LARI », in Proc. LAK 2014, Indianapolis, USA, 2014.
31 Cetis Analytics Series:

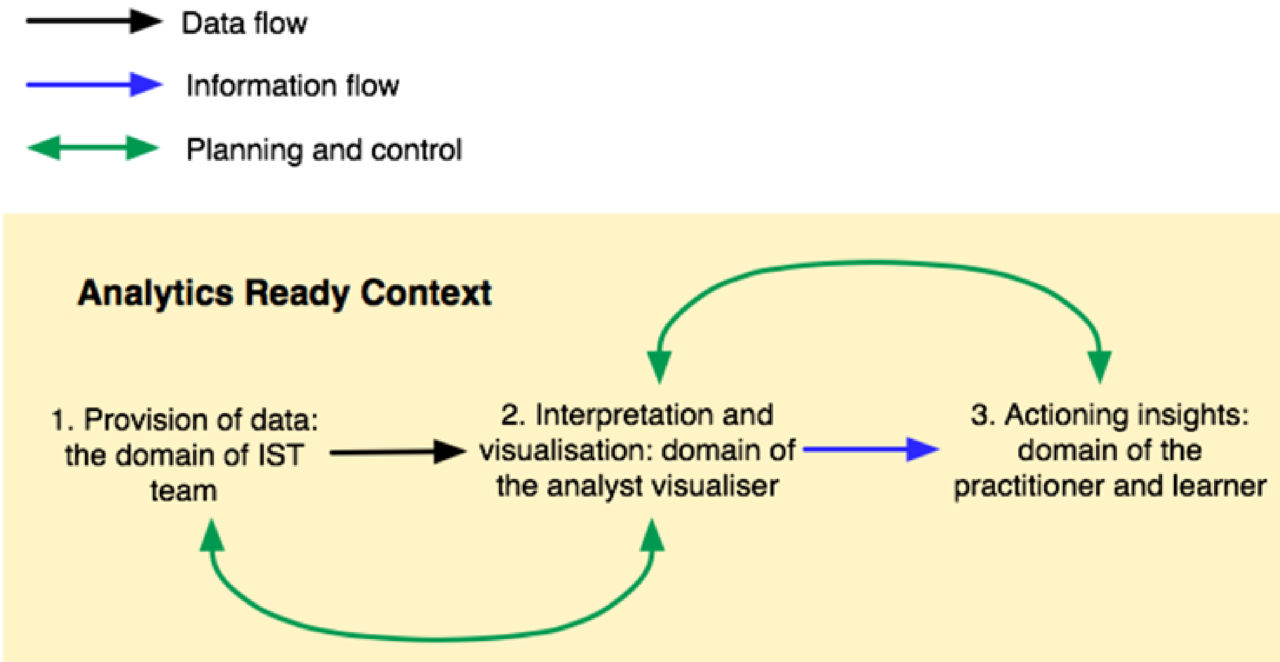


Figure 7 – The three phases of the project

In July 2012, Malcolm Brown, Director of the EDUCAUSE Learning Initiative, wrote in a newsletter: "any institutional program involving learning Analytics must have both: a robust technology to harness and analyze data and effective plans and processes for acting on the results of the analysis." Jack Neill, Sr. Director of Data Analysis at the University of Maryland, in his speech at the pre-conference "Practical Learning Analytics" at EDUCAUSE 2015, describes the data flow according to Figure 8: the data are first collected from multiple and diverse sources, they are then normalized and stored (in centralized way to UMUC), the results of learning Analytics tools are then transmitted to users via reports, indicators or dashboards.

Among other issues, whether to conduct an internal development using an open platform or to acquire a market tool, possibly with a specific integration, is a difficult choice. To clarify this choice, we must at least ask the following four questions:

- Is the problem I am facing really specific? And if it is, to what degree?
- What skills needed for the project do we have in-house? Are they adequate and available to complete the project?
- Which technologies (databases, data warehouse, technical infrastructure...) do we already have in-house?
- How much time is required to develop the data infrastructure?

While many companies offer tools to implement learning Analytics, these are mostly not open software. It is important to be vigilant on a number of points: (1) the transparency of the tools used (avoid the black boxes that come out of Analytics without knowing the method of calculation, and thus avoid obtaining uninterpretable results), (2) the ability to scale tools (crucial point in learning Analytics, where the data flow is enormous) and process multiple heterogeneous data, (3) the nature, quality and customization of the results made to allow to adapt outputs of tools to the real needs of teachers and students and to its ability to evolve, and of course (4) the cost.

In short, a reasonable position could be to:

- Conduct by yourself: if one does not wish to do what you expect, if you cannot do otherwise, and if you possess the necessary human resources and competences.
- Buy: if the market tools realize exactly what you want, if you do not have the internal development expertise, or if you do not have the time or the means.
- Participate in collaborative development, or join a consortium, if you share their goals.

Considering the examples available in the literature, it is clear that in the United-Kingdom or in the Netherlands, and even to some extent in the United States, efforts have been made to engage in collaborative approaches (through JISC, SURF and the Open Learning Initiative), with an orientation towards open tools (e.g., Memorandum of Understanding signed between JISC and APEREO).

During an informal discussion after a keynote at the EDUCAUSE 2015 conference, a speaker involved in a learning Analytics project in some American university gave the following three advises about what he considers the keys of success: (1) have a high level of information about your data and a strong monitoring of the data flow, (2) form the right project team by mixing actors, technicians and experts, and (3) often try to compare “a priori” hypotheses with reality.

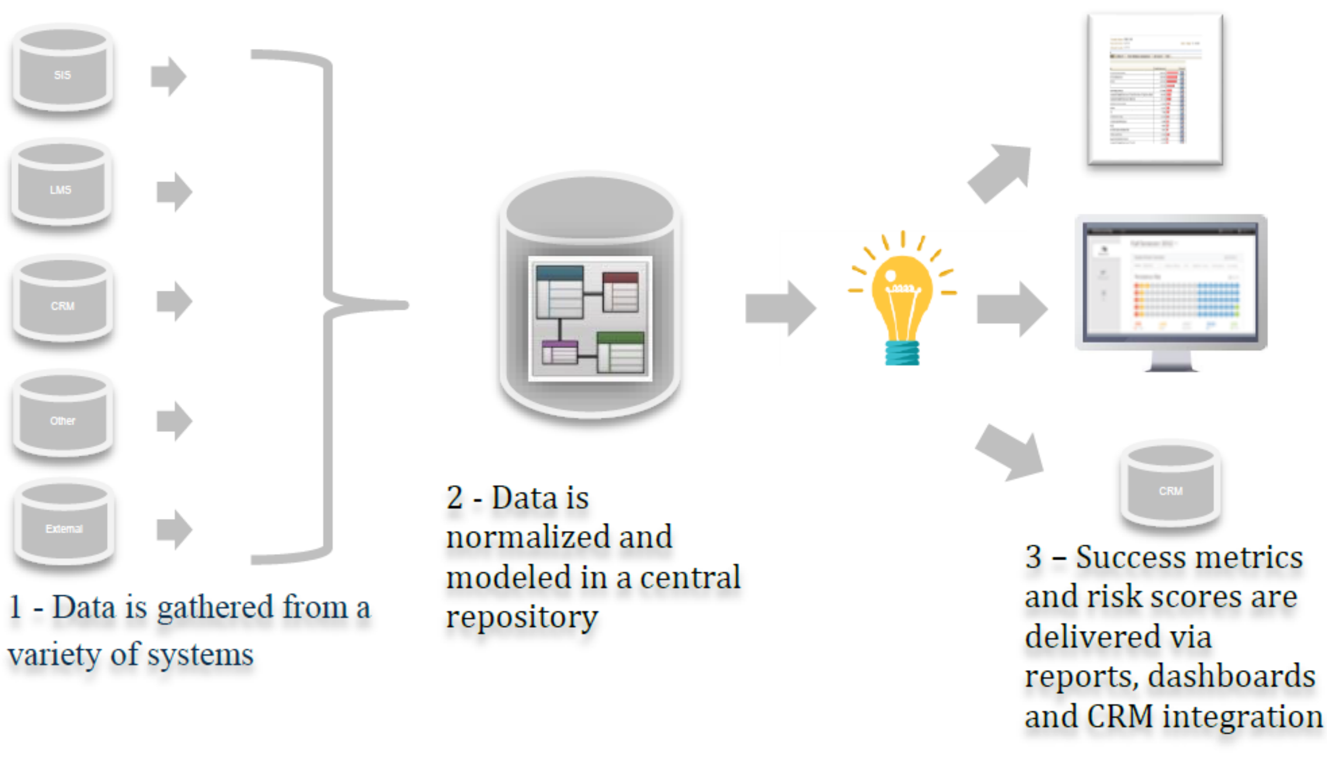


Figure 8 – Data flow

2. Panorama of research activities in the field of learning Analytics

A large part of research papers in learning Analytics deal with statistical analysis of learning (and teaching) data, which fits well the definitions provided in the introduction. Recent examples of such work include Molenaar and Chiu [2], who analyse learning sequences of different social groups, Snow et al. [3], who analyse differences of flexibility of writing style between learners considered as being good writers and other learners, Ferguson and Clow [4], who analyse the level of engagement in MOOCs, Vozniuk et al. [5], who analyse the use of peer evaluation tools, etc.

However, a large number of papers deal with a variety of other very different topics. For instance, Liu et al. [6] focus on modelling learners' misconceptions. Misconceptions reflect a thoughtful yet incorrect concept or strategy of the learners, and are considered by some researchers as one of the major causes of student failure [7]. As a first step towards that goal, Liu at al. limited themselves to the modelling of misconceptions in multiplications and additions of fractions, and extended the concept of knowledge component modelling to incorporate them.

Another example is the work of Adjei et al. [8], who focus on detecting courses' prerequisites. Being able to automate the retrieval of such information would have a great value, as it is often very difficult and time costly to retrieve it manually. Although the authors only have proposed a first simple method and the results may not be considered convincing yet, the idea is very interesting and promising.

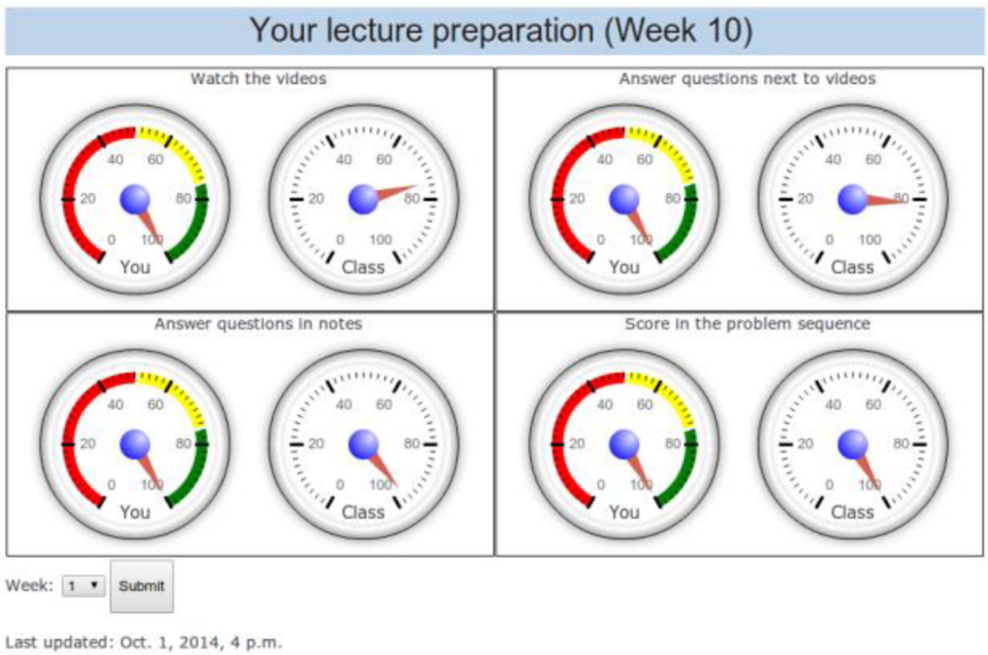


Figure 9 – Example of a dashboard, from (Khan and Pardo, 2016)

Another part of the research in that field deals with the design of dashboards. Dashboards are most often targeted to teachers and learners in higher education [9]. One recent example of a self-monitoring dashboard was proposed by Ruiz et al. [10], which allows students to track their emotions during courses. Another recent example is the work of Khan and Pardo [11], where a self-monitoring dashboard is proposed to higher education students with an emphasis on engagement. The dashboard uses data about video watching, answers to questions next to the videos and in subsequent assessments and displays engagement indicators as shown in Figure 9.

A last very different example is that of Ben David et al. [12], who aim at enhancing student performance and engagement by automatically sequencing questions. The proposed algorithm exploits knowledge about skill acquisition over time, which is incorporated into a Bayesian Knowledge Tracing model.

All of these examples were published in the proceedings of learning Analytics and Knowledge (LAK), the main conference on the topic. Clearly, research published in these proceedings goes far beyond the analysis of learning data to better understand the learning process. In order to have a clearer idea of the boundaries of what learning Analytics is in practice, we decided to browse the articles of that conference.

2.1. Classification of LAK papers and current trends

This section presents the result of a systematic classification and quantification of papers published in the proceedings of the successive editions of the conference learning Analytics and Knowledge (LAK). This conference is the reference of the research field, and we think it is reasonable to assume that it is representative enough of the different types of work being carried out. To date, six editions have taken place and a total of more than 300 papers have been published in the corresponding proceedings.

Our aim is to detect the different dimensions according to which the domain can be classified, and to analyse the evolution of their respective importance year after year. We chose to focus on the last three editions of the conference (LAK 2014, 2015 and 2016) because the acceptance rate just had dramatically dropped at that time[32], supposedly making the quality and focus of the published work higher than in the previous edition.

Some other relatively similar synthesis work has been carried out. For example, Dawson et al. [13] showed a similar evaluation of existing work. The authors identified all the articles of the conference LAK 2011 to 2013, and some additional items of special editions of magazines. However, their analysis focused mainly on the social network of the authors of these articles in relation to the classical categories in the field of information systems. Our work is different in that it focuses on identifying emerging key themes, their distribution and evolution across multiple dimensions on a different time

2.1.1. Subtopics

Learning Analytics covers a large number of possible subtopics. After reviewing the articles of the proceedings of the successive editions of the LAK series of conferences, we identified four main categories:

- Recommendations and notifications;
- Educational data mining: pattern mining, predictive modelling, clustering, etc.;
- Visualization tools such as dashboards;
- Statistical analysis.

32 The acceptance rate went from 39 to 28%.

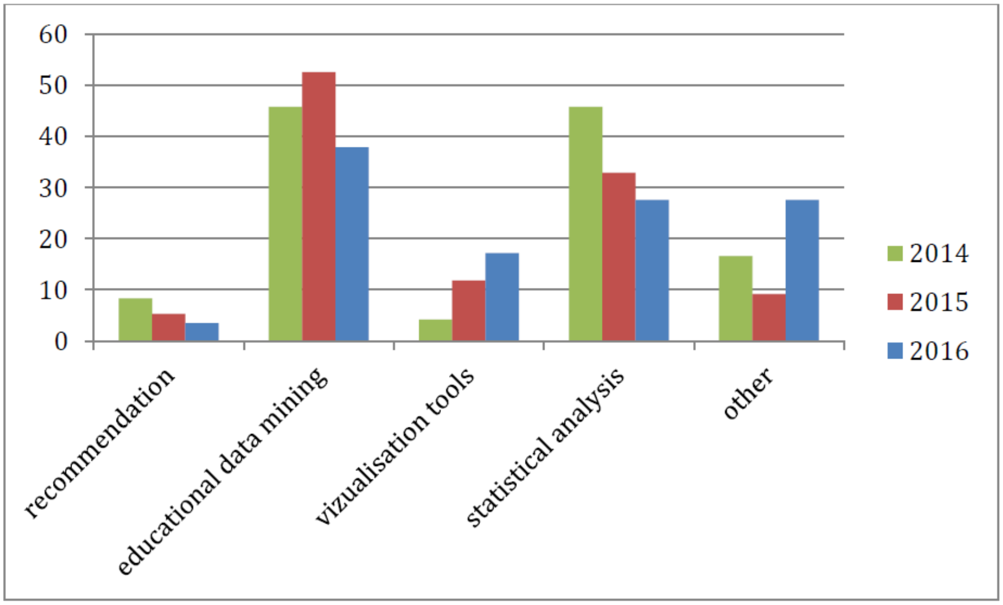


Figure 10 – Evolution of subtopics in proportion of the classified papers

Figure 10 shows the evolution of these subtopics in proportion to the classified papers. Perhaps the most obvious observation that can be made is the importance of educational data mining in 2014 and 2015. Interestingly, the number of papers focusing on that subtopic dropped in 2016. At the same time, visualisation tools emerged at such a point that they now represent the most frequent subtopic.

Moreover, although statistical analysis is supposed to be closest to the definition of learning Analytics provided by the very conference organizers, it slowly tends to take less and less importance. Another surprising phenomenon is the low importance of recommendations and notifications, and the fact that this importance even decreases over the years.

Last, less than 28% of the papers focus on various other subtopics, such as analysing the current state of the research on interactive surfaces [14], some methodology for finding causal relationships [15], advices on reducing the selection bias in real-world educational interventions [16], etc.

2.1.2. Targets

Research in learning Analytics can be intended for at least two different targets: learners and teachers. Note that only two of the four aforementioned subtopics allow such targets: (1) Recommendations and notification and (2) Visualisation tools. Educational data mining and statistical analysis are indeed independent of the possible targets, as the corresponding findings can be exploited to any of them.

Figure 11 shows the evolution of the proportion of work focusing on each of these targets. As can be seen, in 2014, teachers and learners were equally represented. This changed in 2015 and 2016, as much less of the papers focused on the teachers. Notice that in general, other targets are possible and were never yet studied yet. These include institutions and researchers.

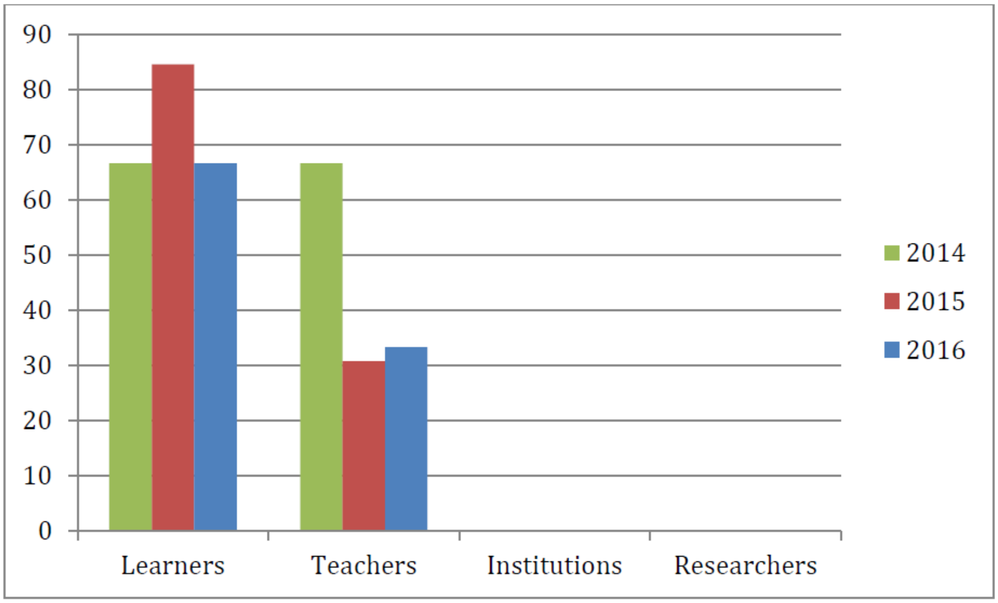


Figure 11 - Evolution of targets in proportion of the classified papers that focus either on recommendations and notifications or on visualisation tools.

2.1.3. Levels

The level of education is another important dimension of learning Analytics. It may correspond to primary education, secondary education, higher education or open education. Figure 12 shows the evolution of levels in proportion of the classified papers. The first observation that can be made is that very little work specifically studies the learning in primary and secondary education (less than 19%), whereas much more work focuses specifically on higher education and open education.

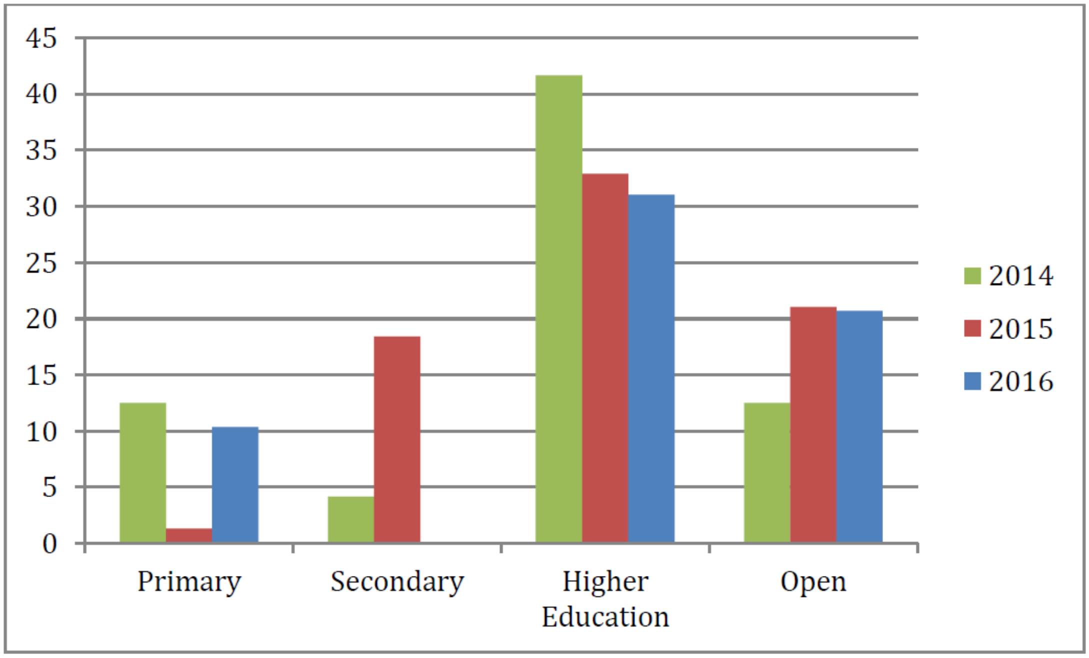


Figure 12 – Evolution of levels in proportion of the classified papers

2.2. Discussion

This classification puts forward several underexplored and unexplored areas of research that could be explored by researchers in the future.

We think that the most critical is that of recommendations and notifications. Much of the current work focuses on analysing available data and building predictive models out of it. However, learning Analytics offers a unique opportunity to let the system automatically act on the educational process.

One noteworthy project in that direction is the PERICLES project which has just ended. The aim of this project was to design a tool that provides recommendations of complementary educational resources or pedagogical activities to the learner, and which were adapted to his history, level and objectives. The resulting recommender system combined usage data from two separate platforms. The first was the personal learning environment of a university, and the second the portal of a provider of open educational resources. The objective was to propose the open resources to the students directly in their private environment.

Perhaps one explanation for that lack of research in that subtopic is that recommendations are particularly difficult in the context of education, as it is much more dangerous to provide bad recommendation to the learners as to the users of, for instance, a music recommendation system. In other words, in the context of education, recommendations can be harmful. We think one way to overcome this difficulty is to shift from the usual paradigm of optimizing recommendations towards accuracy metrics to optimizing towards other criteria that are more suited for the learners, such as engagement, scaffolding, etc.

The second underexplored area we would like to put forward is the research in primary and secondary education. One possible explanation for this is that it is often easier to get data and make experiments in both other contexts (higher education and open education). However, one could argue that the most important steps in education are the first ones: once a learner has dropped out at an early stage such as during primary education, it is much more difficult for him to ever go back on track.

Last, we would like to mention the unexplored possibilities of targeting recommendations, notifications, and visualisation tools to other entities than the learners and the teachers. One could for instance imagine a recommendation system designed for institutions that would help them make decisions. Another possibility is the design of dashboards for the researchers of the community

3. Conclusion

Digital tools have become almost ubiquitous, and have induced new means of understanding and enhancing the learning process. Students' learning can now be optimized and a large number of projects from the emerging field of learning Analytics have illustrated how such tools can be used to actually be beneficial to students, teachers and even institutions. As shown in the previous section, learning Analytics is now an active and pluri-disciplinary research field, bringing promising answers to many challenges that Higher Education institutions are facing.

A learning Analytics project within an institution is not only a research/development activity or a technological issue, but also an organisational issue. Furthermore, the collection of student data induces a number of ethical and privacy issues and challenges that have to be carefully considered and can be addressed in a variety of ways, depending upon the institutional and cultural context. The literature reports that the involvement of students and staff in the approval and adoption of open and transparent "codes of practice" and policies for the use of learning Analytics is positive.

The economic issues related to the internationalisation of education, the development of open education and the increasing demand from the society for lifelong learning are also emergent aspects. The digital transition of higher education is on the road, learning Analytics is a way to support the emergence of an aware, inclusive, personalized, online educational system.

4. References

[1] P. Long and G. Siemens, “1st International Conference on Learning Analytics and Knowledge,” 2011. [Online]. Available: https://tekri.athabascau.ca/Analytics/.

[2] I. Molenaar and M. M. Chiu, “Effects of sequences of socially regulated learning on group performance,” in Proceedings of the Fifth International Conference on Learning Analytics And Knowledge, 2015.

[3] E. L. Snow, L. K. Allen, M. E. Jacovina, C. A. Perret and D. S. McNamara, “You've got style: detecting writing flexibility across time,” in Proceedings of the Fifth International Conference on Learning Analytics And Knowledge, 2015.

[4] R. Ferguson and D. Clow, “Examining engagement: analysing learner subpopulations in massive open online courses (MOOCs),” in Proceedings of the Fifth International Conference on Learning Analytics And Knowledge, 2015.

[5] A. Vozniuk, A. Holzer and D. Gillet, “Peer assessment based on ratings in a social media course,” in Proceedings of the Fourth International Conference on Learning Analytics And Knowledge, 2014.

[6] R. Liu, R. Patel and K. R. Koedinger, “Modeling common misconceptions in learning process data,” in Proceedings of the Sixth International Conference on Learning Analytics And Knowledge., 2016.

[7] D. A. Muller, “Designing effective multimedia for physics education,” University of Sydney, Sydney, 2008.

[8] S. A. Adjei, A. F. Botelho and N. T. Heffernan, “Predicting student performance on post-requisite skills using prerequisite skill data: an alternative method for refining prerequisite skill structures,” in Proceedings of the Sixth International Conference on Learning Analytics And Knowledge, 2016.

[9] B. A. Schwendimann, M. J. Rodríguez-Triana, A. Vozniuk, L. P. Prieto, M. S. Boroujeni, A. Holzer, D. Gillet and P. Dillenbourg, “Understanding learning at a glance: An overview of learning dashboard studies,” in Proceedings of the Sixth International Conference on Learning Analytics And Knowledge, 2016.

[10] S. Ruiz, S. Charleer, M. Urretavizcaya, J. Klerkx, I. Fernández-Castro and E. Duval, “Supporting learning by considering emotions: tracking and visualization a case study,” in Learning Analytics And Knowledge, 2016.

[11] I. Khan and A. Pardo, “Data2U: scalable real time student feedback in active learning environments,” in Proceedings of the Sixth International Conference on Learning Analytics And Knowledge, 2016.

[12] Y. Ben David, A. Segal and Y. K. Gal, “Sequencing educational content in classrooms using Bayesian knowledge tracing,” in Proceedings of the Sixth International Conference on Learning Analytics And Knowledge, 2016.

[13] S. Dawson, D. Gašević, G. Siemens and S. Joksimovic, “Current state and future trends: A citation network analysis of the learning Analytics field,” in Proceedings of the Fourth International Conference on Learning Analytics And Knowledge, 2014.

[14] R. Martinez-Maldonado, B. Schneider, S. Charleer, S. Buckingham Shum, J. Klerkx and E. Duval, “Interactive surfaces and learning Analytics: data, orchestration aspects, pedagogical uses and challenges,” in Learning Analytics And Knowledge, 2016.

[15] [1 K. R. Koedinger, E. A. McLaughlin, Z. Jia, Julianna and N. L. Bier, “Is the doer effect a causal relationship?: how can we tell and why it's important,” in Learning Analytics And Knowledge, 2016.

[[16] C. Brooks, O. Chavez, J. Tritz and S. Teasley, “Reducing selection bias in quasi-experimental educational studies,” in Learning Analytics And Knowledge, 2015.

ICDE - International Council For Open And Distance Education
Lilleakerveien 23, Oslo 0283, Norway

www.icde.org

icde@icde.org