

LACE – CREATING A COMMUNITY ON LEARNER ANALYTICS AND EDUCATIONAL DATA MINING – STRUCTURING THE DISCOURSE

Tore Hoel, Oslo and Akershus University College of Applied Sciences, Norway, Adam R. Cooper, CETIS, UK, Mehrnoosh Vahdat, ITS Infinity, Italy, Fabrizio Cardinali, ITS Infinity, Italy

Abstract

Discussions on MOOCs, on digital textbooks and new authoring and publishing models, on privacy issues related to digital practices, etc. have led to increased interest in Learning Analytics and Educational Data Mining. The European 7th Framework Coordination and Support action project, LACE (Learner Analytics Community Exchange) is devoted to structure this discussion and build communities that explore the potentials and pitfalls of the unprecedented availability of data related to education and learning. This paper gives a basic introduction to Learning Analytics and Educational Data Mining, presenting a generic framework and outlining how this will be used to support community building and discourse on different aspects of these themes within schools, higher education and the workplace. Learning Analytics and Educational Data Mining both only recently brought to the top of the educational technology trend lists, this paper argues that a discourse and co-design approach would be the best to follow for the broad range of identified stakeholders.

Introduction

When translating learning into numbers all kinds of discourse emerge, especially when the promise is that learners, teachers, local authorities, companies and others soon will have tools that make the churning and interpretation of these numbers available for all with a device. Enthusiasts and sceptics form positions, and the battle to come is less exciting for the many stakeholders that are oriented towards step-by-step improvements of learning based on sound evidence, within ethical boundaries. Moving the field of Learning Analytics (LA) and Educational Data Mining (EDM) beyond technology hype, prejudiced resistance and pedagogical straightjacket choice, there is a need for a sound framework identifying the critical dimensions of LA and EDM, so that focussed discourse within the particular stakeholder communities can identify the the most promising designs for improved learning and developing education and training systems.

In January 2014, LACE, a 30 months Support and Coordination action project under the the 7th European Framework Program kicked off with the aims to promote knowledge creation and exchange on LA and EDM; increase the evidence base; contribute to the definition of future directions; and build consensus on interoperability and data sharing. The project acronym expands to Learner Analytics Community Exchange, and LACE is eager to build on existing community activities, e.g., research being done in SOLAR (Society for Learning Analytics Research), and new EU specific targeted research projects; and to make sure that practices in Schools, Higher Education and Workplace learning are part of the conversation. These sectors are the focus areas of the project, in addition to Interoperability and Data Sharing.

The LACE project involves nine core partners¹ covering all educational sectors, with well established practice in research and European projects within technology enhanced learning. From the very beginning the activities of the project is directed towards involving the broader communities in creating an hub of evidence on LA/EDM supporting the discourse. (An Evidence Hub like the one created for Open Educational Resources <http://ci.olnet.org> is one of the deliverables of LACE).

¹ Open Universiteit Nederland, NL, Cetis, the Centre for Educational Technology and Interoperability Standards at the University of Bolton, UK, Institute for Educational Technology at the Open University, UK, Infinity Technology Solutions, IT, Skolverket, the Swedish National Agency for Education, SE, Kennisnet, NL, Høgskolen i Oslo og Akershus, NO, ATiT, Audiovisual Technologies, Informatics and Telecommunications, BE, EDEN, the European Distance Education Network, HU

The backdrop for LACE

It is probably no co-incidence that learning analytics and educational data mining, as for data analytics generally, are the subject of interest and excitement at a time of increasing global competition and large scale economic difficulties. The promise of using data to be more efficient and effective is obvious to many people, given the increased access to data from various sources. Data-driven education may also provide a sound basis for doing things differently, re-inventing parts of our approach to education and training rather than incrementally improving what we do now.

The Society for Learning Analytics Research (SoLAR) has defined Learning Analytics 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' (1). A key element of analytics is that they are not primarily concerned with reporting activity or with theoretical insights. Instead, they focus on the provision of 'actionable intelligence' that can provoke or encourage practical action. In an educational setting, analytics allow learners and educators 'to increase the degree to which our choices are based on evidence rather than myth, prejudice or anecdote' (1,2).

The terms "learning analytics" and "educational data mining" have emerged over the last few years, largely representing the identity of two research communities (3). The LACE project is hesitant to present a too narrow definition of the concepts not excluding any topic of interest to this emerging community. From a LACE perspective learning analytics will be defined by what people do; and it will be emergent and changing. However, it makes good sense to look upon LA/EDM as multi-layered field, presenting a micro, meso and macro perspective on the entities to be described. Finer-grained process data that are now increasingly available for analysis inform higher level policies, as described in Figure 1. This model from the UNESCO Policy Brief on Learning Analytics sees a convergence of the three levels with breadth and depth from macro and meso levels adding power to micro analysis (4).

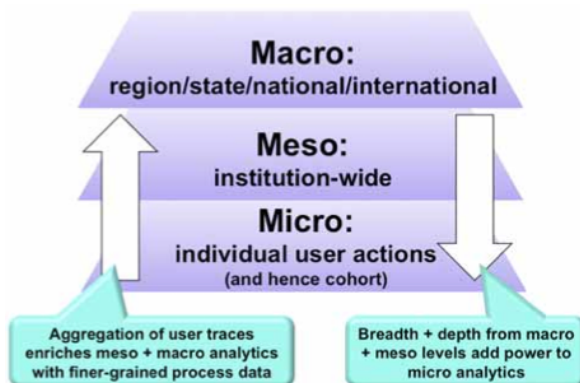


Figure 1 Levels of Learning Analytics (Source: UNESCO Policy Brief, November 2012)

The promises of Learning Analytics and Educational Data Mining may be conceived differently in the three sectors the LACE project will work with.

School education: Education ministries and their agencies are urged to find more effective and more economical systems for schools. Their challenge is made even harder because the school system must anticipate national needs for well-educated and skilled citizens, many of whom will have jobs that we cannot yet imagine. LA and EDM offer a valuable source of policy evidence but may also help to teach and assess 21st Century skills (5).

Higher education: European universities are being forced into new ways of operating and new business models by a combination of the economic situation, government policies and globalisation. Although the picture varies across Europe, university management and staff are struggling to act in the face of new regulatory approaches, competition from the private sector, a changing relationship with the students influenced by fees and diminished employment opportunities, reduced budgets, a feeling of threat from the latest wave of e-learning in the form of the MOOC, and increased perception that university education is failing to deliver the kind of graduates necessary for long term viability of European economies. It is imperative that universities innovate in what they

offer and how they deliver it. LA and EDM offers both new opportunities and a means to shorten the time to understand the effect of change.

Workplace learning: Europe's economic strength is underpinned by manufacturing, agriculture and financial services and it is manufacturing that is where the LACE project focuses attention, in particular on smart manufacturing, which many analysts see as being the means for Europe to reclaim some of the global manufacturing productivity that has recently moved out of the region. To get manufacturing smarter, new educational practices and solutions need to be trialled, shared and adopted at a much faster pace than in the past and driven by data derived from the workplace and from the learning activities. Analytics are far from new to industry; therefore, it will be interesting to explore how LA and EDM would fit with already existing work patterns.

In the following this paper will introduce the generic framework for Learning Analytics developed by Greller and Drachsler (2012), which will serve as a tool for structuring the community discourse supported by LACE. With this tool in mind, the paper will lay out and explore some of the cross-points to be addressed by the project where thematic groupings (e.g., content & resources, support for learning, assessment & performance) are matched against educational sector levels.

A Generic LA Framework

Learning Analytics is about doing much more than before because data and tools are available to us. However, Learning Analytics also raises many non-technical issues about data ownership and openness, ethical use and dangers of abuse, – issues Drachsler and Greller describes as 'softer' issues and problem areas, which also need to be taken into consideration designing a generic LA framework (6). These authors used a general morphological analysis approach (cf. 7) to deduce the critical dimensions involved in a comprehensive discourse on learner analytics.

«With the framework, we take the presumption that responsible designers of analytic processes will not only implement what is technically possible to do and legally allowed (or at least not prohibited), but to consider holistically the outcomes for stakeholders and, even more importantly, the consequences for the data subjects, i.e., the people supplying the data (cf. the section on stakeholders below).» (8)

The framework model (Figure 2) considers six critical dimensions. Each of the dimensions can be subdivided into several instantiations falling into that dimension. For example, the generic "stakeholder" dimension can have instantiations (values) like "learners" and "teachers." The list of instantiations in the diagram is not exhaustive and can be extended on a case-by-case basis. To stay with the above example, commercial service providers and even automated agents could also function as stakeholders in a LA process. It is useful to note that through connecting various (and also multiple) different instantiations of each dimension, concrete use cases can be constructed. The dimensions are "critical" in the sense that each of the six fields of attention is required to have at least one instantiation present in a fully formulated LA design, even if some dimensions are vaguer than others in this respect.

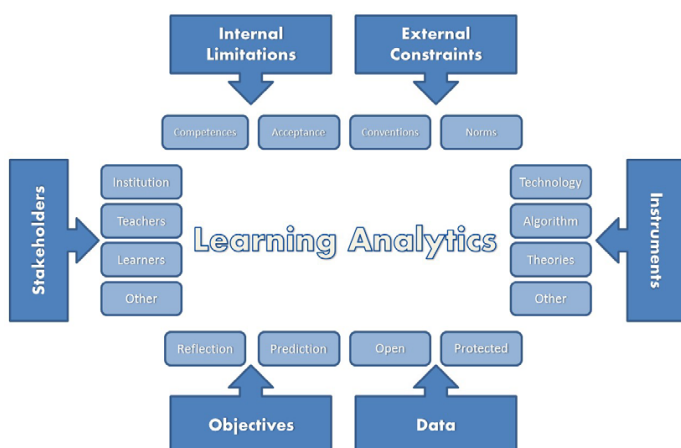


Figure 2. Critical dimensions of learning analytics

In the following the six dimensions are discussed more in detail, keeping in mind the tasks ahead of the newly started LACE project.

- **Stakeholders:** the contributors and beneficiaries of learning analytics. Institutions, teachers, and learners – they could all be suppliers of data and users of data. The stakeholder scenarios we will see in the future are complex and not well researched; many stakeholders have still to realise that they are concerned.

- **Objectives:** set goals that one wants to achieve. The promise of Learning Analytics is to unveil and contextualise so far hidden information out of the educational data and prepare it for the different stakeholders. Greller and Drachsler distinguished two fundamentally different objectives: **prediction and reflection (8)**. UNESCO (2012) listed five examples of LA: LMS/VLE Dashboards, Predictive Analysis, Adaptive Learning Analytics, Social Network Analysis, and Discourse Analytics. These categories of LA find different positions on the prediction - reflection continuum.

- **Data:** the educational datasets and their environment in which they occur and are shared. Data is key for learner analytics in a narrow sense; however, the terms for giving access to and for sharing the data are key to reap the benefits in a broader sense. **Linking datasets** are hugely interesting in terms of technical results, but may be counterproductive if not accepted by the learners. The soft issues of access to data is therefore a crucial aspect of this dimension, e.g., open vs. protected datasets, anonymisation, trust systems, etc.

- **Method:** technologies, algorithms, and theories that carry the analysis. How raw data 'translates' into information depends upon the instruments we use, and their theoretical underpinnings. This field is dynamic and emergent, and it is a huge challenge **to open the 'black box'** and make sure that the users understand the **limitations of their interpretations**.

- **Constraints:** restrictions or potential limitations for anticipated benefits. Even if constraints are an integral part of all dimensions of the framework, restrictions and potential limitations are singled out for dedicated scrutiny. Especially, the ethical and legal aspects of data access and sharing need to be built into new LA/EDMI designs.

- **Competences:** user requirements to exploit the benefits. Tools do deliver impressive visualisations of complex datasets; however, without skills and competencies to interpret the results learner analytics may do more harm than good. Basic numeric and other literacies, as well as ethical understanding are not enough to realise the benefits that learning analytics has to offer (8). What competences is required for optimal exploitation of learning analytics data remains to be defined and could only be the achievement of community discourse the LACE project wants to contribute to.

Learner Analytics Community Exchange

The LACE project is designed to capture the key questions of LA and EDM as they arise, acknowledging that this is a young field of research. Therefore, the project has designed a loose framework for facilitating the conversation leading up to actionable insights into the key challenges. At the intersection between Thematic Group and Sector we can describe an idealised interaction: policy and research shape new practice, practice should inform research, and policy should be informed by research and practice. LACE aims to promote useful interactions at these points but also to make conceptual connections and open channels of communication that will increase knowledge cross-over.

The thematic groups are defined as follows:

Content & Resources will include the production, use and management of resources with educational value. These are likely to be mostly digital resources but need not be so. The opportunities analytics opens up for enhancing the quality of intentionally educational resources, for aiding resource discovery and for the creation of "smart content" all come under this thematic group.

Support for Learning will include analysis to inform or guide both teachers and learners. Support for reflective practice, for course design, for early warning of difficulty, for identification of need, etc are all part of this thematic group.

Assessment & Performance will be concerned with the application or design of objective measures of what a person can do, knows etc. This thematic group includes the use of analytics to enable new forms of assessment, to make assessment in authentic contexts a viable proposition, to increase the status of non-traditional assessment instruments, and to generally enrich the range of assessment and feedback methods.

The intersection between thematic group and sector defines **cross-points that**, in theory, should burst with heated discourse once prodded with the right community exchange instruments of the project. These are project inventories found in most projects, e.g., a community website, social media practice, reports and briefs, etc. In addition, the project will populate the LACE Evidence Hub, which will give a well structured overview of the key challenges, issues, potential solutions, research claims, and the evidence, content resources, and organisational and personal resources that could back up the carefully drafted statements in knowledge system. Drawing heavily on the **Evidence Hub and** ideas gathered during the project a Policy Delphi process will be conducted to expose differences of perception and vision from a group of researcher and practitioner experts selected from the project's liaison organisations and from people who have participated in LACE activities. This will consider views on what is desirable, what is feasible and the obstacles to making what is desirable happen.

It is too early in the project to have a firm opinion of what the cross-points will unveil. However, the frameworks presented give grid for making projections about future discourse. We conclude this paper with some projections related to the 2014 theme of the EDEN conference, e-learning at work and the workplace.

Further discourse on LA & EDM and Workplace learning

Among the 18 themes for the 2014 conference at least 7 directly address LA and EDM. These seven themes can be clustered in three groups, – one related to competence management and assessment; another to scaling up work based learning by ICT; and the third addressing institutional quality management. We will shortly touch upon the LA and EDM issues related to each of these clusters.

Competence management and assessment: Technologies for formalising and exchanging information of competency frameworks (cf. InLOC – Integrating Learning Outcomes and Competences²) allow for systems that help individuals to position their own competences in relation to a norm defined by a company, an industry group, a professional quality body, etc. Gap analysis could give both the employer and the employee means to define actions related to employability, task management, workplace learning and training etc. The locus of intervention could be content and resources (e.g., learning outcomes linked to educational resources); support for learning (e.g., personal learning profile vs. idealised task competence profiles); or assessment and performance (e.g., competency profiles for teams to help selection). When designing systems in this domain it would be advisable to walk through the LA framework presented above in order to identify potential stumbling blocks. For instance, access to data could be dramatically changed if the system is designed from learning vs. a control perspective.

When designing LA systems for the workplace of today one should acknowledge that discipline knowledge is only one critical yardstick in competence management. Learning dispositions and the transferable competencies associated with skilful learning in diverse contexts are increasingly important, and a field that the learning analytics field has yet to engage with (9).

Scaling up work based learning: Traditional approaches to manufacturing and to workforce training are no longer adequate if Europe is to compete globally. The digital marketplaces are developing rapidly, with crowd sourcing and crowd funding schemes, and Produce on Demand and Make to Individual models, – spurred by the introduction of Internet of things, 3D printing and other technological innovations. Talent driven innovation is one of the main drivers of global manufacturing competitiveness, which leaves up-skilling of the workforce imperative. Still missing are tools to support the training process and the assessment of its efficiency. A 2005 survey of a variety of corporate settings confirmed that many companies already were using technology-aided learning methods, and had a strategic plan for e-learning. However, only a small part specified that this plan was efficient and that they could measure the return of investment in training (10).

Learning analytics is an emerging method of learning assessment and appraisal in the workplace. By analytics of learning behaviour of employees, it is possible to understand who needs more training, and in what process. It

² InLOC is a European standardisation project – see blog archive of project lead Simon Grant <http://blogs.cetis.ac.uk/asimong/category/inloc/>

also allows seeing how training can be the most efficient and cost effective. LACE project will the coming two years explore the evidences of Learning Analytics in the workplace, particularly in smart manufacturing.

Quality management: The EDEN conference themes focussing on student guidance services, assessment and evaluation, retention techniques, and performance support all address the main areas of LA research, where we start seeing some results (2). In these areas the main challenge is to make the already existing data available for analysis. Universities and other educational institutions have a long tradition of analysing and reporting data, before anyone put the EDM label on the activity. Even if the objectives are glorious (e.g., retention and support), the data remain unanalysed even if new instruments are brought to the market. Only an approach of discourse and co-design may engage the necessary stakeholder so that the full potential of LA and EDM in quality management can be exploited.

Conclusions

The New Media Consortium Horizon Report 2014 for Higher Education (11) puts up Rise of Data-Driven Learning and Assessment as a mid-range trend driving changes in higher education within 3 – 5 years. The report points out “while interest is considerable, higher education in general has yet to fully embrace these sorts of processes. Privacy and ethics are just beginning to be addressed, but the potential of using data to improve services, student retention, and student success is clearly evident”. The same can be said for schools and workplace learning. Potential is a promise, not real before delivered on. For LA and EDM to deliver, there is no quick fix in a new technology, a LA dashboard or ground breaking algorithm. LA and EDM must be approached from a holistic and systemic perspective, balancing the critical dimensions identified in the LA framework presented in this paper. The LACE project is founded on the assumption that usage context is key to the success of these new technologies. Therefore, the communities need to be engaged in exploring the different contexts. Learning Analytics may well have only one year or less time-to-adoption (11), however, the short history of technology-enhanced learning has shown that availability of a technology is just one factor leading up to successful adoption for most stakeholders.

References

1. Cooper A.R. (2012). What is Analytics? Definition and Essential Characteristics. CETIS Analytics Series, Vol.1, No.: 5.
2. Ferguson R. (2013). Learning Analytics for Open and Distance Education. CEMCA EdTech Notes
3. Siemens, G. & Baker, R.S.J. (2012). Learning Analytics and Educational Data Mining : Towards Communication and Collaboration. In LAK12.
4. Shum, S.B. (2012). Learning analytics. UNESCO Institute for Information Technologies in Education. Policy Brief. Available at <http://iite.unesco.org/pics/publications/en/files/3214711.pdf>, Accessed 2014-02-05
5. Voogt, J. et al. (2010). 21st Century Skills Discussion Paper, University of Twente. Available at: [http://www.internationalsymposiumoneducationalreform.com/storage/21st Century Skills.pdf](http://www.internationalsymposiumoneducationalreform.com/storage/21st%20Century%20Skills.pdf).
6. Drachsler H., Greller W. (2012) The pulse of learning analytics understandings and expectations from the stakeholders. LAK 2013, ACM, <http://dl.acm.org/citation.cfm?id=2330634>
7. Ritchey T. (2013) General Morphological Analysis. A general method for non-quantified modelling. Downloaded from the Swedish Morphological Society (www.swemorph.com)
8. Greller W, Drachsler H. (2012). Translating learning into numbers: A generic framework for learning analytics. International Journal of Educational Technology & Society. 15(3):42–57. <http://lnx-hrl-075v.web.pwo.ou.nl/handle/1820/4506>
9. Buckingham Shum, S., Deakin Crick, R. (2012). Learning dispositions and transferable competencies: pedagogy, modelling and learning analytics. In: 2nd International Conference on Learning Analytics & Knowledge, 29 Apr - 02 May 2012, Vancouver, British Columbia, Canada.
10. Bonk, C., Kim, K.J. & Zeng, T. (2005). Future Directions of Blended Learning in Higher Education and Workplace Learning Settings. In P. Kommers & G. Richards (Eds.), Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2005, pp. 3644-3649. Chesapeake, VA: AACE.
11. Johnson, L., Adams Becker, S., Estrada, V., Freeman, A. (2014). NMC Horizon Report: 2014 Higher Education Edition. Austin, Texas: The New Media Consortium.