Data Sharing for Learning Analytics – designing conceptual artefacts and processes to foster interoperability

Tore HOEL[*] & Weiqin CHEN
Oslo and Akershus University College of Applied Sciences, Norway
*tore.hoel@hioa.no

Abstract: Learning Analytics is based on data from the digital traces left by learning activities. In the controlled environment of a research lab combining data from different sources does not pose many problems. However, when scaling up learning analytics for general use in schools and universities data sharing and interoperability become major challenges. These issues are now being addressed in standardisation settings, both internationally and nationally. A case study of a Norwegian standards project shows that there are considerable conceptual issues emerging when stakeholders representing different interests start working towards consensus on these issues. Based on the case study this paper contributes with a number of conceptual constructs and a process that will make it easier to reach consensus about different aspects related to access to and exchange of data from different sources relevant for analysis of learning and the contexts in which learning occurs.

Keywords: Data Sharing, Interoperability, Learning Analytics, Educational Data Mining, Standardisation

1. Introduction

A common definition of Learning Analytics (LA) is given by the Society for Learning Analytics Research (SOLAR): "Learning analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs" (LAK11, 2011). The definition presupposes that we have a clear idea of what kinds of data are needed to optimise learning and its environments. When moving out of the research lab into the complex field of real life actors representing the vastly diverse interests we find in education we see that it is not only the access to data that represents a challenge; we also lack good concepts to describe the data we want to collect. According to the LA definition it is our understanding of what contributes to learning and optimal learning contexts that should frame our search for data. In the real world we have to do with what we have got, implying that we often start with the data coming out of our learning management systems and our learning assessment systems (Macfadyen & Dawson, 2010; Rienties, Toetenel, & Bryan, 2015; Kitto, Cross, Waters, & Lupton, 2015).

Data sharing can be defined as the release of data for use by others (Cooper & Hoel, 2015). In beginning of 2016, Standards Norway, the main standards organisation of Norway, gathered 'the others' around a table and started work on a technical report on "data sources and conditions for data sharing for learning analytics". The work is to be based on the interest of vendors, school authorities, universities, publishers, and others in the Norwegian market who want to advance the use of LA based on a richer set of data. A provisional scope has been agreed to clarify what are the most important data sources in the Norwegian market; what conditions regulate access and sharing; and to come up with ideas of methods for sharing that will give access to data across actor groups.

In the opening meeting of this project, in which both authors took part, it soon became clear that a data centric approach to data sharing was fraught with subtleties that soon could render the discussion impenetrable. As an example, the concept of a data source proved difficult to use. A data source does not tell much about what information is embedded in the data, which in turn is important to know in order to see if for example entailed personally identifiable information (PII) would make it difficult to share the data. In this meeting it was felt that there was a need for developing a new conceptual toolbox to make the exploration of different data sharing scenarios fruitful.

This paper will explore the discourse space one enters when addressing the data sharing needs for the LA community. Based on this case study we will develop a set of conceptual artefacts that can be used in further work in this particular group and hopefully beyond. The paper is organised as follows: After reviewing related work, we will present a small case study of the Norwegian standards project on data sharing. Based on the requirements identified in the case study, we will construct conceptual artefacts and a process that could be used in this context. The proposals will be discussed and the conclusions will present ideas how this work could be developed further.

2. Related Work

Interoperability and data sharing become issues first when we go beyond LA research and start to explore how LA will influence the agendas for schools, universities and national policy makers. LA being an emergent field of research, it is as expected that till now 'interoperability' and 'data sharing' in the context of LA have been rather absent in the research literature. The European Union LA support and coordination action, LACE, on the other hand had a work package on interoperability and data sharing. In a deliverable on 'Data Sharing Requirements and Roadmap' Cooper and Hoel (2015) reported they were struck by the extent to which the characterisation and ramifications of data sharing had not been worked on.

> The interpretation of data sharing is at present somewhat confused by common conceptions of ownership and related factors, under-developed thinking about the topic, and sometimes a failure to consider it. The increasing use of software hosted in "the cloud" - i.e. Software as a Service (SaaS) - has amplified this situation, but we have also inherited a confusion from the days when most software used in education was running on-site. This confusion relates to ownership and control, and the extent to which the educational establishment has absolute authority or acts as a custodian on behalf of the learner. While it is clear that there is some data which the educational establishment is required to keep, and some of which the learner has no right to change, there has generally been little attention to the details of ownership, control, and custodianship. (Cooper & Hoel, 2015, p. 11)

Within a research context there is a long tradition for data handling with ethical committees and a systematic approach for deposit, sharing, reuse, curation and preservation of data (van den Eynden & Bishop, 2014). Even if some of the processes and technologies are relevant for large scale LA delivery, going beyond a controlled research setting will bring into play a more complex set of actors and systems. Till now it is mainly technical factors that are driving the need for data sharing. "Increasing use of Cloud Computing models of service and IT provision, where expertise or technology is provided by a separate organisation to the education provider, has increased the extent to which data is not only distributed between different IT systems, but is also distributed among legal entities" (Cooper & Hoel, 2015, p. 9). However, as pointed out in the report from the LACE project, the situation for educational institutions is more characterised by Small Data than Big Data, – "you can easily fit your data in a spreadsheet on your laptop computer!" (Cooper & Hoel, 2015, p. 9). While a spreadsheet may be inadequate for captured activity data, it remains true that for most of the potential applications of learning analytics in education and training practice, the useful data will be of a scale well below that of Big Data.

For the practical work the Norwegian standards group sets out to do on data sharing and interoperability Cooper and Hoel (2015) give limited help. They observe that the variety of data that is relevant to learning analytics is indeed potentially very great; however, the LACE report limited the interest primarily to concern data about people and their activity in a learning-related situation. Other data sources, e.g., national and international classification schemas for subject matter of courses or learning resources lack many of the complications of person-related data, and they are undertaken as Open Data initiatives and other projects (Cooper & Hoel, 2015, p. 7). However, as the initial meeting in the standards group showed, these other data sources are often the point of interest, from where the stakeholders start to explore LA data sharing and interoperability. Therefore, it seems to be a gap – not addressed by the current LA research literature – how to bridge between person-related activity data and the other data sources well established in the education community.

One way of proceeding is to actually see what kinds of data are used for LA. In 2012, Chatti, Dyckhoff, Schroeder, and Thüs did a review of recent literature related to LA and related fields and

found that centralized web-based learning systems (e.g. Intelligent Tutoring Systems (ITS), and Learning Management Systems (LMS)) represent the most widely used data source for LA. They further found that most of the current LA applications were oriented toward intelligent tutors or researchers/system designers; the most commonly applied objectives were adaptation and monitoring/analysis; and the most frequently used LA techniques were classification and prediction (Chatti et al., 2012). This pattern, however, they thought would change,

"as the focus of LA will shift toward more open, networked, personalized and lifelong learning environments. LA further requires key stakeholders to address a number of challenges, including questions about handling increasing data volume, heterogeneity, fragmentation, system interoperability, integration, performance, scalability, extensibility, real-time operation, reliability, usability, finding meaningful indicators/metrics and appropriate information visualization, supporting mixed-method approaches (quantitative and qualitative), data privacy, stewardship, ethics, and integration of LA into everyday practice. These challenges will need to be addressed as the understanding of the technical and pedagogical issues surrounding LA evolves" (Chatti et al., 2012).

Not surprisingly, what the study of Chatti et al. (2012) shows is that educational practitioners start with the data they have. In universities one have LMS and some experimentation with ITS. This, however, does not give the full picture of what LA promises to deliver. Only for Social Learning Analytics, what Ferguson and Buckingham Shum (2012) propose as a subset of LA, five distinct approaches are identified: network analytics, discourse analytics, content analytics, dispositions analytics and context analytics. Each of these approaches has its own justification and typical set of data. When we are looking for data interoperability and the possibility to share and merge data sets in a future perspective, we need to look at the different types of LA approaches and see what data sources they build on.

2.1 Focus of this study

The review of related work has established a research gap related to the conceptualisation of data sharing for LA. The perspective in this paper is pragmatic, in the sense we want to facilitate the process of coming up with consensus of data sharing and interoperability for LA within a national context. We see that there is a need for a more concrete discussion of the aims for data sharing than outlined in the LACE report discussed above. Cooper and Hoel (2015) pointed towards "more useful analysis through combination of data from different sources", "sufficient scale of data to determine relevance and quality of ed[ucational] resources", "critical mass of data for learning science research, reproducibility and transparency in LA research", "cross-institutional strategy comparison", "research on the effect of education policy", "social learning informal settings", and "learner data as a teaching and learning resource" as aims for learning analytics data sharing (p. 8). None of these rationales would advance the discussion around the table in the standardisation group under study.

After a short case study into the dynamics of the group, the authors of this paper will design a first draft of a discourse toolbox, which will be tested in the coming meetings in the group. This research is positioned in the first Relevance Cycle of the three research cycles of Design Science (Hevner, 2004; 2007), addressing requirements and field testing. The purpose is to come up with candidate concepts that describe the problems and opportunities in the application domain from a people, organisational systems, and technical systems perspective.

3. Case study of the initial phase of a consensus process

The kickoff of the standard project on data sharing May 2016 was preceded with an invitation to think about the issues from a bird's eye view and to contribute use cases led by four simple questions: Who does you represent? What data sources do you use? What data sources would you like to get? Do you have comments on conditions for sharing? Both activities were carried out as a collaborative writing effort using Google Docs.

From a work group facilitator's point of view, the preparatory work was a disappointment. The high level reflection on data sharing for LA opened up a Pandora's box of everything related to data in education. One contributions argued why open data is important and therefore supported by

government policies. Another reported ongoing work to create a data architecture for higher education in Norway mapping the activities of all service providers in the sector. Just a quick look at the draft document would prove that this attempt to get a high level grasp of challenges would lead nowhere.

Four stakeholders contributed to the first round of use cases: a big vendor, a local school authority, a school agency, and a publisher. Even with the same set of questions and access to each others' contributions the answers vary a lot in scope. We have a list of services and data providers, with a number of concerns and issues related to technical architecture, ownership, sharing culture, what type of data are collected, etc.

The school authority gave an overview of their central databases on curricula and learning goals and official statistics from national authorities. They were also pointing out that they had access to student information data on users, LMS data and user-generated data from a number of applications for digital learning resources. What they wanted to get hold of were roster data, competency information related to learning activities, and local learning goals related to the curriculum.

The school agency was concerned with the availability of the vast datasets managed by The Norwegian Directorate for Education and Training and the Norwegian Centre for Research Data. The data are essential for assessing the quality of education in Norway; however, the datasets are not easily accessible, and provided examples of exchange with these data authorities prove that substantial negotiations are necessary to make these datasets an active open source for learning analytics.

The publisher presented their solution for ebooks, explaining that they stored detailed information on usage patterns of each learning resource (e.g., right and wrong answers to quizzes, did the user check for right answer, how does the learning resource relate to curriculum, etc.).

The kickoff meeting, gathering 22 experts representing all the relevant stakeholders for this work in Norway discussed rationale, scope, output of work and agreed on working procedures before embarking upon technical work. As expected, the initial discussion was dominated by framing activities, positioning the actors (Hoel & Pawlowski, 2012), making sense of the scope of work (Hoel & Mason, 2012), and exploring the stakeholder interests (Hoel & Hollins, 2011). The appointed technical lead tried to drive discussion towards getting a grasp of what data the stakeholders were interested in exchanging, but struggled to get beyond principled views of open access and sweeping reference to categories of data and the data sources mentioned in the shared document.

Much as the facilitator repeatedly referred to use cases as a useful instrument to map concrete and relevant stakeholder interests the discussion never came to the point where for example a publisher would declare: We have these data, which makes this application work today; however, if we get those data form that source we could make a much better learning resource. Participant observation of the discussion made it visible that the group lacked a common conceptual understanding of data sharing and LA. To discuss use cases did not make any sense for the group. Along the LA spectrum from person-related, activity-focussed analytics to more traditional academic analytics (Baepler & Murdoch, 2010) different concepts of data come into play. The group lacked the necessary conceptual common base to engage in a solution oriented discourse.

To create a common ground one needs to reach out – to find a new position. However, the discourse in this start-up meeting also highlighted another prerequisite for moving towards consensus: willingness to be explicit about one's own position and interests. Conceptual tools are not enough; one also need motivation.

Two comments were noticed as pointing towards a common ground on which to build consensus about ways forward. The representative for the data protection authorities said the legal boundaries were not that difficult to map providing one was able to identify *what information* to be exchanged. It was not enough to just to focus on sources of data without knowing what information was represented. The other comment was from a publisher who said: We have only been discussing what data we would want to have. As a publisher we have data *to give*, *or sell*, but we don't know how this could be done.

3.1 Requirements for design

This short case study of the setting up of a standards project highlights the need for conceptual clarification and the design of a process that will deliver consensus on principles that will level the

playing field for the Norwegian LA actors. Also the case study demonstrates the need to bridge the gap identified in the review of related work (Section 2) between person-related activity data and other types of education data.

In the initial phase of the project a survey of available data sources is foreseen. The challenge, however, is to ensure that this survey identifies the relevant data sources that will be part of actual negotiations between actors in the market who want to extend their existing data sets or make more data available for analysis. In doing so, we need a description of the data attributes that are stumbling blocks for exchange.

It is also clear from this case study that the objective of LA needs to be made explicit in order to focus the search for fitting data. LA is still an ill-defined field of interest. The different stakeholders focus on different data types, e.g., along the range from person-related data to aggregated high-level data on different groups' learning results. If a stakeholder's aim with the LA is not stated, it may be easy to define a data source out of scope because it does not fit a dominant stakeholder interest.

4. Design of conceptual artefacts and process

The aim of this design is to better facilitate a standards development process. The process is set up to arrive at a common and negotiated understanding of what data should be shared in the Norwegian market, giving the different stakeholders increased opportunities to develop LA services for all sectors of education by also pointing towards how it could be achieved.

Therefore, the developed artefacts are designed to make the stakeholders sitting around the table willing and able to share enough information so that they can work out what data in the Norwegian educational market could be made available for exchange between actors, on what conditions. Standardisation is a consensus process, and the intended output, a technical report (ISO 5966:1982), should pave the road for practical progress within the community in question. The test of success is whether new data sources are released for use by others as a result of the consensus documented in the technical report.

4.1 Concepts

The conceptual artefacts we propose are designed to answer specific questions.

I) How to declare specific stakeholder's interest in data sharing?
Stakeholder's position is often given by business interests. Many actors are reluctant to discuss their business models, and therefore, in standardisation settings we often see specific business interests hidden behind more general high-level market concerns. One way of making it easier to discuss business drivers for data sharing is to ask stakeholders to position themselves in the LA landscape. The following instruments are proposed:

1. How do you characterise your interest in LA (referring to the definition of LA given by SOLAR) – are you mainly working to improve *learning* or the *contexts*, in which learning occurs?

If you do not want to choose one or the other, think of one or more typical scenario(s) where you are 1) working directly on providing feedback to the learner, or 2) working on different learning contexts (learning resources, learning design, learning tools provision, physical infrastructure, etc) where the learner is more indirectly influenced.

2. In case of *alternative 1 learning*, Figure 1 gives a simple model of the LA cycle, focussing on data and metrics informing learner interventions.
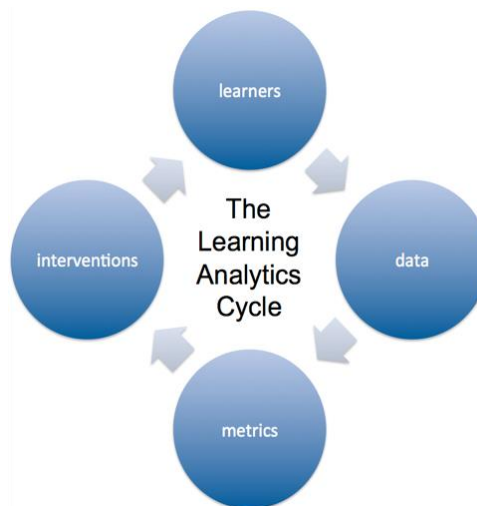
Figure 1. The Learning Analytics Cycle (Clow, 2012)

Starting with *metrics* or analytics, what insight are you looking for in the specific LA use case or scenario? Think of the visualisations, dashboards, or lists coming out of the analytics and note potential insights into learning behaviour, navigation through learning resources, students 'at risk', assessment results, motivation, etc.

Given you have a pretty good idea of your metrics – what you are looking for – a) what data are you using or envisioning use of now; and b) what data would strengthen your analytics if you would be able to collect them?

List data sources of a) and b) in a table and add a column of Ownership/Control. Some of the data, you, e.g., as a vendor or institution, will be in control of. Other data sources have to be released from an external legal body. Categorise your data sources; sort the list; add a third column *Sharing Issues*, and identify issues that could contribute to or block your sharing (of the data you control yourself) or your access (to external data).

3. In the case of *alternative 2 contexts*, Figure 2 gives a template model of learning and its contexts. Revise the model to suit your LA scenario(s): What context do you want to improve?
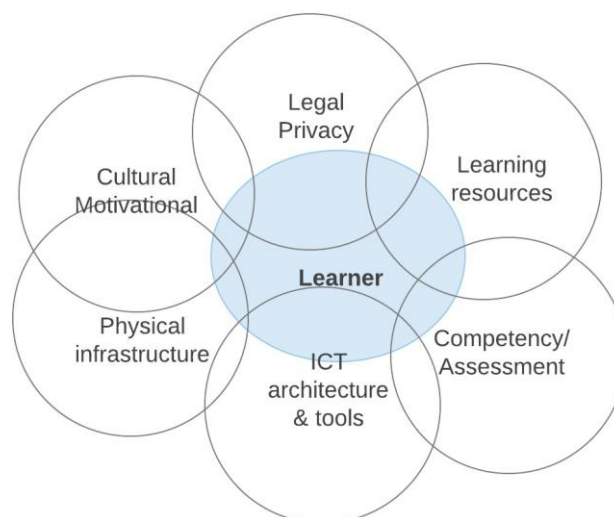

Figure 2. Learning Contexts (template model)

There are a number of data schemas describing learning contexts. E.g., there are metadata standards for describing learning resources and competency structures. Often these standards overlap or interchange, e.g., a learning resource is targeting a specific learning objective.

First, list all data schemas relevant to your chosen learning context. In particular, look for connection points. E.g., in a description of a textbook there is information about class level, which

hooks up to a specification of class structure in schools, which in turn is related to descriptions of curricula for each level, which are broken down in specific learning objectives, and so on.

Second, when you have an idea of which data schemas that are relevant for describing your learning context, create a table with the 2nd column stating who populates the data schemas. Some of the data will be produced in your organization; some will be external. And some of the external data will be more static, e.g., the Norwegian Directorate for Education and Training has a curriculum service that is available for queries via a defined API.

Third, add a column to your table registering sources of *dynamic information* about the learning context, information, which you find crucial for establishing the quality of the context. Look at the interchange points (described as 'hooks' above) for clues; e.g., a data stream that relates activity data coming from a tools log to learning objectives; to a physical installation or artefact; or to a assessment register could provide information that are useful to understand how the learning context performs in supporting learning. In carrying out this third step, you will be challenged to come up with hypothesis about the improvement potentials of the learning context in question.

Fourth, add another column to your table for your ideas on who are the data controllers for the dynamic information you have identified as crucial for your project. Reflect also on the issues that could contribute to or block sharing and access to this data.

II) What are the barriers for access to and sharing of crucial data identified using instrument I?

Use the information gathered in the previous step to list barriers for access to and sharing of the data you find important. Each barrier should be evaluated in terms of how it is related to PII issues. Is it necessary to gather information on persons, e.g., because you would like to merge datasets? Or could you do with sets of aggregated data?

Barriers could be described at different levels, e.g., technical, semantic, organisational, political or legal. Use this classification as a scaffold in creating your list of barriers.

III) What are the enablers for access to and sharing of crucial data identified using instrument I and II?

Based on the information gathered in the previous steps, what are your ideas for solutions? The task at hand is to make different actors interoperate, directly or indirectly in order to improve learning and its contexts. The solutions are found at different levels, the same as for the barriers (e.g., technical, semantic, organisational, political or legal). This classification could also be used to scaffold the brainstorming of enablers. However, it is also useful to think about the broader stakeholder picture for this endeavour. In most societies education is shared responsibility, with actors that take on different tasks. For example, national authorities may be challenged to build 'trust architectures' that could make it easier for actors A and B to share data.

4.1 Process

In preparing the standards work described in the case study (Section 3), a simplified use case approach was chosen. A use case driven approach has the advantage that it helps to cope with the complexity of the requirement analysis process; however, the disadvantage is the lack of synthesis (Regnell, Kimler, & Wesslén, 1995). The standardisation group needs to manage the complexity of different data, diverging data description schemas, data governance, etc.; and synthesis is not that important in the beginning of a standardisation process. What makes us put less emphasis on gathering use cases in the initial stage of this work is the issue of motivation. There is a need to make sure that different stakeholders want to extend their perspective beyond current business models and are willing to expose their future visioning to other stakeholders.

We have identified the challenge of sticking to and developing one's own stakeholder perspective, and the challenge to go to the core of learning analytics (the improvement of learning and its contexts) as the main obstacles for a successful consensus process and output. Therefore, in designing the process (Figure 3) we have made sure these challenges are addressed.

The process model describes a spiral process, starting with mapping stakeholder interests. For the initial round, the model proposes to form two separate subgroups for mapping of stakeholder interest, one with more focus on adaptive learning (vendors, developers, and publishers), and one with

a more academic analytics perspective (school agencies, local authorities and universities). For the second round we suggest that the whole group meets together to compare notes and maintain perspectives and interests throughout the range of LA practices.
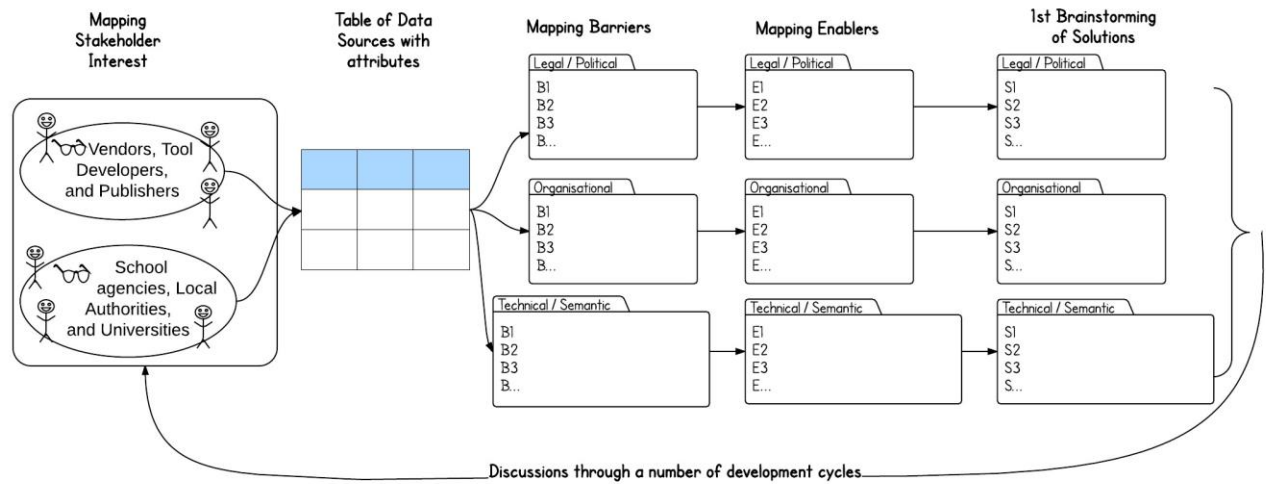

Figure 3. Draft process model of standardisation work related data sharing for LA

The discourse is captured in a shared table, and the joint group proceeds to map barriers and enablers (instrument II and III above). The last subprocess before going back to checking stakeholder interests and the potential for new data sources is the first brainstorming of data sharing solutions. Also for solutions it is suggested that the discussion should be structured according to the different interoperability levels used in the previous discussions.

5. Discussion

Cargill (2011) described standardisation as a poorly understood discipline in practice. "While there are excellent studies of standardization as an economic phenomenon, or as technical a phenomenon, or as a policy initiative, most of these are ex post facto and written from a dispassionate academic view. They are of little help to practitioners who actually are using and creating standards" (Cargill, 2011).

This study is written with the persons actually creating the standard in mind, "working in an area of imperfect knowledge, high economic incentives, changing relationships, and often, short-range planning (Cargill, 2011). At the start of the process much energy is used to get the right stakeholders on board, and agreeing upon rationale and scope. It is the authors' experience from many years of participation in national standardisation that it is unexpectedly hard to move beyond the initial phase of general knowledge sharing into actual technical work based on real stakeholder positions. Often the subject field is complex, and there are lots of technologies in the market waiting to be explored and better understood. With a heterogeneous group of participants it is easy to get stuck in seminar style meetings, where it is undemanding to agree upon trends, but arduous to create consensus on new technical specifications. In the case of this LA technical report it also seems to be an issue that this new field of interest needs to be justified vis-à-vis top management, new customers, and others.

Scoping is key to a successful standardisation process (Hoel & Mason, 2012). If arguing *why* a technology is useful is included in the scope, the scope is clearly to wide. The process should be more focussed on questions of *what* and *how*. Nevertheless, there is a need to know why, at least to keep the participants motivated to move beyond mere knowledge sharing. Therefore, in the process designed in this paper we have defined questions of metrics and analytical outcome to precede the questions of what data could be collected. It is crucial that we are able to base the work on actual needs for analytics originating from each actors' core business.

The scope of the project in our case study includes also *conditions and solutions for data sharing*. In the process we have designed the discussion of solutions to come as a brainstorming

exercise at the end of the process cycle. At the initial stage of this specification work we would like to downplay the role of sharing solutions and conditions till we have a good grasp of what data sources are available. We see in the discussion that some data sources are described as *not available* simply because there are no solutions developed that would make sharing possible. In such cases it is useful to have a short brainstorming about solutions in order to put the source on a roadmap for data sharing. We also see that some data sources are taken for granted, while it would be easy to come up with scenarios that would scatter that impression (e.g., related to re-identification of anonymised data). A brainstorming of solutions would also help in this discussion. The main purpose of the process described in Figure 3 is to identify data sources and start the next phase of mapping sharing conditions and requirements for sharing solutions.

For barriers, enablers and solutions we propose to use the interoperability levels from the European interoperability framework (IDABC, 2004) as a scaffold for the discussion. The framework reminds us that interoperability is not only a technical question; it is also about agreeing upon using the same concepts, harmonising business cultures, agreeing upon common policies, and developing rules of law in order to level the playing ground for a well functioning market.

6. Conclusions and further work

This study has designed conceptual artefacts and a process to support the initial discourse of a standardisation group organised to draft a technical report on data sharing and interoperability for learning analytics. The case study informing the design is set in a Norwegian context; however, both the challenges addressed and the contributions of this research are international in scope. It is one of the paradoxes when big data comes to school that without work on data models and interoperability there are only small data available for learning analytics, "the data [will] remain isolated in self-referencing islands" (Cope & Kalantzis, 2016).

As the field of LA is maturing we will move from big data to meaningful data, where the LA community becomes "more focused on broad research from many data sources and targeting many nuanced questions about what it can deliver" (Merceron, Blikstein, & Siemens, 2015). The challenge for a local market with a mixed stakeholder group, ranging from advanced tool developers looking for an international market for their cutting edge technologies to school authorities wanting to evidence-base their assessment policies, is to agree on what questions to ask. Knowing that it is the questions that lead to fruitful data sources, we need to design a consensus process that pick the low-hanging fruits without losing sight of the big promises of learning analytics.

The developed conceptual artefacts and the process will be tested in the Norwegian standardisation project. Already now we see the need for further development, and we will point to two obvious cases: one related to our understanding of data that are difficult to share, and one related to the use of existing technical infrastructure in Norway.

A data source (or we might use the more precise term data catalogue) contains a number of datasets. Not all of these datasets are problematic in terms of data sharing. But some are, and how do we pinpoint these aspects so that we can start to design solutions for easier access? It is clear to us that a group of these problematic aspects relates to personally identifiable information (PII). However, we would suggest there is a need for a LA specific conceptual model of this phenomenon. We cannot see that this model exists, and we think such a model would have helped the current work on data sharing and interoperability for LA.

Looking ahead to solutions that could support data sharing, we see that Norway have a good technical infrastructure for education that could be used. We have a identity management system now being expanded to include a API gateway[1] connecting data sources and end-user applications. This infrastructure could be used to solve privacy and data protection issues allowing market actors to exchange data without compromising PII (Hoel & Chen, In press; 2016; 2015).

References

---

[1] www.uninett.no/en/service-platform-dataporten

Baepler, P., & Murdoch, C. J. (2010). Academic Analytics and Data Mining in Higher Education. International Journal for the Scholarship of Teaching and Learning,

Cargill, C. F. (2011). Why Standardization Efforts Fail. The Journal of Electronic Publishing, 14(1). http://doi.org/10.3998/3336451.0014.103

Chatti, M. A., Dyckhoff, A. L., & Schroeder, U. (2012). A reference model for learning analytics. International Journal of Technology Enhanced Learning (IJTEL), Vol. 4, No. 5/6 http://doi.org/10.1504/IJTEL.2012.051815

Clow, Doug. (2012). The Learning Analytics Cycle: Closing the Loop Effectively. In

Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK12), Vancouver, BC, Canada, April 29 May 02, 2012, edited by Simon Buckingham

Cooper, A. & Hoel, T. (2015) Data Sharing Requirements and Roadmap. Deliverable 7.2 from LACE project, online at http://www.laceproject.eu/deliverables/d7-2-data-sharing-roadmap.pdf

Cope, B., & Kalantzis, M. (2016). Big Data Comes to School: Implications for Learning, Assessment, and Research. AERA Open, 2(2), 1–20. http://doi.org/10.1177/2332858416641907

Ferguson, R., & Shum, S. B. (2012). Social learning analytics: five approaches. the 2nd International Conference (pp. 23–33). New York, New York, USA: ACM. http://doi.org/10.1145/2330601.2330616

Hevner, A. (2007). A Three Cycle View of Design Science Research. Scandinavian Journal of Information Systems, 2007, 19(2):87-92, 1–6.

Hevner, A., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. Mis Quarterly, 28(1), 75–105.

Hoel, T., & Chen, W. (In press) Privacy and Data Protection in Learning Analytics – Suggestions for Systems Design

Hoel, T., & Chen, W. (2016). Privacy-driven design of learning analytics applications: Exploring the design space of solutions for data sharing and interoperability. Journal of Learning Analytics, 3(1), 139–158. http://dx.doi.org/10.18608/jla.2016.31.9

Hoel, T. & Chen, W. (2015). Privacy in Learning Analytics – Implications for System Architecture — In Watanabe, T. and Seta, K. (Eds.) (2015). Proceedings of the 11th International Conference on Knowledge Management. ISBN 978-4-9908620-0-8

Hoel, T., & Mason, J. (2012). Deficiencies of scope statements in ITLET standardization. In G. Biswas et al. (Eds.) (2012). Proceedings of the 20th International Conference on Computers in Education. Singapore: Asia-Pacific Society for Computers in Education

Hoel, T. & Pawlowski, J.M. (2012). Managing standards development in emergent fields of technology innovation – a proposed model of key processes in ICT standardisation, in Enterprise Interoperability - Proceedings of the Workshops of the Sixth International Conference I-ESA '12, ISTE Ltd, London.

Hoel, T., & Hollins, P. A. (2011). Standards for Learning, Education and Training – a proposal for an improved process (pp. 1–3). Presented at the ICALT 2011. http://doi.org/10.1109/ICALT.2011.146

IDABC. (2004). European interoperability framework for pan-European eGovernment services. Version 1.0. Luxembourg: Office for Official Publications of the European Communities. Retrieved from http://ec.europa.eu/idabc/servlets/Docd552.pdf?id=19529

ISO 5966:1982 Documentation -- Presentation of scientific and technical reports. http://www.iso.org/iso/catalogue_detail.htm?csnumber=12160

Kitto, K., Cross, S., Waters, Z., & Lupton, M. (2015). Learning analytics beyond the LMS (pp. 11–15). Presented at the the Fifth International Conference, New York, New York, USA: ACM Press. http://doi.org/10.1145/2723576.2723627

LAK11 (2011). Call for Papers. 1st International Conference on Learning Analytics and Knowledge. Online at https://tekri.athabascau.ca/analytics/call-papers

Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. Computers and Education, 54(2), 588–599. http://doi.org/10.1016/j.compedu.2009.09.008

Merceron, A., Blikstein, P., & Siemens, G. (2015). Learning analytics: From big data to meaningful data. Journal of Learning Analytics, 2(3), 4–8. http://doi.org/10.18608/jla.2015.23.2

Regnell, B., Kimbler, K., & Wesslen, A. (1995). Improving the use case driven approach to requirements engineering. Presented at the RE '95: Proceedings of the Second IEEE International Symposium on Requirements Engineering, IEEE Computer Society.

Rienties, B., Toetenel, L., & Bryan, A. (2015). "Scaling up" learning design: impact of learning design activities on LMS behavior and performance (pp. 315–319). Presented at the LAK '15: Proceedings of the Fifth International Conference on Learning Analytics And Knowledge, ACM. http://doi.org/10.1145/2723576.2723600

Van den Eynden, V., and Bishop, L. (2014). Incentives and motivations for sharing research data. A Knowledge Exchange Report. Retrieved from: knowledge-exchange.info/Default. aspx?ID=733