

Privacy-driven design of Learning Analytics applications – exploring the design space of solutions for data sharing and interoperability

Tore Hoel

Oslo and Akershus University College
of Applied Sciences
Oslo, Norway
tore.hoel@hioa.no

Weiqin Chen

Oslo and Akershus University College
of Applied Sciences, Oslo
University of Bergen, Bergen, Norway
weiqin.chen@hioa.no

ABSTRACT

Studies have shown that issues of privacy, control of data, and trust are essential to implementation of learning analytics systems. If these issues are not addressed appropriately systems will tend to collapse due to legitimacy crisis, or they will not be implemented in the first place due to resistance from learners, their parents, or their teachers. This paper asks what it means to give priority to privacy in terms of data exchange and application design and offers a conceptual tool, a Learning Analytics Design Space model, to ease the requirement solicitation and design for new learning analytics solutions. The paper argues the case for privacy-driven design as an essential part of learning analytics systems development. A simple model defining a solution as the intersection of an approach, a barrier, and a concern is extended with a process focussing on design justifications to allow for an incremental development of solutions. This research is exploratory of nature, and further validation is needed to prove the usefulness of the Learning Analytics Design Space model.

Categories and Subject Descriptors

I.6.4 [Computing Methodologies]: Model Validation and Analysis

H.1.2 [User/Machine Systems]: Human Factors

J.1 [Administrative Data Processing]: Education

K.4.1 [Public Policy Issues]: Ethics, Privacy, Regulation

General Terms

Design, Human Factors, Standardization, Legal Aspects.

Keywords

Learning Analytics, Privacy, Data Sharing, Trust, Control of data, Privacy by Design, Interoperability

1. INTRODUCTION

Learning analytics (LA) is developing rapidly in higher education, and it is beginning to gain traction in schools, according to foresight analysts [25, 26]. Nevertheless, market players experi-

-ence severe set-backs related to lack of trust in LA systems [30]. A main barrier for mainstream adoption of this technology is revolving around concerns about privacy, control of data and trust [23, 33]. This paper promotes the idea that LA systems development should be based upon a “privacy by design” approach, rather than addressing privacy concerns as an unpleasant afterthought. If systems that have integrated privacy concerns in their designs are prioritised it would help research and development to focus on viable projects, not wasting time and money on blue skies technologies.

Privacy may, however, be defined as out of scope for LA systems and LA interoperability specification development, as one might think that privacy issues are dealt with by front-end systems that provide the data exhaust for analytics. This position is both conceptually and practically flawed. That said, privacy is also an equivocal concept that needs to be understood in context of the emerging LA practices.

“The principles of data protection by design and data protection by default” [13] have recently been built into European and US policies, respectively through the General Data Protection Regulation [14] and Recommendations for Business and Policy-makers from the US Federal Trade Commission [18]. The Privacy by Design (PbD) framework was developed within the Information and Privacy Commission of Ontario, Canada, with goals to ensure “privacy and gaining personal control over one’s information and, for organizations, gaining a sustainable competitive advantage”¹. The PbD framework laid down by Cavoukian [5] encompasses IT systems, accountable business practices, and physical design and networked infrastructures; following seven foundational principles, among them Privacy as the Default Setting; Privacy Embedded into Design; Full functionality; End-to-End Security; Visibility and Transparency; and User-Centricity.

As long as these principles are maintained as high level concepts that are left open to be defined by the organisation seeking “competitive advantage”, the PbD approach will have difficulties leaving any footprint on a particular domain. The principles need to be applied in context, both in terms of domain (in our case learning), and design (i.e., systems engineering) activities. This paper aims to develop a design process model that will make it easier to create privacy-aware designs for learning analytics.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

LAK’15, March 16–20, 2015, Poughkeepsie, NY, USA.

Copyright 2010 ACM 1-58113-000-0/00/0010 ...\$15.00.

¹ PbD - 7 Foundational Principles, see www.privacybydesign.ca/index.php/about-pbd

The paper is organised as follows: A literature review looks into how privacy issues are dealt with in LA, and why contexts are an important backdrop for understanding privacy. Based on this review the concept of a LA Design Space is developed. The model is given a first validation through analysis of a few small case studies on concerns and barriers for data sharing. After reflecting on the results of the walk through of the model, ideas for future work conclude the paper.

2. RELATED WORK

Is privacy recognised as an issue in current LA research? If the 57 papers presented in March 2014 at LAK14, the most representative conference for the LA community, are used as evidence one may say that privacy is recognised, but only superficially so. 12 of 57 papers mentioned privacy, three of them describing how data was anonymised to protect privacy. The rest of the papers were concerned with privacy as a barrier [16]; a restriction for data tracking [10]; and a property of a cluster of stakeholder concerns revolving around risks [9]; however, privacy is clearly an obstacle that could be overcome in order to reap the benefits of LA. “Learners need to be convinced that they [LA systems] are reliable and will improve their learning without intruding into their privacy” [17]. “Many myths surrounding the use of data, privacy infringement and ownership of data need to be dispelled and can be properly modulated once the values of learning analytics are realized” [2]. Some authors reminded that one should be mindful (of privacy) when designing user interfaces [1]. In doing so, another paper pointed out that while ethics and privacy are features of educational data sciences, public entities (at least in the U.S.) are required to adhere to the Family Educational Rights and Privacy Act (FERPA) and other regulations, whereas “in the private sector there are fewer restrictions and less regulations regarding data collection and use” [32]. One paper called for ethical literacy by LA knowledge practitioners, “maintaining an ethical viewpoint and fully incorporating ethics into theory, research, and practice of the LAK discipline” [35].

Privacy is hardly defined in the LAK community papers we have analysed. In order to bring privacy more to the centre of LA application design there is a need to unpack this socio-cultural concept. Privacy in LA is related to how data are used, stored, and exchanged. When data contain information that can be linked to a specific person, we talk about “personal data”. We also talk about “private data”, data that are part of a person’s privacy. The boundaries put around personal and private data are social agreements that depend on who the person is and in what social setting the data are created. A key question is who is the owner of the data. The answer to this question certainly involves the person at hand, but to leave the control to this person alone is often a too simple solution.

Heath, discussing contemporary privacy theory contributions to LA found that the “debate regarding privacy has swung between arguments for and against a particular approach with the limitation theory and control theory dominating” [20]. Control theory focuses on allowing individuals to control their personal information, while limitation theory is concerned with the limitations set on the persons who could gain access to personal information. Heath puts more confidence, however, in theories that highlight contexts as the organising concept, one of the contexts being LA. An international workshop on the future of privacy [8] concluded that there will be an increased acceptance of sharing data for common good, increased social and public

value, with a following likely evolution of the notion of privacy: from “ability to control one’s personal information” (collection, disclosure, use) to “a dynamic process of negotiating personal boundaries in intersubjective relations”. Then a good understanding of the meaning of ‘context’ is needed.

Helen Nissenbaum has moved the privacy debate beyond ‘control’ and ‘limitation’, promoting respect for context as a benchmark for privacy online [31]. Her theory of contextual integrity is a theory of privacy with regard to personal information “because it posits that informational norms model privacy expectations; it asserts that when we find people reacting with surprise, annoyance, indignation, and protest that their privacy has been compromised, we will find that informational norms have been contravened, that contextual integrity has been violated” [31]. Context is an elusive concept that needs to be defined. Nissenbaum has studied privacy contexts that shape privacy policy, i.e., context as technology system or platform; context as business model or business practice; context as sector or industry; and context as social domain.

In the discourse on LA and interoperability it is natural to focus on technical characteristics as the context, e.g., properties defined by respective media, systems, or platforms that shape the character of our activities, transactions and interactions. “If contexts are understood as defined by properties of technical systems and platforms, then respecting contexts will mean adapting policies to these defining properties” [31]. However, Nissenbaum does not think the best solution is to develop privacy context rules for Twitter, Facebook, specific learning applications, etc. She aspires to promote respect for contexts, understood as respect for social domains, as it “offers a better chance than the other three [technology system, business model, or industry sector] for the Principle of Respect for Context to generate positive momentum for meaningful progress in privacy policy and law” [31].

Willis, Campbell and Pistilli [37] seem to be well aligned with Nissenbaum’s contextual integrity theory in their paper exploring the institutional norms related to using big data in higher education, particularly for predictive analytics. They applied a Potter box² approach to ethical reflection, and concluded that “the institution is responsible for developing, refining, and using the massive amount of data it collects to improve student success and retention”. Furthermore, “the institution is responsible for providing a campus climate that is both attractive and engaging and that enhances the likelihood that students will connect with faculty and other students” [37].

From a contextual integrity perspective, the institution may not have violated the informational norm if the roles of the actors involved, e.g., students, teachers, administrators, are acknowledged; the agreed information types were used; and the agreed data flow terms and conditions were followed. Actors, information types, and transmission principles are the three key parameters offered by Nissenbaum for describing a context in terms of integrity and informational norms.

By looking at education as a social domain instantiated in a number of specific contexts the tools provided by Nissenbaum’s privacy theory are well suited to analyse the design space for LA applications, providing privacy is chosen as a key foundation for application development.

² http://en.wikipedia.org/wiki/Potter_Box

The next sections will develop the concept of a design space as a tool for discussing design constraints of LA solutions.

3. A DESIGN SPACE FOR LA DATA SHARING

In looking for the low-hanging fruits of LA Interoperability Hoel and Chen [23] built on Enterprise Architecture theory and came up with a concept of a solution space as the three dimensional model describing concerns, barriers and solutions (Figure 1). In this paper this concept is further developed into a LA design space (LADS). It is understood as a range of potential designs that could solve identified LA problems, e.g., related to privacy, control of data, and trust. These designs are justified according to a design space analysis. MacLean et al. [27] presented design space analysis as an approach to represent design rationale, focussing on three aspects: Questions, Options, and Criteria. Questions are key issues for structuring the space of alternatives; Options are possible alternative answers to the Questions; and Criteria are the basis for evaluating and choosing among the Options.

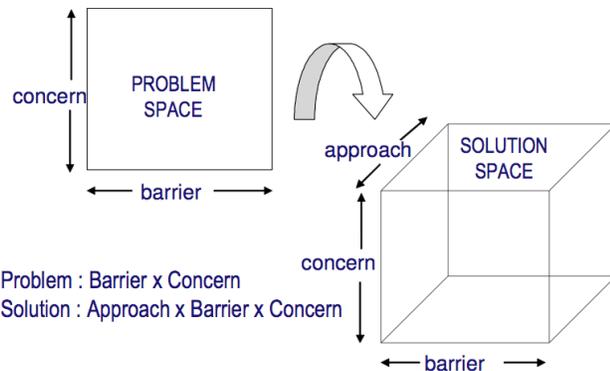


Figure 1. Solutions as the intersection of approaches, barriers, and concerns [6].

This paper will carry out a first development and validation of the LADS model. This research is positioned in the first Relevance Cycle of the three research cycles of Design Science [21, 22], addressing requirements and field-testing. The purpose is to come up with a model that could make the ideas of PbD more relevant for LA solutions promoting data sharing and interoperability.

3.1 Learning Analytics Design Space model

The LA Design Space model is based on a three-step process, identifying Concerns, Barriers, and Design solutions.

The concerns related to data sharing and interoperability revolve around issues of privacy, control of one's own data, and trust to applications and service providers [23]. At this stage and for this conceptual study these issues suffice to populate the model.

The barriers related to data sharing and interoperability are part of the challenges of scaling up LA. As Ferguson et al. [17] observe there are currently few reports in the LA literature of deployment of scale. For the purpose of this paper the barriers are derived from a small case study on available data based on a variety of sources, like educational data specifications, sensor data APIs, activity stream specifications, LA typologies, etc.

The design space is made up of solutions that can be justified through analysis of questions, options and criteria for selection.

In the following we will select some data as input for a first demonstration of the viability of the model.

4. DATA FOR CONSTRUCTING THE PROBLEM SPACE

In order to conduct a first run through of the model we can select the data at hand, not worrying about representativeness. For this paper a series of short case studies was conducted following the data track of LA. LA begins and ends with data; first, by data originating from the learner, and last, by actions of the LA system and follow-up actions by the learner feeding into the next LA cycle. In between, the data have been stored in a standardised format of sorts, and have been subject to data clearance procedures following national, institutional or company rules and regulations.

A study of the data elements of the US Common Education Data [7] concludes that much of the data residing in Student Management Systems or Learning Activity Record Stores are not imbued with privacy issues raised by introduction of new LA practices. Of course, there are issues related to identification of a person that are sensitive; and aggregation of disparate data about a person can always be felt as a threat, especially if one loses trust in the system itself. However, these data have been around in education for time immemorial without causing too much concern. It is the learning process data, sitting in the crossing between organisations, people and learning resources (Figure 2) that now with LA have become so much more important.



Figure 2. CEDS Normalized Data Set Conceptual Model [7].

Process data are, as observed in new LA applications, captured in formats defined in activity stream specifications, e.g., ADL Experience API, Tin-Can, IMS Caliper³. These specifications establish a core language to describe activities through providing information on subject, verb, object, context, etc. On top of these core specifications there are community profiles providing specialised vocabularies for educational settings like schools, higher education, workplace training etc. With a powerful and extensible core language one is, in principle, able to describe any

³ www.adlnet.gov/tla/experience-api, tincanapi.com, www.imsglobal.org/caliper

activity, which opens up the question of what LA practitioners want to describe.

Fergusson and Shum [15] introduced five categories of analytics that make use of five partly overlapping classes of data:

- **Social Network** (analyses relationships, using data about identifiable persons and their activities, e.g., publishing papers, participating in social platforms, etc.)
- **Discourse** (analyses language as a tool for knowledge negotiation and construction, using full-text data from discussion fora, talk, and other written text sources)
- **Content** (analyses user-generated content, using data from web 2.0 applications)
- **Disposition** (analyses intrinsic motivations to learn, using a range of activity data, in principle generated by all the tools used by the learner during day and night)
- **Context** (considers formal and informal learning, based on data describing the contexts within which learning happens, e.g., use of tools, educational setting, groups, etc.)

Most of the different types of analytics described by Ferguson and Shum [15] would not be possible without data from social software, also called Web 2.0 applications. With mobile devices now in nearly any student's pocket, use of social software services is part of everyday life, also on campus or in the classroom. Even if institutional policies should try to restrict use of these services for formal education, they are often used.

Garaizar and Guenaga [19] explored how HTML5 browser APIs could shed some light on how the use of web apps in mobile environments had potential to enhance learning. The APIs to different sensors in wearable computers (mobile phones, wrist bands, watches, etc.) open up a range of new data sources. Table 1 lists the data types used by HTML5 APIs and derives questions as to what pedagogical use or use for learning analytics these data types could have.

Table 1 Data types in HTML5 APIs and their potential use for LA

Data type	Information provided	Potential Questions from a LA point of view
Geolocation: Your geographical position according to your smart phone, accessed by apps, browser (HTML5)	Latitude / Longitude changes	Is the learner at school or at home? Is she commuting? Where does the learning take place?
3D Orientation	Acceleration changes	Is the context suitable for learning?
Battery	Status of battery, charging	Does the battery status affect the learning context? How?
Network	Cost of the	Does the cost of the

information	network access	network access disrupt the learning scenario? How?
Offline & online events	Connectivity status	Which are the problems caused by the lack of continuity in connectivity?
DOM storage: File, Indexed database	Local storage	What did the learner do when she was offline? Did it affect the learning process?
Ambient light	Surrounding light to the learner	Is the learning environment suitable for learning or more suitable for relaxation?
Temperature	Temperature around the learner	Is the learning environment suitable for learning?
Atmospheric pressure	Height above ground	Is the context suitable for learning?
Proximity of objects		Are learning aids accessible to the learner during work with a particular app?
Gestures	Swipe, pinch, twist, etc.	What is the learner focused on?
Blood pressure		What is the physical state of the learner during learning events?
Heart beat		What is the physical state of the learner during learning events?
Perspiration		Is the learner nervous?
getUserMedia	Native Access to audio & video devices	What is the learner looking at? What is she listening to? How is the learning context in terms of space, luminosity, noise, etc.
WebRTC	Send & Receive multimedia between browsers	How can the multimedia streams be collected, stored, analysed and enriched in real time?
WebVVT	Subtitles and audio descriptions	What is the impact of adding supplementary textual information to multimedia streams?
Animations (CSS, SMIL, rAF, SVG, Canvas 2D, WebGL)	Declarative and Procedural animations	What is the impact of adding supplementary visual information to multimedia streams?

Timers (high resolution, user, resource, navigation)	timestamps per millisecond	How long does it take to perform an action (download a learning activity, render a web app, etc.)? Is the learner multitasking? Is she bored? Is she cheating via automatic responses?
DOM 4 mutation observes, drag & drop events, focus	Fine-grained user interactions	Which web controls are easy or hard to use? Which are the gestures and/or complex interactions preferred by learners?
Page visibility, fullscreen, pointer lock	Mono task / multitask scenarios	Is the learner multitasking? How? When? Do mono task / multitask activities enhance learning?
History	History of web session	Is the workflow of the learning app appropriate?

Following the data trail, literally speaking, from the headmaster's filing cabinet to the pocket of the learner has changed our focus of analysis away from the data elements and their potential privacy issues to data in contexts. Privacy is not a unidimensional concept describing the relationship between the data element and the person this element holds information about. By bringing in the context dimension, we see that data belong to more than the person described, and it is the characteristics of the setting (context) that impact the privacy concerns.

The case studies have shown that the context of formal study or teaching is essential as it establishes the boundary for what is within or outside the scope of data available for learning analytics. If this boundary is crossed, e.g., by introducing social software services run by a third party, this can only happen by individual consent from case to case. This is an institutional perspective. From an individual or a third party perspective this boundary may be less definitive, which leads to tensions among different stakeholders in the use of LA to support learning. However, the boundaries between formal and informal learning are far from clear, as Malcolm et al. [28] have demonstrated. They found (and that was before social media took off in learning) "a complete lack of agreement in the literature about what informal, non-formal and formal learning are, or what the boundaries between them might be" [28].

The input for constructing the Problem Space is concerns and barriers. The first workshop on LA at ICCE2014 expanded on the privacy, control, and trust cluster of issues referred to above [23], and mapped concerns [29]. There are concerns that point in the direction of restrictive sharing of data and putting a cap on services that interoperate. However, there are also concerns about not being able to reap all the benefits of LA, "understanding and optimizing learning" [11]. These are benefits that are directly in the interest of the learner, who wish to be in control of her data. We see that we have multiple stakeholders with legitimate interests, and that the eventual solutions will have to balance the interests of more than one party.

Concerning barriers, Educause Center for Applied Research [3] identified in a 2012 study on analytics in higher education four major challenges to achieving success with analytics: affordability, data, culture, and expertise. From an institutional perspective cost is the main obstacle, however, factors like misuse of data, regulations requiring use of data, inaccurate data, and individuals' privacy rights are barriers that higher education leaders worry about in a situation where they are collecting more data than ever before [3].

5. CONSTRUCTING THE LEARNING ANALYTICS PROBLEM SPACE

Based on the concerns and barriers derived from the selected case studies the following problem space for LA data sharing is constructed:

Two concerns are pulling the 'data sharing slider' in opposite directions: Prioritising privacy and individual control of data tends to limit data sharing, while wanting to take advantage of the latest personal learning app on the market is an invitation to tick a number of 'give-access-to' boxes.

The barriers are related to the concept of a 'user in context'. Informal and individual learning leaves the decisions of giving access to personal data to the user, and is a matter of appreciation of benefits; feeling of control; trust in applications, companies, and institutions; etc. In the current situation, individuals seem to be more willing to take risks and go for new and innovative solutions. While formal learning is led by institutions wanting to have 'ethical use of student data policies' in place, and sticking to institutional learning platforms that only use a limited set of data sources for LA. For the institutions, lack of privacy frameworks is a major barrier against data sharing and the use of the sensitive data sources available to commercial LA providers.

The barriers seem to be more socio-cultural or organisational than technical or legal (to use the European interoperability framework dimensions [24]). However, the solutions to the problems will need to address all these interoperability challenges.

6. CONSTRUCTING THE SOLUTION SPACE AND CHOOSING DESIGNS

Concerns and barriers drive the design activities towards solutions. Which solution to choose is dependent upon the design directions given by the design space analysis. The following attempt is a first iteration of the last process of the LADS model, Figure 3.

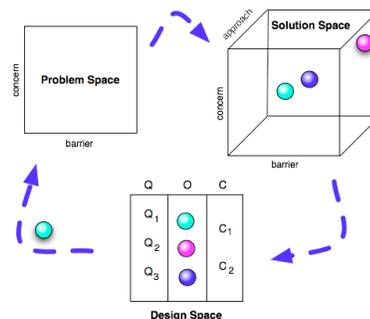


Figure 3. The Learning Analytics Design Space Model.

The solutions are found by addressing the concerns and breaking the barriers, which we in our case define as technical, socio-cultural, and legal. Going for a ‘radical’ alternative, with use of a variety of data sources and a high degree of data sharing, we could see these tentative solutions based on requirements from the case studies:

- **Technical:** Design of a specification allowing user to express detailed conditions for data sharing when signing in to LA applications, with opt-out possibilities
- **Socio-cultural:** To boost trust in LA systems, development of privacy declarations, industry labels guaranteeing adherence to privacy standards, and other means of supporting customer dialogue about privacy
- **Legal:** Ownership and control of data from learning activities are strengthened in national and international law

What solution should be suggested for design? The design space analysis starts with questioning the rationale of a project, as a refinement of the problem space analysis. For our purpose, we maintain the ambitious goal of using applications supporting personalised and adaptive learning. Furthermore, we ask if the solution is safe from ‘loosing face’ through leakage of personal information. And we ask if the solution supports ubiquitous learning by allowing both formal and informal learning by the same application.

It is the criteria for which options to choose that drive the design process based on the identified solutions. The PbD approach advocated above [31] gave priority to the social domain as the context to explore – to see if contextual integrity has been maintained when data are shared. Therefore, we will ask: Does the proposed option pass the test of having been subject to an informed public deliberation on the benefits of LA and the consequences of data sharing for the user as well as for the institution, the service provider, etc?

In the case of *technical solution* proposed above, we see that the design must go beyond a quick technical fix solving the problem once and for all, giving the user absolute control. The institution, e.g., the school, the university, should have a say, also being responsible for catering for the greater good, the class or group, the parents, the society. Therefore, technical solutions should include an element of permanent negotiations, thus requiring solutions that are transparent and simple.

The *legal solution* is also an option that will not get first priority. Of course, solutions must have legal backing. But the privacy concerns about data sharing are not solved by legal measures alone. By the criteria we have chosen, our analysis points to the socio-cultural domain for solutions and design requirements.

A *socio-cultural design solution* must focus on the communication between user and system/service provider. Trust is not a ‘thing’ that is negotiated once and lasts forever; it must be renegotiated again and again. Therefore, especially in an environment that is dynamic and crowded with actors with different interests, large-scale, complex and non-transparent solutions will be challenged. It will be easier to maintain context integrity with smaller solutions. Smaller LA solutions may seem to be a contradiction in terms, as the ideas of big data and data sharing across systems often lead to plans for large-scale solutions, may be with a centralised Learning Record Store aggregating data from a number of systems. Nevertheless, if

maintaining trust is pivotal to LA systems in the current stage of development, our design space analysis concludes that the socio-cultural aspects of negotiating access to data should direct the design of technical solutions, legal frameworks, and implementation. With that as a result of the first design cycle of the LADS model, new concerns and barriers should be mapped in order to arrive, after several iterations, at an implementable design.

7. DISCUSSION

Educational institutions have always used learners’ behaviour and performance data to determine, visualise, and sort strengths and weaknesses of individual learners and groups. What is new with LA, is the ability to process this information in real time and on demand. Furthermore, LA can go far beyond classroom assessment procedures. By doing so, LA is working with data the learner often does not know are being used. LA can be used to compute the relationships between learners based on their interactions, or to compare the commitment of a learner in a course based on time spent on the learning material, or to compare text written by students against pre-existing corpora. Thus, LA affects the privacy rights of learners in a new manner, making it necessary for the learner and the institution to negotiate the boundaries between personal and institutional spaces, between informal and formal learning, and between institutionally provided tools and technology for personal use. As Thomas argued, there is a “need to re-assess the concept of teaching and learning space to encompass both physical and virtual locations, and adapt learning experiences to this new context” [36]. This need is reinforced by the introduction of LA as a support technology. LA is an emerging discipline [34], and most of the technological ideas are still on the drawing table. Therefore, there is a strong need to “do the right thing” from the outset, to avoid setbacks and the need to correct misconceptions and rebuild trust after privacy collapses.

This paper contributes a conceptual tool to ease the requirement solicitation and design for new LA solutions. A simple model defining a solution as the intersection of an approach, a barrier, and a concern was extended with a process focussing on design justifications to allow for an incremental development of solutions. We used Privacy by Design principles to steer the development of ideas for solutions; however, other principles could be used to test alternative design solutions, like pedagogical principles focussing on learning efficacy, learner-centred approaches, ubiquitous learning, etc.

8. CONCLUSIONS AND FURTHER WORK

Privacy Awareness is reported as one of the major features of smart LA when researchers are summarising their experiences ‘from the field’ [12]. LA is a young field, both as research and application design are concerned. New ideas are being launched nearly every day, and there is a need to test these ideas to see if they meet the different stakeholders’ requirements. For example, Kennisnet, a Dutch Governmental School Agency, has chosen PbD principles as a starting point for their design [4]. “Next, we use the open User Managed Access (UMA) standard. The student, or parent for underage students, has a central place and is the owner of his own educational data”, Bomas explains [4]. Will giving the students and parents full ownership to their data using the UMA standard benefit educational goals? In order to find an answer to this question one need to do an analysis of how the

standard is implemented and how the different concerns are addressed.

In this paper we have proposed the LADS model as a tool to answer such questions. The tool allows users to map the problem space and analyse different solutions according to different criteria. The first tentative validation of the model presented in this paper shows that the model has potential to make a requirement discourse on LA applications more fruitful. However, in order to verify this conclusion further testing is necessary. This will be done in the context of the European project on Learning Analytics Community Exchange (LACE), which has identified privacy and ethics as a major theme for community discourse to develop the field of LA.

9. ACKNOWLEDGMENTS

This paper was partly produced with funding from the European Commission Seventh Framework Programme as part of the LACE Project, grant number 619424.

10. REFERENCES

- [1] Aguilar, S. 2014. Perceptions and Use of an Early Warning System During a Higher Education Transition Program. (Feb. 2014), 1–5.
- [2] Arnold, K.E. et al. 2014. Building Institutional Capacities and Competencies for Systemic Learning Analytics Initiatives. (Feb. 2014), 1–4.
- [3] Bichsel, Jacqueline. Analytics in Higher Education: Benefits, Barriers, Progress, and Recommendations (Research Report). Louisville, CO: EDUCAUSE Center for Applied Research, August 2012, available from <http://www.educause.edu/ecar>.
- [4] Bomas, E. (2014). How to give students control of their data. Blog post of August 29th, 2014, online at <http://www.laceproject.eu/blog/give-students-control-data/>
- [5] Cavoukian, A. and Ontario, I.A.P.C. 2014. Privacy by Design.
- [6] Chen, D., & Daclin, N. (2006). Framework for enterprise interoperability. Proc of IFAC Workshop EI2N.
- [7] CEDS. 2014. Common Education Data Standards (CEDS) Version 4 Data Model Guide. Online at <https://ceds.ed.gov/>, retrieved 2015-01-15
- [8] Dartiguepeyrou, C. (2014) The Futures of Privacy, ISBN 978-2-915618-25-9 © Fondation Télécom, Institut Mines-Télécom, February 2014
- [9] Drachsler, H. et al. 2014. The impact of learning analytics on the dutch education system. (New York, New York, USA, 2014), 158–162.
- [10] Drachsler, H. et al. 2014. The Learning Analytics & Knowledge (LAK) Data Challenge 2014. (Feb. 2014), 1–2.
- [11] Duval, E. 2011. Attention please!: learning analytics for visualization and recommendation (pp. 9–17). Presented at the Proceedings of the 1st International Conference on Learning Analytics and Knowledge - LAK '11, New York, USA: ACM Press. doi:10.1145/2090116.2090118
- [12] Ebner, M. et al. 2015. Seven features of smart learning analytics - lessons learned from four years of research with learning analytics. eLearning Papers. (Jan. 2015), 1–5. Retrieved from [http://www.openeducationeuropa.eu/en/article/Assessment-](http://www.openeducationeuropa.eu/en/article/Assessment-certification-and-quality-assurance-in-open-learning_From-field_40_3)
- [13] European Commission. 2012. On the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). COM(2012) 11 final
- [14] European Commission. 2013. Data Protection Directive (Directive 95/46/EC)
- [15] Ferguson, R. and Shum, S.B. 2012. Social learning analytics. (New York, New York, USA, 2012), 23.
- [16] Ferguson, R. et al. 2014. DCLA14: Second International Workshop on Discourse-Centric Learning Analytics.
- [17] Ferguson, R. et al. 2014. Setting Learning Analytics in Context: Overcoming the Barriers to Large-Scale Adoption. (Feb. 2014), 1–3.
- [18] FTC. 2012. Protecting Consumer Privacy in an Era of Rapid Change: Recommendations For Businesses and Policymakers. Online at <http://www.ftc.gov/reports/protecting-consumer-privacy-era-rapid-change-recommendations-businesses-policymakers>, retrived 2015-01-15
- [19] Garaizar, P. and Guenaga, M. 2014. A multimodal learning analytics view of HTML5 APIs: technical benefits and privacy risks. (New York, New York, USA, Oct. 2014), 275–281.
- [20] Heath, J. 2014. Contemporary privacy theory contributions to learning analytics. *Journal of Learning Analytics*. 1, 1 (2014).
- [21] Hevner, A. (2007). A Three Cycle View of Design Science Research. *Scandinavian Journal of Information Systems*, 2007, 19(2):87-92, 1–6.
- [22] Hevner, A., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *Mis Quarterly*, 28(1), 75–105.
- [23] Hoel, T. and Chen, W. 2014. Learning Analytics Interoperability – looking for Low-Hanging Fruits. In Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education
- [24] IDABC. 2004. European Interoperability Framework. Version 1.0. Online at <http://ec.europa.eu/idabc/en/document/3473.html#finalEIF>, retrieved 2015-01-15
- [25] Johnson, L. et al. 2014. Johnson, L., Adams Becker, S., Estrada, V., Freeman, A. 2014. NMC Horizon Report: 2014 Higher Education Edition. Austin, Texas: The New Media Consortium. (Jan. 2014), 1–52.
- [26] Johnson, L., Adams Becker, S., Estrada, V., and Freeman, A. 2014. NMC Horizon Report: 2014 K-12 Edition. Austin, Texas: The New Media Consortium.: <http://cdn.nmc.org/media/2014-nmc-horizon-report-k12-EN.pdf>. Accessed: 2014-08-27.
- [27] MacLean, A. et al. 1991. Questions, Options, and Criteria: Elements of Design Space Analysis. *Human-Computer Interaction*. 6, (1991), 201–250.
- [28] Malcolm, J. et al. 2003. The interrelationships between informal and formal learning. *Journal of Workplace Learning*. 15, 7/8 (Dec. 2003), 313–318.

- [29] Mason, J. et al. (in press) Questions as data: Outcomes and perspectives from the 1st ICCE Learning Analytics Workshop, 2014
- [30] New York Times. 2014. InBloom student data repository to close. News story 2014-04-21, online at <http://bits.blogs.nytimes.com/2014/04/21/inbloom-student-data-repository-to-close>, retrived 2015-01-15
- [31] Nissenbaum, H. 2014. Respect for Context as a Benchmark for Privacy Online: What It Is and Isn't (Jan. 2014), 1–128.
- [32] Piety, P.J. et al. 2014. Educational Data Sciences – Framing Emergent Practices for Analytics of Learning, Organizations, and Systems. (Feb. 2014), 1–10.
- [33] Scheffel, M., Drachsler, H., Stoyanov, S., & Specht, M. (2014). Quality Indicators for Learning Analytics. *Educational Technology & Society*, 17 (4), 117–132.
- [34] Siemens, G. 2013. Learning Analytics The Emergence of a Discipline. *American Behavioral Scientist*. 57, 10 (Oct. 2013), 1380–1400.
- [35] Swenson, J. 2014. Establishing an Ethical Literacy for Learning Analytics. (Feb. 2014), 1–5.
- [36] Thomas, H. (2010). Learning spaces, learning environments and the dis'placement' of learning. *British Journal of Educational Technology*, 41(3), 502–511.
- [37] Willis, J.E., III et al. 2013. Ethics, big data, and analytics: A model for application. (2013).