

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, Norway

tore.hoel@hioa.no

Weiqin Chen

University of Bergen, Norway

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 a 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 focusing on design justifications to allow for an incremental development of solutions. This research is exploratory in nature, and further validation is needed to prove the usefulness of the Learning Analytics Design Space model.

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 many foresight analysts (Johnson et al., 2016; Johnson, Adams Becker, Estrada, & Freeman, 2014a; Johnson, Adams Becker, Estrada, & Freeman, 2014b; Griffiths, Brasher, Clow, Ferguson, & Yuan, 2016). Nevertheless, market players experience severe setbacks related to lack of trust in LA systems (Singer, 2014; Drachsler et al., 2016). A main barrier for mainstream adoption of this technology revolves around concerns about privacy, control of data, and trust (Hoel, Mason, & Chen, 2015; Mason, Hoel, & Chen, in press; Griffiths, Hoel, & Cooper, 2016; Hoel & Chen, 2014, 2015; Cooper and Hoel, 2015; Scheffel, Drachsler, Stoyanov, & Specht, 2014). 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 were prioritized, it would help research and development to focus on viable projects instead of wasting time and money on blue-sky technologies.

Privacy may, however, be defined as beyond the scope of LA systems and LA interoperability specification development (ADL, 2013; IMS Global, 2015), as one might think that privacy issues are dealt with by front-

(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>

end systems that provide the data exhaust for analytics. This position is flawed, both conceptually and practically. First, privacy cannot be handled only by a sign-on process or a consent form; privacy permeates all processes of the LA process cycle (Hoel, Chen, & Cho, 2016). Second, if privacy requirements are not reflected at the time of design, the developed solutions may not deliver according to law or market needs (Hoel & Chen, 2015). That said, privacy is also an equivocal concept that needs to be understood in context of emerging LA practices.

“The principles of data protection by design and data protection by default” (EC, 2012, p. 27) have recently been built into European and US policies, respectively, through the General Data Protection Regulation (Council Directive 95/46/EC) and Recommendations for Business and Policy-makers from the US Federal Trade Commission (FTC, 2012). The privacy-by-design (PbD) framework was developed within the Information and Privacy Commission of Ontario, Canada, with goals of “ensuring privacy and gaining personal control over one’s information and, for organizations, gaining a sustainable competitive advantage” (Cavoukian, 2012, pp. 36–37). The PbD framework laid down by Cavoukian (2012) encompasses IT systems, accountable business practices, physical design, and networked infrastructures and follows these seven foundational principles:

1. Proactive not reactive; preventative not remedial
2. Privacy as the default setting
3. Privacy embedded into design
4. Full functionality – positive-sum, not zero-sum
5. End-to-end security – full lifecycle protection
6. Visibility and transparency – keep it open
7. Respect for user privacy – keep it user-centric (p. 37)

As long as these principles are maintained as high-level concepts left open to be defined by the organization seeking a “competitive advantage,” the PbD approach will have difficulties in 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.

The paper is organized as follows: Section 2 offers a literature review that looks at how privacy has been the focus of research and discourse within the LA community in the last few years. Contexts and context integrity are identified as an important backdrop for understanding privacy. Based the authors’ previous work, an LA Design Space concept is developed and a model offered as a useful discourse artefact for achieving privacy-driven design of LA (Section 3). In Section 4, the current state of the art related to data sharing is described in the case used in Section 5 to construct a Problem Space, a Solution Space, and, based on these constructs, a Design Space analysis of viable solutions for dealing with privacy in LA. The result is discussed in Section 6, and Section 7 concludes with ideas for further work.

(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>

2 RELATED WORK

Is privacy recognized as an issue in current LA research? The yearly international conferences on Learning Analytics and Knowledge (LAK) are a representative outlet for LA research. Looking at the main conference proceedings of LAK '14 and LAK '15, one may say that privacy is recognized, but only superficially so. However, from 2014 to 2015, we see signs of a new approach that not only identifies privacy as a concern but points to privacy solutions at different levels. At LAK '14, 12 of 57 papers mentioned privacy, three of them describing how data was anonymized to protect privacy. The rest of the papers were concerned with privacy as a barrier (Ferguson, De Liddo, Whitelock, de Laat, & Buckingham Shum, 2014a); as a restriction for data tracking (Drachsler, Dietze, Herder, d'Aquin, & Taibi, 2014b); and as a cluster of stakeholder concerns revolving around risks (Drachsler, Stoyanov, & Specht, 2014a). However, privacy is clearly an obstacle that should be overcome in order to reap the benefits of LA since "Learners need to be convinced that [LA systems] are reliable and will improve their learning without intruding into their privacy" (Ferguson et al., 2014b, p. 251). "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" (Arnold et al., 2014, p. 259). Some authors reminded the audience that one should be mindful (of privacy) when designing user interfaces (Aguilar, 2014). In doing so, another paper pointed out that while ethics and privacy are features of educational data sciences, public entities are required to adhere to FERPA and other such regulations, whereas "in the private sector there are fewer restrictions and less regulations regarding data collection and use" (Piety et al., 2014, p. 198). 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" (Swenson, 2014, p. 250).

One year later, at LAK '15, privacy was still not a major theme (mentioned in 10 out of 82 papers), but the issue was put on the agenda by researchers active in European projects in a panel discussion (Ferguson et al., 2015) and a workshop dedicated to ethics and privacy¹ (Drachsler et al., 2015a). The main conference papers of LAK '15 still looked at privacy as a search term (Sekiya, Matsuda, & Yamaguchi, 2015), a course subject (Vogelsang & Ruppertz, 2015), or an abstract concern (Scheffel, Drachsler, & Specht, 2015), which could limit access to data (Wang, Heffernan, & Heffernan, 2015; Drachsler et al., 2015b), or one that must "be addressed given the larger scale of the tools usage compared with pilot studies" when "testing the tool in-the-wild" (Martinez-Maldonado et al., 2015, p. 6).

However, two papers advocated that institutions "must engage more proactively with students, to inform and more directly involve them in the ways in which both individual and aggregated data is being used" (Prinsloo & Slade, 2015, p. 8). Prinsloo and Slade (2015) explored the concept of student *privacy self-management* and issues around consent and the seemingly simple choice to allow students to opt-in or opt-out of having their data tracked. They concluded that the way forward cannot simply be to introduce a choice between opt-in or opt-out as "Only by increasing the transparency around learning analytics

¹ A majority of the contributions to this special issue of JLA (Vol. 3, No. 1) are based on input to this LAK '15 workshop.

(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>

activities will HEIs gain the trust and fuller co-operation of students” (2015, p. 8).

Kitto, Cross, Waters, & Lupton (2015), the authors of the second paper, discussed privacy vs. data ownership and proposed a technical solution, the Connected Learning Analytics Toolkit, as a radically different solution to current systems in the market since “Many of the ethical problems that arise from within the privacy perspective evaporate when students are given full access to their data” (p. 5). Kitto et al. (2015) referenced a work by Pardo and Siemens (2014) that advocates a contextual approach with respect to information privacy; sometimes we want our information to be public, sometimes not.

No doubt, the upcoming LAK '16 conference will move the research frontier on ethics and privacy for LA; so will outputs from the European LACE project, which has published a Review Report on current issues and their solutions (Drachslar et al., 2016a), as well as this special issue of the *Journal of Learning Analytics*. A preprint of a LAK '16 paper by Drachslar and Greller (2016) promotes a checklist approach to trusted learning analytics building on a number of catch phrases (determination, explain, legitimate, involve, consent, anonymize, technical, external) making up the DELICATE checklist. “[W]e would like to encourage the Learning Analytics community to turn the privacy burden into a privacy quality label,” Drachslar and Greller state, seeing the challenges as “a ‘soft’ issue, rooted in human factors, such as angst, scepticism, misunderstandings, and critical concerns” (p. 5). Referencing the authors of this paper (Hoel and Chen), Drachslar and Greller spell out that they “would refrain from solving a weakness in a new learning technology by proposing technical fixes or technological solutions, such as standardization approaches” (2016, p. 5).

In choosing between soft checklists and hard technical fixes, there is a need for a conceptual tool that could help us move from barriers and concerns to well-argued solutions. The aim of this paper is to develop such a conceptual framework. However, before doing so there is still a need to unpack privacy as a socio-cultural concept to bring it more to the centre of LA application design.

2.1 A Contextual Approach to Privacy

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” 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 settings the data are created and shared. A key question revolves around who is the owner of the data. The answer certainly involves the person at hand, but to leave the control to this person alone is often too simple a solution.

Heath (2014), 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” (p. 142). Control theory focuses on allowing individuals to control their personal information, while limitation theory is concerned with the limitations set on those who

(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>

could gain access to personal information. Heath puts more confidence, however, in theories that highlight contexts as the organizing concept, one of the contexts being LA. At an international workshop on the future of privacy, Dartiguepeyrou 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 the “ability to control one’s personal information’ (collection, disclosure, use) to ‘a dynamic process of negotiating personal boundaries in intersubjective relations’” (2014, p. 13). Thus, a good understanding of the meaning of “context” is needed.

Helen Nissenbaum (2014) has moved the privacy debate beyond “control” and “limitation,” promoting respect for context as a benchmark for privacy online. Her theory of contextual integrity is a theory of privacy regarding 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” (Nissenbaum, 2014, p. 25). Context is, however, an elusive concept that needs to be defined. Nissenbaum has studied the 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” (Nissenbaum, 2014, p. 14). 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” (Nissenbaum, 2014, p. 25).

Willis, Campbell, and Pistilli (2013) 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 concluded, “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” (Willis et al., 2013, p. 6). Recent development of codes of ethics by higher educational institutions shows that the educational systems are responding to the challenges to improve the contextual integrity of their students (Sclater, 2016).

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.

(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>

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 analyze the design space for LA applications, providing privacy is chosen as a key foundation for application development.

3 FROM PROBLEMS TO SOLUTIONS: CONSTRUCTING A LEARNING ANALYTICS DESIGN SPACE (LADS) MODEL

This paper will carry out a first development and tentative validation of the LADS model. This research is positioned in the first Relevance Cycle of the three research cycles of Design Science Research (DSR) (Hevner, March, Park, & Ram, 2004; Hevner, 2007), addressing requirements and field-testing. The purpose is to come up with a model that will make the ideas of PbD more relevant for LA solutions promoting data sharing and interoperability. However, the scope of the LADS model is not limited to issues of privacy, control of data and trust. This initial cycle of DSR process focuses on “generating design alternatives and evaluating the alternatives against requirements until a satisfactory design is achieved” (Hevner, 2007, p. 90). In this paper, we do the first design and testing of the LADS model against requirements solicited through community exchange and analysis of cases derived from LA practices. In order to prove the usefulness of the model, rigorous evaluation needs to be done. Some ideas on how this future research could be done are presented in Section 7.

In looking for the low-hanging fruits of LA Interoperability, Hoel and Chen (2014) built on Interoperability and Enterprise Architecture theories and came up with a concept of a solution space. These theories are concerned with how organizations are able to solve problems by communicating and exchanging information, using the information exchanged, and getting access to the functionality of a third system (Chen & Daclin, 2006). The solution space is conceived as a three-dimensional model, describing concerns, barriers, and solutions (Figure 1).

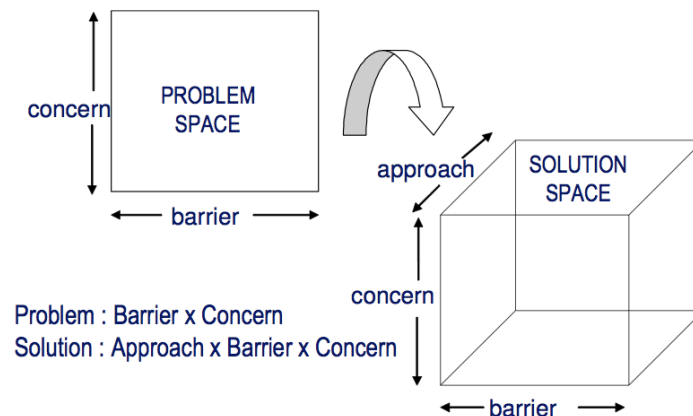


Figure 1: Solutions as the intersection of approaches, barriers, and concerns (Chen & Daclin, 2006).

(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>

In this paper, this concept of a solution space 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., those related to privacy, control of data, and trust. These designs are justified according to a design space analysis. MacLean, Yong, Bellotti, and Moran (1991) presented design space analysis as an approach to represent design rationale, focusing 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.

The LA Design Space model (Figure 2) is based on a three-step process, identifying concerns, barriers, and design solutions. The following walk through the three steps will explain the LADS model.

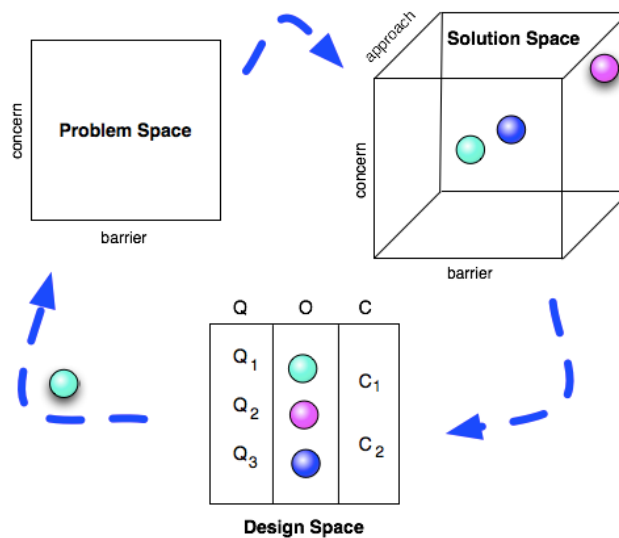


Figure 2: The Learning Analytics Design Space Model.

1. Constructing the problem space: For this paper, the *concerns* are related to data sharing and interoperability, which revolve around issues of privacy, control of one’s own data, and trust in applications and service providers (Hoel & Chen, 2014). The *barriers* related to data sharing and interoperability are part of the challenge of scaling up LA. As Ferguson et al. (2014b) observe, few reports currently exist in the LA literature regarding deployment of scale. Moving from research and pilot environments to large-scale applications could prove difficult due to lack of data for learning analytics (Cooper & Hoel, 2015; Griffiths, Hoel, & Cooper, 2016). For the purpose of this paper we have explored how LA data could be collected (Section 4) to identify barriers and propose solutions.

2. Constructing the solution space: Solutions should be developed along many dimensions, (e.g., technical, organizational, legal, or political), trying out both “soft” and “hard” approaches (see Figure 2 where the solutions are represented by coloured dots.)

(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>

3. Constructing the design space and selecting a first solution: In the last step, the questions derived from the Problem Space analysis are used to analyze the candidate solutions (in Figure 2, see Option column “O” in Design Space), and criteria (C1 and C2) derived through moving from problem to solution to design. These will be used to select one or more solutions (green dots) for further analysis in a continuous development cycle. For the sake of argument, one solution might be that a “technical fix,” e.g., a data-sharing consent dashboard needs to be developed, and that codes of practice and organizational policies were not enough to provide solutions to the identified problems.

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

4 CASES OF DATA SHARING: ISSUES TO BE ANALYZED USING THE LADS MODEL

In order to conduct a first run through of the model, we will identify concerns and barriers selected from a few cases we have built for this paper exploring which data could be available for learning analytics. After examining different aspects of data sharing in this section, in Section 5 we will use the results as input to see if the LADS model is a viable instrument for analysis.

LA begins and ends with data. Data are generated from learner actions and the contexts of learning; then the analytics produces new data, which is used by follow-up actions and interaction with the learner, which in turn produce new data to feed into the next LA cycle. The data are stored in standardized formats of sorts, and are 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 (CEDS, 2014) concludes that much of the data residing in Student Management Systems or Learning Activity Record Stores are not imbued with privacy issues raised by the introduction of new LA practices. Of course, there are sensitive issues related to the identification of a person; and the 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 decades without causing too much concern. It is the learning process data, sitting in the intersection between organizations, people, and learning resources that now have become so much more important.

Process data are, as observed in new LA applications, captured in formats defined in activity stream specifications, e.g., ADL Experience API,² Tin-Can,³ IMS Caliper⁴ (Griffiths, Brasher, Clow, Ferguson, & Yuan,

² www.adlnet.gov/tla/experience-api

³ tincanapi.com

⁴ www.imsglobal.org/caliper

(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>

2016). These specifications establish a core language to describe activities by providing information on subject, verb, object, context, etc. On top of these core specifications, community profiles provide specialized 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 activity, which opens up the question of what LA practitioners want to describe.

Ferguson and Buckingham Shum (2012) introduced five categories of analytics that make use of five partly overlapping classes of data:

- **Social network** (analyzes relationships using data about identifiable persons and their activities, e.g., publishing papers, participating in social platforms, etc.)
- **Discourse** (analyzes language as a tool for knowledge negotiation and construction using full-text data from discussion fora, talk, and other written text sources)
- **Content** (analyzes user-generated content using data from Web 2.0 applications)
- **Disposition** (analyzes intrinsic motivations to learn using a range of activity data, in principle generated by all the tools used by the learner)
- **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 Buckingham Shum (2012) would not be possible without data from social software, also called Web 2.0 applications. With mobile devices now in nearly every student’s pocket, use of social media is part of everyday life, including on campus or in the classroom. Even when institutional policies try to restrict their use in formal education settings, social media still pervades the educational space.

Garaizar and Guenaga (2014) explored how HTML5 browser APIs could shed some light on how the use of Web apps in mobile environments has the potential to enhance learning. The APIs allow web pages to make use of data collected by different sensors, e.g., sensors embedded in wearable computers (mobile phones, wristbands, watches, etc.). This opens up a range of new data sources. Table 1 lists the data types used by HTML5 APIs and derives questions as to what pedagogical or learning analytics uses these data types could potentially have.

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

Data type	Information provided	Potential questions
Geolocation	Latitude / longitude changes	Is the learner at school or at home? Is she commuting? Where does the learning take place?
3-D Orientation	Acceleration changes	Is the context suitable for learning?

(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>

Data type	Information provided	Potential questions
Battery	Status of battery, charging	Does the battery status affect the learning context? How?
Network information	Cost of network access	Does the cost of network access disrupt the learning scenario? How?
Offline and online events	Connectivity status	Which problems are 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	Light surrounding 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 and 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 and receive multimedia between browsers	How can the multimedia streams be collected, stored, analyzed, 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?

(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>

Data type	Information provided	Potential questions
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 and drop events, focus	Fine-grained user interactions	Which web controls are easy or hard to use? Which gestures and/or complex interactions are preferred by learners?
Page visibility, full screen, pointer lock	Single task / multitask scenarios	Is the learner multitasking? How? When? Do single 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 moved our focus of analysis away from the data elements and their potential privacy issues to data in context. Privacy is not a unidimensional concept describing the relationship between the data element and the person about whom this element holds information. By bringing in the context dimension, we see that data belong to more than the person described; it is the characteristics of the setting (context) that impact the privacy concerns.

Exploring these cases of data available for learning analytics, we 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. From an institutional perspective, if this boundary is crossed — e.g., by introducing social software services run by a third party — this can only happen by individual consent on a case-by-case basis. 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, Hodgkinson, and Colley (2003) have demonstrated. They found (*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” (Malcolm et al., 2003, p. 313).

The input for constructing the Problem Space is concerns and barriers. The first workshop on LA at ICCE 2014 expanded on the privacy, control, and trust cluster of issues referred to above (Hoel & Chen, 2014), and mapped concerns (Mason, Hoel, & Chen, in press). Some concerns 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 (Duval, 2011). These benefits are directly in the interest of the learner who wishes to be in control of her data. Since we have multiple stakeholders with legitimate interests, the eventual solutions must balance the interests of all parties.

(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>

Concerning barriers, the Educause Center for Applied Research identified four major challenges to achieving success with analytics in higher education: affordability, data, culture, and expertise (Bichsel, 2012). From an institutional perspective, cost is the main obstacle; however, factors like misuse of data, regulations requiring the use of data, inaccurate data, and individual privacy rights are barriers that higher education leaders worry about since they are collecting more data than ever before (Bichsel, 2012).

Hoel, Mason, and Chen (2015) analyzed a corpus of more than 200 questions gathered by the Learning Analytics Community Exchange⁵ and found that the discussion on data sharing and big data for education is still in an early stage. Conceptual issues dominate and there is still a long way to go in moving towards solutions for technical development and implementation.

5 A FIRST DEMONSTRATION OF THE LEARNING ANALYTICS DESIGN SPACE MODEL

Based on the concerns and barriers derived from the selected cases in Section 4, we construct a Problem Space for LA data sharing. This Problem Space leads to an exploration of solutions, which in turn will be selected as candidates for design.

5.1 Building the Problem Space

From a learner's perspective, two concerns are pulling the "data sharing slider" in opposite directions: prioritizing 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 the appreciation of benefits, feeling of control, trust in applications, companies, institutions, and so on. In the current situation, individuals seem to be more willing to take risks and go for new and innovative solutions (Xu, Luo, Carroll, & Rosson, 2011). While formal learning is led by institutions wanting to have *ethical use of student data policies* in place, they tend to stay with institutional learning platforms that use only a limited set of data sources for LA. For the institutions, lack of privacy frameworks is a major barrier to data sharing and using sensitive data sources that otherwise are only available to commercial LA providers.

The barriers seem to be more socio-cultural or organizational than technical or legal, to use the European interoperability framework dimensions (IDABC, 2004); however, the solutions will need to address all these interoperability challenges.

⁵ www.laceproject.eu

(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>

5.2 Building the Solution Space

Solutions are found by addressing the concerns and breaking down the barriers, which in our case we define as being of a technical, socio-cultural, and legal nature. Going for a “radical” alternative, using a variety of data sources and a high degree of data sharing, we can see these tentative solutions based on requirements from the cases discussed in Section 4:

- **Technical:** design a specification allowing users to express detailed conditions for data sharing when signing up for LA applications, with opt-out possibilities
- **Socio-cultural:** 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:** strengthen ownership and control of data from learning activities in national and international law

The next step is to choose one or more of these alternative solutions for design.

5.3 Design Space Analysis

Which solution should be focused on? 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 personalized and adaptive learning. Furthermore, we ask, is the solution safe from “losing face” through leakage of personal information? And does the solution support ubiquitous learning by allowing both formal and informal learning in the same application?

The criteria for which options to choose drive the design process based on the identified solutions. The privacy-by-design approach advocated by Nissenbaum (2014) gave priority to the social domain as the context to explore — to see if contextual integrity is maintained when data are shared. Therefore, 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, and others?

In the case of the *technical solution* proposed above, the design must go beyond a quick technical fix to solve the problem and give the user absolute control. The institution (school or university) should have a say, since it is also responsible for the greater good, the class or group, the parents, and society. Technical solutions should, therefore, include an element of permanent negotiation, thus requiring simple, transparent solutions (Hoel & Chen, 2015). The *legal solution* is also an option but not the first priority. Of course, solutions must have legal backing, but the privacy concerns surrounding data sharing are not solved by legal measures alone. Our analysis points instead 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

(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>

provider. Trust is not a “thing” that, negotiated once, lasts forever; it must be renegotiated repeatedly. Especially in a dynamic environment crowded with actors with different interests, large-scale, complex, non-transparent solutions will therefore be challenged. It will be easier to maintain context integrity with smaller solutions. Smaller LA solutions may seem a contradiction in terms, as the ideas of big data and data sharing across systems often lead to plans for large-scale solutions, perhaps with a centralized Learning Record Store or data warehouse 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 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.

Table 2: Summary of the first iteration of the LADS model.

Questions	Solutions	Criteria	Design Candidate	Solution
Will student privacy self-management be maintained?	User data sharing consent tool	Promote context integrity		
Will privacy in different contexts be respected?	Data sharing dashboard with consent and opt-out mechanisms	Continuous negotiation between learner, institution, and third parties		
Will different user groups trust the solutions?	Learner/institution dialogue practices	Avoid obfuscation, promote transparency	Solution that prioritizes the socio-cultural aspects for negotiation of access to data for learning analytics	
Will the solutions support ubiquitous learning in both formal and informal settings?	Regulation of data ownership and control through law	Harvest low-hanging fruits		

Table 2 summarizes the first iteration of using the LADS model to form questions and design solutions. This table maps the process illustrated in Figure 2 with examples of problems, solutions, criteria, and a candidate design solution identified for the selected cases in Section 4.

6 DISCUSSION

Educational institutions have always used learner behaviour and performance data to determine, visualize, 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 (Williamson, 2015). LA can be used to compute the relationships between learners based on their interactions, 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 has argued, “learning spaces have to be planned on the strength that different kinds of learning will only emerge once these spaces are used by students” (Thomas, 2010, p. 508). When “much, if not most, learning does not occur in formally designated learning spaces,” it is time to “wrest the locus of control from the traditional conception of learning space planning as the exclusive province of architects and physical facility planners” (Thomas, 2010, pp. 503, 510). This need to re-assess where learning happens is reinforced by the introduction of LA as a support technology. LA is, however, an emerging discipline (Siemens, 2013), and most of the technological ideas are still on the drawing board. 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 focusing on design justifications to allow for the incremental development of solutions. We used privacy-by-design principles to steer the development of ideas toward solutions; however, other principles could be used to test alternative design solutions, like pedagogical principles focusing on learning efficacy, learner-centred approaches, ubiquitous learning, and so on.

7 CONCLUSIONS AND FUTURE RESEARCH

Privacy awareness is reported to be one of the major features of smart LA when researchers summarize their experiences “from the field” (Ebner, Taraghi, & Saranti, 2015). LA is a young field both in research and in application design. New ideas are being launched nearly every day, and there is a need for testing to see if they meet the requirements of different stakeholders. For example, Kennisnet, a Dutch governmental school agency, has chosen PbD principles as a starting point for their new design: “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, 2014). Will giving students and parents full ownership of their data using the UMA standard benefit educational goals? In order to answer this question, one must analyze how the standard is implemented and how the different concerns are

(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>

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 analyze different solutions according to different criteria. The first tentative validation of the model presented in this paper shows that it has the potential to make a requirement discourse on LA applications more fruitful. However, in order to verify this conclusion, further testing is necessary. The European Learning Analytics Community Exchange (LACE) project has identified privacy and ethics as major themes for community discourse to develop the field of LA. This project will be a suitable testing ground for the LADS model.

8 ACKNOWLEDGMENTS

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

REFERENCES

- ADL (Advanced Distributed Learning). (2015). xAPI specification. *Produced by the Experience API Working Group in support of the Office of the Depute Assistant Secretary of Defense (Readiness) Advanced Distributed Learning Initiative*. Retrieved from <https://github.com/adlnet/xAPI-Spec/blob/master/xAPI.md>
- Aguilar, S. (2014). Perceptions and use of an early warning system during a higher education transition program. *Proceedings of the 4th International Conference on Learning Analytics and Knowledge (LAK '14)*, 390–391. <http://dx.doi.org/10.1145/2567574.2567625>
- Arnold, K. E., Lynch, G., Huston, D., Wong, L., Jorn, L., & Olsen, C. W. (2014). Building institutional capacities and competencies for systemic learning analytics initiatives. *Proceedings of the 4th International Conference on Learning Analytics and Knowledge (LAK '14)*, 257–260. <http://dx.doi.org/10.1145/2567574.2567593>
- Bichsel, J. (2012). Analytics in higher education: Benefits, barriers, progress, and recommendations. Louisville, CO: EDUCAUSE Center for Applied Research. <https://net.educause.edu/ir/library/pdf/ERS1207/ers1207.pdf>
- Bomas, E. (2014, 29 August). How to give students control of their data. [Web log post] Retrieved from <http://www.laceproject.eu/blog/give-students-control-data/>
- Cavoukian, A. (2012). *Privacy by design: From rhetoric to reality*. Toronto, ON: Information and Privacy Commissioner of Ontario. Retrieved from <https://www.ipc.on.ca/images/Resources/PbDBook-From-Rhetoric-to-Reality.pdf>
- CEDS. (2014). Common education data standards (CEDS): Version 4 data model guide. Retrieved from <https://ceds.ed.gov/>
- Chen, D., & Daclin, N. (2006). Framework for enterprise interoperability. In *Proceedings of the Workshops and the Doctoral Symposium of the Second IFAC/IFIP I-ESA International Conference: EI2N, WSI, IS-TSPQ*. Retrieved from <http://chen33.free.fr/M2/Elearning/CIGI2009.Chen.final.pdf>
- Cooper, A., & Hoel, T. (2015). Data sharing requirements and roadmap. Public Deliverable D7.2 from the LACE project (European Commission Seventh Framework Programme, grant number 619424).

(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>

- Retrieved from <http://www.laceproject.eu/deliverables/d7-2-data-sharing-roadmap/>
- Dartiguyeyrou, C. (Ed.). (2014). *The futures of privacy*. France: Fondation Télécom, Institut Mines-Télécom. Retrieved from <https://www.fondation-telecom.org/wp-content/uploads/2016/01/2014-The-Futures-of-Privacy-.pdf>
- Drachslar, H., & Greller, W. (2016) Privacy and learning analytics — it's a DELICATE issue: A checklist to establish trusted learning analytics. *Proceedings of the 6th International Conference on Learning Analytics and Knowledge (LAK '16)*, 89-98. <http://dx.doi.org/10.1145/2883851.2883893>
- Drachslar, H., Hoel, T., Greller, W., Kickmeier-Rust, M., Steiner, C., & Griffiths, D. (2016) Ethics and privacy for learning analytics: A review of current issues and their solutions. Retrieved from LACE project website <http://www.laceproject.eu/>
- Drachslar, H., Hoel, T., Scheffel, M., Kismihók, G., Berg, A., Ferguson, R., Chen, W., Cooper, A., & Manderveld, J. (2015a). Ethical and privacy issues in the application of learning analytics. *Proceedings of the 5th International Conference on Learning Analytics and Knowledge (LAK '15)*, 390–391. <http://dx.doi.org/10.1145/2723576.2723642>
- Drachslar, H., Dietze, S., Herder, E., d'Aquin, M., Taibi, D., & Scheffel, M. (2015b) The 3rd LAK data competition. *Proceedings of the 5th International Conference on Learning Analytics and Knowledge (LAK '15)*, 396–397. <http://dx.doi.org/10.1145/2723576.2723641>
- Drachslar, H., Stoyanov, S., & Specht, M. (2014a). The impact of learning analytics on the Dutch education system. *Proceedings of the 4th International Conference on Learning Analytics and Knowledge (LAK '14)*, 158–162. <http://dx.doi.org/10.1145/2567574.2567617>
- Drachslar, H., Dietze, S., Herder, E., d'Aquin, M., & Taibi, D. (2014b). The learning analytics & knowledge (LAK) data challenge 2014. *Proceedings of the 4th International Conference on Learning Analytics and Knowledge (LAK '14)*, 289–290. <http://dx.doi.org/10.1145/2567574.2567630>
- Duval, E. (2011). Attention please!: Learning analytics for visualization and recommendation. *Proceedings of the 1st International Conference on Learning Analytics and Knowledge (LAK '11)*, 9–17. <http://dx.doi.org/10.1145/2090116.2090118>
- Ebner, M., Taraghi, B., & Saranti, A. (2015). Seven features of smart learning analytics: Lessons learned from four years of research with learning analytics. *eLearning Papers*, 40. Retrieved from <http://www.openeducationeuropa.eu/en/article/Seven-features-of-smart-learning-analytics---lessons-learned-from-four-years-of-research-with-learning-analytics->
- EC (European Commission). (2012). *Proposal for a Regulation of the European Parliament and of the Council 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*.
- Council Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ L 281/31
- Ferguson, R., & Buckingham Shum, S. (2012). Social learning analytics: Five approaches. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK '12)*, 23–33. <http://dx.doi.org/10.1145/2330601.2330616>
- Ferguson, R., Cooper, A., Drachslar, H., Kismihók, G., Boyer, A., Tammets, K., & Martínez Monés, A. (2015). Learning analytics: European perspectives. *Proceedings of the 5th International Conference on Learning Analytics and Knowledge (LAK '15)*, 69–72. <http://dx.doi.org/10.1145/2723576.2723637>
- Ferguson, R., De Liddo, D., Whitelock, D., de Laat, M., & Buckingham Shum, S. (2014a). DCLA14: Second international workshop on discourse-centric learning analytics. *Proceedings of the 4th International Conference on Learning Analytics and Knowledge (LAK '14)*, 283–284. <http://dx.doi.org/10.1145/2567574.2567631>

(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>

- Ferguson, R., Clow, D., Macfadyen, L., Essa, A., Dawson, S., & Alexander, S. (2014b). Setting learning analytics in context: Overcoming the barriers to large-scale adoption. *Proceedings of the 4th International Conference on Learning Analytics and Knowledge (LAK '14)*, 251–253. <http://dx.doi.org/10.1145/2567574.2567592>
- Federal Trade Commission (FTC). (2012). *Protecting consumer privacy in an era of rapid change: Recommendations for businesses and policymakers*. Retrieved from the Federal Trade Commission website <http://www.ftc.gov/reports/protecting-consumer-privacy-era-rapid-change-recommendations-businesses-policymakers>
- Garaizar, P., & Guenaga, M. (2014). A multimodal learning analytics view of HTML5 APIs: Technical benefits and privacy risks. *Proceedings of the 2nd International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM '14)* (pp. 275–281). New York: ACM. <http://dx.doi.org/10.1145/2669711.2669911>
- Griffiths, D., Brasher, A., Clow, D., Ferguson, R., & Yuan, L. (2016). “Visions of the Future,” Horizon Report. Public Deliverable D3.2. LACE Project. Retrieved from http://www.laceproject.eu/wp-content/uploads/2016/02/LACE_D3_2.pdf
- Griffiths, D., Hoel, T., & Cooper, A. (2016). *Learning analytics interoperability: Requirements, specifications and adoption*. Public Deliverable D7.4. LACE Project (European Commission Seventh Framework Programme, grant number 619424). Retrieved from <http://www.laceproject.eu/d7-4-learning-analytics-interoperability-requirements-specifications-and-adoption/>
- Heath, J. (2014). Contemporary privacy theory contributions to learning analytics. *Journal of Learning Analytics*, 1(1). Retrieved from <https://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/3339>
- Hevner, A. (2007). A three cycle view of design science research. *Scandinavian Journal of Information Systems*, 19(2), 87–92 (article 4).
- 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. (2014). Learning analytics interoperability: Looking for low-hanging fruits. In C.-C. Liu et al. (Eds.), *Proceedings of the 22nd International Conference on Computers in Education (ICCE 2014)* (pp. 253–264). Japan: Asia-Pacific Society for Computers in Education.
- Hoel, T., & Chen, W. (2015). Privacy in learning analytics: Implications for system architecture. In T. Watanabe & K. Seta (Eds.), *Theory and Practice for Knowledge Management: Proceedings of the 11th International Conference on Knowledge Management (ICKM 2015)*. Retrieved from http://hoel.nu/publications/Hoel_Chen_ICKM15_final_preprint.pdf
- Hoel, T., Chen, W., & Cho, Y.-S. (2016). *Privacy requirements for learning analytics: From policies to technical solutions*. Paper presented at the Workshop on Ethics and Privacy for Learning Analytics at the 6th International Conference on Learning Analytics and Knowledge (LAK '16), 25 April, Edinburgh, UK.
- Hoel, T., Mason, J., & Chen, W. (2015) Data sharing for learning analytics: Questioning the risks and benefits. In H. Ogata et al. (Eds.), *Proceedings of the 23rd International Conference on Computers in Education (ICCE 2015)*. Asia-Pacific Society for Computers in Education. Retrieved from http://hoel.nu/publications/ICCE2015_Hoel_Mason_Chen_final.pdf
- 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>
- IMS Global. (2015). *Caliper Analytics™ Background*. Retrieved from the website of IMS Global Learning Consortium <http://www.imsglobal.org/activity/caliperram>

(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>

- Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., & Hall, C. (2016). *NMC Horizon Report: 2016 Higher Education Edition*. Austin, Texas: The New Media Consortium. Retrieved from <http://cdn.nmc.org/media/2016-nmc-horizon-report-he-EN.pdf>
- Johnson, L., Adams Becker, S., Estrada, V., & Freeman, A. (2014a). *NMC Horizon Report: 2014 Higher Education Edition*. Austin, Texas: The New Media Consortium. Retrieved from <http://www.nmc.org/pdf/2014-nmc-horizon-report-he-EN.pdf>
- Johnson, L., Adams Becker, S., Estrada, V., & Freeman, A. (2014b). *NMC Horizon Report: 2014 K–12 Edition*. Austin, Texas: The New Media Consortium. Retrieved from <http://cdn.nmc.org/media/2014-nmc-horizon-report-k12-EN.pdf>
- Kitto, K., Cross, S., Waters, Z., & Lupton, M. (2015). Learning analytics beyond the LMS: The connected learning analytics toolkit. *Proceedings of the 5th International Conference on Learning Analytics and Knowledge (LAK '15)*, 11–15. <http://dx.doi.org/10.1145/2723576.2723627>
- MacLean, A., Yong, R. M., Bellotti, V. M. E., & Moran, T. P. (1991). Questions, options, and criteria: Elements of design space analysis. *Human–Computer Interaction*, 6, 201–250. http://dx.doi.org/10.1207/s15327051hci0603&4_2
- Malcolm, J., Hodkinson, P., & Colley, H. (2003). The interrelationships between informal and formal learning. *Journal of Workplace Learning*, 15(7/8), 313–318. <http://dx.doi.org/10.1108/13665620310504783>
- Martinez-Maldonado, R., Pardo, A., Mirriahi, N., Yacef, K., Kay, J., & Clayphan, A. (2015). The LATUX workflow: Designing and deploying awareness tools in technology-enabled learning settings. *Proceedings of the 5th International Conference on Learning Analytics and Knowledge (LAK '15)*, 1–10. <http://dx.doi.org/10.1145/2723576.2723583>
- Mason, J., Hoel, T., & Chen, W. (in press). Questions as data: Illuminating the potential of learning analytics through questioning an emergent field. Accepted for publication in *Research and Practice in Technology Enhanced Learning*.
- Nissenbaum, H. (2014). Respect for context as a benchmark for privacy online: What it is and isn't. In C. Dartiguepeyrou (Ed.), *The futures of privacy* (pp. 19–30). France: Fondation Télécom, Institut Mines-Télécom. Retrieved from <https://www.fondation-telecom.org/wp-content/uploads/2016/01/2014-The-Futures-of-Privacy-.pdf>
- Piety, P. J., Hickey, D. T., & Bishop, M. J. (2014). Educational data sciences: Framing emergent practices for analytics of learning, organizations, and systems. *Proceedings of the 4th International Conference on Learning Analytics and Knowledge (LAK '14)*, 193–202. <http://dx.doi.org/10.1145/2567574.2567582>
- Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, 45, 438–450. <http://dx.doi.org/10.1111/bjet.12152>
- Prinsloo, P., & Slade, S. (2015). Student privacy self-management: Implications for learning analytics. *Proceedings of the 5th International Conference on Learning Analytics and Knowledge (LAK '15)*, 83–92. <http://dx.doi.org/10.1145/2723576.2723585>
- Scheffel, M., Drachsler, H., Stoyanov, S., & Specht, M. (2014). Quality indicators for learning analytics. *Educational Technology & Society*, 17(4), 117–132.
- Scheffel, M., Drachsler, H., & Specht, M. (2015). Developing an evaluation framework of quality indicators for learning analytics. *Proceedings of the 5th International Conference on Learning Analytics and Knowledge (LAK '15)*, 16–20. <http://dx.doi.org/10.1145/2723576.2723629>
- Sclater, N. (2016). Developing a code of practice for learning analytics. *Journal of Learning Analytics*, 3(1), 16–42. <http://dx.doi.org/10.18608/jla.2016.31.3>
- Singer, N. (2014, 21 April). InBloom student data repository to close. *New York Times*. Retrieved from

(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>

- <http://bits.blogs.nytimes.com/2014/04/21/inbloom-student-data-repository-to-close>
- Sekiya, T., Matsuda, Y., & Yamaguchi, K. (2015). Curriculum analysis of CS departments based on CS2013 by simplified, supervised LDA. *Proceedings of the 5th International Conference on Learning Analytics and Knowledge (LAK '15)*, 330–339. <http://dx.doi.org/10.1145/2723576.2723594>
- Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, 57(10), 1380–1400. <http://dx.doi.org/10.1177/0002764213498851>
- Swenson, J. (2014). Establishing an ethical literacy for learning analytics. *Proceedings of the 4th International Conference on Learning Analytics and Knowledge (LAK '14)*, 246–250. <http://dx.doi.org/10.1145/2567574.2567613>
- Thomas, H. (2010). Learning spaces, learning environments and the dis‘placement’ of learning. *British Journal of Educational Technology*, 41(3), 502–511. <http://dx.doi.org/10.1111/j.1467-8535.2009.00974.x>
- Vogelsang, T., & Ruppertz, L. (2015). On the validity of peer grading and a cloud teaching assistant system. *Proceedings of the 5th International Conference on Learning Analytics and Knowledge (LAK '15)*, 41–50. <http://dx.doi.org/10.1145/2723576.2723633>
- Wang, Y., Heffernan, N. T., & Heffernan, C. (2015). Towards better affect detectors: Effect of missing skills, class features and common wrong answers. *Proceedings of the 5th International Conference on Learning Analytics and Knowledge (LAK '15)*. 31–35. <http://dx.doi.org/10.1145/2723576.2723618>
- Williamson, B. (2015). Digital education governance: Data visualization, predictive analytics, and “real-time” policy instruments. *Journal of Education Policy*, 31(2), 123–141. <http://dx.doi.org/10.1080/02680939.2015.1035758>
- Willis, J. E., Campbell, J., & Pistilli, M. (2013, 6 May). Ethics, big data, and analytics: A model for application. *Educause Review*. Retrieved from <http://www.educause.edu/ero/article/ethics-big-data-and-analytics-model-application>
- Xu, H., Luo, X. R., Carroll, J. M., & Rosson, M. B. (2011). The personalization privacy paradox: An exploratory study of decision making process for location-aware marketing. *Decision Support Systems*, 51(1), 42–52. <http://dx.doi.org/10.1016/j.dss.2010.11.017>