The Influence of Data Protection and Privacy Frameworks on the Design of Learning Analytics Systems

Tore Hoel
Oslo and Akershus University
PO Box 4 St. Olavs plass
NO-0130 Oslo
Norway
Tore.Hoel@hioa.no

Dai Griffiths
University of Bolton
Deane Road
Bolton BL3 5AB
UK
D.E.Griffiths@bolton.ac.uk

Weiqin Chen
University of Bergen,
PO Box 7807
NO-5020 Bergen
Norway
Weiqin.Chen@uib.no

ABSTRACT

Learning analytics open up a complex landscape of privacy and policy issues, which, in turn, influence how learning analytics systems and practices are designed. Research and development is governed by regulations for data storage and management, and by research ethics. Consequently, when moving solutions out the research labs implementers meet constraints defined in national laws and justified in privacy frameworks. This paper explores how the OECD, APEC and EU privacy frameworks seek to regulate data privacy, with significant implications for the discourse of learning, and ultimately, an impact on the design of tools, architectures and practices that now are on the drawing board. A detailed list of requirements for learning analytics systems is developed, based on the new legal requirements defined in the European General Data Protection Regulation, which from 2018 will be enforced as European law. The paper also gives an initial account of how the privacy discourse in Europe, Japan, South-Korea and China is developing and reflects upon the possible impact of the different privacy frameworks on the design of LA privacy solutions in these countries. This research contributes to knowledge of how concerns about privacy and data protection related to educational data can drive a discourse on new approaches to privacy engineering based on the principles of Privacy by Design. For the LAK community, this study represents the first attempt to conceptualise the issues of privacy and learning analytics in a cross-cultural context. The paper concludes with a plan to follow up this research on privacy policies and learning analytics systems development with a new international study.

CCS Concepts

• Security and privacy→Privacy protections • General and reference→Design • Security and privacy→Social aspects of security and privacy • Security and privacy→Privacy protections • Applied computing→E-learning

LAK '17, March 13 - 17, 2017, Vancouver, BC, Canada

ACM 978-1-4503-4870-6/17/03…$15.00
DOI: http://dx.doi.org/10.1145/3027385.3027414

Keywords

Learning analytics; Privacy frameworks; Data protection; Data protection by design; Data protection by default; Personal information; Learning analytics systems design; Privacy by design; Learning analytics process requirements.

1. INTRODUCTION

1.1 The Task Undertaken by this Paper

As learning analytics in schools, universities and the workplace starts to scale up, concerns about privacy and data protection will also inevitably increase. This was confirmed by the EU funded Learning Analytics Community Exchange (LACE) which asked in the title of a review paper if privacy was a show-stopper for the field [13]. In the paper, examples are provided of Learning Analytics (LA) projects that were stopped or red-flagged because of privacy concerns. At the time that these problems emerged (2013 - 2015), international standards groups were keeping privacy issues out of scope in their work on specifying how activity streams should be expressed, exchanged and stored [1, 23]. These events and processes are well known to the learning analytics community, but, in parallel with this work, there are emerging national and international policies on data protection which are developing and generating requirements for LA which have received less attention. One hopes that the field of LA research has experienced enough set-backs, and has reflected on them sufficiently, to realise that there is a need for a deeper and more systematic treatment of ethical and data protection issues. Assuming that this is the case, the next step is to ask what implications an awareness of privacy issues could have for the design of LA tools and architectures.

The panorama of legal and policy issues raised by new and emerging LA technologies and practices, is immensely varied. Moreover, LA takes place in a wide range of social contexts, within which varying configurations of interest groups seek to guide the development of LA. Consequently, we are faced with a landscape which is hard to understand, or even to survey. In order to make a start on this task, there is a need for a comparative analysis of the legal and policy environment in the different regions of the world, together with an assessment of the varying factors which come to bear on the regulatory framework in different countries. The present paper will contribute to this exploration by analysing emerging international privacy frameworks, which are currently being turned into national laws in various parts of the world. The commitments which these frameworks and laws imply, in turn, determine the requirements for the design of national and international LA systems and architectures. We then present some initial, illustrative, case

studies on what role data protection regulations play in the national discourse among LA experts in selected countries, to discuss how privacy issues would influence the development and application of LA in different national contexts.

1.2 The Complex Landscape of Learning Analytics and Privacy Policy

The range of technical environments for LA is increasingly complex, sometimes putting users in control of the management of their data, and sometimes keeping them completely out for the loop. This, and the range of data sources involved, enmeshes learning analytics with multiple personal and societal issues in ways that are not yet fully analysed. The situation is made still more convoluted by the fact that the rise of learning analytics "is not presented simply as a more effective way to carry out educational activities, but also as a means to transform the context in which the new methods are embedded" [13].

Early adoption of LA has predominantly occurred in the US, where "it is playing an increasing role in determining how many post-secondary education institutions (PSEIs) engage with their students at multiple points in the student journey, as well as in the design of teaching and learning content and delivery" [31]. What is actually being done in PSEIs around the globe to apply student data at an institutional level is hard to establish, as these institutions are to some extent self-regulated, and adoption often follows research initiatives (and research ethics policies) in different university departments. While we wait for international case studies to be published, we can observe that national Ministries of Education (MoEs) have established "big data in education" centres at selected universities, to do research on LA, but also to develop a knowledge base for policy development. Examples include SLATE (University of Bergen, Norway), the Centre for Big Data on Technology-Mediated Education (Beijing Normal University, China), and the Learning Analytics Center (Kyushu University, Japan). In some Western countries LA is diffused to schools from university research via publishers and vendors, who adapt new applications using student data for interactive learning resources, innovative apps for STEM education, etc. However, there are examples of MoEs and school agencies that have more ambitious plans for large-scale adoption, e.g., the Republic of Korea wants to encourage large scale application of LA within a couple of years (Cho, personal communication, 24 September 2016).

When the Norwegian Centre for ICT in Education published its first guide on learning analytics in 2015, it concluded that the implementation of LA would probably be illegal unless a number of principles of data protection were adhered to [6]. The principles were summarised by the Centre as "lawfulness, purpose limitation, data minimisation, data quality, storing and deletion, right to know what information is stored, and information safety". These principles are derived from The Personal Data Act of April 2000 [6], which in turn builds on the European Data Protection Directive (Directive 95/46/EC). The text of the Act describes these principles succinctly: The data controller shall ensure that personal data are processed only if a) the data subject has consented; "b) are used only for explicitly stated purposes that are objectively justified by the activities of the controller, c) are not used subsequently for purposes that are incompatible with the original purpose of the collection, without the consent of the data subject, d) are adequate, relevant and not excessive in relation to the purpose of the processing, and e) are accurate and up-to-date, and are not stored longer than is necessary for the purpose of the processing" [7].

In the guide published on its website, the Norwegian school agency asks "how will the school owner make sure that information only are used for learning and not for other purposes, for example to control pupils and teachers? (..) What is the boundary between information that are relevant for learning and information that are not relevant, but nevertheless are of interest for registration and analysis" [6]? The text alludes to the Centre's limited trust in school owners being "able to maintain the most important principle of data protection: The data subject should be in control of and agree to how their own data are used" [6].

This initial response by a Northern European school agency to a hot new topic may be at the most cautious end of the scale, but legal constraints are something most authorities would consider with great care. Therefore, the authors of this paper suggest that the privacy frameworks that underpin national legal systems should be explored to see what requirements can be extracted, and the implications of these requirements for the design of technical systems and practices serving the needs of the different stakeholders of LA.

Before starting on the task of examining privacy frameworks it is worth sounding a note of caution. The market for LA technologies is global, and the drivers for the diffusion of LA practices are international, such as standards, access to new data sources, new sensor technologies, new pedagogical trends, etc. Privacy frameworks are also international, and often developed by organisations promoting international trade (OECD, APEC). Nevertheless, agreed upon concepts like 'purpose specification', 'collection limitation', 'individual participation', etc. tend to get very different interpretations when they are applied in national and institutional policies, culture and professional practice.

We now turn to an analysis of some key privacy frameworks, and their relationship to LA.

2. PRIVACY REQUIREMENTS FOR LA SYSTEMS

Legal requirements are filtered through a set of national policies, educational culture, infrastructure and organisational factors before designers of LA systems can synthesise them as hard requirements for applications. Figure 1 illustrates how, in the authors' experience, different factors contribute to MoEs' policies. It makes a big difference if a privacy framework is turned into national law (as the case with the European GDPR), compared to the burying of privacy principles in agreements of trade and economic cooperation. However, as many of the principles build on the same ideas, policymakers who want to influence the direction taken in the design of privacy and data protection for LA could find arguments to support their proposals in these agreements.

European countries, Korea and Japan are members of OECD, while China, Japan, and Korea of the countries we have studied are members of APEC, the Asia-Pacific Economic Cooperation. OECD published its first on Guidelines on the Protection of Privacy and Trans-border Flows of Personal Data in 1980 (revised and published at OECD Privacy Framework in 2013 [29]). These influenced the EU Data Protection Directive of 1995 (95/46/EC), which now will be replaced by the EU GDPR, published in May 2016 [10], and designed to "make sure people's right to personal data protection (...) remains effective in the digital age" [11]. APEC published its Privacy Framework in 2005 [3]. APEC has

been working towards updating their framework in time to mark the 10th anniversary of its adoption; however, a current activities list at the APEC website informs that the organisation is still working on the APEC Privacy Framework 2015, and that there are ongoing activities to promote interoperability between the APEC-EU Privacy Rules Systems [2].

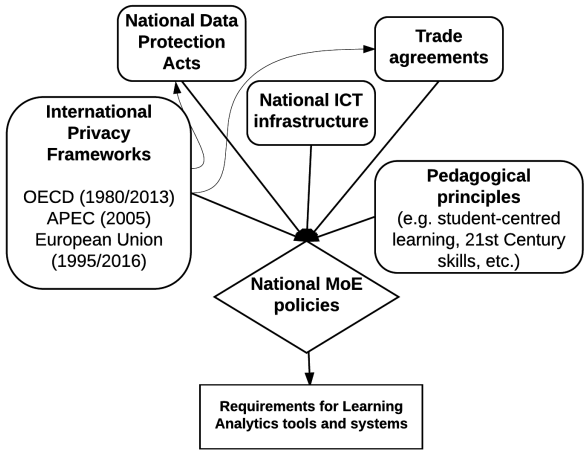

Figure 1. Influence factors impacting national requirements for LA tools and systems

The three frameworks are heavily influenced by each other, and they share many of the same concepts, as shown in Table 1. (In the table, we have listed the principles to show similarities, not to reflect how their order in the framework documents.) All frameworks try to strike a balance between protection of the individual and free flow of trade. One might say that the APEC framework leans more to the latter side, and the EU framework more to the former. Whether this is reflected in national data protection laws is beyond the scope of this study. However, in all countries we have been studying, a discourse on privacy and sharing of personal information for LA should find support in globally shared concepts, including the central concepts of purpose limitation, data minimisation and openness and transparency.

Table 1. Privacy principles as defined in the privacy frameworks of OECD, APEC and EU

OECD	APEC	EU GDPR
	Preventing Harm	Lawfulness, Fairness and Transparency
Collection Limitation	Collection Limitation	Data Minimisation
Purpose Specification	Choice	Purpose Limitation
Use Limitation	Uses of Personal Information	Storage Limitation
Data Quality	Integrity of Personal Information	Integrity and Confidentiality
Openness	Notice	
Individual Participation	Access & Correction	Accuracy
Accountability	Accountability	Accountability
Security Safeguards	Security Safeguards	
		Data Protection by Design and by Default

A further comparison of the three frameworks gives the impression that the GDPR is more updated in respect of meeting the challenges of a digital world (rules regarding breach notification, automated decision making and profiling, data portability, etc.). GDPR is also alone in promoting the principle of Data Protection by Design and by Default. It seems that the APEC framework gives less rights to the individual and gives a higher priority to the interests of the organisation.

2.1 Privacy Framework Requirements related to LA Processes and Pedagogical Requirements

In an exploration of the implications of the European data protection regulations for learning analytics design Hoel and Chen [16] used the LA process lifecycle model (Figure 2) of the international standardisation organisation ISO/IEC JTC1/SC36 as a template for discussing how GDPR requirement would influence systems development. The conclusion was that GDPR had specific requirements that would influence each process (possibly with the exception of Visualisation).

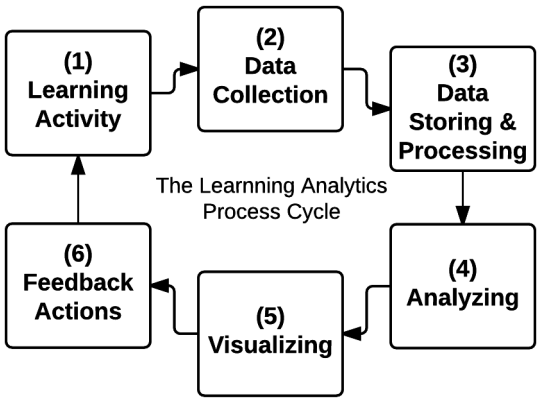

Figure 2. LA processes defined in ISO/IEC 20748-1 [21]

Table 2 gives a summary of the findings of [15], where provisions of the GDPR are mapped to each LA process (Column 2). Data protection by design and default is an all-encompassing requirement that influences all the LA subprocesses. In the present paper, we have added pedagogical requirements derived from mapping LA processes with GDPR requirements (Table 2, Column 3). In a discourse about privacy for LA we would claim it is important to frame legal requirements for system design in terms of pedagogical aims in order to reach out to the educational community. It will be much easier to move systems and tools design forward if educational stakeholders see that their pedagogical interests are met or affected.

Table 2. Summary of analysis of GDPR and pedagogical requirements related to LA processes

LA Processes	GDPR Requirements	Pedagogical Requirements
Learning activity	Give information of processing operation and purpose	Explicit formulation of the scope of LA processes. Choice of metrics that give answers to the pedagogical questions that initiated the LA process.
Data collection	Affirmative action of consent to data collection	Support of learner agency
Data storage and processing	Access to, and rectification or erasure of personal data. Exercise the right to be forgotten. Pseudonymisation and risk assessment	Support of learner agency
Analysis	Meaningful information about the logic involved. Information of profiling, e.g., predictive modeling	Support of learner agency and understanding of learning context
Visualisation	General requirements about transparency and communication	Selection of salient issues for pedagogical intervention
Feedback actions	Information about the significance and envisaged consequences of data processing	Pedagogical intervention, relating actions to pedagogical goals

The main pedagogical grounding of a LA process is centered on the selection of which Learning Activities to analyse and decisions on Feedback Actions. However, the other processes are not pedagogically neutral. If Data Collection, Data Storing and Processing, and Analysis are designed well, they could contribute to building learner agency and a better understanding how data are used in a modern society.

2.2 LA Process Requirements derived from the GDPR

Table 2 gives a high level view of how LA processes would be influenced by one privacy framework, the GDPR now being given legal status as of 2018 in European countries. In order to develop a more detailed list of design requirements for LA systems coming out for privacy frameworks we have used the GDPR as a starting point, in particular the overview of the GDPR developed by UK Information Commissioner's Office (ico.org.uk) [24]. This overview was published to help organisations understand their responsibilities and what rights are given to the individual. Design requirements for LA systems (see below) are developed by transforming legal requirements into systems requirement using the process life cycle model developed by ISO/IEC (Figure 2) as a scaffold. (For simplicity, we will use the term 'learner' for 'end-user of the LA system'. In the list of design requirements for LA systems, the numbers in parenthesis refer to the LA sub-process described in Figure 2.)

Right to be informed

The learner will throughout the full cycle of the LA process (1-6) be able to get information about a) what is the purpose of the LA session for specific activities of learning (1); b) what data are collected (2); c) how the data are stored and processed (3); d) what principles (e.g., predictive models, algorithms) are used for analysing learning data (4); e) what visualisations are used to render results (5); and f) what are the technical (as opposed to human) LA feedback actions that are designed for the particular LA process (6)

Right to access

The learner will throughout the full cycle of the LA process (1-6) be able to access, i.e., read and download, a) personal information (3), b) activity data (2) used for analysis (4), c) stored results of analysis (3)

Right to rectification

The learner will at any time be able to enter into communication with the data controller to launch claims for rectification of personal information (1-3).

Right to erasure

The learner has at any time the opportunity to raise the wish to be forgotten, which means deletion or removal of personal data when there is no compelling reason for continued processing (1-3). In an educational context, this could involve a number of actions, depending on educational level (mandatory or voluntary education) and contract agreements. These are some of possible scenarios: a) LA is not a necessary pedagogical means: All involvement with the LA process is terminated; b) LA is necessary on aggregated data: Only anonymised data are collected (and steps are taken to make sure that re-identification is not possible); c) a time restriction for storage of data is agreed, and all data are erased after completion of module, course, academic semester, degree, etc.; d) LA is an integral part of the course offering and the student is given the option to terminate the course and have his/her data deleted.

Right to restrict processing

This is somewhat similar to the LA attributes described above (Right to erasure), with the difference that the processing is put on hold and the data kept for use in historical analysis, aggregated analysis etc. (3).

The learner could also reserve herself from taking part in specific learning analytics processing.

Right to data portability

Learners have access to their learning activity data, so that when moving to another institution or another tool or LA system the learner can take their data with them for reuse in the new setting (3).

Right to object

There must be a service agreement that informs about the learners' rights to object to any aspect of the LA processes (1-6).

Right related to automated decision making and profiling

Individuals have the right not to be subject to a decision, which is based solely on automated processing; and which produces a legal effect or a similarly significant effect on the individual (4-6). Learning analytics may entail automated decision making and profiling of sorts, and these might be part of the contract between the institution and the individual.

Learners must be able to a) obtain human intervention; b) express their point of view; and c) obtain an explanation of the decision and challenge it.

Accountability and governance

The institution must be able to demonstrate that they have systems in place (policies and procedures) that uphold the protection of personal information and minimise risk of breaches (1-6).

Breach notification

When systems are compromised in any way, learners should be notified (3).

Transfer of data

Only EU relevant for European countries - the GDPR has regulations about transfer of data outside the EU region (3).

Data Protection By Design And By Default

LA systems development should conform to the principle of Data Protection By Design And By Default (1-6).

This exercise of constructing a detailed list of design requirements for LA systems by mapping between provisions in the most recent of the privacy frameworks and individual operations of a LA process cycle raises a number of questions that could be asked to different stakeholders around the world to get a picture of how privacy is conceived in application of LA. This is work that lies in the future. We will use this mapping of how legal and LA system requirements points to a new design space for LA as a backdrop for exploring how the discourse of these issues are held in selected countries now planning educational interventions using LA.

3. PRIVACY DISOURSE IN SELECTED COUNTRIES

A 2015 survey of European citizens' attitudes to data protection [34] concluded that only 15% felt they had complete control over the information they provided online; one in three people (31%) thought they had no control over it at all. Nine out of ten Europeans expressed concern about mobile apps collecting their data without their consent, and seven out of ten worried about the potential use that companies may make of the information disclosed.

This massive concern about data protection among ordinary citizens in Europe is not reflected in the discourse of the international LA research community. The LAK conference is the principal forum of this community, and the place that one would hope to find a response to these concerns, based on research evidence. However, a search for mentions of the 'data protection' in the proceedings of LAK reveals that the term does not appear in the proceedings of 2014 or 2015 [13], and only once in 2016 [8]. A similar lack of proposals for how data protection issues could be handled in an educational context is observable in our brief studies of privacy discourse in some of the countries that are now considering policy on LA.

We have focused in this first phase of this research on European and Asian countries; Europe because of the new Data Protection Regulation, and Asia because of the presence of research initiatives and establishment of organisations that take a national responsibility to promote LA in countries like Japan, Korea and China.

3.1 European Union / European Economic Area

The EU/EEA includes more than 30 countries that vary a great deal in terms of LA readiness and how privacy issues have been discussed in education. The reasons we discuss the region as a whole are twofold. Firstly, three years of community building by the LACE project has provided a good overview of the discourse on privacy. Secondly, in 2018 a new European data protection reform will establish a uniform data protection law for all EU/EEA countries, and this will influence the application of LA in the region.

A LACE review paper on ethics and privacy [13] concluded that learning analytics practice has shifted significantly from the principles of informed consent of the participants, which have to date been the bedrock of research ethics. Recent practices in leveraging data are challenging the concepts of data minimization (focused collection) and consent requirements. The review paper takes as an example the requirements for educational institutions set out in a UK Code of Practice developed by Jisc, a university service provider [30]. Under the Code, institutions do not have to ask students for permission to gather, hold and analyse their data. 'Consent' is reframed to refer to permission to take action on the results of data analysis, a quite different matter from obtaining permission to gather data. In taking this position Jisc does no more than recognise current practice, and indeed the modest proposal that institutions should normally obtain consent to take action on the results of analytics is more than most educational institutions in Europe currently do to obtain consent from their students.

The Open University UK has been a trailblazer in developing institutional policies on ethical use of student data for LA [35]. One of the principles states that "The OU has a responsibility to all stakeholders to use and extract meaning from student data for the benefit of students where feasible". These guidelines from Jisc and OUUK propose that there is an ethical duty on the institution to gather the best data that it can about its learners, to ensure that the service that it provides is as good as it can be. The implication seems to be that if individuals were given the right to opt-out, then this could be seen as unethical, because opting-out reduces the efficacy of learning analytics which can improve the education of others. This interpretation constitutes a radical reframing of the rights and obligations of the individual and the collective that are defined in the analysed privacy frameworks, with potentially profound consequences.

The European LA community discourse on ethical use of student data has not yet been influenced by another major debate on privacy coming out of the revision of the European data protection regulations. It seems inevitable that this will change, as in May 2016 a four year European Union revision process of the 1995 Data Protection Directive (95/46/EC) was concluded with the publishing of the General Data Protection Regulation (GDPR). Consequently the EU/EEA countries have until May 2018 to transpose these regulations into their national law. According to a factsheet from the European Commission (EC), the GDPR "will ensure that you receive clear and understandable information when your personal data is processed. Whenever your consent is required, it will have to be given by means of a clear affirmative action before a company can process your personal data. The new rules will also strengthen individuals' right to be forgotten, which

means that if you no longer want your personal data to be processed, and there is no legitimate reason for a company to keep it, the data shall be deleted" [9].

The "right to be forgotten" has become a topic of widespread public debate online, but the importance of GDPR for the use of data in education remains to be analysed. Hoel and Chen [16] have argued that the regulations will influence development and implementation of LA systems, and potentially strengthen the pedagogical grounding of these systems. The core of the GDPR relates to minimisation of data and use limitation. This restricts data collection to specified purposes and prevents re-purposing. It puts a bar on random collection of users' digital footprints and sharing (selling) them for other – not clearly declared – purposes. This restriction to minimisation and specific use in turn will, one may hope, lead to more focus on the core selling point, i.e. pedagogic application of analytics.

3.2 Japan

Japan has chosen a bottom-up approach to application of educational data for LA, with a number of ministries and agencies launching projects that encourage industry to develop solutions. There is still no public debate on privacy issues related to educational big data, according to the president of the Japanese Society for Learning Analytics (jasla.jp) Professor Yasuhisa Tamura (personal communication, September 2016). However, the various actors involved are monitoring international development in order to understand the importance of current trends, and there is a desire to 'import' guidelines and use cases for learning analytics. MIC, the Ministry of Internal Affairs and Communication, is leading a Smart School pilot project for development of ICT in K-12 school. For 2017, 250 million yen is allocated to develop a LA support system that records learning histories and the results of learning lessons, provides visualizations, improves the quality of teaching and student guidance and class and school management.

Professor Hiroaki Ogata, Director of Learning Analytics Center at Kyushu University, reports that it has been easier to introduce LA in higher education than in K-12 in Japan. His university is currently "practicing LA at university level", and it is the government's position that results from some universities will apply to K-12 in the future (Ogata, personal communication, October 2016).

One factor that will have an influence on how the landscape for LA in Japan develops is that the different ministries take different positions on how LA data could be used for commercial development. The MEXT (Ministry of Education, Culture, Sports, Science and Technology) is seen as rather conservative, being reluctant to disclose educational data, even if it has been anonymized. MIC and METI (Ministry of Economy, Trade and Industry) are eager to give access to LA data for business use by third parties. Professor Tamura sees that the different positions make it difficult to reach national consensus on how to handle LA data in Japan (Tamura, personal communication, September 2016).

3.3 Republic of Korea

In contrast to Japan, Korea will apply a top-down approach, at least for K-12 education. KERIS (Korea Education and Research Information Service) is the agency that is leading work on ICT in education. The 2014 KERIS report on "Prospects for the Application of Learning Analytics" (available in Korean only) [26] is the only official report on LA to date (Oct 2016). This report does not discuss privacy and data protection issues as a concern for development. However, KERIS has been active in developing the new ISO/IEC framework standard on learning analytics interoperability, where privacy policies play an important part [25]. KERIS organised a LASI-ASIA event in September 2016; and in panel discussions it became clear that KERIS has very ambitious ideas of rolling out LA in schools as soon as possible, and that technical development projects have been established to support this initiative. However, it is not clear to an international observer how policies for privacy and data sharing will be handled. Vendors and tools developers have had meetings with the government to discuss access to educational data (Kya Ha Lee, personal communication, Sept 2015). As the CEO of a small software company explained it to the authors, the Ministry of Education is the most conservative, and the avoidance of errors is a cornerstone of Korean culture. Therefore, vendors try to be cautious and present boilerplate conditions for users to accept by ticking a box when signing up to services in order to be in line with Korean privacy legislation (the Korean Government runs a rich advisory service at www.privacy.go.kr).

3.4 China

Development of the Chinese educational system as a whole is traditionally top-down, driven by national campaigns. However, there is considerable room for experimentation and the testing of new trends, including LA, providing it does not distract too much from supporting national curricula. The establishment of LACE China as a community exchange vehicle at Beijing Normal University's Centre for Big Data on Technology-Mediated Education in September 2016 has given insights into Chinese research and discourse on LA. So far, it is these authors' observation that Chinese researchers are now prioritising to find a research focus for LA and to define what parts of the educational system might benefit from LA. Issues of privacy and data protection are recognised as important, but from a Chinese perspective, other issues may be more pressing to discuss in order to leverage the data currently available for analysis.

China does not have a data protection act or a data protection regime. With growing awareness of the dangers of unprecedented and often illegitimate online access to personal information the need to speed up the legislative process is stated in the public debate [36]. It is the examples of excessive collection of personal data from online shopping, unauthorised disclosure of personal information by governmental and commercial institutions, illegal trade in personal information, etc. [4] that drive the expressed needs for a Chinese personal information protection act. How data sharing for LA will be conceptualised in this context remains to be seen. Ongoing software development projects may give a hint, as the LA dashboard application currently under development by a substantial team at the Beijing Advanced Innovation Centre for Future Education at Beijing Normal University.

This project will provide an integrated app and desktop application for viewing data about students and teachers, which is to be made available to both teachers and educational managers at various levels. The authors asked the team what the teachers thought about the application, and we were told that they did not like it, because it made their work more open to management control, but that this was not seen as a problem. This is only one example, but it does suggest that the Chinese educational authorities may feel able to override the concerns of interest groups about the use of data, in order to promote the common good as they conceive it. Interestingly this position is not far removed from that taken by the Open University's policy, discussed above.

4. BENEFICIARIES OF LEARNING ANALYTICS

As indicated in the Korean case, the discourse on LA and privacy will be coloured by stakeholder position, and by the identification of those who are seen as the main beneficiaries of data-driven analysis. LACE has suggested a classification [27] that looks upon *institutional administrators* in relation to activities such as marketing and recruitment, or efficiency and effectiveness measures; *individual learners* to facilitate a greater understanding of their progress and study behaviours; *teachers and support staff* to inform interventions with individuals and groups; and *academic staff* who might wish to adapt existing teaching materials or develop new curricula.

It is not surprising that it is the institutional perspective that is dominant in the cases we have reported. Institutions collect information and ask themselves how they could make better use of the data to promote their goals. LA is still in its infancy, and learners, teachers and academic staff need to see tools and solutions if they are to buy into the promises of analytics. Privacy issues are used at a system or a national level to ward off change (Japan, Korea); however, when individual actions are taken on the results of data analysis (EU), questions about privacy and data protection arise as a natural consequence. A general debate on big data and data protection, partly spurred by introduction of new legislation (EU) would add weight and give direction to this discourse. Hoel and Chen [18] argue that in a European context the GDPR with the new principles of Data Protection by Design and Data Protection by Default will bring the focus of learning analytics back to the learner and will serve as a lever for bringing pedagogy into the discourse on LA.

Figure 2 indicates how the authors perceive the orientation towards beneficiaries of the privacy frameworks and countries that we have analysed, placing them on a continuum between "focus on the individual" and "focus on the organisation".

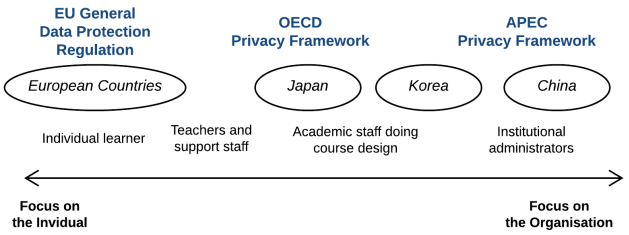

Figure 2. Orientation towards individual and organizational beneficiaries of privacy frameworks and countries

Studies of cultures and organisations [14] are out of scope for this paper; the classification of the countries we have studied is only a first reflection on the scant discourse we have been able to trace where concepts of privacy meet pedagogical and country-specific cultures in an effort to define requirements for LA systems and practices. Following extensive community exchange about LA worldwide, the authors of this paper conclude that LA largely remains on the drawing board, rather than on the commercial shelves ready to be applied. In order to move towards design we ask (also with the help of the model in Figure 2) how privacy frameworks and legal constraints on access to data and data sharing are influencing development of LA in countries such as those that we have studied.

5. DISCUSSION

This paper is premised on the hypothesis that legal constraints, defined in privacy frameworks and data protection acts, will have an impact on LA tools and practices. We have argued that even if the legal situation in this domain varies substantially, for example between European countries and China, it is fruitful to solicit specific design requirements (as reported in Section 2.2) through systematic analysis of technical and organisational implications of legal provisions in the most recent and advanced of the available frameworks (i.e., the GDPR). With these requirements in mind, we have turned to selected countries to see if the observed discourse on LA application and privacy could shed some more light on future development, keeping in mind that there is a global market for both LA tools and practices. It goes without saying that this is an initial exploration, which needs to be extended as different policies on the use of educational data are developed around the world.

5.1 Individual vs. organisational focus

It is interesting to consider whether requirements that are grounded in the wish to protect the individual will have an influence on LA system design in countries with a more collectivistic or organisational focus (ref. Figure 2). Of course, the answer to this question depends on the aims of LA in the different national contexts, and also on whether individual beneficiaries are considered in LA implementation work. In all the countries in this study, we have seen an interest in supporting the learning process per se, giving the individual learner a more adaptive learning environment. Even if the legal backing of a student for having full control of his or her data is completely different between a European country and China, it is our observation that the principles of fairness and transparency, accuracy, notice, and accountability – and maybe also purpose limitation and collection limitation – resonate well with the design criteria heralded by our Asian colleagues. Also LA system designers in cultures that give more value to organisational interests see that without the confidence and trust of end-users, new tools will be repurposed or circumvented if the user only sees them as part of a surveillance apparatus.

5.2 Schools vs. Higher Education

When we consider the different sectors of the educational system it appears that there are differences between K-12 and Higher Education (HE) in their approaches to privacy. Schools may be more susceptible to the influence of legal constraints than HE [15], because of their responsibility for the minors under their care, and because their work takes place under a social and political spotlight. Development in HE is more research driven, and strong role of research ethics rules in that environment may delay the discussion of the ethical and data privacy implications of full scale applications of LA outside the research context. Our case studies from Japan and Korea show that innovative solutions could still be stalled by the tug of war between parties that want data to be open vs. those who want data to be confined to the educational institutions that collect them.

5.3 Data Protection by Design and by Default

Both K-12 and HE institutions need incentives to build privacy requirements into their systems and practices. The principles of Data Protection by Design and by Default (DPbD&D) introduced in the European GDPR could prove to be a vehicle for change [18]. The principles are premised on Privacy by Design (PbD), a term first coined by the Canadian information and privacy commissioner of Ontario, Ann Cavoukian [5]. PbD can be seen as

"an engineering and strategic management approach that commits to selectively and sustainably minimize information systems' privacy risks through technical and governance controls" [32]. With GDPR, from May 2018, this engineering and management approach has the backing of the national European laws.

The way that DPbD&D could be used to change the design discourse can be illustrated with an example mentioned by the Korean vendor who contributed to our case study. She explained that she adhered to the national privacy laws by requiring users to tick a usage agreement form signing up to her company's service. In this she is in line with the way that legal requirements have been met by most learning technologies companies to date, seeking the lowest bar to pass the threshold. With the introduction of DPbD&D, developers of LA tools will not escape with just a simple checkbox form; by default they will have to dig deeper and open up each subprocess for a discussion related to data protection. In doing so, it is our assumption that the discourse on the need for strengthening privacy will focus more on the individual learner as beneficiary of LA (Figure 2), and the discourse will also have a pedagogical grounding (e.g., reasoning about learner agency, see Section 2.2), highlighting benefits of adaptive learning, etc.) [15].

5.4 The window of opportunity for design

This study has confirmed that we are at an important point in time, when choices are being made about where to take LA as a data-driven educational practice. Focusing on privacy issues could be seen as throwing a spanner into the works, by raising impractically complex issues. A stress on privacy issues may also be perceived as being in opposition to the educational or cultural values of a country. But a focus on privacy and data protection also creates the opportunity to achieve the necessary leverage in determining what questions LA should answer. The DPbD&D principles raise relevant questions, but it is the educational community that needs to provide the pedagogical scenarios that make the design of LA possible. The window of opportunity is tight: Will South Korea wait to launch a national LA solution for K-12 until individualised privacy solutions are found, or will the government build one data store for all? Will the Japanese industry come up with solutions that allow third party vendors to analyse LA data? Will European countries realise that the GDPR has given them a tool to move the LA discourse away from only covering technical issues such as the limits to anonymisation, encryption algorithms, and data security mechanisms, and towards supporting learner agency, teaching of learning of 21st century skills, and a more active learner teacher dialogue? In most countries, legal requirements are seen as an abstruse topic, but we propose that they could be used by the educational community as a lever to bring pedagogy into the discourse on LA.

6. RELATED WORK AND RESEARCH GAPS

Issues related to ethics and privacy are on the top of the list of concerns that need to be addressed according to LA researchers and practitioners [28, 22]. In a number of papers Hoel and Chen [20, 17, 21] have also explored what technical solutions a privacy-driven design of LA might lead to. In [19] Hoel, Cho and Chen researched how privacy and data protection requirements would affect all processes of the LA cycle.

Spiekermann and Cranor [33] distinguished two approaches for building privacy-friendly systems, "privacy-by-policy" and "privacy-by-architecture". The former approach focuses on the implementation of the notice and choice principles of fair communications, while the latter minimises the collection of identifiable personal data and emphasising anonymisation and client-side data storage and processing. It is argued that "notice and choice are needed to implement "privacy-by-policy" only where "privacy-by-architecture" cannot be implemented" [12]. In this paper we have just started to unpack the differences between these two approaches by analysing the differences between the APEC and EU privacy frameworks and seeing how the discourse on privacy issues are being conducted in European and Asian countries. The EU changed the direction of data protection by introducing privacy-by-architecture principles to the GDPR. Whether these principles have resonance in the APEC countries, where the tradition more is to define privacy by policy would be an interesting topic for further research, which is needed to understand how privacy approaches, national policies and architectures form tools and practice development in a specific domain as LA.

The momentum caused by the EU revision of the data protection framework and other recent developments (e.g., the Edward Snowdon case) has created an interest in defining privacy as "an integral part of the next wave in the technology revolution" and privacy engineering as new discipline [12]. How privacy engineering will be conceptualised and applied in different parts of the world is an interesting and new area of research, which this paper identifies as a research gap of particular interest to the LA community.

7. CONCLUSIONS AND FURTHER WORK

This paper makes a contribution to knowledge about how concerns about privacy and data protection related to educational data can drive a discourse on privacy engineering for LA. International privacy frameworks developed by OECD, APEC and the EU are identified as an impetus to solicit privacy requirements concerning all parts of a LA process cycle. The role that these requirements will in fact play in the design of LA tools and practices around the world depends on a host of factors discussed in this paper. As this paper is the first, to our knowledge, that explores how legal frameworks might influence LA design in different countries, its main contribution is to identify this topic as an important research area. Such research has the potential to give the LA community a better grasp of how privacy, pedagogy and technical development interact, and what the implications are for interoperability. An international scope is essential in carrying out this work, as there is a global market for learning technologies, and what is developed in one part of the world is rapidly taken up in another.

The exploration of national discourse in this paper is the result of talks with colleagues as part of a community building initiative in September-October 2016. A community of US, Australian, European, Korean, Japanese and Chinese researchers are now ready to start a comparative study on LA and policy, international aspirations, achievements and constraints. This new study will conceptualise different privacy approaches and explore how privacy-by-policy and privacy-by-architecture may impact the beneficiaries of learning analytics.

8. REFERENCES

[1] ADL (Advanced Distributed Learning). 2015. xAPI specification. Produced by the Experience API Working Group in support of the Office of the Deputy Assistant Secretary of Defense (Readiness) Advanced Distributed Learning Initiative. Retrieved from https://github.com/adlnet/xAPI- Spec/blob/master/xAPI.md. Accessed: 2016-06-01.

[2] APEC. Undated. Asia-Pacific Economic Cooperation, Electronic Commerce Steering Group, Current Activities. Retrieved from http://www.apec.org/groups/committee-on-trade-and-investment/electronic-commerce-steering-group.aspx. Accessed: 2016-06-01.

[3] APEC. 2005. Privacy Framework. ISBN 981-05-4471-5

[4] Baidu Encyclopedia. Undated. Personal Information Protection Act (In Chinese). Retrieved from http://baike.baidu.com/subview/2046064/2046064.htm. Accessed: 2016-06-01.

[5] Cavoukian, A. 2012. Privacy by Design: From Rhetoric to Reality. Retrieved from https://www.ipc.on.ca/images/Resources/PbDBook-From-Rhetoric-to-Reality.pdf. Accessed: 2015-02-13.

[6] Dahl, M. Undated. Læringsanalyse. Memorandum published at the Norwegian Centre for ICT in Education website http://iktsenteret.no/ressurser/notat-laeringsanalyse. Accessed: 2016-06-01.

[7] Datatilsynet. Undated. Personal Data Act. Act of 14 April 2000 No. 31 relating to the processing of personal data (Personal Data Act) Retrieved from https://www.datatilsynet.no/English/Regulations/Personal-Data-Act-/. Accessed: 2016-06-01.

[8] Drachsler, H. and Greller, W. 2016. Privacy and Learning Analytics – it's a DELICATE issue. Proceedings of LAK '16, April 20-24, 2016, Edinburgh, UK, 1 ACM 978-1-4503-3417-4/15/03.

[9] European Commission. 2016. How does the data protection reform strengthen citizens' rights?. Retrieved from http://ec.europa.eu/justice/data-protection/document/factsheets_2016/factsheet_dp_reform_citizens_rights_2016_en.pdf. Accessed: 2016-06-01.

[10] European Commission. 2016. REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

[11] European Commission. 2015. Questions and Answers - Data protection reform. Retrieved from http://europa.eu/rapid/press-release_MEMO-15-6385_en.pdf. Accessed: 2016-06-01.

[12] Finneran Dennedy, M., Fox, J., and Finneran, T. 2014. The Privacy Engineers Manifesto: Getting from Policy to Code to QA to Value (1st ed.). Apress, Berkely, CA, USA. ISBN:978-1-4302-6355-5.

[13] Griffiths, D., Drachsler, H., Kickmeier-Rust, M., Steiner, C., Hoel, T., and Greller, W. 2016. Is Privacy a Show-stopper for Learning Analytics? A Review of Current Issues and Solutions. Learning Analytics Review 6. Published by the LACE project. ISSN:2057-7494. Retrieved from http://www.laceproject.eu/learning-analytics-review/privacy-show-stopper. Accessed: 2016-06-01.

[14] Hofstede G., Hofstede G.J., and Minkov M. 2010. Cultures and organizations: Software of the mind, 3rd Edition. USA: McGraw-Hill Publishers.

[15] Hoel, T. and Chen, W. 2016. Implications of the European data protection regulations for learning analytics design. Proceedings of CollabTech 2016 and CRIWG 2016, Kanazawa, Japan, September 14-16, 2016.

[16] Hoel, T. and Chen, W. 2016. Data Sharing for Learning Analytics – designing conceptual artefacts and processes to foster interoperability. In Chen, W. et al. (Eds.) (2016). Proceedings of the 24th International Conference on Computers in Education. India: Asia-Pacific Society for Computers in Education.

[17] Hoel, T. and Chen, W. 2016. Privacy-driven design of learning analytics applications: Exploring the design space of solutions for data sharing and interoperability. Journal of Learning Analytics, 139–158. Retrieved from http://doi.org/10.18608/jla.2016.31.9.

[18] Hoel, T. and Chen, W. 2016. The Principle of Data Protection by Design and Default as a lever for bringing Pedagogy into the Discourse on Learning Analytics. Workshop paper in Chen, W. et al. (Eds.) (2016). Proceedings of the 24th International Conference on Computers in Education. India: Asia-Pacific Society for Computers in Education.

[19] Hoel, T., Chen, W., and Cho, Yong-Sang. 2016. Privacy Requirements for Learning Analytics – from Policies to Technical Solutions. Paper presented at Workshop on Ethics and Privacy for Learning Analytics, Monday, April 25th 2016 at the 6th International Conference on Learning Analytics and Knowledge (LAK '16), Edinburgh, United Kingdom

[20] Hoel, T. and Chen, W. 2015. Privacy in Learning Analytics – Implications for System Architecture — In Watanabe, T. and Seta, K. (Eds.) Proceedings of the 11th International Conference on Knowledge Management. ISBN 978-4-9908620-0-8 Presented at ICKM 15 in Osaka, Japan, 4 - 6 November 2015.

[21] Hoel, T. and Chen, W. 2014. Learning Analytics Interoperability– looking for Low-Hanging Fruits. In Liu, C.-C. et al. (Eds.) Proceedings of the 22nd International Conference on Computers in Education. Japan Asia-Pacific Society for Computers in Education.

[22] Hoel, T., Mason, J., and Chen, W. 2015. Data Sharing for Learning Analytics – Questioning the Risks and Benefits. In Ogata, H. et al. (Eds.) Proceedings of the 23rd International Conference on Computers in Education. China: Asia-Pacific Society for Computers in Education.

[23] IMS Global. 2015. Caliper Analytics Background. Retrieved from http://www.imsglobal.org/activity/caliperram. Accessed: 2016-06-01.

[24] Information Commissioner's Office. Undated. Overview of the General Data Protection Regulation (GDPR). Retrieved from https://ico.org.uk/for-organisations/data-protection-reform/overview-of-the-gdpr/. Accessed: 2016-06-01.

[25] ISO/IEC 20748-1. 2016. Information technology – learning, education, and training – Learning Analytics Interoperability – Part 1: Reference model.

[26] KERIS. 2014. Prospects for the Application of Learning Analytics (in Korean). Retrieved from http://www.keris.or.kr/board/pb_downloadNew.jsp?bbs_num=18451&ix=21990. Accessed: 2016-06-01.

[27] LACE. 2015. What are Learning Analytics? Retrieved from http://www.laceproject.eu/faqs/learning-analytics. Accessed 2016-06-01.

[28] Mason, J., Chen, W., and Hoel, T. 2016. Questions as data: illuminating the potential of learning analytics through questioning an emergent field. Research and Practice in Technology Enhanced Learning, 1–14. Retrieved from http://doi.org/10.1186/s41039-016-0037-1. Accessed: 2016-06-01.

[29] OECD. 2013. Privacy Framework. Retrieved from http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf. Accessed: 2016-06-01.

[30] Sclater, P. and Bailey, P. 2015. Code of practice for learning analytics. Jisc. Retrieved from https://www.jisc.ac.uk/guides/code-of-practice-for-learning-analytics. Accessed: 2016-06-01.

[31] Slade, S. 2016. Applications of Student Data in Higher Education: Issues and Ethical Considerations. Ithaka S+R. Retrieved from http://sr.ithaka.org/?p=283891. Accessed: 2016-06-01.

[32] Spiekermann, S. 2012. The challenges of privacy by design. Communications of the ACM, 55(7), 38–3. Retrieved from http://doi.org/10.1145/2209249.2209263. Accessed: 2016-06-01.

[33] Spiekermann, S., and Cranor, L. F. 2009. Engineering Privacy. Software Engineering, IEEE Transactions on, 35(1), 67–82. DOI:10.1109/TSE.2008.88.

[34] Special Eurobarometer 431. Undated. Data protection. Retrieved from http://ec.europa.eu/public_opinion/archives/eb_special_439_420_en.htm#431. Accessed: 2016-06-01.

[35] The Open University. 2014. Ethical use of Student Data for Learning Analytics Policy FAQs. Retrieved from http://www.open.ac.uk/students/charter/sites/www.open.ac.uk.students.charter/files/files/ec ms/web-content/ethical-student-data-faq.pdf. Accessed: 2016-02-01.

[36] Wei, L. 2015. Personal Information Protection Law under Network Environment. In Chinese. Peking University Legal Information Network. Retrieved from http://article.chinalawinfo.com/ArticleFullText.aspx?ArticleId=89816. Accessed 2016-06-01.