Chen, W. et al. (Eds.) (2016). Proceedings of the 24th International Conference on Computers in Education. India: Asia-Pacific Society for Computers in Education

The Principle of Data Protection by Design and Default as a lever for bringing Pedagogy into the Discourse on Learning Analytics

Tore Hoel & Weiqin Chen
Oslo and Akershus University College of Applied Sciences, Norway
tore.hoel@hioa.no

Abstract: Bringing pedagogy and learner agency to the forefront of the design of learning analytics systems is the central concern of this paper. With the new European data protection regulations now in place we consider their potential to influence systems design from this perspective. In particular, the principles of Data Protection by Design and Date Protection by Default are explored to see if conforming to these principles will bring the focus of learning analytics systems back to the learner. The emerging understanding leads to a model of the relationship between data minimisation and utility within contexts of data sharing, which allows constructing scenarios that highlight how pedagogy and data protection are related and could inform the design of LA solutions. Our analysis suggests that the new data protection regulations will influence development and implementation of learning analytics systems and that the pedagogical grounding of these systems will be strengthened.

Keywords – Learning analytics, Data protection, Data protection by design, Data protection by default, Learning analytics systems design, Privacy by design

1. Introduction

In translating learning into numbers there is a risk to throw the baby out with the bathwater: The learning analytics (LA) dashboards show the answers, but what were the pedagogical questions that initiated the data collection and measuring in there first place? And how do we reconnect with the learner whose agency is essential for even start thinking of objectifying the learning trajectories in a learning grap

In a pedagogically grounded LA design there is a need to identify requirements that put learner agency and pedagogical questions to the forefront of system development (Friend Wise and Williamson Schaffer, 2015). After the metrics for the data collection are decided upon it is easy to lose track of the individual when designing how data should be collected, stored and processed, what models should be used for analysis, and how results should be visualised. When feeding back the results to the learners the individual may be seen merely as a member of class described with traffic light colours (e.g., red = at risk) unless there are requirements requiring tools for a balanced conversation between learner and teacher, or learner and institution. Data protection regulations might be an example of such a requirement that could bring the focus in LA systems design back to the learner. In this paper we will explore the requirements coming out of the recent revision of the European Data Protection Regulations. We ask if these regulations could be used as a lever to bring pedagogy into the LA systems engineering discourse.

Privacy is posed as a potential show-stopper for LA (Hoel and Chen, 2016a). A review of current issues and solutions by the LACE project (Griffiths et al., 2016) gives examples of projects

that have been stopped or given a red flag because of concerns over privacy and data protection. Some developers have chosen to define privacy out of scope, like the groups that have specified how activity streams should be expressed, exchanged and stored (ADL, 2015; IMS Global, 2015). This may be a risky approach. If issues of ethics and privacy are not given sufficient consideration new solutions could easily backfire when questions are asked if they can be trusted.

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). Now the EU/EEA countries have until May 2018 to transpose these regulations into their national law. The introduction of the new GDPR coincides with an interesting and decisive moment in time for the emergent field of LA now soon ready to move out of the research labs into large-scale adoption. When scaling up LA architectures identity management, data storage, anonymisation and peudonymisation, etc. need to be developed and put in place. The GDPR will no doubt influence this development. In particular, the principles of "data protection by design" and "data protection by default" (DPbD&D) now written into the regulations may prove to be influential. However, these principles need to be further specified, and there is a need to understand how DPbD&D may change how we design LA tools and architectures.

A key aim of this paper is to explore the background of the DPbD&D principles and explore their possible impact in the context of LA. This context is obviously more narrow than the scope area of GPDR, and given the particular characteristics of learning, education and training (LET) the principles guiding this regulation might have unmapped effects on learning technology design.

2. Data Protection by Design and by Default in the EU Regulation

The GDPR defines DPbD&D as one of several general data protection principles, i.e., purpose limitation, data minimisation, limited storage periods, data quality, legal basis for processing, processing of special categories of personal data, measures to ensure data security, and the requirements in respect of onward transfers to bodies not bound by the binding corporate rules (EU, 2016, Article 47, d). However, the regulation does not give a clear definition of DPbD&D. Article 25 (EU, 2016) is dedicated to the principle (stated in the title). The first paragraph sets out the challenges a data processing implementer is facing (state of the art; cost, nature, scope, context and purposes of processing, etc.) and points to certain measures to take, e.g., pseudonymisation. The second paragraph reminds the data controller of his/her duties of implementing appropriate technical and organisational measures; and the third paragraph notes that the requirements set out in the previous paragraphs may be demonstrated using an approved certification mechanism.

In order to understand the underpinnings of the DPbD&D principle one needs to search for its roots in the discourse leading up to the GDPR. DPbD&D is premised on Privacy by Design (PbD), a term first coined by the Canadian information and privacy commissioner of Ontario, Ann Cavoukian (2012). In 2009, PbD was introduced by the Article 29 Working Party (a European Advisory Body set up by the EC) as an additional principle to innovate the data protection framework when the European Commission launched its consultation (Article 29 Working Party, 2009).

Privacy by Design is based on 7 "foundational principles" formulated by Cavoukian (2012): 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, and 7) Respect for user privacy – keep it user-centric.

When discussed by Article 29 Working Party, the idea of embedding privacy "already at the planning stage of information-technological procedures and systems" (Article 29 Working Party, 2009) is highlighted. Implementing PbD would require evaluation of several, concrete aspects or objectives, the advisory body explains, in particular, when making decisions about the design of a processing system, its acquisition and the running of such a system. The aspects or objectives mentioned are

> • *Data Minimization*: data processing systems are to be designed and selected in accordance with the aim of collecting, processing or using no personal data at all or as few personal data as possible.
> • *Controllability*: an IT system should provide the data subjects with effective means of control concerning their personal data. The possibilities regarding consent and objection should be supported by technological means.
> • *Transparency*: both developers and operators of IT systems have to ensure that the data subjects are sufficiently informed about the means of operation of the systems. Electronic access / information should be enabled.
> • *User Friendly Systems*: privacy related functions and facilities should be user friendly, i.e. they should provide sufficient help and simple interfaces to be used also by less experienced users.
> • *Data Confidentiality*: it is necessary to design and secure IT systems in a way that only authorised entities have access to personal data.
> • *Data Quality*: data controllers have to support data quality by technical means. Relevant data should be accessible if needed for lawful purposes.
> • *Use Limitation*: IT systems which can be used for different purposes or are run in a multi-user environment (i.e. virtually connected systems, such as data warehouses, cloud computing, digital identifiers) have to guarantee that data and processes serving different tasks or purposes can be segregated from each other in a secure way. (Article 29 Working Party, 2009)

In 2012, the same year the EU put forward its data protection reform "to make Europe fit for the digital age" (Reform of EU data protection rules, nd), Sarah Spiekermann published a viewpoint article stating,

> "heralded by regulators, Privacy by Design holds the promise to solve the digital world's privacy problems. But there are immense challenges, including management commitment and step-by-step methods to integrate privacy into systems" (Spiekermann, 2012).

Spiekermann welcomed that privacy impact assessments were to become mandatory in the new European data protection legislation.

> "However, they must be accompanied by a clear set of criteria for judging their quality as well as sanctions for noncompliance. (..) Most important, [they] need to be made mandatory for the designers of new technologies—the IBMs and SAPs of the world—and not just data controllers or processors who often get system designs off the shelf without a say" (Spiekermann, 2012).

The GDPR holds no criteria, not even a definition of DPbD&D, but nevertheless, states that "the principles of data protection by design and by default should also be taken into consideration in the context of public tenders" (EU, 2016, Recital 78). In order to act upon the principles, a definition is needed. Spiekermann defines PdD as "an engineering *and* strategic management approach that commits to selectively and sustainably minimize information systems' privacy risks through technical and governance controls" (Spiekermann, 2012).

When it comes to LET technologies and learning analytics it is important to address the engineering challenges, as we have described the field as open for innovation (Hoel and Chen, 2016b). This was acknowledged by Spiekermann in 2012 when she observed that privacy scholars still put too much focus on information practices only (such as Web site privacy policies). "Instead, they should further investigate how to build systems in client-centric ways that maximize user control and minimize network or service provider involvement" (Spiekermann, 2012).

The remainder of this paper is dedicated to exploring how the principles of PbD and DPbD&D could be harnessed to develop more privacy proof and more pedagogically grounded LA applications. Firstly, we search the literature for examples of privacy engineering; secondly, we use a simple LA process lifecycle model to scaffold a discussion what impact PbD engineering will have for the LA domain. Finally, we discuss to which extent this direction of design will impact on the pedagogical grounding of the development of LA systems.

.

3. Related Work

It seems that the momentum caused by the EU revision of the data protection framework and other recent developments (e.g., the Edvard Snowdon case) has created an interest in defining *privacy engineering* as a discipline (Finneran Dennedy, Fox, and Finneran, 2014; Oliver, 2014; Spiekermann and Cranor, 2009), building on the thesis "that privacy will be an integral part of the next wave in the technology revolution" (Finneran Dennedy, Fox, and Finneran, 2014). Spiekermann and Cranor (2009) 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. These authors argue that "notice and choice are needed to implement “privacy-by-policy” only where “privacy-by-architecture” cannot be implemented" (Spiekermann and Cranor, 2009). The question is how far the PbD principles goes in supporting the *privacy-by-architecture* approach?

Gürses, Troncoso and Diaz find PbD to be vague and leaving many open questions about their application when engineering systems (Gürses, Troncoso, and Diaz, 2011). They even see the definition given by Cavoukian as recursive ("privacy by design means apply privacy by design"), communicating "to the reader that something needs to be done about privacy from the beginning of systems development, but it is not clear what exactly this privacy matter is nor how it can be translated into design" (Gürses, Troncoso, and Diaz, 2011). Gürses et al. claim that PbD can be reduced to "a series of symbolic activities to assure consumers' confidence, as well as the free flow of information in the marketplace" (Gürses, Troncoso, and Diaz, 2011). Their solution for applying PbD to systems is to include *data minimisation* as the foundational principle because of "the risk inherent in the digital format" (Gürses, Troncoso, and Diaz, 2011).

Discussing *Privacy and Data Protection by Design – from policy to engineering* the EU Agency for Network and Information Security (ENISA) observed, "deploying privacy by design methods might limit the utility of the resulting system. Hence the designer needs to find a trade-off between privacy and utility w.r.t. a certain metric" (Danezis et al., 2014). The example ENISA uses is privacy-friendly statistical databases. Using the PbD principle of prior privacy gives anonymised data with little analytical utility. Even if this approach is popular among the computer science academic community, ENISA claims it is seldom used. What is used, is posterior privacy using utility preservation features and different methods of risk assessment. It is no surprise that ENISA then concludes "anonymisation is a key challenge for the next decade because this tension between privacy and utility will be at the core of the development of the big data business" (Danezis et al., 2014).

One recommendation of ENISA is that standardisation bodies need to include privacy considerations in the standardisation process. In 2014, OASIS (Organization for the Advancement of Structured Information Standards) published a standard on Privacy by Design Documentation for Software Engineers by a committee chaired by Ann Cavoukian (OASIS, 2014). The specification translates the seven PbD principles to conformance requirements for documentation, "to demonstrate

that privacy was considered at each stage of the software development life cycle" (OASIS, 2014). Even if the specification clarifies the PbD principles in a number of sub-principles that are mapped to documentation requirements, it is a way to go to derive succinct engineering principles. However, it is specified that "purposes must be specific and limited, and be amenable to engineering controls", and that the documentation "shall clearly record the purposes for collection and processing, including retention of personal data". Of the engineering controls, it is required that "strict limits should be placed on each phase of data processing lifecycle (..) including limiting collection" (OASIS, 2014).

4. Directions for design of LA systems

In an exploration of the implications of the European data protection regulations for learning analytics design Hoel and Chen (2016c) used the LA process lifecycle model (Figure 1) of ISO/IEC JTC 1/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).

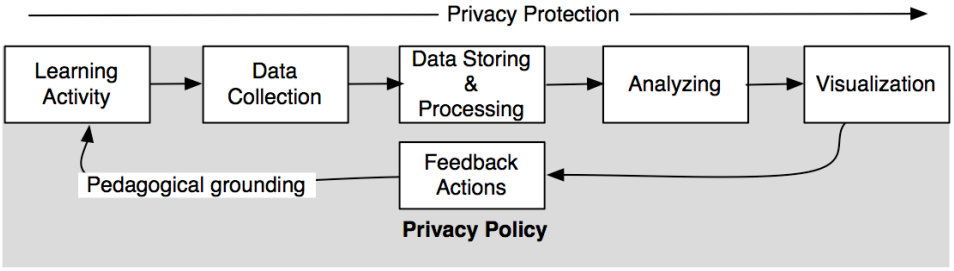

Figure 1. Learning analytics processes as defined in forthcoming ISO/IEC 20748-1 standard with highlighting of processes that need pedagogical grounding.

Table 1 gives a summary of the findings of Hoel and Chen (2016c), 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. If this requirement should be more than symbolic statements (Hoel and Chen, 2016c) DPbD&D needs to be translated into engineering principles and design actions guided by the GDPR requirements identified in Table 1. Data protection, however, is not an end in itself; it is a means to make a pedagogical tool safer to use. In designing the tool one therefore needs to know where pedagogy comes into play in a LA process cycle. In the 3rd column of Table 1 we have explored pedagogical requirements related to each LA subprocess.

Table 1. 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	What is the pedagogical scope of the LA process? Choose metrics that give answers to the pedagogical questions initiating the LA process.

Data Collection	Affirmative action of consent to data collection	Support of learner agency
Data Storing & Processing	Access to, and rectification or erasure of personal data. Exercise the right to be forgotten. Pseudonymisation and risk assessment	Support of learner agency
Analysing	Meaningful information about the logic involved. Information of profiling, e.g., predictive modeling	Support of learner agency and understanding of learning context
Visualising	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

As indicated in Figure 1 the pedagogical grounding of a LA process is centered around selecting which Learning Activities to analyse and deciding about Feedback Actions. However, the other processes are not pedagogical neutral. If Data Collection, Data Storing and Processing, and Analysis are designed well, they could contribute to build learner agency and a better understanding how data are used in a modern society.

In defining a set of *by default* principles systems development should be built upon we see that pedagogical grounding fits well DPbD&D. One needs to specify which questions to address in LA; and this aligns well with purpose specificity and data minimisation, two concerns identified in the review of related work (Section 3) as key principles of DPbD&D. To further unpack the relationships between these concepts we have developed a model describing the relationship between data minimisation and utility in different contexts of data sharing (Figure 2).

With a un-reduced dataset the utility of analysis would be high according the model described in Figure 2. To make a design decision on the trade-off between data minimisation and utility, it is, in particular in the design of LA systems, useful to bring in the concept of a *data sharing context*. This context may be broad or narrow, depending on the size and characteristics of the group having access to data. As an example, when running a LA application for a limited session in a school class the need for data minimisation is low (the personal information is already known to the group), and context of sharing is narrow (the information stays in the group), and one can expect high utility as one can use an unreduced dataset.

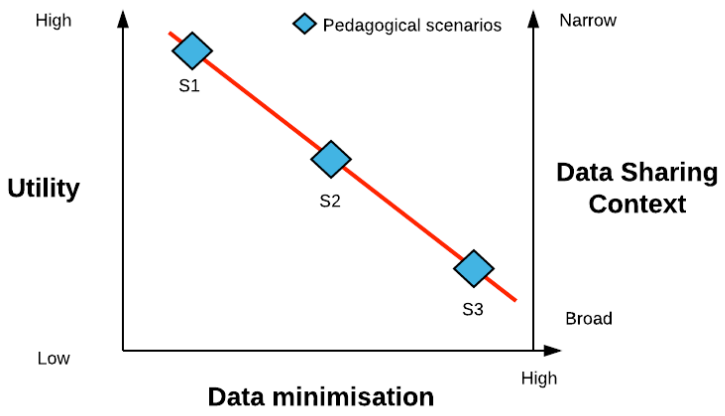

Figure 2. Relationship between data minimisation and utility within contexts of data sharing

What do the principles of data protection by design and default imply when widening the data sharing context? For example, when moving the analytical focus from the class, to the school, to the district, and further to the ministry of education, – how do we decide about the trade-offs between privacy and utility – with respect to what metrics? For LA design the answers to these questions do not lie in advanced anonymisation techniques – at least not as a starting point. Where we would suggest to start is to go back to the pedagogical grounding of the different analytical operations.

This approach is illustrated by some scenarios that have different characteristics according to the model described in Figure 2.

Scenario 1. School class with extensive data sharing: This scenario targets primary education, which is organised in school classes where the students are known to each other and the data comes from systems controlled by the school authority (e.g., learning management systems, learning content systems, subject related tools for math and physics education, etc.).

Aims of LA usage: Allowing a variety of tools to collect rich sets of data.

Pedagogical grounding: Learner agency, with the goal of empowering the student being able to manage own data and understand the benefits and risk of sharing personal information.

DPbD&D implementation: Instead of focussing on limitation of purpose and minimisation of data the focus is on controllability, transparency, and user friendliness.

Requirements for design solutions: Privacy features should be an integral part of school apps. These features should not only give the learner control over own data (where it is used, consent to share with other systems, etc.), but also support learning objectives set out in the curriculum on how to handle and understand data. The solutions should allow experimentation and learning exploration in data management in a data-driven and networked world.

The low level of data minimisation in this scenario should be compensated by architectural features that limit the longevity of the data and use sandboxing technologies keeping the learning and education space separate from life areas that are ruled by other ethical standards. For example, data from pre-university education should not end up in commerce, workplace, military, etc.

Scenario 2. Publisher / vendor / content provider with apps for STEM education: This scenario targets secondary education; however, the data processing is managed by third party vendors that sign up individual students. The vendors want to improve their tools and services by adding data from sources controlled by the students as well as by the school authorities.

Aims of LA usage: To support adaptive learning, giving students immediate and continuous feedback on achievements; giving the teachers a picture of how each students are doing at all time; and giving providers input to improve learning design and content.

Pedagogical grounding: Learner centered approach, giving the learner learning objectives, content and instructions that fit the competency level and unique needs of each learner.

DPbD&D implementation: In addition to the criteria of user controllability, transparency, and user friendliness, this scenario needs to consider issues of data minimisation, data confidentiality, and use limitation.

Requirements for design solutions: One direction of design could be to put all personally identifiable information (PII) under the control of the user (e.g., in a Personal Learning Record Store (PLRS) (Kitto et al., 2015), which the vendor system communicates with using different anonymisation techniques). By using standardised interfaces for competency gap queries, learning objective fulfilment, etc. the vendor tool could make queries to the PLRS without being exposed to any PII. In giving the user full control over own data the principles of purpose limitation and data minimisation are followed. It is up to the vendor tool to ask precise questions using a well defined formats in order to come up with the appropriate learning content, choose the right tests, and give the proper support.

An alternative direction of design would be to use a trusted API broker to handle PII serving certified vendors tokens that make it possible to use different datasets for learner-centered analytics. As there are always possibilities for re-identification using pseudonymised data also these solutions should provide tools for consent and management of learning activity data.

Scenario 3. School authority quality assurance system: This scenario targets school principals, politicians, parents and others that want to influence the contexts in which learning and teaching occur. The relevant data sources are many and varied; however, the need of PII is non-existent.

Aims of LA usage: To monitor the development of learning and education contexts in order to be able to allocate resources, adjust aims, facilitate planning for persons or groups, etc.

Pedagogical grounding: Quality assurance, based on metrics developed by local authorities, ministry of education, and others.

DPbD&D implementation: Removal of PII using state-of-the-art pseudonymisation techniques.

Requirements for design solutions: The design must balance the need for anonymity and analytical utility. A solution with centralised data warehouse where different interests could be allowed to probe the data without prior hypothesis is not in line with the DPbD&B principles. A solution with *a priori* data protection guarantee is also seen as less valuable as the anonymisation techniques will reduce the data quality too much. Instead, a service is proposed based on the following requirements: 1) defined analytical questions and specified groups of users, 2) dynamic brokerage of data sources of relevance, also from distributed PLRSs, and 3) dynamic anonymisation techniques based on advanced risk assessment analysis, which also considers the context of use of the analytics.

The three scenarios give requirements for different designs depending on the constraints described in Figure 2. We see that design decisions based on pedagogical and DPbD&D principles have potential to override the more technical or theoretical relationship in the model between a high degree of utility and low degree of data minimisation. This is an important finding supporting the assumption captured in the title of this paper.

5. Discussion

Legal requirements are potentially decisive in influencing which tools will be implemented in the educational market, especially for primary and secondary education (Hoel and Chen, 2016a). Whether DPbD&D is a legal term or not depends the point of view. It is included in the legal regulations of the European Union; however, the term is ill defined and in order to understand what it means one has to include engineering and management theories. Even if DPbD&D is hard to understand these principles could change the direction of LA tools development. When looking at how legal requirements have been met by learning technologies till now we see a practice of looking for the lowest bar and just doing enough, but no more, to pass that threshold. The use of LMS is a case in point; students tick a box in a usage agreement form the first time they open the tool and that is the first and last time they are exposed to questions about handling of their data. With DPbD&D developers of LA tools will not escape with a simple form with a checkbox; they have *by default* to dig deeper and open up each subprocess for discussion related to data protection.

Our exploration of directions for design of LA systems has shown that there is a need for pedagogical perspectives when DPbD&D is turned into technical requirements. In an educational setting data protection is more than technical issues about limits to anonymisation, encryption algorithms, and data security mechanisms; it is also a matter of supporting learner agency, teaching of learning of 21st century skills, and a more active learner teacher dialogue. The DPbD&D principles raise the questions; however, the educational community has to provide the pedagogical scenarios to make it possible to design LA solutions that transform learning and teaching, not merely pass the legal threshold of handling PII safely.

The three scenarios constructed in the previous section draw on background information we have from participating in Norwegian, European and international work in learning analytics. Many other scenarios could have been constructed. Nevertheless, it is interesting to reflect upon the ideas for solutions that are indicated following a pedagogically inspired DPbD&D approach.

First, it seems that we are looking at local, temporal and distributed solutions, especially in scenarios with extensive sharing of complex datasets. The idea of a national data warehouse where all learning activities are stored and processed is not supported.

Second, technologies that allow the data subject to manage his/her own data, e.g., cloud-based Personal Learning Record Stores, should be explored. The pedagogical requirements strengthen this direction of design even if personal e-portfolio systems (Stefani, Mason, and Pegler, 2007) and personal learning spaces (Dabbagh, and Kitsantas, 2012) have not been a great success neither from a market nor a technical perspective. The introduction of cloud based services in education might make this approach more promising.

Third, trusted 3rd party organisations may play an important role in providing API gateway services that allow both vendors, students and teachers to exchange information without compromising PII. This approach could also be question driven services, where activity data are not exchanged, only the answers to specific questions (e.g., how is the result of this quiz related to the student's performance in general). From a pedagogical perspective the challenge for these kinds of solutions is to prevent them from being a black box you just have to trust, not understanding how it works.

Fourth, governmental engagement may be necessary to develop LA infrastructures as described above. 3rd party organisations may be hard to establish without support from ministries of education. And as the responsibility to amend new data protection regulations rests with the

government, awareness raising about the benefits of DPbD&D in supporting good pedagogical solutions should also be undertaken by ministries of education.

6. Conclusions and Further Work

Design and implementation of privacy requirements in systems requires translation of complex legal, social, ethical, and also pedagogical concerns into systems requirements. The new GDPR of the European Union will serve as a guideline on how to address these concerns. In this paper we have sought to trace the background for the principles of Data Protection by Design and by Default and explored how this principles should be handled by stakeholders of learning analytics systems, a new and emergent field of learning technologies.

The principles of DPbD&D are only two of a number of data protection measures that are now prescribed by law in Europe. Combined they represent an opportunity to make the design of data protection features in LA systems more driven by pedagogical considerations. This paper has explored how GPDR requirements and pedagogical requirement could influence the understanding of the different parts of a learning analytics process cycle. To help this exploration a model was developed of how a central aspect of data protection, data minimisation, is related to utility within different data sharing contexts. Furthermore, a scenario template was developed and three scenarios of different LA systems were presented.

Our analysis suggests that data protection regulations could change the direction of LA systems development if combined with pedagogical requirements. Further work is needed to see how data protection provisions are handled by the LA developers community at large. It would also be interesting to see how ministries of education are planning to engage with the new field of LA and how they would use new legal requirements to influence the use of data-driven technologies in education. More work is needed to see how students, teachers and administrators could be involved in development of LA tools and systems to make the processes more transparent and understandable. It is implied in this paper that issues of data protection and privacy have potential to engage these stakeholders in LA design. From a technical architecture point of view, there is an immediate need to analyse whether current approaches involving learning record stores and data warehouses comply with the new data protection regulations; and, if so, identify what alternative architectures could be designed.

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

Article 29 Working Party (2009). The Future of Privacy. 02356/09/EN WP 168. Retrieved from http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2009/wp168_en.pdf

Cavoukian, A. (2012, June). Privacy by Design: From Rhetoric to Reality. Retrieved February 13, 2015, from https://www.ipc.on.ca/images/Resources/PbDBook-From-Rhetoric-to-Reality.pdf

Dabbagh, N., & Kitsantas, A. (2012). Personal Learning Environments, social media, and self-regulated learning: A natural formula for connecting formal and informal learning. Social Media in Higher Education, 15(1), 3–8.

Danezis, G., Domingo-Ferrer, J., Hansen, M., Hoepman, J.-H., Le Metayer, D., Tirtea, R., & Schiffner, S. (2014). Privacy and Data Protection by Design – from policy to engineering. European Union Agency for Network and Information Security (ENISA). DOI 10.2824/38623

EU (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)

Finneran Dennedy, M., Fox, J., & 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

Friend Wise, A., & Williamson Schaffer, D. (2015). Why theory matters more than ever in the age of big data. Journal of Learning Analytics, 2(2), 5–13. http://doi.org/10.18608/jla.2015.22.2

Griffiths, D., Drachsler,H., Kickmeier-Rust, M., Steiner, C., Hoel, T., 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 http://www.laceproject.eu/learning-analytics- review/privacy-show-stopper

Gürses, S., Troncoso, C., & Diaz, C. (2011). Engineering Privacy by Design. Computers, Privacy Data Protection https://www.cosic.esat.kuleuven.be/publications/article-1542.pdf

Hoel, T. & Chen, W. (2016a). Implications of the European data protection regulations for learning analytics design. Workshop paper accepted for presentation at The International Workshop on Learning Analytics and Educational Data Mining (LAEDM 2016) in conjunction with the International Conference on Collaboration Technologies (CollabTech 2016), Kanazawa, Japan - September 14-16, 2016

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

Hoel, T. & Chen, W. (2016c). Implications of the European data protection regulations for learning analytics design.CollabTech 2016 and CRIWG 2016), Kanazawa, Japan September 14-16, 2016

IMS Global. (2015). Caliper AnalyticsTM Background. Retrieved from the website of IMS Global Learning Consortium http://www.imsglobal.org/activity/caliperram

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

Oliver, I. (2014) Privacy Engineering. A dataflow and ontological approach, ISBN 9978-1497569713

OASIS (2014) Privacy by Design Documentation for Software Engineers Version 1.0. Committee Specification Draft 01. 25 June 2014. http://docs.oasis-open.org/pbd-se/pbd-se/v1.0/csd01/pbd-se-v1.0-csd01.pdf

Reform of EU data protection rules (n/d). Retrieved from http://ec.europa.eu/justice/data-protection/reform/index_en.htm

Spiekermann, S. (2012). The challenges of privacy by design. Communications of the ACM, 55(7), 38–3. http://doi.org/10.1145/2209249.2209263

Spiekermann, S., & Cranor, L. F. (2009). Engineering Privacy. Software Engineering, IEEE Transactions on, 35(1), 67–82. http://doi.org/10.1109/TSE.2008.88

Stefani, L., Mason, R., & Pegler, C. (2007). The educational potential of e-portfolios: Supporting personal development and reflective learning. Routledge.