LA4S White Paper:
Data Protection and Ethics in Learning Analytics

- March 2018 -

- Michael Kickmeier-Rust, Christina Steiner -

The 2018 General Data protection regulation of the EU imposes significant demands on the protection of individual data. This is delicate are in the context of educational organizations and particularly in the context of Learning Analytics and the regulation specifically alienates educators. On the other hand, the prominent appearance of the new regulation in mainstream media raises awareness among educators and providers of technologies for education. At the same time, the EU encourages the establishment of certificates and labels. This white paper attempts to stake the claims for a labelling system in Learning Analytics, specifically from the perspective of smaller educational organizations such as schools and small to medium size

1. The European GDPR 2018 and its Relevance for Education

The EU General Data Protection Regulation (GDPR) replaces the Data Protection Directive 95/46/EC and was designed to harmonize data privacy laws across Europe, to protect and empower all EU citizens data privacy and to reshape the way organizations across the region approach data privacy. The aim of the GDPR is to protect all EU citizens from privacy and data breaches in an increasingly data-driven world that is vastly different from the time in which the 1995 directive was established. Although the key principles of data privacy still hold true to the previous directive, many changes have been proposed to the regulatory policies (https://www.eugdpr.org/key-changes.html).

In the context of schools and small to medium scale universities, data protection might have played a subordinate role in the past. This will change with the new regulations. The following summarizes the key aspects of data protection. The italicized parts emphasize important relations to the educational sector.

The EU general Data Protection Regulation 2018

General principle

Personal data may only be used if there is an actively and freely given, specific and informed consent of the persons concerned. Silence or inactivity should not be interpreted as consent.

In addition, data processing may also be put in place when a "legitimate interest" is given and can be proved , but only on condition that it does not undermine the fundamental rights of the "data subject".

Similar requirements apply to the creation of user profiles, exceptions are made for statistical and research purposes and when using pseudonymous or anonymous data.

Schools and higher education establishments may claim that the collection of data of students is necessary and legitimate. This may, however, conflict with the extended protection of children (see below).

European data protection certificate

The General Data Protection Regulation also proposes the award of a "European Privacy Certificate" in order to increase the trust of users in appropriately certified services and legal

certainty for providers. There is also a clause stating that telecommunications and Internet companies are only allowed to submit data to third-country authorities such as the United States on the basis of European law or similar agreements.

Information rights

The GDPR gives affected persons far-reaching information rights. For example, an obligation to inform about the duration of data storage is provided.

Data protection officer

In principle, a data protection officer of an organization or company must be appointed if it processes data of more than 5.000 affected parties within 12 months. Furthermore, responsible bodies that operate with particularly sensitive data must also appoint a data protection officer.

Typical schools and small universities may not be obliged to install a data protection officer. Sensitive data usually refers to data that reveals racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, genetic data, biometric data for the purpose of uniquely identifying a natural person, and data concerning health or a natural person's sex life and/or sexual orientation. Nonetheless it is advisable to install a data protection officer.

Sanctions

The envisaged European General Data Protection Regulation provides severe maximum penalties for infringements in heights of 100 million euros or five percent of the annual turnover of a company, whichever is the higher. The resulting financial risk is thus quite comparable with that in antitrust violations.

Companies must therefore conduct regular compliance audits in key data processing areas to minimize the risk of costly privacy breaches. They must conduct a risk analysis and an impact assessment and have compliance checked every two years by an external expert.

The GDPR provides a collective action right. The data processing organizations and companies may then be exposed to legal disputes over data protection violations. However, the association may initiate a procedure for infringement of data protection laws not only in court, but also in the case of data protection supervisory authorities, representing at least one person concerned. Affected parties have the right to demand even immaterial damages. In

case of violations of data protection law, the companies thus make not only very high fines but also not inconsiderable private claims for damages.

The claim for compensation for the damage incurred by the data subject due to the unlawful processing of his data, which is directed against the data controller, is extended to order data processors. The person concerned can thus claim damages against several bodies who are jointly and severally liable. Both the responsible entity and the contract processor can avert liability if they can prove that they did not cause the event causing the damage. If several data processing bodies are involved, each body is liable to the party concerned in the amount of the full damage.

Affected parties can also demand compensation for immaterial damages.

Although schools and higher education establishments may not be sentenced to the maximum fines, law suits regarding compensation due to data protection failures.

General information duty in case of data breaches

In the event of data loss, data subjects and, at the same time, the relevant data protection supervisory authority must be informed. Data protection breakdowns are "without unjustified delay", that is within 72 hours, to inform the person concerned.

Requirements for privacy compliance

According to the GDPR data-preserving bodies must maintain data protection policies and suitable technical and organizational measures to prove compliance with the GDPR. There is also an obligation to revise them every 2 years.

The data controller (and also the data processor) must ensure that both the systems used to process personal data and the corresponding work processes are designed to be data-protection-friendly.

This typically applies to the educational sector.

Implementation of impact assessments

The data controller as well as the data processor are obliged to carry out risk analyzes and data protection impact assessments in advance of the planned data processing. In the event that special risks are identified, the supervising authority must be informed, which may even prohibit the planned data processing. The risk analyzes and impact assessments must be

constantly reviewed and examined in the form of a continuous data protection lifecycle management.

This typically applies to the educational sector.

Transparency requirements

The GDPR broadens the transparency requirements. The data controllers are required to provide privacy statements about the data processing they have performed. They must clearly describe how the persons concerned can exercise their rights, e.g., for being informed or for the deletion of data. This privacy policy must be generally understandable and easily accessible.

The GDPR stipulates that the data controller must also inform about how long the processed data are stored.

Establishment of a procedure for the execution of personal rights

Data controllers have to set up electronic procedures by means of which the persons concerned by the data processing have their rights of participation, such as data protection.

At the request of the person concerned, the responsible body must respond within 40 calendar days. In exceptional cases, the answer may be given one month later.

Rights to data deletion and correction

GDPR constitutes a right to request the deletion or correction of data. Were data made public, the organization is obliged to take all reasonable measures to have this data also deleted from third parties too.

This is controversial in the context of educational settings. It is an inherent problem that in educational settings an opt-out decision is often not possible for individuals.

Data portability

Upon request, data controllers must provide data subjects with the data processed by them electronically if they process their data automatically. This should allow the data subjects in particular to transfer data from one organization to another.

Group privilege

Group-internal data transfers between affiliated companies can take place under easier conditions. However, the prerequisite is that an adequate level of data protection is guaranteed. This can happen, for example, through intra-group contractual regulations or industry-wide codes of conduct.

This can be relevant for organizations that are composed of several schools or higher education establishments. For example, organization of private schools.

Adequacy of the level of data protection in third countries

Article 41 of the GDPR provides that the EU Commission can now also determine the adequacy of the protection level for individual territories and certain processing sectors within a third country, to which data may subsequently be transmitted.
For example, individual states can be recognized in a third country.

Binding corporate rules

Art. 43 DSGVO provides that binding corporate rules can also justify data transfer to third countries. Such binding corporate rules must be recognized beforehand. Permission to transfer data must be granted (only) by a single data protection supervisory authority.

Personal reference of IP addresses

IP addresses are not generally classified as personal data. This is very controversial, though.

Protection of children

The data of children under the age of 13 may only be processed on the basis of the consent of the legal guardian. This extends the data protection of children.

2. A framework for Data Protection and an Ethical Code of Conduct

To find a balance between learning analytics research and beneficial uses of data, on the one hand, and individual privacy, on the other hand, Learning Analytics solutions need to appropriately address privacy and data protection principles and comply with relevant legal regulations. The aim is to translate the frameworks and guidelines proposed in the literature to deal with ethical and privacy issues, the different relevant jurisdictions (privacy and data protection laws), and the advice and suggestions from existing ethics boards into a coherent set of requirements for Learning Analytics solutions. These requirements should go beyond outlining philosophical ideals, but should actually be applied as ethical principles and to fed into the design and development of technologies. In line with Schwartz (2011), the requirements shall represent an accountable approach reflecting the specific ethical and data protection issues relevant for the project. They shall provide an appropriate frame for researching and exploring the educational possibilities to benefit from learning analytics without sacrificing privacy (Bomas, 2014).

Concretely, a set of principles relating to privacy, data protection, and ethics has been identified, which form the requirements for a future data protection and ethics labelling format. These principles have been derived from an integration and harmonization of general guidelines for fair information practice relating to personal data, ethical frameworks proposed for big data and learning analytics, complemented by the discussion points of the ethics advice, and in alignment with the aspects of data protection and privacy covered by national and European regulations. The following presents an overview and mapping between the ethical and privacy principles from these different resources. The mapping has been done based on the consideration of the individual principles; principles have been mapped to a common topic if a reasonable overlap and pragmatic matching in the idea behind could be identified. It can be seen that the three different types of sources nicely overlap and cover very similar aspects. The principles and requirements derived for this white paper were formulated based on this integration of privacy and data protection resources and the identification of the relevant aspect covered under each topic. While the way, how these principles are actually applied and implemented may take different forms and may change during project lifetime, compliance with the current laws and regulations shall be ensured at any stage of the project as a main requirement of privacy and data protection.

DATA PRIVACY

The first and overarching requirement in this white paper is data privacy, in line with the fundamental right to data protection as reflected in national regulations and the EU data protection directive (Rodotà, 2009). Collection and use of personal data need to be fair and provide appropriate protection of privacy. Information on privacy and data protection practices should be available and easily understandable.

Users having the feeling their privacy is endangered may show resistance (Greller & Drachsler, 2012). To give them the feeling that their data is used in an acceptable and compliant way, policies and guidelines to protect the data from abuse are needed and need to be communicated. The protection of data with respect to data collection and analysis is ensured by legislation and by additional institutional privacy regulations (Campbell et al., 2007), as represented by the privacy principles at hand.

Users' desire for privacy stands somehow in contrast with initiatives in learning analytics research towards greater openness of educational datasets. Both perspectives are comprehensible – private users may not want to disclose their personal data, at least, wish to ensure controlled disclosure of their data, while learning analytics researchers are aiming at getting access to appropriately large data sets to test and refine their methods. The anonymization and processing of data according to legal requirements is a key factor mediating and harmonizing between these positions. In general, there is strong legal protection of personal data (see also Section 3); sometimes even competing with other legal frameworks, like the Freedom of Information Act in United Kingdom (Greller & Drachsler, 2012).

Key implications for the realization of Learning Analytics:

- Privacy and data protection policy should be available for inspection.
- Reasonable measures should be taken to ensure data privacy, including the use of encryption (where appropriate), password protection, minimal and anonymized data exchange through APIs. Insofar as possible the user should not be able to be identified from data that is stored about them, although information will need to persist in the database about the user's credentials in order for the system to render information about learners (e.g. in a non-anonymous form, to teachers, where appropriate).
- Data for analysis should be stored anonymously, associated only with a 'key' (where appropriate). Data should be stored (electronically or otherwise) anonymously and anonymized at the earliest opportunity, where this is not automated.
- Data use must be in alignment with each of the above policies where data is maintained or passes through the country to which the policy pertains.

PURPOSE AND DATA OWNERSHIP

The purpose and boundaries of a learning analytics application should be clearly defined and available before processing is started. “Processing personal data for undefined and/or unlimited purposes is unlawful” (FRA, 2014, p. 68). In essence, considering learning analytics as a moral practice, learning analytics should aim at supporting learners (e.g. Slade & Prinsloo, 2013; The Open University, 2014). When researching new learning analytics methods, though, focus may be on studying new methods of assessment and on establishing a better understanding of learning processes, in a first instance, without implementing any direct consequences or interventions based on analytics outputs. In this case, establishing and ensuring reasonable accuracy of analytics results (i.e. creating truly actionable knowledge) represents the ethical standard to be addressed first (H. Römer, personal communication, 27 November 2014; Schwartz, 2011), before dealing with ethical questions on the responsibility to act or not act based on the new knowledge gained (e.g. Willis, 2014).

Another relevant ethical aspect is data ownership. It has been argued that in this regard there is a lack of legal clarity, when considering learning analytics applications (Greller & Drachlser, 2012). Traditionally, the data collected about a person, i.e. before anonymization, belongs to the owner of the data collection tool (data client). Meanwhile, there is a trend of considering users as the owners of the data collected about them and institutions are borrowing them for a clearly stated purpose. In learning analytics things get more complicated very quickly, since usually data from a whole population of learners is used to produce a prediction model – and the question arises, who the owner of such kind of model is (Pardo, 2014). So, even if the raw personal data is owned by the user, what about the information derived from it? While for raw learning data there is no issue of copyright, copyright and database rights may be relevant for enhanced learning data (e.g. collations of data, prediction models). The owner of any IPR is typically the institution that has collected (and enhanced) the data (Kay et al., 2012).

The question of data ownership is also further complicated when thinking of the integration of learning data from different sources, which may potentially mean different organizations/data clients. It has been argued that to fully exploit the potential of learning analytics and build a holistic picture of an individual’s learning (e.g. Ferguson, 2012; Dyckhoff, 2011), data integration is needed – e.g. institutionally held student data with learning data from educational tools.

It has been argued that, in fact, the concept and consideration of data ownership may not be most appropriate and helpful, but more relevant are the notions of data controller and data processor as used in data protection regulations (Sclater, 2014). Data controller is a natural or

legal person, or an authority, that processes personal data and determines the purpose of processing. The data subject has the right to be provided with information about the identity of the data controller (including contact details) and purposes of processing. A data processor is a separate legal entity, who processes personal data on behalf of the controller (FRA, 2014).

An adequate specification and documentation of the purpose of data processing needs to be ensured at any stage, and must be made available – together with information about the data controller – for access by the data subject or supervisory authorities.

Some of the key implications for realization include:

- Data collected, processed and used is for the purpose of the analytic algorithms and visual methods only.
- The accuracy and maturity of the analytics should be disclosed to those who use them.
- Where integration between multiple entities exists, the user should be informed about where the data has been in order for the analytic to be produced.

CONSENT

Informing users about the collection of their data and gathering their consent need to be realised as a basic ethical principle and procedure (Greller & Drachsler, 2012). It has been argued that in learning analytics there should be virtually no reasons to waive informing users about the use of their data and to set up a clear policy of informed consent (Slade & Prinsloo, 2014).

The principle of consent refers to giving data subjects the possibility to agree/disagree to a data collection and application. The information provided as a basis for gathering consent should establish a balance between allowing research and protecting users from potential harm and thus, may refer to “a broad definition of the range of potential uses to which a student’s data may be put” (Slade & Prinsloo, 2014).

A learning Analytics solution has to apply appropriate techniques for gathering consent from students and parents, as a legal basis for processing personal data. Thereby, consent shall be collected before any information is collected. Consent need to be free, informed, specific and given unambiguously. Sufficient information needs to be provided to the data subject, to assure he/she is clearly informed about the object and consequences of consenting before taking the decision. Information needs to be precise and easy to understand. Consent given non-explicitly on the basis of inactivity (passive consent from parents) is usually not considered as unambiguous and should be avoided (FRA, 2014; H. Römer, personal communication, 27 November 2014). Although the European regulations do not explicitly mention a general right

to withdraw consent at any time, it is widely presumed and accepted that such right exists (FRA, 2014).

In a Learning Analytics environment consent may be collected by clicking a box on the screen, providing the user the choice to agree/disagree with the collection and use of the data being collected from them. This may have to be complemented by paper-based consent gathering from parents or legal guardians. According to current privacy legislation the collection of consent also needs to be implemented for the use of Cookies. In case of gathering consent online, layered information notices have been suggested as a good solution to provide access to adequate information in concise and more extensive versions (FRA, 2014). The language used for information need to be understandable for the concerning group of individuals and consent needs to be collected in an explicit manner, providing the option of later withdrawal of consent.

The following are important for Learning Analytics:

- Informed consent needs to be given using an opt-in policy before data collection takes place (whether this is electronic or otherwise).
- Users have the right to withdraw their permission for the data to be used at any time, therefore the mechanisms handling data collected electronically about students' and teachers' use of technology must have facilities/processes to exclude specific users' data within analysis.
- Even if consent is not given for data to be collected, users should still have access to the analytic facilities.
- For parts of the system where Cookies are used, an explicit agreement from the user is required for proceeding (e.g. using a modal dialogue), if the Cookie will persist after the end of a session of use.
- All consent requests must be in the local language and contain plain wording.

TRANSPARENCY AND TRUST

Transparency is probably the issue that relates to most concerns in ethical considerations on learning analytics (Pardo & Siemens, 2014). While privacy legislation requires learners' consent for data collection, the principle of transparency goes beyond that. Data subjects (i.e. usually learners, but also teachers) should be given notice about what kind of data is gathered and recorded, and should be provided with information on how the analytic processing is done. Transparency also means to provide information on data management procedures, on how data is dealt with after its primary purpose, and whether information is transmitted to outside an institution. Users should, however, not only be informed about how their data is used outside

and educational institution, but also within the institution (Slade & Prinsloo, 2013). In addition, data subjects should also be made aware of the possible outcomes of the data application and the measures of data protection taken (Willis & Pistilli, 2014).

In a Learning Analytics environment, notice and transparency may be created by posting the respective information unavoidable, understandable, and readily accessible at a prominent location on the website. According to the Fair Information Practice Principles (Federal Trade Commission, 1998), notice of the following information is considered essential to consider data subjects as properly informed: the entity collecting the data, the uses to which the data will be put, potential recipients of data, the type of data collected and data collection method, consequences of refusal, and measures taken to ensure data quality and security. Frequently also information on consumer rights is included. In case of learning analytics, an appropriate and understandable description of the analytic models/procedures should be provided (H. Römer, personal communication, 27 November 2014). Data subjects should be enabled to understand what is happening with their data (FRA, 2014).

Informing users about what kind of data is recorded and for what purpose is not only an important ethical and legal privacy principle, but it is also key to foster trust in data subjects – for learning analytics, and for the educational institution applying it. If users trust the learning analytics technology, because they understand the data application and the (potential) value and usefulness it may have to them, users experience and acceptance is considerably enhanced (Pardo & Siemens, 2014). As a result, the application of the principle of transparency should also include information on the potential benefits (or harms) due to the data application, to raise users' awareness and understanding of the learning analytics approach and, potentially, involve them as active agents in the implementation of learning analytics.

This means:

- Information regarding what data is used and how it is gathered, recorded, processed etc. should be easily accessible within the system, and should be easily understandable.

ACCESS AND CONTROL

In addition to gathering users' consent and providing transparency of when and how data is collected and analyzed, data subjects should be given control of their own data. This means, users should be given access to the data collected from them, and the opportunity to correct them, if necessary. The principle of access and participation is reflected in legislation as a right of the data subject. While giving access is completely in line with the idea of transparency, the

aspect of modifying data is somewhat challenging in learning analytics and only applies to certain types of data – i.e. data from plain observations, but not necessarily summaries or results obtained from data. Procedures for correction or deletion of personal data, if inaccurate, misleading, or outdated, need to be provided to users.

In fact, some authors have even claimed to establish a culture of participation, to consider learners as an agents sharing responsibility for the accuracy, maintenance, and up-to-dateness of their student data; they may even be actively involved in the implementation of learning analytics and help shaping interventions (Slade & Prinsloo, 2013; The Open University, 2014). This requires a clear plan and procedure of communication with learners.

Dashboards and open learner models are approaches of visualizing learning analytics data and results. They are often an inherent part of learning analytics approaches as instruments for reporting and fostering reflection (Bull & Kay, 2010; Verbert, Duval, Klerkx, Govaerts, & Santos, 2013). These visual approaches provide users access to the data whenever and for how long they want and thus, offer transparency to data subjects on the data collected about the learning process (Pardo & Siemens, 2014). More recent approaches of negotiated user models reflects the idea of student control, since the open learner model is used to interactively negotiate and potentially update the content of the learner model. Access and control over data need to be governed by technically implementing appropriate authentication mechanisms and the establishment of an access right structure. Simple and understandable procedures for indicating inaccurate data, for updates or corrections, and for verifying information need to be established and implemented in the management and maintenance of data files.

In the context of the realization of a Learning Analytics system this means:

- All data held about users should be available to inspect.
- Facilities to manage underlying data (create, read, update, delete) need to be provided to users, in alignment with the purposes of the tools. In the case of the open learner model, this is also partially addressed by a negotiation component.
- Users should be authenticated (i.e. their identity ascertained) before access to the outcomes of analytic processes may be observed.
- Different users may have different authorization privileges (e.g. teacher may see their students data, students may wish to grant permission for their peers to see their data). Users will always be able to inspect information which is about them.

ACCOUNTABILITY AND ASSESSMENT

Principles of data protection can only work with appropriate mechanisms to enforce and redress them (FRA, 2014). The institution, department or person responsible or accountable for a learning analytics application and its proper functioning need to be identified. A clear structure of responsibilities of individual stakeholders and persons has been established from the outset of the development and/or application of Learning Analytics solutions.

In addition, the learning analytics process should be evaluated in order to refine data collection, management, and analysis (Pardo & Siemens, 2014). The overarching goal of learning analytics is to better understand learning processes and to optimize and support learning and teaching. This can only be achieved when ensuring correctness of the data and analytics algorithms. In fact, when using learning analytics outputs as a basis for taking decisions, for educational interventions etc., the possibility of error should be taken into account and it should be ensured that these outputs are reasonably accurate (cf. Schwartz, 2011). This accuracy cannot be assumed from the outset of introducing and testing a learning analytics approach. As a result, assessment of a new learning analytics is needed; in the beginning ideally through an approach of pure validation and without any direct consequences for data subjects. This is to avoid any harm to users and softens the ethical claim of an obligation to act on the basis of the newly gained knowledge and the question for defining responsibilities for taking action, as frequently discussed in the literature (e.g. Campbell & Oblinger, 2007).

A constant reviewing and adjusting of analytics methods will increase the accuracy of results and suitability of the learning analytics process and maximize impact (Pardo, 2014; Van Harmelen & Workman, 2012). The importance of the review and revision stage in analytics is also highlighted by Schwartz (2011). Beside that, he also refers to the assessment of the impact of using analytics on the basis of stakeholders trust.

Implications are:

- The limitations and maturity of the visual analytics and the underlying processes should be made clear in the system, or minimally to the end users.
- Regular reviews of analytic processes should take place. If changes are made, minimally end users need to be aware of this, and messages within the software should be considered. User expectations must be managed.

DATA QUALITY

According to different ethics frameworks an appropriate quality of data needs to be ensured (e.g. Federal Trade Commission, 1998; OECD, 2013; Pardo & Siemens, 2014). Data needs to be representative, relevant, accurate and up-to date. Information that is not up-to date cannot

be assumed to be reliable or reflecting the current status of a learner and may thus, lead to wrong conclusions from analytics (The Open University, 2014). An approach of sharing responsibility for the accuracy and maintenance of personal data between educational institution and learner (compare 'Access and Control') is considered reasonable for ensuring an adequate level of data quality.

Especially when gathering and combining data from multiple sources care needs to be taken to use reliable sources. It needs to be acknowledged that the data collected may provide an incomplete picture of the learning process and only represents a snapshot in time and context. Bias and stereotyping need to be prevented by constantly taking into account the incomplete and dynamic nature of individual learning and experience (Slade & Prinsloo, 2014).

Beside an adequate quality of learning raw data, it needs to be ensured that data is used wisely for carrying out integration and analysis. Any interpretation, enhancement, or manipulation of data with the aim of extracting meaning should be grounded on a sound technique; the analytics models should be transparent and available for review and testing.

Implications are:

- New information and inferences added should have almost immediate effect in updating the analytics with relevant, accurate, up-to date information. This may mean that analytics require a clear datestamp stating time/date they were last amended.

DATA MANAGEMENT AND SECURITY

In general, personal data needs to be treated and managed in a sensitive and ethical way. Data must be kept protected and secure at different levels and by adequate measures, in accordance with applicable jurisdictions. Accountability, thus, requires safeguards for data protection; compliance of data processing with data protection regulations needs to be demonstrated (FRA, 2014).

Appropriate measures need to be taken to protect the data against unauthorized access, loss, destruction, or misuse. This includes a clearly defined policy of who is authorized to access the data, to which parts of the data and the application, and which kind of data operations are allowed (Pardo & Siemens, 2014). Processes for redress need to be provided to users in case of any unauthorized access or use of personal data. Preservation and storage of data needs to be aligned with national and EU regulations.

Special attention needs to be paid to this principle, in particular when personally identifiable or even sensitive data is managed. Anonymization is often used as a strategy to foster willingness to disclose data. Beside lower reluctance of user to share their data, data protection

regulations are eased with this kind of data, which provides greater flexibility and possibilities in the data application.

In line with this principle of data management and security, the effective governance and stewardship of data should be ensured and a clear and transparent structure of data shall be established. Security thereby needs to involve measures on a managerial and on a technical level (Federal Trade Commission, 1998; FRA, 2014). On the managerial level, internal organizational rules should be established that cover, for example regular information of employees about data security rules, obligations of confidentiality, a clearly defined structure of responsibilities and competencies in data processing and transfer, training on effective security precautions etc. Technical measures for data security relate to having the right equipment (hardware and software) in place, encryption in data transmission and storage, the use passwords to limit access, data storage on secure servers etc.

For the technical realizations, this means:

- Data should be protected (e.g. with regularly changed passwords to access databases and APIs, and encryption of sensitive information).
- Data should be regularly backed up.
- Account management facilities are needed (e.g. to change passwords)
- Data should be anonymized at the earliest opportunity.
- Appropriate security measures should be implemented on all information transfers between components.

3. Conclusion

It is inevitable that data protection and privacy aspects but also ethical considerations will become a more important and a more “regular” part of education. The fulfillment of all this requirements and demands is not always easy – specifically in typical schools or small to medium scale higher education establishments. This white paper intends to provide a brief insight into the relevant legal regulations in Europe and it provides a framework that is supposed to serve as a scaffolding for planning, developing, installing, and applying Learning Analytics in an organizations daily practice.

Data protection is not (only) a burden and perhaps a show-stopper for learning Analytics, it open new pathways towards a more reflected, fair, and effective use of analytics technologies in education.

4. References

Berg, A. (2013, August 21). Towards a uniform code of ethics and practices for learning analytics [Web log post]. Retrieved from https://www.surfspace.nl/artikel/1311-towards-a-uniform-code-ofethics-and-practices-for-learning-analytics/

Bomas, E. (2014, August 29). How to give students control of their data. [Web log post]. Retrieved from http://www.laceproject.eu/blog/give-students-control-data/

Brusilovsky P., Karagiannidis C. & Sampson D. (2004). Layered evaluation of adaptive learning systems. International Journal of Continuing Engineering Education and Life-Long Learning, 14, 402-421.

Campbell, J. P., DeBlois, P. B. & Oblinger, D. G. (2007). Academic analytics. Educause Review, 42, 1–24.

Cavoukian, A. (2011). Privacy by design. The 7 foundational principles. Implementation and mapping of fair information practices. Information and Privacy Comminssioner of Ontario. Retrieved December 12, 2014 from http://www.ipc.on.ca/images/Resources/pbd-implement-7foundprinciples.pdf

Debatin, B., Lovejoy, J.P., Horn, A.-K., & Hughes, B.N. (2009). Facebook and online privacy: Attitudes, behaviors, and unintended consequences. Journal of Computer-mediated communications, 15, 83108.

Dirndorfer Anderson, T. & Gardiner, G. (2014). What price privacy in a data-intensive world? In: iConference 2014 Proceedings (pp. 1227-1230).

Dyckhoff, A.L. (2011). Implications for learning analytics tools: A meta-analysis of applied research questions. International Journal of Computer Information Systems and Industrial Management Applications, 3, 594-601.

Ellicksoon, P.L. & Hawes, J.A. (1989). An assessment of active versus passive methods for obtaining parental consent. Evaluation Review, 13, 45-55.

Ess, C. & AoIR (2002). Ehtical decision-making and Internet research: Recommendation from the AoIR Ethics Working Committee. AoIR.

Federal Trade Commission (1998). Privacy Online: A report to Congress. Federal Trade Commission. United States of America.

Fisher, C.B. (2013). Decoding the ethics code. A practical guide for psychologists. Thousand Oaks: SAGE Publications.

FRA (2014). Handbook on European data protection law. European Union Agency for Fundamental Rights. Council of Europe. Retrieved December 10, 2014 from http://fra.europa.eu/en/publication/2014/handbook-european-data-protection-law

Friedman, B. (1997). Human values and the design of computer technology. Cambridge, MA: Cambridge University Press.

Greller, W. & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. Educational Technology & Society, 15, 42-57.

Gotterbarn, D. (1999). How the new Software Engineering Code of Ethics affects you. IEEE Software, 16, 58-64.

Gotterbarn, D., Miller, K. & Rogerson, S. (1997). Software Engineering Code of Ethics. Communications of the ACM, 40, 110-118.

Government of Canada (2004). Personal Information Protection and Electronic Documents Act. Canada: Minister of Justice.

Markham, A. & Buchanan, E. (2012). Ethical decision-making and Internet research: Recommendations from the AoIR Ethics Working Committee (Version 2.0). AoIR.

Johnson, L., Adams Becker, S., Estrada, V., & Freeman, A. (2014). NMC Horizon Report: 2014 Higher Education Edition. Austin, Texas: The New Media Consortium.

Kay, D., Korn, N., & Oppenheim, C. (2012). Legal, risk and ethical aspects of analytics in higher education. JISC CETIS Analytics Series: Vol. 1 No. 6. Retrieved October 22, 2014 from http://publications.cetis.ac.uk/c/analytics

Kobsa, A. (2007). Privacy-enhanced web personalization. In P. Brusilovski, A. Kobsa, & W. Nejdl (Eds.), The adaptive web: Methods and strategies of web personalization (pp. 628-670). Berlin: Srpnger.

Levin, A. & Nicholson, M.J. (2005). Privacy law in the United States, the EU and Canada: The allure oft he middle ground. University of Ottawa Law & Technology Journal, 2, 357-395.

Long, P., & Siemens, G. (2011). Penetrating the fog. Analytics in learning and education. EDUCAUSE Review, 46, 30-40.

Moore, S. L. (Ed.) (2008). Special Issue: Practical Approaches to Ethics for Colleges and Universities. New Directions for Higher Education, 2008(142), 1-7.

Movius, L.B. & Krup, N. (2009). U.S. and EU Privacy Policy: Comparison of regulatory approaches. International Journal of Communication, 3, 169-178.

OECD (2013). The OECD Privacy Framework. OECD Publishing.

OECD (2013). Exploring Data-Driven Innovation as a New Source of Growth: Mapping the Policy Issues Raised by "Big Data". OECD Digital Economy Papers No. 222. OECD Publishing.

Pardo, A. (2014). Designing learning analytics experiences. In J.A. Larusson & B. White (eds.), Learning analytics: From research to practice (pp. 15-38). New York: Springer.

Pardo, A. & Siemens, G. (2013). Ethical and privacy principles for learning analytics. British Journal of Educational Technology, 45, 438-450.

Range, L., Embry, T., & MacLeod, T. (2001). Active and passive consent: a comparison of actual research with children. Ethical Human Sciences and Services, 3, 23-31.

Rodotà, S. (2009). Data protection as a fundamental right. In Gutwirth, Y. Poullet, P. de Hert, C. de Terwangne,S. Nouwt (Eds.), Reinventing data protection? (pp. 77-82). Dortdrecht: Springer.

Scheffel, M., Drachsler, H., Stoyanov, S., & Specht, M. (2014). Quality indicators for learning analytics. Educational Technology & Society, 17, 117-132.

Schwartz, P.M. (2011). Privacy, ethics, and analytics. IEEE Security and Privacy, 9, 66-69.

Sclater, N. (2014, October 29). Notes from Utrecht Workshop on Ethics and Privacy Issues in the Application of Learning Analytics [Web log post]. Retrieved from http://analytics.jiscinvolve.org/wp/2014/10/29/notes-from-utrecht-workshop-on-ethics-and-privacyissues-in-the-application-of-learning-analytics/

Slade, S. & Prinsloo, P. (2013). Learning analytics: ethical issues and dilemmas. American Behavioral Scientist, 57, 1509-1528.

Spiekerman, S. & Cranor, L.F. (2009). Engineering privacy. IEEE Transactions on Software Engineering, 35, 67-82.

Stiles, R.J. (2012). Understanding and managing the risks of analytics in higher education: A guide. EDUCAUSE. Retrieved December 9, 2014 from http://net.educause.edu/ir/library/pdf/EPUB1201.pdf

Stutzman, F. & Kramer-Duffield, J. (2010). Friends only: Examining a privacy-enhancing behavior in Facebook. In: Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2010) (pp. 1553-15262). ACM: Atlanta.

Tene, O. & Polonetsky, J. (2013). Big data for all: Privacy and user control in the age of analytics. Northwestern Journal of Technology and Intellectual Property, 11, 239-273.

The Open University (2014, Septemer). Policy on ethical use of student data for learning analytics. Milton Keynes: The Open University. Retrieved December 2, 2014 from http://www.open.ac.uk/students/charter/essential-documents/ethical-use-student-data-learninganalytics-policy

The White House (2012). Consumer data privacy in a networked world: A framework for protecting privacy and promoting innovation in the global digital economy. Washington: The White House.

US Government (2004). Code of Federal Regulations. Education. Family Educational Rights and Privacy. 34 CFR Part 99. Washington: Department of Education.

Van Harmelen, M. & Workman, D. (2014). JISC CETIS Analytics Series: Vol.1 No.3, Analytics for learning and teaching. University of Bolton, 2012. Retrieved August 19, 2014 from http://publications.cetis.ac.uk/2012/516

Verbert, K, Duval, E., Klerkx, J., Govaerts, S., & Santos, J.L. (2013). Learning analytics dashboard applications. American Behavioral Scientist, 57, 1500-1509.

Willis, J.E. (2014, August). Learning analytics and ethics: A framework beyond utilitarianism.

Willis, J.E. & Pistilli, M.D. (2014). Ethical discourse: Guiding the future of learning analytics. EDUCAUSE Review. Retrieved December 1, 2014 from http://www.educause.edu/ero/article/ethical-discourse-guiding-future-learning-analytics

Willis, J.E., Campbell, J.P., & Pistilli, M.D. (2013, May). Ethics, big data, and analytics: A model for application. EDUCAUSE Review. Retreived October 28, 2014 from http://www.educause.edu/ero/article/ethics-big-data-and-analytics-model-application