Privacy and Analytics – it's a DELICATE Issue
A Checklist for Trusted Learning Analytics

Hendrik Drachsler
Welten Institute
Open University of the Netherlands
Heerlen, The Netherlands
hendrik.drachsler@ou.nl

Wolfgang Greller
University of Education
Vienna, Austria
wolfgang.greller@phwien.ac.at

ABSTRACT
The widespread adoption of Learning Analytics (LA) and Educational Data Mining (EDM) has somewhat stagnated recently, and in some prominent cases even been reversed following concerns by governments, stakeholders and civil rights groups about privacy and ethics applied to the handling of personal data. In this ongoing discussion, fears and realities are often indistinguishably mixed up, leading to an atmosphere of uncertainty among potential beneficiaries of Learning Analytics, as well as hesitations among institutional managers who aim to innovate their institution's learning support by implementing data and analytics with a view on improving student success. In this paper, we try to get to the heart of the matter, by analysing the most common views and the propositions made by the LA community to solve them. We conclude the paper with an eight-point checklist named DELICATE that can be applied by researchers, policy makers and institutional managers to facilitate a trusted implementation of Learning Analytics.

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

Keywords
Learning Analytics, data management, educational data mining, implementation, privacy, ethics, trust, legal aspects,

1. INTRODUCTION
In 2011, Learning Analytics has been hailed by the Horizon Report [1] as a revolutionary game-changer for teaching and learning. Rapid moves and developments by enthusiasts have, however, slowed down recently due to rising concerns about the impact analytics has on individuals, their identity and integrity, as has already been identified early on by [2]. These second thoughts about Learning Analytics, in our view, originate from and run in parallel to the fears expressed in the wider context of Internet safety, surveillance, and commercial exploitation of data and labour on the Internet.

Since then, vivid academic discussions are taking place on how to provide acceptable approaches for the institutional adoption of Learning Analytics. A number of initiatives have been created to address issues of ethics and privacy in relation to Learning Analytics. At LAK15, the first workshop on "Ethics and Privacy in Learning Analytics" (EP4LA) has been organised jointly by the EU FP7 project Learning Analytics Community Exchange[1] and the SURF SIG Learning Analytics[2], who also organised similar events at other conferences in the Netherlands (Utrecht), US, (Washington), and France (Paris).

How pertinent the issue is can be seen in prominent recent examples like inBloom in the US [3] below. In spite of receiving more than $100m of grant funding by the Gates and Carnegie foundations, and despite the potential benefits, the inBloom analytics system was closed down for good in April 2014, after parents and pressure groups expressed sincere concerns about the misuse of data, the repurposing of data for commercial interests, as well as general safety from cyber-attacks. Another, very similar case concerned the Snappet foundation in the Netherlands which provided tablets to some 400 public primary schools with pre-installed apps for maths and language learning. Snappet re-purposed the collected usage data and used it to classify and predict individual student success. It then provided schools with information on educational interventions. However, an investigation by the national organisation for the protection of personal data (CBP) identified the collected datasets as 'personal data' that needed to be treated according to the Dutch laws on privacy about individuals. The fact that the data collected affected young children and provided insights into their performance at school has been used by the CBP to classify this data as 'highly sensitive' and demanding the highest privacy standards [4] below.

These two cases demonstrate how sensitive the issue of privacy and ethical use of educational data is, particularly when dealing with underage children. It has to be said, though, that the scares of unprotected privacy do not confine themselves to minors, but appear through all age groups. This shows that ignoring the fears and public perception of the application of analytics, no matter how benevolent the intent, can lead to a lack of acceptance, protests, and even failure of entire Learning Analytics implementations.

Acceptance of technological solutions using data in education depends to a great extent on the data subjects being sufficiently aware of the consequences of using the system, the validity and relevance of the results obtained, and the level of transparency of the data model – e.g., how data are collected, stored, processed and shared (cf. the technology acceptance model [5]). A big challenge for Learning Analytics in this respect is the complexity of the data collection and algorithmic analysis processes. The applied technologies are not trivial and it can be rather difficult to

LAK '16, April 25 - 29, 2016, Edinburgh, United Kingdom

ACM 978-1-4503-4190-5/16/04...$15.00
DOI: http://dx.doi.org/10.1145/2883851.2883893

[1] http://www.laceproject.eu/blog/about-todays-ethics-and-privacy-in-learning-analytics-ep4la/

[2] https://www.surfspace.nl/sig/18-learning-analytics/

provide non-technical educational stakeholders (learners, teachers, managers, and external parties like education authorities or parents) with an understanding of how and what data are being collected, how they are processed, and how reliable the results of the analysis are.

The above cases, and others, justify a thorough and open discussion about the issues surrounding the safe and transparent use of Learning Analytics in an educational context. To contribute to this wider academic exchange is the aim and objective of this paper. We approach this complex issue primarily from an institutional and policy perspective with a view to provide some guidance for educational managers, decision makers, and data curators for K12, Higher Education, and work-based learning when implementing privacy-conform and ethically agreed solutions for Learning Analytics. To this end, we developed an eight point checklist named DELICATE that can serve as a reflection aid.

In the remaining parts of this paper, and leading up to the mentioned DELICATE checklist, we work our way through various relevant criteria of ethics and privacy. We added the criteria from each subsection into a matrix that formed the foundation for the DELICATE checklist Our goal is to provide a practical tool that can be used by implementers to quickly check the privacy risks that the introduction of data processes could throw up, and how to deal with them. We, therefore, first summarise the current state of the art on ethics and privacy for Learning Analytics. Thereafter, we try and clean up the substantial overlap in understanding between these concepts, which often leads to confusion when discussing the implications of Learning Analytics. After clarifying the relevant working concepts, we analyse some of the most prominent fears and doubts about Learning Analytics and try to demystify them. Finally, we conclude the paper with the DELICATE checklist that aims to provide a means to support educational organisations in becoming trusted Learning Analytics users.

2. RELATED WORK

In an empirical expert study, Scheffel et al. [6] identified 'data privacy' as the most important aspect for increasing quality and trust in Learning Analytics. This followed an earlier survey among the emerging Learning Analytics community [7], which showed similar opinions, with about two thirds of the surveyed experts believing that Learning Analytics will affect 'privacy and personal affairs'. Despite the apparent prominence of the issue in public and academic thinking, only few papers have been published in this area to date (e.g., [8][9][10][11][12]), and, even fewer policies or guidelines regarding privacy, legal protection or ethical implications [14][15] were developed and publicised.

While the law relating to personally identifiable information is widely understood, there has been insufficient attention on privacy from a user-centred perspective to reassure users of proper conduct when providing them with data services. There are no clearly defined best practices for the anonymisation, reuse, storage and security of educational data. Still, the need for an ethical and privacy-wise acceptable data analysis has been widely acknowledged, including by some national authorities. In late 2015, the SURF foundation of the Dutch Universities released a very comprehensive review of the legal state of educational data in the Netherlands that is in line with European law and shows ways on how Learning Analytics can take advantage of educational data without affecting the legal rights of individuals [13]. Kennisnet, the Dutch innovation foundation for schools, has also set up an initiative and consultation on "privacy by design" in a user-centric approach[3]. "User-centric" design in this context aims at putting the data subjects in control of the data. In previous attempts to organise academic and educational consent in this area, the UK CETIS institute published a consultation paper [16] and the JISC, more recently, developed a code of practice for Learning Analytics [14].

The user-centric aspect also underlies the most recent paper by Slade & Prinsloo [11], which explores the topic from a student vulnerability side using a substantial set of recent literature. Vulnerability is an interesting perspective to take in the management of privacy and ethics. Unfortunately, in their paper, it is not explained what these vulnerabilities actually are in the context of Learning Analytics. What are the concrete risks beyond, for example, that students may get targeted advertisements? How do these vulnerabilities balance against others – non-analytic ones (that is, for example, the vulnerability of dropping out of school education)? Furthermore, it is not clear from their elaboration whether the authors see vulnerabilities as part of the "learning contract" or fiduciary duty of the institution. Nevertheless, they quite rightly state that "privacy management goes beyond the traditional binary of opting in or opting out" and go on to propose a framework for learner agency [ibid.] that explores facets of control such as privacy self-management, timing and focus, and contextual integrity. With this, they clearly demand a more proactive engagement with students and stakeholders to inform and more directly involve them in the ways individual and aggregated data are being managed and used.

In another state-of-the-art approach, Steiner et al. [9] provide a thorough and clearly articulated piece of work in the context of a EU FP7 project called LEA's box[4]. After analysing previous propositions in depth, they synthesise them into a privacy and data protection framework with eight foundational requirements: (1) data privacy; (2) purpose and data ownership; (3) consent; (4) transparency and trust; (5) access and control; (6) accountability and assessment; (7) data quality; and, (8) data management and security. They conclude that technology and tools developed and used in Learning Analytics contexts need to be in line with these foundations and see them as fundamental requirements for a proper code of conduct.

The review of related recent literature already provided us with a rich set of criteria that have been added into the concept matrix when compiling the DELICATE checklist for Learning Analytics practioners. Elements from this part of the study included in the checklist are questions relating to the purpose of the data collection, data ownership and access, legal grounding, informed consent, anonymisation, but also ethical aspects like the transparency of the data collection process.

3. ETHICS, PRIVACY, AND LEGAL FRAMEWORKS

In order to be able to discuss the challenges of ethics and privacy for Learning Analytics, we first need a better understanding of both concepts as well as their relationship towards each other. Below, we will express our summarised views on what constitutes ethics and what is typically meant when talking about privacy. In short, we can say that ethics is a moral code of norms and conventions that exists in society externally to a person, whereas

[3]https://www.kennisnet.nl/fileadmin/kennisnet/publicatie/Personaliseren_in_het_Leren_een_Internationale_Schets.pdf

[4] http://www.leas-box.eu

privacy is an intrinsic part of a person's identity and integrity. The understanding of what constitutes ethical behaviour varies and fluctuates strongly over time and cultures. Privacy, on the other hand, is first and foremost context bound [16], in that it forms the boundary of one's person or identity against other entities. Thus, the understanding of privacy can diverge greatly between, e.g., persons living in large one-room family households and people living in large space single-occupancies, even when they belong to the same culture at the same time. Perceived violation of privacy can occur, when the ethical code of the surrounding society conflicts with the personal boundaries.

In common language and popular thinking, there exists a substantial overlap between ethics and privacy, which sometimes leads to confusion when discussing the effects of Learning Analytics on either of them. Both concepts have in common that they can be approached from various perspectives, especially sociologically, psychologically, and even philosophically and religiously. They manifest themselves differently in a legalistic context and change over time.

3.1 Ethics

Ethics is the philosophy of moral that involves systematising, defending, and recommending concepts of right and wrong conduct. In that sense, ethics is rather different to privacy. In fact, privacy is a living concept made out of continuous personal boundary negotiations with the surrounding ethical environment.

Research ethics have become a pressing and hot topic in recent years, first and foremost arising from discussions around codes of conduct in the biomedical sciences such as the human genome [19], but also, more recently, in the shape of "responsible research and innovation" (RRI) which is being promoted by the European Commission[5].

Historically, the first basic written principles for ethical research originated from the Nuremberg trials in 1949, and were used to convict leading Nazi medics for their atrocities during the Second World War [16]. This so-called Nuremberg Code is the first manifest for research ethics. It contains ten internationally recognised principles for the experimentation on humans:

1. Data subjects must be voluntary, well-informed, and consent to their research participation.
2. The experiment should aim at positive results for society.
3. It should be based on previous knowledge that justifies the experiment.
4. The experiment should avoid unnecessary physical and mental suffering.
5. It should not be conducted when there is any reason to believe that it implies a risk of death or disabling injury.
6. The risks of the experiment should be in proportion to (that is, not exceed) the expected humanitarian benefits.
7. Preparations and facilities must be provided that adequately protect the subjects against the experiment's risks.
8. The staff who conduct or take part in the experiment must be fully trained and scientifically qualified.
9. The human subjects must be free to immediately quit the experiment at any point when they feel physically or mentally unable to go on.
10. Likewise, the medical staff must stop the experiment at any point when they observe that continuation would be dangerous.

The Nuremberg Code stimulated a major initiative in 1964 to promote responsible research on human subjects for biomedical purposes in the Helsinki Declaration [20]. It represents a set of ethical principles developed by the World Medical Association, but it is widely regarded as the cornerstone in ethical human research. The Helsinki Declaration has later been developed into the Belmont Report [21] in 1978 by the US National Commission for the Protection of Human Subjects of Biomedical and Behavioural Science. Both documents build on the basic principles of the Nuremberg Code. Together, they provide the foundation for the modern, ethically agreed conduct of researchers mainly in the medical fields. These are nowadays also taken to apply to technological research and data collection about human subjects. The basic principles can be summarised as:

- Voluntary participation in research;
- Informed consent of the participants, and, with respect to minors, the informed consent of their parents or guardians;
- Experimental results are for the greater good of society;
- Not putting participants in situations where they might be at risk of harm (either physical or psychological) as a result of participation in the research;
- Protected privacy and confidentiality of the information;
- Option to opt-out;

To make it clear what ethical and what unethical research may constitute it is helpful to recall some prominent examples of unethical research from the past. Two infamous cases have been the *Milgram experiment* at Yale University [22], and the Stanford *Prison experiment* [23]. Those experiments are in their nature very different to the evaluation of any data tools in the way they caused harm to their participants. The negative impact of those experiments has been very direct and even physical to the participants of the experiments.

This is rather different to the effects that Learning Analytics research has on its research subjects. For the sake of argument, however, one could envisage dividing a class of students into control and experimental groups for A/B testing, and then, providing a positive stimulus to the participants of group A and a negative stimulus to the others. Such an experimental setup could be considered as unethical, as it would disadvantage group B. Harm in such a case would not be physical, but by deprivation of beneficial learning opportunities.

To prevent such negative impact from research and data usage, ethical committees – or IRB as they are called in the US – are charged with ensuring the protection of fundamental rights of all subjects participating in experiments (i.e., humans, but also animals). In most Western universities, nowadays, any human-subject research has to pass the ethical committee, and research ethics has also become part of the training of young scientists. From this, we dare say, that the risk of having unethical research being conducted at a Western university can be considered rather limited.

That being said, with the rise of Big Data and cloud computing new ethical challenges emerged, spurring the call for ethical committees at Big Data companies like facebook. A recent trigger was the facebook contagion study from 2014 [24], where a team of company researchers manipulated the newsfeed of over 650,000 facebook users without notification or informed consent.

[5] http://ec.europa.eu/programmes/horizon2020/en/h2020-section/responsible-research-innovation

The reaction to this manipulation has been massive among the user community and beyond. However, ethics is a volatile human made concept and what we see after the facebook study, is, that researchers now discuss the pros and cons of the study. Some people argue that the study has indeed been unethical, but, at the same time, contributed new insights into human behaviour [25]. In this context, it remains questionable whether self-regulation of profit-oriented companies via ethical committees would work in practice.

By contrast, as has been indicated in the previous paragraphs, after decade-long debates and risk assessments, the public education and research systems are much more advanced in applied ethics and reflective practice than private enterprises. Still, as the above cases of inBloom and Snappet have shown, the educational domain is extremely sensitive, and, therefore, universities and also the increasing amount of private companies now operating within education need to be very careful with using student data for Learning Analytics.

It is important to mention, that ethical approval is typically only required for 'experiments'. If a system is rolled out directly and becomes part of the operational infrastructure of an institution, ethical approval is not always sought. The rollout of the new system then has to comply with the privacy and data protection laws according to national legislation.

While producing the DELCIATE checklist, this section on ethical foundations provided several aspects that were added into the matrix, e.g., explaining the added value and intent, balancing risks versus benefits, informed consent, option to opt out, training and qualification of staff.

3.2 Privacy

The right to privacy is a basic human right and an established element of the legal systems in developed countries. Already in 1890, Warren & Brandeis [26] wrote an article about "*The Right to Privacy*", where they explained that privacy is the "right to be let alone", and focused on protecting individuals. This right is often challenged in the context of the yellow press with regards to royals and celebrities.

The concept of privacy as a right to be let alone was further developed by Westin in 1968 [27] who made it clear that new technologies change the balance of power between privacy and societal technologies. From this, Westin went on to specify privacy as the "*right of informational self-determination*" and as a vital part for restricting government surveillance in order to protect democratic processes.

According to Westin [ibid.], each individual is continually engaged in a personal adjustment process in which they balance the desire for privacy with the desire for disclosure and interaction with environmental conditions and social norms. Flaherty [28] took the informational self-determination further and claimed that networked computer systems pose a threat to privacy. He first specified 'data protection' as an aspect of privacy, which involves "*the collection, use, and dissemination of personal information*". This concept forms the foundation for fair information practices used by governments globally. Flaherty promoted the idea of privacy as information control. Roessler [29] later operationalised the right to privacy across three dimensions: 1. Informational privacy, 2. Decisional privacy, 3. Local privacy. It is important to note that privacy is not the same as anonymity or data security. They are related concepts that have an effect on privacy, but do not represent privacy as such.

Another important aspect of privacy especially in the age of Big Data is Contextual Integrity. Contextual Integrity is a concept that has arisen in recent years to provide guidance on how to respond to conflicts between values and interests, and to provide a systematic setting for understanding privacy [30]. It is not proposed as a full definition of privacy, but as a framework for evaluating the flow of information between agents (individuals and other entities) with a particular emphasis on explaining why certain patterns of flow provoke public outcry in the name of privacy (and why some do not). Contextual Integrity defines a context specified by roles, activities, norms, and values that interact with one another. The actors in this context are: senders, receivers, subjects and the attributes are data fields.

Contextual Integrity is very much at odds with the Big Data business model that actually aims to collect and integrate as many data sources as possible and gain new insights from those data through overarching mining and analyses. It uses data that has been collected under different pretexts and circumstances. This *repurposing* of data is totally against the concept of Contextual Integrity as described.

In other works, Hildebrand [31] examines the concept of privacy from a legal perspective and concludes that "privacy concerns the freedom from unreasonable constraints that creates the freedom to reconstruct one's identity". Data mining might therefore be seen as impacting the personal development of identity.

Again, this section on privacy has contributed to our DELICATE matrix and checklist. Relevant things we added from it were: data ownership and user control, transparency about data collection, data protection, the need to renew a data contract as it is a dynamic concept, repurposing vs. contextual integrity, control of access to data, option to opt out.

3.3 Legal Frameworks

The European position towards Learning Analytics has been expressed in the European Commission's report: "New Modes of Learning and Teaching in Higher Education" [32]. In recommendation 14, the Commission clearly stated: *Member States should ensure that legal frameworks allow higher education institutions to collect and analyse learning data. The full and informed consent of students* ***must*** *be a requirement and the data should only be used for educational purposes"*, and, in recommendation 15: "*Online platforms should inform users about their privacy and data protection policy in a clear and understandable way. Individuals should always have the choice to anonymise their data."*[6]

They base these recommendations on the EU Data Protection Directive 95/46/EC [33], i.e., the European law on personal data protection, comparable to the OECD's Fair Information Practice Principles [34]. Both are widely accepted frameworks and are mirrored in the laws of many U.S. states and other nations and international organisations.

Directive 95/46/EC defines personal data as "*any information relating to an identified or identifiable natural person ('Data Subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity.*" The Directive is only applicable if automated processing of 'personal data' is employed or if it is part of a 'filing system' (Article 3). This

[6] http://ec.europa.eu/education/library/reports/modernisation-universities_en.pdf

means, in reverse, that if data is properly anonymised, the EU directive 95/46/EC does not apply. However, experts around the world are adamant that 100% anonymisation is not possible (see Section 4.3 below).

With this Directive, the EC is following the previously mentioned ethical frameworks such as the Nuremberg Code and the Helsinki Declaration by demanding consent, legal compliance, and that research should be beneficial for society. But it also specifies some general obligations in Article 6 such as: *Use of personal data need to be:*

- processed fairly and lawfully;
- for specified, explicit and legitimate purposes;
- safeguarded from secondary use and further processing;
- adequate, relevant and not excessive;
- accurate and up to date;
- stored no longer than necessary;

It is important to highlight that these principles are equally valid whether or not the data subjects provided full consent to access and use their data.

With Article 6, EU Directive 95/46/EC addresses a very important aspect towards Big Data business models [35] in restricting the use of data for limited purposes only. Big Data business models are driven by the collection and storing of infinite amounts of data without an expiry date and for later re-purposing. Thus, most of the time, the data is being collected in different contexts, and found later to be of benefit for other information needs. This is diametrically opposed to existing data legislation like the EU Directive 95/46/EC and the concept of Contextual Integrity as explained in the above section on privacy. The reuse of collected data for other (unspecified) purposes is against the information rights of an individual.

The commercial habit of indiscriminate collection and repurposing of data, therefore, strengthened the idea of adding another legal aspect to the EU Directive 95/46/EC which entails *the right to be forgotten.* This legal concept was put into effect on some occasions in the European Union since 2006. It arose from the desire of individuals to "determine the development of their life in an autonomous way, without being perpetually or periodically stigmatised as a consequence of a specific action performed in the past." [36]. Following the high-profile case of EU vs. Google in 2014, search data can now be removed on request by the user. There are, however, serious challenges connected with the right to removal of personal digital footprints and data shadows – the former being the traces left by users in electronic systems, the latter representing data about an individual left by others (e.g. tagged photos in facebook) [37]. More often than not it is unclear against whom such right can be claimed [ibid.]. Things are being complicated even more as recent national and European legislation works towards more data storage for security purposes, as is apparent in the European Data Retention Directive, or, more formally known as Directive 2006/24/EC [38].

Besides the evolution of the legal frameworks that underlie Learning Analytics, the sector of private education service providers also started initiatives towards self-regulation in a move to appease critics after the inBloom case. Some 187 K12 school service providers in the US have signed a *Privacy Pledge*[7] to protect data gained from educational data subjects and especially the school sector. In the UK, the JISC recently published a first draft of a Code of Practice [14] for Learning Analytics. One of the examples mentioned in the Code is the "Ethical use of Student Data for Learning Analytics Policy" which has been established at the Open University UK [15]. The JISC Code of Practice and the Policy by the Open University are good starting points and blueprints for other implementations to follow. In the Netherlands too, the SURF foundation recently released a guiding paper how to treat educational data in a privacy conform way [13].

This section has been very influential on the compilation of our DELICATE checklist. It reiterates items that we already came across in previous subsections, e.g., informed consent, which has been further strengthened through the legal frameworks mentioned in this part. Other perspectives that have been added to the DELICATE matrixes were: lifespan of data, legal grounding, anonymisation, and, the requirement to limit the scope of the analytics to what is essential and necessary to fulfil a defined purpose.

4. FEARS TOWARDS LEARNING ANALYTICS

Researchers and institutions dealing with Learning Analytics are facing up to privacy as a big concern. Many people are not really aware of the legal boundaries and ethical limits to what they can do within the sphere of privacy protection. Institutions, on the one hand, have a fiduciary duty and need to demonstrate care for the well-being and positive development of students, leading them to success in their studies [17]. On the other hand, there is widespread fear of negative consequences from the application of Learning Analytics, such as negative press or loss of reputation and brand value, or even legal liabilities. Despite the enormous promise of Learning Analytics to innovate and change the educational system, there are hesitations regarding, among other things, the unfair and unjustified discrimination of data subjects; violation of personal privacy rights; unintended and indirect pressure to perform according to artificial indicators; intransparency of the Learning Analytics systems; loss of control due to advanced intelligent systems that force certain decisions; the impossibility to fully anonymise data; safeguarding access to data; and, the reuse of data for non-intended purposes. This is a non-exhaustive list, but we need to take all concerns seriously if we are to establish proper implementations for Learning Analytics.

From what has been said, one could conclude that privacy and legal rights are a burden which we need to combat in the same "formal arena", perhaps by amending the laws. Some people already argued for an adjustment of existing privacy rights to the new age of Big Data, like facebook founder Mark Zuckerberg when he famously stated: "Privacy is dead!" [42]. True, privacy got heavily affected by the latest technology developments and the ever rising computer processing power. But, in fact, these basic human rights have been established for very good reasons and we should firstly care to build technologies that support human learning within the existing ethical and legal boundaries, created many years before the Big Data boom, in the Belmont report and Directive 95/46/EC [21][33].

Similarly, we would refrain from solving a weakness in a new learning technology by proposing technical fixes or technological solutions, such as standardisation approaches, e.g., the emerging IMS Caliper specification[8] (cf. also [12]). Instead, we prefer to see this as a "soft" issue, rooted in human factors, such as angst, scepticism, misunderstandings, and critical concerns. Within this

[7] http://studentprivacypledge.org/?page_id=45

[8] http://www.imsglobal.org/activity/caliperram

paper, we, therefore, aim to provide some answers to the fog of misunderstandings around privacy and legal obligations dealing with the semantics of the concepts and translating them into clear action points through the suggested DELICATE checklist. We hope to show that ethical and privacy supported Learning Analytics is indeed possible, and we would like to encourage the Learning Analytics community to turn the privacy burden into a privacy quality label.

There are a wide variety of anxieties expressed with regards to the analysis, sharing and exploitation of personal data, including learner records and digital traces. These are not confined to education alone. Rather, such fears are transferred from other environments (commercial, governmental, social) with little differentiation of the respective domain. Compared to these, we'd argue, the institutional environments in education can be considered as relatively safe havens. This is due to the considerate nature of the education system being in charge of young people's personal development, the long-standing experience and practice of ethical behaviours (including the promotion of such behaviours to the learners), and the public scrutiny and transparency which goes far beyond any other sector. Learners, therefore, should be able to feel in a safe space, where they are allowed to make mistakes without the fear of consequences or unnecessary publicity.

Below, we summarise the most widespread and most critical topics, concerns or arguments against applied Learning Analytics. These have been collected from participants of the earlier mentioned workshop series EP4LA[9] that took place at different locations between October 2014 and March 2015. The organisers received an overwhelming interest from all over the world, which resulted in six international workshops on ethics and privacy with legal and Learning Analytics experts who discussed current regulations and norms for most pressing questions. We later sourced the results of the workshops to further refine our DELICATE checklist.

4.1 Power-relationship, data/user exploitation

One of the criticisms levelled against analytics and Big Data in general is the asymmetrical power relationship it entails between the data controller and the data subject [17]. This can lead to a feeling of being powerless and exploited. That this concern reaches wider than Learning Analytics is demonstrated by events like the 2015 summit of the International Society of Information Sciences (#IS4IS) in Vienna, dedicated to "the Information Society at the crossroads"[10]. One key argument in this context was the exploitation of digital labour, as is the case with facebook – which is widely regarded as the world's biggest advertising company – monetising user data and contributions for commercial profits [40]. Another criticism is the fact that data models are not neutral, but reflect and reversely influence an "ideal" identity. This means that data subjects are given a (designed) data identity by others, which can cause friction, especially in the context of the global economic and cultural divisions between Western and developing world: "people are operating within structured power relations that they are powerless to contest" [41]. To a smaller extent there is also such an asymmetry in power between Learning Analytics providers and users.

While these concerns are real and serious, things present themselves differently in an institutional learning context. Naturally, there has always been an asymmetrical relationship between the institution, the teacher, and the learner. However, the benevolent fiduciary mission and walled garden of education should install confidence and trust also in the use of digital assets and data. Nevertheless, challenges here are the increased openness of teaching tools, using cloud services and social networks, as well as the pressures to commercialise and commodify education [42].

One of the main objectives we strive for with the DELICATE list and surrounding discussion is to encourage data intensive learning providers to develop into 'Trusted Knowledge Organisations' that can demonstrate a responsible, transparent, and secure treatment of personal data.

4.2 Data ownership

At present, there is no clear regulation for data ownership of any party, i.e. neither the student, the university or a third party provider. Data ownership is a very difficult legal concept. It is assumed, that the digital footprints someone leaves behind belong to the data subject, i.e. the users themselves, as long as it isn't specified differently in the terms of service of the provider of the digital system. An additional factor is that the data subject cannot manage all those tons of data breadcrumbs on a daily basis. Thus, data subjects are in need of having service providers take care of the data storage and management. Furthermore, a single data subject has no ownership rights on an aggregated data model computed out of their digital footprints. But, according to EU Data Protection Directive 95/46/EC article 12, the data subject always has the right to know about all the information that has been collected about them. Just recently, a law student from Austria brought down the so-called Safe Harbour data transfer agreement between the European Union and the United States used by more than 4.000 companies, including Google, Facebook, and IBM[11]. He won a law suit that took more than two years and forced facebook and other Big Data providers to keep all data collected from a data subject in the same country and provide an overview of this data to the users[12].

Data ownership becomes even more complex, however, when we consider the processing of data. If there is a computational model developed from a collection of data traces in a system, can a student still opt-out of such a data model that is being generated around their data traces? In search for answers, data ownership remains a very complicated issue that is mainly dominated by the technical power the developers and service providers offer to the data subjects. There are visions to change this power relationship by enabling individuals to carry and curate their own datasets through "personal data stores". This idea parallels earlier such user-centric solutions, like openID for generic authentication across the web. It could fundamentally turn the whole data movement in education around if it were to become real. One such attempt has already been initiated with the MIT's openPDS[13], which "*allows users to collect, store, and give fine-grained access to their data all while protecting their privacy"*. Kristy Kitto has made another attempt in 2015 in a blog post at the LACE project where she made a call towards a manifesto for data ownership[14].

[9] http://www.laceproject.eu/blog/about-todays-ethics-and-privacy-in-learning-analytics-ep4la/

[10] http://summit.is4is.org/is4is-summit-vienna-2015

[11] http://www.reuters.com/article/2015/10/07/eu-ireland-privacy-schrems-idUSL8N1272Z820151007

[12] http://europe-v-facebook.org

[13] http://openpds.media.mit.edu/

[14] http://www.laceproject.eu/blog/towards-a-manifesto-for-data-ownership/

For the DELICATE checklist we added from this section the demands to ask for consent, and to strictly monitor who has access to data.

4.3 Anonymity and data security

According to PricewaterhouseCoopers[15], on average, around 117,339 attempts of information heists were made per day in 2014. These statistics reveal the severity of the problem of data security. Such attacks bring even big players like Apple, Sony, or Microsoft into trouble and they suffer from not being able to protect their online systems in an adequate way. Cyber-attacks are perhaps an even greater threat for universities and smaller educational providers who have fewer resources to establish and maintain appropriate data security measures. Naturally, this then becomes an issue for the protection of student and staff Learning Analytics data.

Anonymisation is often seen as the "easy way out" of data protection obligations. Institutions use various methods of de-identification to distance data from real identities and allow analysis to proceed [43]. Most of the time, though, data controllers consider replacing identifiers as sufficient to make data anonymous. But many studies have shown that this kind of anonymisation can, at best, be considered as pseudonymisation [44]. Anonymised data can rather easily be de-anonymised when they are merged with other information sources. Famous approaches include de-anonymisation of medical datasets: names of patients were not included but demographic information that could be linked to electoral registers allowed the retrieval of names and contact information for the medical records [45]. Among those relatively easy approaches are also more computational ones as presented in [46][47]. They showed how Netflix users could be re-identified from the dataset of a Netflix competition by combining it with data from the movie platform IMDB. Those and other examples show that robust anonymisation is hardly achievable with the increase of computational methods, data integration and calculation power. Some residual risk of identification has to be taken into account for educational data and Learning Analytics. Data security and encryption, therefore, has a vital role to play in the acceptance of such systems and combined security strategies, including anonymisation, can go a long way to protect individual privacy from illegal access.

As a further measure to make data less sensitive over time and to protect privacy, data degradation has been suggested whereby a timestamp is introduced when data should be deleted in order to prevent further use [48]. This approach lets data decay over time and in that way protects privacy and informational self-determination of data subjects.

It is rather obvious that this section mainly contributed to the criteria matrix of the checklist in the areas of anonymsation and (physical) data security. In that way, it helped weighting the criteria collected in previous sections. Furthermore, timestamping data for expiry or access was seen as another reasonable point to make.

4.4 Privacy and data identity

It can be argued that our (external) identity is made up by the sum of our actions and appearances. Arora [44] critically contrasts the two notions of system identity versus social identity, the latter encapsulating who we are in our social environment, while the former represents the collected digital data. This has become even more sensitive with the collection of personal biometrical information. Logging, tracking and storing individuals online, therefore, can be considered an intrusion into our identity and a violation of our perceived selves, i.e. the information produced by our actions and appearances – exploited for other purposes by third parties without consent or even knowledge. Profiling actions by companies and governments have been under fire for some time. Their impact on the formation and declaration of a person's own identity, the democratic society, and civic self-determination have been highlighted by some scholars (e.g. [31][48]). Repurposing and fragmentation of personal digital identities isn't the only criticism raised in the context of personal integrity, but the fact that individuals are typecast into data identities not of their own choosing and measured against benchmarks, indicators, or algorithmic stereotypes that are out of their control [41]. What is more, if one doesn't fit into a data model, it is often applied on a probabilistic basis [49].

Here again, the special relationship in education can ease the problem. Students are in a "learning contract" with the institution or training provider they sign up with. For the duration of this relationship, the teacher and institution need to be trusted to act responsibly and in the favour of its clients. This is unlike the ambiguous commercial relationship that for-profit enterprises have towards their investors and shareholders versus customers.

The development of DELICATE has been influenced by these thoughts via the call for transparency of data design and collection, control of data usage by third parties, and the data contract as a dynamic concept that needs to be approved by the data subjects in case of relevant changes.

4.5 Transparency and trust

It is often said that lack of transparency can cause unease and concern with data subjects. However, it is rarely defined how this transparency should manifest itself. Commercial providers like Google keep their algorithms secret, and, yet, as long as results are relevant and in line with users' expectations, there is trust in the service, despite it being a black box. On the other hand, Google Takeout, which allows insight and export of the user dataset from all Google services, is of no great use to ordinary end users, as they are unable to understand or re-use that data. Here then, again, we see the asymmetrical power relationship: while transparency of the user leads to commodification of their data, the reverse isn't true for large companies.

The Korean-born German philosopher Byung-Chul Han [50] states that transparency turns social systems of trust into control systems because information is too readily available. Learning Analytics is perceived as making learners/teachers transparent and open for criticism, while keeping the system itself opaque and out of scrutiny. There is widespread anxiety in the education community that data and information retrieved from data subjects may be used against "outliers", and, thus, leads to more conformity and uniformity [2]. As such, analytics is perceived by some as an engine for controlling and correcting behaviours.

The issue in education can best be tackled by being clear and open about the purpose of data collection and the compliance with the existing legal frameworks. A code of conduct can clarify to the stakeholders and data subjects what the intentions for analytics are, how long data is being kept, and what internal procedures are available to contest negative consequences. As an additional measure, the focus of analytics should be put on providing information for human decision making, prediction and self-reflection rather than accountability. Playing analytics results back to the

[15] http://www.pwc.com/gx/en/consulting-services/information-security-survey/key-findings.jhtml

data subject and letting them decide for themselves, whether to ask for pedagogic support and intervention or not, puts the learner in control, something that is anyway desirable as an educational outcome.

This section again emphasises the aim of DELICATE to establish trusted Learning Analytics in support of all stakeholders to make more of their educational eco-system rather than manifesting a monitoring and surveillance system that will destroy any trusted teacher-learner relationship and scares them of making mistakes and learn from them.

5. INTRODUCING DELICATE

The DELCIATE checklist is derived from the intensive study of the legal texts mentioned above, a thorough literature review and several workshops with experts, e.g., the six EP4LA workshops with over 100 experts involved[16]. Relevant information from those sources has been entered into a matrix on ethics and privacy criteria that are relevant for establishing 'trusted Learning Analytics'. The criteria that have been added to the matrix have been further weighted in case they have been mentioned multiple times in the various subsections of our studies. We used this weighting to identify the importance of a criterion to be added to the DELICATE checklist. In a later step, we sorted the collected criteria into clusters and labelled them. Lastly, we presented these clusters of criteria to a group of experts from the LACE project for feedback and verification. In a final step, we designed the DELICATE checklist into a memorable format in order to reach the goal of providing a largely self-explanatory practical tool for establishing trusted Learning Analytics within any data-driven educational organisation.

We strongly believe that trusted Learning Analytics are a key vision to establish a learning organisation that has the bravery to talk about mistakes and failures and learn from them. In order to establish this level of 'trust', regulations need to be in place that guards the personal information rights but also empowers the organisation to gain insights for its improvement. We believe that the DELICATE checklist can be a practical means towards this vision. It can be applied as a quality checklist to review if the data processing procedures fulfil the trusted Learning Analytics regulations of an educational institution. Similar to medical checklists or aviation checklists [51], it safeguards critical processes and objectives by raising the attention and awareness of the involved stakeholders and guides them through the procedure. The DELICATE checklist can be used in the context of a value-sensitive design approach as specified by Friedman in 1997 [52]. Value-sensitive design is the idea that ethical agreements and existing privacy laws need to be embedded where and when it is relevant for the design and usage of a system like Learning Analytics – starting early on in the design and implementation process, and close to where the technology is being rolled out.

The final version of the DELICATE checklist contains eight action points that should be considered by managers and decision makers planning the implementation of Learning Analytics solutions either for their own institution or with an external provider. Figure 1 shows the full overview of the checklist and all its relevant sub questions. The full version of the checklist can be downloaded from the LACE website[17].

[16] http://www.laceproject.eu/blog/about-todays-ethics-and-privacy-in-learning-analytics-ep4la/

[17] http://www.laceproject.eu/ethics-privacy/

6. CONCLUSIONS

In order to use educational data for Learning Analytics in an acceptable and compliant way, and to overcome the fears connected to data aggregation and processing, policies and guidelines need to be developed that protect the data from abuse and ensure treatment in a trustful way.

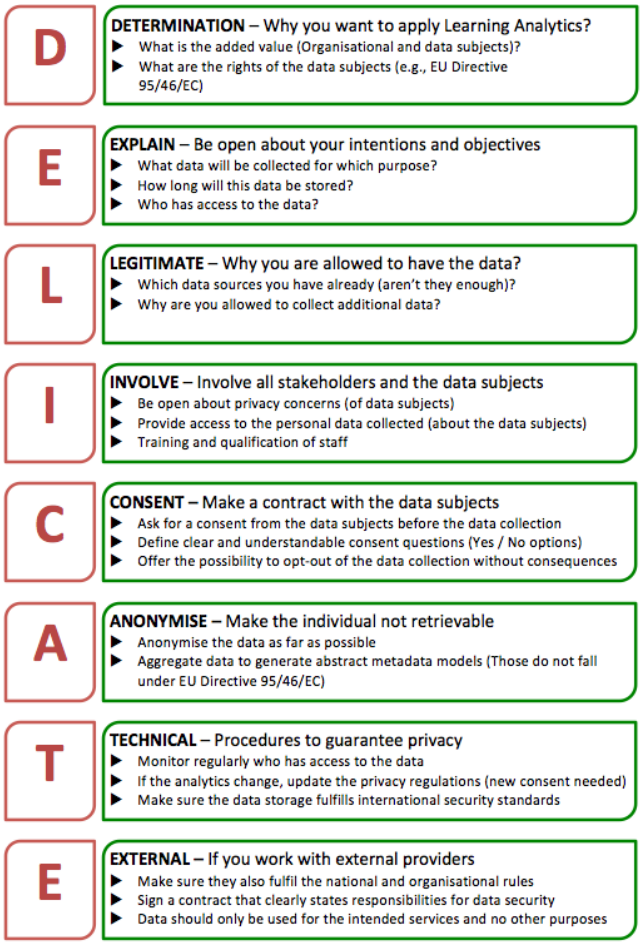

Figure 1: The DELICATE Checklist. © Drachsler & Greller, 2016.

We need to see data protection not as a mere legal requirement, but should embed the care about privacy deeply into Learning Analytics tools and increase the trust of data subjects in these systems. Privacy should not been seen as a burden but rather as a valuable service we can offer to build trusting relationships with our stakeholders. To build this kind of relationship, a high degree of transparency, combined with reassuring explanations referencing relevant legislation like the EU Directive 95/46/EC are needed. Therefore, the "contract" between learners and their educational providers needs to be reviewed and adjusted to reach this level of trust and with it the backing to release the full potential of Learning Analytics.

In conclusion, we believe that Learning Analytics projects should follow a value-sensitive design process, which allows considering ethical and privacy values on the same level as functional requirements. Thereby, the aforementioned ethical considerations are not seen as an unfortunate constraint, but help to develop a system that achieves its aims not only in a technical but also in an ethical and humane manner. At the same time, we request that the education sector as a whole and institutions in particular need to

get better in distinguishing themselves from commercial profit-oriented enterprises and advertise their mission of care and support for individuals to revive trust in the system.

We hope that the DELICATE checklist will be a helpful instrument for any educational institution to demystify the ethics and privacy discussions around Learning Analytics. As we have tried to show in this article, there are ways to design and provide privacy conform Learning Analytics that can benefit all stakeholders and keep control with the users themselves and within the established trusted relationship between them and the institution.

ACKNOWLEDGMENTS

The efforts of Hendrik Drachsler have been partly funded by the EU FP7 LACE-project.eu (grant number 619424). We thank all the enthusiastic participants around the planet who attended the EP4LA workshop series. Those smart discussions supported us in crafting the DELICATE checklist.

REFERENCES

[1] Johnson, L., Smith, R., Willis, H., Levine, A., & Haywood, K. (2011). The 2011 Horizon Report. Austin, Texas: The New Media Consortium.

[2] Greller, W., & Drachsler, H. (2012). Translating Learning into Numbers: A Generic Framework for Learning Analytics. Educational Technology & Society, 15 (3), 42–57.

[3] New York Times, (2014) InBloom Student Data Repository to Close. April 21, 2014. Available at: http://bits.blogs.nytimes.com/2014/04/21/inbloom-student-data-repository-to-close/?_r=0

[4] College Bescherming Persoonsgegevens. Onderzoek CBP naar de verwerking van persoonsgegevens door Snappet. 2014. Available at: https://cbpweb.nl/sites/default/files/downloads/mijn_privacy/rap_2013_snappet.pdf.

[5] Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. Management science, 46(2), 186-204. DOI 10.1287/mnsc.46.2.186.11926.

[6] Scheffel, M., Drachsler, H., Stoyanov S., & Specht, M. (2014). Quality Indicators for Learning Analytics. Educational Technology & Society, 17 (4), 117–132.

[7] Drachsler, H., & Greller, W. (2012). The pulse of learning analytics: understandings and expectations from the stakeholders. In Proceedings of the 2nd international conference on learning analytics and knowledge (pp. 120-129). ACM. DOI 10.1145/2330601.2330634.

[8] Prinsloo, P. & Slade, S. (2013). An evaluation of policy frameworks for addressing ethical considerations in Learning Analytics. In Proceedings of the 3rd Learning Analytics and Knowledge conference, Leuven, Belgium. DOI: 10.1145/2460296.2460344

[9] Slade, S. & Prinsloo, P. (2013). Learning analytics: ethical issues and dilemmas. American Behavioral Scientist, 57 (10), pp. 1510–1529. DOI: 10.1177/0002764213479366

[10] Pardo, A. & Siemens, G. (2014). Ethical and privacy principles for Learning Analytics. British Journal of Educational Technology, 45(3), 438–450. DOI: 10.1111/bjet.12152.

[11] Prinsloo, P. & Slade, S. (2015). Student privacy self-management: implications for Learning Analytics. In Proceedings of the 5th international conference on learning analytics and knowledge Pooghkeepsie, New York, USA. DOI: 10.1145/2723576.2723585

[12] Hoel, T. & Chen, W. (2015). Privacy in Learning Analytics – Implications for System Architecture. In: Watanabe, T. and Seta, K. (Eds.) (2015). Proceedings of the 11th International Conference on Knowledge Management.

[13] Engelfriet, E., Jeunink, E., Maderveld, J. (2015). Handreiking Learning analytics onder de Wet bescherming persoonsgegevens. SURF report. https://www.surf.nl/kennisbank/2015/learning-analytics-onder-de-wet-bescherming-persoonsgegevens.html

[14] Sclater, N., & Bailey, P. (2015). Code of practice for learning analytics. Available at: https://www.jisc.ac.uk/guides/code-of-practice-for-learning-analytics

[15] Open University UK (2014). Policy on Ethical use of Student Data for Learning Analytics. Available: http://www.open.ac.uk/students/charter/sites/www.open.ac.uk.students.charter/files/files/ecms/web-content/ethical-use-of-student-data-policy.pdf

[16] Kay, D., Korn, N., and Oppenheim, C. (2012). CETIS Analytics Series: Legal, Risk and Ethical Aspects of Analytics in Higher Education, serial number: ISSN 2051-9214 Vol 1, No 6. Available at: http://publications.cetis.org.uk/2012/500

[17] Slade, S., & Prinsloo, P. (2015). Student vulnerability, agency and learning analytics: an exploration. Journal of Learning Analytics, Special Issue on Ethics and Privacy.

[18] Steiner, C., Kickmeier-Rust, M.D., and Albert, D. (2015). LEA in Private: A Privacy and Data Protection Framework for a Learning Analytics Toolbox, Journal of Learning Analytics, Special Issue on Ethics and Privacy.

[19] Lauss, G., Bialobrzeski, A., Korkhaus, M., Snell, K., Starkbaum, J., Vermeer, A.E., Weigel, J., Gottweis, H., Helén, I., Taupitz, J. & Dabrock, P. (2013). Beyond Genetic Privacy. Past, Present and Future of Bioinformation Control Regimes. Available at: http://private-gen.eu/uploads/media/PRIVATE_Gen_FINAL-REPORT_2013_02.pdf

[20] World Medical Association. (2001). World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. Bulletin of the World Health Organization, 79(4), 373.

[21] National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, Bethesda, MD. (1978). The Belmont report: Ethical principles and guidelines for the protection of human subjects of research. ERIC Clearinghouse.

[22] Milgram, Stanley. Das Milgram-Experiment. Reinbek: Rowohlt, 1974.

[23] Zimbardo, P. G., Maslach, C., & Haney, C. (2000). Reflections on the Stanford prison experiment: Genesis, transformations, consequences. Obedience to authority: Current perspectives on the Milgram paradigm, 193-237.

[24] Kramer, A. D., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. Proceedings of the National

Academy of Sciences, 111(24), 8788-8790. DOI: 10.1073/pnas.1320040111.

[25] Kleinsman, J., & Buckley, S. (2015). Facebook Study: A Little Bit Unethical But Worth It? Journal of Bioethical inquiry, 12, Issue 2, 179-182, DOI: 10.1007/s11673-015-9621-0.

[26] Warren, S. D., & Brandeis, L. D. (1890). The right to privacy. *Harvard law review*, 193-220.

[27] Westin, A. F. (1968). Privacy and freedom. *Washington and Lee Law Review*, *25*(1), 166.

[28] Flaherty, D. H. (1989). Protecting privacy in surveillance societies: The federal republic of Germany, Sweden, France, Canada, and the United States. UNC Press Books.

[29] Roessler, B. 2005. *The Value of Privacy*, Oxford: Polity Press.

[30] Nissenbaum, H. (2004). Privacy as Contextual Integrity. Washington Law Review 79(1). Available: https://www.nyu.edu/projects/nissenbaum/papers/washingtonlawreview.pdf

[31] Hildebrandt, M. (2006). 3. Privacy and Identity. *Privacy and the criminal law*, 43.

[32] European Commission (2014). New modes of learning and teaching in higher education. Luxembourg: Publications Office of the European Union 2014, 68 pp. ISBN 978-92-79-39789-9 DOI:10.2766/81897 http://ec.europa.eu/education/library/reports/modernisation-universities_en.pdf

[33] European Union: EU Data Protection Directive 95/46/EC. http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:31995L0046.

[34] Organisation for Economic Co-operation and Development. (2002). OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. OECD Publishing. http://www.oecd.org/sti/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm

[35] Mayer-Schönberger, V., & Cukier, K. (2013). Big data: A revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt.

[36] Mantelero, A. (2013). The EU Proposal for a General Data Protection Regulation and the roots of the 'right to be forgotten'. Computer Law & Security Review 29 (3): 229–235. Available at: https://www.academia.edu/3635569/The_EU_Proposal_for_a_General_Data_Protection_Regulation_and_the_roots_of_the_right_to_be_forgotten_

[37] Koops, B. J. (2011). Forgetting footprints, shunning shadows: A critical analysis of the 'right to be forgotten' in big data practice. In: SCRIPTed, Vol. 8, No. 3, pp. 229-256, 2011; Tilburg Law School Research Paper No. 08/2012. Available at: http://dx.doi.org/10.2139/ssrn.1986719

[38] European Union: EU Data Retention Directive 2006/24/EC http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32006L0024

[39] The Guardian (2010), Privacy no longer a social norm, says facebook founder. (January, 11,2010) http://www.theguardian.com/technology/2010/jan/11/facebook-privacy

[40] Fuchs, C. (2015). Digital Labour and Karl Marx. Routledge.

[41] Arora, P. (2015). Bottom of the Data Pyramid: Big Data and the Global South. In: Discover Society, vol. 23. Available at: http://discoversociety.org/2015/08/03/bottom-of-the-data-pyramid-big-data-and-the-global-south/

[42] Casey, J., & Greller, W. (2015). Jane Austen and the Belly of the Beast Part 2 - Language and Power: Commodification, Technology and the Open Agenda in Higher Education. In: ISIS Summit Vienna 2015—The Information Society at the Crossroads. Multidisciplinary Digital Publishing Institute. Available at: http://sciforum.net/conference/isis-summit-vienna-2015/paper/2910/download/pdf

[43] Tene, O., & Polonetsky, J. (2012). Privacy in the age of big data: a time for big decisions. *Stanford Law Review Online*, *64*, 63. Available at: https://www.stanfordlawreview.org/online/privacy-paradox/big-data

[44] Ohm, P. (2010). Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization, 57 UCLA L. Rev. 1701.

[45] Sweeney L. (2000). Simple Demographics Often Identify People Uniquely. Carnegie Mellon, Data Privacy Working Paper 3. http://dataprivacylab.org/projects/identifiability/.

[46] Narayanan, A., Shmatikov, V. (2008). Robust De-anonymization of Large Datasets (How to Break Anonymity of the Netflix Prize Dataset). The University of Texas at Austin February 5, 2008. http://arxiv.org/pdf/cs/0610105.pdf

[47] Narayanan, A., and Felten, E.W. (2014). No silver bullet: De-identification still doesn't work. http://randomwalker.info/publications/no-silver-bullet-de-identification.pdf

[48] Heerde, H.J.W. van (2010). Privacy-aware data management by means of data degradation - Making private data less sensitive over time. Ph.D. thesis, University of Twente, Enschede. http://www.vanheerde.eu/phdthesis.pdf.

[49] Hildebrandt, M. (2008). Profiling and the identity of the European citizen. In: Profiling the European citizen (pp. 303-343). Springer Netherlands. DOI: 10.1007/978-1-4020-6914-7.

[50] Han, B.-C. (2015). The Transparency Society. Stanford University Press. ISBN: 9780804794602

[51] Clay-Williams R., Colligan L. (2015). Back to basics: checklists in aviation and healthcare. BMJ Qual Saf 2015, 24: 428 – 431. DOI:10.1136/bmjqs-2015-003957

[52] Friedman, B., ed. (1997). Human values and the design of computer technology. Cambridge University Press.