An evaluation of policy frameworks for addressing ethical considerations in learning analytics

Paul Prinsloo
University of South Africa
TVW4-69, P O Box 392, Unisa
0003, South Africa
+27 12 429 3683
prinsp@unisa.ac.za

Sharon Slade
Open University
Foxcombe Hall, Boars Hill
Oxford, UK
+44 1865 327000
sharon.slade@open.ac.uk

ABSTRACT
Higher education institutions have collected and analysed student data for years, with their focus largely on reporting and management needs. A range of institutional policies exist which broadly set out the purposes for which data will be used and how data will be protected. The growing advent of learning analytics has seen the uses to which student data is put expanding rapidly. Generally though the policies setting out institutional use of student data have not kept pace with this change.

Institutional policy frameworks should provide not only an enabling environment for the optimal and ethical harvesting and use of data, but also clarify: who benefits and under what conditions, establish conditions for consent and the de-identification of data, and address issues of vulnerability and harm. A directed content analysis of the policy frameworks of two large distance education institutions shows that current policy frameworks do not facilitate the provision of an enabling environment for learning analytics to fulfil its promise.

Categories and Subject Descriptors
K.3.1 [Computers and Education]: Computer Uses in Education - *Distance learning,* K.7.4 [The Computing Profession]: Professional Ethics – *Codes of ethics*

General Terms
Management, Documentation, Security, Legal Aspects.

Keywords
learning analytics, ethics, distance learning, policy

1. INTRODUCTION
The majority of institutions have long employed academic analytics for reporting, operational and financial decision-making, and quality assurance purposes [1,2]. However, while learning analytics is heralded as one of the key trends expected to significantly impact on the shape of higher education within the next few years [3,4,1], many institutional policy frameworks fail to fully reflect the use and ethical implications of learning analytics. The increasing digitisation of learning has resulted in the availability of real-time data on an unprecedented scale, creating new opportunities for harvesting digital trails of students' (non)engagement [5,1,6]. Knowing more about students' learning processes and trajectories allows the potential for higher education institutions to offer personalised and customised curricula, assessment and support to improve learning and retention [1,7].

LAK '13, April 08 - 12 2013, Leuven, Belgium

Realising the promise of learning analytics will require institutions to align their policies with national and international legislative frameworks; to consider the ethical issues inherent in the harvesting, use and dissemination of data and to ensure an enabling environment for adequate resourcing and integration of institutional support.

This paper will analyse the existing policy frameworks of two large distance education institutions according to a set of considerations developed by Slade and Prinsloo [8]. The Open University in the UK (OU) operates largely within a developed world and the University of South Africa (Unisa) within a developing world context, each with considerably different student profiles, business architectures and programme qualification mixes. However, both are specialist open distance learning (ODL) institutions with huge student numbers. Furthermore, both grapple with balancing the tensions inherent in the massification of higher education and openness, and their commitment to quality provision, accreditation and planning interventions to address concerns about throughput rates.

2. LEARNING ANALYTICS AS MORAL PRACTICE - A SOCIO-CRITICAL APPROACH
Defining and addressing ethical issues in learning analytics depend on a number of epistemological and ideological assumptions. Our own epistemological and ideological framework falls within the broad scope of a socio-critical approach which entails being critically aware of the way our cultural, political, social, physical and economic contexts and power-relationships shape our responses to the ethical dilemmas and issues in learning analytics Choosing a specific socio-critical approach allows, inter alia, engagement with both the potential and challenges of learning analytics and recognises the unequal power-relations between students and the institution.

Analysing a diverse range of literature, Slade and Prinsloo [8] propose a number of considerations from which institutions can develop context-specific and appropriate guidelines and policy frameworks. These considerations are:

- Who benefits and under what conditions?
- Conditions for consent, de-identification and opting out
- Vulnerability and harm
- Collection, analyses, access to and storage of data

3. METHODOLOGY

This paper entailed a two-stage qualitative research design. The first stage applies a *directed* content analysis approach [9]. Institutional policies from both the OU and Unisa were identified and cross-analysed to defamiliarise and ensure validity and reliability. Trustworthiness was ensured by member-checking.

Neither institution has policies covering the analysis and use of data with the explicit purpose of understanding, predicting and influencing student learning. Both have specific policies relating to the ethical use of student data for research. All other policies or frameworks with reference to monitoring, surveillance, privacy and security of data, etc., in the two institutions were identified and reviewed for cues or guidelines specifically pertaining to the use, analysis, dissemination and storage of educational data.

The second stage of the research design involved the evaluation of the respective policies of the OU and Unisa against the set of considerations [8] discussed above.

4. AN OVERVIEW OF CURRENT POLICY FRAMEWORKS AT UNISA AND THE OU

There is currently no single policy at either institution which covers, per se, the notion of learning analytics. In order to establish potential gaps or guidelines pertaining to proposed considerations [8], all policies of both institutions relating to data, teaching and learning and research were scrutinized.

4.1 Who Benefits and Under What Conditions?

While authors such as Kruse and Pongsajapan [10] propose a "student-centric" approach to learning analytics, Slade and Prinsloo [8] state that it is crucial that *all* key stakeholders in learning analytics should benefit from learning analytics. Such an approach not only involves students, but everyone involved in delivering and supporting learning.

However, we acknowledge that students are a major stakeholder and should be seen as active agents in defining the purpose and scope of collected data, as well as conditions for its use (e.g. de-identification). While there is some information accepted to fall within the normal scope of the registration agreement between the student and the institution, other categories of information will require informed consent and an active commitment from students to ensure the correctness and currency of data.

4.1.1 Unisa

While the different policies and guidelines pertaining to conducting research using student data address issues of privacy, informed consent, vulnerability, harm and benefit, there is no indication in any Unisa policy or guideline document communicating the harvesting and the conditions of harvesting of data from students to the student community. References are provided for Unisa documents in the public domain.

The **Guidelines for Conducting Research Involving Unisa Staff, Students or Data** [11] explicitly excludes institutional research or "authorised routine data gathering activity, necessary for the efficient administration and operation of Unisa" (p.1).

When students register, they declare that the information they provide is correct and current.

The benefits of harvesting and analysing educational data are not explicitly highlighted in any Unisa policy or guideline document. Students are also not provided with information on how their data is used, by whom and under which conditions.

4.1.2 Open University

The majority of OU policy documents may be accessed through the OU website [12]. The OU **Student Community Charter** includes a number of overarching principles which aim to establish a shared responsibility between University and student. The Charter sets out the University's intention to "anticipate and respond positively to different needs and circumstances" by providing "support that is appropriate for each individual learner and their subject area of study."

The **Retention of Student Data and Records** policy makes clear that some data will be depersonalised and retained for uses relating to management, development and research.

The policy states that personal data owned by the university may be shared with third parties; and conversely personal data owned by other organisations may be shared with the OU. Where the third party is acting as a university agent, the university remains the data controller. The third party must adhere to the university's student data retention and security policies. Where the third party takes ownership of OU student data (e.g. government agencies, sponsors, etc.), the third party becomes the data controller. The data is then subject to that third party's data retention policies.

The alumni office has an objective to both keep alumni up to date with University activities and to pursue donation prospects. The policy setting out the **Retention of Alumni Data and Records** requires alumni contact details and contact history (e.g. donations made) to be held indefinitely.

4.2 Conditions for Consent, De-identification of Data and Opting Out

Considering conditions for consent, de-identification of data and the option to opt-out of the collection of certain types of data refers to the notion of informed consent and transparency, (e.g. information regarding the uses to which their data might be put, algorithms used to analyse data, etc.)

4.2.1 Unisa

While there are guidelines which include explicit conditions for research on Unisa students, there is no guidance in any policy where students are informed either that data will be harvested and used, or which data will be harvested. The **Interception and Surveillance Policy** states that Unisa may monitor or track, although the policy applies specifically to the monitoring and surveillance of employees. The **Data Privacy Policy** defines students as consumers, that is, "any natural person who enters ... into an electronic transaction with a supplier as the end user of the goods or services offered by that supplier; e.g. students" [p.1]. The purpose of the **Data Privacy Policy** is to "protect the privacy of privacy subjects; provide guidelines for the collection, use, disclosure and maintenance of personal information by Unisa; limit Unisa's possible liability for privacy infringement; and educate users on privacy and related rights" (p. 2).

This policy also stresses that "Users may only collect personal information on privacy subjects if such information is necessary for *business purposes of Unisa*, or when the privacy subject has given permission that his/her personal information be collected by or on behalf of the Unisa" (p. 2). This implies that the collection, analysis and use of student data is within the legal parameters of the "business purposes" and is therefore legitimate. The collection of information to inform the business of Unisa is confirmed by the **Guidelines for conducting research involving Unisa staff, students or data** [11] which states that while the purpose of this document is "to provide guidelines for acquiring permission to do

research that involves Unisa staff, students and/or data" (p.1), these guidelines "do not apply to [the Department of Institutional Statistics and Analysis] DISA research approved by the Vice Principal: Research and Innovation, or to duly authorised routine data gathering activity, which is necessary for the efficient administration and operation of Unisa" (p. 2).

None of the existing policies therefore mandates the university to explicitly inform students that their behaviour may be monitored or surveilled, or provides students with the opportunity to opt out of these actions. Students may opt out as objects of research, and the **Policy on Research Ethics** [13] makes informed consent and anonymity non-negotiable (unless the latter is waived by the participant him or herself). **The Students' Charter on Rights and Responsibilities** [14] makes no mention of data privacy, access to own personal data, or to ensuring the correctness and currency of personal data (whether from the perspective of the institution or as a specific responsibility of students).

4.2.2 Open University

Within the **Data Protection policy,** students are informed that "some information, including the information you give us about your ethnic background or a disability, may be used by the University to identify students who require additional support or specific services. We consider disclosure of this information as explicit consent to use this information for this purpose" (p. 1).

The OU **Terms and conditions governing the use of software, tools and content** document sets out that material produced and uploaded to the OU LMS "may be used by The Open University on an irrevocable and perpetual basis and may be incorporated into module material and other content" (p.3).

4.3 Vulnerability and Harm

Vulnerability and harm are defined as implicit or explicit discrimination (whereby a student receives, or does not receive, support based on what might be considered to be a random personal characteristic), the consequences of labelling (on student identity and behaviours) and the validity of regarding student groups based on assumptions made about shared characteristics.

4.3.1 Unisa

If we consider that issues of privacy are directly linked to notions of vulnerability and harm, there are several policies and guidelines dealing with the protection of data and the prevention of harm. The **Unisa Information Security Policy** prescribes a three-tier classification system for information: namely confidential, internal and public use. The purpose of the policy is to "protect Unisa's corporate data and information and any client, employee or student information within its custody or safekeeping by safeguarding its confidentiality, integrity and availability" (p. 4).

Personal information is very broadly defined in the **Data Privacy Policy** as encompassing a variety of data including (but not limited to) information relating to the race, gender, sex, pregnancy, marital status, national, ethnic or social origin, colour, sexual orientation, age, physical or mental health, disability, religion, culture, language and birth of the individual; information relating to the education or the medical, criminal or employment history of the individual or information relating to financial transactions involving the individual; and correspondence sent by the individual that is implicitly or explicitly of a private or confidential nature or further correspondence that would reveal the contents of the original correspondence. Much of the data above might be used within a learning analytics profiling model. This then raises the interesting ethical issue regarding the use of such data to personalise or customise the learning experience without the explicit consent or refusal of consent by students.

Interestingly, the correctness of data provided to students by the institution and vice versa can play a huge role in the scope and permanence of vulnerability and harm. The Unisa **Students' Charter on Rights and Responsibilities** [14] is silent on matters of data privacy, and on the student's responsibility to ensure that data provided to the institution are correct. The **Students' Disciplinary Code** [15] includes a statement under the description of 'misconduct' concerning the provision of "materially false information *about the University*" (p. 4), but excludes provision of materially false or incorrect personal data *to* the institution.

4.3.2 Open University

The OU's **Information, Advice and Guidance Policy** (IAG) highlights OU objectives to empower students to achieve their study goals and to develop independence in their decision-making by "providing timely and targeted IAG to students at key points along the student journey that recognises and is responsive to diverse and distinct need" by "ensuring online information and advice is personalised, accessible, accurate, up to date and applies innovative technology". It aims to provide a service which "respects the needs of the individual student and is in their best interests". The policy does not discuss whether the support is in the best interests of the individual student or the wider cohort to which a student belongs, nor who makes decisions regarding best interests. The lack of clarity on this point suggests the potential for a dual role for the student which appears currently absent in practice. Nor is it clear whether there are further guidelines which determine the point at what tailored IAG becomes untenable in terms of available (and affordable) resource.

The IAG policy explicitly recognises the diversity of student backgrounds and educational experience, and flags that the service delivered will be targeted to "the specific needs of enquirers and students at different stages of their student journey". In this way, there is a tentative attempt at least to be transparent about differential levels of service, about assumptions made about student groupings which may relate to their shared characteristics, and to recognise that these profiles may change over time.

Students are informed, when they first register that they must notify the University within a reasonable time if they change their personal details. When informing the University of a disability which might affect their studies, students must provide further evidence as required. Similarly, if any module requires the student to meet specific conditions, there is a responsibility to inform the University should those conditions no longer be met. The OU's **Fraud Response Policy** makes it clear that any student intentionally and dishonestly making a false representation or dishonestly failing to disclose information is considered to have committed fraud for gain (where gain is assumed to extend beyond the purely financial). The **Code of Practice for Student Discipline** also includes as unacceptable conduct knowingly making a false statement or fraudulently providing information (at registration or when asking for a particular service).

4.4 Collection, Analyses, Access to and Storage of Data

This aspect relates to information about sites (both inside and outside of the LMS) used to gather information and to give informed consent regarding the scope and rights of the institution to harvest, analyse and use data from such sources. Students should be told which information is integral to official

institutional business, which information may be harvested with or without consent, and how they and others can help to ensure the correctness, currency and appropriateness of data. Students should be informed about uses of their data at registration.

4.4.1 Unisa

This consideration is covered extensively by the policy framework at Unisa. The collection, analyses, access to and storage of *student* data is not explicitly mentioned but implied within the broader guidelines on data privacy. For example, the **Data Privacy Policy** focuses explicitly on the institution's responsibility to safeguard staff and operational data rather than a duty for *students* to respect the data of fellow students or staff. The **Information Security Policy** proposes a three-tier classification system for information: namely confidential, internal and public use. The purpose of the policy is to "protect Unisa's corporate data and information and any client, employee or student information within its custody or safekeeping by safeguarding its confidentiality, integrity and availability" (p.4). The same principles are addressed in the **Information Sensitivity Classification Policy** [16] which allows that "a single lapse in information security can have significant long-term consequences" and that "Unisa unduly risks loss of student relationships, loss of public confidence, internal operational disruption, excessive costs and competitive disadvantage" (p.1). This policy states also that its intention is to consistently protect confidential information regardless of its form, the technology used to process it, who handles it, its location, and the stage in its information lifecycle. Under "Access control" the principle of "need to know" is established, meaning "that information must not be disclosed to any person who does not have a legitimate business need for the information" (p.2).

The **Records Management Policy** [17] defines "record" as "recorded information, regardless of format or medium, which has been created, received, used, accessed and maintained by Unisa (and/or predecessors) as evidence and information in pursuance of its legal obligations *or in the transaction of business.* Included are e-mail, electronic records and records other than correspondence" (p.1). Access to records is governed by "the 'sensitivity classifications' allocated to record series and detailed in the **Information Sensitivity Classification Policy.**" [16] Access to records by employees or third parties is dealt with in accordance.

4.4.2 Open University

The OU has a clear **Data Protection Policy** establishing that student records are created and maintained over significant periods of time. The student record includes data collected at registration and throughout the student journey. The data controller is clearly given as The Open University, although it is acknowledged that external service providers may process personal information under strict contractual confidentiality obligations. Students are informed that personal information is used to:

- "process applications;
- provide services (including providing certain online facilities and/or services and sending information about current and future study opportunities with the University);
- conduct research to help plan and improve university services;
- produce statistical information for publication;
- provide information about students to others, in line with legal and government requirements. The OU will transfer personal information outside of the European Economic Area only when necessary safeguards have been secured by contract.
- allow others to provide services to students and alumni"

This policy states that data may be transferred within the University on a "need-to-know" basis to facilitate the provision of academic and other services to students. There is no clear guidance given regarding these academic or other services, nor any explanation of what may define a "need to know" basis.

Within the **Terms and conditions governing the use of software, tools and content** document, students are advised to "check the terms and conditions and privacy policy of any other (external) website you visit". Within the University LMS, students are informed via the Data Protection policy that "Cookies are used so that we can easily recognise you when you return to our websites and, as a result, will enable us to provide you with a better service. We may also track user traffic patterns in order to determine the effectiveness of our website. Information obtained from these cookies will not be used for marketing purposes or released to third parties."(p.4). Students may opt not to receive cookies while browsing the University's website, but would not then be able to access password-protected sites.

The **Open University Data Protection Policy** states that "Information is protected from unauthorised access and we are confident no one will be able to access your personal information unlawfully." Any personal information transmitted from a student's browser to the OU web service, or from the service to the student's browser, is encrypted (as long as the student's web browser supports the Secure Sockets Layer (SSL)). Students are explicitly warned that internet email is not always secure and that they take responsibility for information. Students are also made aware that they may access personalised stored data, as well as an overview of those stakeholders granted access to specific datasets.

The Freedom of Information Code of Practice limits public access to information to non-personal recorded information. However, the student's rights to privacy under the Data Protection Act 1998 outweigh the rights of other individuals to access their information under the Freedom of Information Act 2000 except under certain public interest exemptions.

The **Retention of Student Data and Records Policy** sets out the conditions under which student data is maintained and covers all student data (relating to an identifiable individual), information, records and content relating to university business created by university staff or students. The UK's Data Protection Act 1998 requires that student records are retained only as long as is necessary, and should be accurate and up-to-date. As a student can continue to study modules for many years, the deletion of certain information after a set time with a requirement for the student to re-submit up-to-date information ensures compliance with this principle. The policy states both that "There is an expectation by students, employers and Government agencies … that Universities should retain a permanent core record of student names, the modules and qualifications studied and their outcomes" (p.3), and that "there are records and data which need to be retained whilst a student might continue to study with the OU."

On completion of their qualification, students become OU alumni. The policy on **the Retention of Alumni Data & Records** requires that alumni contact details continue to be held in student record systems to ensure a single instance of accurate information.

Students are also advised under the guidance document **Using Social Networking Tools** that they should take care to avoid activity that infringes another person's privacy (e.g. by posting their contact details without permission).

5. SUMMARY OVERVIEW AND DISCUSSION

Having reviewed the policies of both Unisa and the OU against the proposed considerations [8], it seems clear that the institutions' current policy frameworks largely focus on academic analytics and research with an emphasis on data governance, data security and privacy issues. Since learning analytics is more concerned with learning data at course and departmental level [1], both institutions' policy frameworks appear to lack explicit guidance for the questions, issues and ethical challenges to institutionalise learning analytics.

While both institutions have classification systems for categorising information, neither makes a distinction between the different layers of information harvested from students in terms of consent and opportunities to opt-out. It is accepted that there are certain types of information and analyses (e.g. cohort analyses) that fall within the legitimate scope of business of higher education. There is though an urgent need to approach personal data differently when it is used to categorise learners as at-risk, in need of special support or on different learning trajectories. Currently both institutions employ relatively crude and incomplete data sets to customise learning and support. Both institutions have ample policies and guidelines to protect data and to ensure that data is governed according to national and international legislation.

It is clear from the existing policy frameworks of both that the definition and scope, harvesting and analyses of data is an imbalanced and non-transparent affair.

6. CONCLUSION

Educational data mining is established practice in higher education and the increasing digitisation of education, technological advances, the changing nature and availability of data have huge potential for learning analytics to contribute to our understanding of the different variables impacting on the effectiveness of learning, student success and retention.

Most higher education institutions have existing policy frameworks in response to (inter)national legislative contexts to regulate and govern intellectual property, safeguard data privacy, and regulate access to data. These policy frameworks may not always be sufficient to address the specific ethical challenges in the harvest and analysis of big data in learning analytics.

Approaching learning analytics from a socio-critical perspective [8] suggests gaps in the policy frameworks of two large distance education institutions: the OU and Unisa. Both these institutions' policy frameworks offer extensive protection and regulation with regard to data privacy and protection.

This brief review of both institutions' policy frameworks highlights the irregularity of learning analytics where the institution is the only role-player with decision-making power, determining the scope, definition and use of educational data without the input of other stakeholders.

This research indicates that some higher education institutions' policy frameworks may no longer be sufficient to address the ethical issues in realising the potential of learning analytics.

7. ACKNOWLEDGEMENTS

The authors would like to acknowledge the OU and Unisa for their support, and Aimee Rhead for her preparation work.

8. REFERENCES

[1] Long, P. & Siemens, G. 2011. Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review* September/October, 31-40.

[2] Van Barneveld, A., Arnold, K.E., & Campbell, J.P. 2012. *Analytics in higher education: establishing a common language*. ELI paper 1: 2012. EDUCAUSE. Retrieved from http://net.educause.edu/ir/library/pdf/ELI3026.pdf

[3] Booth, M. 2012. Learning analytics: the new black. *EDUCAUSE Review*, July/August, 52-53.

[4] Johnson, L., Adams, S., & Cummins, M. 2012. *The NMC Horizon Report: 2012* Higher education Edition. Austin, Texas: The New Media Consortium.

[5] Oblinger, D.G. 2012. Let's talk analytics. *EDUCAUSE Review*, July/August, 10-13.

[6] Siemens, G. 2011. Learning analytics: envisioning a research discipline and a domain of practice. Paper presented at LAK12, Vancouver. Retrieved from http://learninganalytics.net/LAK_12_keynote_Siemens.pdf

[7] Subotzky, S., & Prinsloo, P. 2011. Turning the tide: a socio-critical model and framework for improving student success in open distance learning at the University of South Africa. *Distance Education*, 32(2), 177-193.

[8] Slade, S., & Prinsloo, P. 2013. *Learning Analytics: Ethical Issues and Dilemmas.* Manuscript accepted for publication, American Behavioral Scientist.

[9] Hsiu-Fang Hsieh, H-F., & Shannon, S.E. 2005. Three approaches to qualitative content analysis. *Qualitative Health Research, 15,* 1277-1288. DOI: 10.1177/1049732305276687

[10] Kruse, A., & Pongsajapan, R. 2012. Student-centered learning analytics. Retrieved from https://cndls.georgetown.edu/m/documents/thoughtpaper-krusepongsajapan.pdf

[11] Unisa. 2012. Guidelines for conducting research involving Unisa staff, students or data. http://heda.unisa.ac.za/filearchive/Guidelines_Research%20Involving_UNISAstaff.pdf

[12] Open University Essential documents for students http://www8.open.ac.uk/students/essential-documents/

[13] Unisa. 2012. Policy on research ethics. http://cm.unisa.ac.za/contents/departments/res_policies/docs/ResearchEthicsPolicy_apprvCounc_21Sept07.pdf

[14] Unisa. 2007. Students' charter on rights and responsibilities. http://cm.unisa.ac.za/contents/departments/studentaff_policies/docs/StudentCharter_apprvCounc_30Nov07.pdf

[15] Unisa. 2007. Students' disciplinary code. https://my.unisa.ac.za/tool/a87dd927-a9e0-4b59-0012-5ab7d72ca660/contents/courses/docs/StudentDisciplinaryCode_apprvCounc_26Jan08.pdf

[16] Unisa. 2007. Information sensitivity classification policy. http://cm.unisa.ac.za/contents/departments/corp_policies/docs/InfoSensitivClassificAnnexA_apprvCounc_21Sept07.pdf

[17] Unisa. 2007. Records management policy. http://cm.unisa.ac.za/contents/departments/corp_policies/docs/RecordsManagementPolicy_ManCom_15May07.pdf