



## Open Research Online

---

The Open University's repository of research publications and other research outputs

# Ethical oversight of student data in learning analytics: a typology derived from a cross-continental, cross-institutional perspective

## Journal Article

How to cite:

Willis, James; Slade, Sharon and Prinsloo, Paul (2016). Ethical oversight of student data in learning analytics: a typology derived from a cross-continental, cross-institutional perspective. Educational Technology Research and Development (Early Access).

For guidance on citations see [FAQs](#).

© 2016 Association for Educational Communications and Technology

Version: Accepted Manuscript

Link(s) to article on publisher's website:

<http://dx.doi.org/doi:10.1007/s11423-016-9463-4>

---

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

---

[oro.open.ac.uk](http://oro.open.ac.uk)

# Ethical Oversight of Student Data in Learning Analytics: A Typology Derived from a Cross-Continental, Cross-Institutional Perspective

James E. Willis, III\*  
Indiana University  
Jameswillis3@gmail.com

Sharon Slade  
The Open University  
Sharon.Slade@open.ac.uk

Paul Prinsloo  
University of South Africa  
prinssp@unisa.ac.za

## Abstract:

The growth of learning analytics as a means to improve student learning outcomes means that student data is being collected, analyzed, and applied in previously unforeseen ways. As the use of this data continues to shape academic and support interventions, there is increasing need for ethical reflection on *operational* approvals for learning analytics research. Though there are clear processes for vetting studies resulting in publication of student-gathered data, there is little comparable oversight of *internally* generated student-focused research. Increasingly, ethical concerns about the collection and harvesting of student data have been raised, but there is no clear indication how to address or oversee these ethical concerns. In addition, staff members who are not *typical* researchers may be less familiar with approvals processes and the need to demonstrate potential for harm, etc. If current trends point to a range of individuals harvesting and analyzing student data (mostly without students' informed consent or knowledge), how can the real danger of unethical behavior be curbed to mitigate the risk of unintended consequences? A systematic appraisal of the policy frameworks and processes of ethical review at three research institutions (namely, the University of South Africa, the Open University in the United Kingdom, and Indiana University in the United States) provides an opportunity to compare practices, values, and priorities. From this cross-institutional review, a working typology of ethical approaches is suggested within the scope of determining the moral intersection of internal student data usage and application.

**Keywords:** Ethics, learning analytics, application, ethical typology, research protections, review boards

**Conflict of Interest:** The authors declare that they have no conflict of interest.

## Background

Research efforts to improve the effectiveness of teaching and learning have been well-established within modern higher education since the 1960s (Brumbaugh 1960; Cross 1967). Despite various conceptual models and a rich legacy of institutional research, schools are not much closer to understanding student success and retention, nor closer to stemming the flow of

student attrition (Subotzky and Prinsloo 2011; Tinto 2006, 2012). To a large extent, graduate student retention and attrition remains a “black box” (Ehrenberg, Jakubson, Groen, So, and Price 2007), and yet there is a seemingly never-ending need for more data, analysis, and prediction in higher education (Howard, McLaughlin, and Knight 2012; Schildkamp, Lay, and Earl 2013).

In the nexus between the need for more effective teaching and learning and the changing context of higher education, the praxis of learning analytics emerges (Dawson, Gašević, Siemens, and Joksimović 2014). Learning analytics has been defined as “the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs” (Siemens, 2012, p. 4). As such it can include “techniques such as predictive modeling, building learner profiles, personalized and adaptive learning, optimizing learner success, early interventions, social network analysis, concept analysis, and sentiment analysis” (Siemens, 2012, p. 4). (See Siemens and Long, 2011, for a discussion on the difference between academic and learning analytics). Underpinning the various procedural applications of learning analytics is the belief that “unprecedented amounts of digital data about learners’ activities and interests” will allow institutions “to make better use of this data to improve learning outcomes” (Buckingham Shum and Ferguson, 2012). Amidst the hype and potential proffered by researchers and theorists, concerns about the underlying ethical implications of learning analytics increase in prominence (Macfadyen, Dawson, Pardo, and Gašević 2014). Ethical reflection in learning analytics serves a decidedly heterogeneous function because such inquiry necessitates the production of questions cutting across legal, behavioral, procedural, and social boundaries, exposing potential and actual concerns for student, institutional, and faculty roles in the production and use of learning data.

It seems that emerging practices of learning analytics are outpacing institutional regulatory and policy frameworks' ability to provide sufficient guidance and protection to students, who are often seen as *data objects* and passive receivers of assumed benefits. There is also little indication of whether learning analytics development or practice qualifies as *research* and would, as a result, need to get equivalent ethical clearance from institutional review boards (IRBs).

In the context of ethical concerns regarding learning analytics, it is clear that there is an urgent need to establish clarity going forward.

### **Ethical Approaches to Student Data: A Review**

Ethical concerns have been addressed in varying degrees within learning analytics literature from the outset of a range of contemporary practices which use student-level data to predict outcomes, formulate interventions, and reconsider pedagogical, curricular, and learning approaches. Campbell, DeBlois, and Oblinger (2007) identified the *obligation to act* which follows from identifying individual student concerns in the data. Though the ethical implications of the obligations of administrators and instructors were not set out for some time (Willis, Campbell, and Pistilli 2013; Slade and Prinsloo 2013), growing concern emerged over the open disclosure of tracking mechanisms (Ferguson 2012), surveillance (Greller and Drachsler 2012), legal and ethical dimensions of using data (Kay, Korn, and Oppenheim 2012), and information privacy (Macfadyen and Dawson 2012). A motivating factor in the obligation to act is a “fiduciary duty” (Slade and Prinsloo 2013) due to the power imbalance between the student and the providing institution.

Other early concerns were highlighted as natural extensions of the logical outcomes of the research (de Freitas et al. 2014); these approaches were not systematic inquiries into the

potentially harmful effects of using student data. The ethics of opt-in/opt-out and consent (Prinsloo and Slade 2015; Hack 2015a), anonymity of data (de Freitas et al. 2014), and expressly utilitarian ethics in learning analytics (Willis 2014) later emerged. Kruse and Pongsajapan (2012) propose a move from learning analytics as applied *to* students *by* administrators and professors toward systems where students are equipped and enabled with their own data, though this approach also has potential ethical problems.

As learning analytics systems continue to grow in complexity, ethical approaches must continually reassess the risks of the examined data encroaching on student privacy (Yanqing, Guangming, One, and Woolley 2013). Within higher education institutions, ethical practices of review tend to be varied and not easily categorizable (Largent, Grady, Miller, and Wertheimer 2012; Locke, Ovando, and Montecinos, 2016). When activities have an external focus, for example, the publication of results relating to learning analytics, data practices may fall under the review of an institutional review board. However, *internal* practice and processes, such as those associated with operational or scholarship activities, are often murkier (Abbott and Grady 2011; Guta, Nixon, and Wilson 2013). Furthermore, the increasing complexities of aggregated data across platforms, modalities, and courses may actively transgress the boundaries of accepted prior research procedures and protocols (Kelly and Seppälä 2015).

The procedures of opting-in, opting-out, and determining qualification for exempt and expedited review research can vary according to local rules (Kelly and Seppälä 2015; Moxley 2013). For example, pedagogic review may fall under normal education practice whereas procedural review may ensure protection of all research subjects (Hack 2015b). Even though informed consent may stand central to ethical protection of *research* participants, the identification of individuals in online learning environments - where guest IDs, avatars, or screen

names may (or may not) correspond to individuals - can become increasingly difficult (Kelly and Seppälä 2015). Algorithmic decision-making and machine learning raise not only new ethical questions, but also questions regarding oversight and prevention of harm (Willis and Strunk 2015). The storage of online surveys and other online research-oriented activities present increasingly-novel approaches to gathering, analyzing, and disseminating data; such approaches likewise present challenges to IRB's (Buchanan and Ess 2009; Buchanan and Hvizdak 2009). One of the clearest long-term challenges is stewardship of student data (Willis and Strunk 2015).

The modern IRB stands as a third-party approval to ensure the ethical treatment of humans with proper protections in research design (Moskal 2016) and thus commands a central role in research involving humans. The 1979 Belmont Report raised greater societal awareness of research performed on people (Vitak, Stilton, and Ashktorab 2016). Anecdotal evidence suggests that while learning analytics *is* widely considered as research, many practitioners do not see the need to apply for ethical clearance. This is compounded by mission creep in the purview of IRBs which may contaminate and frustrate the research process, with a sense that IRBs often act without oversight (Bledsoe, Sherin, Galinsky, and Headley 2007; Carr 2015).

Learning analytics is dependent on data with much of it ambiently captured in the student cycle of matriculation to post-graduation. The use of this data for purposes of intervention helps shape protocols of learning analytics. Like other social sciences that apply data to intervention, learning analytics can shift attitudes of those receiving treatments, and thus change outcomes. IRBs are rightly positioned to oversee the ethics of research experiments to prevent malfeasance, abuse of participants, and potentially harmful interventionalism.

### **Research Design and Methodology**

This study's research design is a *qualitative* interpretative or hermeneutic multiple-case study (Bos and Tarnai 1999; Yin, 2009) which results in a typology (Kluge 2000; Knobelsdorf 2008). The multiple-case study involved a literature review and a content analysis of three institutions' approaches to ensuring ethical research. The units of analysis were the applicable policy and guiding framework documents.

We followed a *convenient* sampling methodology and selected the three institutions where we as researchers are based, providing us easy access to policy information. The selection of these three institutions on three different continents also enriched the development of a typology due to different levels of maturity in learning analytics practices, functioning of IRBs and the use of student data to inform teaching and learning.

Content analysis is an established methodology in quantitative and qualitative research designs (Bos and Tarnai 1999; Elo and Kyngäs 2007; Hsieh and Shannon 2005) and increasingly a feature in learning analytics (Kovanović, Joksimović, Gašević, Hatala, and Siemens 2015). Qualitative content analysis "is defined as a research method for the subjective interpretation of the content of text data through the systematic classification process of coding and identifying themes or patterns" (Hsieh and Shannon 2005, p.1278). The aim of content analysis here is "to attain a condensed and broad description of the phenomenon, and the outcome of the analysis is concepts or categories describing the phenomenon" resulting in a "model, conceptual system, conceptual map or categories" (Elo and Kyngäs 2007, p. 108).

Aligned with Thomas' (2011) *phronesis* approach to doing case studies, we adopted a dialogical model proposed by Rule and John (2011) in which theory and analytics constructs (resulting from the literature review) and research interact dialogically throughout the research process. "Such an approach acknowledges that theory infuses research in all its aspects,

including the identification and selection of the case, the formulation of research purposes and questions, the survey of literature, the collection and analysis of data, and the presentation and interpretation of findings” (p. 100). A dialogical approach further allowed for the three researchers to engage with a different institution’s policy frameworks, and to member-check the analysis and excursions.

Trustworthiness was ensured by following Rule and John’s (2011) proposal to present thick descriptions, verifying accounts between researchers, the creation of an audit trail and using critical peer checks. It is important to note that this multiple case study design does not provide a basis for generalisation to other contexts but “are generalisable to theoretical propositions” (Yin, 2009, p. 15).

The literature on case studies (e.g. Rule and John 2011; Thomas 2011; Yin 2009) and content analysis (e.g. Bos and Tarnai, 1999; Downe-Wamboldt, 1992; Elo and Kyngäs, 2007; Hsieh and Shannon, 2005) are relatively quiet on the specific ethical issues involved in case study research and content analysis as methodology where the analyses involve only *document* analysis. Scott (2005), for example state that the general principles of social research applies to the analysis of documents.

In this study, a *type* consists of “a set of characteristics that are interrelated and logically connected in regards to content” (Knobelsdorf 2008, par. 24). It is furthermore important to distinguish between classifications which are mutually exclusive and exhaustive, and types which “combine characteristics that are not uniquely and exclusively allocated to it” (Knobelsdorf 2008, par. 26). Because of overlaps in research contexts, these types do not necessarily represent reality but instead provides a heuristic to assist with making sense of complex social phenomena.

Typologies are developed through grouping processes where different elements are coded or identified and clustered according to shared characteristics (internal heterogeneity) and important differences between these groupings/clusters (external heterogeneity) (Kluge 2000). In following Kluge (2000) and Knobelsdorf (2008), this study broadly adopted a four-stage model of empirically-based typification entailing:

1. Developing the relevant dimensions of comparison - established during the analysis in a combination of theoretical knowledge and collected data. The dimensions of comparison resulted from a deductive, directed content analysis approach that entailed identifying key concepts in theoretical frameworks and published research (Elo and Kyngäs, 2007; Hsieh and Shannon, 2005). This informed the second stage of the typification.
2. Empirically grouping cases and regularities. Due to the lack of research on the ethical clearance involved in learning analytics, two main distinguishing features were established - namely, learning analytics as research, and learning analytics as an emerging form of specific research. The combination of the literature review with a content analysis of the research policies at three different institutions revealed other possible types - initially clustered as *learning analytics as something else*. The *deductive*, directed content analysis approach not only analyzed the manifest content *but also latent* content looking for the silences on pertinent issues in the analyzed texts.
3. Analysis of coherence, meaningful relationships and typification.
4. Characterization of types involving a detailed description illustrating unique attributes and overlapping attributes with other types.

As Bowker and Star (1999) suggest, classifications are not neutral or objective and so decisions which result in a classification may remain invisible. By developing a typology for understanding the ethical approvals needed at the intersection between research and learning

analytics, the process becomes more transparent. The application of a typological approach to learning analytics follows directly from recent research exploring typologies in information privacy contexts (Koops, Newell, Timan, Škorvánek, Chokrevski, and Galič forthcoming) and complementary frameworks in the field of learning analytics (Colvin et al. 2016).

### **An Analysis of Existing Institutional Practice**

Given the established position of IRBs within higher education institutions, it is helpful to compare the policy frameworks of ethical review across three research institutions on three continents (namely, Indiana University in the United States, the Open University in the United Kingdom, and the University of South Africa). In the presentation of the analysis, we enter into dialogue with the institution in question's policies and have used the notion of an "excursus" to act as signpost indicating a short digression in the analysis. This allowed us to comment on or respond to specific issues in the analysis without interrupting the line of argumentation in the original document. The excursus therefore "forms a separate parenthetical segment in the discourse" (Redeker, 2000, p. 14). (See for example Bauman, 2004). It is important to note that we use "excursus" as operator in our *presentation* of the analysis of the policy content, and not as methodology of analysis.

#### **Indiana University**

Indiana University (IU) ensures its rigorous research environment is supported by oversight of the IU Human Subjects Office (HSO) which includes the Institutional Review Board. IU requires IRB approval of all human subject studies carried out by anyone affiliated with the university or its affiliates. IU's online portal (2015) includes information about the processes of review, flow charts detailing the levels of review, and links to submit an IRB application. IU requires Collaborative Institutional Training Initiative (CITI) training for

researchers engaged in human research; this training includes ethics training based on the 1979 Belmont Report, historical context for human research, and practices observed by individual institutions.

**Excursus 1.** *Much has changed in regards to human-technology interaction since the 1979 Belmont Report; it is unlikely that an updated consensus document will be developed soon, so it becomes the responsibility of researchers to disclose and practice moral decision-making. As algorithms increasingly impact human learning, and as they lead to further automation in adaptive cycles of deep learning, the line between organic life and algorithmically-altered life blurs. When such algorithms adapt to human change, the original human coder disappears and becomes untraceable.*

The first step of the application process is to determine the level of review. Research involving humans can have various design methodologies which may determine if the review process can be *exempt* (involving normal educational practice not involving any sort of duress or deprivation to resources), *expedited* (involving research that involves identifiable data), or *full board review*.

**Excursus 2.** *The research team may incur additional work if an application is submitted for exemption but is directed to undergo expedited review. When using experimental methodologies, it may be unclear at the outset how the research may change - and whether changes warrant an amended disclosure to the IRB. A deeper question for the IRB's oversight of the processes is the downstream use of data and how its interpretive contexts can change. While the IRB application can include what happens to datasets, is it possible to fully disclose how generative data insights in an open and recursive educational system are handled at some point in the future?*

The second step includes disclosure of each researcher and his/her role, research question(s) asked, methodologies, procedures for collecting and storing data, interview or survey protocols, duration of the study, and other applicable documentation to ensure minimal risk, and consent of participants. Studies involving implied consent (i.e., where documentation is impractical to gather) require clear reasoning for justification. The third step involves submission to the IRB, where the application is assigned to an analyst.

**Excursus 3.** *IRB analysts are broadly-trained to understand various quantitative and qualitative methodologies utilized in human subjects research. However, as learning analytics technologies and methodologies change, it might be rightly asked if this remains the case: Is it possible to fully understand interactions with people in a way that facilitates high-impact research and protects human interests in immediate and downstream data contexts?*

After the fourth step of review, the analyst will either approve the study as it has been submitted or require additional clarification and documentation. In the final step of post-approval, which remains in effect for the duration of the study, the researcher must disclose additional personnel (if any), amendments for changes like altered surveys or consent forms, and notification to the IRB if there are unintended consequences for any research activities.

**Excursus 4.** *An IRB application must have an end date. While in active status, researchers must decide if current activities fall under the disclosed protocols of the approved application; researchers are responsible for disclosing any changes via an amendment. As learning analytics research continues to proliferate in terms of pervasive surveillance of student activities, disclosure of data obtained by researchers may continue to be problematic. For example, if, during the normal course of research, it is uncovered that data indicate a potential harm, is there proper oversight to compel researchers to alert the IRB of possible problems?*

The HSO defines research as being both “systematic” and “contributing to generalizable knowledge” (Indiana University 2015) which means case studies involving individuals do not fall under the need for oversight. The requirement of generalizable information is important because it establishes a difference between publicly-accessible information versus non-published internal studies. Likewise, the HSO’s threshold for human research includes interacting or intervening with individuals or the use of individually-identifiable data.

In 2015, the HSO adopted flexibility options for human research meant to streamline IRB overview practices; the criteria to be deemed *flexed* include research with minimal risk to its participants and not federally funded. For example, study of common educational practice, as long as it is not federally supported, would generally be eligible for flexibility. Extending flexibility in some circumstances may add a layer of complexity in determining how researchers apply to the IRB. The HSO ensures that scenarios are available to help researchers determine classification and oversight level of a study.

**Excursus 5.** *With potential to streamline research oversight, does the flexibility option provide enough review to normal educational practice studies, yet enable methodological practices which may uncover patterns that would otherwise need additional ethical review? One of the main requirements of the flexibility option is lack of federal funding; this sends a strong signal to researchers that ethical oversight may be financially-driven or for the purpose of preventing potential litigation.*

### **The Open University**

The Open University’s (OU) Code of Practice for Research (2014) has a broad scope and cross references several other policies and codes. It applies to *anyone* (whether staff, student, or any other individual) conducting research in the University’s name. The code sets out the

specific responsibilities for unit heads to ensure that all researchers have adequate opportunities for relevant training.

**Excursus 1.** *Although this includes the responsibility to provide access to training, there is no clear responsibility to ensure that staff actually undertake training or demonstrate sufficient understanding of the issues. The onus is on the researchers to maintain the best practice both for themselves and for any other individual on their team.*

Except for “no risk” research, research involving human participants (and/or their data) is subject to approvals from the Human Research Ethics Committee (HREC). In addition, it must comply with the requirements of the Data Protection Act (and details must be registered with the University data protection officer). Data used may be subject to the Freedom of Information Act. A checklist of risk is submitted to support the determination of the level of ethics review required.

**Excursus 2.** *The extent of risk, then, may be considered subjective in part and is determined by an individual who may not be able to judge potential future risk such as that levied by learning analytics (e.g., the prevention of student progression or the withholding of support offered to other students on the basis of their data or predicted behaviors).*

Valid informed consent under HREC requires that potential participants should always be informed in advance, and in terms which are understandable, of any potential benefits, risks, inconvenience, or obligations that might reasonably be expected to influence their willingness to participate.

**Excursus 3.** *The definitions of risk which determine formal oversight are wide-ranging and have emerged from traditional trial-based research. In learning analytics, risk might relate to participants’ ability to give informed consent for inclusion in an ongoing study of their*

*(online) behaviors. Understanding the range of possible outcomes which might emerge from the research may directly impact their study options.*

Consent should be gained consistently, by outlining the purpose and participatory involvement guidelines and must result in a signed consent form. If longer-term data collection is planned, the means of obtaining renewal of consent at appropriate times must be considered. Participants may withdraw their consent at any time and expect to have any provided data destroyed (up to a specified date).

**Excursus 4.** *This inclusion of the need to obtain advance consent creates a potential practical conflict with research conducted on external online sites, which may not always be feasible.*

Research outputs and data must be recorded in a durable, secure, and retrievable form, be appropriately indexed, and comply with any relevant protocols. Retention and archiving of data must comply with external requirements and the terms which ethical approval was granted. Researchers must also comply with obligations to funders. Information on the source of financial support for the research must be transparent and, if published, research outputs should include disclosure of any potential conflicts of interest.

The University is committed to the UK Research Council (RCUK) policy on access to research outputs and believes that the ideas from publicly-funded research should be made accessible for public use and scrutiny, as rapidly and effectively as possible. Data forming the basis of publications must be available for discussion with other researchers.

The assessment of risk and the processes associated with HREC approvals are consistent with UK practice. However, research involving students (including the collection of information from enquirers, students, or alumni) and what may be considered operational practice, such as

the purposes of learning analytics, does *not always* require HREC approvals. Issues of risk are muddied by the less immediate timeframe and the difficulties in assessing future outcomes for those involved. Instead, separate institutional (rather than ethical) approval from a Student Research Project Panel (SRPP) is required. This approval is less stringent: the research methodology should be defined clearly, explained well, and have a good chance of producing the results needed to answer research question/s.

**Excursus 5.** *Unlike the academic research covered by HREC approvals, there is no defined need for staff who engage with student data to complete relevant training. There are fewer formal requirements for student support staff, course administrators, and lecturers who might typically engage in new learning analytics approaches to have relevant skills needed for data manipulation and interpretation.*

For survey-based research involving student data, consent (signed or electronic) must be recorded, although students may also have opted out of being contacted in advance. The invitation to participate should explain why they are being asked, what the research is about, and how the results will be used.

**Excursus 6.** *Research intended to inform learning analytics practice may not necessarily have outcomes which may be clearly defined in advance. Exploration of patterns of study or student success, for example, may lead to interventions which had not been foreseen at the time of seeking initial consent. There is potentially a gap between establishing clear boundaries and trust in terms of purpose and the resulting outcomes in terms of changed practice.*

The OU has fairly wide-ranging and well-defined procedures for managing student data for research purposes. However, there is no current *formal* approvals process for piloting internal projects which involve student data, such as exploratory learning analytics work. Also, there is

no formalized process for approving the transfer of an approach based on research (which may have received formal ethics or institutional approvals) into a *business as usual* approach. There is no recommended guidance to determine whether a particular methodology is robust. Although not consistent, the historical position has been one of assumed consent for both the implementation of new methodologies and consent from students.

The University is beginning to draft a procedure which would work alongside formal policy and approvals processes to be used as *guidance* for both scholarship work relating to learning analytics and for the transfer of learning analytics research into standard institutional practice. It is likely that the resulting process will not require formal approvals for the application of established methodologies to whole cohorts, but *may* require staff to seek SRPP approvals for piloting exploratory approaches which treat subsets of students in different ways, such as a/b testing and new predictive modelling approaches. It has been agreed that interventions arising from learning analytics must be part of the student operational record and systems may need amending in order to do this.

**Excursus 7.** *Whilst the purpose and description of interventions themselves may be recorded, it is not currently possible to provide recorded detail at an individual level of the information which triggered that intervention (in the case of predictive models).*

In 2014, the OU introduced a new policy relating to the ethical uses of student data for learning analytics. The policy is based on eight underlying principles. The development and release of the policy into the public domain has been supplemented by additional materials available to students to increase transparency. In terms of *established* learning analytics practice, the University is moving toward a position of informed consent, whereby students have access to disclosure of University data use, and by registering, have granted their consent.

**Excursus 8.** *Informed consent effectively means established practice requires no formal approvals, although there is considerable guidance available to staff to ensure that the constraints of the approach and the dataset are well understood and ethical. Consideration is also being given to development of a comprehensive record of all models and uses of learning analytics. It is fair to say, though, many students will not engage with formal policy prior to or following registration and so will remain largely unaware of the uses of their data.*

### **University of South Africa**

The University of South Africa (Unisa) has two main policies which do not refer directly to learning analytics but deal with student data, its analysis, and use. The policies are the Policy on Research Ethics (Unisa 2012) and the Data Privacy Policy of Unisa (2014).

The Policy on Research Ethics (2012) covers the established principles for ethical research by referring to the need to warrant integrity, accountability, and rigor and a commitment espousing “the constitutional values of human dignity, equality, social justice and fairness,” and promoting “the internationally recognised moral principles of autonomy, beneficence, non-maleficence and justice” (p.1).

**Excursus 1.** *To ensure ethical adherence, the Policy applies only to approved research defined as the “systematic investigation aimed at the development of, or contribution to, knowledge” (p.3) which would include the collection and analysis of student data. The Policy clarifies that the researcher must have ethical clearance. Where does this leave educators or support staff who collect and analyze student data? What protection do participants have in research that has not been approved?*

In addition, the Policy states, “The publishing of research findings should...not harm research participants” (p.5). While the Policy is concerned about the possible harm triggered by

publication of findings, it does not directly address the considerations of the potential harm through operationalizing analyses.

**Excursus 2.** *While it is clear that participants should not be harmed, researchers must take full responsibility and be held accountable for “all aspects and consequences of their research activities” (p.5). This raises four issues: establishing that harm was caused by the implementation of research findings; determining whether changes in practice driven by learning analytics are well understood; limiting harm by the protection of rights and interests, and the privacy and dignity of research participants, especially those who are vulnerable due to ignorance and powerlessness; regarding learning analytics as research, students may be especially vulnerable.*

In the context of “data sharing,” the Policy states, “As far as possible... relevant findings of the research are taken back to the research participants, institutions or communities in a form and manner that they can understand” (p.8).

**Excursus 3.** *Do students, by virtue of registration and the acceptance of Terms and Conditions, provide tacit consent for their personal data to be collected and analyzed?*

The Policy focuses specifically on research involving human participants, and commits research at Unisa to four moral principles – autonomy (the right to withdraw or opt-out), beneficence, non-maleficence, and justice. The Policy is clear - “Autonomy requires that individuals’ participation should be freely given, based on informed consent and for a specific purpose.” The Policy section on competence, ability, and commitment to research is clear - “Researchers should be both personally and/or professionally qualified for the research that they undertake” (p.10)... and that “Researchers should be honest about their own limitations, competence, belief systems, values and needs” (p.11).

**Excursus 4.** *Do those engaged in learning analytics have the conceptual and theoretical understanding of student success as the result of mostly non-linear, multidimensional, interdependent interactions at different phases in the nexus between student, institution and broader societal factors (Subotzky and Prinsloo 2011)?*

Under the section “Risk minimisation” (p.11) the Policy states, “Researchers should ensure that the actual benefits to be derived by the participants ... clearly outweigh any possible risks, and that participants are subjected only to those risks that are clearly necessary.”

**Excursus 5.** *In light of the fiduciary duty of higher education to support students and to ensure effective teaching and learning strategies and access to student support (Prinsloo and Slade 2014), risk minimisation provides the best rationale for learning analytics as moral practice (Slade and Prinsloo 2013).*

The Policy makes clear that participants are “indispensable and worthy partners in research,” and that the selection of participants should be fair, the “social, cultural and historical background of participants should be taken into consideration in the planning and conduct of research” and research should not “infringe the autonomy of participants by resorting to ... *the promise of unrealistic benefits*” (p.12).

**Excursus 6.** *When students are regarded as partners rather than data objects, what are the implications, unless learning analytics falls outside the scope of traditional research? In experimental design or where benefits or personalization are applied selectively and without student knowledge, how can students be assured fairness? Given that students’ digital data offer potentially flawed proxies, how can the complexity of the students’ backgrounds be considered? When students are invited to share data with the unrealistic promise of benefits, to what extent does this constitute undue influence?*

With regard to “informed consent” (p.13), the Policy makes it clear that “Researchers should respect their right at any stage to refuse to participate in particular aspects of the research or to decide to withdraw their previous given consent without demanding reasons or imposing penalties.” The Policy is emphatic regarding “participants' right to privacy, anonymity and confidentiality gains additional importance in such cases as they do not know the real purpose or objectives for which they are providing information” (p.14).

**Excursus 7.** *If learning analytics is regarded as research, then what is the response to informed consent? If it is uncertain what will be found and what the exact purpose of the research is, how does this impact students’ right to privacy, anonymity, and confidentiality?*

The Data Privacy Policy’s (2014) description of personal information includes, *inter alia*, “the personal opinions, views or preferences of 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” (p.2).

**Excursus 8.** *This suggests that students’ blog posts or responses on institutional learning management systems are included as well as personal communication between lecturers and students or even between students.*

Point 5.15 states, “Personal information of privacy subjects will not be processed *outside the purpose* it was collected for, *without the prior written consent* of the privacy subject involved.”

**Excursus 9.** *This has potentially dramatic implications for learning analytics. Does information provided by students (number of clicks and downloads, blog posts and discussion forums, etc.) fall inherently within the parameter of business information and so does Unisa have a right to harvest and analyze it?*

## **Summing Up: Common Issues**

Common amongst the ethics review processes at Indiana University, the Open University, and the University of South Africa is the core value that human research subjects should receive careful oversight to prevent harm. This applies, also, to learning analytics, where technology, educational intervention, and human participation converge in a dizzying combination.

Other common issues in the ethical review of learning analytics include the following: the unclear definition of “harm;” the possibility that some methodologies of intervention may surpass the understanding of the analysts responsible for approvals; a potential inability of researchers to pre-determine outcomes and potential harm or limitations to students; a difficulty in obtaining advance consent of participants; a lack of disclosure to students to enable them to make an informed judgement of participation in research or application of treatment; and a suggestion that learning analytics practice is not seen as research and is thus not subject to the rigor of IRB approvals.

### **Establishing a Common Framework**

In current iterations of learning analytics, as well as in educated hypotheses about where the processes could develop in the future, there are special challenges to those overseeing ethical review, especially as learning analytics begins to shape changes in pedagogic research (Regan, Baldwin, and Peters 2012). The challenges faced by IRB review might be unique because they combine direct human research and use of digital and physical educational technologies with understanding the nuances of specific methodologies that may pose potential problems. The influx of massive data sets coupled with new research techniques already present an array of potential difficulties assessing ethical approaches (Clark et al. 2015). These challenges may

undermine the spirit of ethical oversight by the IRB and expose the power asymmetries of the IRB in the following ways:

1. *Invasive techniques with built-in obfuscation and redirection.* Schools already have access to student emails which can be mined for qualitative purposes, digital search histories and app usage data, and students' online activities. It is possible that learning analytics systems can surreptitiously persuade students to engage in other activities.
2. *Surveillance.* Similar to invasive techniques, the surveilling of students' activities is possible in the physical world (like entry card swipes) and the digital world (like actively monitoring online activity). Unlike invasive techniques, surveillance might be further demarcated in the methodological approaches insofar as it entails the active watching of students' actions as they are related to learning. This may be justified in IRB applications as aggregated data (whereby individuals may not be identified) or as necessary components of providing customized support.
3. *Questionable intervention.* Practices associated with collecting and utilizing student data also create situations where administrators, faculty, and staff may intervene in students' learning. Ideally, such interventions can help students improve outcomes. The quantity, scope, and sensitivity of such data can be enormous, but how data are applied, to whom it is known, its storage location(s), and the potential for abuse should give pause to IRBs.
4. *Ideological power and biases.* Data interpretations are subject to biases. If learning analytics efforts are driven by efforts for student retention (i.e., tuition monies), the shape of the interpretative procedures is quite different than a project effort driven by a desire to provide students with detailed information to facilitate study choice. Questions about how the data are stored, in what format it is disclosed to parties (i.e., in rough data sets or dashboards), and the

methods employed to redact or seal particular data points can all present unique issues with ideological power and bias.

To address these challenges directly, a typological framework provides a dynamic tool for review boards, researchers, tool developers, and institutions alike and is heuristically conceived to guide parties through potential ethical implications. The purpose of the multi-institutional analysis coupled with a proposed typology is to support and encourage future use of internal student data rather than to constrain research. Theoretical constructs include concern for the moral and legal protection of student data, development of a typographic approach relevant for research design, planning, implementation, and measurement, and distinctions between the boundaries of application and research as independent schemas:

1. If we accept learning analytics as *research*, it will necessarily fall under the jurisdiction of the IRBs with associated challenges and concerns;
2. If we accept learning analytics as *an emerging specific form of research* needing oversight, we may also acknowledge that current processes or frameworks are not suitable;
3. If we accept learning analytics as *a practice which falls outside the traditional notion of research*, we suggest four possibilities:
  - Learning analytics as the *scholarship of teaching and learning*;
  - Learning analytics as *dynamic, synchronous, and asynchronous sense-making*;
  - Learning analytics as an *automated process*;
  - Learning analytics as a *participatory process and collaborative sense-making*.

For additional surveillance classifications, Kitchin (2013) describes three types of data surveillance (directed data: generated by digital forms of surveillance pre-determined by a person; automated data: collected online automatically; and volunteered data: gifted by users)

and Knox (2010) describes three different surveillance scopes (panoptic: visible and automated collection which can alter behavior; rhizomatic: multi-directional data collection with flattened hierarchies; and predictive: present and past data collection to extend temporal reach into the future).

Table 1 presents a view of the three perspectives of learning analytics (as research, as emerging research and as “other” practice) and outlines the characteristics of each, together with a description of existing approvals processes, relevant stakeholders and related outcomes. The typological view aims to more explicitly identify existing gaps and issues for future consideration.

Rather than conceiving of learning analytics as a specific set of technologies or innovations, *Learning Analytics As...* indicates a purposeful opening of extant and emerging descriptors for processes and procedures. This also encourages adaption to a wider array of constituencies like centers for teaching and learning, information technology, administration, and other bodies engaged in the processes of examining learning data. Subsequently, a *definition* is provided to further clarify how learning analytics are conceived as data processes. The definition helps distinguish the *Learning Analytics As* categorization from other descriptors. This distinguishing is furthermore made via the suggestion of a *surveillance type* modelled on Kitchin’s (2013) way of distinguishing how data is gathered, stored, examined, and used either by human agents or by automated processes (that are, presumably, still examined by human agents at points in the process). The suggestion of a surveillance type is an application of Kitchin’s (2013) approach and is meant to clarify the definition, not impose a static category on the type of learning analytics.

*Foundation Assumptions* are intended to interrogate potentially latent implications of data use and analysis. This functionally works with *and* against the definition of the type of learning analytics types for the express purpose of offering additional clarifications and challenging current ways data is analyzed. To further demonstrate this tension, we offer a surveillance scope elaborated by Knox (2010). In contrast to the type of surveillance elaborated by Kitchin (2013), surveillance scope is an attempt to describe the outcome of data monitoring and usage. These outcomes both contribute to the definition of the type of learning analytics processes, but also challenge how the data is used as particular outcomes of analysis.

Three additional categories, *approval*, *actors*, and *outcome*, help further delineate the extant and emerging processes of learning analytics as human processes, which contrast slightly with the aforementioned descriptors which are more data-intensive. The *approval* category indicates how institutions undertake obtaining official recognition and clearance for the research processes and how individuals assent or consent to research with their data. The *actor* category indicates who the research is done to and the possible effects of that research; in this instance, particular attention is paid to students as being arbiters of their own data insofar as they are objects of data and data analysis. Finally, the *outcome* category is an attempt to predict or suggest how the various processes can be realized in the “real world” of intervention, actionable data, and the publication of research methodologies and findings.

**Table 1: Towards a Typology of Different Ethical Approaches**

<b>LEARNING ANALYTICS AS</b>	<b>DEFINITION and SURVEILLANCE TYPE (Kitchin 2013)</b>	<b>FOUNDATIONAL ASSUMPTIONS and SURVEILLANCE SCOPE (Knox 2010)</b>	<b>APPROVAL</b>	<b>ACTORS</b>	<b>OUTCOME</b>
<i>Research</i>	<p>Defining, organizing, and utilizing extant data points to make sense of correlated points, predict outcomes, and/or intervene in the learning process.</p> <p><i>Surveillance type:</i> directed</p>	<p>Many extant data points exist as a matter of course.</p> <p>Data collected is lawful, protected by the institution, and requires approval to use for publication or external purposes.</p> <p>Generally, data released externally is aggregated and de-identified.</p> <p><i>Surveillance scope:</i> panoptic</p>	<p>Formal processes.</p> <p>Students must provide consent.</p>	<p>Everyone involved in the gathering, analysis, and interpretation of data, for whatever purpose.</p> <p>Students mostly only as data objects and not as collaborators or participants.</p>	<p>Primarily publications and reports.</p> <p>Mission creep.</p> <p>Frustration.</p> <p>Takes too long.</p> <p>Does not necessarily protect.</p>

<p><i>An emerging specific form of research</i></p>	<p>Same as <i>research</i>, but methods, character, and function(s) being developed may not correspond with existing ethical frameworks.</p> <p><i>Surveillance type:</i> directed, automated, and/or volunteered</p>	<p>Current oversight procedures control for static research design that allows for minimal iteration.</p> <p>Dynamic review needed to capture how research design and implementation may change.</p> <p><i>Surveillance scope:</i> panoptic, rhizomatic, and/or predictive</p>	<p>Undefined, unclear.</p> <p>From the literature, it is clear the processes and assumptions do not allow for this emerging form of research that is process and practice oriented.</p>	<p>Students seen as data objects and providers of information, though may be iteratively incorporated as research collaborators and/or participants.</p>	<p>Publications and learning intervention.</p> <p>Methodologies must account for changes, adaptations, and iteration.</p> <p>Protection can be problematic due to potentially unforeseen changes or complications in methodology.</p>
<p><i>Scholarship of teaching and learning</i></p>	<p>Part of the normal sense-making of everyone involved in the development of curricula, pedagogy, student support and assessment, mostly historical, post de facto.</p> <p><i>Surveillance type:</i> directed</p>	<p>The lecturer is main actor – not including other stakeholders in the process.</p> <p><i>Surveillance scope:</i> panoptic and/or predictive</p>	<p>Undefined, unclear. Consent is not normally requested.</p>	<p>Still sees students as passive data objects and providers of information. Evaluation of any changed approach can be complex.</p>	<p>Changes to teaching and learning design; proactive tracking of students to facilitate targeted interventions.</p>

<p><i>Dynamic, synchronous and asynchronous sense-making</i></p>	<p>Falls under the broader scope of scholarship of teaching and learning with an emphasis on both the instantaneous and historical collection and analysis of student data, and the timing of the interventions.</p> <p><i>Surveillance type:</i> directed and/or automated</p>	<p>As learning increasingly moves online and digital the need to observe and act <i>timeously</i> and personalize learning or structure immediate interventions has become the new normal.</p> <p><i>Surveillance scope:</i> rhizomatic and/or predictive</p>	<p>Undefined, unclear.</p>	<p>Still sees students as data objects and providers of information</p>	<p>Insight into behaviors and outcomes which may lead to future or immediate changes to teaching and learning design and delivery.</p>
<p><i>Automated</i></p>	<p>Outcomes are determined on the basis of prediction.</p> <p><i>Surveillance type:</i> automatic and/or volunteered</p>	<p>Future outcomes are linked to past behaviors and outcomes (students like you).</p> <p><i>Surveillance scope:</i> panoptic and/or predictive</p>	<p>Undefined, unclear.</p>	<p>Students seen as data objects with more defined labels attached.</p>	<p>Interventions based on knowledge of historical students.</p> <p>Can be problematic in terms of algorithmic reliability, robustness of data, and application of predictions as <i>fact</i>.</p>

<p><i>Participatory process and collaborative sense-making</i></p>	<p>Methods and purpose are explored with all relevant stakeholders taking an active role.</p> <p><i>Surveillance type:</i> directed, automated, and/or volunteered</p>	<p>Education is an open and recursive system – resulting in the need for continuous, context-specific sense-making.</p> <p>All stakeholders should be involved – students, researchers, teaching staff, administrative staff, regional staff, etc.</p> <p><i>Surveillance scope:</i> rhizomatic and/or predictive</p>	<p>All stakeholders involved – may need broad, blanket consensus at the beginning of each course or as part of the teaching and learning policy of the institution – overseen by the highest decision making body.</p>	<p>Students are more than data objects or the providers of consent. They are and should be active collaborators in the sharing, interpretation, governance, quality assurance and integrity of data (as proposed by the Precision Medicine Initiative (2015)).</p>	<p>Continuous improvement, responsiveness in improving institutional effectiveness and efficiency, the allocation of resources, and student retention and success.</p>
--	--	---	--	--	--

## Conclusions

Learning analytics has clear potential to positively impact teaching and learning as well as contribute towards the more effective allocation of resources and return on investment. However, ethical concerns often remain unaddressed. Traditionally, IRBs assume responsibility to oversee and ensure that research is ethical and that participants are informed and provide consent. Given the range of concerns and real potential for harm, one institutional response might be to subject all learning analytics to IRB approval. This would, however, contribute to the mission creep and increasing bureaucratization of research, and solidify concerns that learning analytics practice does not fit comfortably with traditional descriptions of researchers, research, informed consent, and approval.

The analysis of the finer points of IRB rules opens the dialogue for specific guidelines as they can apply to the growing fields of using student data for actionable insight. The use of excursus allows for a unique way to take a short digression and do a deep analysis of a particular policy point while leaving the methodology and line of argumentation intact. In this way, excursion as a “signpost” provides a bridge to examine the multitude of potential concerns in very different systems of student data use in schools in the United States, the United Kingdom, and South Africa. The proposed typology attempts to disentangle the dilemma of the need for oversight of the processes that monitor learning analytics amidst concerns that this may cause further mission creep by IRBs and frustrate the purpose of learning analytics. As such, in following “the balancing act of classifying” (Bowker and Star 1999, p.324), the proposed typology is a “living” heuristic (p. 326), open for reconfiguration and reconstruction.

Despite its limitations, the typology does assist in providing a map or lens of a phenomenon which might offer researchers and institutions insight into the existing gaps in

approval processes relating to both research and practice and so highlights a further need to consider some of the outstanding issues. The unearthing of such ethical issues helps not only research practice concerns, but also helps bolster the integrity of using student data to assist in the learning process. It is hoped that consideration of such a typology might encourage institutions to develop more extensive approvals guidelines or frameworks for evaluating current and future practice which in turn encourage an awareness of the ethical issues relating to widespread adoption of learning analytics.

## References

- Abbott, L. & Grady, C. (2012). A systematic review of the empirical literature evaluating IRBs: What we know and what we still need to learn. *Journal of Empirical Research on Human Research Ethics* 6(1), 3-19.
- Bauman, Z. (2004). *Wasted lives. Modernity and its outcasts*. Cambridge, UK: Polity Press.
- Bledsoe, C. H., Sherin, B., Galinsky, A. G., & Headley, N. M. (2007). Regulating creativity: research and survival in the IRB iron cage. *Nw. UL Rev.*, 101, 593.
- Buckingham Shum, S. & Ferguson, R. (2012). Social learning analytics. *Educational Technology & Society* 15(3), 3-26.
- Brumbaugh, A. J. (1960). Research designed to improve institutions of higher learning. Retrieved from <http://files.eric.ed.gov/fulltext/ED017141.pdf>
- Bos, W. & Tarnai, C. (1999). Content analysis in empirical social research. *International Journal of Educational Research* 31, 659-671.
- Bowker, G. C., & Star, S. L. (1999). *Sorting things out: Classification and its consequences*. London, UK: The MIT Press.
- Buchanan, E. A. & Ess, C. M. (2009). Internet research ethics and the institutional review board: Current practices and issues. *SIGCAS Computers and Society* 39(3), 43-49.
- Buchanan, E. A. & Hvizdak, E. E. (2009). Online survey tools: Ethical and methodological concerns of human research ethics committees. *Journal of Empirical Research on Human Research Ethics* 4(2), 37-48.
- Campbell, J., DeBlois, P. B., & Oblinger, D. G. (2007). Academic analytics: A new tool for a new era. *EDUCAUSE Review*.

- Carr, C. T. (2015). Spotlight on ethics: institutional review boards as systemic bullies. *Journal of Higher Education Policy and Management*, 37(1), 14-29.
- Clark, K., Duckham, M., Guillemin, M., Hunter, A., McVernon, J., O’Keefe, C., Pitkin, C., Praver, S., Sinnott, R., Warr, D., & Waycott, J. (2015). *Guidelines for the ethical use of digital data in human research*. Carlton Connect Initiative, The University of Melbourne, Melbourne.
- Cross, K. P. (1967). When will research improve education? Retrieved from <http://eric.ed.gov/?id=ED025206>
- Colvin, C., Rogers, T., Wade, A., Dawson, S., Gašević, D., Buckingham Shum, S., Nelson, K., Alexander, S., Lockyer, L., Kennedy, G., Corrin, L., & Fisher, J. (2016). *Student retention and learning analytics: a snapshot of Australian practices and a framework for advancement*. Canberra, ACT: Australian Government Office for Learning and Teaching Retrieved from <http://www.olt.gov.au/project-student-retention-and-learning-analytics-snapshot-current-australian-practices-and-framework>.
- Dawson, S., Gašević, D., Siemens, G., & Joksimović, S. (2014). Current state and future trends: A citation network analysis of the learning analytics field. In *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*, 231-240. ACM.
- Downe-Wamboldt, B. (1992) Content analysis: method, applications, and issues, *Health Care for Women International*, 13(3), pp. 313—321. DOI: 10.1080/07399339209516006.
- de Freitas, S., Gibson, D., Plessis, C. D., Halloran, P., Williams, E., Ambrose, M., Dunwell, I. & Arnab, S. (2014). Foundations of dynamic learning analytics: Using university student data to increase retention. *British Journal of Educational Technology* 46(6), 1175-1188.

- Ehrenberg, R. G., Jakubson, G. H., Groen, J. A., So, E., & Price, J. (2007). Inside the black box of doctoral education: What program characteristics influence doctoral students' attrition and graduation probabilities? *Educational Evaluation and Policy Analysis*, 29(2), 134-150.
- Elo, S., & Kyngäs, H. (2007). The qualitative content analysis process. *Journal of Advanced Nursing* 62(1), 107-115, doi: 10.1111/j.1365-2648.2007.04569.x.
- Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5-6), 304-317.
- Greller, W. & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Journal of Educational Technology and Society*, 15(3), 42-57.
- Guta, A., Nixon, S.A., & Wilson, M.G. (2013). Resisting the seduction of “ethics creep”: Using Foucault to surface complexity and contradiction in research ethics review. *Social Science and Medicine* 98, 301-310.
- Hack, C. (2015a). Applying learning analytics to smart learning — ethics and policy. In Andrew Middleton (Ed.), *Smart Learning - teaching and learning with smartphones and tablets in post-compulsory education* (57-62). Retrieved from <http://melsig.shu.ac.uk/wpcontent/uploads/2015/04/Smart-Learning.pdf>
- Hack, C. (2015b, April 30). Does pedagogic research require ethical review? [Web log]. Retrieved from: <https://www.heacademy.ac.uk/does-pedagogic-research-require-ethical-review>
- Howard, R. D., McLaughlin, G. W., & Knight, W. E. (2012). *The handbook of institutional research*. San Francisco, CA: Jossey-Bass.

- Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis, *Qualitative Health Research* 15(9), 1277-1288. doi: 10.1177/1049732305276687.
- Indiana University. (2015). Office of Research Compliance. Retrieved from:  
<http://researchcompliance.iu.edu/hso/index.html>
- Kay, D., Korn, N. & Oppenheim, C. (2012). Legal, risk and ethical aspects of analytics of analytics in higher education. *CETIS Analytics Series* 1(6), 1-30.
- Kelly, A. E. & Seppälä, M. (2015). Changing policies concerning student privacy and ethics in online education. *International Journal of Information and Education Technology* 6(8), 652-656.
- Kitchin, R. (2013). Big data and human geography: Opportunities, challenges and risks. *Dialogues in Human Geography* 3(3), 262-267.
- Kluge, S. (2000). Empirically grounded construction of types and typologies in qualitative social research. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research* 1(1).
- Knobelsdorf, M. (2008). A typology of CS students' preconditions for learning. In *Proceedings of the 8th International Conference on Computing Education Research* (62-71). ACM.
- Knox, D. (2010). Spies in the house of learning: A typology of surveillance in online learning environments. In *Proceedings from e-Learning: The Horizon and Beyond conference* (St. John's, Newfoundland, Canada).
- Koops, B-J., Newell, B. C., Timan, T., Škorvánek, I., Chokrevski, T., & Galič, M. (forthcoming). A typology of privacy. *University of Pennsylvania Journal of International Law*. Retrieved from [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2754043](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2754043)
- Kovanović, V., Joksimović, S., Gašević, D., Hatala, M., & Siemens, G. (2015). Content Analytics: the definition, scope, and an overview of published research. *Handbook of*

- Learning Analytics*. Retrieved from [http://www.researchgate.net/profile/Vitomir\\_Kovanovic/publication/283462803\\_Content\\_Analytics\\_the\\_definition\\_scope\\_and\\_an\\_overview\\_of\\_published\\_research/links/5639163008ae4624b75efb36.pdf](http://www.researchgate.net/profile/Vitomir_Kovanovic/publication/283462803_Content_Analytics_the_definition_scope_and_an_overview_of_published_research/links/5639163008ae4624b75efb36.pdf)
- Kruse, A. & Pongsajapan, R. (2012). Student-centered learning analytics. CNDLS Thought Paper. 1-12. Retrieved from: <https://cndls.georgetown.edu/m/documents/thoughtpaper-krusepongsajapan.pdf>
- Largent, E. A., Grady, C., Miller, F. G., & Wertheimer, A. (2012). Money, coercion, and undue inducement: A survey of attitudes about payments to research participants. *IRB* 34(1), 1-8.
- Locke, S., Ovando, C. J., & Montecinos, C. (2016). Institutional power and the IRB: Saving souls or silencing the other in international field work. In Bhopal, K. and Deuchar, R. (Eds.) *Researching Marginalized Groups*. New York, NY: Routledge.
- Macfadyen, L. P. & Dawson, S. (2012). Numbers are not enough. Why e-learning analytics failed to inform an institutional strategic plan. *Educational Technology and Society*, 15(3), 149-163.
- Macfadyn, L. P., Dawson, S., Pardo, A., & Gašević, D. (2014). Embracing big data in complex educational systems: The learning analytics imperative and the policy challenge. *Research and Practice in Assessment* 9(2), 17-28.
- Moskal, P. D. (2016). Incorporating the scholarship of teaching and learning (SoTL) into instruction. In C.D. Dziuban, A.G. Picciano, C.R. Graham, and P.D. Moskal (Eds.), *Conducting Research in Online and Blended Environments: New Pedagogical Frontiers*, New York: Routledge.

- Moxley, J. (2013). Big data, learning analytics, and social assessment. *The Journal of Writing Assessment* 6(1).
- Open University. (2013). Code of practice for research at the Open University. Retrieved from <http://www.open.ac.uk/research/main/sites/www.open.ac.uk.research.main/files/files/ecms/web-content/CoP-amended-after-Senate-Feb-2014-Final-version-updated-Dec-2014-for-external-use-FINAL.pdf>
- Open University (2014). Policy on ethical use of student data for learning analytics. Retrieved from <https://learn3.open.ac.uk/mod/url/view.php?id=85812>
- Precision Medicine Initiative (2015). Privacy and Trust Principles. Retrieved from <https://www.whitehouse.gov/precision-medicine>
- Prinsloo, P. (2014, October 22). Mene, mene, tekkel, upharsin: researcher identity and performance. Inaugural lecture at the University of South Africa. Retrieved from [http://www.researchgate.net/profile/Paul\\_Prinsloo/publication/267395307\\_Mene\\_mene\\_tekkel\\_upharsin\\_researcher\\_identity\\_and\\_performance/links/544f2f200cf29473161bf642.pdf](http://www.researchgate.net/profile/Paul_Prinsloo/publication/267395307_Mene_mene_tekkel_upharsin_researcher_identity_and_performance/links/544f2f200cf29473161bf642.pdf)
- Prinsloo, P. & Slade, S. (2013). An evaluation of policy frameworks for addressing ethical considerations in learning analytics. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, 240-244. ACM.
- Prinsloo, P. & Slade, S. (2014). Student data privacy and institutional accountability in an age of surveillance. In Menon, M. E., Terkla, D. G., and Gibbs, P. (Eds.) *Using Data to Improve Higher Education*. Rotterdam, The Netherlands: Sense Publishers.

- Prinsloo, P. & Slade, S. (2015). Student privacy self-management: Implications for learning analytics. In *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge*, 83-92. ACM.
- Redeker, G. (2000). Coherence and structure in text and discourse. *Abduction, Belief and Context in Dialogue*, 233-263. Retrieved from <http://www.let.rug.nl/~redeker/redeker2000.pdf>
- Regan, J., Baldwin, M. A., & Peters, L. (2012). Ethical issues in pedagogical research. *Journal of Pedagogic Research* 2(3), 44-54.
- Rule, P. & John, V. (2011) *Case study research*, Pretoria: Van Schaik Publishers.
- Schildkamp, K., Lay, M. K., & Earl, L. (Eds). 2013. *Data-based decision making in education. Challenges and opportunities*. London, UK: Springer.
- Scott, J. (Ed.). (2005) *Documentary research*, London, UK: Sage.
- Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE review*, 46(5), 30-40.
- Siemens, G. (2012). Learning analytics: Envisioning a research discipline and a domain of practice. *Second International Conference on Learning Analytics & Knowledge*, Vancouver, BC, Canada, 29 April - 2 May 2012.
- Slade, S. & Prinsloo, P. (2013). Learning analytics ethical issues and dilemmas. *American Behavioral Scientist*, 57(10), 1510-1529.
- Subotzky, G., & 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.
- Tinto, V. (2006). Research and practice of student retention: What next? *Journal of College Student Retention*, 8, 1-19.

- Tinto, V. (2012). Enhancing student success: Taking the classroom success seriously. *The International Journal of the First Year in Higher Education*, 3(1), 1.
- Thomas, G. (2011) *How to do your case study. A guide for students and researchers*, London, UK: Sage.
- Unisa. (2012). Policy on research ethics. Retrieved from [http://www.unisa.ac.za/contents/faculties/service\\_dept/ice/docs/Ethics%20and%20research\\_Policy\\_2012.pdf](http://www.unisa.ac.za/contents/faculties/service_dept/ice/docs/Ethics%20and%20research_Policy_2012.pdf)
- Unisa. (2014). Data privacy policy. Approved by Council, November. Pretoria, South Africa: Unisa.
- Vitak, J., Shilton, K., & Ashktorab, Z. (2016). Beyond the Belmont principles: Ethical challenges, practices, and beliefs in the online data research community. Forthcoming for *CSCW 2016*, 27 February - 2 March. Retrieved from: [http://www.researchgate.net/profile/Jessica\\_Vitak/publication/282348909\\_Beyond\\_the\\_Belmont\\_Principles\\_Ethical\\_Challenges\\_Practices\\_and\\_Beliefs\\_in\\_the\\_Online\\_Data\\_Research\\_Community/links/560d69bd08ae6cf68153efcd.pdf](http://www.researchgate.net/profile/Jessica_Vitak/publication/282348909_Beyond_the_Belmont_Principles_Ethical_Challenges_Practices_and_Beliefs_in_the_Online_Data_Research_Community/links/560d69bd08ae6cf68153efcd.pdf)
- Willis, III., J. E. (2014). Learning analytics and ethics: A framework beyond utilitarianism. *EDUCAUSE Review*.
- Willis III, J. E., Campbell, J., & Pistilli, M. (2013). Ethics, big data, and analytics: A model for application. *EDUCAUSE Review Online*.
- Willis, III. J. E. & Strunk, V. A. (2015). Ethical responsibilities of preserving academicians in an age of mechanized learning: Balancing the demands of educating at capacity and preserving human interactivity. In White, J. and Searle, R. (Eds.) *Rethinking Machine Ethics in the Age of Ubiquitous Technology* (166-195). Hershey, PA: IGI Global.

- Yanqing, D., Guangming, C., One, V., & Woolley, M. (2013). Big data in higher education: An action research on managing student engagement with business intelligence, *Second International Conference on Emerging Research Paradigm in Business and Social Science*, Middlesex University, Dubai, 26-28 November 2013.
- Yin, R.K. (2009) *Case study research. Design and methods*, 4<sup>th</sup> edition, Applied Social Research Methods Series, Volume 5. London, UK: Sage.

### **Acknowledgments**

The authors would like to thank Dr. Viktoria A. Strunk for providing feedback and suggested edits on early drafts of this paper. The authors are also appreciative of three reviewers' close reading and anonymous peer review feedback.