Jones International Journal of Educational Technology in Higher Education (2019) 16:24
https://doi.org/10.1186/s41239-019-0155-0

International Journal of Educational Technology in Higher Education

RESEARCH ARTICLE Open Access

Learning analytics and higher education: a proposed model for establishing informed consent mechanisms to promote student privacy and autonomy

Kyle M. L. Jones

Correspondence: kmlj@iupui.edu; kylejones@thecorkboard.org
Department of Library and Information Science, School of Informatics and Computing, Indiana University–Indianapolis (IUPUI), Indianapolis, Indiana, USA

Abstract

By tracking, aggregating, and analyzing student profiles along with students' digital and analog behaviors captured in information systems, universities are beginning to open the black box of education using learning analytics technologies. However, the increase in and usage of sensitive and personal student data present unique privacy concerns. I argue that privacy-as-control of personal information is autonomy promoting, and that students should be informed about these information flows and to what ends their institution is using them. Informed consent is one mechanism by which to accomplish these goals, but Big Data practices challenge the efficacy of this strategy. To ensure the usefulness of informed consent, I argue for the development of Platform for Privacy Preferences (P3P) technology and assert that privacy dashboards will enable student control and consent mechanisms, while providing an opportunity for institutions to justify their practices according to existing norms and values.

Keywords: Higher education, Learning analytics, Student privacy, Autonomy, Informed consent

Introduction

Big Data is a 'cultural, technological, and scholarly phenomenon' (Boyd & Crawford, 2012, p. 663) that transcends boundaries; consequently, researchers and pundits alike have had a hard time establishing a 'rigorous definition' (Mayer-Schönberger & Cukier, 2013, p. 6).[1] Big Data generally allows for 'things one can do at a large scale that cannot be done at a smaller one, to extract new insights or create new forms of value' due to new flows of data and information derived from observing human behaviors or information disclosures by individuals (Mayer-Schönberger & Cukier, 2013, p. 6). This has proven to be valuable in many contexts (e.g., commerce, national security, etc.), and higher education is now pursuing its own Big Data agenda to mine for insights into student behaviors, learning processes, and institutional practices using learning analytics technology.

Much like Big Data, there exists no commonly accepted definition of learning analytics (for sundry definitions, see Dawson, Heathcote, & Poole, 2010; van Barneveld, Arnold, &

Campbell, 2012). However, it is often understood as 'the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising [*sic*] learning and the environments in which it occurs' (Long & Siemens, 2011, p. 33).

While emerging learning analytics practices hold some promise to improve higher education, they are morally complicated and raise ethical questions, especially around student privacy. Since learning analytics often rely on aggregating significant amounts of sensitive and personal student data from a complex network of information flows, it raises an important question as to whether or not students have a right to limit data analysis practices and express their privacy preferences as means to controlling their personal data and information.

I begin the paper with an overview of learning analytics. I follow this part with a discussion on privacy theory, especially as it relates to information control and how such controls support and extend individual autonomy. Informed consent has historically been the mechanism by which we try to control information about ourselves, so I consider its role in expressing our privacy preferences and its limitations in the age of Big Data. Next, I highlight the many ways students unknowingly disclose data and information to their institution and third parties without the ability to control such disclosures. Finally, I propose a model for establishing informed consent mechanisms to promote student privacy and autonomy using P3P technology and privacy dashboards in ways that balance student and institutional interests.

Big data and higher education

New pathways for higher education policy and the learning sciences are opening up due to the growth of interconnected databases in data warehouses. Many learning analytics advocates believe capturing, archiving, and analyzing student profiles and behaviors will lead to improved institutional decision making, advancements in learning outcomes for at-risk students, greater trust in institutions due to the disclosure of data, and significant evolutions in pedagogy, among other things (Long & Siemens, 2011). To support these ends, universities are actively aggregating student data to support an array of learning analytics initiatives, which I address in this section.

Opening the black box of learning with student data and learning analytics

A complex assemblage of information and educational technology drives colleges and universities, and it has brought about a new phenomenon: The datafication of learning (Mayer-Schönberger & Cukier, 2014a). Each bit and byte, once aggregated and analyzed, may hold potential to reveal impressive new insights into student learning behaviors and outcomes. In the hands of educators, data-based visualizations of *how* and *what* a student is learning can assist instructors to develop customized instructional strategies and curricula. Each student represents a potential source of data, and considering that 21 million students enrolled in American higher education institutions in 2012 (National Center for Education Statistics, 2013), universities have a latent trove of data ready for Big Data projects.

Beyond the individual student level, there also exists opportunities for institutions to share their disparate datasets (see Unizin, 2015) or even link data at a federal level (see

Kolowich, 2013), which presents further opportunities for analytical insights at an even larger scale. Eleven research-intensive universities and two state systems are members of Unizin's consortium, which according to its CEO and COO aims to '[p] articipate in the creation of the world's largest learning laboratory' by creating a 'data pool' that 'would allow institutions to take a scholarly and practical approach to critical questions around student performance' (Littleworth & Qazi, 2017). Joining the consortium provides an institution access to a data warehouse in which over 720,000 students may exist as data points.[2] At the time of this writing, the warehouse reportedly held all the data created in the consortium's central learning management system, Canvas; however, there is a possibility to enhance analytics by aggregating data from other sources (e.g., admissions records) and Unizin tools (Qazi, 2017; Raths, 2016).

Learning management system analytics

The most common application of learning analytics technology is in the context of an institution's learning management system (LMS). LMSs are traditionally used to support online or hybrid teaching environments, within which students interact with various learning objects and work collaboratively. For example, students take quizzes; submit assignments; read assigned materials, such as journal articles and other electronic texts (eTexts or eBooks); and interact with their peers in discussion forums and wikis.

Learning analytics systems capture student behaviors, which are commonly referred to as the 'digital breadcrumbs' students leave throughout the system within LMSs as students navigate and interact with their peers and the digital space (Norris, 2011). In the recent past, it was a 'slow and cumbersome' process to export LMS data for analysis, but it is increasingly the case that common LMS systems include data extraction tools alongside their analytic products (Brown, Dehoney, & Millichamp, 2015; Macfadyen & Dawson, 2010, p. 590). The analytics can descriptively detail the date, time, and duration of students' digital movements, including if, when, and for how long they read an electronic text (e.g., eBook or PDF article) or took an online quiz. Other statistics detail a student's overall completion rate of a course, whether or not a student is predicted to succeed in the course, and map the strength of a student's peer-to-peer/peer-to-instructor network using social network analysis. LMSs embedded with learning analytics tools use data visualization techniques to create information dashboards from which instructors can infer how to intervene in a student's education, while other systems allow students, themselves, the ability to monitor their own progress using similar dashboards.[3] Some systems automatically intervene with algorithms, which send status updates or e-mails to students and instructors alike, notifying both parties of potential problems.

LMS-based learning analytics are informed by student data from other campus systems, including commonly used student information systems (SISs). SISs hold a majority of the information students disclose on their applications for admission, their enrollment records, and their academic history. Over time, their digital records may be augmented with other information, including financial aid awards, involvement on campus, disciplinary and criminal reports, and personal health information.

eAdvising analytics

eAdvising systems are another area ripe for learning analytics. Austin Peay State University's eAdvising system includes a recommendation engine that suggests courses

based on students' academic profiles and considers their course path with the past success of peers like them (Denley, 2012). Other eAdvising systems warn students when they stray from their chosen path, blocking them completely from registering for courses if they fail to return to a pre-determined set of courses; or if students are deemed to be 'at risk,' professional advisors give them priority advising attention (California State University Long Beach, 2014; Lewis, 2011; Parry, 2012).

eAdvising analytics rely heavily on data held within institutional SISs. The historical academic information, especially past ACT and SAT scores, alongside current academic information, such as course grades and enrollment records, are crucial for predictive eAdvising analytics systems. eAdvising systems, like Campus Labs' Beacon system, pull supplemental data from sources like personality profiles, specialized entrance exams and surveys, and geolocation information from student ID card swipes or WiFi-connected device beacons. For example, Beacon's survey questions automatically create an alert for resident assistants in campus housing if students indicate that they are having trouble making friends, and geolocation tracking information is available to advisors for them to assess a student's engagement on campus (Campus Labs, 2014).

Institutional analytics

While learning analytics applications typically focus attention on individual courses and learners, there is a growing market for institution-wide analytic applications. Brightspace, Blackboard, and Instructure, all prominent educational technology companies, offer learning analytics solutions that allow institutional researchers and other administrators access to data and dashboards that compare student activity and learning metrics within and between courses, departments, and colleges across a university.[4]

Institution-wide learning analytics afford administrators the ability to drill down into segmented and longitudinal student data. Doing so helps an institution develop reports concerning student performance with respect to learning outcomes, departmental performance measures, and instructor performance over time. These measures and more, some argue, help an institution and its individual departments respond to stakeholder pressures to demonstrate institutional effectiveness and more easily meet government reporting requirements (Glass, 2013; Long & Siemens, 2011).

Edge-case analytics using social and biometric data

Leading thinkers in the learning analytics field argue that a student's 'every click, every Tweet or Facebook status update, every social interaction, and every page read online' leaves a 'digital footprint' (Long & Siemens, 2011, p. 32) that can 'render visible' (Buckingham Shum & Ferguson, 2012, p. 5) unseen social learning behaviors. This 'smorgasbord' (Diaz & Brown, 2012, p. 13) approach to data aggregation motivates novel approaches to learning analytics and encourages 'fishing expeditions' (Mayer-Schönberger & Cukier, 2013, p. 29) within the data for new insights and trends.

Learning analytics advocates have yet to demonstrate the efficacy of social analytics at scale, but emerging projects point to some potential uses. Some institutions are monitoring and mining their students' use of Facebook (see Ho, 2011; Hoover, 2012), while other institutions even scan RFID chips in student IDs at lecture halls and classrooms in order to correlate attendance with classroom performance (Brazy, 2010; O'Connor, 2010). If universities track student movements using geolocation data and map interpersonal

connections, they can begin to understand the social lives of students, their relationships, and the web of personal networks on campus, which Matt Pittinsky (formerly of Blackboard) believes is a 'very useful layer of data …. [that shows] evidence of social integration' (Parry, 2012, para. 57), an important indicator of academic success.

Institutions and researchers are also exploring the role of biometric data in learning analytics. Advocates of biometrics for learning analytics argue that measurements of a student's 'heart rate, body temperature, ambient luminosity, [location and movement],' among other things can be useful for understanding attention, stress, and sleep patterns, which hold the potential to determine circumstances that impede or aid learning (Arriba Pérez, Santos, & Rodriguez, 2016, p. 43). When biometrics and the analytics resulting from them are shared with learners, initial research indicates that such information may help individuals self-regulate their attention behaviors (Spann, Schaeffer, & Siemens, 2017).

To these ends, the Bill and Melinda Gates Foundation, an outspoken proponent of data-driven education, funded the development of an 'engagement pedometer,' a biometric bracelet that tracks electrical charges in a student's sympathetic nervous system (Simon, 2012). By way of analytics that analyze each bracelet's data, instructors can see a student's engagement level (or lack thereof) in real time. While this and other similar projects have not reached the mainstream, they foreshadow the role biometric data can play in learning analytics projects (see Alcorn, 2013; Schiller, 2015).

Learning analytics and privacy as control of one's data and information

If institutions continue to develop data analytics projects and infrastructures in order to capture sensitive, comprehensive student data, the obligation to do so responsibly will increase as well. Even with noble and good ends in mind—namely improving learning (however defined)—learning analytics practices surveil and intervene in student lives. Consequently, learning analytics, like many Big Data practices, are rife with privacy problems and ethical quandaries, which continue to grow in complexity (Johnson, Adams Becker, Estrada, & Freeman, 2015).

The question then is whether or not those who design learning analytics systems and support its ends will provide students privacy protections. Evidence in the literature suggests that learning analytics highlight 'blind spots' (Greller & Drachsler, 2012, p. 50) in institutional policy and 'poses some new boundary conditions' (Pardo & Siemens, 2014, p. 442) around student data and privacy, which may negatively affect the future success of learning analytics if left unaddressed (Siemens, 2012). One such question concerns the degree to which students should control information about themselves; I turn to this for the remainder of the article.

Privacy as control of information

Big Data practices often raise significant privacy issues, which have sparked academic and public debate with fervor and intensity last seen in the 1970s when concerns erupted regarding government data banks (see Lyon, 2014; Marr, 2015). The rise of data collection in and of itself is concerning, but the advancing pace of predictive analytics and their role in public and private life pushes against accepted normative, ethical, and legal privacy boundaries in ways unforeseen and unknown

(Crawford & Schultz, 2014). As such, the scholarly conversation surrounding Big Data and privacy, especially information privacy, is multifaceted and reflects various theories and approaches to privacy problems.

Privacy as a form of information control is a dominant theme in scholarly literature, serves as the basis for legal doctrine, and has informed important Supreme Court decisions (Nissenbaum, 2010; Solove, 2008). According to Alan Westin's (1967, p. 7) seminal text, *Privacy and Freedom*, privacy is an individual's 'right to determine for themselves when, how, and to what extent information about them is communicated to others.' A control approach to privacy assumes not that information is absent in others' minds, but that we can determine who can access information about ourselves and limit to whom and under what conditions it is disclosed (Fried, 1968; Froomkin, 2000; Nissenbaum, 2010).

Privacy-as-control is biased towards individual choice and treats information as a part of one's person. In many respects, individual information control treats personal information as a Lockean property right (Solove, 2008). By acknowledging that individuals have the right to choose how others can access and use their information, this privacy perspective advances the idea that information 'flows naturally from selfhood' (Solove, 2008, p. 26), thus 'every Man has a *Property* in his own *Person'* and that property should be respected as being part and parcel to one's self (Locke, 1689, emphasis and capitalization in original).

Losing control

Big Data practices present unique issues that are dissolving our control over personal information. The technological mélange of ubiquitous sensors, devices, networks, and applications around and embedded in our lives continue to surreptitiously capture data about us. These data are valuable, which has prompted companies, institutions, and especially data brokers, whose under-regulated industry often fails to protect individuals against consequential data leaks (Roderick, 2014; see Cowley, Bernard, & Hakim, 2017), to build data-mining infrastructures.

When identifiable data are aggregated and analyzed, lives become more transparent to those with the data while their data practices grow more opaque and influential. This is what Richards and King (2013) call the Transparency Paradox. While we may wish to keep information private by expecting companies to de-identify data, the connected nature of databases and the power of analytic technologies often makes deidentification efforts futile (Ohm, 2010). Richards and King (2013) identify this as the Identity Paradox. And institutions and organizations continue to grow their privilege and power over individuals by exploiting their personal information, while the same individuals are left with few options to rein in flows of personal information. This is Richards and King's (2013) final paradox: The Power Paradox.

The risk of each paradox would be lessened if individuals had more control over their personal information. However, institutional bureaucracy, corporate policy, and legal jargon adds to a Kafkaesque nexus that makes such information control processes unapproachable, much less useful (Solove, 2004; Tene & Polonetsky, 2013). Without some checks on personal information flows and the development

of digital dossiers, individuals will have little say in how powerful entities use identifiable information (Solove, 2004).

Harms to autonomy

What is problematic about people losing control over their information is the effect Big Data and other data-driven practices have on autonomy (Goldman, 1999). Autonomous individuals are self-governing, which is to say that they are able to incorporate their 'values and reasons' (Rubel & Jones, 2016, p. 148) into rational decision-making processes according to their will (Kant, 1785). Society cares about protecting one's autonomy because it 'shows respect for the person' (Marx, 1999, p. 63).

Autonomy and information privacy are often interlinked. According to Rubel and Jones (2016), three discrete types of connections exist between the concepts. First, privacy may be an *object* of autonomy, which is to say that individuals may choose to seek information privacy or not. Second, privacy may be a *condition* of autonomy. Here, privacy serves 'a fundamental and ineliminable role' (Alfino & Mayes, 2003, p. 6) in autonomy by protecting individuals from undue intrusions into spheres of life that could limit 'individual conscience' (Richards, 2008, p. 404)—such as developing intellectually, forming moral constructions, and assessing social values—or influence one's decisions to the point that they are not fully one's own (Bloustein, 1964; Reiman, 1976). Finally, privacy may *promote* autonomy. When organizations and institutions respect information privacy expectations and allow information to flow according to those expectations, they advance autonomist aims; however, when these same actors hide information, use information to deceive, or employ information practices to interfere and manipulate individual lives, they reduce autonomy.

Due to predictive capabilities and the direct influence Big Data practices have in daily life, these emerging data-based technologies present real threats to individual autonomy. Many Big Data practices aim to capture as much of the human experience as possible, including physical, mental, and emotional activity. In doing so, individuals are taken from a corporeal whole and transformed into binary code as 'data doubles' with the purpose of changing 'the body into pure information, such that it can be rendered more mobile and comparable' (Haggerty & Ericson, 2000, p. 611). The problem is that the data double fails to be a 'comprehensive or representative' reflection of human life, yet powerful actors use it to influence a person's behavior (Andrejevic & Gates, 2014, p. 191). Where autonomy is concerned, organizations and institutions who analyze data doubles rarely promote autonomy by failing to describe the construction of the algorithm, the information on which it relies, and how and when analytic technologies nudge humans to accomplish specific ends—which may not be in an individual's best interests. When we know that such practices are occurring in digital spaces that help us intellectually develop (e.g., when we search for information or read eBooks), we may 'guard our words [and] our thoughts' (Richards, 2015, p. 101); thus, the surveillance minimizes our autonomy.

The role of informed consent in expressing privacy choices

Informed consent, or 'notice and choice,' is the process by which individuals are notified of how a secondary party, such as organizations (like a business) or institutions (like a university), will use information about them (Tene & Polonetsky, 2013, p. 260). It also informs them of their rights to privacy, as well as the express rights the second

party retains regarding the information. After being informed of rights and information practices, individuals can then choose whether or not to agree—to *consent*—to the terms in front of them and enter into a relationship with the second party or not. However, even though informed consent acts as 'the gold standard for privacy protection' (National Research Council, 2007, p. 48), it is not a panacea for privacy problems (Flaherty, 1999).

Rarely are individuals fully aware of what they are agreeing to. In our current data brokerage climate, Adam Moore (2010) argues that the benefits we gain from consent to one set of information-based services are far outnumbered by the harms that can accrue when the same information is sold later on. Furthermore, consent implies awareness of how our information will be used, but we can rarely envision the downstream uses, the unequal benefit to the second and third parties to whom we disclose information, and the potential consequences for our privacy (Hui & Png, 2006; Marx, 1999). Also concerning is the fact that informed consent procedures are usually biased towards those who seek out personal information. It is also often the case that individuals must choose to opt-out of *inclusive* information gathering practices, not opt-in, which produces the effect that more information is gathered than necessary.

Data miners do not shoulder full responsibility for the weaknesses of informed consent; some of it rests with individuals. Acquisti's (2004) work on informed consent behaviors revealed that individuals desire immediate gratification and are more willing to opt-in to inclusive information practices in part because it requires them to do less work to protect their privacy and limit information disclosures. This want for gratification is more quickly satiated when companies provide a sense of control––even if this is not the case—that motivates individuals to consent (Brandimarte, Acquisti, & Loewenstein, 2013). Through this lens we can see how informed consent can become a predatory structure that does not benefit individuals nor promotes their ability 'to make meaningful, uncoerced choices' (Goldman, 1999, p. 103) through negotiation of information disclosure terms.

Big Data practices add additional challenges to informed consent mechanisms in ways that create informational and technological issues, some of which may be insurmountable. It is increasingly the case that informed consent continues to '[groan] under the weight' of dynamic and complex assemblages of systems, information flows, and data-driven practices; consequently, new approaches to informed consent are necessary in the Big Data era if we are to recapture the value informed consent once held for protecting privacy (Barocas & Nissenbaum, 2014, p. 64). Going forward, I recommend a novel approach to improving informed consent after first illustrating the many ways students unwittingly disclose data and information about themselves to higher education actors.

Disclosing and using data without student consent

Historically, higher education institutions have *failed* to promote informed consent practices within and outside of classrooms, using paternalistic justifications to warrant their information practices (Connelly, 2000). But when students were recently asked about data practices in higher education, they made compelling statements in favor of personal data control and the need for fair and useful informed consent processes (Slade & Prinsloo, 2014). The discrepancy between what institutions think they can do with student data and what students expect is done with their data may 'rupture the

fragile balance of respect and trust upon which this relationship is founded' (Beattie, Woodley, & Souter, 2014, p. 424). By highlighting information practices in higher education, this section details when and how students disclose data and information about themselves without ever being informed about the analytic purposes to which they may be put by their institution.

Comprehensive profiles

One driving motivation of those who advocate for learning analytic technologies is to understand how different populations of students learn. In order to accomplish this, institutions must develop comprehensive profiles about learners. For businesses who share the same goal, they look outward and purchase data profiles from data brokers. For higher education institutions, they look inward and mine the trove of information gleaned from admissions materials and applications.

The information students reveal about themselves on applications for admission, and materials in support of their applications, is not trivial; in fact, it is often sensitive and telling. Admission applications include questions related to a student's academic achievement, including transcripts and standardized test scores; professional ambitions; demographic and socioeconomic information; and family networks and their academic achievement level, among other things. For example, ACT and SAT documents include information about the types of activities students participated in while in high school, along with the kinds of social activities they plan to engage in while in higher education.[5,6] Some applications ask for descriptive essays related to the student's reading habits and cultural interests, and even the prospective student's disciplinary and criminal history. Others may even solicit answers regarding the student's religion, sexual orientation, and gender identity (see Caldwell, 2012; Hoover, 2011; Steinberg, 2010). In total, this information serves to build comprehensive individual profiles.

By building data-rich student profiles, universities set the foundation on which to run analytical tests and develop predictions. Where admission offices are concerned, institutional actors can compare data profiles of applicants with segments of the existing student body to develop predictive scores of the applicant's potential for success, and thus better inform the student enrollment process (Goff & Shaffer, 2014). After students enroll in their institution of choice, learning analytics technologies often correlate their digital and analog behaviors with specific segments of their respective profiles (e.g., GPA, race, gender, etc.); in fact, the efficacy of most learning analytics applications would markedly decrease if it were not for the ability to compare a student's digital trails with the wealth of information acquired from admissions applications. And while data profiles borne from admissions applications are rich, they become even more so as other sources of student data are grafted on as students interact with institutional information systems.

The problem is that it is unlikely that higher education institutions fully inform their prospective students about how the details of their lives revealed on admissions applications will be used and by whom. Clearly, students *expect* that these applications will inform admissions decisions, but they fail to intuit downstream uses and institutions do not explicitly explain information practices that are reliant on this store of personal data. In fact, applications for admission, the point at which we may expect universities to establish informed consent, may not even express student privacy rights, especially with regard to information control; many institutions even claim a property right to prospective students'

information.[7] This practice is especially problematic considering that students may feel that they have no option but to reveal all of the sensitive details about their lives, as there is always the chance they will be denied admission if they fail to provide information.

Classroom disclosures

Besides the application for admission, students also reveal sensitive information about themselves by creating profiles on third-party applications their institutions and instructors often require them to use in courses. Students are not routinely informed of the ways in which the companies responsible for these learning platforms use and protect the information students disclose as users.[8] Consider the example of Piazza, a company that offers question and answer functionality as a stand-alone application or with direct integration into common LMSs. Over 750,000 students at 1,000 institutions in 70 countries use Piazza to share information about themselves, access course materials, and communicate with their peers, instructors, and teaching assistants (J. Gilmartin [Piazza representative], personal communication; Piazza, n.d.). Data derived from students—including disclosures about their class history, internships, majors, and expected graduation year—have helped Piazza to build a secondary service, Piazza Careers. This service enables technology companies to court students for jobs if they fit a specific profile, that is after the companies purchase access to Piazza Career's store of student data-based analytics and other services (Piazza Careers, n.d.).

Higher education institutions often enter into contracts with third-party educational technology services in order to get access to useful teaching and learning applications; in return, educational technology companies get access to valuable student data. Some may assume that students are aware of already or can find out how these applications scrape user profiles for information to build secondary tools and services, but this is not accurate. While institutions often negotiate terms of service agreements on behalf of their students, the details of those agreements are opaque and not always readily or publicly accessible.

Simply because policies or memoranda of understanding exist that detail how student data should be used, we cannot assume that such agreements work to the benefit of students. In fact, a lack of transparency regarding these agreements and a failure to fully inform students about how third-party companies use their data raises immediate concerns and questions. It may be that institutions are withholding information about data practices to keep student privacy concerns at bay, concerns that could potentially derail beneficial contracts with vendors.

Universities may claim that hinderances to student information flows, like requiring informed consent, impede necessary institutional practices, like instruction or even day-to-day business activities. In fact, §99.31 of FERPA, the Family Educational Rights and Privacy Act, (1974), allows the institution to disclose private, identifiable student information—without informing students—to anyone within the institution who has a 'legitimate educational interest' or to a third party who provides 'institutional services or functions,' like an educational technology company.[9] But as we saw with the Piazza example, third parties can use student data for their own benefit.

Continuously tracked

With the rise of Big Data in higher education, universities will continually track a students' digital and physical movements and activities, and students will unknowingly disclose information about themselves on a daily basis. What is most problematic about these types of data disclosures is that the technology that enables them seems benign and beneficial. Students are not aware of the complex web of data capture technologies that store, aggregate, and analyze their information. Yet, there are particular types of data tracking that students may—and arguably should—be informed about to empower them to make informed decisions in their life.

Tracking technologies that capture geolocation, temporal data, and metadata raise serious concerns. Systems that can map in real time (or closely to) students' physical and/or digital location and the time of their movements or activities disturbs our normative expectations and riles up our concerns regarding 'dataveillance' (Clark, 1987). It is plausible that universities will use geolocation tracking to incentivize less social and more academically-oriented movements, like visiting the library, in order to improve learning outcomes.[10] And special categories of students may come under higher scrutiny than others, such as minorities who have received diversity scholarships or student-athletes who are already under constant surveillance where their social media is concerned (see Reed, 2013). In both cases, students may more closely regulate their behaviors due to concerns about how their data trails could be used against them (Hier, 2003).

Analytic technologies that assess a student's social well-being and affective state may also impact a student's expectation of privacy. Text mining, social network analysis, and biometric devices that observe and analyze data trails can monitor a students' level of engagement with their courses, discover whether or not they are socially connected with peers, and reveal if they are experiencing emotional issues, which some argue justifies institutional overrides of individual privacy (Prinsloo & Slade, 2017; Sclater, 2016). In effect, it makes typically invisible states of being and doing highly visible to any number of institutional actors with access rights. Yet, anyone who has had the privilege of experiencing college would balk at these revelations, as these formative years are often a time for identity development and exploration, socially and intellectually. Students may rightfully be worried that the data and insights mined from it will become a part of their permanent educational record and lead to decontextualized decision making (see Mayer-Schönberger & Cukier, 2014b). As evidence to this point, Stanford University students discovered that their institution logged when they used their ID cards to unlock doors; this information led to student backlash, substantiating that these are not unfounded concerns (see Pérez-Peña, 2015).

Building an informed consent model for learning analytics

Institutions retain the freedom to develop policies and practices in support of student privacy: FERPA is the policy 'floor' and not the 'ceiling' of how institutions should regulate and safeguard student information flows (Family Policy Compliance Office, 2011, p. 5; Rubel & Jones, 2016). In this section, I propose that institutions should use these freedoms to develop a technologically-enhanced informed consent mechanism using data privacy dashboards built on top of a technical identity layer. This model, I argue, considers the weaknesses of informed consent in the age of Big Data, and it

prompts institutions to explicitly justify how and when their information practices run afoul of existing norms in order to procure student consent.

The emerging student voice

From an institutional perspective, informed consent may run counter to the ends to which universities use learning analytics as a means. Recall the statistician's mantra: More data, more power. Informed consent opens up opportunities for limited access to and limited coverage about student life; consequently, students may reduce the efficacy of learning analytics by expressing their privacy preferences for greater control over identifiable data (Danezis et al., 2014 in Hoel & Chen, 2016; Slade & Galpin, 2012).

In my conversations with institutional actors, both for other research projects on learning analytics and in my daily interactions with administrators and staff, this argument—that institutions need all available student data to act in students' best interests—is often followed up with the position that students do not care about privacy in the first place, thus robust privacy protections are neither needed nor worth the effort. Emerging empirical evidence refutes this argument. Students are 'weirded out' by institutional surveillance (Roberts, Howell, Seaman, & Gibson, 2016, p. 8), have expressed support for informed consents processes (Roberts et al., 2016), are unaware of how their institution protects their privacy (Fisher, Valenzuela, & Whale, 2014), and argue that they should be able to limit data sharing for learning analytics (Ifenthaler & Schumacher, 2016).

Pushing forward with learning analytics *without* considering student privacy preferences—or ignoring such preferences all together—is foolhardy and morally suspect. I will not go as far to say that privacy-lite learning analytics initiatives are meant to do harm, in fact they are most likely well-intentioned but misplaced paternalistic actions (Jones, 2017). However, not considering student privacy preferences runs counter to norms of respecting individual autonomy and expressions thereof in choice making. In the long run, neglecting the emerging student voice weakens the foundation on which learning analytics are being developed (Beattie et al., 2014; Roberts et al., 2016). The question, then, is how to pursue informed consent mechanisms.

Informed consent in an age of big data

Big Data practices that disclose and capture data and information *across* contexts pose significant problems for informed consent. The volume of data and constant evolution of information flows makes it nigh impossible to effectively deploy informed consent mechanisms. Any hope that one's identity is protected by anonymization practices is dashed by the fact that aggregating enough data can tell tales about one's identity in ways that allow powerful actors to 'control and steer' individuals even without knowing their full identity (Gutwirth & De Hert, 2008, p. 289 in Barocas & Nissenbaum, 2014). Standard informed consent mechanisms cannot comprehensively detail the relationship between the data subject and the data miner, nor can they fully capture the attributes that characterize data and information flows; as such, their efficacy is limited (Barocas & Nissenbaum, 2014). However, there is still some hope for informed consent within some contexts—including higher education.

Big Data information flows are hard to track and manage. They create a web of connections between a variety of actors and entities in ways that often ignore norms, disregard transmission principles, and do not heed contextual values. But in universities,

flows of student information *are* trackable, manageable, and—when given proper care—can maintain harmony with extant norms. The central problem is that higher education institutions have not evolved their identity infrastructures while building capacity for data warehousing and analytics. Universities need to advance these infrastructures before they can begin to educate students about the purposes of identifiable data flows and support student privacy preferences.

Maximizing the identity layer

If the goal is to promote student choice over how their identifiable data flows, to whom, under particular conditions, and towards specific ends, then the first step is to clearly attribute data to students. Once these connections are accurately made, students will have the opportunity to express their choice over how their data flows using technical means.

Some will argue that this is a poor starting point. They may state that identifiable data should not be gathered for learning analytics purposes without student consent in the first place. While this position has its merits, it is untenable. Institutions *do* need identifiable data for legitimate business and educational purposes. But more importantly, the default state of institutional infrastructures is to identify students, authenticate their credentials, and use those credentials to authorize access to a variety of systems.

Identity management technologies, such as active directory services and single sign-on protocols, serve as the gatekeepers to student information systems, online learning applications, and to a campus's networks, among many other systems (Bruhn, Gettes, & West, 2003). These identity management systems create an identity layer in campus data infrastructures that connects identifiable students to flows of data and information. The default state of identification presents a significant opportunity to enhance the identity layer by adding on protocols that enable the expression of privacy preferences and forcing systems to respect such preferences downstream. The Platform for Privacy Preferences (P3P) protocol serves as a model for maximizing the existing identity layer.

The platform for privacy preferences (P3P) model

The World Wide Web Consortium (W3C) developed the Platform for Privacy Preferences (P3P) protocol in the early 2000s (W3C 2007). About the protocol, Lorrie Cranor (2003)—one of the lead architects of P3P—writes:

> [P3P] specifies a standard computer-readable format for Web site privacy policies. P3P-enabled Web browsers read policies published in P3P format and compare them with user-specified privacy settings. Thus, users can rely on their agents to read and evaluate privacy policies on their behalf. Furthermore, the standardized multiple-choice format of P3P policies facilitates direct comparisons between policies and the automatic generation of standard-format human-readable privacy notices. (p. 50)

Lawrence Lessig (2006) generally describes P3P as a machine-readable protocol that enables technologies to communicate, assess, and respect individual privacy choices set in applications and digital tools. Users set their privacy preferences in their web browser; the browser, acting as the agent, interprets the privacy policies of the website; and the browser then determines whether or not the website respects users' privacy

preferences (Cranor, Egelman, Sheng, McDonald, & Chowdhury, 2008). When the policies are congruent with the preferences, the user engages with the website; but when the two are incongruent, the browser warns the user of the privacy preference mismatch, blocks the cookies, and requests user input for how to proceed. Researchers also expanded P3P to improve privacy policy accessibility using simplified language and browsable matrices, including standardized 'nutrition label' notices to transform privacy policies into intelligible, actionable information for users (Kelley, Bresee, Cranor, & Reeder, 2009).

P3P ultimately failed. Major web companies, such as Google, ended up routing around user privacy preferences with hacks and browsers, like Microsoft's Internet Explorer (IE), never fully embraced the P3P protocol (Fulton, 2012). And even though IE did have some P3P capabilities, anecdotal evidence suggests that users were not fully aware of the privacy-enhancing capabilities (See Cranor, 2012a at footnote 38). Reflecting on the demise of P3P, Cranor (2012b) writes that a major reason for the low adoption rate of P3P stemmed from the fact that P3P was an optional, self-regulatory privacy standard without any teeth; there was simply little to no incentive to respect users' privacy preferences. The protocol, however, was a technical achievement. It proved that individuals could set privacy preferences, web applications could communicate their privacy policies in intelligible ways, and users would be the final arbiters in choosing whether or not to disclose information about themselves.

We can imagine scenarios where P3P technology could regulate the flow of student information according to student expectations for learning analytics. For instance, in an eAdvising system that uses geolocation tracking to determine student interactions with learning spaces (e.g., libraries, writing and other tutoring centers), students may wish for either the data to not be retained at all or for such information to remain undisclosed to their advisors. Informed by P3P technology, that information would be held securely within the data warehouse and remain undisclosed to this particular actor type. Similarly, students may be ok with the disclosure of identifiable learning management system interaction data to instructors, but with the limitation that such data does not include their IP address. The P3P would interpret these rules, disclose the appropriate data, and withhold the restricted data accordingly.[11]

Very little work has been done to date to capitalize on the existing identity layer to build P3P-like protocols; this is especially true for the United States. The work that has been accomplished has centered in Europe. Cooper and Hoel (2015) highlight Norwegian education, which at a national level adopted Feide, a federated identity management system, for use in primary, secondary, and higher education institutions. According to their report, '[the university] … register [s] and authenticate [s] their members[, and] the service providers define their access rules' (p. 53). After the implementation of Feide, Connect, an interoperability layer, was added to enable secure data transfer using standardized APIs and support the expression of privacy preferences. When Norwegian students initiate relationships with third-party service providers through Connect, they voluntary consent to particularized data practices but retain the right to opt out. If students choose to opt out, the service provider is directed to delete identifiable data. The university may, as well, require students to consent to certain service providers and their data practices.

Building student privacy preferences into data dashboards

A Platform for Privacy Preferences (P3P) protocol provides the means by which student privacy preferences are respected, but it does not enable the process of *informing* students of information practices nor the ability to *consent* to such practices by setting privacy preferences. For that to occur, student privacy dashboards need to be built (like Connect), which can be integrated into existing data dashboards.

As previously mentioned, learning analytics technology shares its statistical findings and predictions with *institutional actors* through visualizations (e.g., charts, trend lines, etc.). But, in order to promote self-awareness and encourage reflection among *learners*, some proponents of learning analytics advocate for creating data dashboards specifically for students (Clow, 2012; Duval et al., 2012). Data dashboards enable self-management over learning, and they also serve as a model for how informed consent could be improved.

Improving existing data dashboards with privacy preference settings would provide a central location where students would be informed about information practices that use their data and give them opportunities to opt out of personal data flows, possibly at a granular level. Privacy-promoting dashboards could include improved matrices and so-called nutrition label privacy policies, like what was developed for P3P. With such applications, students could learn about identifiable data flows and the ends to which they are put, dictate how they are informed (e.g., e-mail or text) about new data flows, and use toggle-like switches to determine what aspects of their information and data should be used for very specific purposes. Furthermore, privacy elements of data dashboards could archive and provide simple access to relevant information policies, as well as important communications from their institution regarding privacy concerns (e.g., data breaches).

Foregrounding norms, values, and expectations

While student data dashboards with privacy preference setting affordances empower students control over their information, they also benefit universities. In some cases, institutions will need to set defaults that allow for particular types of information flows. In order to achieve these ends, higher education administrators should have the ability to turn off and on some student data controls, or deny certain choices altogether. Thus, we arrive at an important point: What justifies overriding student privacy preferences?

Like other Big Data practices, 'the purposes to which data is being put [for learning analytics], who will have access to it, and how identities are being protected' remains opaque to students (Sclater, 2014, p. 20). Opaque information practices breed distrust, interfere with the development of interpersonal relationships, and motivate individuals to guard information about themselves. So, we can expect that if higher education institutions continue to obfuscate and hide how they use student data for learning analytics, student backlash is likely to occur that will harm the progress of educational data-mining initiatives. When data dashboards inform students about how their institution uses their data and for what purposes, harmful opacity will be reduced by lessening concerns about worrisome abuses brought about by analytics and trust will remain in the 'tripartite relationship between learner, teacher and educational institution' (Beattie et al., 2014, p. 424).

Students will generally expect their university to use standard academic information and some personal information about them in order to administer instruction, provide resources, and operate the institution, among other things. However, the literature

suggests that learning analytics are pushing—if not exceeding—norm boundaries in ways that make students uncomfortable with emerging data practices. Surveilling physical and digital student behaviors, for instance, are practices that do not track with normative expectations, nor are they clearly justifiable. These situations highlight when institutions have an opportunity to use data dashboards to inform students about the motivations behind edge-case learning analytics and seek consent. Students can then respond to institutional justifications by setting their privacy preferences in a data dashboard.

To maximize the utility of data dashboards built to support privacy, institutional efforts have to be made to educate students about the motivations driving educational data-mining practices and demonstrate how such practices are in alignment with the norms, values, and expectations of higher education. One such way to facilitate student privacy preferences *and* enable institutions to argue for more or less restrictions on student information flows is to embed a justified choice architecture into the dashboard (see Thaler, Sunstein, & Balz, 2012). Choice architecture would 'nudge' students towards particular privacy choices; at the same time, institutions could set default choices with a justifiable argument for why a particular choice is preferable. If it is the expectation that dashboards capture *everything* about how student data and information will be used and to what ends, dashboards will be unwieldy and overwhelm students with too many communications, and effectively void the usefulness of this informed consent mechanism. Justified choice architecture works against this particular problem.

Conclusion

In this article, I presented a position that learning analytics highlight existing privacy issues and present new ones related to students' inability to control how institutions use data and information about themselves. By improving the existing technical identity layer with P3P technology and creating privacy dashboards that enable student privacy preference setting, I argued that 1) students will be more fully informed about how their institution uses identifiable data and information and to what ends, and 2) will gain purposeful controls over information flows. This proposed model of informed consent ultimately works to support student privacy and autonomy.

Some readers of this article may disagree with my conception of privacy-as-control, and I agree that there are a number of other fruitful ways to address student privacy as it relates to learning analytics (see Heath, 2014). However, I argue that the central question regarding student control over identifiable data remains crucial, especially given the increasingly sensitive ways that institutions use Big Data practices to direct and intervene in student lives. If individual autonomy is something we value in a society that espouses liberalism, we need to consider ways to support autonomy—informed consent is one such way.

An additional counterargument that this article may raise concerns the position that institutions do not need to seek informed consent at all. Some may argue that legal frameworks (e.g., FERPA) and regulatory processes (e.g., institutional review of research) nullify this obligation or already account for the potential harms. However, FERPA's 'legitimate educational interest' loophole, which allows for nearly unfettered data aggregation, analysis, and disclosure to 'school officials' (institutional actors and, often, educational technology companies), requires no informed consent practices.

Additionally, institutional review boards (IRBs) often view learning analytics projects as forms of assessment, program evaluation, or operational research; IRBs do not need to review these projects and do not require informed consent. Consequently, universities grant themselves an 'ethical review waiver' (Griffiths, 2017, p. 559). In summary, the structures in place are not motivating institutional actors to develop informed consent mechanisms (see Willis, Slade,, & Prinsloo, 2016). Inaction with regard to informed consent is not justifiable. Failing to develop some way of procuring consent, using either the model I proposed or otherwise, signals disrespect for students to live their lives according to their own values and in support of their interests.

The work I presented in this article is a conceptual model, so its efficacy is unknown and is inherently limited. Next, human computer-interaction researchers and interface designers could test the feasibility and potential impact of the model by building mock interfaces that simulate information controls. Using students as research participants, data should be gathered to, among other things, determine student perceptions of such controls, how perceptions fluctuate based on data and information type and source, and test student reactions to various messages from institutions justifying data and information uses along with default settings. Additionally, systems developers should investigate the technical construction of existing institutional identity layers to determine whether or not these layers are adaptable to enable student information controls. This work could benefit from multi-institution investigations supported by higher education information technology organizations, such as EDUCAUSE and the Coalition for Networked Information. At the least, if colleges and universities find the model presented in this article to be worthwhile, they should review current systems to determine if they enable student privacy controls, and they should prioritize working with vendors of technologies who build such controls into their applications.

Endnotes

[1]I refer to 'Big Data' as a socio-technical phenomenon, like boyd and Crawford, and not *just* a large dataset. Therefore, I use the singular form. When I write of 'data' generally, I use the plural form

[2]Seven hundred twenty thousand is the estimated combined total enrollment of all Unizin member institutions, which was calculated by adding enrollment numbers from Carnegie Classification institution profiles and fact books for systems.

[3]For examples of a variety dashboard designs within and outside of LMSs, see Park and Jo 2015; Verbert, Duval, Klerkx, Govaerts, and Santos 2013; Verbert et al. 2013

[4]Brightspace is a suite of learning technology applications formerly known as Desire2-Learn. Desire2Learn rebranded itself in July of 2014, simplifying its corporate name to D2L and rebranding its products under the Brightspace moniker (Schaffhauser, 2014).

[5]Strauss (2017) provides an overview of the variety of personal information that the College Board, the owner and distributor of the ACT and PSAT tests, gathers as part of the assessment process. What is perhaps more concerning than the acts of *gathering* and *disclosing* this information is the fact that the College Board *sells* the information to third parties—including colleges and universities—which added to $834 million in net assets in 2015 (Dudley, 2016; Rivard, 2014).

[6]My thanks to CM for sharing his thoughts on this point.

[7]A simple Google search (university AND ownership "application for admission" site: .edu) will return many ownership policies regarding admissions materials.

[8]Instructors informed by the Quality Matters rubric are required by standard 6.5 to provide links to 'privacy policies for all external tools required in the course.' In a separate study of over 7000 syllabi for online library and information science courses, I saw this standard applied just once.

[9]As a federal regulation, FERPA provides students privacy rights and demarcates institutional responsibilities and privileges with regard to student information.

[10]As documented in Jones and Salo (2018), libraries are tracking student interactions with library resources in order to intervene when library usage is low. These actions are based on the assumption that the act of intervening will lead to more library use and increased learning outcomes, which may not be the case: Students may increase their library use to decrease the burden of surveillance by librarians and instructors, but not towards the end of improving their learning.

[11]Even if the student wishes for some data to remain undisclosed, such preferences would of course be nullified if a judicial order or subpoena were to be issued in order to gain access to identifiable student data (see Rubel and Jones, 2016). My thanks to AA for reminding me of this important point.

Abbreviations

LMS: Learning Management System; P3P: Platform for Privacy Preferences; SISs: Student Information Systems; W3C: World Wide Web Consortium

Acknowledgements

I thank the various private viewers who provided input and copyediting, which undoubtedly made this a better paper. I would also like to thank my research assistants, Laura Summers and Meredith Kostek, for their support. Finally, I drafted some of this work when I was finishing up my doctoral research, which Kristin Eschenfelder and Alan Rubel reviewed. As always, I am thankful to them for their guidance.

Originality and review status

The author confirms that this work is original and has not been published elsewhere, nor is it currently under consideration for publication elsewhere.

Author's contributions

The author is solely responsible for this article. The author read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Competing interests

The author declares that he has no competing interests.

Received: 18 February 2019 Accepted: 3 June 2019
Published online: 02 July 2019

References

Acquisti, A. (2004). Privacy in electronic commerce and economics of immediate gratification. In *Proceedings of the ACM Electronic Commerce Conference, USA*, (pp. 21–29). https://doi.org/10.1145/988772.988777.

Alcorn, S. (2013). *Facial recognition in the classroom tells teachers when students are spacing*. Fast Company Retrieved from http://www.fastcoexist.com/3018861/facial-recognition-in-the-classroom-tells-teachers-when-students-are-spacing.

Alfino, M., & Mayes, G. R. (2003). Reconstructing the right to privacy. *Social Theory and Practice*, *29*(1), 1–18 Retrieved from http://www.jstor.org/stable/23559211.

Andrejevic, M., & Gates, K. (2014). Big data surveillance: Introduction. *Surveillance & Society*, *12*(2), 185–196 Retrieved from https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/bds_ed.

Arriba Pérez, F., Santos, J. M., & Rodriguez, M. C. (2016). Analytics of biometric data from wearable devices to support teaching and learning activities. *Journal of Information Systems Engineering & Management*, *1*(1), 41–54. https://doi.org/10.20897/lectito.201608.

Barocas, S., & Nissenbaum, H. (2014). Big data's end run around anonymity and consent. In J. Lane, V. Stodden, S. Bender, & H. Nissenbaum (Eds.), *Privacy, big data, and the public good: Frameworks for engagement*, (pp. 44–75). New York: Cambridge University Press.

Beattie, S., Woodley, C., & Souter, K. (2014). Creepy analytics and learner data rights. In *Proceedings of Rhetoric and Reality: Critical Perspectives on Educational Technology, NZ*, (pp. 421–425). Retrieved from http://research.moodle.net/84/1/69-Beattie.pdf.

Bloustein, E. J. (1964). Privacy as an aspect of human dignity: An answer to dean Prosser. *New York University Law Review*, *39*(1964), 962–1007.

Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, *15*(5), 662–679. https://doi.org/10.1080/1369118X.2012.678878.

Brandimarte, L., Acquisti, A., & Loewenstein, G. (2013). Misplaced confidences: Privacy and the control paradox. *Social Psychology and Personality Science*, *4*(3), 340–347. https://doi.org/10.1177/1948550612455931.

Brazy, D. (2010). Ariz. College to position sensors to check class attendance. The Badger Herald Retrieved from http://badgerherald.com/news/2010/05/04/ariz-college-to-posi/.

Brown, M., Dehoney, J., & Millichamp, N. (2015). The next generation digital learning environment: A report on research. In *EDUCASE Learning Initiative (ELI)* Retrieved from https://library.educause.edu/resources/2015/4/the-next-generation-digital-learning-environment-a-report-on-research.

Bruhn, M., Gettes, M., & West, A. (2003). Identity and access management and security in higher education. *Educause Quarterly*, *26*(4), 12–16 Retrieved from https://er.educause.edu/~/media/files/articles/2003/10/eqm0342.pdf?la=en.

Buckingham Shum, S., & Ferguson, R. (2012). Social learning analytics. *Educational Technology & Society*, *15*(3), 3–26. Retrieved from https://www.jstor.org/stable/jeductechsoci.15.3.3.

Caldwell, T. (2012). More college students may be asked to declare sexual orientation. The New York Times Retrieved from http://thechoice.blogs.nytimes.com/2012/03/14/sexual-orientation-university-of-california/.

California State University Long Beach. (2014). (EAB) Predictive analytics. Retrieved from https://web.archive.org/web/20170819031412/http://web.csulb.edu/depts/enrollment/staff_reference/e_advising/index.html.

Campus Labs. (2014). Case study: Northern Arizona University–Using data to improve student retention. Retrieved from Internet Archive website: https://web.archive.org/web/20150613222812/, http://www.campuslabs.com/pdf/caseStudy-NAU.pdf.

Clark, R. (1987). Information technology and dataveillance. Retrieved from http://rogerclarke.com/DV/CACM88.html.

Clow, D. (2012). The learning analytics cycle: Closing the loop effectively. In *Proceedings of the Second International Conference on Learning Analytics and Knowledge, USA*, (pp. 134–138). https://doi.org/10.1145/2330601.2330636.

Connelly, R. J. (2000). Intentional learning: The need for explicit informed consent in higher education. *The Journal of General Education*, *49*(3), 211–230 Retrieved from http://www.jstor.org/stable/27797469.

Cooper, A., & Hoel, T. (2015). *Data sharing requirements and roadmap: Deliverable 7.2 (Report)*. Retrieved from Learning Analytics Community Exchange website: http://www.laceproject.eu/deliverables/d7-2-data-sharing-roadmap.pdf.

Cowley, S., Bernard, T. S., & Hakim, D. (2017). Equifax breach prompts scrutiny, but new rules may not follow. The New York Times Retrieved from https://www.nytimes.com/2017/09/15/business/equifax-data-breach-regulation.html.

Cranor, L. F. (2003). P3P: Making privacy policies more useful. *IEEE Security & Privacy*, *99*(6), 50–55. https://doi.org/10.1109/MSECP.2003.1253568.

Cranor, L. F. (2012a). Necessary but not sufficient: Standardized mechanisms for privacy notice and choice. *Journal on Telecommunications and High Technology Law*, *10*(2), 273–308 Retrieved from http://www.jthtl.org/content/articles/V10I2/JTHTLv10i2_Cranor.pdf.

Cranor, L. F. (2012b). P3P is dead, long live P3P! In *This Thing [web log]* Retrieved from http://lorrie.cranor.org/blog/2012/12/03/p3p-is-dead-long-live-p3p/.

Cranor, L. F., Egelman, S., Sheng, S., McDonald, A. M., & Chowdhury, A. (2008). P3P deployment on websites. *Electronic Commerce Research and Applications*, *7*(3), 274–293. https://doi.org/10.1016/j.elerap.2008.04.003.

Crawford, K., & Schultz, J. (2014). Big data and due process: Toward a framework to redress predictive privacy harms. *Boston College Law Review*, *55*(1), 93–128.

Danezis, G., Domingo-Ferrer, J., Hansen, M., Hoepman, J.-H., Le Metayer, D., Tirtea, R., & Schiffner, S. (2014). *Privacy and data protection by design: From policy to engineering (report)*. Retrieved from the Publications Office of the European Union website: https://doi.org/10.2824/38623.

Dawson, S., Heathcote, L., & Poole, G. (2010). Harnessing ICT potential: The adoption and analysis of ICT systems for enhancing the student learning experience. *International Journal of Educational Management*, *24*(2), 116–128. https://doi.org/10.1108/09513541011020936.

Denley, T. (2012). Advising by algorithm. The New York Times Retrieved from http://www.nytimes.com/interactive/2012/07/18/education/edlife/student-advising-by-algorithm.html.

Diaz, V., Brown, M. (2012). Learning analytics: A report on the ELI focus session (report ELI paper no. 2). Retrieved from EDUCAUSE Learning Initiative website: https://library.educause.edu/resources/2012/5/learning-analytics-a-report-on-the-eli-focus-session.

Dudley, R. (2016). College board faces rocky path after CEO pushes new vision for SAT. Reuters Retrieved from http://www.reuters.com/investigates/special-report/college-sat-coleman/.

Duval, E., Klerkx, J., Verbert, K., Nagel, T., Govaerts, S., Parra, G., … Vandeputte, B. (2012). Learning dashboards and learnscapes. In *Proceedings of CHI 2012, USA*, (pp. 1–5) Retrieved from https://lirias.kuleuven.be/bitstream/123456789/344525/1/eist2012_submission_6-2.pdf.

Family Educational Rights and Privacy Act, 34 C.F.R. § 99 (1974).

Family Policy Compliance Office. (2011). The family educational rights privacy act: Guidance for reasonable methods and written agreements. Retrieved from Department of Education Family Policy Compliance Office website: https://www2.ed.gov/policy/gen/guid/fpco/pdf/reasonablemtd_agreement.pdf.

Fisher, J., Valenzuela, F.-R., Whale, S. (2014). *Learning analytics: A bottom-up approach to enhancing and evaluating students' online learning* (Report). Retrieved from Australian Government Office for Learning and Teaching website: https://rune.une.edu.au/web/handle/1959.11/15261.

Flaherty, D. H. (1999). Visions of privacy: Past, present, and future. In C. J. Bennett, & R. Grant (Eds.), *Visions of privacy: Policy choices for the digital age*, (pp. 19–38). Toronto: University of Toronto Press.

Fried, C. (1968). Privacy. *Yale Law Journal*, *77*(3), 475–493. https://doi.org/10.2307/794941.
Froomkin, A. M. (2000). The death of privacy? *Stanford Law Review*, *52*(5), 1461–1543. https://doi.org/10.2307/1229519.
Fulton, S. (2012, February 23). Expert: Microsoft's P3P 'ineffective,' Google's privacy bypass unhelpful. *ReadWrite*. Retrieved from https://readwrite.com/2012/02/23/expert-microsofts-p3p-ineffect/
Glass, A. (2013). *The state of higher education 2013 (Report)*. Retrieved from OECD website: https://www.oecd.org/edu/imhe/thestateofhighereducation2013.htm.
Goff, J. W., & Shaffer, C. M. (2014). Big data's impact on college admission practices and recruitment strategies. In J. E. Lane (Ed.), *Building a smarter university: Big data, innovation, and analytics*, (pp. 93–120). Albany: SUNY Press.
Goldman, J. (1999). Privacy and individual empowerment in the interactive age. In C. J. Bennett, & R. Grant (Eds.), *Visions of privacy: Policy choices for the digital age*, (pp. 96–115). Toronto: University of Toronto Press.
Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Journal of Educational Technology & Society*, *15*(3), 42–57.
Griffiths, D. (2017). An ethical waiver for learning analytics? In *12th European Conference on Technology Enhanced Learning (EC-TEL 2017), Estonia*, (pp. 557–560). https://doi.org/10.1007/978-3-319-66610-5_62.
Gutwirth, S., & De Hert, P. (2008). Regulating profiling in a democratic constitutional state. In M. Hildebrandt, & S. Gutwirth (Eds.), *Profiling the European citizen: Cross disciplinary prospectives*, (pp. 271–302). Dordrecht: Springer.
Haggerty, K. D., & Ericson, R. V. (2000). The surveillant assemblage. *The British Journal of Sociology*, *51*(4), 605–622. https://doi.org/10.1080/00071310020015280.
Heath, J. (2014). Contemporary privacy theory contributions. *Journal of Learning Analytics*, *1*(1), 140–149 Retrieved from http://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/3339.
Hier, S. P. (2003). Probing the surveillant assemblage: On the dialectics of surveillance practices as processes of social control. *Surveillance & Society*, *1*(3), 399–411 Retrieved from https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/3347/3309.
Ho, C. (2011). Companies tracking college athletes' tweets, Facebook posts go after local universities. The Washington Post Retrieved from http://www.washingtonpost.com/business/capitalbusiness/companies-tracking-college-athletes-tweets-facebook-posts-go-after-local-universities/2011/10/10/gIQAyHZ9oL_story.html.
Hoel, T., & Chen, W. (2016). The principle of data protection by design and default as a lever for bringing pedagogy into the discourse on learning analytics. In *Proceedings of the 24th conference on computers in education, India*, (pp. 113–121). Retrieved from http://topicmaps.estandard.no/files/Hoel_Chen_LAEDM_ICCE16_final.pdf.
Hoover, E. (2011). Elmhurst college will ask applicants about sexual orientation. The Chronicle of Higher Education Retrieved from http://chronicle.com/blogs/headcount/elmhurst-college-will-ask-applicants-about-sexual-orientation/28553.
Hoover, E. (2012). Facebook meets predictive analytics. The Chronicle of Higher Education Retrieved from http://chronicle.com/blogs/headcount/facebook-meets-predictive-analytics/32770.
Hui, K., & Png, I. P. L. (2006). The economics of privacy. In T. Hendershott (Ed.), *Handbook of information systems and economics*, (pp. 471–497). North Holland: Elsevier Science.
Ifenthaler, D., & Schumacher, C. (2016). Student perceptions of privacy principles for learning analytics. *Education Technology Research and Development*, *64*(5), 923–938. https://doi.org/10.1007/s11423-016-9477-y.
Johnson, L., Adams Becker, S., Estrada, V., & Freeman, A. (2015). *The NMC horizon report: 2015 higher education edition*. Austin: The New Media Consortium.
Jones, K. M. L. (2017). Learning analytics and its paternalistic influences. In P. Zaphiris & A. Ioannou (Eds.). *Lecture Notes in Computer Science, Learning and Collaboration Technologies: Technology in Education (LCT 2017, HCI International 2017)* (pp. 281-292). Cham, SZ: Springer. https://doi.org/10.1007/978-3-319-58515-4_22
Jones, K. M. L. & Salo, D. (2018). Learning analytics and the academic library: Professional ethics commitments as a crossroads. *College & Research Libraries*, 79(3), 304–323. Retrieved from https://crl.acrl.org/index.php/crl/article/view/16603
Kant, E. (1785). *Groundwork of the metaphysics of morals*. Retrieved from http://www.gutenberg.org/ebooks/5682
Kelley, P. G., Bresee, J., Cranor, L. F., & Reeder, R. W. (2009). *A "nutrition label" for privacy. Proceedings of the 5th Symposium on Usable Privacy and Security, USA*, 1–12. https://doi.org/10.1145/1572532.1572538
Kolowich, S. (2013). The new intelligence. Insider Higher Ed Retrieved from http://www.insidehighered.com/news/2013/01/25/arizona-st-and-knewtons-grand-experiment-adaptive-learning.
Lessig, L. (2006). *Code: Version 2.0*. New York: Basic Books.
Lewis, B. (2011). New initiative advances ASU's effort to enhance student success. ASU News Retrieved from https://asunews.asu.edu/20111012_eAdvisor_expansion.
Littleworth, R., & Qazi, A. (2017). The power of a higher education consortium. *Educause Review* Retrieved from http://er.educause.edu/articles/2017/8/the-power-of-a-higher-education-consortium.
Locke, J. (1689). Second treatise on civil government. Retrieved from http://press-pubs.uchicago.edu/founders/documents/v1ch16s3.html.
Long, P., & Siemens, G. (2011). Penetrating the fog: Analytics in learning and education. *Educause Review*, *46*(5), 30–40 Retrieved from https://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-and-education.
Lyon, D. (2014). Surveillance, Snowden, and big data: Capacities, consequences, critique. *Big Data & Society*, *1*(2), 1–13. https://doi.org/10.1177/2053951714541861.
Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an 'early warning system' for educators: A proof of concept. *Computers and Education*, *54*(2), 588–599. https://doi.org/10.1016/j.compedu.2009.09.008.
Marr, B. (2015). Barbie wants to chat with your child—But is big data listening in? Forbes Retrieved from https://www.forbes.com/sites/bernardmarr/2015/12/17/barbie-wants-to-chat-with-your-child-but-is-big-data-listening-in.
Marx, G. T. (1999). Ethics for the new surveillance. In C. J. Bennett, & R. Grant (Eds.), *Visions of privacy: Policy choices for the digital age*, (pp. 39–67). Toronto: University of Toronto Press.
Mayer-Schönberger, V., & Cukier, K. (2013). Big data: A revolution that will transform how we live, work, and think. New York: Houghton Mifflin Harcourt.
W3C. (2007). Platform for Privacy Preferences (P3P) project. Retrieved from https://www.w3.org/P3P/
Mayer-Schönberger, V., & Cukier, K. (2014a). *Learning with big data: The future of education*. New York: Houghton Mifflin Harcourt.

Mayer-Schönberger, V., & Cukier, K. (2014b). *Your high school transcript could haunt you forever: How big data could create an inescapable "permanent record."*. The Atlantic Retrieved from https://www.theatlantic.com/education/archive/2014/03/your-high-school-transcript-could-haunt-you-forever/284346/.

Moore, A. D. (2010). *Privacy rights: Moral and legal foundations*. University Park, Pennsylvania: The Pennsylvania State University press.

National Center for Education Statistics. (2013). Digest of education statistics: 2012. Retrieved from National Center for Education Statistics website: http://nces.ed.gov/programs/digest/d12/.

National Research Council (2007). *Engaging Privacy and Information Technology in a Digital Age*. Washington, DC: National Academies Press. https://doi.org/10.17226/11896.

Nissenbaum, H. (2010). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford, CA: Stanford University Press.

Norris, D. (2011). *7 things you should know about first-generation learning analytics (report)*. Retrieved from EDUCAUSE website: https://library.educause.edu/resources/2011/12/7-things-you-should-know-about-firstgeneration-learning-analytics.

O'Connor, M. C. (2010). Northern Arizona University to use existing RFID student cards for attendance tracking. *RFID Journal* Retrieved from http://www.rfidjournal.com/articles/view?7628.

Ohm, P. (2010). Broken promises of privacy: Responding to the surprising failure of anonymization. *UCLA Law Review*, *57*(6), 1701–1777 Retrieved from http://www.uclalawreview.org/pdf/57-6-3.pdf.

Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, *45*(3), 438–450. https://doi.org/10.1111/bjet.12152.

Park, Y., & Jo, I-H. (2015). Development of the learning analytics dashboard to support students' learning performance. *Journal of Universal Computer Science*, 21(1), 110–133. https://doi.org/10.3217/jucs-021-01-0110

Parry, M. (2012). Big data on campus. The New York Times Retrieved from http://www.nytimes.com/2012/07/22/education/edlife/colleges-awakening-to-the-opportunities-of-data-mining.html.

Pérez-Peña, R. (2015). Students gain access to files on admission to Stanford. The New York Times Retrieved from http://www.nytimes.com/2015/01/17/us/students-gain-access-to-files-on-admission-to-stanford.html.

Piazza. (n.d.). To infinity and beyond. Retrieved from https://web.archive.org/web/20150424161953/https://recruiting.piazza.com/insights.

Piazza Careers. (n.d.). Transforming college recruiting. Retrieved from https://recruiting.piazza.com.

Prinsloo, P., & Slade, S. (2017). An elephant in the learning analytics room: The obligation to act. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference, Canada*, (pp. 46–55). https://doi.org/10.1145/3027385.3027406.

Qazi, A. (2017). Data management: The fundamentals [blog post]. Retrieved from http://unizin.org/2017/06/06/data-management-the-fundamentals/.

Raths, D. (2016). Indiana CIO Brad Wheeler provides update on Unizin. Campus Technology Retrieved from https://campustechnology.com/articles/2016/07/18/indiana-cio-brad-wheeler-provides-update-on-unizin.aspx.

Reed, S. (2013). Four areas of collegiate student-athlete privacy invasion. *Communication & Sport*. https://doi.org/10.1177/2167479513510910.

Reiman, J. H. (1976). Privacy, intimacy, and personhood. *Philosophy and Public Affairs*, *6*(1), 26–44 Retrieved from http://www.jstor.org/stable/2265060.

Richards, N. (2015). *Intellectual privacy*. New York: Oxford University Press.

Richards, N. M. (2008). Intellectual privacy. *Texas Law Review*, *87*(2), 387–446.

Richards, N. M., & King, J. H. (2013). Three paradoxes of big data. *Stanford Law Review Online*, *66*, 41–46 Retrieved from https://www.stanfordlawreview.org/online/privacy-and-big-data-three-paradoxes-of-big-data/.

Rivard, R. (2014). Predicting where students go. Inside Higher Ed Retrieved from https://www.insidehighered.com/news/2014/09/19/colleges-now-often-rely-data-rather-gut-hunt-students.

Roberts, L. D., Howell, J. A., Seaman, K., & Gibson, D. C. (2016). Student attitudes toward learning analytics in higher education: "The Fitbit version of the learning world.". *Frontiers in Psychology*, *7*(1959), 1–11. https://doi.org/10.3389/fpsyg.2016.01959.

Roderick, L. (2014). Discipline and power in the digital age: The case of the US consumer data broker industry. *Critical Sociology*, *40*(5), 729–746. https://doi.org/10.1177/0896920513501350.

Rubel, A. & Jones, K. M. L. (2016). Student privacy in learning analytics: An information ethics perspective. *The Information Society*, 32(2), 143–159. https://doi.org/10.1080/01972243.2016.1130502

Schaffhauser, D. (2014). D2L unveils revamped platform 'Brightspace' with adaptive learning. *THE Journal* Retrieved from http://thejournal.com/articles/2014/07/14/d2l-intros-revamped-platform-brightspace-with-adaptive-learning.aspx.

Schiller, B. (2015). It'll be a lot harder to cut class with this classroom facial-recognition app. Fast Company Retrieved from http://www.fastcoexist.com/3042445/itll-be-a-lot-harder-to-cut-class-with-this-classroom-facial-recognition-app.

Sclater, N. (2014). *Code of practice for learning analytics: A literature review of the ethical and legal issues (Jisc report)*. Retrieved from Jisc website: http://repository.jisc.ac.uk/5661/1/Learning_Analytics_A-_Literature_Review.pdf.

Sclater, N. (2016). Developing a code of practice for learning analytics. *Journal of Learning Analytics*, *3*(1), 16–42. https://doi.org/10.18608/jla.2016.31.3.

Siemens, G. (2012). Learning analytics: Envisioning a research discipline and a domain of practice. In *Proceedings of the Second International Conference on Learning Analytics and Knowledge, USA*, (pp. 4–8). https://doi.org/10.1145/2330601.2330605.

Simon, S. (2012). Biosensors to monitor students' attentiveness. Chicago Tribune Retrieved from http://articles.chicagotribune.com/2012-06-12/news/sns-rt-us-usa-education-gatesbre85c018-20120612_1_gates-foundation-veteran-english-teacher-teachers-feedback.

Slade, S., & Galpin, F. (2012). Learning analytics and higher education: Ethical perspectives. In *Proceedings of the Second International Conference on Learning Analytics and Knowledge, USA*, (pp. 16–17). https://doi.org/10.1145/2330601.2330610.

Slade, S., & Prinsloo, P. (2014). Student perspectives on the use of their data: Between intrusion, surveillance and care. In *Proceedings of the European distance and E-learning network, UK*, (pp. 291–300) Retrieved from http://oro.open.ac.uk/41229/.

Solove, D. J. (2008). *Understanding privacy*. Cambridge, MA: Harvard University Press.

Solove, D. J. (2004). *The digital person:* Technology and privacy in the information age. New York, NY: New York University Press.

Spann, C. A., Schaeffer, J., & Siemens, G. (2017). Expanding the scope of learning analytics data: Preliminary findings on attention and self-regulation using wearable technology. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference, Canada*, (pp. 203–207) Retrieved from https://dl.acm.org/citation.cfm?id=3027427.

Steinberg, J. (2010). University of Pennsylvania tries outreach based on sexual orientation. The New York Times Retrieved from http://thechoice.blogs.nytimes.com/2010/02/26/penn/.

Strauss, V. (2017). How the SAT and PSAT collect personal data on students—and what the College Board does with it. The Washington Post Retrieved from https://www.washingtonpost.com/news/answer-sheet/wp/2017/03/30/how-the-sat-and-psat-collect-personal-data-on-students-and-what-the-college-board-does-with-it.

Tene, O., & Polonetsky, J. (2013). Big data for all: Privacy and user control in the age of analytics. *Northwestern Journal of Technology and Intellectual Privacy*, *11*(5), 239–273 Retrieved from http://scholarlycommons.law.northwestern.edu/njtip/vol11/iss5/1/.

Thaler, R. H., Sunstein, C. R., & Balz, J. P. (2012). Choice architecture. In E. Shafir (Ed.), *The behavioral foundations of public policy*, (pp. 428–439). Princeton, NJ: Princeton University Press.

Unizin. (2015). About. Retrieved from http://unizin.org/about/.

van Barneveld, A., Arnold, K. E., Campbell, J. P. (2012). *Analytics in higher education: establishing a common language (report no. ELI3026)*. Retrieved from EDUCAUSE learning initiative website: https://library.educause.edu/resources/2012/1/analytics-in-higher-education-establishing-a-common-language.

Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist*, 57(10), 1500–1509. https://doi.org/10.1177/0002764213479363

Westin, A. F. (1967). *Privacy and freedom*. New York, NY: Atheneum.

Willis, J. E., III, Slade, S., & Prinsloo, P. (2016). Ethical oversight of student data in learning analytics: A typology derived from a cross-continental, cross-institutional perspective. *Educational Technology Research and Development*, *64*(5), 881–901. https://doi.org/10.1007/s11423-016-9463-4.

Publisher's Note