In Their Own Words:
Student Perspectives on Privacy and Library Participation in Learning Analytics Initiatives

*Kyle M. L. Jones, Michael R. Perry, Abigail Goben, Andrew Asher, Kristin A. Briney, M. Brooke Robertshaw, and Dorothea Salo**

Introduction

Colleges and universities continue to adopt networked information and communication technologies at an unprecedented rate, effectively enmeshing their campuses in an invisible layer of ubiquitous systems. The modern higher education institution relies on networked technologies to serve an array of needs, such as administering complex institutions, supporting interpersonal communications and scheduling, facilitating face-to-face and online learning experiences, and campus security, among other things. Each time faculty, staff, and students gain access to, click on resources within, and submit information to these systems, their actions and content are transformed into data; and since many institutional technologies require credentials for access, much of these data also identify individuals.

To maximize the analytic value that may be mined from an institution's data, colleges and universities have begun to strategically aggregate data previously stored in an array of application-specific databases into inclusive data warehouses. With these warehouses, higher education institutions have a stronger understanding of what data are available to them; consequently, they can develop insights into specific behaviors and outcomes—and craft interventions—using analytic strategies adopted from data science and mold them to serve educational and administrative needs. Such strategies fall under the umbrella term "learning analytics."

Learning analytics is defined as the "measurement, collection, analysis, and reporting of [student and other data] for the purposes of understanding and optimising learning and the environments in which it occurs."[1] With learning analytics, institutions claim to be more prepared to describe (what is happening?), diagnose (why did it happen?), and predict (what is likely to happen?) factors that influence, enhance, or impede student learning, as well as prescribe (what should we do about it?) data-based interventions.[2] Some learning analytics applications use descriptive statistics to illuminate student behaviors, while others employ predictive statistics to determine the possibility of a particular outcome. The contexts in which learning analytics applications and tools are found are varied. Instructors use learning analytics to visualize student progress in learning management systems using data dashboards.[3] Advisors may use predictive measures to determine if a student is unlikely to succeed in a course or program. Institutional researchers are building complex models to more finely measure and predict student retention and graduation rates.[4] And academic librarians are adopting learning analytics to

* *Kyle M. L. Jones is an Assistant Professor at Indiana University-Indianapolis (IUPUI), kmlj@iupui.edu; Michael R. Perry is Head of Assessment and Planning, Northwestern University, michael.perry@northwestern.edu; Abigail Goben is an Associate Professor, University of Illinois at Chicago, agoben@uic.edu; Andrew Asher is an Assessment Librarian, Indiana University-Bloomington, asherand@indiana.edu; Kristin A. Briney is a Data Services Librarian, University of Wisconsin-Milwaukee, briney@uwm.edu; M. Brooke Robertshaw is an Assistant Professor, Oregon State University, Brooke.Robertshaw@oregonstate.edu; and Dorothea Salo is a Faculty Associate, University of Wisconsin-Madison, salo@wisc.edu.*

identify relationships between information access, use, and services with learning outcomes.[5]

Academic library participation in learning analytics continues to increase due to the decades worth of pressure to justify budget expenditures and prove their value. With these conditions in mind, the Association for College and Research Libraries' "Value of Academic Libraries" report recommended a realignment of library practices and evaluation strategies with the demands of academic administrators and their accountability measures.[6] With this report and grants awarded to and managed by ACRL, the Assessment in Action program provided case studies and numerous research papers identifying how librarians were implementing the capture of learning analytics at their institutions.[7]

On the surface, the rapid flow of data and information within and across a campus, and the subsequent aggregation of data and information, may seem benign. Furthermore, requiring secure access to networks and their resources suggests good stewardship and security of vast information technology complexes. Few have so far questioned that campus systems are anything but good for furthering the educational mission of an institution and supporting its research aims. In this era of Big Data, higher education stakeholders might simply expect institutions to apply advanced analytic practices. Efficacy and benefits may arise from learning analytics at the small scale, though widespread and longitudinal evidence remains as yet unseen.[8] However, technological advancements raise serious questions about appropriate roles for these systems and how downstream uses of the aggregated data may lead to specific harms, including but not limited to autonomy-reducing effects created by digital surveillance, unjustifiable invasions of privacy due to data aggregation and dredging, and erosions in higher education's foundational values.

The literature is evolving to address multiple facets of student privacy, which is to be expected, but there still remains one major gap. The aggregation of identifiable data representing a student's personal and educational behaviors immediately raises access, use, and control concerns—all standard privacy issues. The extant research generally provides normative analyses and policy recommendations, which are insightful and useful to a degree. However, they overwhelmingly fail to consider the perspective of the primary targets of learning analytics: students.

There may be two reasons for this. One, it is plausible that most assume that students care about their privacy; therefore, privacy-positive arguments are correctly oriented. Second, and contrary to the first, it may be that student privacy research does not care about the student perspective, assuming it is underinformed, misinformed, and/or irrelevant. This gap in the literature needs research and the findings could be immensely useful for policy makers, system designers, administrators, faculty, librarians, and others involved in learning analytics technologies and practices.

To fill the identified gap in the literature, we have begun a three-year, three-phase, grant-funded project solely focused on identifying various aspects of student perspectives of their privacy in relation to learning analytics technologies and practices within higher education (broadly) and as employed by academic libraries (specifically). This paper presents some initial findings and emerging themes from phase one of this project, which interviewed 100-plus undergraduate students at eight higher education institutions.

Literature Review

A Student's Growing Digital Dossier

As selective abstractions, data only represent partial phenomenological aspects.[9] To some proponents of learning analytics, this is a weakness; limited data means limited insights. Data are pre-factual and pre-analytical, yet they serve as the foundation for analytic practices. Therefore, the reasoning goes, the more data is made accessible, the better the chances are that data analysis will lead to actionable insights by mining data stores.[10] To overcome this shortcoming, analytics advocates state that the mass aggregation of data is the only process by which

to maximize analytic insights and capture student life in exhaustive, fine-grained detail. To this end, a cottage industry of edTech companies has emerged to provide data analytic services to higher education institutions, and the latter are maturing their data infrastructures to support learning analytics efforts.

An institution's vast network of information systems can capture students' analog behaviors, digital interactions, social networks, physical characteristics (i.e., biomarkers), and a host of other data points representing their personal and academic lives. Students reveal much of this information themselves, primarily on their application for admission where they reveal personal, academic, and financial information. With each form they fill out or LMS course that they navigate during their higher education tenure, these *structured*, indexical data make it fairly easy to develop the beginning of a student's digital dossiers. But as *unstructured* data are combined, the dossier becomes more comprehensive and granular. Daniel Solove reminds us that each click reveals a student's behavior; each sensor interaction captures a location and moment in time; and each bit of data exhaust created as a byproduct of system interactions develops rich metadata to include in digital record describing student life—all of which adds to a student's dossier and serves as a dataset against which to run analytics and store predictions about a student's potential for success.[11] Given the increasing ability of institutions to capture and analytically intervene in student life, there has been a groundswell of popular and academic literature on student privacy.

Particularized Student Privacy Issues

For every piece extolling the virtues of big data in higher education, there is an article highlighting the potential privacy harms. Consider this pair of articles in *The Chronicle of Higher Education*. Brown highlights how each student is a potential test subject for learning analytics[12]; Blumenstyk counters that while this may be so, treating students thusly raises serious ethical concerns regarding geolocation tracking and surveillance practices.[13] Those arguing for the big data in higher education continue to beat on the drum of efficiency and effectiveness, asserting that data analytics will save sparse resources, help manage bureaucratic institutions, and demonstrate to stakeholders the relevance and value of a college degree.[14] While the popular literature is raising student privacy concerns, it tends to skim the surface of the issues, highlighting aspects that are "creepy" or "Orwellian" without digging into particularized concerns.[15]

Arguably, one of the reasons for a lack of cogent and direct writing on student privacy—and privacy generally—is that it is an unwieldy concept. While a privacy invasion may be recognizable when experienced by an individual, it is challenging to disentangle and understand its many facets concretely.[16] Also, given that privacy protections and expectations thereof are often context dependent, it is near impossible to establish a "unitary definition" that can serve as a foundation for policy, law, and social expectations.[17] Regardless of its ambiguity, a few definitions are generally accepted—and are challenged by learning analytics practices.

We argue that three approaches to privacy are useful for understanding student privacy, especially in relation to learning analytics practices. The first includes an individual's right to be let alone, or a right to protect oneself from intrusion.[18] The second addresses limited access to self, or the ability to seclude one's emotions, ideas, and behaviors from others.[19] Third, and finally, privacy is protected when individuals can control personal information. Westin's privacy definition in this category argues that individuals should be able to determine "when, how, and to what extent information about them is communicated to others."[20] We will address specific student privacy concerns below using these three privacy lenses.

A Student's Right to Be Let Alone

The advent of new technologies challenges current ideas of what privacy should be: So it was with the Eastman Kodak Company's camera in 1884, and so it is with learning analytics. Writing in response to Kodak's camera,

Warren and Brandeis stated that privacy is found "in the peace of mind or the relief afforded by the ability to prevent"[21] disclosure of personal information, especially "thoughts, sentiments, and emotions."[22] Whether or not students can exercise such a right partly depends on their ability to express agency with regard to personal information.

From application to admission through to graduation, students are increasingly losing the ability to find relief from data and information collection. Students are required to reveal sensitive details about their past life and future ambitions, in addition to a host of biographic information, simply to be *considered* for admission—they are never guaranteed anything in return for these information disclosures. Additionally, this may include the mining of both their social media as well as that of their peers, as admissions offices seek to review materials never intended by the candidate as a formal part of the application packet. And as they interact with information systems, many of which are compulsory to use as part of one's education, pulling out of the higher education information technology complex is less an option and more a privacy pipe dream. Williamson further reminds that learning analytics, as a form of digital governance, often require students to submit to analytics, change their behaviors, and act in ways that serve desirable ends set out by those who wield data infrastructures and algorithms—doing otherwise marks one as "suboptimal" in this sociotechnical system.[23]

If students lack the agency necessary to express their preference to be let alone, a question of coercion emerges. Playing with this question, Johnson considers the development of systems that support *strong* coercion (coercion with a penalty risk), writing:

> *Rarely are such systems outright coercive, but one could imagine developing such systems by, for instance, linking student activity data from a learning management system to financial aid awards. Rather than relying on end-of-semester grades, an institution might condition aid on keeping up on work performed throughout the semester: reading materials accessed, assignments completed, and so forth.*[24]

Prinsloo and Slade contend that *weak* coercion is the greater concern. While facing no clear risk of penalty, students are pressured to disclose information and respond to interventions, especially when they are in a position of vulnerability. They cite nudging messages and predictive measures, especially related to poor outcomes, as instances of weak coercion.[25] Others argue that if students wish to find relief from learning analytics and related coercive data practices, their only option may be to enroll at another institution where such practices are less invasive.[26] However, such an option does not exist for many students who lack certain financial resources and social privileges.

A Student's Right to Limit Access to Oneself

While closely related conceptually to a right to be let alone, a right to limit access to oneself presents a unique view. The former concerns *disclosure*, whereas the latter is concerned with an individual's ability "to decide for [oneself] to what extent they shall be the *subject* of public observation and discussion"[27] (emphasis added). The question then is whether or not students can protect themselves "from unwanted access by others—either physical access, personal information, or attention."[28] Addressing this question in a learning-analytics context also requires us to consider human and non-human actors alike.

One question is whether and how institutions are restricting access to students' digital dossiers, and additionally what rights students have to make these decisions for themselves. While data security measures do not guarantee privacy—for one could build a highly secure data system inclusive of privacy-invading information—they do provide barriers against unapproved access to private information. Unfortunately, "there are no

clearly defined best practices for the anonymisation [*sic*], reuse, storage and security of educational data"[29] Each institution is left to its own devices to develop appropriate policy and create or adopt technological systems that can account for differential access and audit information use by particular actors. Therefore, each institution decides how and to what degree students are able to limit access to themselves.

The issue of access-as-privacy goes beyond simply permitting Person A's access to Person B's private information to how individuals are made *subject* to observation and discussion. With learning analytics, student life is "made into" and "considered as" data for the sole purpose of description, analysis, and intervention.[30] Student data is aggregated, transformed into individual and group profiles—sometimes deidentified—and subjected to a battery of analytical tests. Administrators, institutional researchers, faculty, advisors, and increasingly librarians access the derived data to inform their respective practices. Third parties, or "school officials," often have access to student data and information to provide institutions' information services and systems.[31] When all these things are considered, it becomes clear that there is a considerable privacy issue where limited access is concerned.

A Student's Right to Control Information About Oneself

Privacy-as-control remains a dominant privacy perspective. On the one hand, personal control over information is a worthy goal because it empowers an individual to place restrictions on how information about oneself flows, to whom, and under what conditions. Information controls respect people's status as "autonomous individuals who can critically think and act to further their own good" (Johnson, 2017, p. 79). On the other hand, regardless of the underlying moral considerations, individuals come to expect control options because of the prominence of the Fair Information Practice Principles, or FIPPs, in federal and international law, organization policy, and terms of service agreements.[32] The principles include:

1. Transparency of record systems of personal data.
2. The right to notice about such record systems.
3. The right to prevent personal data from being used for new purposes without consent.
4. The right to correct or amend one's records.
5. Responsibilities on the holders of data to prevent its misuse.

Over time, these principles have been adopted and amended to explicitly provide a control right, which the FIPPs only imply. For instance, the 2012 Consumer Privacy Bill of Rights states that "Consumers have a right to exercise control over what personal data companies collect from them and how they use it."[33] A number of privacy theorists have suggested various reasons for why privacy-as-control is untenable,[34] especially since the emergence of big data practices,[35] but there still remains a normative expectation surrounding information controls and others suggest that the FIPPs remain relevant but are in need of review.[36] With regard to students and learning analytics, "a delicate balance between control and limits needs to be achieved"[37] but may remain a missed goal.

As the FIPPs state, consent is an aspect of control. If students choose not to consent to learning analytics, then their information remains in their control. However, when learning analytics practices become "part of the operational infrastructure of an institution,"[38] they tend to qualify as evaluation and not research; therefore, the ethical standard is lower and consent requirements do not apply.[39] Consequently, students may become subjects of learning analytics without their knowledge, and even when they wish otherwise.

In the United States, higher education institutions who receive federal funding are subject to the Family Educational Rights and Privacy Act, or FERPA. The law dictates that students have a right to inspect and review their educational record, challenge the record's accuracy and request corrections, and prevent access to their

personally identifiable information (PII) to third parties without consent. However, Zeide points out that there are "numerous exceptions" to the rule, which allows institutions to disclose PII widely to institutional actors and to third parties contracted to provide services to the institution—all with minimal federal oversight or requirements to be transparent about their data practices.[40]

Responding to the overall *lack* of control students have over how institutions access and make use of their PII, some have called for reforms to control mechanisms, information notices, and consent procedures. However, Zeide argues that:

> *Notice and consent models rarely result in informed, voluntary user acceptance. Notice is either too complex to comprehend or too vague to provide an adequate sense of data recipients' information practices. The sheer quantity of data recipients and constantly changing policies make the consideration of terms of use and privacy policies impossibly time consuming.*[41]

Hoel and Chen suggest that the architectures of learning analytics systems should be motivated by privacy-by-design strategies, which may limit the downstream privacy issues, and perhaps resolve some of Zeide's criticisms.[42] Building on work by these researchers, Jones suggests that student privacy dashboards could provide the technical affordances necessary to inform students of learning analytics practices, seek their consent, and enable them to control the flow of PII at a granular level. They may also shore up trust issues that can emerge among students when they are unaware of who accesses and uses their PII.[43]

Student Perspectives of Their Privacy

While the majority of the research has focused on students as a research subject, very few papers have addressed student perceptions of privacy directly. Roberts, Howell, Seaman, and Gibson found that undergraduate students had a fairly positive view of learning analytics, yet were concerned that it could be dehumanizing, autonomy-reducing, and an invasion of privacy.[44] In a mixed-methods study, Schumacher and Ifenthaler found that positive views towards learning analytics depended in part on their ability to manage their privacy, which complemented their earlier work suggesting that students will share identifiable data if learning analytics return useful information.[45] Prinsloo and Slade explored student privacy self-management in relation to massive open online courses (MOOCs).[46] They argue that explicit opt-in and opt-out options may be insufficient for the majority of learning analytics uses in higher education and do not adequately address the need for greater transparency to improve understanding and trust. This need for active consent was also identified by Bomas, who further suggests tools such as an online dashboard where a student who—for underage students—their parents are able to make specific decisions about data capture, access, and uses.[47] These research initiatives begin to lay the groundwork for understanding various aspects of student perceptions of their privacy, but more work is needed, especially in relation to library participation in learning analytics.

Research Methods

The Research Agenda

The study we discuss below highlights initial findings from the first of three phases in a three-year, IMLS-funded research agenda. The research investigates student privacy perspectives and expectations as they relate to learning analytics practices, especially when academic libraries participate.

The research is motivated by the following general questions:

RQ 1: What privacy issues do students identify when informed about library learning analytics initiatives, practices, data types, and data sources?

RQ 2: How do the identified privacy issues map to particular goals of learning analytics initiatives by specific stakeholders (e.g., librarians, instructors, advisors, etc.)?

RQ 3: How do privacy perceptions change according to relevant student demographics and academic experiences?

RQ 4: With regard to their privacy expectations, what library and non-library learning analytics scenarios are acceptable to students, how do they explain the variations in acceptability, and what recommendations would students make to resolve existing privacy problems?

We are in the final stages of phase one, for which we conducted student interviews. For phase two, we will create a taxonomy of privacy problems from the interview data and then deploy a multi-institutional survey at the eight institutions to further investigate these issues and how students with diverse demographic and academic backgrounds perceive them. Phase three will use expert-vetted scenarios to explore how students negotiate their privacy values and expectations in relation to institutional stakeholder needs. Readers can find out more about the ongoing research and related resources at http://datadoubles.org.

This research project began with interviews to give a voice to students and a platform to talk about their unique experiences and values "from their own perspective and in their own words."[48] Additionally, we believed that interviews would enable researchers and students to work together at "site[s] of construction"[49] about learning analytics, which would help them assess previous knowledge and build new understandings of this emerging sociotechnical reality.

Interviewing Recruitment and Structure

The authors and four research assistants conducted 120 interviews with undergraduate students at eight United States institutions, including:

1. Indiana University-Bloomington; Bloomington, Indiana
2. Indiana University-Indianapolis (IUPUI); Indianapolis, Indiana
3. Linn-Benton Community College; Albany, Oregon
4. Northwestern University; Evanston, Illinois
5. Oregon State University; Corvallis, Oregon
6. University of Illinois at Chicago; Chicago, Illinois
7. University of Wisconsin-Madison; Madison, Wisconsin
8. University of Wisconsin-Milwaukee; Milwaukee, Wisconsin

Undergraduate students were recruited using a combination of methods based on institutional requirements and restrictions. These methods included email sampling, posted recruitment flyers, and contacting local student groups. Only students who self-identified as undergraduates over the age of 18 were permitted to participate in the study. Since phase one of the research did not aim to create generalizable results based on student demographics, the researchers did not ask participants for personal information besides their student status. Researchers provided students a $10 Amazon electronic gift card for their participation.

Each institution ran 15 semi-structured interviews, covering five thematic interview protocols (three interviews per protocol). The protocol themes included:

1. Privacy (generally)
2. Data sharing and use
3. Data protections
4. Awareness of and reactions to learning analytics
5. Libraries and learning analytics

Protocols averaged 10 questions. Each protocol shared three core questions that researchers asked of all 120 participants. In addition to the protocol questions, interviewers asked probing and follow-up questions as necessary to elicit comprehensive responses and build researcher-participant rapport.[50] Each researcher's institutional review board reviewed and approved the interview protocols and the research design.

Data Capture and Analysis

The research team provided students three ways to participate in an interview: face-to-face, on the phone, and using the web conferencing system Zoom. For, face-to-face interviews, the team allowed students to choose the place of the interview but also recommended locations (e.g., researcher's office, library conference room). Webcams were not used in Zoom sessions. Researchers digitally recorded interview audio for transcription and analysis purposes. On average, interviews lasted 29 minutes.

Data analysis followed two stages. Researchers wrote case summaries to support the first stage of data analysis and provide a succinct overview of the facts of the interview. Case summaries are "systematic, ordered" summaries of "what is characteristic to the given case," or interview captured in the transcript.[51] In the second stage, researchers line-by-line coded transcripts in Dedoose, a web-based and collaborative qualitative research application, as part of the open coding stage.[52]

This paper reflects early findings from the interviews, so there are important limitations regarding data analysis about which readers should be aware. First, the focus of this paper is primarily on data from the "libraries and learning analytics" thematic protocol and the core questions; researchers only reviewed case summaries and transcripts for this theme and sections related to the core questions. Second, the findings below represent emerging themes and concepts after open coding; more selective coding techniques will lead to greater refinement.

Findings

Information Access

Students are generally unaware of the data and information their institutions have access to about themselves. Interviews prompted students to list examples of student information, but this proved difficult. As one student said, "I don't know what information [my institution] is necessarily taking from me." Probing questions elicited responses indicating that students expect their institution to record demographic information (e.g., names, addresses, and phone numbers), financial aid information, and academic information, such as the courses in which they enrolled and the grades they earned. Once students began identifying types of information to which their institution had access, they would also begin exploring information sources. Often, students recognized that a learning management system was an environment that could capture, as one participant said, "every move a student is making." Other students recognized that using their student ID card or connecting to campus WiFi may produce data as well.

Researchers prompted participants to reflect on their university's library and what information it might have access to about students. Students overwhelmingly believed that their library tracked which physical materials they checked out. Some students suggested that librarians have access to academic information as well, including a student's enrollment records and stated program of study—but not their course grades and grade point average (GPA). To a lesser degree, other students suggested that libraries know when students interact with information systems and what they search for within those systems, such as the library website and journal databases.

Responses to institutional and library access to student information were, on the whole, positive. Students saw that benefits could accrue for themselves, their peers, and their institution from accessing and analyzing student information. For instance, one student said that information access and analysis "doesn't really bother

me at all" because of the ways that information could be used to improve resource usage and provide personalized resources. A common theme across interviews suggests that students recognize that when they are taking advantage of institutional services and information systems, it is plausible—if not an actual reality—that information is being created about them and made accessible for analytical purposes. Even though students were fairly unconcerned that their institution and library had access to identifiable information, this should not be misconstrued as an unlimited data access free-for-all.

Information Restrictions

Students expressed some nuanced arguments about when access should be restricted, especially concerning third parties. First, some students recognized that their liberal perspective on information access should not determine the privacy rights of their peers. In the context of talking about the institution accessing her online searching behaviors, one student said:

> *So, me personally, I don't search things that are really like too out there. I could see where someone else might do that and that would be like a problem; I definitely understand that. Maybe, for me since I personally don't do anything that would embarrass me, I wouldn't be worried about the school seeing; it's not a problem for me… for other students who might not be able to do what I'm doing, give them some privacy. Because who knows if [when] they go home that they even have WiFi, so they come to the school to utilize WiFi and do whatever they do in their free time.*

What is particularly telling about this excerpt is that the student recognizes that a lack of privilege and access to certain socio-material resources plays a part in how students use online resources and are afforded the option to protect their privacy.

Second, some information is especially sensitive, and that information deserves rigorous access limitations. Here, students primarily referred to GPA, but there were other notable examples. For instance, one student expressed concern about wider access to her mental healthcare records:

> *Because if [my institution] had the intention of using my data to create better programs or better educational tools, then I'm all for it, you know.… But I could also see certain things that are tracked, maybe being a little embarrassing. I initially didn't go [to the counseling center] for a long time because I was embarrassed, because I knew that the university was going to be able to track that and look at my record and say, "Oh yeah, she's been going to counseling." And maybe if they wanted to, they could somehow find out what exactly it was that I was talking to the therapist about.*

This student's concern maps to a larger theme regarding access justifications. Students claimed that uses of student information need to map to the institution's responsibility to support its student body and access should be limited to individuals who can make such a claim.

Finally, and most significant to note, students negatively looked upon providing entities outside of the institution access to student data. Researchers asked students "is it acceptable that your institution shares data and information about you outside the institution?" To this question, one student said: "Emphatic no. No. I just don't see how that's helpful for us." Another student said that providing access to his data was "an abuse of privacy." Researchers expressed to students that institutions often partner with third parties to provide a technology platform, such as learning management systems, and information services, like library databases. After being informed of this fact, students provided more nuanced responses. For instance, one student said:

> *So I guess that's a good point. So if you're actually helping the students get an article that they need for like a research paper, then I feel like that would be fine…. [If] it's not benefitting the student in any way or may harm them, then I think that's where you kind of have to draw the line.*

Student concerns about third-party data access lessened under two conditions. First, when data are presented in statistical form without identifying characteristics, students perceived fewer chances of downstream harms. Second, students expressed that they would like the opportunity to review the reasons for such sharing and the ability to consent to specific data practices. They argued that the data could contain sensitive information; and there was a worry that in the hands of the government, the data could be used against students, such as those with Deferred Action for Childhood Arrivals (DACA) status. Additional interview data exists about informed consent that we will address in detail in future studies but could not get to due to space limitations in this piece.

Library Privacy

Researchers engaged students in specific conversations about library participation in learning analytics initiatives with an emphasis on their privacy expectations. Even though students were positive about library learning analytics, they did express a number of questions about the practices, especially since they had never been informed that their library had access to or was analyzing certain types of data. A student responded, "I guess it comes back to what type of data are they looking at? Is it more personal stuff or is it solely looking at, well, what are they researching for this, or what are they doing?" Their concerns, however, were allayed in part by their trust in their library.

Students conveyed that their library was a "good" entity, one in which they placed trust to use data and information in ways that would not break that trust. One student's comment reflected that of her peers in that she compared her trust in her institution and the library in much different terms than major social media companies:

> *And, I feel like since there's kind of like a trust in libraries, like libraries should kind of like return that like gesture in a way. Like, I trust the institution of a library. I don't trust Facebook. So, it's like when you're on Facebook, you know it's kind of like the Wild West…. Yeah, so I feel like the types of data that like a university would collect is different, the purpose of it is completely different. The implicit kind of feeling that people have towards libraries, the library should not betray that trust.*

To maintain that trust—and the student's belief that their privacy remained protected—students expressed that libraries should have "protocols" and "data classifications" in place, which would define sensitive data types and guide data uses. Some students argued that the privacy concerns they did have would be significantly reduced if the library practices did *not* include identifiable data. Deidentified data would also reduce student anxiety that the library, and by extension the university, could target and treat specific students unfairly.

Discussion

Students' lack of awareness regarding the data and information their institutions have access to and may use for learning analytic practices signals a significant area for development on two intersecting fronts. First, the data signal that students lack a degree of privacy literacy; this is an education issue. Second, student desires to be informed and consent to specific data practices is a policy issue. Where libraries are concerned, the degree to which students will remain trusting of their learning analytics practices—and libraries in general—may depend on how library and information science professionals address these two fronts.

As researchers worked through their interview protocols, students became more forthcoming with their perspectives and privacy expectations. However, many students struggled to even come up with ideas regarding basic information to which their institution may have access and use for analytic purposes. We argue that interviews served as an educational experience for many students. The discursive process created an opportunity for them to increase their awareness of learning analytics and enhance their privacy literacy. But, ultimately, this responsibility rests with institutions and their libraries.

The data indicate that students' lack awareness about their institution's data practices stems in part from the latter's minimal—if not nonexistent—informed consent processes. While often viewed as a legal or regulatory responsibility, informed consent also serves as an educational tool. By informing students of an institution's data practices, protections, and their related privacy rights, student understanding of how particular data practices intersect with privacy issues is likely to increase. It is essential that institutions and their libraries develop informed consent policies in light of this educational need. Moreover, as the data indicate, student willingness to include their data *in* learning analytics depends in part on improved understanding *about* learning analytics.

Conclusion

This paper has presented the preliminary findings of phase one of student-centered research, wherein we conducted 120 surveys across eight academic institutions in order to explore undergraduate perceptions of privacy in relation to library learning analytics. This initial data suggests that students trust their institutions and see some value in data gathering to inform service improvement, but that they would like to provide active consent, better understand the scope of data captured, and see issues with unilateral data capture that may affect them or their peers.

Since this work represents emerging findings, there are some limitations. We cannot claim full transferability of concepts, though the findings we have discussed are credible.[53] With more time, the research team will be able to create expansive descriptions and more rigorous concepts from the data, which will boost transferability.[54] There is also some indication in the data that we may have a convenience sampling-like issue.[55] That is that some students may have been motivated to participate because they wanted to express strong privacy perspectives and possibly influence the findings. However, there is other evidence in the data suggesting that this is a limited issue; and when considered in whole, there is adequate variation in the sample. As further analysis continues, the research team seeks to identify areas where students demonstrate both least and greatest variance in their opinions about privacy, which will enhance the survey questions that will be part of the next stage of the research project.

While higher education and academic libraries have plunged headlong into big data, data analytics, predictive analytics, and learning analytics, as yet there is still little that is understood about student perceptions about these activities. In order to engage the students and ensure their autonomy and ability to advocate for their preferences to engage or remain anonymous within the data captured, further research is needed to understand their privacy concerns and their desire for transparency.

Acknowledgments

This project was made possible in part by the Institute of Museum and Library Services (LG-96-18-0044-18). The views, findings, conclusions or recommendations expressed in this conference proceeding do not necessarily represent those of the Institute of Museum and Library Services. The team thanks its research assistants for their support: Tanisha Afnan (Indiana University-Indianapolis), Aubree Tillett (University of Wisconsin-Milwaukee), Susan Rowe (Oregon State University), Ashwed Patil (Indiana University-Bloomington), and Donovan Pogue (Indiana University-Bloomington). Finally, the team thanks the undergraduate students who volunteered their time to participate in this study.

Endnotes

1. Phil Long and George Siemens, "Penetrating the Fog: Analytics in Learning and Education," *EDUCAUSE Review* 2011, no. September/October (2011): 30–40.
2. M Minelli, M Chambers, and A Dhiraj, *Big Data, Big Analytics. Hoboken* (NJ: John Wiley & Sons, 2013).
3. Amy Deschenes, "What Does the Dashboard Tell Us?," in *Computers in Libraries 2014*, n.d.
4. Travis T. York, Charles Gibson, and Susan Rankin, "Defining and Measuring Academic Success—Practical Assessment, Research & Evaluation" 20, no. 5 (2015), http://www.pareonline.net/getvn.asp?v=20&n=5; Gloria Crisp, Erin Doran, and Nicole A. Salis Reyes, "Predicting Graduation Rates at 4-Year Broad Access Institutions Using a Bayesian Modeling Approach," *Research in Higher Education*, April 2017, 1–23, https://doi.org/10.1007/s11162-017-9459-x; Alfred Essa and Hanan Ayad, "Improving Student Success Using Predictive Models and Data Visualisations,"
5. Eric Ackermann, *Putting Assessment into Action : Selected Projects from the First Cohort of the Assessment in Action Grant*, 2015; Graham Stone, Dave Pattern, and Bryony Ramsden, "Does Library Use Affect Student Attainment? A Preliminary Report on the Library Impact Data Project," *LIBER Quarterly* 21, no. 1 (October 2011): 5, https://doi.org/10.18352/lq.8005; *Assessment in Action: Academic Libraries and Student Success*, 2018, http://www.ala.org/acrl/AiA; Scott Walter, "Communicating Value through Strategic Engagement: Promoting Awareness of the 'Value of Libraries' through Alignment across Academic, Student, and Administrative Affairs," *Library Management* 39, no. 3/4 (December 29, 2017): 154–65, https://doi.org/10.1108/LM-09-2017-0093.
6. Megan Oakleaf et al., "Value of Academic Libraries: A Comprehensive Research Review and Report" (Association of College and Research Libraries, 2010), www.acrl.ala.org/value.
7. Karen Brown and Kara J. Malenfant, "Academic Library Impact on Student Learning and Success: Findings from Assessment in Action Team Projects" (ACRL, 2017), http://www.ala.org/acrl/sites/ala.org.acrl/files/content/issues/value/findings_y3.pdf.
8. Zacharoula Papamitsiou and Anastasios A. Economides, "Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence," *Journal of Educational Technology & Society* 17, no. 4 (2014): 49–64; Maren Scheffel et al., "Quality Indicators for Learning Analytics," *Journal of Educational Technology & Society* 17, no. 4 (2014): 117–32.
9. Rob Kitchin and Martin Dodge, *Code/Space: Software and Everyday Life* (Cambridge, Mass.: MIT Press, 2014).
10. Rosenberg, D., "Data before the Fact," in *"Raw Data" Is an Oxymoron*, ed. Lisa Gitelman (Cambridge, Massachusetts: The MIT Press, 2013), http://ieeexplore.ieee.org/xpl/bkabstractplus.jsp?bkn=6451327.
11. Daniel J. Solove, *The Digital Person: Technology and Privacy in the Information Age.* (Fredericksburg: New York University Press, 2006).
12. Sarah Brown, "Where Every Student Is a Potential Data Point," *The Chronicle of Higher Education*, April 9, 2017, https://www.chronicle.com/article/Where-Every-Student-Is-a/239712.
13. Goldie Blumenstyk, "Big Data Is Getting Bigger. So Are the Privacy and Ethical Questions.," *The Chronicle of Higher Education*, July 31, 2018, https://www.chronicle.com/article/Big-Data-Is-Getting-Bigger-So/244099.
14. Jason E. Lane and B. Alex Finsel, "Fostering Smarter Colleges and Universities," in *Building a Smarter University: Big Data, Innovation, and Analytics*, ed. Lane, JE (Albany, NY: State University of New York pres, 2014), 3–26.
15. Marc Parry, "Big Data On Campus," *The New York Times*, July 18, 2012, sec. Education Life, https://www.nytimes.com/2012/07/22/education/edlife/colleges-awakening-to-the-opportunities-of-data-mining.html.
16. Herman T. Tavani, "Privacy Online," SIGCAS Comput. Soc. 29, no. 4 (December 1999): 11–19, https://doi.org/10.1145/572199.572203.
17. Lillian BeVier, "Information About Individuals in the Hands of Government: Some Reflections on Mechanisms for Privacy Protection," *William & Mary Bill of Rights Journal* 4, no. 2 (February 1, 1995): 455; Judith Wagner DeCew, *In Pursuit of Privacy: Law, Ethics, and the Rise of Technology*, 1 edition (Ithaca: Cornell University Press, 1997); Helen Fay Nissenbaum, *Privacy in Context : Technology, Policy, and the Integrity of Social Life* (Stanford, CA: Stanford Law Books, 2010).
18. Samuel D. Warren and Louis D. Brandeis, "The Right to Privacy," *Harvard Law Review* 4, no. 5 (1890): 193–220, https://doi.org/10.2307/1321160.
19. Daniel J. Solove, *Understanding Privacy*, 2/28/10 edition (Cambridge, Massachusetts London, England: Harvard University Press, 2010).
20. Alan F. Westin, *Privacy and Freedom*, Second Printing edition (Atheneum, 1967), 7.
21. Warren and Brandeis, "The Right to Privacy," 200.
22. Warren and Brandeis, 205.
23. Ben Williamson, *Big Data in Education: The Digital Future of Learning, Policy and Practice*, 1 edition (Thousand Oaks, CA: SAGE Publications Ltd, 2017), 127.
24. Jeffrey Alan Johnson, "Ethics and Justice in Learning Analytics," *New Directions for Higher Education* 2017, no. 179 (2017): 79, https://doi.org/10.1002/he.20245.
25. Paul Prinsloo and Sharon Slade, "Student Vulnerability, Agency and Learning Analytics: An Exploration," *Journal of Learning Analytics* 3, no. 1 (April 23, 2016): 159–182–159–182, https://doi.org/10.18608/jla.2016.31.10.
26. David Kay, Naomi Korn, and Charles Oppenheim, "Vol.1 No6.: Legal, Risk and Ethical Aspects of Analytics in Higher Education" (CETIS for JISC), accessed February 14, 2019, https://digital.fundacionceibal.edu.uy/jspui/bitstream/123456789/282/1/Legal-Risk-and-Ethical-Aspects-of-Analytics-in-Higher-Education-Vol1-No6.pdf.

27. EL Godkin, "Libel and Its Legal Remedy," in *Journal of Social Science: Containing the Transactions of the American Association* (Leypoldt & Holt, 1880), 80.
28. Sissela Bok, *Secrets: On the Ethics of Concealment and Revelation* (Knopf Doubleday Publishing Group, 2011), 10–11.
29. Hendrik Drachsler and Wolfgang Greller, "Privacy and Analytics—It's a DELICATE Issue. A Checklist to Establish Trusted Learning Analytics" (ACM, 2016), 90, http://dspace.ou.nl/handle/1820/6381.
30. Kyle M. L. Jones and Chase McCoy, "Reconsidering Data in Learning Analytics: Opportunities for Critical Research Using a Documentation Studies Framework," *Learning, Media and Technology* 44, no. 1 (January 2, 2019): 52–63, https://doi.org/10.1080/17439884.2018.1556216.
31. Brenda Leong, "Who Exactly IS a 'School Official' Anyway?," Future of Privacy Forum, January 19, 2016, https://fpf.org/2016/01/19/who-exactly-is-a-school-official-anyway/.
32. Robert Gellman, "Fair Information Practices: A Basic History," *SSRN Electronic Journal*, 2014, https://doi.org/10.2139/ssrn.2415020.
33. The White House, "Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy," *White House, Washington, DC*, 2012, 1.
34. Gary T. Marx, "Ethics for the New Surveillance," *The Information Society* 14, no. 3 (August 1, 1998): 171–85, https://doi.org/10.1080/019722498128809; Daniel J. Solove, "Introduction: Privacy Self-Management and the Consent Dilemma," *Harv. L. Rev.* 126 (2012): 1880.
35. Kate Crawford and Jason Schultz, "Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms," *Boston College Law Review* 55, no. 1 (2014): 37.
36. O Tene and J Polonetsky, "Big Data for All: Privacy and User Control in the Age of Analytics," *Nw J Tech & Intell Prop*, 2012, http://heinonlinebackup.com/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/nwteintp11&section=20.
37. Abelardo Pardo and George Siemens, "Ethical and Privacy Principles for Learning Analytics," *British Journal of Educational Technology* 45, no. 3 (2014): 440.
38. Drachsler and Greller, "Privacy and Analytics—It's a DELICATE Issue. A Checklist to Establish Trusted Learning Analytics," 92.
39. James E. Willis, Sharon Slade, and Paul Prinsloo, "Ethical Oversight of Student Data in Learning Analytics: A Typology Derived from a Cross-Continental, Cross-Institutional Perspective," *Educational Technology Research and Development* 64, no. 5 (October 1, 2016): 881–901, https://doi.org/10.1007/s11423-016-9463-4.
40. Elana Zeide, "Student Privacy Principles for the Age of Big Data: Moving Beyond FERPA and FIPPS," *Drexel Law Review* 8 (2016 2015): 356.
41. Zeide, "Student Privacy Principles for the Age of Big Data," 382.
42. Tore Hoel and Weiqin Chen, "Privacy-Driven Design of Learning Analytics Applications—Exploring the Design Space of Solutions for Data Sharing and Interoperability," *Journal of Learning Analytics* 3, no. 1 (April 23, 2016): 139–158–139–158, https://doi.org/10.18608/jla.2016.31.9.
43. Kyle M. L. Jones, "Learning Analytics and Higher Education: A Proposed Model for Establishing Informed Consent Mechanisms to Promote Student Privacy and Autonomy," [Unpublished] February 2019.
44. Lynne D. Roberts et al., "Student Attitudes toward Learning Analytics in Higher Education: 'The Fitbit Version of the Learning World,'" Frontiers in Psychology 7 (2016), doi:10.3389/fpsyg.2016.01959.
45. Clara Schumacher and Dirk Ifenthaler, "Features Students Really Expect from Learning Analytics," *Computers in Human Behavior* 78 (January 1, 2018): 397–407, doi:10.1016/j.chb.2017.06.030.; Dirk Ifenthaler and Clara Schumacher, "Student Perceptions of Privacy Principles for Learning Analytics," *Educational Technology Research and Development* 64, no. 5 (October 1, 2016): 923–38, doi:10.1007/s11423-016-9477-y.
46. Paul Prinsloo and Sharon Slade, "Student Privacy Self-Management," in The Fifth International Conference (New York, New York, USA: ACM Press, 2015), 83–92, https://doi.org/10.1145/2723576.2723585.
47. Bomas, Erwin, "How to Give Students Control of Their Data," *LACE—Learning Analytics Community Exchange* (blog), August 29, 2014, http://www.laceproject.eu/blog/give-students-control-data/.
48. Steinar Kvale, *Doing Interviews* (London: SAGE, 2011), 11.
49. Kvale, 21.
50. Ayres, Lioness, "Semi-Structured Interview," in *The SAGE Encyclopedia of Qualitative Research Methods* (Thousand Oaks: SAGE Publications, Inc., 2008), 811–12, https://doi.org/10.4135/9781412963909.
51. Udo Kuckartz, *Qualitative Text Analysis: A Guide to Methods, Practice and Using Software*, First edition (Los Angeles: SAGE Publications Ltd, 2014), 52.
52. L. Benaquisto, "Codes and Coding," in *The SAGE Encyclopedia of Qualitative Research Methods*, ed. Lisa Given (2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc., 2008), https://doi.org/10.4135/9781412963909.n48.
53. Devon Jensen, "Credibility," in *The SAGE Encyclopedia of Qualitative Research Methods* (Thousand Oaks: SAGE Publications, Inc., 2008), 139–139, https://doi.org/10.4135/9781412963909.
54. Devon Jensen, "Transferability," in *The SAGE Encyclopedia of Qualitative Research Methods* (Thousand Oaks: SAGE Publications, Inc., 2008), 886–87, https://doi.org/10.4135/9781412963909.
55. Kristie Saumure and Lisa M. Given, "Convenience Sample," in *The SAGE Encyclopedia of Qualitative Research Methods* (Thousand Oaks: SAGE Publications, Inc., 2008), 125–125, https://doi.org/10.4135/9781412963909.