Data Doubles: Student Perceptions of Privacy and Learning Analytics in Higher Education

CNI Fall 2019 Meeting - December 9, 2019

Andrew Asher
Associate Professor and Assessment Librarian
Indian University – Bloomington
asherand@Indiana.edu

Michael Perry
Head of Assessment & Planning
Northwestern University
michael.perry@northwestern.edu
The Data Double research team is comprised of scholars and practitioners at eight institutions:

- Kyle M. L. Jones (PI) - Indian University – Indianapolis (IUPUI)
- Andrew Asher - Indiana University - Bloomington
- Kristin Briney - California Institute of Technology
- Abigail Goben - University of Illinois at Chicago
- Michael Perry - Northwestern University
- Mariana Regalado - Brooklyn College, CUNY
- Dorothea Salo - University of Wisconsin - Madison
- Maura Smale - New York City College of Technology, CUNY

Find out more about each research at datadoubles.org/team
This project was made possible in part by the Institute of Museum and Library Services (IMLS) (LG-96-18-0044-18)

The views, findings, conclusions or recommendations expressed in this website and other artifacts stemming from this research project do not necessarily represent those of the Institute of Museum and Library Services
Grant Overview

A student-centered, three-year research project to study student perspectives of privacy issues associated with academic library participation in learning analytics initiatives

- Phase One: Student Interviews
- Phase Two: Multi-institution Student Survey
- Phase Three: Scenario-Based Focus Groups
Guiding Research 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?
Major Deliverables

• A comprehensive, peer-reviewed literature review on library learning analytics and student privacy issues

• Peer-reviewed research from each phase (targeting open-access journals when feasible)

• Research update presentations primarily at practitioner conferences

• A Toolkit to be published at OSF, which will include presentations, pre-prints, curated reading lists, and research protocols (e.g., interview questions, coding schemes, survey, etc.)
What are Learning Analytics?

“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.”

Why is Student Privacy an Issue?

- Infrastructures supporting learning analytics makes visible student demographics, behaviors, and intellectual processes

- Learning analytics enables powerful actors to describe and predict student life

- Related practices enable human actors and systems to intervene in student life and influence their decision making

- There are serious access and control issues, among other things
Phase One: Semi-Structured Student Interviews
Interviews

- Semi-structured interviews with undergraduate students

- Number of interviews:
  - Targeted: 120 (15 interviews per site (8 sites))
  - Achieved: 112

- Students were provided $10 as an electronic Amazon gift card for their participation
Interview Sites

- Indiana University-Bloomington; Bloomington, Indiana
- Indiana University-Indianapolis (IUPUI); Indianapolis, Indiana
- Linn-Benton Community College; Albany, Oregon
- Northwestern University; Evanston, Illinois
- Oregon State University; Corvallis, Oregon
- University of Illinois at Chicago; Chicago, Illinois
- University of Wisconsin-Madison; Madison, Wisconsin
- University of Wisconsin-Milwaukee; Milwaukee, Wisconsin
Structure

- Each interview asked three core questions then focused on a specific theme
- Themes included:
  - Privacy (generally)
  - Data sharing and use
  - Data protections
  - Awareness of and reactions to learning analytics
  - Libraries and learning analytics
- Interviewers asked probing and follow-up questions as necessary to elicit comprehensive responses and build researcher-participant rapport
- Interviews averaged around 30 minutes
Phase One: Semi-Structured Student Interviews

Findings
Information Access by Institutions

- Students are generally unaware of the data and information to which their institutions have access when initially asked.

- However, probing questions revealed that students could identify a variety of demographic, academic, financial, and other information to which their institutions may have access.

- Some students were able to identify that campus information systems were likely to track their digital behaviors.
Information Access by Libraries

- 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.
Perceptions of Information Access by Libraries

• 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

• 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
Information Restrictions

- Students expressed some nuanced arguments about when access should be restricted, especially concerning third parties.

- Some students recognized that their liberal perspective on information access should not determine the privacy rights of their peers.

- Some information is especially sensitive, and that information deserves rigorous access limitations.

- Students primarily referred to GPA, but there were other notable examples about healthcare records.
Information Restrictions and Third Parties

- Students negatively looked upon providing entities outside of the institution access to student data

- 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

- Student concerns about third-party data access lessened under two conditions:

  1) When data are presented in statistical form without identifying characteristics, students perceived fewer chances of downstream harms

  2) Students expressed that they would like the opportunity to review the reasons for such sharing and the ability to consent to specific data practices
Library Privacy & Trust

• 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.

• Their concerns, however, were allayed in part by their trust in their library.

• 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.
Privacy Literacy

• Students became more forthcoming with their perspectives and privacy expectations as they worked through interviews

• 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

• 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
A Lack of Awareness and Informed Consent

• 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.

• 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.

• Treat informed consent as an educational process—not a policy burden.
Phase Two: Student Surveys
Phase Two: Student Surveys

- Statistically Representative Sample from 8 institutions:
  - Indiana University-Bloomington
  - Indiana University-Indianapolis (IUPUI)
  - Linn-Benton Community College (Albany, Oregon)
  - Northwestern University
  - University of Illinois at Chicago
  - University of Wisconsin-Madison
  - University of Wisconsin-Milwaukee
  - City University of New York (CUNY), Brooklyn College
Phase Two: Student Surveys

- 8 Thematic Modules
  - Knowledge about Learning Analytics
  - Data Sharing Practices
  - Access & Limits
  - Rights
  - Consent
  - Privacy
  - Trust
- Demographic Data
Survey Development

- Brainstorming Questions (Research Team)
- Question Culling, Editing, & Revision (Research Team)
- Cognitive Interviews (Students)
- Subject Matter Expert Panel Review
- Final Revision
Survey Validation

- Content Validity
  - Cognitive interviews evaluating definitions & question interpretation (qualitative)
  - Subject Matter Expert (SME) panel evaluating questions’ contribution to understanding modules (quantitative)
- Construct Validity
  - Confirmatory Factor Analysis
  - Test administration of survey (Northwestern, in-progress)
  - Post-hoc on full survey (spring 2020)
Instrument & materials will be available at:

https://osf.io/d7f3g/
Questions?

asherand@Indiana.edu  
michael.perry@northwestern.edu