2023

Student Privacy and Learning Analytics: Investigating the Application of Privacy within a Student Success Information System in Higher Education

Mary Francis

STUDENT PRIVACY AND LEARNING ANALYTICS: INVESTIGATING THE APPLICATION OF PRIVACY WITHIN A STUDENT SUCCESS INFORMATION SYSTEM IN HIGHER EDUCATION

By

Mary Francis

B.A., The Franciscan University –2004
M.A., University of Iowa – 2006
M.S., University of South Dakota – 2013

A Dissertation Submitted in Partial Fulfillment of
the Requirements for the Degree of
Doctor of Education

Division of Educational Administration
Educational Administration Program
Adult and Higher Education
In the Graduate School
University of South Dakota
May 2023

DOCTORAL COMMITTEE

The members of the Committee appointed to examine
the Dissertation of Mary Francis
find it satisfactory and recommend that it be accepted.

DocuSigned by:
Mejai Bola Mike Avoseh
C2A395A44C93465...

Chairperson

DocuSigned by:
Karen Card
D766CF1D5F9F436...

DocuSigned by:
Lisa Newland
D88A8C7A8EA947A...

DocuSigned by:
Kevin Streff
DFE9CC7B27D9456...

ABSTRACT

Learning analytics are starting to become standardized in higher education as institutions use the techniques of Big Data analytics to make decisions to help them reach their goals. The widespread use of student information brings forth ethical concerns primarily in relation to privacy. While the overarching ethical issues related to learning analytics are discussed in the literature, there has been a call for more studies to examine how they are put into practice. This case study used interviews and other data resources to determine how privacy is addressed within a student success information system at a public institution of higher education. During the inductive coding process three main themes emerged related to the connection between FERPA and privacy, methods to maintain privacy, and students' connection with their data. A deductive coding process was also undertaken to determine how the institution addressed the privacy principles put forth in the larger privacy literature. Overall, the findings showed the institution had a minimal understanding of privacy concerns related to learning analytics. This was not unexpected given the length of time the system had been in use at the institution. Recommendations for the institution include developing policies and procedures to guide their use of learning analytics moving forward.

Mejai Bola Mike Avoseh
Dissertation Advisor Dr. Mejai Bola Mike Avoseh

ACKNOWLEDGEMENTS

I would like to thank my dissertation committee for their feedback on this study. Special thanks to Dr. Avoseh for serving as chair overseeing the project, Dr. Card for reading several drafts and providing comments as my methodologist, Dr. Newland for providing detailed comments on my methods, and Dr. Streff for working with me over the years on various privacy research topics.

I also want to acknowledge my family: Christopher, Henry, and Hazel just so they can see their names in print. :)

Table of Contents

List of Tables

List of Figures

Figures PAGE

CHAPTER 1

Introduction

Institutions of higher education are being called to demonstrate their effectiveness amid the additional requirements of efficiency and maintaining costs. Meanwhile, technological advances have allowed for the gathering and analysis of data to aid decision making. The conjuncture of these two circumstances have made learning analytics a critical component for many institutions. Learning analytics is the use of the big data techniques that are utilized within the business sector but with the goal of improved educational experiences.

Big data is the analysis of data to make decisions. Pence (2014-2015) notes how big data is possible due to the volume, variety, and velocity of data available. Big data has an ever-increasing impact upon daily interactions as it 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-Schonberger & Cukier, 2013, p. 6). The result of big data is the commercialization of information.

While the techniques are similar, there is an important difference between commercial big data analysis and learning analytics. As noted by Rubel and Jones (2016), for learning analytics to have the biggest impact, the student data must be connected to the individual. In big data analytics, the information can be used in the aggregate. This differentiation makes learning analytics a more personalized process which raises additional concerns related to the ethical use of such data.

There are varied uses to which learning analytics are applied. Some of the most common uses as noted by the Society for Learning Analytics Research (SOLAR) (2021) include:

prediction of students' academic success; identification of at-risk students; supporting student development of lifelong learning skills and strategies; provision of personalized and timely feedback to students regarding their learning; supporting development of important skills such as collaboration, critical thinking, communication and creativity; develop student awareness by supporting self-reflection; and support quality learning and teaching by providing empirical evidence on the success of pedagogical innovations. These goals are achieved through the implementation of various analytic strategies including descriptive analytics which provides insights into the past, diagnosis analytics which looks at why something happened, and prescriptive analytics which offers suggestions as to what will happen in the future (SOLAR, 2021).

However, learning analytics is relatively new in higher education and with it comes questions and concerns as to how it is implemented. Proponents of learning analytics highlight the ability to use data to increase the learning experience of students resulting in enhanced education. Yet, there remains concerns as to how these processes may provoke unintended consequences. While there are varied ethical considerations in collecting, analyzing, and using data, one of the most pressing concerns is student privacy. Hoel and Chen (2016) provide the logic for the importance of studying how privacy is addressed in learning analytics. They note that while institutions have long analyzed behavior and performance to make changes, learning analytics has changed how that process is done and the impact that it can have upon individuals. This new process then necessitates a new agreement between student and institution in relation to practice and how goals are met.

Institutions do recognize that privacy is a concern. The *EDUCAUSE 2020 top 10 IT issues* report placed privacy second on the list after security (Grajek, 2020). Burns (2020) notes that while institutions recognize the importance of privacy they must work to develop and improve policies and procedures dealing with student information. With no dedicated federal law or guidance on how to address privacy concerns within learning analytics, it has been left to each institution to develop their own approach. Petersen (2012) recommends that "a common ethical framework for the development of campus privacy policies and practices can be found in the Fair Information Practice Principles" (p. 48). The idea of using privacy principles to evaluate privacy is appropriate "because of the near impossibility of measuring privacy itself […] almost all empirical privacy research in the social sciences relies on measurement of a privacy-related proxy of some sort" (Smith, et al., 2011). This study shall use a case study to see how a small public institution of higher education in the Midwest addresses the privacy of students within a student success information system.

Opportunity for Change

This study will examine how student privacy is addressed within a student information system focused on student success. The data within such systems are frequently utilized in learning analytics. As a developing field of study, learning analytics research will mature only through studies looking at all aspects of the field. This case study will provide an in-depth analysis of one institution that has implemented a specific system. A case study requires data to be collected from a range of sources which ensures a full picture can be developed related to the topic.

Theoretical Framework

Alan Westin put forth an approach to privacy in his 1967 book *Privacy and freedom* that continues to influence privacy research and application to the current day (Austin, 2019; Margulis, 2011; Rizza et al., 2012). Westin's (1967) book begins with a historic and cultural view of privacy which extends to the then current issues related to how technology impacts privacy. Some of the major components that Westin (1967) addresses about privacy includes how individuals and groups control access to themselves, the fact that the need for privacy varies based on different situations, and the fact that individuals can experience both too much and too little privacy. Overall, Westin believes that privacy is necessary for individuals to achieve self-realization. His oft-cited definition of privacy reads in full:

> Privacy is the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others. Viewed in terms of the relation of the individual to social participation, privacy is the voluntary and temporary withdrawal of a person from the general society through physical or psychological means, either in a state of solitude or small-group intimacy or, when among larger groups, in a condition of anonymity or reserve. The individual's desire for privacy is never absolute, since participation in society is an equally powerful desire. Thus, each individual is continually engaged in a personal adjustment process in which he balances the desire for privacy with the desire for disclosure and communication of himself to others, in light of the environmental conditions and social norms set by the society in which he lives. The individual does so in the face of pressures from the curiosity of others and from the processes of surveillance that every society sets in order to enforce its social norms. (Westin, 1967, p. 7)

Background

In one of the earliest influential descriptions of privacy, Samuel Warren and Louis Brandeis (1890) offer an approach to privacy that focuses on the rights of an individual to be let alone specifically focusing on the physical aspects of privacy. This focus on physical space, personal body autonomy, and restrictions of covert governmental surveillance was the primary focus of privacy until Westin put forth his book in 1967. The forward of the book notes that the impetus for this work came from the Committee on Science and Law under the Association of the Bar of the City of New York who provided the support for an inquiry into the impact of modern technology on society (Westin, 1967). Westin led and organized the research which was conducted over four years and resulted in the final publication.

The publication led to new thoughts about privacy and resulted in several researchers testing and applying the ideas found in Westin's work. Margulis (2003) provided a meta-analysis of the research done looking at the privacy states and functions put forth by Westin and found that they supported and extended his theory. His approach still resonates with individuals as Kwasny et al. (2008) found in their study of privacy beliefs that Westin's approach to privacy fit into the ideas of young adults better than those of older adults.

There are several features of Westin's approach that has insured its endurance through the years, and which makes it applicable in the current technological age. First, he posits that privacy occurs at the individual, group, and organization or institution level. This statement provides nuance to the complex interplay of privacy found within different levels. Westin's (1967) approach to privacy also focuses on information privacy rather than physical privacy. This consideration that information can be private outside of the individual is critical as computer systems store an ever-increasing amount of personal information. Finally, in relation to this study, Westin (1967) puts forth the hypothesis that groups may have a harder time maintaining their privacy than an individual. Westin's work resonates as his ideas, established in 1967, could be heard to this day, "American society […] faces the task of keeping this tradition [privacy] meaningful when technological change promises to give public and private authorities the physical power to do what a combination of physical and socio-legal restraints had denied to them" (p. 22).

The next major impact upon privacy literature was the development of the privacy regulation theory put forth by Irwin Altman in 1975. His theory details five properties of privacy and moves the thought of privacy beyond just the individual into the larger society. The ideas from these major contributors to privacy can be classified into different perspectives of privacy as put for by Tavani (2007). He notes that the nonintrusion and seclusion perspectives would include the work of Warren and Brandeis. The perspectives of control and limitation would include the work of Westin and Altman. Control and limitation remain a strong consideration for those researching and designing privacy in systems (Jones, 2019).

Not only are there various overall conceptions of privacy, but various disciples approach privacy with different lenses. Pavlou (2011) notes how the law literature considers privacy as a right or entitlement, social psychology literature considers privacy as a state of limited access or isolation, and information systems literature views privacy as control over information. Smith et al., (2011) reviewed several articles across disciplines and developed two broad categories in which to place the definitions. Value-based definitions view privacy as a human right integral to that society's moral system. Cognate-based definitions refer to an individual's mind, perceptions, and cognition rather than a strict moral value.

The variations and focuses given to privacy are continued to this day with researchers working to build upon and understand privacy. Some of the more recent influential scholars looking at privacy include: Solove, Nissenbaum, and Cavoukian. Solove is a law professor who studies privacy and has written mass market books on the connection between privacy and information technology. Nissenbaum developed the theory of contextual integrity. Cavoukian created an approach to systems engineering which puts privacy in the forefront called privacy by design.

Westin's Privacy Framework

Westin (1967) put forth a framework to consider privacy where he details four states of privacy which serve as the *how* of maintaining privacy and four functions of privacy which detail the *why* of privacy.

States of Privacy

1. Solitude: where the individual is separated from others and free of all observation. This is the most complete state of privacy.
2. Intimacy: where an individual is part of a small unit that allows for close relationships. Intimacy is necessary to meet the basic need of human contact.
3. Anonymity: where an individual is in a public space but is free from identification and surveillance. Knowledge or fear of being observed in public spaces results in a negative response by the individual. Anonymity also includes sharing information without it being connected to a specific individual.
4. Reserve: where an individual's desire to limit communication about themselves is respected by the discretion of others.

Functions of Privacy

1. Personal autonomy: the ability to avoid manipulation, domination, or exposure from others.
2. Emotional release: the ability to let go of emotions and tensions built up from social roles within a private space.

3. Self-evaluation: having time to integrate experiences into meaningful patterns and structures and exert individuality on events.

4. Limited and protected communication: set boundaries for interpersonal relations as well as take part in communication with trusted individuals.

Austin (2019) notes how Westin's states of privacy are about information and how we select a state by limiting the amount of information we will share with others.

Conceptual Framework: Privacy Principles

Privacy remains a universally accepted idea without a universally accepted definition. Lacking a concrete definition, principles can be utilized to consider aspects of privacy. A principle is a shared value upon which regulations, rules, and standards can be built for the protection and advancement of the stated objective. Principles also serve as an appropriate structure to view a concept due to their foundational aspect. Principles retain their relevance throughout time allowing for change in technology and context. In maintaining the privacy of personal data, this trait is especially important as technological advances allow for new approaches and methods to maintain privacy. The past 50 years has seen the development of several data privacy principles created to "identify problematic practices" (Wright & Rabb, 2014).

These data privacy principles have been delineated through consideration of various approaches to privacy. Westin's (1967) listing of privacy states and functions continues to impact the approach and theory surrounding privacy by providing a high-level view of privacy. Austin (2019) notes that Westin's work has influenced the development of data protection law with their integrated principles and the Fair Information Practice Principles developed by the United States Department of Health, Education, and Welfare (1973).

Major Compilations of Data Privacy Principles

As noted earlier, there have been several compilations of data privacy principles generated by national, international, and organizational groups over the years. While no single listing has become the standard, they each provide additional insight into the discussion and analysis of privacy. Some of the more influential published listings of privacy principles include:

- Fair information practice principles (United States Department of Health, Education, and Welfare, 1973).
- OECD Guidelines on the protection of privacy and transborder flows of personal data. (Organization for Economic Co-operation and Development, 1980).
- APEC privacy framework (Asia-Pacific Economic Cooperation, 2005).
- ISO29100 (International Organization for Standardization, (2011).

Data Privacy Principles

The list of data privacy principles used for this research come from a prior study conducted by the author (Francis et al., 2020). In developing this list, several compilations of principles were consulted. Each principle was defined to include a single concept. Due to the broad expansiveness found in language, definitions focused on being concise and direct. Descriptions of each principle are included below.

- Notice: Subjects will be informed of data collection and use policies
- Retention: Data will be removed when no longer required
- Minimization: Limit the amount of data collected, processed, and stored
- Use restriction: Data can only be used for defined and accepted purposes
- Security: Data is handled in accordance with appropriate security principles
- Quality: Data is accurate and kept up to date

- Access: Subjects have the right to know what personal data is being held about them
- Participation: Allow data to be corrected and deleted by the subjects of the data whenever appropriate
- Enforcement: Data holders must comply with applicable policies, laws, and standards
- Consolidation: Consolidation of databases containing personal data cannot be done
- Consent: Enable data subjects to agree to data collection
- Transparency: Make all data collection, use, storage, and deletion as transparent as possible with clear and understandable language used to explain all privacy-related policies
- Context: Apply the context of the jurisdiction one operates in into privacy policies
- Accountability: Privacy policies must be developed that clearly describe the practices and procedures related to the management of personal data
- Identifiability: Data subjects have the option of remaining anonymous or using a pseudonym
- Sensitivity: Treat all data collected, used, stored, and destroyed in manners appropriate to the sensitivity level of the data
- Information flow: Enable the communication of personal information across multiple contexts including international, governmental, economic, and social
- Identifiers: Strong identifiers are only used when necessary
- Disclosure: Make known any data transference to new parties
- Confidentiality: Maintain confidentiality of data throughout processes and beyond
- Breach: Subjects must be informed immediately of any data breach involving their personal data (Francis, 2020, p. 4370).

Research Questions

This study will look specifically at how privacy is addressed within a student success information system through an analysis of interviews with the system administrator, trainings offered on the system, interviews with faculty member who use the system, documentation from the company, documentation from the institution, and interviews with a company representative. All of these sources of data will provide insight into how privacy is addressed. This understanding of how privacy is currently addressed then allows for a consideration of any changes that should be made. While this case study does not represent the practice of all institutions, it does provide an example that can be considered for those implementing their own student success information systems and utilizing learning analytic approaches on the ensuing data.

The following research question will guide this study:

- How is the concept of privacy addressed in relation to a student success information system within an institution of public higher education?

Sub questions include:

- How were the policies and procedures related to student privacy within the system developed and implemented?
- How is the need for privacy balanced against the institution's functional data needs?
- How does the institution weigh individual privacy rights against group benefits?

Significance of the Study

Learning analytics continues to become a mainstay within higher education. As administrators look to prove the impact of education on a public that questions the benefit of higher education and as efficient decisions must be made due to shrinking budgets, learning analytics provides measures that can be used in highlighting benefits of education and making institutional decisions. It is critical that institutions develop considered approaches when developing learning analytic programs.

This study provides officials at institutions of higher education with an appreciation for the importance of student privacy. With recognition of what privacy entails, officials will be able to develop policies and procedures that meet the privacy rights of students. The case study will provide an example of how privacy is considered at one institution in order to get a picture of how privacy is being addressed and how it could be further developed. The data gathered from the company will provide an understanding of what the systems currently allow for in relation to privacy. The knowledge about privacy also provides them with needs as they negotiate with companies in the future. This study also provides the field with a glimpse at how privacy is being currently addressed in a student success information system.

Learning analytics itself is a new discipline with the 2010 *Horizon Report* discussing the use of advanced computational methods and data visualization techniques which is the foundation of learning analytics (Johnson, et al., 2010). Then the first International Conference on Learning Analytics and Knowledge occurred in Canada in 2011. There is therefore a need for in-depth research in the field. Within their systematic review of articles looking at the ethical concerns related to learning analytics, Parkman and McGrath (2021) note that a majority of the research conducted focused on respondents' perceptions and attitudes related to learning analytics rather than the actual use of the systems. They recommended that research looking at "how ethical principles, guidelines, or codes of practices in LA [learning analytics] are put into practice will help us gain a more grounded understanding of how these instruments work in everyday higher education" (p. 13). Lang and Knight (2019) also stress the need for specific case studies looking at ethics and learning analytics. This work will help address that need by providing a case study looking at how privacy principles are addressed with the student success information system by a small public institution of higher education in the Midwest.

Definition of Terms

The following definitions are provided to ensure uniformity and understanding of these terms throughout the study. The researcher developed all definitions not accompanied by a citation. While it is acknowledged that privacy does not have a universal definition, for this study, the operational definition of privacy is included below.

Learning analytics. "The measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs" (Long & Siemens, 2011, p. 33). The data used in learning analytics comes from a variety of sources such as student information systems, student success information systems, and learning management systems.

Privacy. Restriction of access to an individual's personal information.

Student information system. Electronic databases that store student information and allow for such information to be analyzed. These systems often focus on maintaining registrar level information such as grades and admission test scores.

Student success information system. Electronic databases that collect and store student information. These systems often allow individuals and departments from across campus to supply details about the student such as missing assignments and tutoring recommendations.

Organization of the Study

Chapter 1 provides an introduction, opportunity for change, theoretical framework, research questions, the significance of the study, and definition of terms. Chapter 2 provides a review of literature related to privacy and learning analytics. It includes an overview of learning analytics including: benefits, challenges and concerns, and ethical considerations. The chapter also looks of data privacy, providing a discussion related to the difficulties in defining privacy, looking at how privacy changes related to context, understanding privacy principles and privacy harms, and looking at how privacy affects groups as well as individuals. Chapter 2 also provides an overview of laws impacting student data and privacy. Finally, the chapter provides an overview of current models and approaches used in designing, implementing, and evaluating learning analytic systems. The approach, methodology, data collection procedures, site details, and data analysis process are outlined in Chapter 3. Chapter 4 will examine the themes uncovered in the analysis of the data. A summary of the study and conclusions formulated from the findings, a discussion, and recommendations for practice and further study will be addressed in Chapter 5.

CHAPTER 2

Literature Review

This chapter will review the literature related the learning analytics and privacy. The major topics that will be addressed include: learning analytics with subsections providing details on an overview of learning analytics, benefits, challenges/concerns, and ethics of learning analytics; defining data privacy with subsections providing details on considering privacy within context, privacy principles, privacy harms, and the privacy of groups; laws impacting the privacy of students including FERPA, GDPR, and CCPA; and current models and approaches to evaluation learning analytics.

Learning Analytics

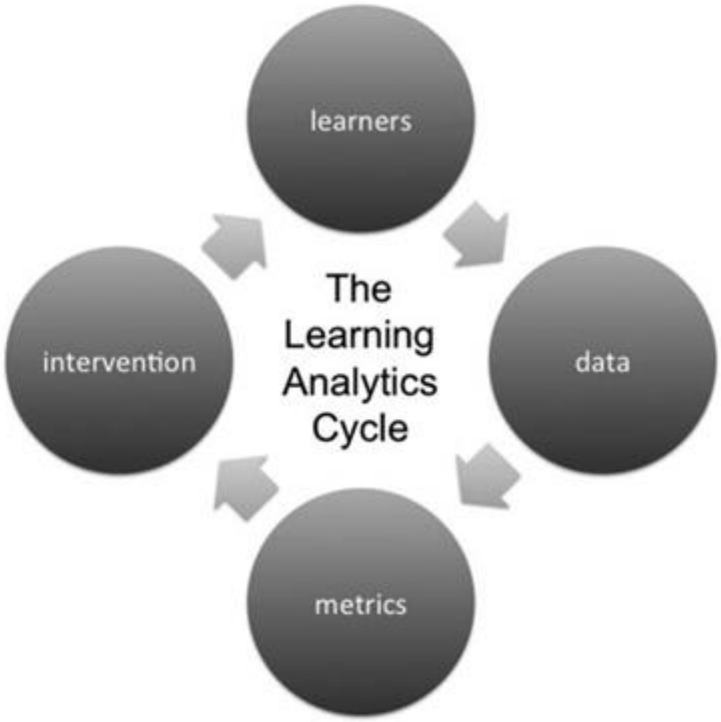
Overview

Student data has long been used to make decisions both on the microlevel in specific classrooms and at the macrolevel in how the institution operates. In the past, this data has come from faculty use of student grades and in-room discussion while institutions looked at yearly retention and graduation rates. While the use of data is not new, the current interest in learning analytics is due to the conjuncture of several trends: the volume of data that is collected, the ability to store that data, the computational capacity now available to institutions, the increase in visualization tools, and the increased demand to analyze and use big data (Slade and Prinsloo, 2013; Siemens, 2013). The use of learning analytics in higher education is a critical area of study as “the quantification, measurement, comparison, and evaluation of the performance of institutions, staff, students, and the sector as a whole is intensifying and expanding rapidly” (Williamson et al., 2020). One of the most frequently used definitions of learning analytics comes from Siemens (2013), “learning analytics is the measurement, collection, analysis, and reporting of data about learners and their contexts, for the purposes of understanding and optimizing learning and the environment in which it occurs" (p. 1382).

With the increased push to utilize learning analytics, institutions must implement structured plans to ensure successful programs. The Data Quality Campaign (2019) provide four policy priorities that they recommend institutions consider as they implement learning analytics: measure what matters, be transparent and earn trust, make data use possible, and guarantee access and protect privacy.

There are several general steps within the process of learning analytics: first data is collected, then the data is aggregated from various sources and examined to find patterns, next actionable decisions are made based on the information, and finally the results are reviewed and revisions made to the models (Steiner et al., 2016). Clow (2013) provides a visual representation in his Learning Analytics Cycle (insert figure 1). These broad steps then allow for institutions to customize their processes.

Figure 1. Clow's Learning Analytics Cycle



Fisher et al. (2020) provide three broad levels of how information can be gathered before use in learning analytics. Microlevel data are fine-grained interactions that are often collected automatically during student interactions with educational systems such as learning management systems (LMS). Mesolevel data comes from analysis of students' writing whether it is from an LMS or social media interactions. Macrolevel data is collected at an institutional level and is tracked over several years such as student demographics and course schedules. This gathered information can be used in several broad categories: predictive modeling where variables and outcomes are tracked to find patterns that can then be applied with future students, social network analysis where relationships between students are tracked via functions such as discussion forums, usage tracking which considers what students do within an LMS or other online environment, content analysis and semantic analysis which provides analysis of text created by students, and recommendation engines which provide suggestions to individuals based on the performance of past students (Clow, 2013).

Benefits

There have been a range of claims related to how the learning analytics available through student information systems, student success information systems, LMS, and other systems will improve education. These include enhanced learning experiences (Long and Siemens, 2011), supported self-regulated learning (Kim et al., 2018), improvement of student learning services (Knight et al., 2014), development of prediction analytics for at-risk students (Saqr et al., 2017), and help for at-risk students (Pardo & Siemens, 2014). Long & Siemens (2011) also discuss the benefits of improved institutional decision making, advancements in learning outcomes for at-risk students, greater trust in institutions due to the disclosure of data, significant evolutions in pedagogy, sense making of complex topics, increase organizational productivity, and provide learners insights into their learning. Foster and Francis (2020) conducted a systematic literature review looking at 34 learning analytics studies that focused on the stated goals of retention, academic performance, and engagement. They found that a majority of the studies reported an increase in student outcomes related to those goals.

Challenges/Concerns

While the goal of learning analytics is to improve the learning experience for students, this does not negate the fact that the gathering of data on students poses risks and challenges. Robertshaw and Asher (2019) conducted a meta-analysis on studies that used learning analytics to look for correlations between library use and instruction impacting student success. Their stated purpose was to determine whether the benefits of using learning analytics outweighed possible harms. They found no to small effect sizes and cautioned that learning analytics may not reach a level of beneficence. They also noted, however, the increasing push for learning analytics within higher education and suggested more robust studies to determine their true impact.

There has been a long-held concern with the impact that technology could have on privacy with Westin (1967) noting how “data surveillance – the maintenance of such detailed daily and cumulative records of each individual’s personal transactions that computerized systems can reconstruct his acts and use such data for social control even without direct physical surveillance” (p. 59). This statement was made during a time where memory and computing power was limited. The current ability of computers to engage in big data analytics has only heightened the ability of systems to gather, collate, and develop profiles of individuals. This accumulation of data often occurs without the individual realizing how specific data points are being combined and utilized.

Selwyn (2019) poses several possible consequences related to learning analytics including: a reduced understanding of education, ignoring the broader social contexts of education, reducing students’ and teachers’ capacity for informed decision-making, a means of surveillance rather than support, a source of performativity, disadvantaging large numbers of people, and serving institutional rather than individual interests. Privacy concerns related to data analytics occur throughout the data lifecycle: from the collection and control of the data, to issues related to who owns the data, through the use of data, and finally related to the maintenance and deletion of the data.

Collection and Personal Control.

Jones (2019) provides a look at how historically the concept that an individual has a right to control who knows specific information about themselves has long been central to privacy. Jones (2019) then goes on to detail how loss of that control results in privacy harms including attacks on personal autonomy. This loss of control may occur either through institutions not asking for consent or by asking for consent without providing a clear description as to how the data will be used. Slade and Prinsloo (2013) add that when asking for consent institutions should also provide details on the possible benefits and harms that may result from sharing or not sharing the information. In deciding how to use information, learning analytics must also answer the question as to who will benefit from the data. Will the analytics support individual students or the institution (Rubel & Jones, 2016)?

There are still concerns after consent is given as it is then possible for analytics to combine those disparate data points in order to uncover more than the individual would have originally intended. Learning analytics is meant to provide insights into the future actions of students thereby allowing appropriate interventions. This requires knowing enough information about an individual to foretell future behavior. This deep understanding of an individual would delve into the privacy of the identity of that individual. Pence (2015) provides various examples of how using anonymous data points, researchers were able to identify a specific individual.

Collection also deals with what type of information is collected. In discussing the types of personal data, Tavani (2007) notes that it is the context rather than the data itself that determines whether something should be considered public or private. Slade and Tait (2019) also note how all information should not be included in learning analytics. Institutions must be deliberate on the types of data collected rather than collecting everything with the hope that some meaning will later be found. For those institutions utilizing analytics developed by third party vendors, they must worry about *black box* analytics where they have no control or understanding as to what data is being used in developing the results provided (Oakleaf, 2016; SoLAR, 2021). Another consideration occurs when institutions start to gather data outside of the systems they control. Slade and Prinsloo (2013) note the challenges of confirming identity when using data from social media profiles.

Ownership of Data.

The collection of data raises questions related to the ownership of that data. Tucker and Long (2018) raise concerns with the influence that student information systems, student success information systems, LMS, and other systems may have upon education and how these third parties may use student data. Brown and Klein (2020) found in their analysis of privacy policies that institutions often treat student data as artifacts for the institution to use rather than personal information connected to individuals. Slade and Tait (2019) note how data is not something that a student owns or creates but rather makes up who they are and thus they must have a critical voice of how it is used. To that end they suggest that institutions act as temporary stewards of the data. Siemens (2013) extends the question of ownership to the analyzed data. Who owns the findings which come from privately inputted data? Can these findings be shared without consent?

Use of Data.

Jones (2019) looks at how institutions may gather consent to use data for one purpose, but then use that same data for other purposes without getting the individuals’ consent. Analysis of large datasets results in the development of patterns based on a range of characteristics. This can result in individuals being put into groups with others sharing an undetermined similarity (Mittelstadt, 2017). This grouping gives the individual an identity, e.g., *high-risk*, that will then impact how they are viewed and the actions they are allowed to take. They may be provided with additional benefits or restrictions limiting their personal autonomy.

Another concern that arises within large dataset analysis is the use of personal information such as race or religion in making groupings. Predictive analytics also may lead to bias in the development of algorithms. The selection of specific data points may cause inherent bias in the predictions made by the system. Discrimination can occur with biased datasets (Romei & Ruggieri, 2014). Ferguson (2019) provides an overview of how equality and justice can be impacted when implementing learning analytics.

Context also has an impact on how data is used. Nissenbaum's (2011) contextual integrity theory highlights how information that is gathered in one context cannot be used in another context without a full understanding of the original context. This impacts not only the meaning of the information, but also has impact upon consent by the individual who may have approved of the gathering of the data is in first instance but does not approve in the second instance.

Maintenance and Deletion of Data.

Maintenance of data at one level involves the technical concerns related to having a secure system that operates as appropriate. This concern is discussed heavily in the computer science and database management literature. One aspect of a secure system is that only specific authorized individuals will have access to the data held within (Slade & Prinsloo, 2013). Within the learning analytics literature the concern related to maintenance often deals with the quality of information within the system. One method for quality checks is to allow access. Slade and Tait (2019) note that students should have access to their raw and analyzed data in order to make corrections as necessary.

In the past with paper academic records, the storing of data was physically restricted. Now with almost unlimited electronic storage capabilities, the question becomes how long student data should be held. Slade and Prinsloo (2013) detail how in gathering data, learning analytics provide a snapshot in time and it as such there may be time for data to be removed from the system especially if students determine that such data does not reflect their current identity.

The deletion of an individual data point for a student will have bearing on whether the information is meant to aid the individual student or contribute to a larger dataset for future analytics.

Ethical.

Many of the challenges with learning analytics have an ethical component. Ferguson (2019) compared 30 identified concerns related to learning analytics to ethical challenges to develop six challenges faced by learning analytics. These include:

> Challenge one: Use data and analytics whenever they can contribute to learner success, ensuring that the analytics take into account all that is known about learning and teaching.
>
> Challenge two: Equip learners and educators with data literacy skills, so they are sufficiently informed to give or withhold consent to the use of data and analytics.
>
> Challenge three: Take a proactive approach to safeguarding in an increasingly data-driven society, identifying potential risks, and taking action to limit them.
>
> Challenge four: Work towards increased equality and justice, expanding awareness of ways in which analytics have the potential to increase or decrease these.
>
> Challenge five: Increase understanding of the value, ownership, and control of data.
>
> Challenge six: Increase the agency of learners and educators in relation to the use and understanding of educational data. (p. 28).

Ferguson et al. (2016) provide a listing of learning analytics challenges that have an ethical dimension. Of the 21 challenges put forth, 14 are addressed in the data privacy principles listed earlier. Many of the others relate more specifically to the management of the data. Table one includes the Ferguson et al. (2016) list of challenges with those related to the privacy principles bolded.

Table 1
Learning Analytics Challenges with Ethical Dimensions

1.	Use data to benefit learners
2.	**Provide accurate and timely data**
3.	**Ensure accuracy and validity of analyzed results**
4.	**Offer opportunities to correct data and analysis**
5.	Ensure results are comprehensive to end users
6.	Present data/results in a way that supports learning
7.	**Gain informed consent**
8.	**Safeguard individuals' interests and rights**
9.	Provide additional safeguards for vulnerable individuals
10.	**Publicize mechanisms for complaint and correction of errors**
11.	Share insights and findings across digital divides
12.	**Comply with the law**
13.	**Ensure that data collection, usage, and involvement of third parties are transparent**
14.	**Integrate data from different sources with care**
15.	Manage and care for data responsibly
16.	**Consider how, and with whom, data will be accessible**
17.	**Ensure data are held securely**
18.	**Limit time for which data are held and before destruction and for which data is valid**
19.	Clarify ownership of data
20.	**Anonymize and de-identify individuals**
21.	**Provide additional safeguards for sensitive data**

Ethics of Learning Analytics

One of the major concerns related to learning analytics are ethical considerations that arise when collecting, using, and storing student data. Some authors have looked specifically at privacy concerns in relation to ethics. Slade and Prinsloo (2013) note how the ethical considerations stemming from the increased use of learning analytics come from issues related to privacy and determining who owns the data that is collected. Gasevic, Dawson, and Jovanovic (2016) note that while privacy and ethics have been a concern related to learning analytics since their development they have not been explored fully in the literature.

There are also authors who have looked at general ethical concerns related to learning analytics which either mention privacy as a broad category or refer to aspects related to privacy. Ferguson et al. (2016) put forth nine ethical goals related to learning analytics including: student success; trustworthy educational institutions; respect for private and group assets; respect for property rights; educators and educational institutions that safeguard those in their care; equal access to education; laws that are fair, equally applied, and observed; freedom from threat; and integrity of self. Slade and Prinsloo (2013) provide three broad categories of ethical concerns related to learning analytics that include: the location and interpretation of data; informed consent, privacy, and the de-identification of data; and the management, classification, and storage of data. Khalil and Ebner (2016) summarized five categories of ethical issues including: transparency of data collection, usage, and involvement of third parties; anonymization and de-identification of individuals; ownership of data; data accessibility and accuracy of the analyzed results; and security of the examined datasets and student records from any threat. Steiner et al. (2016) put forth the following areas that have ethical issues: privacy; informed consent, transparency, and de-identification of data; location and interpretation of data; management, classification, and storage of data; data ownership; possibility of error; and role of knowing and the obligation to act. Slade and Tait (2019) provided a list of core ethical issues that should be considered across all regions of the world: data ownership and control, transparency, accessibility of data, validity and reliability of data, institutional responsibility and obligation to act, communications, cultural values, inclusion, consent, and student agency and responsibility. SoLAR (2021) notes these ethical concerns with learning analytics: privacy, opaque *black box* algorithms, basing classifications on biased datasets, and incorrectly predicting someone’s behavior. The following table provides a comparison of the studies noted above. Direct mentions of privacy are bolded while topics related to privacy are in italics.

Table 2
Comparison of Ethical Concerns

Ferguson et al. (2016)	Slade and Prinsloo (2013)	Khalil and Ebner (2016)	Steiner et al. (2016)	Slade and Tait (2019)	SoLAR (2021)
• student success • trustworthy educational institutions • *respect for private and group assets* • respect for property rights • educators and educational institutions that safeguard those in their care • equal access to education • laws that are fair, equally applied, and observed • freedom from threat • *integrity of self*	• the location and interpretation of data • *informed consent* • **privacy and the de-identification of data** • the management, classification, and storage of data.	• *transparency of data collection, usage, and involvement of third parties* • *anonymization and de-identification of individuals* • *ownership of data* • data accessibility and accuracy of the analyzed results • security of the examined datasets and student records from any threat	• **privacy** • *informed consent, transparency, and de-identification of data* • location and interpretation of data • management, classification, and storage of data • *data ownership* • possibility of error • role of knowing and the obligation to act	• *data ownership and control* • transparency • accessibility of data • validity and reliability of data • institutional responsibility and obligation to act • communications • cultural values • inclusion • *consent* • *student agency and responsibility*	• **privacy** • opaque *black box* algorithms • basing classifications on biased datasets • incorrectly predicting someone's behavior

These listings provide similar and singular examples of ethical topics to consider in relation to learning analytics. There are researchers who have considered the ethical aspects of learning analytics. Pargman and McGrath (2021) conducted a systematic literature review looking at which ethical topics have been addressed in studies that look at the ethics of learning analytics. The ethical topics addressed include: transparency, privacy, informed consent, responsibility, minimizing adverse impacts, validity, and enabling interventions.

Defining Data Privacy

Privacy remains a concept that at one level is understood, yet at another level cannot be completely comprehended as it is approached differently by each individual within a range of contexts. The most frequently agreed upon definition of privacy is the fact that there is no universal definition of privacy (Allen, 1988 and Margulis, 2011). Pavlou (2011) contents that this ambiguity comes from the fact that privacy is a complex concept that can be addressed from a range of different disciplines and perspectives. Within their review of research looking at information privacy, Smith et al., (2011) conclude that there is no definition of privacy that crosses all disciplines. Solove (2008) notes that “privacy is not reducible to a singular essence; it is a plurality of different things that do not share one element in common but that nevertheless bear a resemblance to each other” (p. 756).

The desire for privacy can be seen throughout human history with Westin (1967) providing an overview of numerous studies looking at how privacy is achieved in various cultures. Westin (1967) notes how the desire for privacy can actually be traced back to a need for privacy as seen in studies of animals focusing on population and health. A more recent look at privacy shows that the public desire for privacy increased throughout the mid-twentieth century to the start of the twenty-first century (Westin, 2003). There is a recognition that privacy is important for the well-being of individuals and the groups in which they live.

This recognition has resulted in privacy being recognized as a basic human right in numerous international agreements. The Universal Declaration of Human Rights put out by the United Nations in 1948 notes “no one shall be subjected to arbitrary interference with his privacy, family, home, or correspondence” (United Nations General Assembly [UN], 1948, p. 4). The International Covenant on Civil and Political Rights reaffirmed this earlier statement as long as the situation was lawful (UN, 1966a). The rights related to privacy in the American Convention on Human Rights includes topics such as honor, dignity, interference of private life, and protection of the law (Inter-American Specialized Conference on Human Rights, 1969).

Aside from its recognition as a basic human right, privacy is seen as necessary for a range of other human endeavors. Rubel and Jones (2016) provide a discussion of how privacy is critical for personal autonomy. Privacy is also recognized as a requirement for the development of new ideas (Richards, 2008). For many, privacy is also seen as a critical component within a free, democratic society.

However, even though privacy is internationally viewed as necessary and recognized as a legally protected right, there is not a universally recognized definition of privacy (Castelli, 2014; Weimann & Nagel, 2012). Instead, definitions of privacy will vary depending on context such as time period, cultural norms, physical location, and field of study. In considering this trait of privacy, Solove looks to “conceptualize privacy from the bottom up rather than the top down, from particular contexts rather than the abstract” (2002, p. 1092). Nissenbaum (2011) holds that privacy should be put into larger social contexts.

The various definitions of privacy can be put into broad categories that describe how privacy is approached. One of the earliest legal descriptions of privacy can be seen in Warren and Brandeis's seminal work *The Right to Privacy* (1890). Their broad definition of privacy is the *right to be left alone.* Tavani (2007) describes how this approach to privacy related to physical access has evolved to concerns related to interference in decision making and currently to concerns associated with personal information.

Tavani (2007) classifies this approach to privacy as nonintrusion. He then describes additional approaches to privacy including: seclusion where privacy occurs when one is alone, control where privacy can be found only if one has control over personal information, limitation where privacy entails limiting personal information in various contexts, and restricted access/limited control which has three components related to the concept, justification, and management of privacy. Moor (1997) also felt that the control/restricted access approach to privacy was the best approach when dealing with online information as it allows individuals some control over what information they share yet also recognizes that total control is not possible thus it is critical to restrict access to information to certain individuals within various contexts.

Privacy within Context

While many approaches to privacy focus on ideas related to limitation and control, Nissenbaum (2014) articulated the importance of context. She proposed focusing on a respect for context when considering privacy rather than developing specific rules for assorted contexts. Her model of contextual integrity considers two major cultural norms, the appropriateness of providing the information within the situation and the distribution of the information after it has been shared (Nissenbaum, 2004). Pavlou (2011) also highlights how context will impact the meaning and use of privacy. He notes how different disciplines focus on specific facets of privacy: right or entitlement in law, limited access or isolation in social psychology, and control in information systems. Heath (2014) provides an analysis of how Nissenbaum's approach to privacy can be applied within the context of learning analytics.

As a component of context, culture also impacts the application of privacy with societal norms providing a basis on what is typically acceptable (Margulis, 2003). Austin (2019) notes how the increasing infrastructure of surveillance may change socially accepted norms related to expectations of privacy and with that change it may be harder for an individual to decline the pressure to share information (Austin, 2019). This cultural pressure impacts learning analytics if one student does not give consent amid a group of other students who provide their information. There are also social norms related to the types of information people gather in various contexts. In an educational situation where students are consistently providing their thoughts and answers in order to earn a grade, they are socially primed to provide information to the institution.

This cultural context of privacy also has bearing upon social equity. Individual often provide companies with access to personal information in exchange for free or reduced pricing of their services. It is then harder for low-income individuals to choose privacy. Elvy (2017) describes two models that jeopardize the privacy of low-income individuals. The first is personal data economy where companies pay individuals for their data, a strategy will be appeal more strongly to low-income individuals. The other is pay for privacy where companies charge a fee to prevent data from being collected and use. Again this disadvantages those who cannot afford such a right.

Context also contains the idea of who has access to information. Austin (2019) notes that it is not the quantitative amount of information available that causes privacy concerns, but rather who has access to that information and their relationship with the individual which requires a consideration of privacy. Slade and Prinsloo (2013) enjoin that we should be “critically aware of the way our cultural, political, social, physical and economic contexts and power-relationships shape our responses to the ethical dilemmas and issues in learning analytics” (p. 3).

Privacy Principles

Many models related to privacy focus on choice, yet this is not universally accepted as the best approach with researchers noting that privacy must look beyond this singular approach (Austin, 2019; Solove, 2013; Tavani & Moor, 2001). Austin (2019) notes that “FIPPs provide a more robust set of principles that enable control over personal information than simply notice and choice” (p. 74). Recognizing the importance of privacy along with the fact that there is not an agreed upon approach to address privacy, there needs to be a method to evaluate how an individual can maintain their privacy. This is where privacy principles come into play. A principle is a shared value upon which regulations, rules, and standards can be built for the protection and advancement of the stated objective. In Wright and Raab’s (2014) in-depth discussion of privacy principles they note the importance of principles “because they form the basis for the formulation of questions that organizations can use to determine whether their new technology, system, project or policy might pose risks to one or more types of privacy” (p. 287).

Just as there has been a history of recognizing the right of privacy there has also been attempts to detail the principles which that right entails. One of the earliest compilations of privacy principles related to information data is the US Department of Health Education and Welfare’s Fair Information Practice Principles (FIPPs) from 1973. Another influential set of principles was put together and ratified by the Organization for Economic Co-operation and Development (OECD) in 1980. The OECD principles come about in response to protect data crossing national borders. One of the contributors of the OECD principles discusses the process and why this document continues to serve as a basis for data protection to the current day (Kirby, 2011). The ability of principles to withstand advances in technology and cultural change highlight that they are a useful tool in the evaluation of privacy practices.

For this research, the privacy principles used within the case study coding came from a prior study conducted by the author (Francis et al., 2020). In generating the list of overarching principles, several lists of principles were consulted. Some of these included: Fair information practice principles (United States Department of Health, Education, & Welfare, 1973), OECD Guidelines on the protection of privacy and transborder flows of personal data (Organization for Economic Co-operation and Development, 1980), Generally accepted privacy principles (Schroeder & Cohen, 2011), OASIS's Privacy Management Reference Model and Methodology (OASIS, 2016), and the U.S. Department of Education's Model Terms of Service (Privacy Technical Assistance Center, 2016).

Slade and Prinsloo (2013) provide a listing of principles and considerations specifically focused on dealing with information from learning analytics including: learning analytics as moral practice, students as agents, student identity and performance are temporal dynamic constructs, student success is a complex and multidimensional phenomenon, transparency, and higher education cannot afford to not use data. The application of the considerations they lay out closely align with the privacy principles noted. Pardo and Siemens (2014) also put forth specific principles related to learning analytics. They note how their principles may contain additional considerations within the larger categories of transparency, student control, security, and accountability and assessment.

Slade and Prinsloo (2013) note how it would be almost impossible to create a universal set of privacy guidelines to cover all context, yet they go on to state how it is possible to develop a listing of general principles which can then be used by the institutions to develop their context specific guidelines. This approach to find and address the common denominator within privacy allows for some consistency across education as a whole. Wright and Raab (2014) note how principles provide organizations with a clear understanding of the activities and expectations they should follow in relation to privacy. Ifenthaler and Schumacher (2015) note that "it is important to understand the implications of privacy principles to ensure that implemented systems are able to facilitate learning, instruction, and academic decision-making and do not impair students perceptions of privacy" (p. 16).

Privacy Harms

The discussion of privacy so far as focused on its importance, ways to conceptualize it, and principles to consider in its maintenance. Privacy harms are seen when those safeguards fail. Solove (2006) developed a taxonomy of privacy issues that can occur when data privacy is compromised. He laid out harms in relation to stages of the data information lifecycle. When information is collected, harms could come due to surveillance and interrogation. During information processing, harm comes from aggregation, identification, insecurity, secondary use, and exclusion. Information dissemination can see problems related to breach of confidentiality, disclosure, exposure, increased accessibility, blackmail, appropriation, and distortion. Finally, there could be an invasion of information from intrusion or decisional interference.

Citron and Solove (2021) expanded upon Solove's earlier work to describe the negative effects of a loss of privacy through a categorization of privacy harms. This work is necessary as court cases often require a demonstration of harm when making judgement. They detailed fourteen harms including: physical, economic, reputational, emotional, relationship, chilling effect, discrimination, thwarted expectations, control, data quality, informed choice, vulnerability, disturbance, and autonomy.

Another approach to consider privacy harms is provided by Calo (2011) who defines two categories of harms: subjective harms which result from perception and feelings of unwanted mental states such as embarrassment, fear, and anxiety and objective harms which result from the use of private information leading to a negative action such as identity theft. Harms can also be connected to specific privacy principles. Wright and Raab (2014) provide a listing of example harms when a privacy principle is not met.

Privacy of Groups

In his 1967 work, Westin notes that privacy not only relates to individuals, but also organizations. He notes how organizations need privacy beyond the privacy rights held by the individuals within the organization. Westin (1967) also explores how organizations have the right to maintain privacy in relation to protected communications whereby an organization gathers personal information from individuals. They must have the right to keep such information confidential and used only for the purpose for which it was collected. Margulis (2003) notes that groups of individuals may have a harder time maintaining their privacy than individuals.

The impact of group privacy within learning analytics is especially important as the benefits that come from the collected data often develops only due to the ability to collect data points from a large number of individuals. Loi and Christen (2020) make a case for the importance of group privacy by showing how often in cases of big data analytics an individual does not know that they have been put into a group and as such do not have the ability to consent or understand how their information will be used. This inclusion in ad hoc groups determined by computer analysis of characteristics infringes on the privacy of individuals. Mittelstadt (2017) looks at the rights that should be given to such groups which have been largely ignored in laws and policies which focus solely on the individual.

Laws Impacting Privacy of Students

Pardo and Siemens (2014) note that privacy and data ownership is at a low legal maturity level. Prinsloo and Slade (2016) also note how privacy policies have also lagged behind technological and cultural changes. While several years have passed since their analysis, there is still limited legal guidance on how privacy should be considered. While there are currently over 120 state laws addressing student data in some form, most of these focus on K-12 education rather than higher education (Student Privacy Compass, 2021).

The Family Education Rights and Privacy Act from 1974 remains the strongest legal code related to student data in higher education. More recent laws connected to data privacy relate to commerce and business dealings. While these have no legal oversight within higher education, the European Union’s General Data Protection Regulation and California Consumer Protection Act are discussed below as they are considered the new standard by which data privacy laws will be developed. Also, as consumers come to expect the privacy protections offered under these laws, they will demand the same protections in other context such as their education. Kay et al (2012) also note that because of the special mission of education, schools must consider ethical concerns beyond what is required in law.

FERPA – Family Education Rights and Privacy Act

The Family Educational Rights and Privacy Act (FERPA) was enacted in 1974 to establish a national policy for educational records to simplify concerns that arose due to varied state laws. The act defines academic records as "those records that contain information directly related to a student and which are maintained by an educational agency or institution or by a party acting for the agency or institution" (FERPA, 1974). Control over these educational records is provided to parents and students over eighteen years old. In a legal review of the Act, Daggett (2008) noted it contains four essential requirements. Individuals have the right to review their academic records. They may also request corrections if an error has been made. The individuals must be provided an annual notice of their rights. Finally, they must provide consent for the institution to release their records.

While this last right appears to put control of private information into the hands of the individual or their parent, Parks (2017) notes that in practice institutions hold almost all the power in regard to controlling academic records. There are two main components of the act that cause concerns related to individual privacy. First, is that FERPA focuses on how institutions are able to share information rather than how they use that information internally (Slade and Prinsloo, 2013). This means that institutions are able to share data from department to department. For example, admissions material can be shared with academic departments or student support services.

Secondly the restriction on sharing information is stretched as the law also allows the institution to share identifiable student information to anyone in the institution with a "legitimate educational interest" or to a third party who provides "institutional services or functions" (FERPA, 1974). This clause allowing third party access to student data allows student information systems, student success information systems, LMS, and other systems the ability to gather and utilize information for learning analytics. This means students are not required to provide consent for the sharing of their academic records as long as the institution determines that the collection and use of such data serves an educational function. Parks (2017) notes how with the increase in data collection and analysis it is almost impossible for students to avoid having their information used by the institutions limiting their power in asserting their privacy rights.

Given the fact that FERPA was established in 1974 it focuses mainly on access to physical records. Current academic records are maintained and stored electronically, yet there have been no updates to the law to reflect that change. In her review of FERPA in relation to Big Data, Parks (2017) concludes that FERPA is unable to address the legal and ethical concerns around use of student data. Daggett (2008) also concludes in a review of several legal cases connected with FERPA that “Students' privacy is not well-protected, and schools have disincentives to comply and uncertainty about just what compliance requires” (p. 113). Because the law as written does not provide protection for the students, some argue that the institutions themselves must step in to offer for comprehensive ethical consideration (Prinsloo & Slade, 2015; Tene & Polonetsky, 2013). Jones (2019) notes that FERPA should be considered the *floor* rather than the *ceiling* related to the safeguarding of student privacy.

GDPR – General Data Protection Regulation

The General Data Protection Regulation (GDPR) was put into effect in 2018 as a set of legal guidelines for the collection and processing of personal information from individuals who live in the European Union (EU). The guidelines also address the transfer of personal data outside the boundaries of the EU. The regulation must be followed by organizations that process the personal data of EU citizens or who provide goods and services to those individuals.

Brown and Klein (2020) analyzed 151 privacy statements from 78 institutions of higher education. Of those documents, 54 policies provided information related to the GDPR. Nearly all those policies noted how students did not fall under GDPR even if they were European Union (EU) citizens or the institution offered instruction within the EU. Brown and Klein (2020) went further in their analysis to note that those policies that did reference GDPR often changed the language in their policies to no longer refer to students or individuals but rather *data subjects* which is a term used within the regulation. They noted this change had the effect of removing individual rights and focused instead on data as a concept apart from a person. This idea of how the data about an individual forms a *data double* through a datafication process of individuals is discussed in more depth by Williamson et al. (2020).

CCPA – California Consumer Privacy Act

The California Consumer Privacy Act (CCPA) became effective in 2020. It was meant as a way to improve privacy rights and consumer protection for California residents. It gives more control to consumers in relation to how businesses collect and use their personal information. There are six rights provided to Californians in the bill: the right to know what personal information a business collects about them, the right to know whether a business has sold or disclosed their personal information, the right to request a business to stop selling their personal information, the right to access their personal information, the right to take private action in relation to privacy, and the prevention of businesses from denying equal service or prices for individuals who enact their rights under the law. While the CCPA provides additional privacy rights and expectations for individuals in relation to commercial interactions, it does not address data within educational settings. The only mention of education information within the bill refers to FERPA (California Consumer Privacy Act, 2018).

Current Models and Approaches to Evaluate Learning Analytics

Privacy Integration into Learning Analytic Systems

There have been various guidelines and structures proposed to guide the design of student information systems and student success information systems to account for privacy needs (Bellotti, 1997; Hoel & Chen, 2016; Horvitz, 1999; Jensen et al., 2005; Kitto et al., 2015; Langheinrich, 2001; Steiner et al., 2016). Many of these new processes try to operate under a privacy by design framework as put forth by Cavoukian (2012) which follows seven principles: proactive not reactive, preventative not remedial; privacy as the default setting; privacy embedded into design; full functionality – positive sum, not zero-sum; end-to-end security – full lifecycle protection; visibility and transparency – keep it open; and respect for user privacy – keep it user-centric. However, in designing systems, privacy is often simply one component and not a critical one. These proposals more frequently focus on minimizing privacy violations rather than proactively protecting individual privacy.

Learning Analytics for Institutions

It is important for institutions to establish policies and practices beyond the technical integration of privacy into learning analytic systems. Prinsloo and Slade (2013) reviewed the polices of two institutions in relation to ethical concerns and learning analytics. They found that institutions were not keeping up with the new abilities of learning analytics. To address the new state of learning analytics, institutions can develop and adopt codes of practice to guide the implementation of learning analytics on their campuses. Welse and McKinney (2015) discuss the need for codes of practice noting that their development not only maximize effectiveness and minimize risk, but they also build trust between the institution and its constituents through transparency. There are several guides that have been developed to assist institutions as they develop learning analytics programs. The DELICATE checklist asks a series of questions for institutions planning on implementing learning analytics (Drachsler & Greller, 2016). Cormack (2016) offers a framework when applying learning analytics that separates analysis of data from the intervention providing students with the option to make informed choices. The U.S. Department of Education also offers a guide to help institutions evaluate the privacy rights of students by analyzing student information systems' terms of service (Privacy Technical Assistance Center, 2016). In the U.K. there is a *Code of practice for learning analytics* that includes responsibility, transparency and consent, privacy, validity, access, enabling positive interventions, minimizing adverse impacts, and stewardship of data (Sclater & Bailey, 2018). A more general tool developed by EDUCAUSE, the *Higher education community vendor assessment toolkit*, offers a questionnaire to measure vendor risk related to issues such as privacy and security (EDUCAUSE, 2021).

Example Institutional Codes of Ethics and Policies

For those institutions looking to establish a code of practice related to their learning analytics system, Sclater (2016) provides an in-depth discussion and guide. His article discusses the process used by Jisc to develop their code of practice and then details what should be considered by other institutions. Included in the guide are questions related to ethical, legal, and logistical issues; a list of stakeholders with responsibilities; and a proposed action plan with steps to take in developing a code of practice. The article also includes the broad categories of details that should be included in a code of practice including: introduction, responsibility, transparency and consent, privacy, validity, access, enabling positive interventions, minimizing adverse impacts, and stewardship of data.

There are several codes of practice and institutional polices related to learning analytics that may be used as reference by those institutions looking to develop such documents. The following table provides some examples.

Table 3

Example codes of practice and policies related to learning analytics

Name	URL
Stanford CAROL & Ithaka S+R: Responsible Use of Student Data in Higher Education	https://ru.stanford.edu/
Jisc: Code of Practice for Learning Analytics	https://www.jisc.ac.uk/guides/code-of-practice-for-learning-analytics
Charles Sturt University: Learning Analytics Code of Practice	http://www.csu.edu.au/__data/assets/pdf_file/0010/2507824/2016-CSU-Learning-Analytics-Code-of-Practice_v3-3.pdf
The Open University: Ethical Use of Student Data for Learning Analytics	https://help.open.ac.uk/documents/policies/ethical-use-of-student-data
University of California: Learning Data Privacy Principles	https://www.ets.berkeley.edu/sites/default/files/general/uc_learning_data_principles_final03.05.2018.pdf
CUNY: Resolution Affirming the Privacy of Learning Data and Principles for Working with Third-party Vendors.	http://www1.cuny.edu/sites/cunyufs/committees/senate/standing/libraries-it/meetings-2019-2020/

Chapter 3

Methodology

This chapter contains the approaches, processes, and procedures that guided this research study. Chapter 3 includes information regarding the following: a) the purpose of the study, b) research questions, c) research method, d) site selection, e) data collection, f) data analysis methods, g) researcher background, h) verification of the study, and i) summary.

Opportunity for Change

This study examined how student privacy is addressed within a student information system focused on student success. The data within such systems are frequently utilized through learning analytics. As a developing field of study, learning analytics research will mature only through individuals conducting studies looking at all aspects of the field. This case study provides an in-depth analysis of one institution that has implemented a specific student success information system (system). The case study required data to be collected from a range of sources which ensures a full picture can be developed related to the topic.

This study looked specifically at how privacy is addressed within the system through an analysis of interviews with the system administrator, trainings offered on the system, interviews with the faculty members who use the system, documentation from the company, documentation from the institution, and an interview with a company representative. All of these sources of data provided insight into how privacy is addressed. This understanding of how privacy is currently addressed then allowed for the development of recommendations for the institution. While this case study does not represent the practice of all institutions, it does provide an example that can be considered for those implementing their own student success information systems and utilizing learning analytic approaches on the ensuing data.

Research Questions

The following research question guided this study:

- How is the concept of privacy addressed in relation to a student success information system within an institution of public higher education?

Sub questions include:

- How were the policies and procedures related to student privacy within the system developed and implemented?
- How is the need for privacy balanced against the institution's functional data needs?
- How does the institution weigh individual privacy rights against group benefits?

While this case study is exploratory, given the review of literature there are two propositions related to the findings of the study. These include:

- Privacy will not be addressed fully in the learning analytics system due to a focus on institutional rather than student benefit.
- Privacy will not be addressed fully in the learning analytics system due to a belief that the academic benefits from the learning analytics process outweighs the need for individual student privacy.

Research Method

This study was a single case, descriptive case study. Yin (2009) offers three situations in which a case study is the preferred research method. These include asking *how* or *why* questions, the researcher cannot control events, and the topic deals with a contemporary phenomenon in specific contexts. All three of these situations were present in the current study. First, the research question guiding this study asked how privacy is addressed within a student success information system. Next, as an exploratory look at the concept, the researcher was not able to direct specific actions or decisions related to the issue.

Finally, Yin (2009) notes that a case study is an empirical inquiry that looks at a contemporary phenomenon within a specific context because the boundaries between phenomenon and context are not clear. For this study, it was more powerful to consider the situation of privacy within a student success information system at an operating educational institution. In this way, it was possible to view the topic within the large context of the institution. The impact of outside variables on the idea of privacy must be considered to get a full picture of the concept rather than through a sterile, isolated look at privacy on its own. The case study allowed for a holistic look at the processes and everyday events that impact the consideration of privacy within the student success information system.

The rationale for deciding to perform a single case study focused on the idea that the specific institution selected for the case would be a representative case. The institution selected as the case study is a typical situation, and as noted by Yin (2009), such cases can be informative about the experiences of an average institution. By gathering details on how the topic is addressed by this institution it is possible to garner insights into how it is addressed elsewhere.

Site Selection

The institution of higher education (IHE) chosen was used to represent the abstraction of student privacy within a student success information system. The IHE is a public university within a statewide system in the Midwest. The IHE has an FTE enrollment of ~2,000 students. It offers associate through doctoral degrees with a special focus on technology. The unit of analysis for the study was the institution’s student success information system. This location was chosen as it provided me with access to the individuals involved in the maintenance of the system, user level access to the system, and institutional documents related to the system. I was not a user of the system before this study which allowed for a more objective consideration of the use of the system.

The IHE has been working with the company since 2018 with the official launch of the system in March 2020. The system provides a systematic method to collect data on students as well as serves as a means of communication between students, faculty, and other departments on campus. The company has over 500 educational institutions using the system which highlights how this study may be applicable to other institutions and situations.

Data Collection

This study was a single instrumental case study. As an instrumental case study, this study used a particular case in order "to gain a broader appreciation of an issue or phenomenon" (Crowe et al., 2011, p. 2). The intent was to use one bounded case in order to describe the topic of student privacy. The purpose of this study was to describe how student privacy is addressed in relation to a specific student success information system used for learning analytics by the IHE through analysis of interviews with the system director, faculty, administration, and related documentation.

As a case study, data was gathered from multiple sources in order to allow for triangulation of results. Yin (2009) discusses how triangulation of data occurs when the facts of a case study are supported by more than one piece of evidence. This is one of the strengths of case studies in that they allow for internal confirmation of the research findings allowing me to make a stronger case for their conclusions. This triangulation helps with the research validity and credibility of the study. This confirmation of findings from different sources also provides confirmability (Shenton, 63).

Interview Data Collection

Semi-structured interviews were conducted with the system director at the IHE. These interviews provided background information on the history of use of the system at the institution, the institutional goals of using the system, and how student privacy has been considered in the implementation of the system. Two interviews were held with this individual with coding occurring after each interview in order to ensure a comprehensive understanding of the system was developed.

From the system director, snowball sampling was employed in order to get the contact information of faculty members and administrators who were users of the system. These individuals were interviewed to pose questions related to why they use the system, their thoughts about the system, and any concerns they had related to privacy.

A semi-structured interview was also conducted with a representative from the company. This interview provided details on how the company approaches student privacy and the structural systems which they put into place to allow for institutions using their product to provide privacy. Questions were also asked to highlight how the company suggests institutions implement the system in order to get benefits. After the interviews, the transcripts were shared with the participants to order to allow for member checking which impacts validity and credibility.

Semi-structured interviews allow for a fluid approach to data gathering by allowing the researcher to both ask set questions and explore topics in more depth depending on the interaction and engagement with the interviewee and their answers. Documentation on the interview protocol includes: the invite to participate in the study (Appendix A), consent form

(Appendix b), interview questions for system director (Appendix C), questions for the company representative (Appendix D), and questions for system users (Appendix E).

Table 4

Interviews conducted

Interviewee	**Connection with system**	**Times interviewed**	**Pseudonym**
System Director	Oversaw the implementation of the system and manages the daily operations necessary for the continued use of the system. Provides training.	2	Judy
Administrator #1	Part of the team that implemented the system.	1	Paul
Administrator #2	Utilizes the system as an administrator looking for trends and data to reach decisions.	1	John
Administrator #3	Part of the team that implemented the system. Utilizes the system as an administrator as well as a faculty member.	1	Jim
Faculty #1	Utilizes the system within their roles as a faculty member teaching courses and as a student advisor.	1	Bob
Faculty #2	Utilizes the system within their roles as a faculty member teaching courses and as a student advisor.	1	Bill
Faculty #3	Utilizes the system within their roles as a faculty member teaching courses and as a student advisor.	1	Rose
System Representative	Employee of the system's company. Aids institutions utilizing the system.	1	Chris

Resource Data Collection

In addition to an interview with a representative from the company, data was also gathered through the analysis of company supplied material. This included training materials and their online help center. These documents were reviewed to find references to student privacy and privacy principles. Institutional documents were also analyzed. These included: any written reports and updates on the system, emails sent about the system, policies and procedures related to the system, and documentation created by the institution.

Finally, data was gathered from the training sessions offered by the institution for the faculty and staff members who utilize the system. Analyzing these training sessions not only provided information on how the institution addressed student privacy, but it also was used to gain insight into whether faculty and staff considered student privacy as noted through any questions or comments made. Overall, 53 additional data resources were analyzed.

Ethical Considerations

Approval for this case study was obtained from the University of South Dakota's IRB panel. Confidentiality of the interviewees was maintained. Participants gave consent to have the audio of the interviews recorded. Transcripts were first generated automatically using the Zoom captioning capabilities. These transcripts were then reviewed and confirmed with identifying names removed and the inclusion of pseudonyms. The recordings were deleted with the transcripts will be kept for three years on a protected computer in compliance with the IRB requirements.

Data Analysis

Two approaches were taken to the coding of the resources analyzed in this case study. The first was the use of open coding using an inductive approach. In this process, themes were developed through an open reading of the sources. The data was reviewed several times to allow for refinement of the final themes shared in the reporting of results. The themes developed by the open coding were analyzed by considering how they relate to the research questions noted earlier. The themes were also discussed with the system director at the IHE to get feedback and provide confirmation that the topics were appropriate.

After the inductive coding, the data sources were also analyzed through a deductive coding process using descriptive coding. In deductive coding, a codebook is developed before analyzing the content by drawing from past research and theory on the topic. Descriptive coding allows for the summarization of the content of the text into a description that encapsulates the idea. The listing of privacy principles served as the descriptive codes used in this part of the analysis. This predetermined list of concepts fits well into descriptive coding in that the data sources may not use specific terms but rather relate to ideas that fall within the categories.

When structuring the timing of the coding, the inductive coding process was done first to avoid limiting the content analysis as cautioned by Creswell and Poth (2018). The inductive coding utilized a categorical aggregation approach where a collection of instances were analyzed from the data in order to develop issue-relevant meanings. This approach worked well as the interviews each provide a unique instance of working with the system and through this analysis similar themes were uncovered. These themes were then confirmed through analysis of the data resources.

Researcher Background

Within qualitative research, the researcher takes on active part within the study. Because my engagement with the participants and resources directly impacted the results, it is important to provide clarity on my background and relationship with the topic. I selected this topic due to my interest and past research on the topic of privacy. Since January 2019, I have been a part of a privacy research lab on my campus. Through that lab, I have researched, authored, and co-authored assorted conference papers and journal articles looking at the place of privacy within international law, librarianship, and data privacy management. This work has reaffirmed my believe in the importance of protecting and maintaining individual privacy. Overall, I fall into the category of strict privacy where I do not share personal information and decline to participate in many technological functions that I believe infringe on my right to privacy.

Recognizing my belief in the importance of privacy, I was conscientious in my interactions with the interviewees within the study. When posing questions and comments I avoided making comments or using a vocal tone that implied judgement related to the privacy protections or lack thereof which were found within the system. I highlighted with the participants that I was not looking to judge the system but rather to determine what is currently in place within procedures and practices.

My prior research, especially related to the privacy principles, was an aid in the descriptive coding process as I analyzed the data. Privacy is not a concrete item, but rather is made up of several concepts. Understanding these concepts allowed me to pull connections to meaning from spoken comments or written resources.

Verification of Study

When considering issues related to the impact of a study, it is important to address issues related to reliability and validity. There are several aspects of research design which can address these issues. Yin (2009) looks at how case studies can address construct validity, internal validity, external validity, and reliability. In relation to construct validity, this study used multiple sources of evidence, established a chain of evidence, and had key informants review drafts of the report. To address internal validity, this study matched patterns during data analysis and used logic models. External validity was addressed by placing the study within the literature and theory surrounding the topic of student privacy in learning analytics. Finally, reliability was addressed by providing a detailed account of the research protocol used within the case study.

In addition to Yin (2009), there are other authors who also provide insights on how to address validity and reliability issues when conducting qualitative research. Shenton (2004) provides qualitative specific language related to these validation concepts. He refers to internal validity as credibility, external validity as transferability, reliability as dependability, and confirmability as objectivity. This study will implement several of the suggestions offered by Shenton (2004) within those areas. In relation to credibility: the study followed established research methods within case studies, the researcher had a familiarity with the culture of the organization under study, there was a triangulation of data, tactics were used to ensure honesty from participants, member checks were done with the participants, and the results were examined in relation to past studies. Transferability was addressed by providing detailed descriptions of the organization that was studied and the methods used to collect the data. This detail also helped address the dependability of the study. Finally, confirmability was addressed both through triangulation of the data and personal reflection by the researcher related to thoughts, prior beliefs, and expectations regarding the study.

Creswell and Poth (2018) also provide guidance for individuals to consider when validating case study research. They provide six questions in their evaluation of case studies including: identify the case studied, present a rationale, describe the case in detail, articulate the themes, report assertions or generalizations, and embed researcher reflexivity or self-disclosure. These details correspond to many of the strategies put forth by both Yin (2009) and Shenton (2004) and were addressed within the study.

Summary

This study used a single case study approach to answer the question, how is student privacy addressed in a student success information system. Data was collected and triangulated from numerous sources within a representative institution. A descriptive and open coding process were used to find themes related to the research question. Throughout the data collection and analysis, assorted techniques were implemented to raise the validity and reliability of the study.

Chapter 4

Results

This chapter presents the findings of the case study herein. It first presents a background on the institution's use of the student success information system. Then, this chapter presents the three main themes found during the inductive coding process and provides detail on how these themes fit within the larger literature on the topic of learning analytics and privacy. The chapter then details how the case study addressed the privacy principles through a deductive coding process. This chapter ends with discussion and suggestions for in this area.

Background of Student Success Information System

The IHE contracted with the company to implement the student success information system. In one of their early training videos, the IHE provided an overview of the system to faculty noting it

> Helps advisors and support teams quickly and easily reach out to students in need of extra guidance, connects everyone on campus, from deans and faculty to financial aid, tutoring, and residential life in a collaborative network to support students. Empowers students themselves with the tools they need to stay on track and plan their entire college journey. And gives leadership the insights they need to make informed, strategic decisions and build a culture of student success. (Artifact 15)

The system was able to meet these claims due to the use of student data. The IHE highlighted that "student data is one of the most important tools we have to foster student success" (Artifact 55).

The company itself also highlighted the ways the system could assist institutions by providing "comprehensive technology that links the administrators and faculty, staff and advisors in a coordinated care network to support students from enrollment to graduation and beyond" (Artifact 39). The company noted "the notion behind the platform is to assist the connection to each other and data. Assisting in making sound decisions in a user-friendly system" (Artifact 39).

With the purpose of the system in mind, the IHE contracted with the company and set up the behind-the-scenes structure of the system. The system was ready to be rolled out to the campus in the spring of 2020. This was when the IHE moved to remote learning with the expansion of the COVID-19 safety protocols. This was both a benefit and detriment to the new system. The timing was a positive in the fact that the new system allowed for additional communication capabilities between faculty and students. Faculty could now text students to check-in on their health and academic concerns. The push to all online courses also highlighted those additional communication features as face-to-face options were no longer available. These features of the system resulted in an early adoption of the system by some faculty members.

However, rolling out a new system during such a time of upheaval also caused issues. Faculty were not given an introduction and full training to the system before they went remote. This meant there were issues with faculty not understanding the need, functions, and use of the system. The Fall 2020 trainings that were offered at the beginning of the next school year saw faculty asking what exactly this system was.

Since the initial roll out, the IHE offered various training sessions on the use of specific features of the system. These were offered virtually, in a hybrid format, and via one-on-one training. Video recordings of the trainings were available for faculty to view, and email notices went out to faculty when they needed to engage with the system; for example, at the beginning of each semester, the faculty were requested to complete a progress report on students during the first weeks of class in order to note whether the students were attending or engaging with course content.

The IHE used the system to accomplish several institutional needs. First, it was used as a progress check early in the semester to determine if a student attended or participated in a course. This helped in correcting and ensuring registration records. The institution also used the alerts and communications sent within the system to make decisions. An email sent by the provost's office noted, "The primary source I have to make decisions about students' continued enrollment and respond to complaints from students and parents is your alerts and comments" (Artifact 59). Finally, the IHE used the system to set up appointments with various offices such as advising, student housing, and financial aid. While allowing for a more focused, unified, and systematic approach, this also meant the students had little choice on the use of the system.

Inductive Coding to Answer Research Questions

After conducting interviews and gathering resources from the IHE and the company, an inductive coding process was utilized to uncover themes. Overall, three main themes were discovered in relation to the main research question: *How is the concept of privacy addressed in relation to a student success information system within an institution of public higher education?* The first theme was that privacy could be contained within the institution's adherence to the Family Educational Rights and Privacy Act (FERPA). The second theme highlighted specific methods used to maintain privacy including limiting access to information based on individual roles and ensuring technological security protocols. The third, final theme highlighted concerns raised about the relationship between students and their data.

FERPA Means Privacy

FERPA and other mentions of legal limitations were addressed across all the data sources including interviews, company documents, and institutional documents. The company often provided a default mention of FERPA to provide a warning to institutions as they worked with different data sources and features such as adding student demographic information to the system or adding notes to a student's record. "Do not do this unless you are aware of your institution's IT policies on data imports, privacy, FERPA, and other relevant policies" (Artifact 22). As mentioned in the company's training guides and help center, institutions needed to be aware of FERPA as "any information you enter into [the system] pertaining to a student becomes part of their official student record. It may be subpoenaed by the student as outlined in the Family Education Rights and Privacy Act (FERPA)" (Artifact 27). The company avoided providing specific guidance on how to comply with FERPA regulations.

This generic mention of FERPA was also seen in the training materials put out by the IHE. When talking about using the system, it was noted, "We do have FERPA as a law and something that we follow" (Artifact 56). While it was not described in any detail, FERPA was invoked by describing how the system was initially set up by one of the administrators who took part in that process, "[The institution] adheres to the system level FERPA policy" (Artifact 49). The IHE made a conscious effort to address FERPA. It was also noted how the IHE went beyond some definitions of directory information from FERPA and included student emails as personally identifiable information. Paul noted, "Ours is a little bit more restrictive than the [governing] board level" (Artifact 49).

Even though FERPA was frequently mentioned, most documents and interviewees provided a cursory understanding of what FERPA entailed. The most detailed note on FERPA was from an email sent to faculty members noting “FERPA expressly allows for sharing of students’ educational records with staff who have a legitimate educational interest in providing a service that benefits students” (Artifact 6). However, in the other trainings and documents reviewed, FERPA was acknowledged as important in relation to student privacy, but specific aspects of the law were not discussed. There was a general consensus in the IHE interviews that FERPA was being followed. “Part of the tight lockdown on access to the information was […] concern about FERPA compliance” (Artifact 57). This belief in the adherence to FERPA led to a belief that student privacy was being protected. After describing FERPA, John noted, “I think those are the basic things that keep student information secure” (Artifact 50). Rose, after being asked about privacy, stated, “We just got an email today saying that the FERPA instruction were online” (Artifact 60).

The literature did not stress compliance with FERPA since the activities involved with learning analytics were permissible under the law. While FERPA required students’ consent to share their academic records with a third party, it did not impact any sharing within the institution. This exception was highlighted in the resources provided by the IHE. Beyond internal use, Parks (2017) noted that institutions were “Free to share any information in a student’s academic record with any third party that they designate a ‘school official’” (p. 26). This then addressed any concerns about the external company collecting student data.

Given the fact that FERPA was not actually impacting the use of student data in this situation, it was concerning that it appeared so frequently within the discussion of student privacy at the IHE. Some interviewees focused almost solely on FERPA as the answer to privacy. While there are numerous other issues related to student privacy, for many of the interviewees, FERPA appeared to be the extent of the individuals’ knowledge on privacy. There appeared to be the sincere belief that student privacy was addressed. Interviewees at the IHE did not show any awareness of the additional privacy issues related to learning analytics. While the IHE was acting in good faith by addressing current FERPA requirements, the literature called for movement beyond the law. Parks (2017) concluded "FERPA is unable to address many legal and ethical concerns around current uses of student data" (p. 24). Due to this, some authors called for institutions to move beyond FERPA (Jones 2019; Prinsloo & Slade, 2015; and Tene & Polonetsky, 2013).

Methods to Maintain Privacy

While FERPA was seen as a general aspect of privacy, when asked about more specific measures taken to protect privacy, several of the interviewees were unable to come up with additional items. Bob noted, "I don't know of any. I really don't" (Artifact 47). John had a similar response, "I don't know" (Artifact 50). Bill also shared this level of understanding, "Nope, I don't. I would have no idea" (Artifact 58).

The users' lack of understanding related to student privacy is an important consideration moving forward for the IHE. While some of the literature on learning analytics considered the place of faculty members within the system, to the best of my knowledge were not any studies that looked at faculty and privacy specifically. These answers provided examples of some users' experience and knowledge.

For those interviewees who did have ideas on measures that were taken to promote and protect student privacy, there were two main methods that were discussed. The first dealt with limiting the type of student data accessible to individuals based on their roles. Rose stated,

> I always assumed anything that I was limited to was based upon a law. I assume, that whoever set this up does so with as much positive intent as possible. That faculty and

> administrative staff know the law, and that we've done a good job of informing people of what they can and cannot show. So for that reason, I guess I just assume that whatever I don't have access to is a legal restriction, not somebody just restricting it because it's not necessary." (Artifact 60)

This role-based access to the system was one of the foundational considerations when the system was set up. "Permission access points were migrated from the shared student information system in terms of roles" (Artifact 49). Jim described the setup of the roles noting they were

> Very thoughtful about what sort of information should be collected and who has access. […] That was very deliberate. And there was actually a lot of conversation about that, and that's also why in the beginning it was restricted so much, was in protection of students. (Artifact 57)

Within the ethical guidelines put out by the IHE was a section telling users to "access only student data that is relevant to your role. Some […] may allow access […] outside your role […]. By using data related to your role, you can make the greatest impact and maintain compliance with federal guidelines, such as FERPA" (Artifact 55). The company itself also noted the importance of roles, describing how the system

> Provides the granular permissions necessary to ensure that only educational representatives that have legitimate need and right to see a student's information (courses scheduled, credit accumulation, degree progression, etc.) can access that information. The system provides role-based access, allowing access to certain data to only those users with sufficient privileges. (Artifact 34)

The importance of the use of roles to limit access to types of data mirrors the privacy and learning analytics literature which talks extensively on limitation of data. Pavlou (2011) and

Moor (1997) both noted, in their discussions of privacy, that it should be up to each individual to decide who has the ability to access their data. Austin (2019) built on this idea noting that it was not the quantitative amount of information available that caused privacy concerns, but rather who had access to that information and their relationship with the individual. Within the learning analytics literature, Slade and Prinsloo (2013) noted how limiting access to data to authorized individuals was one feature of a secure system. This was seen within the documents and interviews of the case study. This understanding of the importance of access was highlighted in the UK's *Code of Practice for Learning Analytics* included access as one of the features institutions should consider when establishing a learning analytics system (Sclater & Bailey, 2018). While developed in the UK, this code covered general aspects of learning analytics while allowing for individualized implementation based on location and local needs.

At the IHE, privacy was also seen within a lens of technology security. The director of system noted, "There's a lot of stops that we have in place that would try to make it so that nobody would just get in, you know it's all in a single sign on" (Artifact 48). John noted, when asked about privacy measures in the system, "It is password protected, so I think those are the basic things that keep student information secure" (Artifact 50). Users of the system were also encouraged to use standard security practices. During training, it was noted to users that when using the system in places where others could view your computer "Don't leave it open […] or walk away" (Artifact 56). The company also provided details related to the technological security aspects of the service,

> All emails stored in the […] platform are encrypted at rest which prevents unauthorized access or theft in the unlikely event that the raw data is accessed by unauthorized agents.

> The encryption keys are stored separately from the data and are updated on a regular basis. (Artifact 33)

The importance of the technical security issues addressed in the interviews was also seen in the learning analytics literature. Privacy and technical security were always tightly bound as security was required due to the need for privacy. In Cavoukian's (2012) privacy by design framework, she included end-to-end security as one of her seven principles. Due to the connection between privacy and security, it was often mentioned in lists of ethical concerns related to learning analytics (Khalil and Ebner, 2016; Slade and Prinsloo, 2013; Steiner et al., 2016). Pardo and Siemens (2014) also noted how security impacted learning analytics.

Students' Connection with Their Data

In general, the administrators and users of the system did not have concerns about student privacy. "It's kind of the unspoken expectation that we respect that kind of stuff" (Artifact 47). Jim noted, "We have to trust the people that you hire. That they'll use the material in the right way" (Artifact 57). These comments highlighted an idea that privacy was only a problem when misused by faculty and staff to benefit themselves. There was no nuanced belief that privacy could be violated without a specific breach or harm. For example, there was no mention of concern with the type of data gathered or how long it was maintained. Overall, the system was seen as an institutional good, and as such the procedures were also seen as good.

There were a couple of suggestions offered, however, that interviewees felt would increase the privacy of the system. Bob noted "I don't know if students know what kind of system this is and if they've signed off saying, 'I'm okay with that'" (Artifact 47). This concern with student consent moved beyond basic agreement to deep comprehension with a desire for students to “really understand it with all the other things going on in their life when they first arrive on campus” (Artifact 47).

Creating a system that allowed students to provide informed consent was critical in the literature surrounding learning analytics (Slade and Prinsloo, 2013; Slade and Tait, 2019; and Steiner et al., 2016). Jones (2019) noted how, historically, the idea that an individual had a right to control who knows specific information about themselves had long been central to privacy. This loss of control might occur either through institutions not asking for consent or by asking for consent without providing a clear description as to how the data would be used. The requirement of having a clear description was also noted in the interview as consent without understanding does not allow for meaningful consent. Slade and Prinsloo (2013) added that when asking for consent, institutions should also provide details on the possible benefits and harms resulting from sharing or not sharing the information.

Another concern brought forward by faculty was the ability to correct data if necessary. Rose noted, “It’d be cool if students could update certain aspects like demographic or stuff that wouldn’t require another person to integrate” (Artifact 60). The ability of an individual to correct information was important as it addressed two components related to privacy. First, this meant that students had access to the data connected to them. This openness and transparency on what information the institution was collecting about them allowed for greater trust and cooperation. Second, the ability to make corrections allowed students to keep a more accurate record of themselves. Slade and Tait (2019) noted that students should have access to their raw and analyzed data so they could make corrections as necessary. Ferguson et al. (2016) also included, in their list of challenges with learning analytics, that institutions should offer opportunities to correct data, but added that this process should be publicized so students would know it. This ability to make corrections is codified within FERPA with students being able to correct errors in their educational records (Daggett, 2008).

Research Sub-questions

Overall, there were three sub-questions posed in this case study. In reviewing the interviews and other documents for answers to these sub-questions, what became apparent was how new the IHE was to learning analytics and student success information systems. The answers showed that little thought was given to the below questions at this point in the lifecycle of the program. Given the length of time the system was in place at the IHE, this maturity level was not unexpected, and it did provide the IHE with the opportunity to dedicate focused time and discussion on how the system would move forward. This conscious consideration of the systematic use of the learning analytics features of the system is critical for a successful future implementation.

How were the policies and procedures related to student privacy within the system developed and implemented?

Ensuring that the new system address the policies of the IHE was an important consideration in its implementation. When asked if there were policies or procedures that were required in setting up the student success system, Judy noted, “They [the Board of Regents] let each school sort of implement in their own way, which is good” (Artifact 48). Paul noted, however, that there were some specific Board of Regents (BOR) and institutional policies that impacted student data. In looking through the policies, three were found to be related to students and privacy. First, the IHE had a policy on the Privacy of Student Records. Second, the BOR had a policy on Public Access to Student Directory Information and, third a policy on the Confidentiality of Student Records. These policies focused primarily on FERPA requirements.

Beyond the official policies related to the system, the IHE developed a document titled *Ethical and Impactful Use of Data in Support of Student Success*, which covered the three broad areas of: (a) the importance of using data in working with students; (b) ethical use of data; and (c) guidelines for how to use, interpret, and apply data (Artifact 55). This document provided more of a guiding force in the application of the system. It had specific guidelines for faculty and staff regarding using student data.

When reviewing the official policies, their general nature was evident. They provided broad guidelines on the restriction of student records and data within the abstract rather than specifically within a student success system. Further, these policies were updated many years ago. One policy was last updated in 1993 and the other two were updated in 2011. Notably, the first International Conference on Learning Analytics and Knowledge occurred in Canada in 2011. This again meant that there was no connection between the policies and the abilities of learning analytics. Big data analytics and even types of faculty/student communication changed over the ensuing 11 years. This gap between policy and application was not unique. Prinsloo and Slade (2013) found institutions were not keeping up with the new abilities of learning analytics.

How is the need for privacy balanced against the institution's functional data needs?

Within this case study of student information within a student success system herein, privacy could relate to the individual's ability to choose whether to, and how much, they engaged in the system. When asked about the requirement of students to use the system, Judy noted, "They can opt out of like text messages for instance" (Artifact 48). It was also possible for students to "withhold public directory information by notifying the Office of the Registrar in writing" (Artifact 52). However, the students could not opt out of the system completely. The system kept track of items such as GPA, contact information, ACT scores, etc. The students also could not control if a faculty or staff member opened a case on them and included notes connected to their files.

The concern that the institution would push its interests over the students was discussed by Selwyn (2019). Rubel and Jones (2016) noted that institutions must also answer the question of who would benefit from the data. Would the analytics support individual students or the institution? While the lack of choice on whether to engage with the system or not did point to a focus on the institution rather than the student in this case study herein, the interviews and documents related to the system showed a deep concern for helping the student. “When the data shows areas for improvement, take action on that information” (Artifact 55). When asked about the goals of the system, John stated,

> Obviously increasing student learning outcomes. Another I would hope it would just also improve the general quality of experience they have here […] help them be healthier, happier, more fulfilled in their four years here, so when they leave here they’re better individuals. (Artifact 50)

Judy added that through use of the system, “My ultimate goal is that it does connect to the right resources, it does create an opportunity for timely feedback and follow up” (Artifact 48).

Some resources reviewed even placed the student above the institution. “Use data to support students. Sometimes the best action for a student may not be the most beneficial to your department or office, but we should act in a student’s best interests” (Artifact 55). This desire to put students first might cause extra work for the institution to provide additional resources. Judy noted, “What I hate doing is saying ‘hey we saw that you have this potential red flag that we have no support backing that up’” (Artifact 48). Overall, the IHE put forth a belief that it was not an either-or situation in relation to who benefitted, rather as Bob stated, “I think it could be very useful and mutually be a win win for the student and a win win for the institution” (Artifact 47).

The company itself also highlighted how the system could assist individual students. The system’s “student interface helps you to connect with students on their terms, building belonging and a deeper sense of purpose at your institution” (Artifact 16). In describing the system, the company’s literature noted that it was able to “unite staff to deliver proactive, holistic support to students” (Artifact 16).

The idea that an institution had the responsibility to act upon any information that it collected and generated when utilizing learning analytics was discussed in the literature by Slade and Tait (2019), Pargman and McGrath (2021) and Sclater and Bailey (2018), note especially when an intervention was shown to help a student, it could be considered negligent to deny access to that resource. The ability for the IHE to follow up with a student when they miss a success marker was a concern of Judy who posed, “What’s the level of follow up to support those students (who are marked as not meeting a goal)?” (Artifact 61). This quote showed the concern at the IHE not only for meeting goals of the institution but also supporting individual students.

How does the institution weight individual privacy rights against group benefits?

The IHE used the student success system more for its communication functionality rather than its learning analytics capabilities. It was within the learning analytics that the group benefits were found as individual data points were analyzed to predict behavior and provided more proactive suggestions to students. These group benefits were then taken in the aggregate to reach institutional goals related to retention and graduation rates. Given that the learning analytics were not fully utilized, there was not enough information to fully answer this question.

Deductive Coding of Privacy Principles

While the interviews and other resources were first coded using an inductive process to find themes to address the research question and sub-questions, the data was also deductively coded to determine if and how the IHE addressed the privacy principles.

Notice: Subjects will be informed of data collection and use policies.

The IHE did not address this principle.

Retention: Data will be removed when no longer required.

There was no time limit given on data retention. Data continued to be added to the system with no structure for removal after a certain period.

Minimization: Limit the amount of data collected, processed, and stored.

In the interviews, the quantity of data seemed to be limited more by system abilities rather than institutional decisions. When asked what data was stored in the system, Judy quipped, "What isn't in the system?" (Artifact 48). This lack of minimization is a function of the idea behind big data analytics, where more data allows for the combination of more variables to find connections. Within two faculty interviews, comments were made that it would be better to integrate the system with the IHE's learning management system to integrate course grades and other data together.

Use restriction: Data can only be used for defined and accepted purposes.

No policies or procedures were explicated established laying out how the data would be used. Broad categories of individual student success, retention, and graduation rates were discussed, but there was no plan to meet those specific goals. The company also provided broad descriptions of data usage mentioning general goals rather than specific usage. "Student success analytics and predictive modeling help you understand which interventions are working and how to best adjust your strategy” (Artifact 16).

Security: Data is handled in accordance with appropriate security principles. Security was part of data management. Security measures such as multi-factor authentication were in place to sign into the system while the system itself met standard security protocols. The company noted it aligned its policies with the ISO 27001 framework which is the world’s best-known standard for information security management systems.

Quality: Data is accurate and kept up to date. There was some concern brought forward in the interviews that data within the system might not be correct. This stemmed from the initial uploading of data into the system from an older student information system. This concern was seen to be less of a problem as the system was used and data was inputted directly. The IHE did caution users “if the data does not align with your internal records […] verify the integrity of the data” (Artifact 55).

Access: Subjects have the right to know what personal data is being held about them. Student access to their data within the system could be varied. In their instructional material, the company noted “students cannot view their issued ad hoc alerts […]. However, if your institution has configured the alert notification emails, then students may get an email after those alerts are issued, and the email may include the alert reason” (Artifact 38). This limitation of access was also confirmed by Judy who said, they could “create a note about them. […] I can share with the student or not share with the student” (Artifact 48). However, as highlighted by the company, “Any information you enter into [the system] pertaining to a student becomes part of their official student record. It may be subpoenaed by the student as outlined in the Family Education Rights and Privacy Act (FERPA)” (Artifact 25). Students’ ability to request their information as required under FERPA did mean that they had access to their data; however, the systems set in place obscured the students' process of requesting their data.

Participation: Allow data to be corrected and deleted by the subjects of the data whenever appropriate.

Students had the ability to correct some data within the system. This was limited to factual data such as contact information. One faculty member did bring forth a concern that students should be able to correct data that was incorrect.

Enforcement: Data holders must comply with applicable policies, laws, and standards.

The training and setup documents provided by the company reinforced the need for institutions to base their usage of the system within appropriate laws and policies: "Review your institution's privacy policy and make sure to change access to this report if needed, based on the policies and laws in your location" (Artifact 26). They highlighted that "you are solely responsible for your compliance with FERPA" (Artifact 29). The company also provided information to users within their privacy policy which noted, in part,

> This means that the main responsibility for data privacy compliance lies with your institution as a 'data controller.' It also means that your institution's privacy policy governs the use of your personal information (rather than ours). Your institution determines what information we collect through our products and services and how it is used, and we process your information according to your institution's explicit instructions. (Artifact 21)

The IHE itself also ensured it followed FERPA guidelines as seen through the interviews and addressed in its official polices.

Consolidation: Consolidation of databases containing personal data cannot be done.

The student success information system pulled much of its student data from the IHE’s student information system. Demographic information, grades, test results, etc. were all taken from this system. This was the only connection with the system. Other sources of student information such as the learning management system were not connected.

Consent: Enable data subjects to agree to data collection.

Consent was often a major component when people thought about privacy. In work looking at ethical concerns related to learning analytics, consent occurred frequently (Slade and Prinsloo, 2013; Slade and Tait, 2019; and Steiner et al., 2016). Knowing its importance, the company required the institution to “hereby represent and warrant that for all student records that you disclose (“Student Records”), you have obtained the appropriate consent” (Artifact 29). The company also warned the institutions that

> Many privacy laws require a service provider to get an individual’s consent before using technologies that track that individual’s behavior, which is why we encourage you to make sure that students have provided consent for cookies and other tracking before you send mail. (Artifact 31)

The IHE discussed consent within their *Policy of Student Records*, noting that “no individual or organization outside the institution shall have access to nor will the institution disclose information from the students’ educational records without the written consent of students” (Artifact 52). The policy then included a list of exceptions which included “employees of the institution with an expressed educational interest” and “organizations conducting studies for, or on behalf of, educational agencies or institutions for the purpose of developing, validating, or administering predictive tests for non-solicitous purposes” (Artifact 52). These exceptions allowed student records within the system to be shared without requiring consent.

Transparency: Make all data collection, use, storage, and deletion as transparent as possible with clear and understandable language used to explain all privacy-related policies.

The IHE did not address this principle as there were no specific policies related to the aforementioned.

Context: Apply the context of the jurisdiction one operates in into privacy policies.

The gathering and use of the student data was clearly placed within the context of an educational situation. The goals of the data usage as well as ensuring compliance with FERPA highlighted the context.

Accountability: Privacy policies must be developed that clearly describe the practices and procedures related to the management of personal data.

The policies in place at the IHE were general rather than addressing specific issues of privacy within the system.

Identifiability: Data subjects have the option of remaining anonymous or using a pseudonym.

Students did not have the option to remain anonymous nor to use a pseudonym. In order for the system to have many of its intended benefits, individuals had to be identifiable.

Sensitivity: Treat all data collected, used, stored, and destroyed in manners appropriate to the sensitivity level of the data.

The role-based access to the system allowed different types of data to be available as appropriate. For example, information related to financial aid or counseling was restricted to those specific offices.

Information flow: Enable the communication of personal information across multiple contexts including international, governmental, economic, and social.

The IHE did not address this principle.

Identifiers: Strong identifiers are only used when necessary.

When adding student information into the system, the IHE did not include social security numbers. It was limited to names and student ID numbers.

Disclosure: Make known any data transference to new parties.

The data within the system could be shared with the company with the IHE's permission. This data was then sometimes compared to national datasets such as IPEDS to look at trends. The company does not transfer this student data to third parties.

Breach: Subjects must be informed immediately of any data breach involving their personal data.

The IHE did not address this principle.

Confidentiality: Maintain confidentiality of data throughout processes and beyond.

The IHE highlighted the importance of confidentiality. It noted, "Keep student data confidential. You are entrusted with this information in your capacity as an […] employee, so this data should only be used in support of student success" (Artifact 55).

Overall, the IHE was mixed on how it addressed the above-mentioned privacy principles. The principles that were not addressed or were only partially considered would benefit from the development of policies and procedures related to the system. Simply developing a policy related to the student success information system would result in the IHE meeting certain principles such as accountability and transparency while other principles would then be met by the content within the policy such as stating what happens in case of a breach.

It is not unexpected that the IHE did not fully meet all the privacy principles. While higher education has long established policies and procedures related to student information, learning analytics opened a range of new privacy issues that have not been considered or discussed at scale. Learning analytics itself is more than just privacy. In Ferguson et al. (2016)'s list of ethical challenges related to learning analytics, 14 out of the 21 related to issues found in privacy principles. These were listed as challenges because there was not an easy answer to address the issues. The privacy principles do provide, however, a listing of considerations that the IHE could review in their use of the system.

Discussion

Policies Lead to Trust

The IHE within this case study was at an important juncture. They were utilizing the student success information system for a little over two years. During that time, they started getting students and faculty comfortable with using the system as a communication tool. As they start to prepare to take advantage of the learning analytics functionality, they have the opportunity to take a strategic approach by developing policies and codes of practice that address all aspects of the system, including privacy. Developing clear and straightforward policies with the input of stakeholders including students provides a framework for successful implementation due to the trust and buy-in that would be established (Long & Siemens, 2011).

While the current policies provided a baseline and the ethical use document created by the IHE provided some specific contexts, it would be best for the IHE to develop and adopt specific policies related to the use of student data within the student success information system. These policies, while addressing student privacy, are also needed to define other considerations within the system. Sclater (2016) included in his guide on developing a code of practice the following items: (a) questions related to ethical, legal, and logistical issues; (b) a list of stakeholders with responsibilities; (c) and a proposed action plan with steps to take in developing a code of practice. In developing these policies, the IHE would be able to have conversations related to the larger impact of the system.

In developing their code of practice for learning analytics for Jisc, a digital, data, and technology agency that focuses on education, research, and innovation, Sclater & Bailey (2018) highlighted the need for institutions to have complete transparency in their use of learning analytics including purpose, data collected, processes, and how the data would be used. Pavlou (2011) listed a number of research studies that looked at the relationship between trust and privacy. Overall, individuals showed less concern with privacy if they established trust with the institution. Paylou noted that studies showed “trust is usually viewed as a stronger predictor of behavior that often mediates the relationship between information privacy concerns and willingness to transact” (2011, p. 983).

Austin (2019) highlighted how it was not simply enough to provide individuals with choice in relation to their privacy. Her essay noted that while personal control had long been held as a critical component of maintaining privacy, it was not enough if the choices had to occur within situations that did not provide real options. She proposed providing an environment where meaningful choices could be made, allowing for various states of privacy. Her critique of individual control through informed consent was also highlighted by Jones (2019), who noted that individuals did not truly understand what they were consenting to or how their information would be utilized. Choice and options were not addressed meaningfully by the IHE. While students could opt-out of receiving text messages and could limit how much they used the system personally, they were not able to remove themselves from the system in whole. Students’ data was included in the system and to engage with some departments on campus, they had to utilize the system.

By focusing exclusively on users' choices, it puts the onus on individuals to protect their privacy rather than developing structures into the systems themselves that support privacy. For student data systems to establish trust, they must build privacy into their systems. This "requires a shift from focusing on particular informational interaction between individuals and others and taking a more systemic view of the informational environment to ask whether it generally supports privacy" (Austin, 2019, p. 56).

Hoel and Chen (2016) noted that trust was one of the main barriers related to the adoption of learning analytics. For students to feel comfortable sharing their information they must trust that the institution would use that information appropriately. Several researchers noted how trust was critical for the ongoing use of learning analytics (Cormack, 2016; Green & Baumal, 2019; Pardo & Siemens, 2014; Prinsloo & Slade, 2015; Steiner et al., 2016). Rubel and Jones (2016) discussed how transparency about learning analytics provided personal autonomy and trust in the system. They suggested syllabi should include statements noting the use of learning analytics and the end result of that use.

Currently, documentation produced by the IHE for students and faculty and staff focused on how to use the system rather than the result of using the system. It is important to understand what students think of learning analytics as it is their information that is being used. For institution to "push forward with learning analytics *without* considering student privacy preferences – or ignoring such preferences all together – is foolhardy and morally suspect" (Jones, 2019, p. 12).

Not only is it critical to include students in the use of learning analytics, but Ifenthaler and Schumacher (2016) also noted the need to include other stakeholders such as instructors and instructional designers to ensure that the data collected actually supported student learning. The ability to collect information did not mean that it is necessary or proper to do so. Notable, the IHE did consider what information to collect when setting up the system. Specifically, the process that was used brought in different individuals which allowed for various viewpoints. As an example, by bringing in different viewpoint, the students' birthday was not included in the system. While it is possible to include it as a data point, someone spoke up about how it should not be included.

Implications for IHE

Currently the IHE has opportunities to improve the level of privacy considerations they make for student data. A number of privacy principles could be addressed more fully. Those changes could be made at an administrative level such as developing a process for data breaches. However, the IHE must also work with the faculty and staff users of the system to promote appropriate approaches to protecting student privacy. Official policies may have little impact if the day-to-day procedures do not follow best practices.

In looking at student privacy within the system, there needs to be increased training and understanding related to all aspects of privacy. Compliance with FERPA does not mean that all privacy concerns are addressed. To be truly proactive, the IHE needs to take a more comprehensive look at how students' privacy needs can be met throughout the life cycle of gathering, using, and analyzing student information.

The interviews within the case study herein highlighted little concern among users on the privacy issues related to the access and use of student information. This lack of concern may result in cavalier use of the information and system. It is important for the IHE to provide guidance for faculty on how to appropriately use the student information and include possible consequences for misuse. The consequences vary from simple loss of confidentiality to financial and psychological harm which may result in negative publicity or legal action. All consequences, however, will mean a loss of trust which as noted earlier, is critical to ensuring a successful implementation of student success information systems.

Recommendations, Limitations, and Future Work

The use of student data for learning analytics will continue to expand; as Judy noted, "I think that it's getting more and more robust and there's more and more things that we can do with it" (Artifact 48). Higher education institutions would be best served by ensuring that the system meet not only the academic needs of the students and the institution, but also the personal privacy needs of the students as well.

One of the main recommendations for higher education institutions, such as the IHE, as they begin a more focused push with learning analytics is to spend time developing policies and procedures. This focused time will allow for buy-in and uncover possible issues that can be addressed early in the process. When such processes are developed, special attention should be paid to how student privacy can be impacted by learning analytics. The IHE herein appeared to be unaware of some of the privacy concerns addressed in the learning analytics literature such as black box algorithms, where the criteria used to make decisions are unknown by the institution (Oakleaf, 2016; SoLAR, 2021) and biased analytics where the data points selected for analysis might cause inherent bias in the results (Romei & Ruggieri, 2014). Thus, before the policies and procedures are developed, it may be necessary for the system director and administrators of higher education institutions to review prior research on ethical concerns with learning analytics as well as review some best practice examples from other institutions that are further along in their implementation of learning analytics.

Another recommendation for higher education institutions would be to establish and communicate clear goals related to using the system. While the interviews uncovered similar goals shared by the administrators and faculty users of the system, there were differences in the importance the different individuals placed on the goals. This variety would result in different foci when using the system, which then leads to variety in the importance placed on student data and privacy. This means higher education institutions must create clear communication plans surrounding their learning analytic systems in order to meet the goals ascribed to those systems.

While this case study herein was undertaken to fulfill a gap in the literature as noted by Lang and Knight (2019) and Parkman and McGrath (2021) calling for specific case studies looking at learning analytic systems, additional studies could provide more insight into how institutions are implementing learning analytics. The case study herein provided a look at an IHE that had a functional communications component of a student success information system but had not fully implemented the learning analytics component. To address this limitation, additional case studies could consider institutions within different stages of implementing learning analytics. Such studies could provide insight into practices that have gone well or not.

The case study herein also provided an example of how faculty understand privacy related to learning analytics. This area could see additional research with surveys of faculty members to understand what faculty know and think about learning analytics. Lastly, this case study looked at interviews and documents on what users thought about the system. Future work can consider how users implement and engage with the system by observing faculty and students as they use the system to uncover possible privacy issues.

Chapter 5

This chapter is written as an article that will be submitted to the *Journal of Leaning Analytics*. The journal is an open-access, peer-reviewed publication put out by the Society for Learning Analytics Research (SoLAR). The length restriction for research papers submitted to the journal is 9,000 words inclusive of the abstract, key words, tables/figures, acknowledgements, and reference list.

Abstract

This case study will seek to answer the question, how is the concept of privacy addressed in relation to a student success information system within an institution of public higher education. Three themes were found within the inductive coding process which used interviews, documentation, and videos as data resources. Overall, the case study shows an institution in the early stages of implementing learning analytics and provides suggestions for how it can be more proactive in implementing privacy considerations within the development of policies and procedures.

Introduction and Literature Review

Institutions of higher education are being called to demonstrate their effectiveness amid the additional requirements of efficiency and maintaining costs. Meanwhile, technological advances have allowed for the gathering and analysis of data to aid decision making. The conjuncture of these two circumstances have made learning analytics a critical component for many institutions. Learning analytics is the use of the big data techniques that are utilized within the business sector but with the goal of improved educational experiences.

While the techniques are similar, there is an important difference between commercial big data analysis and learning analytics. As noted by Rubel and Jones (2016), for learning analytics to have the biggest impact, the student data must be connected to the individual. In big data analytics, the information can be used in the aggregate. This differentiation makes learning analytics a more personalized process which raises additional concerns related to the ethical use of such data.

Proponents of learning analytics highlight the ability to use data to increase the learning experience of students resulting in enhanced education. Yet, there remains concerns as to how these processes may provoke unintended consequences. While there are varied ethical considerations in collecting, analyzing, and using data, one of the most pressing concerns is student privacy. Hoel and Chen (2016) provide the logic for the importance of studying how privacy is addressed in learning analytics. They note that while institutions have long analyzed behavior and performance to make changes, learning analytics has changed how that process is done and the impact that it can have upon individuals. This new process then necessitates a new agreement between student and institution in relation to practice and how goals are met.

Institutions do recognize that privacy is a concern. The *EDUCAUSE 2020 top 10 IT issues* report placed privacy second on the list after security (Grajek, 2020). Burns (2020) notes that while institutions recognize the importance of privacy they must work to develop and improve policies and procedures dealing with student information. With no dedicated federal law or guidance on how to address privacy concerns within learning analytics, it has been left to each institution to develop their own approach. This study shall use a case study to see how a small public institution of higher education in the Midwest addresses the privacy of students within a student success information system. As a developing field of study, learning analytics research will mature only through studies looking at all aspects of the field. This case study will provide an in-depth analysis of one institution that has implemented a specific system.

The following research question will guide this study:

- How is the concept of privacy addressed in relation to a student success information system within an institution of public higher education?

Significance of the Study

Within their systematic review of articles looking at the ethical concerns related to learning analytics, Parkman and McGrath (2021) note that a majority of the research conducted focused on respondents' perceptions and attitudes related to learning analytics rather than the actual use of the systems. They recommended that research looking at "how ethical principles, guidelines, or codes of practices in LA [learning analytics] are put into practice will help us gain a more grounded understanding of how these instruments work in everyday higher education" (p. 13). Lang and Knight (2019) also stress the need for specific case studies looking at ethics and learning analytics. This work will help address that need by providing a case study looking at how privacy is addressed with the student success information system.

Learning Analytics

Overview

Student data has long been used to make decisions both on the microlevel in specific classrooms and at the macrolevel in how the institution operates. In the past, this data has come from faculty use of student grades and in-room discussion while institutions looked at yearly retention and graduation rates. While the use of data is not new, the current interest in learning analytics is due to the conjuncture of several trends: the volume of data that is collected, the ability to store that data, the computational capacity now available to institutions, the increase in visualization tools, and the increased demand to analyze and use big data (Slade and Prinsloo, 2013; Siemens, 2013). One of the most frequently used definitions of learning analytics comes from Siemens (2013), "learning analytics is the measurement, collection, analysis, and reporting of data about learners and their contexts, for the purposes of understanding and optimizing learning and the environment in which it occurs" (p. 1382).

With the increased push to utilize learning analytics, institutions must implement structured plans to ensure successful programs. The Data Quality Campaign (2019) provide four policy priorities that they recommend institutions consider as they implement learning analytics:

measure what matters, be transparent and earn trust, make data use possible, and guarantee access and protect privacy.

Benefits

There have been a range of claims related to how the learning analytics available through student information systems, student success information systems, learning management systems, and other systems will improve education. These include enhanced learning experiences (Long and Siemens, 2011), supported self-regulated learning (Kim et al., 2018), improvement of student learning services (Knight et al., 2014), development of prediction analytics for at-risk students (Saqr et al., 2017), and help for at-risk students (Pardo & Siemens, 2014). Long & Siemens (2011) also discuss the benefits of improved institutional decision making, advancements in learning outcomes for at-risk students, greater trust in institutions due to the disclosure of data, significant evolutions in pedagogy, sense making of complex topics, increase organizational productivity, and provide learners insights into their learning. Foster and Francis (2020) conducted a systematic literature review looking at 34 learning analytics studies that focused on the stated goals of retention, academic performance, and engagement. They found that a majority of the studies reported an increase in student outcomes related to those goals.

Challenges/Concerns

While the goal of learning analytics is to improve the learning experience for students, this does not negate the fact that the gathering of data on students poses risks and challenges. Selwyn (2019) poses several possible consequences related to learning analytics including: a reduced understanding of education, ignoring the broader social contexts of education, reducing students' and teachers' capacity for informed decision-making, a means of surveillance rather than support, a source of performativity, disadvantaging large numbers of people, and serving institutional rather than individual interests. Privacy concerns related to data analytics occur throughout the data lifecycle.

One of the major concerns related to learning analytics are ethical considerations that arise when collecting, using, and storing student data. Some authors have looked specifically at privacy concerns in relation to ethics. Slade and Prinsloo (2013) note how the ethical considerations stemming from the increased use of learning analytics come from issues related to privacy and determining who owns the data that is collected. Gasevic, Dawson, and Jovanovic (2016) note that while privacy and ethics have been a concern related to learning analytics since their development they have not been explored fully in the literature.

Pargman and McGrath (2021) conducted a systematic literature review looking at which ethical topics have been addressed in studies that look at the ethics of learning analytics. The ethical topics addressed include: transparency, privacy, informed consent, responsibility, minimizing adverse impacts, validity, and enabling interventions. There are also authors who have looked at general ethical concerns related to learning analytics which either mention privacy as a broad category or refer to aspects related to privacy. The following table provides a comparison of those studies. Direct mentions of privacy are bolded while topics related to privacy are in italics.

Table 2
Comparison of Ethical Concerns

Ferguson et al. (2016)	Slade and Prinsloo (2013)	Khalil and Ebner (2016)	Steiner et al. (2016)	Slade and Tait (2019)	SoLAR (2021)
• student success • trustworthy educational institutions • *respect for private and group assets* • respect for property rights • educators and educational institutions that safeguard those in their care • equal access to education • laws that are fair, equally applied, and observed • freedom from threat • *integrity of self*	• the location and interpretation of data • *informed consent* • **privacy and the de-identification of data** • the management, classification, and storage of data.	• *transparency of data collection, usage, and involvement of third parties* • *anonymization and de-identification of individuals* • *ownership of data* • data accessibility and accuracy of the analyzed results • security of the examined datasets and student records from any threat	• **privacy** • *informed consent, transparency, and de-identification of data* • location and interpretation of data • management, classification, and storage of data • *data ownership* • possibility of error • role of knowing and the obligation to act	• *data ownership and control* • transparency • accessibility of data • validity and reliability of data • institutional responsibility and obligation to act • communications • cultural values • inclusion • *consent* • *student agency and responsibility*	• **privacy** • opaque *black box* algorithms • basing classifications on biased datasets • incorrectly predicting someone's behavior

Defining Data Privacy

Privacy remains a concept that at one level is understood, yet at another level cannot be completely comprehended as it is approached differently by each individual within a range of contexts. The most frequently agreed upon definition of privacy is the fact that there is no universal definition of privacy (Allen, 1988; Castelli, 2014; Margulis, 2011; and Weimann & Nagel, 2012). Pavlou (2011) contents that this ambiguity comes from the fact that privacy is a complex concept that can be addressed from a range of different disciplines and perspectives. Within their review of research looking at information privacy, Smith et al., (2011) conclude that there is no definition of privacy that crosses all disciplines. Solove (2008) notes that “privacy is not reducible to a singular essence; it is a plurality of different things that do not share one element in common but that nevertheless bear a resemblance to each other” (p. 756).

Instead, definitions of privacy will vary depending on context such as time period, cultural norms, physical location, and field of study. In considering this trait of privacy, Solove looks to “conceptualize privacy from the bottom up rather than the top down, from particular contexts rather than the abstract” (2002, p. 1092). Nissenbaum (2011) holds that privacy should be put into larger social contexts. While it is acknowledged that privacy does not have a universal definition, for this study, the operational definition of privacy is *restriction of access to an individual’s personal information.*

Current Models to Integrate Privacy into Learning Analytics

There have been various guidelines and structures proposed to guide the design of student information systems and student success information systems to account for privacy needs (Bellotti, 1997; Hoel & Chen, 2016; Horvitz, 1999; Jensen et al., 2005; Kitto et al., 2015; Langheinrich, 2001; Steiner et al., 2016). Many of these new processes try to operate under a privacy by design framework as put forth by Cavoukian (2012). However, in designing systems, privacy is often simply one component and not a critical one. These proposals more frequently focus on minimizing privacy violations rather than proactively protecting individual privacy.

It is important for institutions to establish policies and practices beyond the technical integration of privacy into learning analytic systems. Prinsloo and Slade (2013) reviewed the polices of two institutions in relation to ethical concerns and learning analytics. They found that institutions were not keeping up with the new abilities of learning analytics. There are several guides that have been developed to assist institutions as they develop learning analytics programs (Cormack, 2016; Drachsler & Greller, 2016; Privacy Technical Assistance Center, 2016; Sclater & Bailey, 2018).

Institutions can also develop and adopt codes of practice to guide the implementation of learning analytics on their campuses. Welse and McKinney (2015) discuss the need for codes of practice noting that their development not only maximize effectiveness and minimize risk, but they also build trust between the institution and its constituents through transparency. For those institutions looking to establish a code of practice related to their learning analytics system, Sclater (2016) provides an in-depth discussion and guide. His article discusses the process used by Jisc to develop their code of practice and then details what should be considered by other institutions.

Methods

This study was a single case, descriptive case study. Yin (2009) offers three situations in which a case study is the preferred research method. These include asking *how* or *why* questions, the researcher cannot control events, and the topic deals with a contemporary phenomenon in specific contexts. All three of these situations were present in the current study. The rationale for deciding to perform a single case study focused on the idea that the specific institution selected for the case would be a representative case. The institution selected as the case study is a typical situation, and as noted by Yin (2009), such cases can be informative about the experiences of an average institution.

Site Selection

The institution of higher education (IHE) chosen was used to represent the abstraction of student privacy within a student success information system. The IHE is a public university within a statewide system in the Midwest. The IHE has an FTE enrollment of ~2,000 students. It offers associate through doctoral degrees with a special focus on technology. The unit of analysis for the study was the institution's student success information system.

The IHE has been working with the company since 2018 with the official launch of the system in March 2020. The system provides a systematic method to collect data on students as well as serves as a means of communication between students, faculty, and other departments on campus. The company has over 500 educational institutions using the system which highlights how this study may be applicable to other institutions and situations.

Data Collection

As a case study, data was gathered from multiple sources in order to allow for triangulation of results. Yin (2009) discusses how triangulation of data occurs when the facts of a case study are supported by more than one piece of evidence. This is one of the strengths of case studies in that they allow for internal confirmation of the research findings allowing me to make a stronger case for their conclusions. This triangulation helps with the research validity and credibility of the study. This confirmation of findings from different sources also provides confirmability (Shenton, 63).

Interview Data Collection

Semi-structured interviews were conducted with the system director at the IHE. Two interviews were held with this individual with coding occurring after each interview in order to ensure a comprehensive understanding of the system was developed. From the system director, snowball sampling was employed in order to get the contact information of faculty members and administrators who were users of the system. A semi-structured interview was also conducted with a representative from the company. After the interviews, the transcripts were shared with the participants to order to allow for member checking which impacts validity and credibility. Overall, 9 total interviews were held with 8 individuals.

Resource Data Collection

In addition to an interview with a representative from the company, data was also gathered through the analysis of company supplied material. This included training materials and their online help center. Institutional documents were also analyzed. These included: any written reports and updates on the system, emails sent about the system, policies and procedures related to the system, documentation created by the institution, and training sessions offered by the institution for the faculty and staff members who utilize the system. Overall, 53 additional data resources were analyzed.

Ethical Considerations

Approval for this case study was obtained from the University of South Dakota's IRB panel. Confidentiality of the interviewees was maintained. The audio of the interviews was recorded. Transcripts were generated from the recordings with the inclusion of pseudonyms after which time the recordings were deleted. The transcripts will be kept for three years on a protected computer in compliance with the IRB requirements.

Data Analysis

The data was analyzed using open coding within an inductive approach. In this process, themes were developed through an open reading of the sources. The data was reviewed several times to allow for refinement of the final themes shared in the reporting of results. The themes developed by the open coding were analyzed by considering how they relate to the research questions noted earlier. The inductive coding utilized a categorical aggregation approach where a collection of instances were analyzed from the data in order to develop issue-relevant meanings. This approach worked well as the interviews each provide a unique instance of working with the system and through this analysis similar themes were uncovered. These themes were then confirmed through analysis of the data resources.

Researcher Background

Within qualitative research, the researcher takes on active part within the study. Because my engagement with the participants and resources directly impacted the results, it is important to provide clarity on my background and relationship with the topic. I selected this topic due to my interest and past research on the topic of privacy. Since January 2019, I have been a part of a privacy research lab on my campus. This work has reaffirmed my believe in the importance of protecting and maintaining individual privacy. Recognizing my belief in the importance of privacy, I was conscientious in my interactions with the interviewees within the study to remain impartial.

Verification of Study

When considering issues related to the impact of a study, it is important to address issues related to reliability and validity. There are several aspects of research design which can address these issues. Yin (2009) looks at how case studies can address construct validity, internal validity, external validity, and reliability. In relation to construct validity, this study used multiple sources of evidence, established a chain of evidence, and had key informants review drafts of the report. To address internal validity, this study matched patterns during data analysis and used logic models. External validity was addressed by placing the study within the literature and theory surrounding the topic of student privacy in learning analytics. Finally, reliability was addressed by providing a detailed account of the research protocol used within the case study.

Results and Discussion

Background of Student Success Information System

The IHE contracted with the company to implement the student success information system. In one of their early training videos, the IHE provided an overview of the system to faculty noting it

> Helps advisors and support teams quickly and easily reach out to students in need of extra guidance, connects everyone on campus, from deans and faculty to financial aid, tutoring, and residential life in a collaborative network to support students. Empowers students themselves with the tools they need to stay on track and plan their entire college journey. And gives leadership the insights they need to make informed, strategic decisions and build a culture of student success. (Artifact 15)

The system was able to meet these claims due to the use of student data. The IHE highlighted that "student data is one of the most important tools we have to foster student success" (Artifact 55).

With the purpose of the system in mind, the IHE contracted with the company and set up the behind-the-scenes structure of the system. The system was ready to be rolled out to the campus in the spring of 2020. This was when the IHE moved to remote learning with the expansion of the COVID-19 safety protocols. This was both a benefit and detriment to the new system. The timing was a positive in the fact that the new system allowed for additional communication capabilities between faculty and students. Faculty could now text students to check-in on their health and academic concerns. The push to all online courses also highlighted those additional communication features as face-to-face options were no longer available. These features of the system resulted in an early adoption of the system by some faculty members.

However, rolling out a new system during such a time of upheaval also caused issues. Faculty were not given an introduction and full training to the system before they went remote. This meant there were issues with faculty not understanding the need, functions, and use of the system. The Fall 2020 trainings that were offered at the beginning of the next school year saw faculty asking what exactly this system was.

Since the initial roll out, the IHE offered various training sessions on the use of specific features of the system. These were offered virtually, in a hybrid format, and via one-on-one training. Video recordings of the trainings were available for faculty to view, and email notices went out to faculty when they needed to engage with the system; for example, at the beginning of each semester, the faculty were requested to complete a progress report on students during the first weeks of class in order to note whether the students were attending or engaging with course content.

The IHE used the system to accomplish several institutional needs. First, it is used as a progress check early in the semester to determine if a student attended or participated in a course. This helps in correcting and ensuring appropriate registration records. The institution also used the alerts and communications sent within the system to make decisions. An email sent by the provost's office noted, "The primary source I have to make decisions about students' continued enrollment and respond to complaints from students and parents is your alerts and comments" (Artifact 59). Finally, the IHE used the system to set up appointments with various offices such as advising, student housing, and financial aid. While allowing for a more focused, unified, and systematic approach, this also meant the students have little choice on the use of the system.

Inductive Coding to Answer Research Questions

After conducting interviews and gathering resources from the IHE and the company, an inductive coding process was utilized to uncover themes. Overall, three main themes were discovered in relation to the main research question: *How is the concept of privacy addressed in relation to a student success information system within an institution of public higher education?* The first theme was that privacy could be contained within the institution's adherence to the Family Educational Rights and Privacy Act (FERPA). The second theme highlighted specific methods used to maintain privacy including limiting access to information based on individual roles and ensuring technological security protocols. The third, final theme highlighted concerns raised about the relationship between students and their data.

FERPA Means Privacy

FERPA and other mentions of legal limitations were addressed across all the data sources including interviews, company documents, and institutional documents. The company often provided a default mention of FERPA to provide a warning to institutions as they worked with different data sources and features such as adding student demographic information to the system or adding notes to a student's record. "Do not do this unless you are aware of your institution's IT policies on data imports, privacy, FERPA, and other relevant policies" (Artifact 22). As mentioned in the company's training guides and help center, institutions needed to be aware of FERPA as "any information you enter into [the system] pertaining to a student becomes part of their official student record. It may be subpoenaed by the student as outlined in the Family Education Rights and Privacy Act (FERPA)" (Artifact 27). The company avoided providing specific guidance on how to comply with FERPA regulations.

This generic mention of FERPA was also seen in the training materials put out by the IHE. When talking about using the system, it was noted, "We do have FERPA as a law and something that we follow" (Artifact 56). While it was not described in any detail, FERPA was invoked by describing how the system was initially set up by one of the administrators who took part in that process, "[The institution] adheres to the system level FERPA policy" (Artifact 49). The IHE made a conscious effort to address FERPA. It was also noted how the IHE went beyond some definitions of directory information from FERPA and included student emails as personally identifiable information. Paul noted, "Ours is a little bit more restrictive than the [governing] board level" (Artifact 49).

Even though FERPA was frequently mentioned, most documents and interviewees provided a cursory understanding of what FERPA entailed. The most detailed note on FERPA was from an email sent to faculty members noting "FERPA expressly allows for sharing of students' educational records with staff who have a legitimate educational interest in providing a service that benefits students" (Artifact 6). However, in the other trainings and documents reviewed, FERPA was acknowledged as important in relation to student privacy, but specific aspects of the law were not discussed. There was a general consensus in the IHE interviews that FERPA was being followed. "Part of the tight lockdown on access to the information was […] concern about FERPA compliance" (Artifact 57). This belief in the adherence to FERPA led to a belief that student privacy was being protected. After describing FERPA, John noted, "I think those are the basic things that keep student information secure" (Artifact 50). Rose, after being asked about privacy, stated, “We just got an email today saying that the FERPA instruction were online” (Artifact 60).

The literature did not stress compliance with FERPA since the activities involved with learning analytics were permissible under the law. While FERPA required students’ consent to share their academic records with a third party, it did not impact any sharing within the institution. This exception was highlighted in the resources provided by the IHE. Beyond internal use, Parks (2017) noted that institutions were “Free to share any information in a student’s academic record with any third party that they designate a ‘school official’” (p. 26). This then addressed any concerns about the external company collecting student data.

Given the fact that FERPA was not actually impacting the use of student data in this situation, it was concerning that it appeared so frequently within the discussion of student privacy at the IHE. Some interviewees focused almost solely on FERPA as the answer to privacy. While there are numerous other issues related to student privacy, for many of the interviewees, FERPA appeared to be the extent of the individuals’ knowledge on privacy. There appeared to be the sincere belief that student privacy was addressed. Interviewees at the IHE did not show any awareness of the additional privacy issues related to learning analytics. While the IHE was acting in good faith by addressing current FERPA requirements, the literature called for movement beyond the law. Parks (2017) concluded “FERPA is unable to address many legal and ethical concerns around current uses of student data” (p. 24). Due to this, some authors called for institutions to move beyond FERPA (Jones 2019; Prinsloo & Slade, 2015; and Tene & Polonetsky, 2013).

Methods to Maintain Privacy

While FERPA was seen as a general aspect of privacy, when asked about more specific measures taken to protect privacy, several of the interviewees were unable to come up with additional items. Bob noted, “I don’t know of any. I really don’t” (Artifact 47). John had a similar response, “I don’t know” (Artifact 50). Bill also shared this level of understanding, “Nope, I don’t. I would have no idea” (Artifact 58).

The users’ lack of understanding related to student privacy is an important consideration moving forward for the IHE. While some of the literature on learning analytics considered the place of faculty members within the system, to the best of my knowledge were not any studies that looked at faculty and privacy specifically. These answers provided examples of some users’ experience and knowledge.

For those interviewees who did have ideas on measures that were taken to promote and protect student privacy, there were two main methods that were discussed. The first dealt with limiting the type of student data accessible to individuals based on their roles. Rose stated,

> I always assumed anything that I was limited to was based upon a law. I assume, that whoever set this up does so with as much positive intent as possible. That faculty and administrative staff know the law, and that we’ve done a good job of informing people of what they can and cannot show. So for that reason, I guess I just assume that whatever I don’t have access to is a legal restriction, not somebody just restricting it because it’s not necessary.” (Artifact 60)

This role-based access to the system was one of the foundational considerations when the system was set up. “Permission access points were migrated from the shared student information system in terms of roles” (Artifact 49). Jim described the setup of the roles noting they were

> Very thoughtful about what sort of information should be collected and who has access. […] That was very deliberate. And there was actually a lot of conversation about that, and that's also why in the beginning it was restricted so much, was in protection of students. (Artifact 57)

Within the ethical guidelines put out by the IHE was a section telling users to "access only student data that is relevant to your role. Some […] may allow access […] outside your role […]. By using data related to your role, you can make the greatest impact and maintain compliance with federal guidelines, such as FERPA" (Artifact 55). The company itself also noted the importance of roles, describing how the system

> Provides the granular permissions necessary to ensure that only educational representatives that have legitimate need and right to see a student's information (courses scheduled, credit accumulation, degree progression, etc.) can access that information. The system provides role-based access, allowing access to certain data to only those users with sufficient privileges. (Artifact 34)

The importance of the use of roles to limit access to types of data mirrors the privacy and learning analytics literature which talks extensively on limitation of data. Pavlou (2011) and Moor (1997) both noted, in their discussions of privacy, that it should be up to each individual to decide who has the ability to access their data. Austin (2019) built on this idea noting that it was not the quantitative amount of information available that caused privacy concerns, but rather who had access to that information and their relationship with the individual. Within the learning analytics literature, Slade and Prinsloo (2013) noted how limiting access to data to authorized individuals was one feature of a secure system. This was seen within the documents and interviews of the case study. This understanding of the importance of access was highlighted in the UK's *Code of Practice for Learning Analytics* included access as one of the features institutions should consider when establishing a learning analytics system (Sclater & Bailey, 2018). While developed in the UK, this code covered general aspects of learning analytics while allowing for individualized implementation based on location and local needs.

At the IHE, privacy was also seen within a lens of technology security. The director of system noted, "There's a lot of stops that we have in place that would try to make it so that nobody would just get in, you know it's all in a single sign on" (Artifact 48). John noted, when asked about privacy measures in the system, "It is password protected, so I think those are the basic things that keep student information secure" (Artifact 50). Users of the system were also encouraged to use standard security practices. During training, it was noted to users that when using the system in places where others could view your computer "Don't leave it open […] or walk away" (Artifact 56). The company also provided details related to the technological security aspects of the service,

> All emails stored in the […] platform are encrypted at rest which prevents unauthorized access or theft in the unlikely event that the raw data is accessed by unauthorized agents. The encryption keys are stored separately from the data and are updated on a regular basis. (Artifact 33)

The importance of the technical security issues addressed in the interviews was also seen in the learning analytics literature. Privacy and technical security were always tightly bound as security was required due to the need for privacy. In Cavoukian's (2012) privacy by design framework, she included end-to-end security as one of her seven principles. Due to the connection between privacy and security, it was often mentioned in lists of ethical concerns related to learning analytics (Khalil and Ebner, 2016; Slade and Prinsloo, 2013; Steiner et al., 2016). Pardo and Siemens (2014) also noted how security impacted learning analytics.

Students' Connection with Their Data

In general, the administrators and users of the system did not have concerns about student privacy. "It's kind of the unspoken expectation that we respect that kind of stuff" (Artifact 47). Jim noted, "We have to trust the people that you hire. That they'll use the material in the right way" (Artifact 57). These comments highlighted an idea that privacy was only a problem when misused by faculty and staff to benefit themselves. There was no nuanced belief that privacy could be violated without a specific breach or harm. For example, there was no mention of concern with the type of data gathered or how long it was maintained. Overall, the system was seen as an institutional good, and as such the procedures were also seen as good.

There were a couple of suggestions offered, however, that interviewees felt would increase the privacy of the system. Bob noted "I don't know if students know what kind of system this is and if they've signed off saying, 'I'm okay with that'" (Artifact 47). This concern with student consent moved beyond basic agreement to deep comprehension with a desire for students to "really understand it with all the other things going on in their life when they first arrive on campus" (Artifact 47).

Creating a system that allowed students to provide informed consent was critical in the literature surrounding learning analytics (Slade and Prinsloo, 2013; Slade and Tait, 2019; and Steiner et al., 2016). Jones (2019) noted how, historically, the idea that an individual had a right to control who knows specific information about themselves had long been central to privacy. This loss of control might occur either through institutions not asking for consent or by asking for consent without providing a clear description as to how the data would be used. The requirement of having a clear description was also noted in the interview as consent without understanding does not allow for meaningful consent. Slade and Prinsloo (2013) added that when asking for consent, institutions should also provide details on the possible benefits and harms resulting from sharing or not sharing the information.

Another concern brought forward by faculty was the ability to correct data if necessary. Rose noted, “It’d be cool if students could update certain aspects like demographic or stuff that wouldn’t require another person to integrate” (Artifact 60). The ability of an individual to correct information was important as it addressed two components related to privacy. First, this meant that students had access to the data connected to them. This openness and transparency on what information the institution was collecting about them allowed for greater trust and cooperation. Second, the ability to make corrections allowed students to keep a more accurate record of themselves. Slade and Tait (2019) noted that students should have access to their raw and analyzed data so they could make corrections as necessary. Ferguson et al. (2016) also included, in their list of challenges with learning analytics, that institutions should offer opportunities to correct data, but added that this process should be publicized so students would know it. This ability to make corrections is codified within FERPA with students being able to correct errors in their educational records (Daggett, 2008).

Policies Lead to Trust

The IHE within this case study was at an important juncture. They were utilizing the student success information system for a little over two years. During that time, they started getting students and faculty comfortable with using the system as a communication tool. As they start to prepare to take full advantage of the learning analytics functionality, they have the opportunity to take a strategic approach by developing policies and codes of practice that address all aspects of the system, including privacy. Developing clear and straightforward policies with the input of stakeholders including students provides a framework for successful implementation due to the trust and buy-in that would be established (Long & Siemens, 2011).

While the current policies provided a baseline and the ethical use document created by the IHE provided some specific contexts, it would be best for the IHE to develop and adopt specific policies related to the use of student data within the student success information system. These policies, while addressing student privacy, are also needed to define other considerations within the system such as: questions related to ethical, legal, and logistical issues; a list of stakeholders with responsibilities; and a proposed action plan with steps to take in developing a code of practice (Sclater, 2016). In developing these policies, the institution would be able to have conversations related to the larger impact of the system.

In developing their code of practice for learning analytics for Jisc, a digital, data, and technology agency that focuses on education, research, and innovation, Sclater & Bailey (2018) highlighted the need for institutions to have complete transparency in their use of learning analytics including purpose, data collected, processes, and how the data would be used. This transparency helps to establish trust. Pavlou (2011) listed a number of research studies that looked at the relationship between trust and privacy. Overall, individuals showed less concern with privacy if they had established trust with the institution.

Austin (2019) highlighted how it was not simply enough to provide individuals with choice in relation to their privacy. Her essay noted that while personal control had long been held as a critical component of maintaining privacy, it was not enough if the choices must occur within situations that do not provide real options. She proposed providing an environment where meaningful choices could be made, allowing for various states of privacy. Her critique of individual control through informed consent was also highlighted by Jones (2019), who noted that individuals did not truly understand what they were consenting to or how their information would be utilized. Choice and options were not addressed meaningfully by the IHE. While students could opt-out of receiving text messages and could limit how much they used the system personally, they were not able to remove themselves from the system as a whole. Their data was included in the system and in order to engage with some departments on campus, they had to utilize the system.

By focusing exclusively on users' choices, it puts the onus on individuals to protect their privacy rather than developing structures into the systems themselves that support privacy. For student data systems to establish trust, they must build privacy into their systems. This "requires a shift from focusing on particular informational interaction between individuals and others and taking a more systemic view of the informational environment to ask whether it generally supports privacy" (Austin, 2019, p. 56).

Hoel and Chen (2016) noted that trust was one of the main barriers related to the adoption of learning analytics. For students to feel comfortable sharing their information they must trust that the institution would use that information appropriately. Several researchers noted how trust was critical for the ongoing use of learning analytics (Cormack, 2016; Green & Baumal, 2019; Pardo & Siemens, 2014; Prinsloo & Slade, 2015; Steiner et al., 2016). Rubel and Jones (2016) discussed how transparency about learning analytics provided personal autonomy and trust in the system. They suggested syllabi should include statements noting the use of learning analytics and the end result of that use.

Currently, documentation produced by the IHE for students and faculty and staff focused on how to use the system rather than the result of using the system. It is important to understand what students think of learning analytics as it was their information that is being used. For institution to “push forward with learning analytics *without* considering student privacy preferences – or ignoring such preferences all together – is foolhardy and morally suspect” (Jones, 2019, p. 12).

Not only is it critical to include students in the use of learning analytics, but Ifenthaler and Schumacher (2016) also noted the need to include other stakeholders such as instructors and instructional designers in order to ensure that the data collected was actually supporting student learning. The ability to collect information did not mean that it is necessary or proper to do so. The IHE did consider what information to collect when setting up the system. The process that was used brought in different individuals which allowed for various viewpoints. As an example, by bringing in different viewpoint, the students’ birthday was not included in the system. While it is possible to include it as a data point, someone spoke up about how it should not be included.

Conclusion

Implications for IHE

Currently the IHE has opportunities to improve the level of privacy considerations they make for student data. Those changes could be made at an administrative level such as developing a process for data breaches. However, the IHE must also work with the faculty and staff users of the system to promote appropriate approaches to protecting student privacy. Official policies may have little impact if the day-to-day procedures do not follow best practice.

The interviews within the case study highlighted little concern among users on the privacy issues related to the access and use of student information. This lack of concern may result in cavalier use of the information and system. It is important for the IHE to provide guidance for faculty on how to appropriately use the student information and include possible consequences for misuse. The consequences vary from simple loss of confidentiality to financial and psychological harm which may result in negative publicity or legal action. All consequences, however, will mean a loss of trust which as noted earlier is critical in ensuring a successful implementation of student success information systems.

Recommendations, Limitations, and Future Work

The use of student data for learning analytics will continue to expand; as Judy noted, “I think that it’s getting more and more robust and there’s more and more things that we can do with it” (Artifact 48). Higher education institutions would be best served by ensuring that the system meet not only the academic needs of the students and the institution, but also the personal privacy needs of the students as well.

One of the main recommendations for higher education institutions, such as the IHE, as they begin a more focused push with learning analytics is to spend time developing policies and procedures. This focused time will allow for buy-in and uncover possible issues that can be addressed early in the process. When such processes are developed, special attention should be paid to how student privacy can be impacted by learning analytics. The IHE herein appeared to be unaware of some of the privacy concerns addressed in the learning analytics literature such as black box algorithms, where the criteria used to make decisions are unknown by the institution (Oakleaf, 2016; SoLAR, 2021) and biased analytics where the data points selected for analysis might cause inherent bias in the results (Romei & Ruggieri, 2014). Thus, before the policies and procedures are developed, it may be necessary for the system director and administrators of higher education institutions to review prior research on ethical concerns with learning analytics as well as review some best practice examples from other institutions that are further along in their implementation of learning analytics.

Another recommendation for higher education institutions would be to establish and communicate clear goals related to using the system. While the interviews uncovered similar goals shared by the administrators and faculty users of the system, there were differences in the importance the different individuals placed on the goals. This variety would result in different foci when using the system, which then leads to variety in the importance placed on student data and privacy. This means higher education institutions must create clear communication plans surrounding their learning analytic systems in order to meet the goals ascribed to those systems.

While this case study herein was undertaken to fulfill a gap in the literature as noted by Lang and Knight (2019) and Parkman and McGrath (2021) calling for specific case studies looking at learning analytic systems, additional studies could provide more insight into how institutions are implementing learning analytics. The case study herein provided a look at an IHE that had a functional communications component of a student success information system but had not fully implemented the learning analytics component. To address this limitation, additional case studies could consider institutions within different stages of implementing learning analytics. Such studies could provide insight into practices that have gone well or not.

The case study herein also provided an example of how faculty understand privacy related to learning analytics. This area could see additional research with surveys of faculty members to understand what faculty know and think about learning analytics. Lastly, this case study looked at interviews and documents on what users thought about the system. Future work can consider how users implement and engage with the system by observing faculty and students as they use the system to uncover possible privacy issues.

References

Allen, A. (1988). *Uneasy access: Privacy for women in a free society.* Rowman & Littlefield.

Austin, L. M. (2019). Re-reading Westin. *Theoretical Inquiries in Law, 20*(1), 53-81.

Bellotti, V. (1997). Design for privacy in multimedia computing and communications environments. In P. Agre & M. Rotenberg (Eds.), *Technology and privacy: The new landscape.* (pp. 63-98). MIT Press.

Burns, S. (2020, November 19). The evolving landscape of data privacy in higher education. *EDUCAUSE.* https://www.educause.edu/ecar/research-publications/the-evolving-landscape-of-data-privacy-in-higher-education/introduction

Castelli, C. (2014, October 3). Privacy engineering concepts avoid defining privacy. *Inside cybersecurity*. https://insidecybersecurity.com/share/1850

Cavoukian, A. (2012). *Privacy by design: From rhetoric to reality.* Information and privacy commissioner of Ontario.

Cormack, A. (2016). A data protection framework for learning analytics. *Journal of Learning Analytics, 3*(1), 91-106.

Data Quality Campaign. (2019). Time to act: Connecting policy to practice to make data work for students. https://dataqualitycampaign.org/resource/time-to-act-2019/

Daggett, L. M. (2008). FERPA in the twenty-first century: Failure to effectively regulate privacy for all students. *Catholic University Law Review, 58*, 59-113.

Family Educational Rights and Privacy Act of 1974, 20 U.S.C. § 1232g (1974).

Ferguson, R., Hoel, T., Scheffel, M., Drachsler, H. (2016). Guest editorial: Ethics and privacy in learning analytics. *Journal of Learning Analytics, 3*(1), 5-15.

Foster, C., & Francis, P. (2020). A systematic review on the deployment and effectiveness of data analytics in higher education to improve student outcomes. *Assessment & Evaluation in Higher Education, 45*(6), 822-841.

Gasevic, D., Dawson, S., & Jovanovic, J. (2016). Ethics and privacy as enablers of learning analytics. *Journal of Learning Analytics, 3*(1), 1-4.

Grajek, S. (2020). Top 10 IT issues 2020: The drive to digital transformation begins. *EDUCAUSE Review Special Report.* https://er.educause.edu/-/media/files/articles/2020/1/er20sr201.pdf

Green, P., & Baumal, B. (2019). Legal, ethical and privacy issues affecting data sharing among Ontario's higher education instititutions in interinstitutional collaboration. *College Quarterly, 22*(2).

Hoel, T., & Chen, W. (2016). Privacy-driven design of learning analytics applications: Exploring the design space of solutions for data sharing and interoperability. *Journal of Learning Analytics, 3*(1), 139-158.

Horvitz, E. (1999). Principles of mixed-initiative user interfaces. *Proceedings of CHI'99.* (pp. 159-166).

Ifenthaler, D., & Schumacher, C. (2016). Student perceptions of privacy principles for learning analytics. *Education Technology Research Development, 64,* 923-938.

Jensen, C., Tullio, J., Potts, C., & Mynatt, E. D. (2005). STRAP: A structured analysis framework for privacy. *GVU Technical Report.*

Jones, K.M.L. (2019). Learning analytics and higher education: A proposed model for establishing informed consent mechanisms to promote student privacy and autonomy. *International Journal of Education Technology in Higher Education, 16*(24), 1-22.

Khalil, M., & Ebner, M. (2016). De-Identification in Learning Analytics. *Journal of Learning Analytics*, *3*(1), 129–138. https://doi.org/10.18608/jla.2016.31.8

Kim, D., Yoon, M., Jo, I. H., & Branch, R. M. (2018). Learning analytics to support self-regulated learning in asynchronous online courses: A case study at a women's university in South Korea. *Computer & Education*, *127*, 233-251.

Kitto, K., Cross, S., Walters, Z., & Lupton, M. (2015). Learning analytics beyond the LMS: The connected learning analytics toolkit. *Proceedings of the 5th internatioanl conference on learning analytics and knowledge.* 11-15.

Knight, S., Buckingham Shum, S., & Littleton, K. (2014). Epistemology, assessment, pedagogy: Where learning meets analytics in the middle space. *Journal of Learning Analytics, 1*(2), 23-47.

Langheinrich, M. (2001). Privacy by design – principles of privacy-aware ubiquitous systems. *Proceedings of Ubicomp 2001.* (pp. 273-291).

Long, P., & Siemens. G. (2011). Penetrating the fog: Analytics in learning and education. *Educause Review*, *46*(5), 30-40.

Margulis, S. T. (2011). Three theories of privacy: An overview. In S. Trepte & L. Reinecke (Eds.), *Privacy online* (pp. 9-17). Springer.

Moor, J. H. (1997). Towards a theory of privacy in the information age. *Computers & Society*, *27*(3), 27-32.

Nissenbaum, H. (2011). Contextual approach to privacy online. *Dædalus*, *140*(4), 32-48.

Oakleaf, M. (2016). Getting ready and getting started: Academic librarian involvement in institutional learning analytics initiatives. *The Journal of Academic Librarianship*, *42*(4), 472–475.

Pardo, A., & Siemens, G. (2014). Ethicial and privacy principles for learning analytics. *British Journal of Educational Technology, 45*(3), 438-450.

Pargman, T. C., & McGrath, C. (2021). Mapping the ethics of learning analytics in higher educaiton: A systematic literature revew of empirical research. *Journal of Learning Analytics,* early access, 1-17. https://doi.org/10.18608/jla.2021.1

Parks, C. (2017). Beyond complicance: Students and FERPA in the age of big data. *Journal of Intellectual Freedom and Privacy, 2*(2), 23-33.

Pavlou, P. A. (2011). State of the information privacy literature: Where are we now and where should we go?. *MIS Quarterly, 35*(4), 977-988.

Prinsloo, P., & Slade, S. (2013). An evaluation of policy frameworks for addressing ethical considerations in learning anlaytics. *Proceedings of the third international conference on learning analytics and knowledge,* 240-244.

Prinsloo, P., & Slade, S. (2015). Student privacy self-management: Implications for learing analytics. *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge.* New York.

Privacy Technical Assistance Center. (2016). *Student privacy while using online educational services: Model terms of service.* https://studentprivacy.ed.gov/sites/default/files/resource_document/file/TOS_Guidance_Mar2016.pdf

Romei, A., & Ruggieri, S. (2014). A multidisciplinary survey on discrimination analysis. *The Knowledge Engineering Review, 29*(5), 582-638.

Rubel, A., & Jones, K. (2016). Student privacy in learning analytics: An information ethics perspective. *Information Society, 32*(2), 143-159.

Saqr, M., Fors, U., & Tedre, M. (2017). How learning analytics can early predict under-achieving students ina blended medical education course. *Medical Teacher, 39*(7), 757-767.

Sclater, N. (2016). Developing a code of practice for learning analytics. *Journal of Learning Analytics, 3*(1), 16-42.

Sclater, N., & Bailey, P. (2018). *Code of practice for learning analytics*. Jisc. https://repository.jisc.ac.uk/6985/1/Code_of_Practice_for_learning_analytics.pdf

Selwyn, N. (2019). What's the problem with learning analytics?. *Journal of Learning Analytics, 6*(3), 11-19.

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Educaiton for Inforamtion*, *22*, 63-75.

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist, 57*(10), 1380-1400.

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist, 57*(10), 1509-1528.

Slade, S., & Tait, A. (2019). Global guidelines: Ethics in learning analytics. *Internatioanl Council for Open and Distance Education.* https://www.aace.org/review/global-guidelines-ethics-in-learning-analytics/

Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: An interdisciplinary review. *MIS Quarterly, 35*(4), 989-1015.

Society for Learning Analytics Research (SOLAR). (2021). What is learnng analytics. https://www.solaresearch.org/about/what-is-learning-analytics/

Solove, D. (2002). Conceptualizing privacy. *California Law Review, 90*(4), 1087-1155. https://doi.org/10.2307/3481326

Solove, D. J. (2008). "I've got nothing to hide" and other misunderstandings of privacy. *San Diego Law Review, 44*, 745-772.

Steiner, C. M., Kirkmeier-Rust, M. D., & Albert, D. (2016). LEA in private: A privacy and data protection framework for a learning analytics toolbox. *Journal of Learning Analytics, 3*(1), 66-90.

Tene, O., & Polonetsky, J. (2013). Big data for all: Privacy and user control in the age of analytics. *Northwestern Journal of Technology and Intellectual Property, 11*(5), 239-274.

Weimann, T., & Nagel, D. (2012). Agreeing on a definition for data protection in a globalized world. *IEEE Technology and Society Magazine*, *31*(4), 39-42.

Welsh, S., & McKinney, S. (2015). Clearing the fog: A learning analytics code of practice. *Globally connected, digitally enabled.* Proceedings of ASCILITE, 241-245.

Yin, R. K. (2009). *Case study research: Design and methods.* Sage.

References

Allen, A. (1988). *Uneasy access: Privacy for women in a free society.* Rowman & Littlefield.

Altman, I. (1975). *The Environment and Social Behavior.* Brooks/Cole.

Asia-Pacific Economic Cooperation. (2005). *APEC Privacy framework.* https://www.apec.org/publications/2005/12/apec-privacy-framework

Austin, L. M. (2019). Re-reading Westin. *Theoretical Inquiries in Law, 20*(1), 53-81.

Bellotti, V. (1997). Design for privacy in multimedia computing and communications environments. In P. Agre & M. Rotenberg (Eds.), *Technology and privacy: The new landscape.* (pp. 63-98). MIT Press.

Brown, M., & Klein, C. (2020). Whose data? Which rights? Whose power? A policy discourse analaysis of student privacy policy documents. *The Journal of Higher Education, 91*(7), 1149-1178.

Burns, S. (2020, November 19). The evolving landscape of data privacy in higher education. *EDUCAUSE.* https://www.educause.edu/ecar/research-publications/the-evolving-landscape-of-data-privacy-in-higher-education/introduction

California Consumer Privacy Act. 2018. Cal. Legis. Serv. Ch. 55 (A.B. 375) (WEST). https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201720180AB375

Calo, R. (2011). The boundaries of privacy harm. *Indiana Law Journal, 86*(3).

Castelli, C. (2014, October 3). Privacy engineering concepts avoid defining privacy. *Inside cybersecurity*. https://insidecybersecurity.com/share/1850

Cavoukian, A. (2012). *Privacy by design: From rhetoric to reality.* Information and privacy commissioner of Ontario.

Cavoukian, A. (2011). *Privacy by design: The 7 foundational principles*. Information and privacy commissioner of Ontario, Canada. https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf

Citron, D. K. & Solove, D. J. (2021). Privacy harms. *GWU Legal Studies Research Paper No. 2021-11, GWU Law School Public Law Research Paper No. 2021-11*, 3-54. http://dx.doi.org/10.2139/ssrn.3782222

Clow, D. (2013). An overview of learning analytics. *Teaching in Higher Education, 18*(6), 683-695.

Cormack, A. (2016). A data protection framework for learning analytics. *Journal of Learning Analytics, 3*(1), 91-106.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches*. SAGE.

Crowe, S., Cresswell, K., Roberson, A., Huby, G., Avery, A., & Sheikh, A. (2011). The case study approach. *BMC Medical Research Methodology, 11*(100), 1-9.

Data Quality Campaign. (2019). Time to act: Connecting policy to practice to make data work for students. https://dataqualitycampaign.org/resource/time-to-act-2019/

Daggett, L. M. (2008). FERPA in the twenty-first century: Failure to effectively regulate privacy for all students. *Catholic University Law Review, 58*, 59-113.

EDUCAUSE. (2021). *Higher Education Community Vendor Assessment Toolkit*. https://library.educause.edu/resources/2020/4/higher-education-community-vendor-assessment-toolkit

Elvy, S. A. (2017). Paying for privacy and the personal data economy. *Columbia Law Review, 117*, 1369-1460.

Family Educational Rights and Privacy Act of 1974, 20 U.S.C. § 1232g (1974).

Ferguson, R. (2019). Ethical challenges for learning analytics. *Journal of Learning Analytics, 6*(3), 25-30.

Ferguson, R., Hoel, T., Scheffel, M., Drachsler, H. (2016). Guest editorial: Ethics and privacy in learning analytics. *Journal of Learning Analytics, 3*(1), 5-15.

Fischer, C., Pardos, Z. A., Baker, R. S., Williams, J. J., Smyth, P., Yu, R., Slater, S., Baker, R., & Warschauer, M. (2020). Mining big data in education: Affordances and challenges. *Review of Research in Education, 44*, 130-160.

Foster, C., & Francis, P. (2020). A systematic review on the deployment and effectiveness of data analytics in higher education to improve student outcomes. *Assessment & Evaluation in Higher Education, 45*(6), 822-841.

Francis, M., Covert, Q., Steinhagen, D., & Streff, K. (2020). *An Inventory of International Privacy Principles: A 14 Country Analysis.* HICSS conference 53. Hawaii.

Gasevic, D., Dawson, S., & Jovanovic, J. (2016). Ethics and privacy as enablers of learning analytics. *Journal of Learning Analytics, 3*(1), 1-4.

Grajek, S. (2020). Top 10 IT issues 2020: The drive to digital transformation begins. *EDUCAUSE Review Special Report.* https://er.educause.edu/-/media/files/articles/2020/1/er20sr201.pdf

Green, P., & Baumal, B. (2019). Legal, ethical and privacy issues affecting data sharing among Ontario’s higher education instititutions in interinstitutional collaboration. *College Quarterly, 22*(2).

Heath, J. (2014). Contemporary privacy theory contributions to learning analytics. *Journal of Learning Analytics, 1*(1), 140-149.

Hoel, T., & Chen, W. (2016). Privacy-driven design of learning analytics applications: Exploring the design space of solutions for data sharing and interoperability. *Journal of Learning Analytics, 3*(1), 139-158.

Horvitz, E. (1999). Principles of mixed-initiative user interfaces. *Proceedings of CHI'99.* (pp. 159-166).

Ifenthaler, D., & Schumacher, C. (2015). Divulging personal information within learning analytics systems. *12th International conference on cognition and exploratory learning in digital age.* 11-18.

Ifenthaler, D., & Schumacher, C. (2016). Student perceptions of privacy principles for learning analytics. *Education Technology Research Development, 64,* 923-938.

Inter-American Specialized Conference on Human Rights. (1969). *American convention on human rights*. Costa Rica. https://www.cidh.oas.org/basicos/english/basic3.american%20convention.htm

International Orgainzation for Standardization. (2011). *Information technology – Security techniques – Privacy framework, ISO/IEC 29100.*

Jensen, C., Tullio, J., Potts, C., & Mynatt, E. D. (2005). STRAP: A structured analysis framework for privacy. *GVU Technical Report.*

Johnson, L., Levine, A., Smith, R., and Stone, S. 2010. *The 2010 horizon report*. Austin, TX: The New Media Consortium.

Jones, K.M.L. (2019). Learning analytics and higher education: A proposed model for establishing informed consent mechanisms to promote student privacy and autonomy. *International Journal of Education Technology in Higher Education, 16*(24), 1-22.

Kay, D., Korn, N., & Oppenheim, C. (2012). Legal, risk and ethical aspects of analytics in higher education. *JISC CETIS Analytics Series*, *1*(6). 1-30.

Khalil, M., & Ebner, M. (2016). De-Identification in Learning Analytics. *Journal of Learning Analytics*, *3*(1), 129–138. https://doi.org/10.18608/jla.2016.31.8

Kim, D., Yoon, M., Jo, I. H., & Branch, R. M. (2018). Learning analytics to support self-regulated learning in asynchronous online courses: A case study at a women's university in South Korea. *Computer & Education*, *127*, 233-251.

Kirby, M. (2011). The history, achievement and future of the 1980 OECD guidelines on privacy. *International Data Privacy Law*, *1*(1), 6-14.

Kitto, K., Cross, S., Walters, Z., & Lupton, M. (2015). Learning analytics beyond the LMS: The connected learning analytics toolkit. *Proceedings of the 5th internatioanl conference on learning analytics and knowledge.* 11-15.

Kitto, K., & Knight, S. (2019). Practical ethics for building learning analytics. *British Jounral of Educational Technology, 50*(6), 2855-2870.

Knight, S., Buckingham Shum, S., & Littleton, K. (2014). Epistemology, assessment, pedagogy: Where learning meets analytics in the middle space. *Journal of Learning Analytics, 1*(2), 23-47.

Kwasny, M. N., Caine, K. E., Rogers, W. A., & Fisk, A. D. (2008). Privacy and technology: Folk definitions and perspectives. *Proceedings of the SIGCHI conference of human factor computer systems.* 3291-3296. https://doi:10.1145/1358628.1358846

Langheinrich, M. (2001). Privacy by design – principles of privacy-aware ubiquitous systems. *Proceedings of Ubicomp 2001.* (pp. 273-291).

Loi, M., & Christen, M. (2020). Two concepts of group privacy. *Philosophy & Technology, 33,* 207-224.

Long, P., & Siemens. G. (2011). Penetrating the fog: Analytics in learning and education. *Educause Review, 46*(5), 30-40.

Margulis, S. T. (2003). On the status and contribution of Westin's and Altman's theories of privacy. *Journal of Social Issues, 59*(2), 411-429.

Margulis, S. T. (2011). Three theories of privacy: An overview. In S. Trepte & L. Reinecke (Eds.), *Privacy online* (pp. 9-17). Springer.

Mayer-Schonberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think.* Houghton Mifflin Harcourt.

Mittelstadt, B. (2017). From individual to group privacy in big data analytics. *Philosophy of Technology, 30,* 475-494.

Moor, J. H. (1997). Towards a theory of privacy in the information age. *Computers & Society, 27*(3), 27-32.

Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review, 79,* 101-139.

Nissenbaum, H. (2011). Contextual approach to privacy online. *Dædalus, 140*(4), 32-48.

Nissenbaum, H. (2014). Respect for context as a benchmark for privacy online: What it is and isn't. In C. Dartiguepeyrou (Ed.), *The futures of privacy* (pp 19-30). Fondation Telecom, Institut Mines-Telecom.

Oakleaf, M. (2016). Getting ready and getting started: Academic librarian involvement in institutional learning analytics initiatives. *The Journal of Academic Librarianship, 42*(4), 472–475.

OASIS. *Privacy management reference model and methodology (PMRM) Version 1.0*. (2016). http://docs.oasis-open.org/pmrm/PMRM/v1.0/cs02/PMRM-v1.0-cs02.html#_Toc452024302

Organization for Economic Co-operation and Development. (1980). *OECD Guidelines on the protection of privacy and transborder flows of personal data*. https://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm

Pardo, A., & Siemens, G. (2014). Ethicial and privacy principles for learning analytics. *British Journal of Educational Technology, 45*(3), 438-450.

Pargman, T. C., & McGrath, C. (2021). Mapping the ethics of learning analytics in higher educaiton: A systematic literature revew of empirical research. *Journal of Learning Analytics*, early access, 1-17. https://doi.org/10.18608/jla.2021.1

Parks, C. (2017). Beyond complicance: Students and FERPA in the age of big data. *Journal of Intellectual Freedom and Privacy, 2*(2), 23-33.

Pavlou, P. A. (2011). State of the information privacy literature: Where are we now and where should we go?. *MIS Quarterly, 35*(4), 977-988.

Pence, H. E. (2014-2015). What is big data and why is it important?. *Journal of Educational Technology Systems, 43*(2), 159-171.

Pence, H. E. (2015). Will big data mean the end of privacy?. *Journal of Educational Technology Systems, 44*(2), 253-267.

Petersen, R. J. (2012). Policy dimentions of analytics in higher education. *EDUCAUSE Review,* 47(4), 44-49.

Prinsloo, P., & Slade, S. (2013). An evaluation of policy frameworks for addressing ethical considerations in learning anlaytics. *Proceedings of the third international conference on learning analytics and knowledge,* 240-244.

Prinsloo, P., & Slade, S. (2015). Student privacy self-management: Implications for learing analytics. *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge.* New York.

Prinsloo, P., & Slade, S. (2016). Student vulnerability, agency, and learning analytics: An exploration. *Journal of Learning Analytics, 3*(1), 159-182.

Privacy Technical Assistance Center. (2016). *Student privacy while using online educational services: Model terms of service.* https://studentprivacy.ed.gov/sites/default/files/resource_document/file/TOS_Guidance_Mar2016.pdf

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

Rizza, C., Curvelo, P., Crespo, I., Chiaramello, M., Ghezzi, A., Pereira, A. G. (2012). Welcome to airstrip one: Someone is (surely) taking care of your privacy. *International association for development of the information society international conference.*

Robertshaw, M. B., & Asher, A. (2019). Unethical numbers? A meta-analysis of library learning analytics studies. *Library Trends, 68*(1), 76-101.

Romei, A., & Ruggieri, S. (2014). A multidisciplinary survey on discrimination analysis. *The Knowledge Engineering Review, 29*(5), 582-638.

Rubel, A., & Jones, K. (2016). Student privacy in learning analytics: An information ethics perspective. *Information Society, 32*(2), 143-159.

Saqr, M., Fors, U., & Tedre, M. (2017). How learning analytics can early predict under-achieving students ina blended medical education course. *Medical Teacher, 39*(7), 757-767.

Schroeder, D., & Cohen, N. A. (2011). GAPP targets privacy risks. *Journal of Accountancy, 212*(1), 52-56.

Sclater, N. (2016). Developing a code of practice for learning analytics. *Journal of Learning Analytics, 3*(1), 16-42.

Sclater, N., & Bailey, P. (2018). *Code of practice for learning analytics*. Jisc. https://repository.jisc.ac.uk/6985/1/Code_of_Practice_for_learning_analytics.pdf

Selwyn, N. (2019). What's the problem with learning analytics?. *Journal of Learning Analytics, 6*(3), 11-19.

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Educaiton for Inforamtion, 22*, 63-75.

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist, 57*(10), 1380-1400.

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist, 57*(10), 1509-1528.

Slade, S., & Tait, A. (2019). Global guidelines: Ethics in learning analytics. *Internatioanl Council for Open and Distance Education.* https://www.aace.org/review/global-guidelines-ethics-in-learning-analytics/

Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: An interdisciplinary review. *MIS Quarterly, 35*(4), 989-1015.

Society for Learning Analytics Research (SOLAR). (2021). What is learnng analytics. https://www.solaresearch.org/about/what-is-learning-analytics/

Solove, D. (2002). Conceptualizing privacy. *California Law Review, 90*(4), 1087-1155. https://doi.org/10.2307/3481326

Solove, D. J. (2006). A taxonomy of privacy. *Univerity of Pennsylvania Law Review, 154*(3), 477-560.

Solove, D. J. (2008). "I've got nothing to hide" and other misunderstandings of privacy. *San Diego Law Review, 44*, 745-772.

Steiner, C. M., Kirkmeier-Rust, M. D., & Albert, D. (2016). LEA in private: A privacy and data protection framework for a learning analytics toolbox. *Journal of Learning Analytics, 3*(1), 66-90.

Student Privacy Compass. (2021). State Student Privacy Laws. https://studentprivacycompass.org/state-laws/

Tavani, H. T. (2007). Philosophical theories of privacy: implications for an adequate online privacy policy. *Metaphilosophy, 38*(1), 1-22.

Tavani, H. T., & Moor, J. H. (2001). Privacy protection, control of inforamtion, and privacy-enhancing technologies. *Computers and Society, 31*(1), 6-11.

Tene, O., & Polonetsky, J. (2013). Big data for all: Privacy and user control in the age of analytics. *Northwestern Journal of Technology and Intellectual Property, 11*(5), 239-274.

United Nations General Assembly. (1948). *The universal declaration of human rights*. Paris. https://www.un.org/sites/un2.un.org/files/udhr.pdf

United Nations General Assembly. (1966a). *International covenant on civil and political rights* https://treaties.un.org/doc/publication/unts/volume%20999/volume-999-i-14668-english.pdf

United States Department of Health, Education, and Welfare. (1973). *Records computers and the rights of citizens*. https://www.justice.gov/opcl/docs/rec-com-rights.pdf

Warren, S. D., & Brandeis, L. D. (1890). The right to privacy. *Harvard Law Review, 4*(5), 193-220. https://doi.org/10.2307/1321160

Weimann, T., & Nagel, D. (2012). Agreeing on a definition for data protection in a globalized world. *IEEE Technology and Society Magazine, 31*(4), 39-42.

Welsh, S., & McKinney, S. (2015). Clearing the fog: A learning analytics code of practice. *Globally connected, digitally enabled.* Proceedings of ASCILITE, 241-245.

Westin, A. F. (1967). *Privacy and freedom.* Atheneum.

Westin, A. F. (2003). Social and political dimensions of privacy. *Journal of Social Issues, 59*(2), 431-453.

Williamson, B., Bayne, S., & Shay, S. (2020). The datafication of teaching in higher education: Critical issues and perspectives. *Teaching in Higher Education*, *25*(4), 351–365.

Wright, D., & Raab, C. (2014). Privacy principles, risks and harms. *International Review of Law, Computers & Technology*, *28*(3), 277-298. https://doi.org/10.1080/13600869.2014.913874

Yin, R. K. (2009). *Case study research: Design and methods.* Sage.

Appendix A

Invite to Participate in Research

Date:

Dear :

You are invited to participate in a research study. The purpose of the study is to explore how privacy is addressed within the policies, procedures, and practices surrounding the use of the [company] system.

We are inviting you to be in this study because you are a user or connected to the operation of the [company] student success system.

If you agree to participate, you will be asked to take part in a personal interview to discuss your knowledge and use of the [company] system. Interviews will be recorded in order to create transcripts for analysis. Recording will be deleted after the creation of the transcript. Interviews will take approximately one hour with the possibility of additional requests for interview sessions as questions arise. Participants may also be asked to provide access to institutional resources related to the system for example: policies, training materials, procedures, etc. In writing up the details of the study, we will do so in such a way that you cannot be identified.

We will keep the information you provide confidential; however, federal regulatory agencies and the University of South Dakota Institutional Review Board (a committee that reviews and approves research studies) may inspect and copy records pertaining to this research.

There are no known risks from being in this study, and you will not benefit personally. However, we hope that others may benefit in the future from what we learn as a result of this study.

Your participation in this research study is completely voluntary. If you decide not to be in this study, or if you stop participating at any time, you will not be penalized or lose any benefits for which you are otherwise entitled.

If you have any questions, concerns or complaints now or later, you may contact us at the number below. If you have any questions about your rights as a human subject, complaints, concerns or wish to talk to someone who is independent of the research, contact the Office for Human Subjects Protections at 605-658-3743. Thank you for your time.

Mary Francis
Mundt Library
Dakota State University
605-256-5845

Appendix B

Informed Consent Form for Interview

UNIVERSITY OF SOUTH DAKOTA

Institutional Review Board

Informed Consent Statement

Title of Project:	Student Success, Privacy, and Institutional Goals: A Case Study Investing the Application of Privacy within a Student Success Data System at a Small, Public Institution within the Midwest
Principal Investigator:	Mejai Bola Mike Avoseh, Delzell Education Center, Vermillion, SD 57069 (605) 658-6617 mejai.avoseh@usd.edu
Other Investigators:	Mary Francis, Delzell Education Center, Vermillion, SD 57069 (605) 256-5845 mary.francis@coyotes.usd.edu

Invitation to be Part of a Research Study

You are invited to participate in a research study. In order to participate, you must be a user or connected to the operation of the [company] student success information system. Taking part in this research project is voluntary. Please take time to read this entire form and ask questions before deciding whether to take part in this research project.

What is the study about and why are we doing it?

The purpose of the study is to explore how privacy is addressed within the policies, procedures, and practices surrounding the use of the [company] system. About seven people will take part in this research.

What will happen if you take part in this study?

If you agree to take part in this study, you will be asked to take part in a personal interview to discuss your knowledge and use of the [company] system. Interviews will be recorded in order to create transcripts for analysis. Recording will be deleted after the creation of the transcript. Interviews will take approximately one hour with the possibility of additional requests for interview sessions as questions arise. Participants may also be asked to provide access to institutional resources related to the system for example: policies, training materials, procedures, etc.

What risks might result from being in this study?

There are no risks in participating in this research beyond those experienced in everyday life.

How could you benefit from this study?

You might benefit from this study through exposure to ideas developed during the interviews about the [company] system. Others might benefit due to changes in procedure surrounding the use of the [company] system resulting from the discussions held during the interviews.

How will we protect your information?

The records of this study will be kept confidential to the extent permitted by law. Any report published with the results of this study will remain confidential and will be disclosed only with your permission or as required by law. To protect your privacy, we will not include any information that could identify you. We will protect the confidentiality of the research data by asking you to use a pseudonym during the recording of the interviews, removing any identifying names in the interview transcripts, deleting the interview recordings after the creation of the transcript, and storing the transcript on a protected computer for three years after which it will be deleted.

It is possible that other people may need to see the information we collect about you. These people work for the University of South Dakota and other agencies as required by law or allowed by federal regulations.

Your Participation in this Study is Voluntary

It is totally up to you to decide to be in this research study. Participating in this study is voluntary. Even if you decide to be part of the study now, you may change your mind and stop at any time. You do not have to answer any questions you do not want to answer.

Contact Information for the Study Team and Questions about the Research

The researchers conducting this study are Mary Francis and Mejai Avoseh. You may ask any questions you have now. If you later have questions, concerns, or complaints about the research please contact Mary Francis at (605) 256-5845 or Mejai Avoseh at (605) 658-6617 during the day.

If you have questions regarding your rights as a research subject, you may contact The University of South Dakota- Office of Human Subjects Protection at (605) 658-3743. You may also call this number with problems, complaints, or concerns about the research. Please call this number if you cannot reach research staff, or you wish to talk with someone who is an informed individual who is independent of the research team.

Your Consent

Before agreeing to be part of the research, please be sure that you understand what the study is about. Keep this copy of this document for your records. If you have any questions about the study later, you can contact the study team using the information provided above.

Appendix C

Interview protocol [company] *system administrator*

Thank you for taking the time to meet with me today. My dissertation study is looking at how student privacy is addressed within the [company] system and your experience with the system will help in discovering an answer to that question. Please take some time to read the consent form.

To aid in the analysis of the data I am gathering, I will be recording our discussion in order to create a transcript of what was said. I will share the transcript after it is completed so that you can confirm that your ideas came across correctly.

Background on system:

Can you share some background on the decision to purchase and use the system?

What does the system do or allow an institution to do?

What happened at the board level to choose and implement the system?

What policies or procedures were developed to guide the use of the system?

Did COVID impact the roll out of the system?

Current use of system:

What features are used in the system? Predictive analytics? Student action prompts?

What benefits have been seen using the system? Specific examples? Numbers?

What is being done to promote use of the system? Faculty? Students? Staff?

What message is being used to promote the system?

Can you provide some examples of the procedures to do some common functions within the system?

What data is stored in the system? How long is it stored?

Does the system integrate with other systems on campus? Which ones? How?

Goals of system:

What are the overarching goals of the system? Could you rank how important each goal is overall?

Are we using the system to its fullest extent? What more could be done? Will we get to that level?

Privacy within system:

What measures are taken to promote and protect student privacy?

How or was student privacy considered when the system was set up?

Do students have any options in whether or how they use the system?

Wrap-up:

Do you have some faculty members who are heavy users of the system that you think would be willing to talk to me?

What else would you like to mention that I haven't asked about?

Thank you for taking the time to talk to me today. If you have any questions or things you would like to add, please feel free to contact me.

Appendix D

Interview protocol [company] *system representative*

Thank you for taking the time to meet with me today. My dissertation study is looking at how student privacy is addressed within the [company] system and your experience with the system will help in discovering an answer to that question. Please take some time to read the consent form.

To aid in the analysis of the data I am gathering, I will be recording our discussion in order to create a transcript of what was said. I will share the transcript after it is completed so that you can confirm that your ideas came across correctly.

Background on system:

What does the system do or allow an institution to do?

How many institutions use the system?

Who are your major competitors?

Are your services similar to others offered by other companies? What makes yours different?

Current use of system:

What features are used in the system? Predictive analytics? Student action prompts?

What benefits have been seen using the system? Specific examples? Numbers?

What message is being used to promote the system?

Goals of system:

What are the overarching goals of the system? Could you rank how important each goal is overall?

Privacy within system:

What measures are taken to promote and protect student privacy?

Are there features that are built into the system that promote privacy?

Are there features that an institution can implement to promote privacy?

Wrap-up:

What else would you like to mention that I haven't asked about?

Thank you for taking the time to talk to me today. If you have any questions or things you would like to add, please feel free to contact me.

Appendix E

Interview protocol [company] *system user*

Thank you for taking the time to meet with me today. My dissertation study is looking at how student privacy is addressed within the [company] system and your experience with the system will help in discovering an answer to that question. Please take some time to read the consent form.

To aid in the analysis of the data I am gathering, I will be recording our discussion in order to create a transcript of what was said. I will share the transcript after it is completed so that you can confirm that your ideas came across correctly.

Current use of system:

What features do you use in the system?

What benefits have you seen using the system? Specific examples? Numbers?

Can you provide some examples of the procedures to do some common functions within the system?

What additional information would you like to see in the system? Is there more data that would help you?

What do you think about the level of access to information in the system? Should users have more/less/different?

Can you talk about the ease of use of the system?

Goals of system:

What are the overarching goals of the system? Could you rank how important each goal is overall?

Are we using the system to its fullest extent? What more could be done? Will we get to that level?

Privacy within system:

What measures are taken to promote and protect student privacy?

Wrap-up:

What else would you like to mention that I haven't asked about?

Thank you for taking the time to talk to me today. If you have any questions or things you would like to add, please feel free to contact me.