Big Data in Student Data Analytics: Higher Education Policy Implications for Student Autonomy, Privacy, Equity, and Educational Value

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Marcia Jean Ham, M.Ed.

Graduate Program in Educational Studies

The Ohio State University

2021

Dissertation Committee

Bryan Warnick, Advisor

Rick Voithofer, Advisor

Jackie Blount

Abstract

Leveraging big data for student data analytics is increasingly integrated throughout university operations from admissions to advising to teaching and learning. Though the possibilities are exciting to consider, they are not without risks to student autonomy, privacy, equity, and educational value. There has been little research showing how universities address such ethical issues in their student data policies and procedures to date though privacy and security policies are abundant. Though privacy and security policies that students sign cover institutions legally, there is more that can be done to support the ethical use of student data analytics at higher education institutions.

This exploratory study addressed why it is important to support the four values of autonomy, privacy, equity, and educational value within university student data analytics policies and procedures. A rationale for focusing on these values was discussed through the lens of Paulo Freire's *Pedagogy of the Oppressed.* A comparative case analysis of data analytics policies and procedures at two large, public universities provided insight into what they emphasized and where risks to student autonomy, privacy, equity, and educational value existed. This study concluded with recommendations of how institutional leadership can use proposed principles of ethical student data analytics to evaluate their own policies and procedures and amend risks that are uncovered through analysis.

Acknowledgments

During the past six years I've been lucky to have remarkable advisors and mentors who have helped me see my work through various perspectives and coached me in how best to channel my inquiries. Dr. Bryan Warnick, it was your Educational Philosophy course that first challenged my preconceived notions of what is good education. I was excited to learn more about various philosophies of education and analyze current educational issues from different philosophical perspectives. Thank you, also, for your support as I presented at OVPES 2017, my first philosophy conference. Dr. Rick Voithofer, our conversations "noodling" on different ideas I had for research were invaluable for me as I narrowed down my vast curiosity about data and technology in education to a manageable dissertation focus. I am grateful for the experience co-authoring "Ethical Issues and Potential Unintended Consequences of Data-Based Decision Making" with you. Dr. Jackie Blount, your seminar helped me see the complexities of equity in education and inspired me down a path of research looking at the impact of data on educational equity. Thank you for all of your work creating an innovative experience for the entire class to co-author and publish a white-paper, "Educational Inequality in Ohio." Thank you Dr. David Staley for the many conversations we had contemplating new paths for the future of higher education.

I also want to thank my parents, Louise and Ed, for their unwavering support during this journey. Thank you also to my friends who were always there with words of encouragement. You always believed in me and I'm forever grateful.

Vita

1985-1989……………………...Tabb High School
1989-1993……………………...B.S. Secondary Social Science Education,
University of South Florida
1994-1995……………………...Outdoor Education Instructor, YMCA Storer Camps
1995-1997……………………...Teacher, Social Studies Department, East Middle School
1997-2000……………………...Teacher, Social Studies Department,
East Kentwood High School-Freshman Campus
1998-2000……………………...M.Ed. Educational Technology,
Grand Valley State University
2000-2008……………………...Teacher, Social Studies Department, Trenton High School
2004-2008……………………...Department Head, Social Studies Department,
Trenton High School
20008-2013…………………….Instructional Designer, Baker College Online
2010-2014……………………...Instructor, Baker College Online
2013-2016……………………...Sr. Instructional Designer and Faculty Development
Coordinator, Office of Distance Education and eLearning,
The Ohio State University
2016-2019……………………...Professional Development Manager, Office of Distance
Education and eLearning, The Ohio State University
2017-Present………………….Online Instructor, EHE Dennis Learning Center,
The Ohio State University
2019-Present………………….Learning Analytics Consultant, The Ohio State University
2019-Present………………….Steering Committee member, EDUCAUSE Student
Success Analytics Community Group

Publications

Arnold, K., Feng, L., Ham, M., Miller, A. (2020, May 28). EDUCAUSE COVID-19 QuickPoll Results: Student Success Analytics. *EDUCAUSE Review*. https://er.educause.edu/blogs/2020/5/educause-covid-19-quickpoll-results-student-success-analytics

Voithofer, R., & Ham, M. (2019). Ethical issues and potential unintended consequences of data-based decision making. In B. Khan, J. R. Corbeil, & M. E. Corbeil (Eds.),

Responsible analytics and data mining in education: Global perspectives on quality, support, and decision making. (pp. 213-228). Routledge.

Ham, M. (2018). Theories of innovation adoption and real-world case analyses. In Correia, A. (author), *Driving educational change: Innovations in action.* Pressbooks. https://ohiostate.pressbooks.pub/drivechange/

Dueker, S., Hawk, N., Nelson, J. L., Nelson, M. J., Walterbusch, T., & Wargelin, L. A. (Eds.). Aras, A., Ayoub, S. S., Bhaktha, N., Brady, A. C., Bussell, K., Dueker, S., Hall, S., Ham, M., Hawk, N., Jenkins, B. R., Kearney, R. C., Kim, Y., Kulesza, A. E., Kuznetcova, I., Lo, M., Mayo, J. W., Mulyadin, T., Nelson, J. L., Nelson, M. J., Paulson, J. L., Rivera, M. D., Robinson, S., Shi, Y., Tiba, E., Tornwall, J., Van Jura, M. J., Walterbusch, T., Wargelin, L. A., & Zhao, K. (2017). *Educational inequality in Ohio, 2016-17* (Research Methodology Center White Paper No. 1). Retrieved from https://rmc.ehe.osu.edu/files/2015/10/EHE-RMC-White-Paper-001_Educational-inequality-in-Ohio_March17.pdf

Fields of Study

Major Field: Educational Studies

Table of Contents

List of Figures

Chapter 1. Introduction

The Research Problem

Beyond conducting research, universities are in the business of educating students. To that end, many universities have been turning toward the use of data to inform decisions they make on how to allocate resources to best meet the needs of their student body. The prevailing notion is that the more a university knows about their students, the better they can provide the services and support needed to facilitate student success. During the first two decades of the 21st century, the amount of data produced by and about each student every day has been staggering—to the point of being called “big data” which is characterized by its enormous volume, variety, velocity, and value to end users. Universities are spending significant portions of their budget on systems that enable the collection, storage, and analysis of all of that big data. For the purposes of this study, data analysis to inform decisions impacting the student educational experience is referred to as student data analytics. There is obvious anticipation about the potential benefits for students if the results of analytics are applied "well”—the definition of which varies by individual and institutional entity—yet there are significant red flags related to student autonomy, data privacy, educational equity, and educational value.

Background of the Problem

The amount of student data being collected, stored, and used by colleges and universities has been met with both alarm and excitement. There is excitement about the possibilities for improving student educational experiences while also alarm over the potential risks to students which may have lasting consequences. Today is it becoming routine in higher education to see data analytics used to help leaders address challenges such as improving student performance, outcomes, and persistence (Picciano, 2012). Analytics assists in the building of models for personalized instruction, mapping learning domains, evaluating learning support from a learning management system (LMS), and scientific discovery about students (Baker, 2010). Inherent with each of these uses for data analytics is a risk to student autonomy, privacy, equity, and educational value.

There are four main types of data analytics that are common in higher education: academic analytics, descriptive analytics, predictive analytics, and learning analytics. Academic analytics encapsulates all data analytics at an institutional level that impact administration, research, management and resource allocation and is used to support strategic decision-making. An important component in academic analytics is descriptive analytics which analyzes historical data about students, research, teaching practices, and administrative processes to identify meaningful patterns from gathered samples. While descriptive analytics focuses on reporting information from the past, predictive analytics aims at identifying trends and associations between related variables in order to estimate the potential of future events or outcomes. The principal objective of predictive analytics is to identify future risks and opportunities based on analyzing trends and identifying associations that otherwise might be missed if relying solely on descriptive analytics (Daniel, 2015). Used in planning for the future, predictive analytics can discover answers to who, what, when, and where in order to help institutions decide their desired outcomes and answer the questions of why they chose a particular outcome and how to go about achieving the outcome (Rajni & Malaya, 2015) such as implementing academic interventions for students identified as "at-risk" through early alert systems to help boost graduation rates (Tampke, 2013). More specifically, predictive analytics can provide insight to help answer such questions as which students will enroll in certain programs and courses, which students will require extra support resources to graduate, and which programs are trending up and those that will be obsolete within a certain period of time based on industry demands (Rajni & Malaya, 2015).

Feeding predictive analytic models is not only descriptive data but also learning data, or learning analytics. Learning analytics is "the collection, analysis, use, and appropriate dissemination of student-generated, actionable data with the purpose of creating appropriate cognitive, administrative, and effective support for learners (Slade & Prinsloo, 2013, p. 1512). Learning analytics focuses primarily on the teaching and learning experiences within a course or a program and using grades and other learner related data to improve student success. Learning analytics is of particular interest to instructors as its purpose is to optimize learning and the learning environment. It operates by collecting, measuring, analyzing, and reporting data on learners and their behaviors within the course learning management system (LMS). At the teaching and learning level, learning analytics is concerned with improving learner success. At a broad institutional level, learning analytics techniques and software are often used to improve overall organizational effectiveness by maximizing processes and workflow through the examination of academic and institutional data (Daniel, 2015). Although learning analytics can be thought of as data analytics for the express purpose of analyzing activities pertaining to learning—including interventions for students deemed at risk of failing—the primary focus of learning analytics is often to inform methods for improving student retention rates (Picciano, 2012).

With such a scope of use for student data analytics, policies concerning data security, privacy, and ethical use would be expected at the institutional level. However, institutional data policies tend to address the security of student data and levels of access with little if any mention of ethical use (Roberts, Howell, Seaman, & Gibson, 2016). Institutions can use student data to promote their goals as leaders see fit. For example, they can use data analytics to plan effective programs and strategies to support students who are predicted to struggle, but with help could achieve success or institutions could use that same data to cull the predicted struggling students from their student body (Ekowo & Palmer, 2016). The latter solution would be less costly for the institution while also serving to increase their enrollment and graduation rates thus helping them earn a higher national ranking against competing colleges and universities, but many would deem it unethical. Instead of helping struggling students achieve success through academic assistance programs, those students are not just left to their own reconnaissance but actually pushed away by the institution from the resources that would benefit them. Hence, practices such as this are cases an institution may want to ensure does not happen. As the above scenario demonstrates, while an action or behavior may be legal, that does not necessarily mean it is ethical. For the purposes of this dissertation, ethical arguments are made from the framework of Paulo Freire's *Pedagogy of the Oppressed* and assume an intent of supporting student autonomy, privacy, equity, and educational value through engagement in policy debates and decision-making.

Higher education institutions have been collecting large amounts of student data for many years. The access and use of that data has been protected by institutional data policies. When investigating data policies at institutions around the United States, it is difficult to find policies available on the institutional websites. If policies can be found, then the access to many of them is restricted to those with institutional email addresses or other proof of affiliation. In this manner, data policies are hidden for many outside the institutions—what they contain regarding content and procedures are secret from outsiders. Even for those who can access the policies, the documents may be difficult to understand due to a proliferation of legal jargon. Based on anecdotal evidence stemming from my informal conversations with leaders in higher education—specifically within institutional technology (IT)—what many policies don't address is data use. The privacy and security of data has been primary in higher education data governance. This is evident in how types of data are structured by levels of sensitivity and basing restriction of access to certain data according to those levels along with the emphasis on adhering to regulations established by the Family Educational Rights and Privacy Act (FERPA). It is also evident through the increased inclusion of Fair Information Practice Principles (FIPPS) in data governance conversations across higher education institutions. How users of the data, once accessed, manipulate and apply the analysis of the data has not been a primary concern from the Office of the Chief Information Officer (OCIO). In some cases, data is collected but not immediately put to use. It is stored for some possible future use not currently identified by the institution (Hubin, Hirsch, & Ham, 2017). Informal conversations with leaders at institutions who manage student learning data has highlighted that there are no overarching laws, nor consistent institutional policies governing the length of time data can be kept. Each institution is governed by state regulations if they exist or they establish their own timeline for data retention depending on their predicted needs for the data.

The types of data being stored by institutions is growing and changing. Institutions have expanded beyond simple historical and academic data and are now also collecting swipe card data, WiFi usage, and LMS activity data (Hubin et al., 2017). With all of this data being collected and stored, institutions seem to be aware of important considerations to be made regarding student data use. The increasing use of predictive analytic systems in recruitment and enrollment management, academic advising and retention, and personalized learning seems to also have led some to start questioning whether advanced data practices might have unintended consequences. Such consequences may be limiting student autonomy over their choice of academic program, directing their learning experiences within an LMS without affording them opportunity to select their own learning path, or inadvertently opening doors for some groups of students while closing them to others based on historical and potentially biased data (Ekowo & Palmer, 2016; Herold, 2017; Johnson, 2014; Kurzweil & Stevens, 2018).

There are conversations happening informally in both the administrative and scholarly arenas about the ethical challenges around the use of data as the ability for organizations to collect data from many sources evolves. The Unizin Consortium is an institutional member driven collaboration between fourteen colleges and universities to improve access, affordability, and learner success among their institutions. Unizin is working on data management solutions for member institutions (Scott, 2017) to help them develop strategies and tools to pool de-identified student data collectively for research purposes. There have been informal conversations among Unizin representatives from the member institutions about ethical use of data in general and a desire to open up more dialog around this topic. Some of the challenges brought up in these conversations include how to bridge the understanding of ethical data analytics between the technically focused system programmers and data analysts and the pedagogically focused faculty and administrators. Another question raised often in conversation is how to operationalize learning analytics at a large university such as those making up the Unizin membership.

Recent articles have demonstrated interest in the subject of ethical data analytics, approaching it from different angles. In their research article, "Learning Analytics: Ethical Issues and Dilemmas", Sharon Slade and Paul Prinsloo (2013) used a socio-critical lens from which to analyze ethical implications of learning analytics in higher education with regard to location and interpretation of data, informed consent, privacy, de-identification of data, and classification and management of data. Ben Williamson (2016) wrote in his article "Digital Education Governance: Data Visualization, Predictive Analytics, and Real-time Policy Instruments" that the development of predictive algorithms has been problematic in that bias has been inadvertently built into the code libraries—large files of code created by coders and passed around coder networks online to develop software—which lead to results that can perpetuate discriminatory practices.

Another example of the heightened interest in ethical use of student data is the November 2017 issue of *Theory and Research in Education* which was entirely devoted to data use and data ethics. In this issue, Tammy Harel Ben Shahar explored the implications of information and communication technologies (ICT) on distributive justice in education in her article "Educational Justice and Big Data". As it pertains to the educational landscape, distributive justice examines questions of how educational resources should be distributed. ICT in education was defined by Shahar (2017) as "the use of electronic devices equipped with interactive platforms and applications that enable students to access learning material, perform educational tasks, and communicate with their teacher and peers and enable teachers to assign these materials and tasks, and evaluate them" (p. 307). ICT also includes the production of vast amounts of student and teacher data at a granular level as a result of their interactions with the technologies. Data on task performance, time on task, queries using search engines, and content of communications within group discussions or emails are some of the data produced and collected through the use of ICT (Shahar, 2017).

Questions around the use of one type of ICT—adaptive learning technologies to personalize learning—have also gained attention, most notably in the k12 arena although the concerns are not limited to k12 education. Concerns arise because there is little empirical evidence to suggest that personalized learning through the use of adaptive learning systems has a significant effect on student learning. Trying to fit each individual's learning experience into a mold of what might be considered typical for most students can undermine the unique skills, abilities, and learning preferences of each student (Herold, 2017). Pushing forward with such initiatives without due examination could lead to educational inequities among students even within the same class. (It is worth noting that for the purposes of this dissertation, equity refers to the distribution of resources fairly if not exactly equally as would be the case if we were to analyze these issues within the context of equality; Stone, 2012.) For example, if the algorithms—the programmed steps to accomplish a task (Khan Academy, n.d.) behind the adaptive learning systems are created based upon the behaviors of "typical students" then those students who fit that mold will benefit from the adaptive learning experience while those students who do not fit the mold will not receive such benefits and may even be overlooked if the instructor relies on the system to provide individualized assistance while they attend to other matters such as grading or lesson-planning. Thus, students experience inequities as the students fitting the norm of typical learning behaviors—as defined by the data and algorithms used to program the adaptive learning system—receive the assistance they need from the adaptive learning system while the students not fitting the norm of typical learning behaviors do not receive what they need to achieve success.

A pronounced critique of adaptive learning concerns the reliance of large technology companies needing access to vast amounts of student data in order for the adaptive learning systems to function as intended (Herold, 2017). This need has led to schools justifying collecting a variety of student data so they have it when it is needed as they look at transforming educational delivery models with adaptive learning. Along with the concerns over bias in algorithms used behind the scenes of adaptive learning, security of the data used in programming caused some angst as word of the data security breach at Edmodo, a popular digital learning platform, made national news. In the consumer market, the issues around data privacy and use have come to light after widespread data harvesting—the gathering of information from websites to organize in a database for later analysis (Import.io, 2019)—and application to create personalized user interfaces on company websites such as Facebook and online shopping sites were made known to the public. The fear for educators is that the same problem might happen as we see institutions scrambling to adopt adaptive learning technologies before considering the structures behind them, how they operate, the data needed, and whether the benefits to student learning outweigh the risk to student autonomy, equity, and privacy (Herold, 2017).

There are two key drivers behind the push for robust student data and its complex analysis: accountability to stakeholders and competition for students. Since 1967 when the Carnegie Commission on Higher Education was formed—and eventually transitioned into the Carnegie Council on Policy Studies in Higher Education in 1974—universities have been researched by external stakeholders such as corporations and government to discover institutional information from fiscal health to curriculum to graduation rates. Today, universities annually contribute standardized data to the Integrated Postsecondary Education Data Systems (IPEDS) and accrediting agencies keep a watchful eye on the curriculum in certain areas to ensure standards are being met (Thelin, 2011). The other application of data is in efforts to attract and enroll students. The beginning of the competition between higher education institutions for students can be traced back to the creation of IPEDS and when the Carnegie Corporation decided that the classification of data within IPEDS was too vague and that new classifications of the different types of institutions needed to be created (Thelin, 2011). In 1983, the first *US News and World Report: College rankings* report was published (College Rank, 2016) propelling rankings into a very public sphere of influence to attract future students. The following explores how universities are handling and using student data, focusing specifically on how it is used to inform policies and practices that impact the services offered to students and improve their overall educational experience.

Leveraging Student Data

Student data is being leveraged to inform decision-making to manage enrollments, inform academic advising, and guide personalized learning. Before looking at the specific use cases for student data, it is important to clarify some of the common terminology used in discussions around student data applications.

Terminology. The term “big data” has been used often in the media and many articles have been written about how companies such as Target, Amazon, Google, and Facebook are using big data to track consumer activities. Big data has reached higher education as well and needs to be understood in order to be used to benefit institutions and students while mitigating potential unintended consequences.

Big data is commonly described by the nature of its volume, velocity, variety, and value—the "4Vs" of big data. Volume refers to the amount of data which, with big data, is massive. Velocity is the speed at which the data is processed. Today, the speed is faster than anything we've seen before (Cai & Zhu, 2015) as data is collected continuously and analyzed by automated systems in real-time (Williamson, 2016). Because of the vast amount of data collected and analyzed at high rates of speed, there is great diversity (i.e. variety) in the types of data processed—audio, video, text, numeric, etcetera—which require expansive data processing systems. Value is the fourth "V" of big data although it is not solely a descriptor of big data as it pertains to smaller data as well (Cai & Zhu, 2015). The value of data increases as the data store increases. With big data, instead of storing data about groups, data about individuals can be analyzed and used to fine tune personalized experiences. Another approach to the value of big data is in simply having the data even if the purpose of having the data is not yet known. Some researchers disagree with this perspective and think that knowing the questions to be answered should still help determine what data to collect even though the technology exists to extract more than what's needed (Vorhies, 2014). One example of big data in higher education is the recording of each individual transaction made by students every day and storing those transactions to be analyzed. There are several thousand transactions generated each day by every single student on campus or interacting with the university website or in the learning management systems (LMS) which makes this a prime example of big data (Picciano, 2012).

Once collected, data is stored in warehouses and mined—analyzed for patterns, associations, and trends that may be missed through traditional analysis methods (Import.io, 2019)—in order to inform decision-making (Picciano, 2012). Kumar and Chadha (2011) point specifically to using data mining to aid in curriculum planning, predicting student registration and performance, detecting cheating online, and identifying anomalies or false data. Course recommender systems also rely on data mining to operate effectively. Recent trends point most often to data mining for predicting student success and personalizing the student learning experience. Data mining allows for student progress to be tracked which makes possible course recommendations during in-person or virtual advising sessions. Such recommendations can be made based on student performance in previous courses in a similar manner to recommender systems used by the entertainment industry which provide suggestions a user receives through platforms like Netflix based on their previous viewings. Based on mined data, some systems have the capability to initiate interventions to influence student action. Universities are also increasingly looking at modeling social networks around campus—student social integration, facilities usage, administrative data and other personal networking data—to study trends that could help predict retention (Johnson, 2014).

Mined data is eventually fed into some sort of analysis application. Although defined differently by administrators, vendors, and educators, the core definition of data analytics is "the use of data to determine courses of action especially where there is a high volume of transactions” (Picciano, 2012, p. 13). Data analytics is becoming more pervasive in decision-making by higher education leaders and it is important to understand the different types of analytics run for various purposes. No matter what the question is for which an answer is being sought, all analytics are powered by algorithms. Algorithms are mathematical calculations with the purpose of processing big data to arrive at a recommendation or conclusion (Clayton & Halliday, 2017). Algorithms are programmed by people and thus are subject to being influenced by the biases of the programmers (Williamson, 2016) and the decision makers who specify the algorithms. This is important to remember when exploring the different types of data analytics and how their results should never be unquestioningly accepted and acted upon. The following discussion explores specific strategies colleges and universities have implemented using predictive and learning analytics to inform their recruitment and enrollment practices, improve their academic advising tactics and retention rates, and personalize the student learning experience.

Applications for Recruitment, Enrollment, and Retention. Universities are applying predictive analytics to refine their enrollment management strategies in order to recruit prospective students and retain current students. It is common for universities to leverage data to streamline recruitment practices in order to direct efforts toward students that show the highest chance of enrolling and achieving success at their institution. Historical data can help in the identification of specific factors that historically have led to students enrolling in courses once admitted to the university. Predictive analytics then combines that data with profiles of prospective students in order to call out those predicted to enroll which gives recruiting officers a pool of potential students to target. A focused strategy such as this improves graduation rates while saving the institution money. Wichita State University has leveraged data analytics for this purpose by having the prospective students compared to the historical data based on variables such as race, gender, ethnicity, standardized test scores, high school grades, and whether they are first generation college students in order to assign a score from 0-100. Based on the score, officials can allocate resources according to those students most likely to attend the university (Klasik, 2013). Franklin & Marshall College used admissions and financial aid data to conduct a retrospective analysis in order to assess whether their merit-based programs actually attracted the students they wanted at their college (Biemiller, 2017).

Beyond recruitment, universities turn to student data analytics to retain existing students. Georgia State University tracked outcomes data to make data driven decisions regarding financial and academic assistance for students. GSU has mined past student performance data since 2003 to provide extra support for courses that student tend to struggle through (Kurzweil & Stevens, 2018). Some institutions, like Howard Community College in Maryland and Jacksonville State University in Alabama, have broadened their use of predictive analytics beyond targeted student recruiting by employing it to anticipate financial needs of incoming and returning students, identifying those who will most likely make use of financial aid if offered. Specifically, HCC and JSU made budgetary adjustments along with admission and financial aid policy revisions based on forecasts of future enrollments run through predictive analytics (Ekowo & Palmer, 2016).

Applications for Academic Advising. Predictive analytics is advancing rapidly in academic advising where many college and university campuses have pioneered the use of student data analytics and predictive analytics in order to support students through the use of early alert systems—intervention programs to improve student success rates—and guidance in selecting degree programs for majors. Some institutions have employed predictive analytics to help provide interventions for students as soon as they begin showing signs of academic struggle that could lead to a dip in performance (Willey, 2018). Early alert systems have been used by academic advisors in order to identify students struggling in their courses and provide targeted interventions before they fall too far behind to recover.

The following are a few cases where student data has been used at different institutions for a variety of purposes in the area of academic advising. Austin Peay State University implemented a course recommender system called Degree Compass that worked in conjunction with My Future—a program that mines student data to identify those courses central to success in each degree program and helps students select programs and provides career information for each program. Together, Degree Compass and My Future could identify the majors where students will be most likely to achieve success based on predictive analytics. Advisors at Georgia State tracked students' progression through their courses and majors using the Graduation and Progression Success system (GPS). GPS was created from millions of student grades earned from the previous decade in order to make predictions about current and incoming students' achievement.

Although there hasn't been real empirical evidence to show that early alert systems have a direct impact on student retention (Cuseo, n.d.) academic advisors at Temple University have felt otherwise. They have seen retention and graduation rates at their university dramatically improve since implementing a new system in 2008. Between 2001 and 2014 the student retention rates between the freshman and sophomore years rose by twelve percent and the four-year graduation rate rose by twenty-four percent while the six-year graduation rate rose by eleven percent. The university did not report implementing other measures along with the early alert system so with these results, the Temple University advisors believed the perceived intrusiveness of the early alert system was worth managing through what they called aggressive advising tactics—they were no longer waiting for the students to come to them (Ekowo & Palmer, 2016; Felton, 2016).

Regardless of the dearth of empirical evidence suggesting that early alert systems have been able to significantly increase student retention, advisors may be more comfortable intervening for a student flagged as at-risk if the early alert system pulled data from current student behavioral indicators rather than just demographic and pre-college performance. The behavioral data being used by these systems is pulled from a variety of contexts such as failing to register for classes, poor classroom performance, not renewing financial aid for housing and so on. This suggests one of the potential strengths of an early alert system in that it can help bring together entire support teams from offices across the university to assist students demonstrating a risk of dropping out (Cuseo, n.d.).

Applications for Personalized Learning. A blending of academic advising and teaching and learning happens often as predictive analytics, early alert systems, and data dashboards are starting to be used across platforms to communicate between students, instructors, and advisors. In some cases this personalized attention is aided through learning analytics dashboards—user interfaces where instructors view student data in a condensed format—which give instructors the ability to notice who is falling short of meeting certain performance criteria so they can reach out to them early for assistance. These dashboards tend to run alongside the course learning management system providing easy access to timely data which instructors of large enrollment classes find quite valuable (Ekowo & Palmer, 2016; Oxman & Wong, 2014). One university that incorporated the use of dashboards in the learning environment was the University of Iowa with Elements of Success in their General Chemistry I courses. Elements of Success is a home-grown predictive analytics system that combines past performance data of prior students with data on current students—backgrounds, time spent on homework, their understanding of the course content—in order to provide every student with a dashboard displaying how they are performing compared to fellow students in the same class. Elements of Success can also forecast an individual student's final grade should they remain on their current performance path (Biemiller, 2017).

Pierce College—a community college in Washington state—had a wholly different strategy for using student data dashboards to improve student performance. They analyzed data by comparing student performance across different sections of a course to see if the assigned instructor had any bearing on the student success outcomes. The dashboards displayed data on their students' performance in their particular course section and also how those students were performing in other courses. Although just in the pilot stage, having access to transparent data allowed instructors within a program to have open dialog with each other about grading practices and collaborate on plans to improve the student experiences in their courses with the goal of improving completion rates (Gose, 2017).

Personalizing learning experiences is a growing trend especially in online learning and for traditional, large enrollment, face-to-face courses using e-learning course sites (Ekowo & Palmer, 2016). Adaptive learning technologies—sometimes referred to as adaptive learning courseware—used along with a learning management system (LMS) such as Canvas or Blackboard are one way to customize learning activities for each student in the course. Adaptive learning systems function by tracking individual student activity and progress in the LMS and automatically adjusting subsequent activities provided to the student based on their performance on previous learning activities (Oxman & Wong, 2014).

Several universities have leveraged adaptive learning technologies in different ways. Georgia State University used an adaptive learning system in their introductory math courses which were delivered in a hybrid format—the courses were delivered partially online and partially face-to-face in a physical classroom. The introductory math courses had high enrollments and the adaptive learning system helped address the individual learning needs of each student in those large courses (Ekowo & Palmer, 2016). Arizona State University used ALEKS—McGraw-Hill Education's adaptive learning courseware—in their Global Freshman Academy math courses as well as their online and traditional on-campus math courses. Each student's learning experience was tailored to their specific learning needs by the program (Lestch, 2017). Colorado Technical University has used intellipath—its own adaptive learning system—to provide personalized learning experiences to students (Ekowo & Palmer, 2016). Intellipath alters how a course is delivered for each, individual student based on the student's demonstrated abilities (Becker, Cummins, Davis, Freeman, Hall Giesinger, & Ananthanarayanan, 2017). As a student works through learning content in the learning management system (LMS), intellipath presents learning pathways in the form of practice questions based on how the student is answering previous questions. This allows students to move more quickly through content they know well and spend extra time practicing in areas in which they struggle (Ekowo & Palmer, 2016). While most adaptive learning experiences have been designed for STEM subjects, the University of Georgia has been conquering new ground by working on a system that would bring adaptive learning to English composition courses. This adaptive system will guide students through foundational concepts to begin with and allow them to progress at their own pace through the more advanced concepts in composition as they are ready (Becker et al., 2017).

There are many other advances happening using adaptive learning courseware. The technology is expanding beyond individual learning experiences to be used in collaborative environments as well. There is promise in the ability of advanced tools to "automatically sort users into groups with shared interests and recommend information sources based on user interests and browsing habits" (Becker et al., 2017, p. 39). That said, a foundation for heightened interest in adaptive learning courseware is that research has shown that students feel more in control of their learning, experience greater enjoyment in the course and have more confidence in mastering the content when adaptive learning technologies are employed in a course (Becker et al., 2017).

Challenges and Focus Areas

As previously mentioned, the amount of data collected about students is growing and not simply historical or strictly related to activities in the LMS. Today, data is being collected regarding a variety of student activities such as when students swipe IDs at dining halls, libraries, student services, and in tracking their locations when accessing the campus Wi-Fi network (Hubin et al., 2017). Location tracking has become more important at many institutions during the COVID-19 pandemic to help ensure students, faculty, and staff maintain established protocols for a safe on-campus environment. In addition to managing safety during the pandemic, an example of this type of data collection is occurring in California. The California State University system, consisting of twenty-three campuses, will pilot an application by Degree Analytics that uses network data collected from student mobile devices as they move around campus using the Wi-Fi network. The hope is that this data will provide insight into what services students are using around campus and for how long and at what times. From this data, better informed decisions about what services and amenities to allocate resources to can be made (Blumenstyk, 2018). The end goal for the plethora of data collected by universities is to provide more well-rounded data to produce a more complete picture of the factors contributing to student retention and success, or lack thereof, giving the institution information on how to meet the students' needs (Hubin et al., 2017).

Problem Statement

Current literature and discussion regarding student data analytics focus on technical strategies and system capabilities along with how higher education institutions are using it most often for recruitment and enrollment, and advising purposes. Strategies for learning analytics solutions can also be found in the literature, though the strategies are typically narrow in scope and specific in purpose to the needs of a course, program or department. The topic of ethics is discussed theoretically with regard to data analytics where the call to action, if there is one, is for institutions to adopt guiding principles for learning analytics specifically or for broad-level privacy principles. What guides analytics work with student data is not just a vision but the boundaries of what is allowed by law and policy at the university—along with state, federal, and, increasingly, international regulations—but also a vision for applying analytics ethically.

Very little literature exists regarding how policy can address ethics with regard to student data analytics. What is plain to see throughout the literature pertaining to student data analytics is how great the reliance on it is for various reasons. Without an ethical compass and review processes guiding policy-making and data process decisions, student data analytics could lead to unintended consequences with far-reaching ramifications for students. Potential ramifications might include limiting student autonomy in selecting their own academic path which impacts career trajectory and a student's sense of control over their future—not just while enrolled but also after graduation. Other unintended consequences could manifest in a perpetuation of inequitable educational experiences if the predictive analytics behind systems for academic advising and adaptive learning use data sets and algorithms that include bias. Inequities experienced along racial, gender, socioeconomic or other lines in higher education can impact the opportunities that are available after graduation for different groups of graduates. Students may also become so accustomed or even apathetic to the university's access to so much of their data that after graduation they continue to willingly grant access to their data to requesting organizations without question. The following review of ethical issues higher education institutions face when using student data analytics focuses broadly on issues related to student autonomy, privacy, equity, and educational value. Student autonomy relates to the degree to which a person is in control of their own experiences and outcomes. Privacy refers to the extent to which a person or organization knows personal information about another. Equity is the manner in which a person is afforded resources they need to achieve their definition of success. Educational value refers to the value that educating students in the ways in which their data is used by others holds for them, their institution, and society as a whole. Educational value is fortified through student engagement in policy debates and decision-making processes. (Each of these ethical issues will be discussed in greater detail in chapter three.) To develop this literature, this study dives deep into the policies and procedures at two universities to see how written policies and procedures address the ethical use of student data analytics in order to mitigate potential harm to their students.

The Study

Institutions of higher education, whether small or large, public or private, are collecting and storing student data, and using data analytics with the intention of effectively supporting the enrollment and retention goals of the institution while providing high quality student experiences from individualized academic advising to personalized learning experiences. However, institutional decision makers are just beginning to explore how the benefits and risks to student autonomy, data privacy, educational equity, and educational value should be considered in the development of data policies and practices.

In order to propose a set of ethical principles around data policy, this dissertation provides a historical overview of data in higher education, insights into the ethical issues involved with student data analytics in higher education, a comparative case study of the data policies and procedures at two institutions, and an ethical analysis of the policies and procedures that the institutions employ. The historical analysis examines higher education institutions in the United States—starting in the 1700s—in the types of data they collected, how they used it and the development of computerized systems to help process the data. The review of ethical issues higher education institutions face when using student data analytics focuses broadly on issues related to student autonomy, privacy, equity, and educational value. The case study explores two large universities and their data policies that specifically pertain to student data along with published information about procedures around the collection and use of student data to see how ethics is addressed. The ethical analysis pays specific attention to how the policies and procedures employed at each institution support or hinder student privacy, autonomy, equity, and educational value.

Purpose of the Study

The purpose of this study is to explore how universities address the ethical risks inherent in student data analytics through their policies and procedures related to the collection, use, and protection of student data. The study investigates specific ways universities are using student data analytics and the drivers behind those strategies. Questions about the ethical implications of strategies employed by the subject universities in this comparative case study are discussed with the understanding—based on each institution's mission statement—that the educational purpose of the university is to provide a liberal education that stimulates critical thinking (Clayton & Halliday, 2017), prepares students for employment with necessary skills, and supports diversity, individuality, and freedom (Pasquerella, 2019). Common themes and gaps in policy and procedures are identified when considering the potential for unintended consequences of student data analytics related to supporting the goals of a liberal education and promoting four key values: student autonomy, student data privacy, educational equity, and the educational value of student data analytics. Ultimately, recommendations are provided to help university leaders and policy-makers shape future policies and procedures to support each of these key values.

Research Questions

The overarching research question for this study was: How are institutions of higher education writing institutional data policies and procedures that address the ethical complexities of student data analytics in an era of big data in order to protect the institution and its students from potential unintended consequences? The study begins with a historical analysis to answer the questions of the types of data collected and how it was used from the mid-eighteenth century to the early twenty-first century. Also explored is the automation through computer technology of data collection and processing tasks and the impact on how higher education institutions function. The following are questions that guided the research:

1. How has data analytics developed through the history of higher education in the United States?
2. What principles should guide the ethical use of student data as big data in terms of autonomy, privacy, equity, and educational value?

3. At universities using big data, what current institutional policies and procedures are in place with respect to student data?
4. How should those policies be evaluated in terms of the principles developed in answer to question two?

Significance of the Study

The purpose of this study is to present a broad outline of ethical principles that institutions need to address around data analytics policy to ensure student autonomy, privacy, equity, and educational value while minimizing risks and unintended negative consequences. Identifying common themes related to ethics and how mitigation of unintended consequences associated with the misuse of student data is already being addressed adequately is an initial step. Identifying gaps where policies and procedures are lacking in attention to such concerns can aid in the planning and focus of future studies. Gaps where certain ethical risks are overlooked should be the target areas for future study. A better understanding of these gaps will also provide a foundation for higher education institutions to develop strategies to close existing gaps within their institution. Through exploration of the research questions, recommendations for how university leaders and policy-makers can support the four key values of a liberal education while advancing the use of student data analytics are provided. These recommendations can be used as a baseline to craft policy that addresses the ethical risks associated with student data analytics while also describing goals for implementing new policies and practices.

The heart of this study is about students. Increasing student autonomy, addressing student privacy issues, closing equity gaps, and fortifying the education about and engagement with data policy helps produce graduates who will be prepared to succeed in their chosen field while also feeling a greater sense of personal agency over their future within each community of influence they belong. Beyond this, graduates will understand how they fit in the larger scheme of the data world and how their actions can promote the ethical use of data on a personal level and more broadly in their local, regional, and global communities.

An appropriate place to begin this study is in the past. Looking at the history of how the analysis of student data in higher education evolved since the beginning of higher education in the United States can help identify why policies and processes are what they are today. Past policies and practices by institutional administrations explain current educational inequities while past data practices help explain the need for action to combat bias embedded within data sets and algorithms. Examining the increasing complexities of data analytics over time demonstrates the need for new perspectives and efforts toward data privacy. As history shows the world becoming more reliant on big data and analytics, the importance of developing students who understand their role as producers, owners, and users of data with autonomy to engage as policy influencers with an eye toward ensuring the ethical use of data becomes evident. So begins the historical exploration in the next chapter.

Chapter 2. The Historical Use of Data in Higher Education

Introduction

Higher education has a rich history in America, a history characterized by changes in the organization and function of its institutions. The purpose and mission of these institutions has evolved through the centuries from who they serve, why they exist, and how they operate in order to carry out their goals. Data has been an integral part in many aspects of the business of higher education to one degree or another. Most recently big data has played a greater role informing leaders on the decisions they make guiding the path forward for their institution and the students they serve. The growth of data-driven decision making has tracked according to the growth of data collection and processing capabilities—all of which seem to be following Moore's Law which predicts the doubling of the amount of digital computing power that can be bought for a certain amount of money every eighteen months according to many in the field (Brynjolfsson & McAfee, 2014). The combined influx of data with greater computing power has led to many changes in how institutions handle their data and higher education has been working to adapt at an ever-increasing pace to the changing needs of society since it began in the United States. Looking into how data has been used by higher education as computing power has increased provides insight into the value it holds for institutions. It also can provide insight into how certain groups of people benefited from the data while others may have struggled. As society advances technologically, it will be imperative for leaders in higher education to consider their data policies and procedures to ensure they are aligned with their mission and with the goals of a democratic society. The data they leverage must support the institution while also supporting the students in their quest for success.

The following discussion highlights the rapidly changing landscape of data use in higher education. It demonstrates the need for institutions to consider how they will continue to use student data while safeguarding autonomy, privacy, equity, and its educational value. This chapter explores this evolution of data use in American higher education from the beginning of the nation to today. Eras of notable changes in computer and data use in higher education and their impact on academic administration and learning is discussed beginning with the time before computers—the 1700s through the early 1900s. From there, the era when computers arrived on the scene at higher education institutions is discussed. This is followed by a look at how the arrival of the Internet impacted institutions. The final era explored is when institutions began having to content with big data. Studying how data analytics evolved within higher education enables policy-makers and institutional leaders to make decisions informed partly from what has and has not worked in the past and why. This type of knowledge can be useful when assessing new data-related innovations for viability and adoptability—as history shows that some data and computer innovations were not adopted easily or leveraged to their full extent by colleges and universities. By understanding the challenges, successes, and failures of past data analytics related initiatives, policy-makers and institutional leaders can be better prepared to apply lessons learned to present and future decision-making efforts.

The Changing Role of Information in Higher Education

Data Eras

Before Computers (1700s - Early 1900s). In colonial America, it is estimated that just one percent of the population attended university—nearly all of them white men mainly from the northeast and wealthy Protestant families. The few colleges and universities that existed in the 1700s were for the elite in society and were not about training people for careers but rather preparing students to be educated members of society able to maintain their elite social status. In fact, degree completion was not important among students and professors, so students matriculated in and then left after a couple of years, ready to move on as clergymen or to follow other pursuits. The 1800s saw an increase in the number of colleges and universities vying for students (Rosen, 2011) although even by the 1880s only twenty-six institutions had more than two hundred students (Carey, 2015). Records were kept on each student and their course of studies while they were enrolled and grades were given during in-person recitations as paper was still very expensive.

In the late 1800s, the purpose for higher education changed when Vermont representative Justin Smith Morrill advocated for the Morrill Land-Grant Act which was signed into law by President Lincoln. The act granted rights to federal land to states in the western territories with the stipulation that at least one college be established for the purpose of educating the industrial class in the mechanical arts and other practical subjects. The reason behind the Morrill Land-Grant Act was to help grow a knowledgeable working class that would advance the American economy as the nation expanded westward and became more industrialized. Many of the land-grant universities that were created grew to become some of the largest and most productive institutions in America. The 1800s also saw the rise of the research university—a university primarily focused on individual research carried out by its professorate with students serving in apprenticeship roles, learning research methods from their professors. In 1876 Johns Hopkins was the first research university established in America. The university was organized into departments depending on research disciplines, an organizational model followed by many subsequent universities. The third purpose for a university that evolved during the 1800s was to provide a liberal arts education so people could grow as critical thinkers and knowledgeable in classic literature and philosophy (Carey, 2015).

Beginning in the 1880s, the industrial economy emphasized the importance of a quality college education. During this time there was a push toward creating efficiency and coherence within and among universities as part of the larger trending Progressive movement. American Progressivism brought with it an emphasis on ratings, rankings, and reputations of universities. As a result, it was during this time period when the Association of American Universities (AAU) was created with the agenda of developing processes for ensuring quality standards and standardization across institutions as it was feared that American institutions were lacking relative to their European counterparts. Soon after the formation of the AAU, the College Entrance Examination Board was established with the purpose of developing reliable standardized college admissions tests. Universities started to institute entrance requirements with a minimum threshold for admittance including minimum scores on college entrance exams. Secondary schools responded by changing their curriculum to prepare students to meet the new entrance requirements and coding certain high school curriculum as being college preparatory which was documented on high school transcripts (Thelin, 2011).

Scientific management became popular in business in the early 20th century as managers were looking to improve efficiencies and turn larger profits. This management trend soon found its way into education. As the demand for educational accountability increased, administrators adopted scientific management principles to operate schools more efficiently (Trujillo, 2014). This trend was seen early with departmental divisions in universities with deans and specialized faculty and staff. The mirror to the industrial business complex was further demonstrated by the view of faculty, even though highly regarded for their research, as labor in the service of management who were the deans and college presidents. Faculty were expected to answer to management and the board of trustees even as they enjoyed a certain amount of self-determination in their research, networking, and academic freedom (Thelin, 2011).

During this era, data was collected about students at American universities relative to demographics, enrollments, and grades and reported out annually by university presidents. In the early 1900s, attending college grew in popularity with the understanding that a good education would prepare a student for the rigors of managing in an industrial economy. The pressure was great to demonstrate educational effectiveness as competition for students mounted. As popular and important attending college was seen during the early 20th century, data showed that access to a college education was still elusive for many due in part to issues of affordability. Enrollment data showed that only five percent of the American population eighteen to twenty-two years old were enrolled in college and were, for the most part, white males. Tuition and fees at many colleges was unaffordable for most American families. Some institutions in the west such as Stanford University and the University of California implemented a "no tuition" policy. Although enrollment data was captured by these and other universities across the eastern and midwestern states that were known to draw students from financially well-off families, there was not sufficient data collected to understand how varying tuition rates affected the socioeconomic make-up of the student body at these different universities (Thelin, 2011).

Universities touted their ability to develop cohesive student bodies through academic and extra-curricular programs. Presidents looked to enrollment data to report on the effectiveness of their programs in retaining and graduating students each year. However, reports of the data were not always accurate in their depiction of what was really happening. A common practice in these reports across universities at the time was to include new students in the enrollment data when reporting on a particular graduating class's retention rate. For example, if a university wanted to show the enrollment data for the Freshman class of 1910 over the next four years through graduation in 1913, they would want to only include enrollments in 1911, 1912, and 1913 from that original cohort of students. What many universities did however was include new, incoming sophomores in the 1911 data—and new, incoming juniors to the 1912 data and so on—to boost their numbers in the report thus making it look like the university was more effective in developing a cohesiveness among their students than was the reality. The reality of the situation at many universities was a high drop-out rate when considering original Freshman student cohorts (Thelin, 2011). This does not tell the entire story of the effectiveness of higher education in successfully graduating students, though, as there was also evidence of many students transferring between institutions. These reports, however, were for the purpose of demonstrating a campus environment that fostered student cohesiveness which was, in actuality, lacking more than the official reports let on.

After World War I, Americans became more interested in higher education and enrollments increased significantly from 250,000 to 1.3 million between World War I and World War II. Higher education was no longer seen as only for the elite but rather for the masses. By 1937 nearly fifteen percent of Americans ages eighteen to twenty were enrolled in college. As enrollments increased and greater investments were made in colleges and universities, private organizations overseeing quality measures such as the Carnegie Foundation for the Advancement of Teaching (CFAT) and the Rockefeller Foundation's General Education board collaborated with the United States Bureau of Education collecting and analyzing data with the goal of demonstrating the effectiveness of university systems across the country. Most of the data analysis was done by the highly funded private organizations while the reporting out of findings was relegated to the federal bureau. Between 1920 and 1940, the work of these private agencies proved effective in strengthening the standardization and implementation of standards at universities—work that had begun back in the late 19th century (Thelin, 2011).

The 1930s and 1940s were the decades of the "mega-university" which were state universities with tens of thousands of students such as the University of California which claimed an enrollment of around twenty-five thousand and The Ohio State University with nineteen thousand. With the enormity of state universities came the corporatization of the business functions of the institutions starting with the membership of the board of regents and the university president. Traditionally those roles were filled by former members of the clergy or academics. During this period more of those positions began being filled by corporate philanthropists which also brought a corporate mindset to the operation of the university and a keen eye to the financing of projects and revenue creation. Ironically, the 1930s also witnessed substantial increases in tuition costs. According a one national survey, the average tuition fee rose from $70 (roughly $601 in 2000 dollars) in 1920 to $133 (roughly $1,143 in 2000 dollars) in 1940. As expensive as college tuition had become, it did not seem to deter students from attending. Students became creative in finding ways to cut their personal expenses by living off campus together in shared housing and limiting the amount of food they ate (Thelin, 2011).

During World War II, student enrollment dipped slightly then ballooned after the war as a result of the GI bill and increased state funding. By the 1949-50 academic year, total student enrollment in higher education had reached nearly 2.7 million—an increase of roughly eighty percent in one decade. Universities were then faced with the problem of expanding campuses to handle the swell of students. As a result, different types of post-secondary institutions came into being such as for-profit vocational and trade schools along with junior colleges which eventually became community colleges. With the increased demand for a college education, universities had to develop a means for processing a greater number of student applications with improved efficiency. Adding to this challenge was the fact that the applications from former military personnel using the GI bill were without traditional college preparatory records and high school transcripts. Admissions officers turned to standardized tests such as the Scholastic Aptitude Tests (SAT) and the American College Testing (ACT) service to aid with admissions decisions. As test scores allowed efficiency in admissions decision-making they also provided leverage for universities to increase their selectivity of students by requiring higher test scores for consideration. Admissions officers viewed the scores as important data that could be combined with other achievement measures like grade-point-averages to predict performance in college (Thelin, 2011).

Computers go to College (1960s - 1980s). Research and development efforts out of Stanford, the Massachusetts Institute of Technology, University of Pennsylvania, and the University of Illinois at Urbana-Champaign during the 1950s ushered in the era of computers in business applications (Picciano, 2012). However, not until the 1960s did administration offices begin operating computers for information processing of data such as student registrations, grades, and payroll. This change allowed the same information that had been tracked by hand before to be processed at far greater speed which was immensely helpful for large institutions. Smaller colleges like Bennington College were disinclined to make the switch to data processing via computers and opted to continue enlisting administrators to hand-process data such as results from questionnaires students filled out to evaluate their instructors each term (Rourke & Brooks, 1964). Computer technology also became integrated into American college classrooms and labs by the 1960s although the systems used then would be considered rudimentary compared to computer technology today. For example, data was collected and stored using Hollerith (punched) cards, sequential magnetic tape files, and massive mainframe computers (Picciano, 2012).

As computers were making their way into the classrooms and administrative offices on college campuses in the 1960s, educators and administrative personnel were faced with a changing work environment due to the new technologies and several universities made concerted efforts to provide training in order to ease the transition. Courses and workshops were not the only avenue for learning about the new computer technologies disrupting life as workers knew it. Publications were distributed about computers and user groups started up like the Project on Information Processing (PIP) by the National Science Teachers Association and the Educational Data Processing Newsletter by the Education Data System Corporation. Education institutions, especially large research universities, may have benefited by applying an integrated systems approach to staff training—sharing information across units as new technologies were integrated—but the departmental silos and the culture and structures unique to each functional unit of the university prevented such cost effective and time saving strategies from being leveraged during this time period (Bare, 1966).

Coinciding with the increasing integration of computers in higher education during this time was expanded interest in the potential for using data to inform policy through more vigorous analysis in the areas of admissions, curriculum development and facilities planning (Rourke & Brooks, 1964). The ability to use computers to automate administrative functions was welcomed emphatically by university administrative personnel. The vast amount of data being collected by large universities had become unwieldy and the idea of leveraging computers to handle all of the data efficiently was worth the money spent on the technology and necessary training. Because of the unsurpassed efficiency achievable by computers, there was general buy-in to the integration of computers to analyze data collected on students. Computers were seen as a "logical extension of man's capability" (Bare, 1966, p. 440). One state university in Massachusetts worked on developing an automated system for data processing to allow the university to apply more accurate criteria in the selection of students than was possible using conventional testing methods. Although slow to take off in the mid-1960s, some universities began investigating the running of "what-if" scenarios using computer programs to better inform decision-making (Rourke & Brooks, 1964).

As new as the data processing technology was in the 1960s university administrative offices, there were already concerns being raised about the potential effects computer use to process student information—applications, registrations, and class schedules to name a few—would have on depersonalizing the student experience (Rourke & Brooks, 1964). However, by automating the clerical tasks such as data processing, many registrar and admissions officers felt that they would actually have more time to spend working with students personally to provide advising and counseling them through their own program decisions, thus affording greater personalization. Smaller colleges were not necessarily expressing the same welcoming sentiments toward adopting computers for student information processing even while acknowledging the benefits to efficiency. For those colleges, the automation of registration processing removed power from individual departments to make their own decisions when selecting students for their programs. Departments were accustomed to being able to exercise their own discretion when making those decisions and were reluctant to give that up no matter what the efficiency benefits might have been (Rourke & Brooks, 1964).

In addition to using computers for analyzing administrative data, computers also started being included in some teaching tasks. However, using computers to automate teaching was greeted with much skepticism as instructors were warry of what they viewed as the invasion of their classroom, curricular, and teaching domains by computers. One obstacle to incorporating computers in teaching—even if instructors were willing—was the number of people needed to develop instructional systems and the huge costs involved. The evaluation of these instructional systems was yet another barrier to wide scale adoption as it proved difficult to identify, agree upon, and quantify the variables needed to assess the success or failure of the systems. This was the complete opposite of the experience administrative officers had when evaluating the benefits provided by computers operating in an administrative capacity where the student data was readily quantifiable such as demographics, grades, attendance, achievement scores, drop-out rates. Because of this, most administrators were comfortable producing reports demonstrating the return on investment in administrative computing technology (Bare, 1966).

Advances in computer technology continued through the late 1960s and 1970s. During this time, administrative records began being stored on magnetic discs. In the late 1970s and 1980s, mini- and micro- computers once again changed administrative processes. Student records and processing applications were moved off of mainframe computers and into these smaller, desktop machines (Picciano, 2012), easily incorporated into administrative offices for ready access to information and reports at any given time.

The Internet Arrives (1990s - Early 2000s). In the 1990s, data was used and analyzed for the purposes of decision-making regarding admissions, revenue, expenditures, matriculation, capacity, financial aid, and instructional transactions. The Internet brought about another significant shift in data processing with the introduction of web-based applications that allowed for more sophisticated methods for collecting and interacting with information. Even with these greater technological capabilities, administrators remained focused on data inputs and producing reports (Burrell, 2017). The wastefulness of the information management capabilities many universities had in their possession continued through the late 1990s. There were several reasons why higher education stalled in capitalizing on their technology resources even as businesses and the masses outside of higher education forged ahead; those reasons related to the silos of departments and operational units common at universities, each with their own collection process for data related to students, finance, and personnel. For example, the admissions office collected relevant data on admitted students while the registrar's office tracked student progress through their program of study and the related data. The financial aid office collected data on loans, payments plans, and other such information while the cashier's office handled all payment data. There were, of course, redundancies of data and one could wonder why not share student records between offices and departments. The problem was that each office used a separate database that was not designed to be shared and data would not match up so it could not be analyzed across platforms even if the database administrators were willing to do so, which most were not. They feared the electronic information getting into the wrong hands outside of their department and what might happen as a result (Edirisooriya, 2000).

Of course, none of these roadblocks to sharing data was new with the use of computers in higher education administration. Even in the early 20th century, university personnel recorded information by hand on ledgers using their own method for coding the data and setting up record fields. When computers became integrated into the daily operations of university administration, those independent data recording systems simply transferred over into the electronic databases. If a university wanted to transform all of the independently curated data from the separate databases under a central standard, it would require considerable time—a rare commodity in the world of university administrators—and a complete change in university culture to embrace the concept and practice of shared ownership of data. For those reasons, while adoption of advanced technologies to improve efficiencies happened as a rapid pace outside of higher education, within institutions there continued to be a significant lag before newer innovations were adopted (Edirisooriya, 2000).

Innovations were advancing not just in computer capabilities for processing data through web applications—even if they were not applied—but in the way education was being delivered. It was during the 1990s and early 2000s that online learning came into being. This profoundly altered the way faculty and students engaged in the act of teaching and learning. It also created a tidal wave of student data. Online courses were delivered through course management systems and universities were able to collect data on each instructional transaction made on the part of the student as well as the instructor. Never before had the administrative functions of data collection been brought so close to the learning process. Universities were soon able to transform their data-driven decision-making processes—popularized in the 1980s and 1990s—with the vast amounts of student data collected coinciding with the increasing capabilities of the hardware and software used to run "what-if" scenarios. No longer were decisions made based on static data only but on the more complex analysis afforded from integrated systems (Picciano, 2012).

Big Data – Big Impact (2000 - Present). During the first two decades of the 21st century, data has been used by universities to analyze enrollments and re-enrollments, academic performance, financial information relating to student retention, student progress, and academic outcomes. The popularity of social networking and mobile technology have effectively transformed administrative functions into twenty-four hours, seven-days-a-week operations (Picciano, 2012). Data is increasingly being collected on student behaviors like library use, tutoring services, LMS activity, co-curricular activity, and on-campus purchasing with student ID cards. Location data as students connect to the university network via Wi-Fi hotspots around campus is also collected and stored even if there is no immediate application (Hubin et al., 2017). The collection and storage of all of this data on students can be argued as justified when considering that the university has taken on a more parental role than was customary in the past. This extended role is in response to pressures from parents of undergraduates for more services and amenities such as sports complexes, luxury accommodations, state of the art labs and classrooms, and expanded student life services such as counseling and financial services. Due to declining state funding, universities are in greater competition for students than ever before. Universities are also reacting to pressures from governing and accrediting bodies along with society as a whole to be held accountable to the quality of education and results their graduates have when searching for jobs after graduation (Carey, 2015; Rosen, 2011).

Parents, accreditation organizations and other stakeholders have scrutinized universities and begun holding institutions accountable for their outputs—a quality education and successful graduates who are able to find jobs after graduation. To demonstrate their effectiveness in achieving that mission, universities are turning to technology advances in data processing and analytics. The more complex data systems developed in the 2000s allowed for the monitoring of student re-enrollments from term to term, academic performance, and financial information to better understand student retention, progression, and achievement of academic outcomes. The data collected by universities in the early 21st century is more diverse and the systems needed to collect and process this big data are able to allow vast amounts of disaggregated data and the various types of data to be collected at increasingly faster rates while operating at lower costs than ever before. From a business perspective, the powerful capability of these new data systems directly impacts the strategies employed at universities by allowing leaders to analyze more robust data in order to better assess student engagement, predict their success, improve the retention rates at the university (Burrell, 2017), and generally make more effective decisions.

With the advent of big data and the complex systems that process the big data came the establishment of two research communities interested in the educational contexts around the use of big data: Educational Data Mining (EDM) and Learning Analytics and Knowledge (LAK). The goals of both communities are the same: to advance education by improving assessments, examining how problems in education are understood, and driving intervention planning and decision-making. There are differences between them in techniques and methods used by each community. The International Educational Data Mining Society (IEDMS) defines EDM as "an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings and using those methods to better understand students, and the setting which they learn in" (Siemens & Baker, 2012, para.5). The Society for Learning Analytics Research (SoLAR) defines LAK as "the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs" (Siemens & Baker, 2012, para.6). There are distinct differences between these two communities even though they have shared goals. EDM focuses on automated discovery through the use of algorithms in data analytic systems and the primary interest for personalized and adaptive learning is the automation processes of the systems without human intervention. LAK has greater interest on the use of human judgement related to data. Their focus regarding adaptive learning is in how the systems empower instructors and students through information.

The research frameworks used by each community differ as well. EDM employs a reductionistic framework emphasizing the research of specific system components and their relationships while LAK takes a more holistic approach studying the whole system. Although the focus of their research activities is defined differently here, both communities rely on the same data and research skill sets and both necessarily include research using automated discovery and human judgement to some extent (Siemens & Baker, 2012).

Both EDM and LAK communities of practice and research play important roles in the advancement of big data use in education and Siemens and Baker (2012) call for a collaborative relationship between them. Because of their similar goals and knowledge bases and differences in research focus, they are uniquely positioned to increase total knowledge built between the two organizations by working together. This potential collaboration is important to help inform the developers of educational learning analytics systems in order for their products to be based on sound research and algorithms. The collaborative work from EDM and LAK would also help inform practitioners who interpret the analytics for the purpose of decision-making. The two organizations working together could develop a comprehensive guide of best practices for using data analytics in education as the field moves toward data-driven decision-making processes (Siemens & Baker, 2012).

Conclusion

Information has played an important role in American higher education since the 1700s. This role has grown in importance and impact through the centuries as technological improvements have been introduced into colleges and universities. Social and political events outside of higher education have also affected the ways information is used to run institutions and meet the needs of students according to changing societal pressures. Themes related to changes in society impacting higher education appeared in this chapter (e.g. scientific management, growing computing power, and the Internet). The need to improve efficiency of data processing is a main driver behind the adoption of data systems and advanced computer technologies. Data for operations and admissions efficiency has been more readily adopted due to the ease of demonstrating the benefits to administrative work. However, data for use in teaching and learning has been slower to adoption as demonstrating the effectiveness of learning analytics has proven more challenging. Another theme from this chapter was the competition among institutions and the use of data to tell the story promoting the value of higher education at different institutions. Related to this theme, knowing how societal pressures have impacted the formation and growth of higher education in the past can help leaders understand how best to respond to similar pressures in the future. For example, the creation of college rankings resulted in an unintended consequence of colleges and universities competing for students by racing to upscale facilities and faculty rosters at increasingly staggering costs to institutions and passed on to students (Carey, 2015; Rosen, 2011). Insight into this history can help institutional leaders reflect on how best to advance their institutions and student education without getting caught in the fray of competition to the point of unintentionally harming students.

The history of technology can also provide insight for future planning. By analyzing the manner in which technological advancements around student data use were successfully adopted and when they faltered can help current higher education leadership in planning for necessary cultural change in leveraging advanced data analytics systems. For instance, the effectiveness of the efforts to train and educate staff and faculty on how to use computers when brought into classrooms and administrative offices in the 1960s (Bare, 1966) can help inform leaders today faced with the challenges of training staff and faculty on best practices for student data analytics. The historical reasons for the development of siloed departments (Carey, 2015) and their inadvertent hinderance to data sharing between departments due to varying collection and coding techniques (Edirisooriya, 2000) can provide insight into how today's leaders can develop big data systems that benefit all units of an institution while maintaining data security and privacy. As big data and analytics systems continue to expand and evolve, a challenge for colleges and universities will be implementing strategies for efficient and effective student data analytics in order to keep pace with stakeholder expectations and policy demands while also protecting and supporting the development of their students. The next chapter discusses what it means for a college or university to build upon lessons from the past to develop policies and procedures around big data for student data analytics. Four ethical lenses—student autonomy, data privacy, educational equity, and educational value—will provide the framework for guiding principles policy-makers can integrate into student data policies and procedures.

Chapter 3. Ethical Issues of Big Data in Higher Education

"Functionally, oppression is domesticating. To no longer be prey to its force, one must emerge from it and turn upon it. This can be done only by means of the praxis: reflection and action upon the world in order to transform it" (Freire, 2018, p. 51).

Introduction

From the historical analysis in chapter two, we can see that since the turn of the century, higher education has experienced many changes in how institutions operate and how education is delivered and evaluated. Increasing reliance on collection and analysis of student data has allowed universities to better meet the needs of students as cohorts and as individuals. With all of the excitement around the potential of student data analytics, there are reasons to proceed with caution, assessing not just the benefits that could come out of the analytics but also the potential unintended consequences impacting student autonomy, privacy, equity, and educational value.

This chapter provides a review of the ethical use of student data in higher education according to these four values. This analysis provides a background for the guiding principles for each value—presented later in the chapter. These principles can help institutional leaders and policy makers get ahead of the ever-advancing capabilities of data analytic systems in order to define how they want student data to be leveraged and where the threshold lies between its ethical and unethical use. It would be easy to assume that college administrators make decisions with the intention of improving the student experience or at the very least to do no harm. Mentioned in chapter two, Burrell (2017) asserted that with the advent of big data and corresponding new data systems employed by universities, leaders are now able to analyze more robust data in order to better assess student engagement, predict their success, and improve the retention rates. However, in one infamous case, the president of Mount Saint Mary's University, a small, private Catholic university located in Emmitsburg, Maryland, issued a survey to students without informing them how their answers would be used. The president's goal was to improve the university's retention rate and so used the survey answers to drive out those students who were not predicted to succeed. A result of this action showed the institution's retention numbers more favorably in national rankings reports. This case stirred public outcry culminating in the president leaving the university (Ekowo & Palmer, 2016). Most cases of student data use at universities are not so cut-and-dried when assessing for whether decisions made by administrators, staff, and faculty are ethical, though.

Decisions regarding the use of student data that inadvertently harm certain groups or individuals are made in the name of helping to improve the university (Ekowo & Palmer, 2016). With the aim of improving the university from the student level to the organizational level, how can it be judged what decisions are ethical and which are not? What actually is ethical use of data and what would it look like if a university followed ethical principles when considering the abundance of student data collected, stored, and analyzed by increasingly sophisticated systems? These are difficult questions to answer as they depend on how one perceives the goals of higher education in a democratic society in relationship to the value and potential harm caused by the collection and use of student data. The following explores issues of data ethics in higher education relative to four key values: student autonomy, privacy, equity, and education. These four values are most important as they are the foundation of ethical practice in education.

The topic of ethics is enormous and complex so it is important to limit what will be addressed in the following discussion. First, for the purposes of this dissertation, ethics pertains to standards of right and wrong in relation to individual rights, obligations, societal betterment, fairness, and virtues such as honesty and compassion (Velasquez, Andre, Shanks S.J., & Meyer, 2010). Each of these aspects of ethics play a role in this discussion. Second, although pertinent to understanding various ethical viewpoints, different philosophies on education will be explored only to the extent that they provide context for an argument. The following discussion will provide a basic framework to guide ethical policy decisions and action. The intention of this chapter is to identify the ethical issues around student autonomy, privacy, equity, and educational value that should be considered before action is taken by a university to create student data policies and procedures.

Student Autonomy

Developing autonomy is important to the mission of any institution supporting democracy. Autonomous citizens are important because they have a knowledge base with critical inquiry skills that enable them to identify, question, and challenge policies when they detect wrongdoing or injustice and know how to foster necessary change. As such, they engage as active participants in society to shape the future of communities through civic, social, and commercial enterprise. Autonomous citizens ensure that a democracy is truly governed by its people and not displaced by a few people or entities that would wield control for their own self-interests.

Autonomy refers to the extent to which a person is in control of their own experiences and outcomes. A thorough conceptualization of autonomy addresses the components of personal agency and self-efficacy. Personal agency is closely related to autonomy in that it requires self-awareness of one's ability to take ownership and control of their destiny along with the will to do so (Bandura, 2006). Autonomy is also closely related to self-efficacy which is having an awareness of one's ability to make decisions and take action. A person needs to have a sense of self-efficacy along with a willingness to act in order to have autonomy (Tilfarlioglu & Doğan, 2011).

Based on this conception of autonomy, big data poses a threat to a healthy citizenry in a democracy in the following ways. Due in large part to the nature of big data—its volume, variety, velocity, and value—the processes behind its application are not readily transparent to those whose data contribute to the applications thus creating a black box. This situation plays out every day across many arenas. Consider that before big data, individuals had more control over what others knew about them and the information they shared with those they encountered daily. Businesses, schools, and governmental bodies would ask for information and individuals would consciously either accept or decline to provide the information. When shopping at a retail store, products were advertised and placed according to results of basic market research of what inventory showed to be in demand or what the store would hope to influence buyers to purchase. Today, with big data being collected through the many devices people use—for example, cell phones, key cards, voice assistants, even kitchen appliances—it is virtually impossible for someone to go through one day without releasing information to others about their activities, habits, or environment—consider networked thermostats and vacuum cleaners that map a home. This collected data is analyzed to get to know consumers' likes and dislikes and behavior patterns without consumers leaving their homes in order to target marketing efforts to more efficiently influence purchasing behavior. The threat of the black box is that in many cases, people aren't cognizant of when, how, where, or what data of theirs is being harvested. The big data systems used to collect and process all of the data would be inordinately difficult to explain to consumers as it is even difficult for those working with the systems to fully understand all of the intricacies involved. This makes is difficult for the average end-user/data-contributor to question and challenge such systems. This can lead to a situation of continued manipulation of citizens who lack autonomy over what is known about them and how it is used. “By means of manipulation, the dominant elites try to conform the masses to their objectives. And the greater the political immaturity of these people (…) the more easily the latter can be manipulated by those who do not wish to lose their power” (Freire, 2018, p. 147). By harvesting, analyzing, and using data to influence behavior, companies and other institutions are limiting the amount of autonomy individuals have over their lives without those individuals being aware.

Another looming threat to a citizens in a democracy which is a lack of informed consent about their data use. “We can’t hope to prevent the collection or creation of inappropriate or inaccurate databases. But we can ensure that the use of that data by employers, insurers and other decision makers is made clear to us when we are affected by it. Without such notification, we may be stigmatized by secret digital judgments" (Pasquale, 2015b, para. 11). The crux of the problem is that there are ramifications for providing consent for the use of personal data. Though legal standards are met with privacy policies presented to users of applications, often users do not truly understand everything for which they provide their consent. The policies presented are long, written in legal jargon and require sign-off by users at the point at which the user needs to have access to the application—time to process the policy is limited at the moment of need. Thus, consent is often given without truly being informed. From this perspective, one could argue that the soliciting of informed consent is a form of coercion instead of consent. It is tempting to say that individuals who are aware of such practices still have autonomy to stop doing business with a company they don't want collecting their data or to disconnect their networked devices altogether but they would do so at the cost of inconvenience or even complete disablement of service. In many cases, such disconnection would put a person at a significant disadvantage in society not just socially but professionally as well. In academia, students may feel they have no choice but to provide consent in order to access applications and software necessary to complete coursework. In reflection, today there may be limitations on the extent to which a person can exercise autonomy without causing themselves personal, social, professional, or—in the case of students—academic harm.

The development of autonomy among students can be supported or hindered by the actions and policies implemented by higher education institutions. When considering the use of student data to enhance the student experience and boost achievement, a thoughtful approach to arriving at viable decisions should include the mission of the university in the context of what sort of graduates they want to produce. The following discusses several examples of how autonomy can supported and—sometimes simultaneously—threatened by university policies and procedures. To begin, parents of young undergraduate students have expectations that universities will protect and nurture their students academically, emotionally, socially, and physically while they are in the physical care of the institution. Attempts by universities to meet these expectations can be seen through the building of luxury dorms, state or the art fitness facilities, providing student counseling and consultation services, and the wide variety of services and entertainment options available to students (Bowen, 2013; Selingo, 2013). In order to provide for the welfare of the whole person of each student, universities raise a case for increased usage of student data analytics—including predictive analytics—in order to help address student needs before they step onto campus. The paternal role universities have been thrust into seems to support the notion of university as caretaker of its students and makes the argument to collect as much information on their students as possible—the more data that is collected and analyzed, the better able to assist students according to their specific needs. Due to potential insights drawn from student data analytics, universities may be seen as knowing what is best for students and thus adopt practices such as nudging students in particular academic directions based on the results of data analytics if not outrightly directing students down certain paths. Driving such practices is the continued pressure that began during the Industrial Revolution (Thelin, 2011) for higher education institutions to demonstrate their value by increasing graduation rates and shortening the time-to-degree for students (Clayton & Halliday, 2017). However, before administrators jump onto the collect-it-all bandwagon in the name of reporting on their value, the potential impact on the development of student autonomy should be considered.

Many university students go about their lives without thinking of how their activity is tracked or why. Every day, students generate thousands of data points for a university as they move around campus and engage in academic and non-academic activities. Location data from students using the campus Wi-Fi through their mobile devices and computers is collected while student ID card swipes at campus stores, cafeterias, libraries, and other student service offices provide large amounts of data to universities. In data terms, individuals can be seen as clusters of data sets such as demographic, academic, financial, social, and health and wellness to name a few. Some of the information gathered helps decision-makers when planning for the future while other data simply sits in storage for potential use later (Hubin et al., 2017). The accuracy and integrity of the data relates to data security and is considered a technical issue. How much of this data universities have legal access to and how it is used relates to data privacy which is a legal issue (Robinson, 2018). Both aspects of data are generally found in student data policies that students sign as part of their admissions process—whether they pay attention to the data policy they sign is another matter. By legal standards, the university can say they informed the student about the use of their data and the student consented when they signed the document. However, if a student does not take the time to read and understand the policy—most likely written in legal jargon—have they really given their informed consent?

Some may argue that students today are accustomed to a certain lack of privacy due to the operational characteristics of k12 schools and parental involvement in their lives (Burbles & Callister Jr., 2000). Their willingness to share so much of themselves on social media and its pervasiveness in their lives would lead one to wonder if privacy is that much of a concern to students. In an October 2016 online survey of 1,000 undergraduate students—500 4-year college students and 500 2-year college students—participants indicated that they were willing to share more personal information if it meant they would get better services in return. Student had an expectation that their data would be used to improve their experiences before, during, and after college with regard to admissions, financial aid, academics, student life, and career services. In particular, students wanted their data to be used by their colleges or universities to help them achieve academic success, find resources for financial assistance, discover student life organizations that they might enjoy, support them in finding career success through internships and finding career opportunities after graduation. Another interesting finding from this survey was that students indicated that they would gladly share more personal information about themselves such as hobbies, social media likes, shoe sizes, and pets' names if it would mean they'd have a greater chance of admission acceptance (Ellucian, 2016).

Although students may be willing to share such personal information in order to obtain better service, it should not translate into meaning that universities have free reign over the use of their data. Students supply certain information willingly with expectations of a specific use and return that will benefit them. They provide their consent to data use based on the information provided by the institution on how their data will be used and in a manner that is understandable by the students. Slade and Prinsloo (2013) accuse universities of not informing students of how their data is collected and used. Providing policies written in legal jargon that students cannot easily understand is not the same as informing them. In their research, Ekowo and Palmer (2016) discovered that students want their universities to be transparent about how they use student data so they—the students—are aware of what they are agreeing to when signing consent forms. This practice of obtaining informed consent is necessary to support student autonomy. If informed consent is obtained for a specific purpose yet the data is then used for alternative purposes without consent, student trust in the university suffers along with their feelings of control over the use of their data resulting in a breakdown in their sense of autonomy.

Fostering autonomy in students can be difficult when providing academic advising and learning experiences that make use of data analytics. For example, analytics can be used to run adaptive learning programs helping to focus students on areas of the curriculum where they demonstrate a need for extra practice while moving on from areas where they show proficiency. Using student data for adaptive learning purposes supports liberal educational ideals in that each student receives up-to-the-moment individualized instruction but it also runs counter to those ideals in that student choice of their learning experiences—the questions and practice exercises or readings and videos to engage with—is limited by the direction dictated by the algorithms running the adaptive learning program. Herold (2017) cites this particular concern as educators espousing a progressive philosophy of education worry that adaptive learning will lower student autonomy in their learning when they are unable to choose which problems to practice or revisit. They believe that this limitation could stifle intrinsic motivation—an important aspect for deep learning to occur. Thus, a consequence of such adaptive learning programs is that students lose control over where they choose to spend their study time as the system makes that decision for them. With the system crunching data and presenting students with the next steps in their learning process, students miss out on the opportunity to do a meta-analysis of their own learning needs and constructing a learning plan for themselves—arguably a valuable skill to develop. Results from a study by Roberts, Howell, Seaman, and Gibson (2016) reported that students do not want learning analytics systems—such as adaptive learning systems—to diminish their autonomy in making decisions about their education or for it to be such a crutch that they become less self-reliant resulting in problems with coping in the workforce where similar systems are not used by employers to track and motivate employees. Although there is justification for being hopeful about future use of adaptive learning, rushing into enterprise adoption of such systems with the belief that it will give students greater control of their learning is premature without an understanding of the effectiveness of these systems as well as the monetary and student costs associated with them (Herold, 2017).

Adaptive learning programs are relatively new in their application in higher education learning management systems (LMS) but targeted interventions for students such as emails from faculty based on learning analytics is not. Advisors, faculty, and administrators have access to student data from the LMS of each course in which a student is enrolled which raises questions on how the learning data is used to boost student success. Further results from the Roberts et al. (2016) study suggested that students were divided when it came to assessing whether learning analytics was more of a problem or benefit to their success. The students saw the potential benefit of receiving personalized services by support personnel like tutors when learning analytics was used to identify struggling students or even providing high achieving students with further personalized resources. Some of the students also saw the potential for increased motivation if they received messages of encouragement similar to those they received on their fitness trackers while at the same time being leery of having their performance compared with other students. They expressed that while some students may find this type of comparison motivating, others may actually be demotivated. Students participating in the study worried that the practice of displaying where students ranked among their peers could divide them through competition—especially among the top tier of students—rather than promote a sense of community among the student body. The potential for continuous messaging from instructors, tutors, and advisors about their performance—whether good or bad—might increase stress felt by students and ultimately do more damage that good. This concern from the students extended to the use of dashboards where they had a constant reminder of their performance every time they logged in to class in the learning management system. The study participants also had concerns that they might be judged based on past student cohorts by instructors and other academic staff. They didn't want their future learning opportunities negatively affected by preconceived notions formed by faculty based on such data.

Academic advising is an area capitalizing on the use of learning analytics and predictive analytics to help students in selecting programs and courses of study. Choosing a major can be a daunting decision for many students and they rely on the assistance of their academic advisor to help them make that choice. With access to a predictive analytics program, advisors have a powerful tool at their disposal to help them provide the best advice possible to students. They can point students in the direction of majors that align with their aptitudes and interests which can help students feel autonomous in their selection of major as the choices presented are manageable to review. If students had to select from the entirety of majors offered at the institution without any indication of which they have an aptitude for, they may feel overwhelmed to the point of indecision or poor decisions. There is risk, however, in the over-reliance on the predictive systems. If students feel that their only choices are those presented by the system without an opportunity to explore options on their own, then their autonomy is effectively stripped from them. To ensure that student autonomy is supported when leveraging their data, it should be done in such a manner that allows students to investigate options the data doesn't automatically steer them toward. Students should understand that they are ultimately the ones to choose the path they want to take.

Advisors must exercise care in how they use the results from predictive analytics along with how they communicate with their advisees. Motivation is an important piece of autonomy and the messages students receive from advisors and faculty intervening when the learning analytics show a decline in performance can serve to either increase or decrease motivation to improve as seen by the results of the Roberts et al. (2016) study. Crafting outreach messages in a positive tone rather than stating outright that they are at risk of failing can have a lasting impact on how a student views themselves and their ability to succeed in the course or program (D. Hooker, personal communication, April 27, 2017; Ekowo & Palmer, 2016). Additionally, targeted messaging to provide individualized support may be construed by students not receiving similar messaging as suspect. Per the Roberts et al. (2016) study, students wondered how they might feel if they knew classmates were receiving additional messages of support from their advisors or instructors that they were not and if the extra intervention would inadvertently give those classmates an unfair academic advantage. In such an instance, although the messaging is not directed at a particular student, the one left out might experience demotivation as a result. This example raises the question of how to ensure educational equity when leveraging student data to improve the student experience which will be discussed later in this chapter.

Throughout the above examples, the issue of transparency and black boxing of systems affecting student learning experiences and decision-making about their academic experiences is evident. A final example to consider in academic advising relates to the use of predictive systems in advising students regarding program enrollment and course selection. If students are not informed about how predictive analytics works during the advising activities described, they may be more apt to succumb to the results of the analytics. Advisees may internalize the power differential between them and their advisor holding the analytic results and thus readily accept the results rather than feel empowered to question the results and make their own decisions. Autonomous students would be informed of where the results came from and allowed to explore all options they have a curiosity toward—whether presented in the analytic results or not.

Principles of Student Autonomy

A university intent on developing autonomous students should present policies that aim to balance two diametric forces: the paternalistic nature often present when providing for the holistic needs of students so they can achieve academic success while also supporting their ability to make their own decisions in carving out a future of their own design. Principles that should guide data policies that support student autonomy include the following:

- Processes and activities leveraging student data are designed and implemented in a manner that strategically supports the development of self-efficacy and agency among students—both needed for autonomy to exist.
- Information about how student data is collected and used by the institution is made transparent and communicated in a manner easily understood by students. Additionally, students are informed of their rights pertaining to their data collection, use, and access in such a way as to support their ability to readily act upon those rights if they so choose.
- Consent to collect and use student data is solicited from students consistently and with full transparency of the rationale along with the implications of consent or

non-consent. Informed consent is obtained whether the data is used for research or other learning or operational interest at the institution.

Privacy

Very much related to the concept of autonomy is the notion of privacy. It is impossible to ensure one’s privacy without one having the autonomy to control what is and is not known about them by another. Beyond personal preference for sharing intimate—or not so intimate—details with others, privacy is important to the wellbeing of a democratic citizenship. The more information others know about a person, the greater their capacity is to manipulate that person’s thoughts, emotions, and behaviors. This threatens the development of autonomous citizens which was argued earlier as being important in a democratic society. Maintaining control over what others know about oneself by exercising one’s right to privacy is a key strategy in maintaining one’s autonomy. Privacy in a democracy should include protection from intrusion by others physically and virtually, the ability to control access to one’s information, and be contextually appropriate based on norms and expectations.

There was once a time when the concept of personal privacy meant that there were aspects of a person’s being and life that others only knew about if that person decided to share with others. Control over who knew what about someone was in the hands of that someone for the most part. Privacy could be maintained by limiting who was invited into a person’s home, what information was shared by that person with others, and by giving expressed consent to businesses to disclose personal information on a case by case basis. Even within a home, family members could seek out privacy by retreating to a space away from the rest of the family. Teenagers notoriously sought private refuge in their bedrooms free of prying eyes.

The introduction and growth of the Internet changed such notions of privacy as nearly everything one does today is tracked in the form of data which is stored and used by those collecting the information. Activities, behaviors, even thoughts and emotions are all gathered as data points through the clicks of online searching and shopping. Data indicating geographic location from global positioning systems (GPS) on mobile devices track one's movement about the planet and multitudes of file uploads and posts to social media provide rich views into the lives of many users. Common household appliances and electronics increasingly have become Internet of Things (IoT) enabled which allow the devices to collect user data in order to provide convenience—consider the refrigerator that keeps a grocery list based on inventory in the appliance and uploads that list to a grocery store which has the order delivered on a certain day. Many people have purchased one or more personal assistants—Amazon's Echo and Google Home to name a couple—to help them keep schedules and lists, answer questions, call friends, and connect to other smart home devices such as thermostats, lights, and home security systems all to make life more convenient by using voice commands. The trade-off is a certain amount of privacy which is being sacrificed—if not fully knowingly—with every connected device and app that is used.

What is meant by saying that the sacrificing of personal privacy is not knowingly done is that, in many instances, people give up their privacy for convenience without reflective thought. The convenience they receive is so enticing to busy people and families that thinking about the implications of sharing their personal data with a service designed to make their lives easier does not happen. If it does, the convenience factor outweighs the value of privacy. For instance, customers made aware of the data being collected and why may often be still be willing to trade their data for the customized experience and the conveniences offered by businesses such as holding credit card information to auto-fill a payment page the next time an order on a website is placed (Burbles & Callister Jr., 2000). Another hinderance to knowing the privacy implications of technology use is the method for which user consent is sought by suppliers of applications and technology. For each device and application, a person uses, a privacy policy and, in some cases, a consent form is supplied for the user to either outrightly accept or deny—a denial resulting in not being able to use the service—or the privacy policy simply states that by using the application or device the user agrees to the policy. Privacy policies generally are lengthy and written in legal jargon that many users either don't understand and/or don't have time to read through. Thus, consent is usually given with little understanding of what the consent actually means. Adding another layer of complexity to the issue is the sharing of user data with third party applications by the provider the user gave consent to initially. In many cases, the policy around the sharing of data is buried within the privacy policy where it won't be noticed by most people.

According to Nissenbaum (2004), there are three principles of privacy that have traditionally guided judgement of whether privacy has been violated or not:

1. "Protecting the privacy of individuals against intrusive government agents" which is protected in several amendments of the United States Constitution. Although

privacy itself is not specifically called out, limiting the power of government over individuals is a goal of the U.S. Constitution and used as justification in legal matters related to privacy infringement by government (pp. 125-126).

2. "Restricting access to intimate, sensitive, or confidential information." Such examples include the Family Educational Rights and Privacy Act (FERPA) and the Health Information Portability and Accountability Act (HIPAA) (pp. 128-129).
3. "Curtailing intrusions into spaces or spheres deemed private or personal." A person's ownership of space in their home is specifically protected through Amendments three and four of the U.S. Constitution Bill of Rights (pp. 129-130).

The three privacy principles relate readily to privacy matters in a traditional sense but advances in technologies such as public surveillance and online communication and data collection add a complexity to privacy the three principles fall short in addressing. To what extent does a person have a reasonable expectation of privacy when they are acting in a public sphere? It is generally accepted in legal circles that when a person steps into a public arena that they do not have a reasonable expectation to privacy; they cannot expect others around them not to notice them and what they do or say. This is how organizations defend the use of advanced technologies in public places—such a facial recognition for surveillance—when they are accused of invading personal privacy (Nissenbaum, 2004).

Privacy has another facet to it in that people tend to expect that the information about them which they agree to share in a particular situation—registering to use an application on their mobile phone for instance—will be used for the purposes of delivering that service and for that use only. This contextual aspect of privacy expectations has become more important as data collection, sharing, and analysis by service providers has increased (Nissenbaum, 2010). Nissenbaum (2004, 2010) describes privacy in terms of contextual integrity in that social norms create the public spheres in which people act. Personal privacy is determined not just by the three principles discussed above but also by what norms say is appropriate information to be shared about a person and also how, when, and to whom the information is distributed. Both appropriateness and information distribution norms must be satisfied for contextual integrity of privacy to be met.

Daniel Solove seems to support the concept of contextual integrity in his taxonomy of privacy—especially when considering the dissemination of information. In the taxonomy, Solove (2006) puts forth what he determines are "four basic groups of harmful activities: (1) information collection, (2) information processing, (3) information dissemination, and (4) invasion. Each of these groups consists of different related subgroups of harmful activities" (p. 488). A person creates data through their activities—most often while engaging with Internet connected technology—and the data is collected. The data is then processed in some manner either known or unknown to the person who produced the data. The data is often then disseminated to others—sometimes unknowingly to the data producer and without their consent. The fourth group of harmful activities relates to intrusive actions—such as hacking and spying—and actions that interfere with decisions—coercion for instance. Intrusive activities do not necessarily require data, but most often do.

Nissenbaum (2004) mentions the three principles of privacy which coincide with what Waldman describes as the "rights-based approach" to privacy which are the building blocks of trust among those sharing data with those collecting and using the data. However, with advances in big data and powerful analytics, policies that address who can access data, how it is stored, and what consent has been given by the data producer does not satisfactorily broach the nuances of how data can be used to put together profiles and used in a myriad of manner that impacts people going about their everyday lives (Waldman, 2018). There may be surveillance data collected through cameras in public areas such as parking lots, city streets, and subway platforms which is considered public and that the public expects and accepts to be used to help keep those areas safe. There may be data that is shared with websites by users agreeing to the "terms of use". In each instance, people have certain expectations of how their data is used and when the data is used for other endeavors outside of the expected uses, a breach of trust occurs. This demonstrates the shortcomings of a rights-based approach to privacy as companies and organizations with big data and advanced technology use data carte blanche for targeted advertising, predictive analytics, and automated decision-making which feels like a breach of privacy for users even though it extends beyond the traditional notions of the right to privacy.

With big data comes a need to approach privacy with a more holistic lens that considers that the use of data outside of the agreed upon terms of use is interpreted as an invasion of privacy to many. Privacy can be breached—or feel breached by users—when smaller bits of data freely given to various websites is aggregated together to create a picture where identities and personal details about a person can be gleaned even when the specifics were never shared by an individual. Another particular characteristic of contemporary notions of privacy is user consent to privacy policies when using a website, application, or service. The idea is that users have control over what they want to share and can opt to not give consent for the collection and use of their data. In policy terms, the user has autonomy over the decision to share their data or not. In reality, many websites, applications, and services are considered needs today and if a user does not accept the terms of use or privacy policy, they will have difficulty engaging fully in society. In reality, users are put in a position of forced acceptance of privacy policies they may not agree with or want to accept just so they can conduct business, schooling, or other personal activities (Waldman, 2018). This is the world many university undergraduates have grown up in. When they became university students, there were certain protections of access to their data such as the Family Education Rights and Privacy Act (FERPA) and institutional data policies, but there was little thought by students about the actual data that is collected and used by the university (Roberts et al., 2016).

There is, of course, institutional value in the protection of student data. It builds student trust in their institution. This trust is important in fostering a relationship that allows the institution to provide support for learning and personal safety to students thus helping students achieve their own educational goals. Imagine the fallout that could occur in the form of student transfers to other universities and fewer new-student applications if a university failed to protect student data. The financial impact could be detrimental to the institution.

There is also institutional value in using student data to protect students themselves. Harking back to the previous discussion on student autonomy, universities must strike a balance between allowing students autonomy in decision making and protecting them in a more traditional paternal role. Certainly there is value in using student data to guide them in directions that have a greater chance for success. This leads to higher retention and graduation rate which is good for the university's ranking as well as beneficial for the students. Using student data to reach out to students who may show indicators of potentially harming themselves or others is another benefit for students as well as the institution. Using and protecting student data to support students with the intention of helping them succeed and become their best selves demonstrates to students that their university cares about them which can go a long way in building trust. Alternatively, there is value to maintaining personal space—for students to have a space where they can be without data being collected about them and where the university will not know or interfere with their activity is valued by students. Private, personal space allows students to relax and let their guard down—to feel free to *be* without surveillance, analysis, or judgement. This is important in that it allows students to recharge and be their authentic selves allowing the development of their autonomy. However, this circles back to the challenge universities have in answering how to allow for personal, private space while also protecting students from poor decisions they may make which put them in harmful situations. That is the balancing act with which universities must contend.

Big data about students provides another value to universities in that it is the fuel enabling learning analytics, predictive analytics and adaptive learning software to run effectively. These systems are complex and require the collaborative efforts between university IT professionals and third-party vendors. FERPA protects the data used behind these systems but due to the increasing amount of data, the speed of its collection and analysis, the complexity of the data, the storage of the data, and the analytic systems, there is heightened concerns around privacy and security of the data. Ensuring that those working with third-party vendors on these systems have legitimate, educational reasons for access to student data adds another layer of awareness and caution for universities. Adaptive learning systems draw extensive critique as many are provided by third-party vendors requiring access to large amounts of student data in order for the systems to function. Such a need provides universities with justification for gathering a variety of data on their students so they have it when needed.

Data privacy has become a hot topic in the consumer market after news of widespread data harvesting and application to create personalized user interfaces on company websites like Facebook and Amazon (Bloomberg, 2018; Diaz, 2020; Weissman, 2018). There is concern that a similar use of student data could happen as universities rush to adopt adaptive learning software before properly vetting their operating structures, the data required for them to run, and whether the benefits to student learning are worth the risk to student autonomy, privacy, and equitable learning experiences (Herold, 2017). The University of California Los Angeles's Chief Privacy Officer has gone so far as to recommend the facilitation of ethical and appropriate use of student data specifically through university data governance structures (Ekowo & Palmer, 2016). If ethical use guidelines were included in institutional data policies, should student awareness be included in the language of those policies and to what extent should the use of student data be divulged to students? Is it even possible to explain all the ways data is used without creating a monster policy document that most likely will not be read by students, thus circling back to the question of whether the student signature on such a document would really indicate them giving informed consent?

Another consideration pertains to the accuracy of data the university collects and uses about their students. If students are not provided with opportunities to inspect, have input, challenge or have their information corrected, universities run the risk of becoming micro black-boxed societies where the activities of the students are monitored and used to make decisions that impact their academic lives without the students being aware (Pasquale, 2015a, 2015b). The goal of universities when using student data should be empowering students in becoming the architects of their own educational experiences and futures so they are active participants in the decision-making process and fully aware of the data and systems assisting them along the way.

Principles of Privacy

There is an obvious tension universities contend with regarding the need to improve institutional efficiencies and effectiveness of programs through the use of student data with the need to respect student data privacy. The tricky question to answer is how to inform students of their privacy rights and implications of informed consent when the nature of privacy today is so complex and situational due to the characteristics of big data. To explain the implications of various privacy choices students may make in different situations would be inexhaustible and impossible to predict with accuracy for each student situation. Additionally, few students—or anyone else for that matter—would have the time or patience to read and comprehend for any practical purpose such an attempt at an explanation. However, this does not mean universities should give up trying to inform students of privacy rights and issues. It's not that students care less today about privacy than they did years ago; it is difficult for students to know how to care about privacy if they don't know the issues around it, though.

For universities to address privacy when crafting student data policies and procedures, the following principles should be included:

- Methods for student data collection and use by the university, along with when and why the data is collected and analyzed, are explained in a concise manner, free of legalese, and written for student consumption.
- Students are informed of when and how they can opt in or out of having their data used by the institution and the immediate implications of their choice so that they are able to give true informed consent if desired.
- The students' right to privacy is of prime importance and is protected by the university through the use of deidentified and anonymous data whenever possible, using identifiable data only when necessary.

Equity

Equity is at the core of the development and maintenance of a healthy democracy. Equity ensures that everyone is able to participate fully in a democracy. A democratic society is at risk when there is equality yet inequitable access or treatment hindering citizens' ability to participate. With an aim toward developing graduates who will engage as democratic citizens, universities have an interest in supporting initiatives and programs that promote educational equity among students. Big data analytics allows for identification of individual educational needs in a manner not before possible thus enabling universities to implement strategies for providing more equitable educational opportunities for all students. In the following discussion equity refers to the fair allocation of resources which does not necessarily mean equal distribution as would be the case if exploring issues of equality (Stone, 2012). Equity in education relates to each student receiving what they need in order to achieve success (Peters, 1969). It is different from equality of opportunity as equal opportunity means that each student has the same opportunity. Looking at an example, each enrolled university student has equal opportunity to take general education courses. However, for students with learning disabilities the actual learning experience in the courses may not be equitable as they will undoubtedly struggle through learning the content more than their peers without learning disabilities. This is why accommodations are made for students registered as needing assistance for learning disabilities. If the accommodations were not provided for those students, their odds of achieving success would be lower than their peers without learning disabilities.

Another frame for equality in education is equality of outcomes which means that each student's measure of success is the same (e.g. everyone graduates with a bachelor's degree) (Curren, 1995). However, some students enroll in higher education with the goal of earning an associate degree and others with the goal of earning a master's degree and so forth. Each individual enrolls as a student with a different goal that determines success for them. Equity is seen as providing individuals what they need to help them be successful. Equity leads to equality of outcomes if the outcome is student success—defined differently by each individual student. For the purposes of this dissertation, student success is measured by the obtaining a degree or credential in an area of study of a student's choosing that is interesting to them and which leads them down a fulfilling—personally, professionally, financially—career path after graduation. (Incidentally, this generally aligns with an institution's definition of student success as high graduation rates improve the raking of an institution and help bring in future student applicants which generate revenue.) Students are the architects of their futures and as such are the ones who define their own measures of success while enrolled at their institution. The institution has the obligation to provide equitable education for each student to achieve what they have defined for themselves as success which does not necessarily mean that each student receives equal treatment. What one student needs to succeed can be quite different than what another needs to succeed given their self-determined definition of success. In sum, equality of opportunity and resources leads to educational inequity. Equity requires that students are treated unequally and is dependent on their specific needs. This is justice in educational opportunity. Treating all students equally, regardless of their unique needs would be an injustice (Peters, 1969).

Data analytics may be used to help identify student needs which can help universities in providing resources to assist students based on identified needs. Some students require financial assistance, others tutoring in specific subjects or emotional support services. The following discusses some of the key issues around data analytics and equity in higher education.

In her article, “Educational Justice and Big Data”, Shahar (2017) examines the impact of information and communications technologies (ICT) on the equitable distribution of educational resources, or distributive justice. ICT in education is defined by Shahar as “the use of electronic devices equipped with interactive platforms and applications that enable students to access learning material, perform educational tasks, and communicate with their teacher and peers and enable teachers to assign these materials and tasks, and evaluate them” (2017, p. 307). One can imagine rightfully that such systems, adaptive learning included, create vast amounts of student and instructor data at very granular levels. Each interaction within the system from performance and time on tasks, search engine queries, and communications in class and group discussions are just some of the data points produced and collected through ICT systems. The intention of collecting, mining, and analyzing such data is to better inform decision making both at the course level and at the institutional policy-making level.

Proponents of ICT argue that it has the potential for solving the issue of distributing scarce resources—high quality instruction—more fairly as ICT can take over tasks to individualize instruction through adaptive learning courseware and provide personalized feedback. Through ICT, more students can have access to high quality faculty teaching than was possible before. The counter argument to ICT solving all the problems of scarcity of quality instruction is that such systems are too costly for some institutions. Thus, inequity of learning experiences remains a challenge (Shahar, 2017).

The interest in ICT—and adaptive learning in particular—has been met with skepticism around the effectiveness and capacities of these systems. This is not unlike the struggles encountered in the mid-20th century when learning technologies started being introduced into classrooms. It was difficult to assess the effectiveness of those systems which slowed their adoption (Bare, 1966). Today, there are few empirical studies that provide evidence that personalized learning through adaptive learning systems actually has any significant positive impact on student learning. Additionally, critics of adaptive learning voice anxiety over the possibility of instructors being replaced by digital technology and data mining systems. Another criticism is that the individual's talents, skills, and learning preferences are discounted entirely by the software running the system (Herold, 2017). For these reasons, further study of adaptive learning programs to better understand how they function to provide personalized direction to an educational experience for each student user is warranted in order to alleviate the potential for educational inequity among classmates.

The allure of adaptive courseware—which delivers individualized learning experiences—is that it allows for each students to learn at their own pace, receiving the coaching and support they need as determined by the system analytics, when they need it. With adaptive courseware, while working through a lesson, each student receives a different learning experience depending on their answers to questions or how long they spend working on a problem; the courseware adjusts subsequent practice problems and supplied resources accordingly. Ideally, each student receives the unique help they need to succeed. Systems like adaptive courseware which rely on algorithms to run predictions and take action on those predictions hold promise while also raising some warning bells around bias and serving non-normative students. Risks to educational equity resulting from student data analytics can arise from bias within the data sets used in predictive analytics—including those used in adaptive courseware—resulting in perpetuation of bias and inequity. As noted in an earlier chapter, this could manifest in adaptive learning environments as the algorithms running the courseware are built with the average student profile in mind. These algorithms are created from vast data sets from previous students which provides the courseware with the typical paths for success to draw from based upon student behaviors in the system. The trouble lies in that those students who do not conform to the norm of what the average student and their track for learning success is can ultimately receive a learning experience that does not provide the same benefit as it does others.

Risks to educational equity also exist within recruitment, enrollment and retention management at higher education institutions. Colleges and universities are in an increasingly critical competition for students every year as state funding diminishes and the need for tuition revenue becomes more important. Since the inception of college rankings reports such as the very popular US News and World Report: Best colleges report, higher education institutions have scrambled every year to receive a favorable ranking to boost student applications (Bowen, 2013; Rosen, 2011). A key factor in receiving a high ranking is showing high retention rates. In order to improve student retention, universities have adjusted their recruiting techniques and enrollment practices to make use of predictive analytics in order to help ensure that incoming students have the greatest chance of completing their degree—yet another key statistic that boosts ranking scores. If one of the goals of a liberal education is equitable access to higher education (Clayton & Halliday, 2017), then some of the uses of predictive analytics described here should be questioned.

Colleges and universities often purchase the results of annual college interest surveys as a starting point to help them target their recruiting efforts (Ekowo & Palmer, 2016). They also turn to big data from previous student cohorts to gain an understanding of the characteristics of students who have enrolled in the past. The University of Oklahoma employed this strategy when it developed predictive models to assess the likelihood that a student would enroll and used the results of the analysis to focus their resources on targeted recruitment of those students who showed the greatest promise of enrolling. The university boasted having the "largest, and most academically prepared student body ever" as a result of using predictive analytics for strategizing recruitment (Mariani, 2018, para. 5). The ethical issue such a tactic brings to bear is that it perpetuates the cycle of exclusion from the opportunity to be considered for admission thus alienating an entire population of students and possibly condemning them to a less-fulfilling life if a fulfilling life entails receiving quality education or at least having the opportunity to receive such an education.

While some may argue that everyone has the same opportunity available to them to apply for admission to a university, that belief can be countered with a different perspective. Consider the argument that there are students in marginalized populations who would not consider applying to attend a university thinking that they are not qualified. They've never been told to consider applying so it hasn't entered their minds or, even worse, they've only heard messages reinforcing a stereotype that people like them or from their neighborhood or social class aren't college material. They may also feel that it is financially out of reach for them to attend a university—a problem for many for over a century (Thelin, 2011). Universities have access to data that help direct recruitment efforts toward students they believe will apply and do well at their institution. Unfortunately, while efficient, this strategy ignores potential in many overlooked populations. Equitable recruitment practices would leverage student data to help recruiters provide underserved populations with the information and messages about applying that would help empower students to see a possible future they would have otherwise dismissed.

Reliance on predictive analytics to target recruiting efforts is attractive since it can help streamline processes and potentially save institutions money. Another tactic employed by some institutions—and mentioned earlier in this chapter—is reaching out to high school students based on their answers to college interest surveys. This tactic can be effective but also lead to students getting excited about, enrolling in and then ultimately discovering that they are mismatched to the university they thought they wanted to attend. Unfortunately, it can be very costly for a student to transfer schools and some may choose to continue down a degree path that doesn't suit them because they can't afford to transfer (Bowen, 2013). Another potential and unfortunate result from targeted recruiting efforts such as those mentioned here is that there will be students who are overlooked in the process who would have excelled at the recruiting university. These students may not indicate interest on the college interest surveys or not apply on their own volition because they don't think they belong at a particular university. If the university doesn't actively reach out to such students, then they could miss out on recruiting potentially excellent academic talent. Making the situation worse is the missed opportunity of those students to better themselves, their future, and ultimately society. For universities that have adopted a mission of developing an educated population that will contribute knowledge for the advancement of society, excluding certain student populations from recruiting efforts because the algorithms behind predictive analytics indicate that it is more cost effective to direct attention to students who are sure to enroll is a disservice to the overlooked student and society. Viewing this issue from another perspective, those students who do not apply to rigorous or prestigious universities because they assume they don't belong when in fact they would do well, miss out on significant life-long earnings potential after graduation and the impact is even greater among minority students than white students (Selingo, 2013).

Compounding the effect is that students who choose a perceived lesser university than a more prestigious one they could have attended have been shown in studies to take a longer time to graduate. The assumption is that attending more selective universities provides closer working relationships with faculty, greater peer pressure to achieve, higher expectations for graduating, and overall greater support resources for students (Bowen, 2013). Further studies actually show that by attending more selective universities the overall likelihood of graduating at all improves, regardless of the time-to-degree (Tough, 2014). All of this combined with the fact that many qualified students who choose not to apply to prestigious universities tend to be from lower socioeconomic backgrounds (Bowen, 2013), it is easy to see how the practice of using predictive analytics to identify only those students most likely to enroll perpetuates a cycle of educational inequity. Although not mentioned in the literature as having been done in the past or as currently being done, one way to avoid the risk of under-educating students would be for recruiters to run predictive analytics to discover those students that would have otherwise been neglected.

Once students are admitted to a university and enroll in classes, the goal of the institution becomes retaining those students from year to year. Strategies university leaders employ to increase retention numbers are diverse and relate to the values of the institutions. The obvious strategy for showing improved retention data is to recruit and enroll more students which has the added benefit of bringing in more revenue to the university. However, not all institutions solve the challenge of boosting retention by enrolling more students. As mentioned earlier, Mount St. Mary's University used predictive analytics to drive out students who were predicted to not succeed at the school (Ekowo & Palmer, 2016). Some universities take a more preemptive approach to securing higher retention rates by culling out those students who would have difficulties paying the high tuition fees and drop out due to financial strain (AtlanticLIVE, 2017).

There are universities that have implemented far different tactics for boosting retention rates by providing programs and services to support students who struggle in their classes. One example is the University of Texas - Austin which analyzes the SAT scores, class ranks, parent educational background among other data points to identify incoming students who have less than a forty percent chance of graduating on time. Instead of pushing those students out of school or leaving them to sink or swim on their own, the university created a leadership program designed to help keep them motivated through tough times and also enrolls them in smaller class sections so they receive the one on one support needed to be successful (Tough, 2014). Programs like those at the University of Texas - Austin cost the institution money and many schools—especially when faced with looming debt and lower state funding—are reluctant to follow suit. The reluctance is seen in wealthier, private schools as well. Many elite Ivy League universities have been known to turn away Pell grant—Federal needs-based grants for undergraduates—recipients due to the institutional costs of running services to support those students (AtlanticLIVE, 2017) which would cut into their budgets for improving the amenities on campus and hiring premier faculty in order to maintain their elite status (Selingo, 2013).

The tactics employed by elite, selective universities to maintain their status, rankings, and wealth discussed above are part of what contributes to the growing gap in socioeconomic status between classes. The more education a person obtains and from more prestigious universities, the better chance that individual will gain greater position, status and wealth over the course of their lifetime. However, some may argue that the more information admissions officers are able to examine about applicants beyond high school transcripts, SAT and ACT scores, the status of the high school they attended, advanced placement or honors courses they took, and any other academic data available, the greater the possibility that education becomes more democratized. If admissions decisions were based on analytics run from not just academic and demographic data—an evaluation system relied upon since the mid-20th century (Thelin, 2011)—but also incorporated data pertaining to the applicant's online shopping habits, social media posts, search engine activity, non-HIPAA protected health indicators, and other online habits to predict the student's potential for success (Shahar, 2017), that may help move the application evaluation process to more fairly consider students who otherwise would be turned down outright. Perhaps it would only solidify the positioning effect of elite universities instead.

The problem of an uneducated or undereducated population is the long-term effect on a democratic society as a whole. Although there is plenty of buzz in the media about the declining value of a Bachelor's degree in the workforce, many jobs today require a degree and there is concern that the United States is short the number of graduates needed to fill those open positions by nearly three million (Rosen, 2011). There are benefits to society as more people gain higher education degrees such as lower unemployment rates—per the three million currently open positions requiring a college degree—and lower poverty rates leading to decreased dependency on social welfare programs, an increased tax base, and overall improved health by living healthier lifestyles (Ikpa, 2016). The value of educating the masses was first seen with the establishment of the Morrill Land-Grant Act in the late 1800s and through the Industrial Revolution to train a workforce to power new businesses and strengthen the American economy

(Thelin, 2011). The need for an educated workforce has not declined so it could be argued that universities, no matter what their rankings, have an obligation to educate as many interested students as possible for the betterment of the collective society.

Mentioned earlier was the great interest university leaders have in predictive analytics because of its potential to help students navigate their academic choices and challenges. Problems arise if excitement surrounding the application of predictive analytics leads to an over-reliance on the results and to the continuation of discriminatory practices and stigma. This can happen when the system pulls information from data bases on variables such as race, socioeconomic status, gender and others to devise selective interventions or to target recruiting efforts toward certain students. In other words, "predictive tools can (…) produce discriminatory results because they include demographic data that can mirror past discrimination included in historical data" (Ekowo & Palmer, 2016, p. 13). If relying heavily on unvetted algorithms pulling from demographic data, predictive models run the risk of further solidifying disparities in college achievement between groups of students as some would be directed toward particular majors based on their race or socioeconomic status which would only perpetuate societal inequities of opportunity (Ekowo & Palmer, 2016). As one of the guiding principles for ethical use of student data that was presented at the 2016 Stanford CAROL and Ithaka S+R conference stated, actions taken by universities based on data analytics should open opportunities for students, not close them (Kurzweil & Stevens, 2018).

In his 2016 article in The Atlantic, Mikhail Zinshteyn broached the subject of universities relying too heavily on historical data of former students in order to make predictions about current students' risks for failure. In the article, Mark Milliron, co-founder of Civitas Learning—an education predictive analytics firm—pointed out that the data collected and analyzed on past students to inform decisions about future students is presented based on the average student from past graduating classes (Zinshteyn, 2016). This is problematic given that historical data can include bias. However, it is now possible to gain a more complete picture of all students passing through an institution providing better insight to guide decision-making. Interestingly enough, Zinshteyn did not address the potential to continue making decisions that inadvertently favor one group of students over another if the algorithms used with the data aren't also checked for bias.

University officials look to the vast amount of student data they are able to collect as a means of developing a holistic picture of each student, but the data may be incomplete or in some cases inaccurate. When this happens, misinterpretations can be made—often without being detected—which have lasting repercussions for students. When leaders employ analytic systems, it is with the hope that they will be able to run analyses without bias. Although these analytic systems have capabilities beyond anything that has been seen in the past, they still rely on databases and algorithms that are programmed and maintained by humans making the entire system implicitly susceptible to a certain amount of bias (Slade & Prinsloo, 2013; Williamson, 2016). The data behind the algorithms are also not neutral or objective which is the assumption made by many relying on the data. Results of analytics are reported in visualizations from which meanings are derived. It is important for those using the reports to understand the extent to which the algorithms include biases and political assumptions as these directly impact the results. By understanding the inherent flaws of analytic systems used to inform decision-making, policy-makers and other university leaders are better able to apply human agency in determining the best course of action to improve student learning experiences while also advancing the mission of their institution (Prinsloo, 2017).

Williamson (2016) wrote that the development of predictive algorithms has been problematic in that bias has been inadvertently built into the code libraries—large files of code created by coders and passed around coder networks online to develop software—which lead to results that can lead to continued discriminatory practices. With all of the new student behavioral data that is now available to university decision-makers, it may seem that bias in admissions, program enrollment, and advising would be eradicated. However, big data might actually perpetuate systems of inequity or create new ones. The algorithms that rely on big data include many data points and are built from the code libraries—some created decades ago with old data sets and are being passed around via the Internet between programmers—which may contain biased information. These algorithms generally are not fully understood by many external to the programming field and even those within the programming field. Assumptions are made that data sets have been vetted for quality and by those collecting the data leaving programmers to focus on creating the requested algorithms for specific purposes. Efforts to develop systems to collect big data and store it may not include vetting the data sets for bias. Programmers may not be aware or feel their job is to look for bias in the data they use to create algorithms either. Those using the results of the analytics may unconsciously bring their own biases to bear on how they interpret and report results. Because of this lack of understanding, a virtual black box forms around the data analytics process. Considering this peculiar nature of big data analytics, it is extraordinarily difficult to identify inequities in the analysis that can lead to injustices in educational decision-making (Shahar, 2017). To help curtail the continuation of this proverbial black box around data processes, institutions may consider one of the guidelines for ethical use of student data from the 2016 Stanford CAROL and Ithaka S+R conference that stated that processes and evaluation of the data systems should be clear and transparent even for students (Kurzweil & Stevens, 2018).

It is helpful to remember that algorithmic systems are a continuum of human and algorithmic interactions—they are not run completely by algorithms without any human intervention nor are they completely reliant on human intervention to function. As such, unintended consequences are unavoidable (Prinsloo, 2017). Big data and the use of algorithms to help organizations use the data to make decisions is here to stay so trying to circumvent unintended consequences by avoiding the use of such systems is not a viable option nor is it an ethical option considering the potential benefits these systems can provide students. Another of the guiding principles for ethical use of student data presented at the 2016 Stanford CAROL and Ithaka S+R conference pronounced that universities had an obligation to analyze student data in order to improve the learning environment and advance general knowledge (Kurzweil & Stevens, 2018). Having a tool with the capability to help students yet not using it is no better than deliberately sabotaging their progress. For that reason, leaders of universities using student data analytics must continue pressing forward while keeping a vigilant eye on how it is used. Latour (2012) likens data analytic systems that leverage algorithms to children. Once a child is born, there is no turning back. A child will do things the parent wants of them but will also eventually do things the parent dislikes. When that happens, a parent cannot just walk away. The parent must deal with the child's behavior one way or another. Such is the case with data analytics systems and those making decisions about their use.

Principles of Equity

While universities cannot simply ignore the potential for leveraging big data analytics, there are certain tensions that university leaders and policy-makers must balance when deciding how to write and implement data policies and procedures with regard to educational equity. Institutions may look toward efficiency and effectiveness of efforts in order to boost student recruitment and retention numbers which must be balanced with the provision of various student support resources. These focus areas may appear opposing at first, however, meeting the support needs of students can actually help institutions meet or exceed recruitment and retention goals. A significant tension lies with the cost of providing equitable support strategies. Leveraging student data and measuring the effectiveness of strategies is important to ensure a balance between institutional needs while providing the best support for students to succeed.

No matter what the specific strategies an institution employs to support equity through the use of student data, there are a few key ethical principles that should be included in policy-making decisions:

- The aim of student data analytics is to provide equitable support mechanisms so that all students can achieve their academic goals.
- Data sets, code libraries, and the algorithms behind student data analytic programs undergo formal audits for bias at consistent and contextually appropriate intervals. Programming and analytic team members at all functional levels are empowered to identify and rectify issues that would result in inequity among students.
- Training around methods for interpretation of analytic results that mitigates bias and inequities are required of all faculty and staff who leverage student data analytics.

Educational Value

There is an educational value student data and the policies and processes around student data analytics hold for students and universities alike. The driving philosophy behind ethical arguments for educational value presented here is that of Paulo Freire, specifically from *Pedagogy of the Oppressed.* This seems an appropriate time to explain the rationale for analyzing ethical issues through a Freirian perspective for this dissertation. Freire's *Pedagogy of the Oppressed* provides an interesting perspective for analyzing the power structures that exist around data production, ownership, and use within a university and society at large. Brynjolfsson and McAffee (2014) predicted the impending widening of the income gap as those who own technology and data will gain much economic and social influence while those who do not will struggle. Expounding on this prediction, as big data becomes more of a factor in everyday life with data owners having the potential to influence the thoughts and behaviors of each individual data producer, it leads a democratic society down a path of great division between those who have ownership and control and those who are blindly led by them—in other words, the oppressive, ruling elite and the oppressed. In an effort to mitigate such a dismal fate, this dissertation attempts to demonstrate how the integration of Freirian concepts in higher education student data policy and procedures can empower students so they are better equipped to engage as knowledgeable citizens in pertinent debates about data use.

It is important to note the context in which *Pedagogy of the Oppressed* was written. It was born, in part, out of Freire's experience living under the oppressive rule of the Brazilian government after the coup d'éte of 1964 (Nelson, Potrac, Groom, & Maskrey, 2016). Freire's focus in this particular work was on a population suffering oppression from their government in a manner that hindered their education and their ability to break free of their socio-economic status and condition so they would continue to be dependent on and in service to their governmental leaders and the wealthy elite in their society. The manner in which education is structured in such a society perpetuates inequalities and servitude to those with social and political power. Only by transforming education to be less paternalistic and more student-focused and student-driven can the oppressed be free to become fully human—Feire's concept of *humanization*—and the oppressors to become free from their own oppressive expectations. "This, then, is the great humanistic and historical task of the oppressed: to liberate themselves and their oppressors as well" (Freire, 2018, p. 44). To be clear, this chapter—nor subsequent chapters—is not discussing students and education in an outwardly, politically oppressive context. This examination is of an application of Freire's pedagogy in the context of student data analytics at universities such as Michigan State University and the University of Texas - Austin—the subjects of the following case study. In this context, questions of who the “oppressed” are and why could they be considered “oppressed” and by whom they are “oppressed” should first be addressed. “Any situation in which ‘A’ objectively exploits ‘B’ or hinders his and her pursuit of self-affirmation as a responsible person is one of oppression” (Freire, 2018, p. 55). In the context of student data analytics and for the purposes of this discussion—and the subsequent case study analysis—undergraduate students in higher education in the United States can be viewed as potentially oppressed. The institutions they attend can be viewed as exhibiting potentially oppressive tendencies. Exploring the relationship between these roles will shed light on why the relationship between students and their university can be viewed as the oppressed within systems that can inadvertently oppress in the context of student data analytics.

How exactly does educational value come in to play with regard to university students and the policies and processes around their data? There is value in students understanding the nature of their data and the role it plays in their educational experience as well as in the operations of their institution. Student understanding about their data enhances their ability to engage in informed decision-making about how student data can and should be used or not used by their university. This is important in developing students who will transfer their experiences engaging in policy discussions and impacting decisions at the university to the greater democratic society. Students who understand how their data is used and the value it holds and who have felt empowered to affect change, will be the ones to support the ethical use of data in a democratic society at large. The alternative is that graduates will embark into their careers and society unaware of how others use their data to influence their behavior and as such will be completely powerless to liberate themselves from such covert oppression. This circles back to the importance of student autonomy, equity, and privacy discussed earlier as they support the educational value of student data—each component building upon the others, allowing for students to learn and be empowered to critically engage in data policy debates.

This is a value-add for the university as well as it leads to balanced and just policy-making when the student voice is an integral part of the process. It can be helpful to think of educational value along a continuum. On one end of the educational continuum, student data policy and processes are completely hidden from students while students go about their daily lives unaware of the data they create and how it is used by the university. At the other end of the continuum, students not only have a comprehension of the data they create and the student data policies at their institution, they are also actively engaged and are important partners in the policy-making process. Universities would be providing an important service for their students if they offered an inquiry-based and participatory education about their data and how it is used by the university as well as in the greater society as discussed earlier. As university leadership considers the most effective strategy to raise the educational value of student data analytics, there will likely be questions that arise about what the value-add is for students to know about data policies and what manner of policy engagement would have the greatest educational benefit for students. The following discussion explores the philosophy and application of Paulo Freire's *Pedagogy of the Oppressed* to help answer these questions.

Looking at how students are dehumanized when considering the ways in which their data is used by universities they attend is where this discussion begins. The power differential between university faculty and students is one example. The students feel ignorant on some level—this tends to be more pronounced in first and second year students than third and fourth year or graduate students—and as such defer to the professor who they believe has the knowledge and urge the professor to tell them what is true to save time and critical thinking on their own part. They do not feel they have the knowledge and background to come to answers on their own or together as cohorts (Freire, 2018). "The oppressed, who have adapted to the structure of domination in which they are immersed, and have become resigned to it, are inhibited from waging the struggle for freedom so long as they feel incapable of running the risks it requires" (Freire, 2018, p. 47). In the case of academic advising where advisors leverage predictive analytics to help guide students down paths where the data indicates they'd experience greater success, students can feel particularly vulnerable to being steered in directions by advisors they feel must know more about what they should do than they, themselves, would. Appreciating the tendency toward this power differential, advisors must practice strategies of engaging with their advisees that support their awakening to how their data is used and their autonomy to use the analytic results as they see fit.

If the oppressed—students in this case—remain unaware of their oppression, they continue down a fatalistic path of acceptance of their situation (Freire, 2018). For instance, as long as students are unaware of how their personal data is used by their university, they accept the use of it as a given and do not resist. “The more completely they accept the passive role imposed on them, the more they tend simply to adopt to the world as it is and to the fragmented view of reality deposited in them” (Freire, 2018, p. 73). Breaking that cycle of blind acquiescence happens slowly and in small steps. This awakening in students is important as it is through their realization of having passively accepted the status quo as a given that real change can occur. With that, it would be understandable for university leadership to attempt initiating an awakening in students through didactic strategies meant to educate students in the use of personal data at the university and in society in general. However, Freire (2018) explains that if the oppressive system gives liberation to the oppressed, this is freedom without liberation as it did not come from the praxis of self-reflection, awakening to an understanding of the world in all dimension—economic, political, social—and action from the oppressed to help set conditions for producing a new life. These three goals are necessary for someone to lead a truly self-managed life—to be autonomous. Thus, the oppressed may be free but not liberated as they still view themselves as less-than compared to the former power-holder—the institution in this case—thus they are still dehumanized by their own selves as they continue to internalize the oppressive forces and feel dehumanized by the power system because they were the ones to grant what they saw as liberation—in actuality, it was freedom without liberation. “The generosity of the oppressors is nourished by an unjust order, which must be maintained in order to justify that generosity” (Freire, 2018, p. 60). In the case of a university, if strategies were employed to teach students about how the institution and society use their personal data and gave them a voice in some policy discussions—those that leadership allowed students to participate in—then the students would experience a certain amount of freedom to engage but not necessarily liberation. Liberation would require students to have first reflected upon the use of their data, awakened to the potential consequences of their data being used by others and the consequences of their own ignorance, and have initiated action on their own behalf resulting from their self-reflection and awakening. If university leadership controls what is taught to students about their personal data and how it is used while controlling when and how students are able to engage in conversations impacting policy decisions, the students are still oppressed. Students should be a part of reimagining and designing how they engage with universities concerning their data. Supporting this idea is another guiding principle for the ethical use of student data presented at the 2016 Stanford CAROL Ithaka S+R conference: the ownership of data should be shared among all stakeholders of the data including the creators, managers, stewards, and users (Kurzweil & Stevens, 2018). Prinsloo (2017) advises that students should be kept informed and involved in decisions around the use of algorithmic systems in learning. He urges decision-makers to remain critical of the institutional versus student benefits and the power relationships involved around student data use while also not avoiding the use of information from the analytics. It is also important to articulate the provisions made to ensure accountability for ethical practices. If, during planning sessions, the administrators and faculty control the conversation—assuming that students are brought to the table to have their voices heard—students will defer to the administrators and faculty as they are seen as knowing more about the issues being discussed than the students. The dialog university leadership have with students must be authentic, trusting them as equals, for liberation to occur.

The awakening piece of the discussion to this point is what Freire referred to as conscientization—the rise of awareness of what is happening and how. It is the process of gaining the "necessary critical thinking tools so that students, instead of internalizing their oppression, understand how institutions of power work to deny them equality of treatment, access, and justice" (Macedo, 2018, p. 17). Literacy was to be used to help adult learners reach conscientization. Freire of course was referring to reading literacy; however, this concept can be extrapolated to data literacy today. As mentioned earlier, the role of public universities is assumed to be preparing students for successful careers while developing them as critical thinkers under an academic system promoting diversity, individuality and freedom (Clayton & Halliday, 2017; Pasquerella, 2019). To do this in today's world requires data literacy in addition to traditional reading literacy. Achieving conscientization means that the oppressed are able to express in their own language what is happening and how. As universities are increasingly turning to big data to inform decision making and policies that affect students, less of the inner workings around the use of student data is being understood by the producers of the data—students in this case—and those making policy regarding data use. The data processes using big data have become black boxed (Pasquale, 2015a). This black box system lies squarely at the far end of the educational value continuum where policy and processes are hidden from students. Perhaps this black boxing is an inadvertent result of university leaders not knowing where to begin when communicating to students about how the vast amounts of data collected about them are being used. Perhaps it boils down to a paternal instinct to say, "trust us; we know what's best for you and will use your data insofar as it helps you achieve success". Improving efforts toward data literacy is an important step universities can take toward breaking through the black box of data policy and procedures and starting students down the road of liberation from data ignorance.

Principles of Educational Value

Developing policy that promotes the educational value of student data for students may be met with some resistance. The tension stems from the tug of what policy-makers and university leaders want to do and the challenges of actually doing it. One driving force is the desire to allow students the full opportunity to achieve conscientization around student data use and engage in the relevant policy debates and decision-making processes. The opposing force is the pressing question of how an institution would actually achieve this goal. This question may loom larger for large universities than for smaller, private institutions though the tension still exists. Student data policy needs to balance this tug between promoting educational value of student data analytics and operational needs of the institution. A starting point is to develop student data policy that includes the following ethical principles:

- Student data policy and processes around the use of student data support the educational value for students, promoting the humanization of all students.

- Policy decision-making processes include the student voice as an important contributor to the creation of policy and procedures affecting the collection and use of their data.
- Increasing data literacy through the education of students about the collection and use of their data leverages inquiry-based strategies and ongoing dialog between students and university faculty, staff, and administrators.

Conclusion

The many challenges discussed in this chapter may leave higher education leaders and policy-makers wondering how to go about addressing ethical considerations around student autonomy, privacy, equity, and educational value through their student data analytics policies and procedures. Factors related to the ethical and operational tensions discussed in this chapter make each institution unique. Thus, the particular path an institution takes to meet these challenges will vary. Regardless of the differing contexts, each institution developing student data policy can integrate the following principles as an important first step in strategically supporting the development of student autonomy, privacy, equity, and educational value:

- Processes and activities leveraging student data are designed and implemented in a manner that strategically supports the development of self-efficacy and agency among students.
- Information about how student data is collected and used by the institution is made transparent and communicated in a manner easily understood by students. Additionally, students are informed of their rights pertaining to their data

collection, use, and access in such a way as to support their ability to readily act upon those rights if they so choose.

- Consent to collect and use student data is solicited from students consistently and with full transparency of the rationale along with the implications of consent or non-consent. Informed consent is obtained whether the data is used for research or other learning or operational interest at the institution.
- Methods for student data collection and used by the university, along with when and why the data is collected and analyzed, is explained in a concise manner, free of legalese, and meant for student consumption.
- Students are informed of when and how they can opt in or out of having their data used by the institution and the immediate implications so that they are able to give true informed consent if desired.
- The students' right to privacy is of prime importance and is protected by the university through the use of deidentified and anonymous data whenever possible, using identifiable data only when necessary.
- The aim of student data analytics is to provide equitable support mechanisms so that all students can achieve their academic goals.
- Data sets, code libraries, and the algorithms behind student data analytic programs undergo formal audits for bias at consistent and contextually appropriate intervals. Programming and analytic team members at all functional levels are empowered to identify and rectify issues that would result in inequity among students.

- Training around methods for interpretation of analytic results that mitigates bias and inequities are required of all staff and faculty who leverage student data analytics.
- Student data policy and processes around the use of student data support the educational value for students, promoting the humanization of all students.
- Policy decision-making processes include the student voice as an important contributor to the creation of policy and procedures affecting the collection and use of their data.
- Increasing data literacy through the education of students about the collection and use of their data leverages inquiry-based strategies and ongoing dialog between students and university faculty, staff, and administrators.

Looking at an example of what the inclusion or exclusion of these four values looks like within university student data policies and procedures is a helpful next step. The following case study analyzes the policies and procedural documents and notices related to the use of student data analytics at two universities to see how the four values of a liberal education that supports a democratic society have been addressed and where gaps exist.

Chapter 4: Design and Methodology

Overview

The previous chapters provided background information about the historical use of student data by higher education institutions and why increasingly bigger data—while promising in its potential—poses a threat to student autonomy, privacy, equity, and educational value if left unaddressed. With that, an exploration into the policies and procedures universities set in place around the use of student data analytics was appropriate to see examples of how possible threats to these four values were being addressed. Guiding the exploratory research was the overarching question: How are institutions of higher education writing institutional data policies and procedures that address the ethical complexities of student data analytics in an era of big data in order to protect the institution and its students from potential unintended consequences? By looking into actual university policies and procedures, assessment of the trajectory of future research that could be valuable in testing hypotheses around ethics and student data analytics in higher education can be done (Dudovskiy, n.d.; Reiter, 2017). Because relatively little research has been conducted in this area, a logical strategy to begin with was an exploration of current policies and procedures to see how ethics is addressed or present with regard to student data analytics.

Research Design

The study was done as a structured comparative case study of current data policies and procedures around the use of student data at two public universities. A structured comparative case study is when "the researcher writes general questions that reflect the research objective and that these questions are asked of each case under study to guide and standardize data collection, thereby making systematic comparison and cumulation of the findings of the case possible" (George & Bennett, 2005, p. 67). The rationale for using a structured comparative case study was due to the study being focused only on the policies and procedures each institution had in place to handle student data analytics. Guiding questions were used for analyzing each case so the information gathered was comparable (George & Bennett, 2005). The purpose of this structured comparative case study was to identify how university data policies and operations are addressing potential ethical risks involved with using student data and to what extent institutional policy-makers address ethical issues in student data policies and procedures.

Analysis of each of the two cases was conducted from the perspective of Paulo Freire's *Pedagogy of the Oppressed.* In this sense, this study applied Freirian philosophy to explore the selected cases rather than using the cases to develop, revise, or test a theory (Kaarbo & Beasley, 1999). Questions developed to guide analysis focused on the four key values of a liberal education in a democratic society discussed earlier. Areas within policies and procedural documents and webpages pertaining to the use of student data were explored for how each of the values were supported—directly or indirectly—along with how they were threatened. Evidence of a threatened value could manifest overtly through stated policies or through omission of direct support and within the context with other policies and processes.

The process followed for analyzing the cases is described by Creswell and Poth (2018) as "within-case analysis" to analyze each case separately with the guiding focus questions, identifying themes and thematic gaps within each case. Following the within-case-analysis was a cross-case analysis using the same focus questions to analyze thematic trends and gaps across both cases. (Thematic gaps also surfaced as a result of the literature review in chapter three which included a general analysis of current research, conference proceedings, and webcasts pertaining to student data ethics. These uncovered the ethical issues around the use of student data analytics that are of concern among leaders, researchers, and policy-makers at higher education institutions.) Data collected according to the focus questions was recorded and coded for analysis. The guiding focus questions and coding procedures are discussed later in this chapter.

Sample

Criteria

Selection of the universities included for study, Michigan State University and the University of Texas - Austin, was based on a few key factors. The first factor was comparability of institution size due to the need for cases where big data is a factor in student data analytics. Both selected universities have large undergraduate student populations—as of fall 2018 Michigan State reported over 39,000 students and the University of Texas - Austin reported over 40,000 students (Kowarski, 2019). Another criterion for inclusion in the study was the extent to which information about student data policies and procedures was accessible. Information sources used in the study were official written data policies and content posted within the web pages of various units across campus directly involved with the security, privacy, and operations around student data.

Rationale

Another key factor in case selection was the degree to which I engage in collaborative work with staff and faculty of potentially selected universities during the normal course daily work at The Ohio State University. The goal was to select case universities with whom I did not have—nor would have in the foreseeable near future—a collaborative working relationship. Considering the number of universities to include in the sample, a small number of cases is normal and useful for exploratory research as no generalizations are to be made based on the results. The goal was to broaden current understanding of the intersection of ethics and student data policy and procedures in higher education—an area where little research has been done to date—in order to gain a picture of where future research studies might be most valuable (Reiter, 2017).

Procedure

Data Collection

Data was collected from university website pages and publicly available policy documents. (The reason behind researching based on information publicly available was to explore the transparency with which information about student data analytics was shared at each university in the study. This also explains the rationale for excluding interviews with members of each university community.) Data collection began by exploring the policies and web pages published by the OCIO and other university IT units as this is where much of university data related to student activity on the learning management system (LMS) resides and where users of technology and university networks are supported. A line by line examination of the data privacy statements and policies was done. An exploration of the websites for the Office of the Registrar and undergraduate academic advising at each institution came next. From that point, the websites for each of the colleges at both universities were examined for any outward facing information regarding the use of student data by the individual colleges. The reason for looking into each college was to follow up on statements found in the university data privacy policies and statements that indicated that some colleges may have their own privacy policies in place. Examining each college website would uncover the extent to which colleges adopted their own data privacy policies and to what degree they varied from the university privacy policy. Investigation into where and how learning analytics was specifically discussed was done for each university with varying results. Finally, a look at websites with information about diversity, equity, and inclusion at both universities was important in order to see the extent to which information on how student data was being used to promote equity and how it was shared. Web pages were saved and policy documents downloaded into a qualitative data analysis program—MAXQDA—for later coding and analysis.

Instruments

The instrument for conducting research was the Internet for web searching through university website pages and links to policy documents. The instrument used for document and webpage analysis was MAXQDA—a qualitative data analysis program. All documents and webpages collected from the university websites were uploaded to

MAXQDA and sorted by university for detailed coding and analysis. Microsoft Excel was used to save backup copies of query results from MAXQDA and EndNote X9 was used to collect reference information on each webpage and policy document analyzed in the study.

Coding

A code system was created with two separate coding procedures for data analysis: descriptive and values coding. Descriptive coding was done first to identify issues and items the institutions address specifically in their policies and documentation of procedures regarding the use of student data analytics. Values coding was done according to the values identified as the focus of research—privacy, autonomy, and equity—with an emphasis on teasing out parts of the institutions' policies and procedures that impact student data privacy, autonomy, and educational equity.

Code construction was based on several guiding questions for the research. The following guiding research questions drove the development and organization of the descriptive codes. These descriptive questions were a necessary first step in identifying where certain information was located and how it was presented before addressing the values questions.

- How is information about data policies and procedures related to student data at the university communicated to stakeholders? This question was meant to assess transparency of information which aligns to the values of student autonomy and privacy.

- How do academic programs address student data analytics? Because universities have historically been comprised of siloed departments with their own cultural norms (Bare, 1966; Edirisooriya, 2000), this question assessed whether student data analytics was being leveraged at the individual program level—apart from a university-level practice—in a strategic way. If so, it was meant to uncover how the four values were or were not apparent.
- What is communicated about the use of student data by academic advising, the registrar's office, and IT? These are common units that leverage student data heavily. It was important to analyze each for how they communicate with students as each impacts a student's educational experience greatly.
- In what manner are the benefits of student data analytics to the institution explained compared with the benefits of student data analytics to students? The intent of this question was to evaluate the extent to which the focus was on the benefits to the institution versus the students, uncovering ethical tensions discussed in chapter three.
- Where are student rights explained pertaining to their personal data? The purpose of this question was to arrive at how the values of autonomy and privacy were supported or threatened through the communication—or the lack thereof—to students about their rights.

The values coding schema was determined by the following guiding research questions:

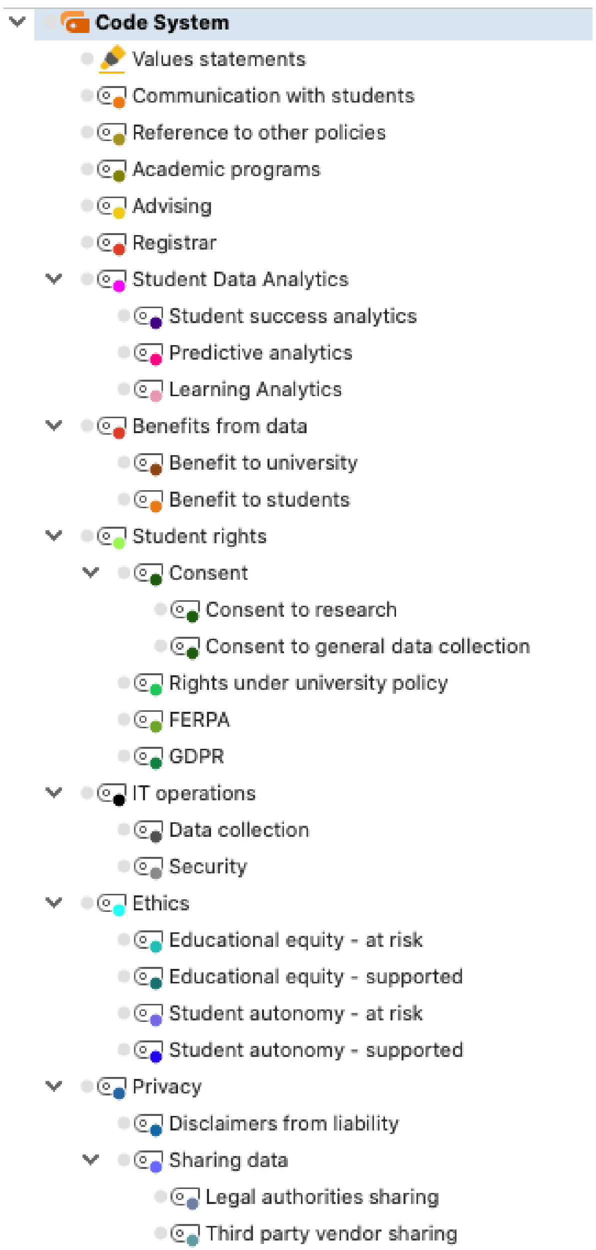
- How are issues affecting student autonomy addressed relevant to the use of student data analytics? With this question, each policy document and procedure uncovered on the university website was analyzed for how autonomy was directly or indirectly addressed or not addressed when it could have been.

- How are issues affecting student data privacy addressed relevant to the use of student data analytics? With this question, each policy document and procedure uncovered on the university website was analyzed for how privacy was directly or indirectly addressed or not addressed when it could have been.
- How are issues affecting educational equity addressed relevant to the use of student data analytics? With this question, each policy document and procedure uncovered on the university website was analyzed for how equity was directly or indirectly addressed or not addressed when it could have been.

Broader codes were parceled into smaller subcodes as warranted. Figure 1 shows the coding system created to code the collected data.

Figure 1

Coding system used in MAXQDA



Note: The colored dots to the left of each code hold no specific meaning but were used to help see various code segments from within each document and webpage recalled in MAXQDA.

Analysis

Document analysis was conducted by university grouping. All collected data was organized into two groups—one for Michigan State University and the other group for University of Texas-Austin. Starting with MSU, data was coded according to the descriptive codes, allowing for coded segment overlap, after which UT-Austin data was similarly coded with descriptive codes. Once the descriptive coding was complete, the coding process was repeated using values coding—first with MSU data then with UT-Austin data. During the process of researching documents, policies, and web pages, notations were inserted for key values-related discussion points to return to during analysis after coding was completed. Those notations were included in the coding scheme and labeled as "Values statements" (see Figure 1).

Once all coding was complete the within-case analysis of the data proceeded beginning with MSU. For each of the guiding questions for the descriptive coding, a query was run within MAXQDA by first selecting all the data elements from MSU while also selecting the codes related to each particular guiding question independently of the others for descriptive coding. The results of the query displayed all the data elements from MSU that included the selected codes and the specific coded segments within the data elements. Coded segments were organized by data elements automatically by MAXQDA. A number at the top of the results screen displayed how many data elements contained the code(s) in the query along with the number of coded segments in each data element. Scrolling down the results screen to view the specific coded segments allowed for analysis of the specific information contained in each segment.

This query process was run for each of the descriptive guiding questions for MSU and UT-Austin. For the codes without subcodes, only one query was necessary. Those codes with subcodes required several different queries to potentially reveal information that would be missed if only looking at the parent code. For example, the parent code *Student Data Analytics* contained three subcodes—*Student Success Analytics, Predictive Analytics, and Learning Analytics*—each needing a separate query run in addition to an overall query for *Student Data Analytics* to help view the entire picture. This type of querying was run for the parent code *Benefits from data* which had two subcodes—*Benefit to university* and *Benefit to students.* It was also run for the parent code *IT operations* which had the subcodes *Data collection* and *Security.* The *IT operations* parent code also needed to have its own query run because, unlike the *Student Data Analytics* and *Benefits from data* parent codes, the *IT operations* parent code had some coded segments of its own indicating there was information about the use of student data pertinent to IT operations that did not fit in either the *Data collection* or *Security* subcodes but also was not defined enough to warrant a third subcode.

The same query process was executed for the values guiding questions for each university. Each parent value code—*Student Rights, Ethics, Privacy* had multiple subcodes. *Consent* was one subcode for *Student rights* and it also contained two subcodes of its own—*consent to research* and *consent to general data collection. Sharing data* was one subcode under the parent code *Privacy* and had two subcodes of its own—*legal* *authorities sharing* and *third-party vendor sharing.* The parent code *Ethics* encompassed subcodes pertaining to educational equity and student autonomy—where each was being supported within the data and where each was at risk. Because each parent value code had coded segments apart from the coded segments in the subcodes, queries were run for parent codes individually then run for each subcode embedded within the parent code. Additionally, queries were run for the entire code package within a parent code (e.g. *Privacy* with each of its embedded subcodes was run as one query). For each of the queries run, an Excel file was saved as a backup in case the MAXQDA program faltered or the files became inaccessible due to potential unforeseen technology glitches.

Conclusion

Through the comparison of the categorical results from and between each coding method, major themes—those that were specifically addressed as well as those that were not—began to surface. The analysis of results is presented in the next chapter. The ethical analysis addressing the gaps uncovered through the case study analysis is discussed in the final chapter.

Chapter 5. Results

Overview

This chapter presents the results from the comparative case study featuring Michigan State University and the University of Texas - Austin. The importance of this case study and its results lies in seeing how these two large universities leverage big student data to support student autonomy, privacy, equity, and educational value. The results also uncover areas within policies and practices that threaten the development of these values among their students and, as such, should be addressed. Results from each within-case analysis are shared by answering the overarching research question—How are institutions of higher education writing institutional data policies and procedures that address the ethical complexities of student data analytics in an era of big data in order to protect the institution and its students from potential unintended consequences?—through addressing more specific guiding research questions. The guiding questions are grouped according to whether they are descriptive in nature or values-based. General themes found through the analysis are reflected upon. Identified gaps where information about student data analytics policy and practice would have been expected but was lacking will be addressed in the ethical analysis in chapter six. The conclusion of each within-case analysis addresses the question of how the educational value of student data analytics is promoted at the university based upon the results of the analysis of policies and procedures. Following the results of each university within-case analysis, the results of the subsequent cross-case analysis of common themes between both universities is discussed. The chapter concludes with a discussion of the limitations of the study along with recommendations for further research.

Within-Case Analysis: Michigan State University

Michigan State University is a public, land-grant university—the first in the country—whose main campus is located in a suburban area of Lansing, Michigan. MSU has an admissions acceptance rate of 78% and the total undergraduate enrollment for autumn 2018 was 39,423. In the U.S. News 2020 Rankings of Best Colleges, MSU ranked number 84 among national universities. Popular majors among undergraduate students include Business, Management, Marketing, Communication, Journalism, Biological and Biomedical Sciences, Social Sciences and Engineering. MSU reports a freshman retention rate of 92% (U.S. & World Report, L.P., 2020a).

Descriptive Inquiry

Question 1. How is information about data policies and procedures related to student data at the university communicated to stakeholders?

To begin answering this question it is helpful to understand the three facets making up the student data analytics in the context of this research. The first facet is student success analytics—analytics for the purpose of helping students complete their degrees and graduate. The second facet is predictive analytics—analytics used to predict outcomes for current or potential students based on algorithms built with data collected from prior students over many years. The third facet of student data analytics is learning analytics—analytics using student data from the learning management system (LMS) or other instructional programs that students engage with during the course learning process to help students improve performance in specific courses or programs of study.

Information about the use of student success analytics was discussed in articles written about MSU's Hub for Innovation in Learning and Technology. Through the Hub, faculty and staff receive personalized support for designing and facilitating their courses using innovative technologies to improve student success. When student success analytics was specifically discussed, the focus was on retention and graduation rates. Predictive analytics was not mentioned through the investigation of MSU web pages and policy statements with the exception of the Student Success Dashboard—a tool used by academic advisors to identify students based on various thresholds. No details about the tool were discussed on the Technology at MSU webpage though a brief description of how the tool is iteratively developed in collaboration with faculty and staff from each college as certain performance thresholds are determined was available along with a statement about the training of advisors in the ethical use of the data and interviewing techniques to support students (Technology at MSU, n.d.-b).

There were two areas of the MSU website where information about learning analytics was discussed: The Hub for Innovation in Learning and Technology and the Technology at MSU News page. One of the ways the Hub serves instructors and staff is by supporting them in their learning analytics endeavors. The focus with learning analytics is to analyze course-level data to develop mechanisms that help students improve performance in specific courses. Part of that process at MSU's Hub is to identify pain points inadvertently built into courses that may hinder students from learning. "As much as it pains me as a teacher to admit, the classroom—my classroom—is sometimes a bridge, sometimes a curb. We can construct with students the forms of mobility they need to pass under bridges and over curbs if we can see them more clearly (Grabill, 2017, para. 10).”

In 2017 MSU began using a tool for learning analytics though the LMS, Desire to Learn (D2L). Course administrators and editors have access to many advanced analytics meant to aid in gaining insight into where and how students are achieving and where they need help. Using the D2L Insights portal, an instructor or college administrator can learn the tools being used in the course by students, how much time they spend in certain areas of the course, how many discussion posts they write, when and how often students log into the course, and conduct quiz item and question analyses. The portal also displays “academically at-risk students filtered by the calculated grade with a selected risk threshold (Halick, 2017, Running a Report section).” Users of the Insights portal are able to download reports for further analysis (Halick, 2017).

MSU communicates with students about how their data is used through the various webpages on the university website. The onus is on the students to look through the pages to find the information about data policies and procedures and their rights under laws such as the Family Education Right and Reporting Act (FERPA). At MSU there are three privacy statements with information directed toward students. The overarching university Privacy Statement explains briefly how information is collected, used, and retained by the university along with the use of links to third-party websites, information about the European Union’s General Data Protection Regulation (GDPR) with regard to research data, and consent. The privacy statement found on the Technology at MSU website is a truncated version of the University Privacy Statement—information about the use and retention of personal data and the GDPR have been omitted. It does contain the same explanation found in the University Privacy Statement that any changes made to the statement will be indicated within the statement itself (i.e. no additional communication to students will be sent when changes to the privacy statement are made). A disclaimer clause is included in the privacy statement on the Technology at MSU site that is not found on the University Privacy Statement. The disclaimer explains that "Neither Michigan State University, nor any of its units, programs, employees, agents, or individual trustees, shall be liable for any improper or incorrect use of information obtained through the use of this Site" (Tech. at MSU, 2015, Disclaimer section; Michigan State University, 2018). There is a separate privacy statement pertaining to the Student Success Dashboard which explains to students the information collected by the system, how it is used, and how the information is protected. The statement indicates that students have some manner of choice in what information they want to be used by the system (Tech. at MSU, 2016).

A characteristic that stood out in the documents and web pages discussing student data analytics in one degree or another was the number of times other state and federal legal acts or policies were referenced and, in many cases, linked from the information on MSU's website. Most of the legal acts and policies provided more details regarding when and how certain information can be shared by the university and with whom (Halick, 2017; Tech. at MSU, 2012, 2015, n.d.-a; Office of the Registrar, n.d.-a, n.d.-b; MSU, 2018, n.d.-b). In the MSU Privacy Statement, links to third-party vendor privacy policies are included—the link to Google's privacy practices along with information on how students can opt out of being tracked by Google Analytics can be found roughly halfway through the document (Tech. at MSU, 2015). Occasionally, links to other university policies are included on web pages such as the link to the MSU Institutional Data Policy found on the Technology at MSU news page announcing the availability of the D2L Insights portal (Halick, 2017). If a student chose to follow the link to the MSU Institutional Data Policy from the article about D2L Insights they'd find themselves presented with more links to other laws, policies, and appendices with additional detailed information within that policy (Tech. at MSU, n.d.-a). MSU shares with student their rights under FERPA and how they can control access to their student directory information through the Office of the Registrar's website. Within the FERPA notification, students are given instructions for filing a grievance regarding the information on file in their student academic records (MSU, n.d.-b).

Question 2. How do academic programs address student data analytics?

The support and maintenance of systems for keeping student records and for academic programs is provided through the Registrar's Office (O. of R., n.d.-a). Individual college websites post a link to the MSU Privacy Statement in a footer on their main page. The footer is the same on every college website providing consistency across all colleges. How student data analytics is leveraged by particular colleges and programs is not publicly accessible through college websites beyond the University Privacy Statement link at the bottom of each web page (Michigan State University, n.d.-a).

Question 3. What is communicated about the use of student data by academic advising, the registrar's office, and IT?

The Registrar's Office website describes the abundance of data it maintains for reporting and support for colleges and academic programs along with advising and IT. Their mission, "steward academic records and data accurately, effectively, and efficiently," is posted prominently on the About Us page. A detailed list of the responsibilities of the Registrar's Office are listed beneath the mission, vision, and values statements (O. of R., n.d.-a). From the main page for the Registrar's Office, users can find information about programs, advising, and student resources for academic records, personal information, and forms students may need for managing records information for a variety of reasons. Under the Programs and Policies menu tab on the Registrar's Office main web page students can access information about FERPA (Michigan State University, n.d.-c).

General information about how student data is used by the academic advising office can be found on the Registrar's Office website. The use of student information for maintaining electronic student academic records for monitoring academic progress, transfer credits, degree audits, progress reporting are listed as ways the advising office uses student data (O. of R., n.d.-a). More detailed information is described regarding the use of the Student Success Dashboard by advisors. Students can learn about the information provided to advisors through the dashboard and how advisors use it to identify opportunities for students, reach out to those who are struggling, and that advisors receive training on the ethical use of student data and strategies for advising with the dashboard (Tech. at MSU, n.d.-b).

Information about the collection and use of student data at the university through IT was found on several webpages and documents. Two main policies governing the use of technology and student data are the MSU Institutional Data Policy and the Acceptable Use Policy for MSU Information Technology Resources. The former, although publicly available, is written for a faculty/staff audience in that it describes procedures for authorizing access and protecting student data (Tech. at MSU, n.d.-a). The latter is written for the broader MSU IT user audience—faculty, staff, students—and focuses on communicating the expectations and responsibilities for use of MSU technology. A notice that the university has the right to "inspect Users' records, accounts, and devices as needed to fulfill its legal obligation to operate and administer any MSU IT resource" is included in the Acceptable Use Policy (Tech. at MSU, 2012, section 6.1.4).

Specifics on the types of student data collected by the university and how they are collected are communicated mainly in the three data privacy statements—the University Privacy Statement, the Technology at MSU Privacy Statement, and the Student Success Dashboard Privacy Statement. Within all three privacy statements, information about the use of cookies and analytics are shared along with the information collected to create log files and how those log files are used to maintain the system (Tech. at MSU, 2015, 2016; MSU, 2018). The Student Success Dashboard privacy statement provides unique information about collection of data from the user interface pertaining to communications through email initiated through the dashboard system itself, advising appointment information, and advising case management information. Student records are not collected by the dashboard, but that information is retrievable through the dashboard system for academic advising (2016).

Data security is addressed within the various privacy statements discussed above though mainly within the institutional data policy. Within the MSU Institutional Data Policy, particulars about various levels of data sensitivity—such as public versus confidential—are explained along with how access is managed. Members of the university community are called upon to take ownership of following best practices to keep Institutional Data secure. Generalities regarding data management are addressed—usage, storage, transfer, disposal. Security provisions and training for those accessing different types of data are explained including access to financial data—specific information provided in Michigan's Identity Theft Protection Act is referenced. More detailed information regarding best practices for data disposal and effective practices for information security is also provided in appendices to the document (Tech. at MSU, n.d.-a).

Question 4. In what manner are the benefits of student data analytics to the institution explained compared with the benefits of student data analytics to students?

Messaging regarding the benefits to the university derived from the use of student data for analytics is nearly twice that of the direct benefits for the students. The web pages and privacy statement for the Student Success Dashboard discusses the efficiencies to the advising process, improvements to programs and services at the university, and monitoring the integrity of the dashboard (Tech. at MSU, 2016, n.d.-b). Other sites discussing student success analytics express the benefits of enhanced reporting and visualizations through advanced learning analytics (Halick, 2017).

At the broader, institutional level, the emphasis appears to be on benefits for governance processes along with external and internal reporting (Tech. at MSU, n.d.-a; MSU, n.d.-b). There are brief mentions of direct student benefits from student data analytics in the University Privacy Statement and the information announcing the availability of the D2L Insights portal however, most of the messaging around direct student benefits appears in the information about the Student Success Dashboard for advising—the ease of scheduling appointments, helping them make informed decisions, improving student retention and graduations rates, and reducing time to graduation (Tech. at MSU, n.d.-b).

Question 5. Where are student rights explained pertaining to their personal data?

When explaining the rights of students on various university web pages, FERPA is referred to often and most notably on the Registrar's Office website. A specific menu option to retrieve FERPA information is provided off the main webpage and includes links to a PDF of the FERPA regulation, information for students along with parents and links to related materials about the rights of students (O. of R., n.d.-c). What is not posted on the Registrar Office website or in any of the linked policy documents is information regarding the European Union's General Data Protection Regulation (GDPR) as it might pertain to data relevant to students in or from the European Union (MSU, n.d.-c).

Much is written about student rights to control access to their directory information (MSU, n.d.-b) and consent for use of university collected data through the LMS, the institutional website, and other university systems (O. of R., n.d.-b; MSU, n.d.-b). Student consent is specifically stated as being required for others to access or the Registrar Office to release student record or directory information (O. of R., n.d.-b; MSU, n.d.-b). In all three privacy statements explored, each express that only data students knowingly and voluntarily share will be collected by MSU. However, those privacy statements also explain the automatic collection of data through the use of cookies, third-party tracking technologies, and server logs for which consent is given when students use the website (Tech. at MSU, 2015, 2016; MSU, 2018).

Values Inquiry

Question 6. How are issues affecting educational equity addressed relevant to the use of student data analytics?

There are two different ways educational equity is assessed for this study: where equity is supported through established policies and practices and where policies and practices could actually threaten educational equity. Not surprisingly, the relatively few informational pieces addressing actions and policies around student data analytics supporting educational equity were found within news about the use of learning analytics at the MSU Hub for Innovation in Learning and Technology. A discussion was posted on their site about efforts to use learning analytics to (1) identify barriers to success inadvertently designed within the curriculum and course delivery processes and (2) being able to identify struggling students early and provide helpful interventions (Grabill, 2017).

With the release of the D2L Insights portal, a news article described the benefit of faculty being able to analyze quiz questions for fairness although the emphasis appeared to be on the ability to distinguish which questions separate the top students from the rest of the students in a course (Halick, 2017), not necessarily whether bias is built into the question giving some students an unfair advantage over their classmates. Within the same article about the D2L Insights portal was an explanation of how students can be identified as at-risk by the tool—this is done through a calculated grade with a programmed risk threshold (Halick, 2017).

Question 7. How are issues affecting student autonomy addressed relevant to the use of student data analytics?

Comparatively speaking, there were roughly three times as many references to issues related to student autonomy as there were to educational equity among the documents and web pages discussing student data use and analytics. Much of the protected autonomy students have pertains to their right to control who has access to their academic and directory information. There is a difference in how students experience control over their data in each instance. Academic records are protected from being accessed by others—including parents and guardians—unless express permission is granted by the student. Conversely, directory information is considered public and students are required to submit an opt-out request in order to restrict access to their student directory information (O. of R., n.d.-a, n.d.-b; MSU, n.d.-b).

The privacy statements analyzed for this study each explained the use of cookies for improving website performance and the user experience along with Google Analytics for tracking usage of different areas of the website. Student autonomy is supported through the explanation of how students can adjust their cookie settings in their web browser and opt-out of Google Analytics if they choose not to have their data collected for that purpose (Tech. at MSU, 2015, 2016; MSU, 2018). This, however, raises a dichotomous conundrum that hinders student autonomy for some but not all students. There are some students who have the knowledge of how to adjust cookie settings in their web browser and opt-out of Google Analytics, however, there are many students who may not know how to do that and would need instructions provided to them if they wanted to limit the use of cookies and Google Analytics. The MSU Privacy Statement (2018) provides a link to “more information about Google’s privacy practices and information on how to opt out of Google Analytics tracking” of student web browsing activities (Third-party Tracking Technology section). What the student experiences after clicking the provided link is a page on Google’s website displaying many links to various articles and policies such as the GDPR and California Consumer Privacy Act. A student would need to go through four more clicks to finally arrive at instructions on how to adjust their cookie settings in Chrome—providing that’s the browser they use—to limit the reach of Google Analytics. Five clicks in total would be required for a student to finally reach those instructions if they click on the correct links each time. The opportunities for students to get lost among the web pages on the Google site are immense. Students may give up in frustration.

Also posted in the MSU Privacy Statement—and related to the use of cookies by the university—is a notice informing students of the following: "You have the option of disabling or not accepting cookies by changing the preferences on your browser. If you opt to disable cookies, you will still be able to use certain sections of our Site. However, you will not be able to use any parts of this Site that require a login" (MSU, 2018, Cookies section). While at first glance this notice appears to support student autonomy, it is yet another instance where those students who have the technical knowledge of how browser privacy options work and can be adjusted are able to maintain a level of privacy that other students would not—not without assistance from technical support which is not offered in the notice. There is another issue at play here in that if students disable cookies, they will be unable to use any portion of the university website that requires a login—the LMS being the main password protected site students would need to engage in their courses while also limiting students' ability to access library materials online. Although students are being told that they have the option to disable cookies and thus protect their privacy to a certain extent, if they do that, they will struggle to engage as students and thus experience a type of coercion to forgo a level of privacy to which they otherwise would not agree.

Although students have FERPA protecting their rights concerning access to their educational information and are able to control access to their student directory information if they choose, there are limits to the data they can control access to and when. As mentioned earlier, student academic records are not shared with parents or guardians without student consent. The exception to this is when there is a health or safety emergency or where university policies regarding alcohol and illegal substances or other violation of federal, state or local laws is concerned. The university may also disclose student identifiable data to third parties providing services to or conducting research on behalf of the university (O. of R., n.d.-b). In the context of this study, these exceptions are not determined to be risks to student autonomy as they are in place to protect the student body from physical threats to themselves and their peers.

Question 8. How are issues affecting student data privacy addressed relevant to the use of student data analytics?

Privacy is a significant focus in any information shared about the D2L Insights portal, the Student Success Dashboard, and in documents discussing technology use and access of student data. The main crux is how and when student data is shared and not shared with others within and outside of the institution (Halick, 2017; Tech. at MSU, 2012, 2016; O. of R., n.d.-a, n.d.-b). Information contained in privacy statements and on other sites discussing student data use follow privacy laws and regulations such as FERPA and, in some cases, reference the law itself. Details are shared about the right of the university to share student identifiable data as required for university operations, assisting students with academic records needed by university administration and financial aid (Tech. at MSU, 2015; O. of R., n.d.-b; MSU, n.d.-b). Sharing of information with external vendors is considered necessary for the proper operations of university services—an example would be sharing data with the Educational Advisory Board (EAB), the vendor partner who developed the predictive model that powers the Student Success Dashboard from MSU student data (Tech. at MSU, n.d.-b). MSU acknowledges their obligation to provide personally identifiable information (PII) under certain requests from legal authorities (Tech. at MSU, 2015; MSU, 2018, n.d.-b) and references the Michigan Freedom of Information Act (MFIA) as the basis for sharing under such circumstances (Tech. at MSU, 2012).

There are a few disclaimers noted within all three of the privacy statements analyzed in this study. There is one line stating that other university units and websites may have their own, separate privacy policies in accordance with their individual needs (Tech. at MSU, 2015, 2016; MSU, 2018). In further analysis of specific college websites, the MSU Privacy Statement was consistently referenced without alterations except on the Technology at MSU website as discussed earlier (Tech. at MSU, 2015; MSU, n.d.-a). MSU does not assume liability for the practices on third-party websites. Additionally, "neither Michigan State University, nor any of its units, programs, employees, agents, or individual trustees, shall be liable for any improper or incorrect use of information obtained through the use of this Site" (Tech. at MSU, 2015, Disclaimer section, 2016, Disclaimer section).

Question 9. From the results of the analysis of coded segments of webpages, documents, and forms, how is the educational value of student data analytics being promoted at the university?

Educational value pertains to the benefit students and the institution gain from students learning about how their data is collected and used by the university and engaging in conversations regarding student data policy and policy decisions. Efforts are being made by Michigan State University to educate students on how their data is collected, used, and protected through the Technology at MSU website and in the University Privacy Statement. Explanations on what students can do to limit access to their directory information and their rights under FERPA can be found through the Registrar Office website (MSU, n.d.-c). However, although FERPA information is easily found on the Registrar Office website, information about how to opt out of sharing directory information requires students to dig deep through the long table of contents of the Academic Programs Catalog—provided they know to look there—for the link to Michigan State University Access to Student Information then scroll through an abundance of small text to find one line under the “Definitions” heading informing students that they can change access by visiting the StuInfo website—which is under password protection (O. of R., n.d.-b). Regarding student engagement in student data policy discussions and policy-making, there was no evidence of such discussions and activities occurring at MSU.

Information shared about student success analytics and learning analytics tools is general. Posted information about the D2L Insights portal is directed toward a faculty and staff audience as the tool appears to be designed for use by faculty and staff, not students themselves. Although it is the students’ data feeding the analytics in D2L Insights, students appear to not be informed about how their instructors may be using the tool (Halick, 2017). The webpage providing information about the Student Success Dashboard gives a basic overview of what the tools is, how it is iteratively developed from years of past student data and how it provides efficiencies in academic advising for students and advisors. What is interesting is that only through searching the University website for "predictive analytics" did the information page for the Student Success Dashboard appear. The academic advising homepage provides a link for students to click on if they want to schedule an appointment. The link directs students to a password protected web portal developed by the EAB—which is the entity who developed the predictive model powering the Student Success Dashboard (Neighborhood Student Success Collaborative, n.d.; Tech. at MSU, n.d.-b). As it appears, there is some information explaining how predictive analytics is being used at the university for student success initiatives though that information is not readily available to students.

Themes

During the analysis of policies and practices communicated on the MSU website, a few themes emerged across the various webpages and documents. (1) There is a strong emphasis on communicating privacy protections specific to accessing and sharing of student data. References to legal statutes such as FERPA were included often along with state laws and other university policies pertaining to the privacy and security of student data. Student rights through FERPA were communicated with special note regarding the control students can exercise over the sharing of their directory information. (2) Rationale for the collection of student data focused mainly on the benefit it brought to the university through research, improved efficiency in operations management, and greater academic reporting—each historically significant to universities as they have long fought to demonstrate value to remain competitive in the educational marketplace (Bare, 1966; Thelin, 2011; Trujillo, 2014). (3) Communication about the use of predictive analytics centered on the benefit for the academic advising process rather than from the perspective of how students would benefit. (4) Learning analytics—although highlighted in articles and on the MSU website—is in its infancy through work being done by faculty and staff in partnership with The Hub for Innovation in Learning and Technology and through use of the D2L Insights portal.

Within-Case Analysis: University of Texas - Austin

The University of Texas - Austin is a public university with its campus located in downtown Austin, Texas. UT-Austin has an admissions acceptance rate of 39% and the total undergraduate enrollment for autumn 2018 was 40,804. In the U.S. News 2020 Rankings of Best Colleges, UT-Austin ranked number 48 among national universities. Popular majors among undergraduate students include Business, Marketing, Management, Communication, Journalism, Biological and Biomedical Sciences, Social Sciences and Engineering. UT-Austin boasts a freshman retention rate of 95% (U.S. News & World Report, L.P., 2020b).

Descriptive Inquiry

Question 1. How is information about data policies and procedures related to student data at the university communicated to stakeholders?

As described previously in the discussion about MSU, in the context of this study, student data analytics consists of student success analytics, predictive analytics, and learning analytics. UT-Austin employs several systems for various goals in this area. One program is the 360 Connections program in which all first-year students are required to participate. Data from nearly 8,000 incoming students is used to assign them to small cohorts of 20 or less to help them make quality social connections to help them not feel alone on the large UT-Austin campus.

Advisors have several systems they rely on to assist their advisees at all levels. The Progress to Degree Dashboard is used to see the percentage calculation a student has achieved progressing toward their degree and to reach out to students who are falling behind or off track. Once students have begun their coursework at the university, data is analyzed according to key indicators to identify changes that may occur to indicate a student is getting off track to their target graduation date. An example indicator would be if a student does poorly in a program gateway course. When identified early, academic advisors can reach out to such students to provide guidance toward degree pathways where they might experience greater success (Student Success Initiatives, 2017). Data about student progress toward degree is also used to manage course registration slot times. Those students who are closer to graduating get earlier registration times than others. This provides another incentive for students to keep up progress on their degree path (S.S.I., 2017). Data is also used each term to run course availability analysis in order to determine which courses need to run and ensure there are enough seats available in order to satisfy the needs of students so they can stay on track to graduate on time. Statistical forecasts are run for incoming Freshman classes to ensure that each student gets registered for a full course load before classes start.

UT-Austin has put into place several strategies for identifying at risk students before they even begin their first year through the use of data dashboards. Explained in the UT Diversity & Inclusion Action Plan progress updates for students (2018b), the

Student Program Database (SPD) “provides an online tool for the colleges and schools to manage their rosters of students assigned to success programs and identifies unassigned students so that they can be invited to participate in a program” (3.2 Action Item). Success programs are academic learning communities that approximately 25% of incoming students are invited to participate. More than a decade of historical academic and demographic student data is used in predictive models using 12 key factors to identify students from the incoming Freshman class who may struggle the most to graduate in four years. That information is used to assign them to academic learning communities and identify support programs that would help them achieve success (University of Texas at Austin, 2018b; S.S.I., 2017). Regarding the admissions review process, The Environmental Context Dashboard aids in the holistic review of applications by considering the context of students’ opportunities when evaluating student performance (UT-Austin, 2018b).

Although some information about the Interactive Degree Audit tool that students and advisors can use to track a student’s progress toward degree was shared through a news item on the Provost’s website (S.S.I., 2017), the ultimate authoritative document for students at UT-Austin is the Catalog. Information about undergraduate and graduate education along with the Law School and Medical School is located in the Catalog. Colleges and units can add information to the Catalog. The section of particular interest for this study was the General Information section as it is where policy around technology and student data in various forms and for different purposes is discussed (University of Texas at Austin Office of the Registrar, 2020). Under General Information

Appendix C, information pertaining to educational records—specifically academic records and directory information is shared (UT-Austin, 2020).

The Web Privacy Policy is linked to from every university web application and site. The front page is set up with easily navigable links to different content sections of the policy. Within the policy, students are informed about the use of Google Analytics and their ability to opt-out of having their data collected by the application though they need to follow a link to a Google site in order to do so. However, the link opens the exact webpage students need to download a Chrome plugin that will keep their browsing data from being sent to Google Analytics (IT@UT, 2019). Students do not need to hunt through volumes of information in order to exercise this right if they so choose. Students are also assured that under particular circumstances the University may be required to provide notice to affected individuals or certain governing authorities if a data breach results in disclosure of personal data (IT@UT, 2019)."

The source of much information regarding data protections, rights, responsibilities, and use is the Acceptable Use Policy for University Students. Within this policy document, students can learn about their responsibilities when it comes to using personal and university devices for university business, securing data, along with the privacy expectations they are afforded regarding their emails and files. Tips on how students can protect their information, privacy and respect that of others—including respecting copyright—are shared in a manner students should be able to understand with limited legal jargon. However, this policy is lengthy which can be a significant barrier to student understanding of the policy. The user agreement is signed on an annual basis

(Information Security Office, n.d.-a), assumingly to satisfy the need to document informed consent to the policy by each enrolled student.

Within many of the policy documents and webpages conveying information about student data use were links and references to several federal and state laws, codes, and policies along with other related university policies. In the Information Uses and Security Policy, 11 federal and state statutes and regulations were listed along with 16 supplemental UT-Austin policies, standards, guidelines, procedures, and forms all with active links to the written documents. References to the GDPR were also included in the Web Privacy Policy (UT-Austin, 2020; 2019; I.S.O., n.d.-a, n.d.-b). The plethora of legal and policy references within these particular university policies may do more to protect the university from potential lawsuits than actually inform students of rights or responsibilities.

Question 2. How do academic programs address student data analytics?

The information regarding how academic programs use data is posted in the Web Privacy Policy and refers to the information used for admissions into certain programs. Although the university has guidelines for how long student data is kept by the university after applying, there are some departments that are subject to laws stipulating different requirements for information storage. The departments that fall within that category are not listed (IT@UT, 2019).

Analyzing the homepage of each college at the university, all but one—the School of Law—include a link to the Web Privacy Policy at the bottom of their site. Eleven of the 19 college homepages also include a link to "Site Policies" which opens a page of links to several other university policies. Each college website includes the Web Privacy Policy and Site Policies links in a special page footer. The School of Law does not include a page footer nor any of the seemingly standard policy links that are present on every other college homepage.

Question 3. What is communicated about the use of student data by academic advising, the registrar's office, and IT?

Data collection practices are discussed in several documents and sites. Information about collecting data to identify and support underrepresented student populations is discussed in the University Diversity and Inclusion Plan updates (UT-Austin, 2018b). With regard to academic advising, students are informed in a general sense on how student data analytics is used by academic advisors via different applications. For example, an explanation is provided about the Interactive Degree Audit (IDA) system and how it allows students, advisors, and administrators to view an overall picture of how well the student is on track to graduate in a timely manner and help facilitate conversations between students and advisors. Mentioned earlier, the Progress to Degree Dashboard helps advisors identify struggling students early and reach out to them to get them back on track before they fall too far behind to catch up. In conjunction with the Progress to Degree Dashboard are post-matriculation assessments. The university acknowledges that a student's path to graduation may change during their time at UT-Austin. Periodic post-matriculation assessments are another strategy for academic advisors to identify students who may falter in key gateway courses which can sidetrack their best intentions. Advisors are able to reach out sooner rather than later to such students to help them with alternative choices—getting the students back on the path to success (S.S.I., 2017).

Other sources of information pertaining to student data can be found on the website for the Office of the Registrar whose main function pertaining to student data is maintaining academic records and official transcripts. The Office of the Registrar also shares information about the student data kept and maintained for academic record and the student directory (UT-Austin, 2020). Processes for using student data to improve student retention and graduation rates is shared through the Provost's website (S.S.I., 2017).

Information related to surveillance of student electronic activity is addressed in a few areas. The IT department does not monitor file content or email content of students. However—during the course of normal IT operations of system management—administrators may become aware of file content and need to take appropriate action if the content warrants. This is the case also regarding personal web pages—the university does not monitor them for content but will take necessary action if it becomes aware of unlawful activity posted on personal Web pages (I.S.O., n.d.-a). The Web Privacy Policy—linked to from nearly every web page on the university website—serves to inform students, faculty, and staff about how the university's collection, maintenance, and use of personal information regardless of the purpose of the data (IT@UT, 2019). It is a significant resource of information pertaining to the collection and use of student data is communicated. Students can learn where and how along with what data is collected by the university IT department, how it is collected, and why. The Web Privacy Policy explains that purchase points where students need to input credit card or bank numbers through the university website use encryption to secure the card information. The policy also discusses expectations that users will contact the Chief Information Security Officer if they feel a data breach has occurred. The Web Privacy Policy seems to make an attempt to put the students' minds at ease by explaining how the data collection is not meant to be used for surveillance but rather to assist in maintaining the operations of various technologies used by students (IT@UT, 2019).

Information security measures are explained in great detail in the Information Resources Acceptable Use Policy which must be signed by anyone handling university data—including research data. This policy is 74 pages in length and includes several references and links to related laws and policies along with minute details of how the university secures data and the roles of each entity that handles data. Although available to students through the Information Security Office website, it is not meant for the general student population beyond those who will be working with institutional data and researchers (I.S.O., n.d.-b). The policy pertaining to information security that has the general student body as the target audience is the Acceptable Use Policy for University Students which can also be found on the Information Security Office website (Tech. at MSU, 2012).

Question 4. In what manner are the benefits of student data analytics to the institution explained compared with the benefits of student data analytics to students?

The many benefits to UT-Austin from the collection and use of student data are explained mainly in the Web Privacy Policy. The university aggregates data from various IT sources for statistical reporting and research purposes. The university processes student data to meet contractual obligations to third-party partners and vendors along with processing for admissions purposes. Personal student data is analyzed in aggregate to help ensure the university web runs smoothly.

IT processes that leverage student data to improve website efficiency are done with the student experience in mind. An efficiently operating website means students experience fewer points of frustration during the course of their study activities. For example, the use of cookies on the university webpages is explained to students with the option for students to adjust their cookie settings should they not want to be tracked. One benefit to students of allowing cookies from the university is that they will not need to continually re-enter their password for each web application they use (e.g. library services, the LMS, research sites). Student user data is also analyzed to monitor the popularity of different university web pages and to make changes in order to better serve the students who frequent those pages. (IT@UT, 2019). This is not just a benefit for the functioning of the university website but also serves to help improve the student user experience.

Several student benefits derive from the use of student data analytics for academic advising and outreach purposes. Predictive analytics is used to identify incoming Freshmen who may struggle to graduate within four years. Advisors are able to use this information to reach out to those students and invite them to learning communities—

required for all Freshmen—where they are most likely to receive the extra connection and support they need in order to succeed academically. The Student Program Database, Interactive Degree Audit, and monitoring progress toward degree through the post-matriculation assessment process—all explained earlier—each leverage student data analytics to help connect advisors with students to help keep students on track to succeed academically and graduate on time. Also key in this goal is the course availability analysis conducted in order to identify the demand for courses each term and ensure enough seats are available for students, especially those nearing the end of their degree programs to help them graduate on time (S.S.I., 2017).

Question 5. Where are student rights explained pertaining to their personal data?

Most of the rights held by student with regard to the collection and use of their data are explained in the Web Privacy Policy. There, students are advised of their rights to access their data, have inaccuracies corrected, and, in some cases, have data deleted or stop the processing of the data. Students can expect that data used for university research purposes will not be repurposed by researchers for their own research purposes. Students can expect that they will receive notice if a breach of their data occurs.

Regarding student educational records, students have the right to access those records and challenge inaccuracies. Student rights are protected by FERPA which is described in several policies and UT-Austin web pages where students visit to conduct university business such as accessing their academic records and making changes to their directory information (UT-Austin, 2020). Where FERPA rights are explained, a link to the actual policy is generally provided (IT@UT, 2019; I.S.O., n.d.-b; Texas One Stop, n.d.). The rights for students located in the European Union (EU) when data is collected fall under the General Data Protection Regulation (GDPR). However, the U.S. federal government and the state of Texas have their own laws concerning student data processing and data retention. It is acknowledged that the university may have valid reasons for maintaining student data even though they may conflict with the GDPR. "As a general rule, in cases where Texas or Federal law conflict with the laws of other countries in regard to the processing, use or maintenance of a data subject's personal information, including provisions of the GDPR, the University will treat Texas and Federal law as controlling" (IT@UT, 2019, section 19).

Pertaining specifically to student consent, students are informed that they can provide consent to have data shared by the university such as when requesting academic records be shared with another institution, and that they can rescind their consent at any time. Students are also informed that UT-Austin, being a research institution, sends out many electronic surveys for various research purposes. Assurance is given that personal information gathered from surveys for research will not be used for purposes other than the original intended research (IT@UT, 2019).

Values Inquiry

Question 6. How are issues affecting educational equity addressed relevant to the use of student data analytics?

UT-Austin has applied data analytics to efforts designed to promote educational equity in recent years with the adoption of the University Diversity and Inclusion Action

Plan (UDAIP) in 2017—the culmination of a year of collaborative work between students, faculty, and administrators. The plan outlines goals for increasing diversity among the student body in general as well as in certain program areas notably lacking in diversity (UT-Austin, 2018a). Each year updates are reported regarding progress being made toward each of the goals in the plan. Reports on the strategies the Office of Admissions has implemented to target and reaching out to underrepresented populations in their recruitment efforts were included in recent updates. The use of predictive analytics to identify students who may struggle to graduate has been helpful in providing support for incoming students from underrepresented populations including first-generation college students (UT-Austin, 2018b). Data mining and predictive analytics are used by the university to not only identify students who may struggle to graduate and students from marginalized populations for recruitment and outreach efforts but also to identify barriers to certain programs and courses that may be more problematic for some student populations than others. That information is then used to redesign programming (S.S.I., 2017).

There are two areas found within university policies that pose a threat to educational equity. One is found in the Web Privacy Policy with regard to the option to decline the use of Google Analytics. The link to information about Google Analytics opt-out browser add-on is provided and the subsequent directions from Google seem relatively easy to follow; it does require that a student have some understanding of what a browser add-on is and how to use it once installed. The other thing to note is that the paragraph in the Web Privacy Policy discussing Google Analytics does not mention that one way to opt-out is to use a different browser such as Firefox. Those students with a certain level of knowledge about how browsers function will understand that as an option without it being stated. Another interesting item in the Web Privacy Policy relates to the cookie statement. There, students are provided an explanation as to what cookies are and how they make their browsing experience smoother but there is no description of how they can opt-out of cookies regardless of browser (IT@UT, 2019). Again, for students with a certain level of computer savvy, they have the ability to control what data the university learns about their browsing activities whereas those students without that background knowledge don't have the information or even know that the option exists to control cookies on their devices. This results in those students having no option but to relinquish a certain level of privacy to which they may otherwise object.

The other area discovered among the many listed policy points in the university Catalog that could impact educational equity lies in the sanction, "withholding an official transcript or degree may be imposed upon a student who fails to pay a debt owed to the University. The sanction is lifted when the student has paid the debt" (UT-Austin, 2020, Appendix C: Chapter 9, section 9-203). One can agree that the university has the right to expect payment for providing an education to its students. However, a sanction such as this may present an additional hardship to students who lack the financial resources to complete payment for their courses or materials.

Question 7. How are issues affecting student autonomy addressed relevant to the use of student data analytics?

There are several, nominal means by which students can exercise autonomy and control over their data. Students are able to restrict the directory information that is shared about them and are able to do so through the Texas One Stop website. Additionally, there they can access their academic records and submit requests for changes (Texas One Stop, n.d.). The sharing of student data from and with other institutions is done only when explicitly requested by students (IT@UT, 2019).

Additionally, FERPA information and links to the policy are provided on many university web pages and policy documents available to students (IT@UT, 2019; I.S.O., n.d.-a, n.d.-b; Texas One Stop, n.d.). Students are given the necessary contact information should they want to file a FERPA non-compliance complaint (UT-Austin, 2020). With regard to research efforts by the university, students must give consent within the context of the research study and can withdraw consent without repercussion if desired (IT@UT, 2019).

There are areas where autonomy is at risk, however. Reflecting back on the opt-out option for cookies and Google Analytics, this is an effort to provide students some semblance of autonomy over when data pertaining to their browsing activity is collected and analyzed (IT@UT, 2019). However, hearkening back to the previous discussion about the risk to educational equity, that autonomy can only be enjoyed by students who are able to take advantage of the option. Another area of risk to student autonomy concerns the student directory. Directory information about students is considered public by default. If a student does not want their information public, they must initiate the request to block it from view (UT-Austin, 2020). Although the ability to adjust the privacy of directory information supports student autonomy, it also threatens it for the students who do not realize their information is public and for those students who do not know how to change the sharing status of their information.

Question 8. How are issues affecting student data privacy addressed relevant to the use of student data analytics?

Explained within the Web Privacy Policy is how the university IT custodians use student data analytics to help fine tune the university website and that the data used is aggregated and anonymized. Raw server log data is only shared with custodians of particular websites. These logs are used to help determine popularity and use of different websites so adjustments can be made accordingly. As mentioned previously in this chapter, the Acceptable Use Policy for Students notes that the university does not surveil student email communications and personal websites for content providing students some assurance of privacy (I.S.O., n.d.-a). Regarding research practices, data used for analytics in university research endeavors is “subject to appropriate safeguards, including the use of data minimization and pseudonyms when possible” (IT@UT, 2019, section 6). Personal data used for research is anonymized whenever possible without compromising the purpose of the research (IT@UT, 2019).

Question 9. From the results of analysis of coded segments of webpages, documents, and forms, how is the educational value of student data analytics being promoted at the university?

The plethora of information presented about the collection, use, protection of student data is a significant effort from the university to teach students about their data, its value to them and to the university and how they can and cannot control certain aspects of it according to various federal and state laws and university policies. However, at no point is there evidence of students being involved in discussions related to student data use policies and practices. Students are receiving information with little opportunity to engage in critical conversations related to what they learn about the use of their data by the university or to impact policies and practices. A beacon of hope that this might change in the future is that students were included in the creation of the university Diversity and Inclusion Plan during the 2016-17 academic year—the plan was adopted in 2017 (UT-Austin, 2018a).

In the state of Texas, the Top 10 Percent Law was passed by the state legislature after admissions practices at the University of Texas using race as a factor in acceptance and denial were deemed unconstitutional in 1996. This meant that the top 10% of students from every high school in the state were guaranteed admission to any UT system institution, including the UT - Austin. (Currently the threshold for admission into the more elite UT-Austin is the top 7% from any Texas high school). Students from middle and upper socioeconomic districts had higher SAT and ACT scores and were better prepared for the academic rigors at UT-Austin while students from lower-socioeconomic districts struggled and had higher attrition rates. The lack of a sense of belongingness among those students was identified as a contributing factor for those students leaving the university without a degree. This situation partly accounts for the development of the Student Success Dashboard (Tough, 2014).

With the use of predictive analytics through the Student Success Dashboard, incoming Freshmen who may experience impediments to academic success are identified and placed into learning communities and programs—described earlier—designed to provide them with extra academic support. From the identified students, those who also come from families below a certain income threshold—as data shows that students from lower income households have an even greater attrition rate than their peers—are enrolled in the University Leadership Network (ULN) which meets regularly to help students develop leadership skills. Reported results of this program show an increase in student retention and graduation rates among the participating students. An interesting aspect to this program is that students are not informed about why they receive invitations to participate or why they are enrolled in classes with lower student numbers or in the ULN. The rationale for not informing students about the selection process is that the university does not want participating students to feel singled out from the rest of the incoming student body; they want to assist those students in achieving academic success while also helping them to feel that they belong at the university. The underlying assumption is that if the students know that they were identified for a special program due to their socio-economic background and analytics predicting their potential for failure from several other data elements that it may fuel a self-fulfilling prophecy for the target student population causing them to see themselves as not being good enough and not belonging at the UT-Austin thus leading many to potentially leave the university before graduating (Tough, 2014). As noble as that rationale may be, it does little to promote the educational value of student data analytics. Keeping processes using student data analytics which affect students' educational experience hidden from those students overlooks an opportunity to expand their understanding of the value of their data and how it is used to assist in their academic success.

Themes

Two main themes arose from the analysis of UT-Austin's data policies and practices as posted through their website: (1) the strategic use of predictive analytics in recruitment, admissions, and advising activities with an emphasis on improving diversity and inclusion among the student body and academic programs and (2) transparency of policies and processes related to student data collection and use through detailed communication with students. The leveraging of predictive analytics in order to achieve this goal while subsequently improving educational equity through various tools and strategies highlighted how advanced analytics could be applied. The commitment by the university to improve the diversity of its student body with the creation and adoption of the UDAIP and its annual progress updates posted publicly for anyone to read demonstrated the focus of the university to improve in this area.

It is evident that transparency about how student data is collected, used, and protected is a primary focus of the university. Transparency is manifested in the communication campaign around student data mainly regarding IT processes, including data security, and data privacy. Students are required to sign an extensive acceptable use policy which includes details about what is acceptable and what is not regarding the use of university technology and data resources. The Web Privacy Policy is linked to from nearly every university website including college website homepages. Links to FERPA and other state and federal laws and university policies regarding student data protection and student rights can be found in several websites and policy documents on the UT-Austin website including the Texas One Stop site where students visit frequently to register for classes, request transcripts, and file for financial aid among other things (Texas One Stop, n.d.).

Cross-Case Analysis

Themes

When considering the cases of MSU and UT-Austin together, there are themes of student data analytics that are shared by each university. Evident through the communications in policy documents and webpages on both university websites is a focus on data privacy and security. On both university homepages are links to privacy policies and links to the privacy policies are included in footers on every university department and application homepage. Communication with students regarding the privacy policy, acceptable use of university technology and data resources along with student rights pertaining to their data err on the side of over-loading information in an effort to ensure transparency.

Both universities demonstrate a propensity to include extensive references to federal and state laws along with related university policies. Active links to the referenced laws and policies are found from within the university privacy policies, acceptable use policies, and other sources of information about student data use commonly found through the Registrar Office website. Of the external laws referenced by each university, FERPA is given the greatest attention. Summaries of general principles from FERPA are provided on different student-facing webpages—primarily webpages managed by the IT department or the Registrar Office. Links to the actual FERPA document are included alongside the summaries and also included in any other reference to the collection, storage, and use of student data whether it pertains to the operations of the university website, university research, or student educational records and directory information.

Although there are similarities in general themes around student data analytics and how they are addressed by each institution, there are some noteworthy differences. One difference is where each university places emphasis on the use of predictive analytics and how they leverage the capability. Although information about the use of predictive analytics at MSU is light, what is described focuses on its use for academic advising to help advisors identify struggling students (Tech. at MSU, n.d.-b). At UT-Austin, a great deal of focus regarding the use of predictive analytics is not only on how it is used for academic advising in a similar manner to the strategy at MSU, but also on how various tools and programs have been developed to aid in recruitment of students, student enrollment and placement into various programs designed to support them socially and academically, and to help increase diversity and inclusion within the student body (UT-Austin, 2018b; S.S.I., 2017).

Limitations and Recommendations for Further Study

While a great deal of information was uncovered through this study, being that it was an exploratory study, there was not a theory to be tested nor a theory developed. The purpose of the study was to uncover areas where more focused research would prove valuable to higher education leaders. Because this was a comparative case study for exploratory purposes, a small sample size was used. With only two universities included in the study, generalizations were unable to be declared. While both universities were similar in size, mission, and programming to ensure the institutions were comparable in their access to big data about their students when it came to analyzing the results, the characteristics of student data policies and practices at higher education institutions with different institutional profiles may have significant differences from the two universities included in this study.

The data analyzed for this study consisted solely of qualitative data gathered from analysis of policies and practices shared publicly through and linked to from the MSU and UT-Austin websites. One reason for this was to assess how transparent each university was in communicating their policies and practices pertaining to student data use and analytics. The limitation of this strategy was that it omitted information from policy work that may be underway at the universities but hasn't been formally approved yet. Such work could be interesting to include in analysis to gain an understanding of the future of student data analytics at each university. Future studies could also include interviews with leaders in the various departments dealing most often with student data analytics—the Registrar Office, the IT department, academic advising, admissions—to gain insight into how policies and practices came to be and what plans the institution has for the future of student data analytics. Another research component that could be included in future studies is a student survey regarding their understanding of policies and practices by their institution around the use of their data. University leaders may find such information helpful when planning for the future of student data analytics and communication with students about policies.

Chapter 6 - Discussion and Conclusion

Overview

The results of the comparative case study with Michigan State University and the University of Texas - Austin discussed in chapter five highlight a few themes unique to each and themes that were common to both institutions. The importance of student data in providing support for student educational experiences is evident though strategies for doing so vary between the two universities. This chapter provides an ethical analysis of how each university addresses the four key values discussed throughout this dissertation—autonomy, privacy, equity, and educational value. Before introducing the ethical analysis, I want to take time to emphasize current trends happening with student data that make the following analysis particularly important as university leaders must start addressing potential unintended consequences from the increased use of student data. The problem with not addressing this in policies and procedures now is that—as discussed earlier in this dissertation—technological capabilities may advance so rapidly that policy will not be able to catch up if neglected. In that case, the data and technology capabilities will drive data analytics processes which will, in effect, determine policy and the norms for acceptable practice—whether ethically desirable or not.

One of the greatest things of value students bring to an institution from the time they are admitted to the university to the time they graduate is their data. Their data—related to academic performance, geolocation, swipe-card usage, learning behavior, social activities, and more—all help the university make strategic decisions as an institution. Considering the competition universities have faced since the 1800s for students (Rosen, 2011) and continue to face today (Carey, 2015), having access to large amounts and a variety of student data to analyze for current and future trends can help improve efficiencies in recruitment, aid in student success and retention (Baker, 2010; Picciano, 2012), and provide services determined to be of greatest importance for students (Bowen, 2013; Rosen, 2011; Selingo, 2013).

At times, events in society can cause certain data to become more valuable. For example, the 2020 COVID-19 pandemic revealed more about the value of data as educational institutions became interested in tracking exposure and contraction of the virus among students, faculty, and staff. This information could be gathered through a process of contact tracing and there were a few technology companies—notably Apple and Google—pushing to develop mobile applications that would assist with doing just that. Through the use of students' mobile phones and the use of the downloaded app, a record of whether students had contracted the virus or had been exposed could be kept. As institutions implemented recommendations by the Center for Disease Control (CDC) to establish classroom, dining hall, and other physical space models that promoted social distancing and other protection measures, they could potentially cross-reference contact tracing information through mobile applications with student swipe card data to track virus spread (Vasudevan & Panthagani, 2020). There were many issues concerning privacy and ethics as institutions considered the adoption of such tracking strategies.

The future of data analytics in higher education is quickly approaching, making it imperative for leaders to be continuously mindful of the changes happening in various segments of society and the advances in technology impacting society as what happens outside of higher education always impacts what happens within it. The COVID-19 pandemic was unforeseen and global in its impact yet had real implications for higher education and the collection and use of student health data. There are, however, noticeable data trends happening across different sectors of society that higher education leaders can investigate now for potential influence on the future of higher education. For example, connecting with businesses and governments on how they are adjusting their data privacy policies in the wake of the European Union's General Data Protection Policy (GDPR) will help inform universities on their strategy to accommodate the requirements outlined in the policy themselves.

Another trend on the horizon is the integration of adaptive learning systems in order to make learning activities within an LMS more personalized. This personalized learning will also be accessible anytime, anywhere through mobile devices. Data from those mobile devices, while users are engaged in managed learning environments, will be collected and analyzed for insights on how to improve the learning experience by instructional designers and instructors. Administrators will be better able to understand who their students are—such as their learning habits and study location preferences—based on LMS activity and location tracked through Wi-Fi hotspot usage around campus. This type of data, combined with student ID swipe card data collected at libraries, dining halls, and other student services will paint a fairly detailed picture of each individual student enabling universities to further personalize their educational experience.

Turning attention to the business side of running a university, data-informed decision-making will increase in importance and expectation. Any strategic plan created by university officials will be expected to have been derived from the results of thorough data analytics. Additionally, transparency of the data and processes for decision-making will be imperative as more stakeholders become increasingly aware of data use by universities. Those stakeholders (e.g. students, parents, faculty, staff) will demand to know how and why their data is being used. Methods for securing informed consent from stakeholders to use their data will constantly need revision as big data systems continue to change.

As alluded to above, privacy will continue to be of great concern when it comes to big data. Data-owner rights have already become a primary focus of university IT and OCIOs thanks to policies such as the EU's GDPR and California's Data Privacy Law. Informed consent will play a larger role in student data policy decisions than it has in the past. "CIOs must educate campus leaders and stakeholders to be cognizant of the ethics of data collection with regard to the quality of data, privacy, security and ownership. Current policies, decision making, and data governance structures may not be adequate to address the ethics of big data utilization" (Burrell, 2017, p. 4). The point being made here is that big data about students in higher education is growing every day and the type of information that is valuable to institutions continues to expand as the challenges in society diffuse throughout higher education. Thus, it is increasingly essential for leaders and policymakers to analyze student data policies and practices according to the mission and values of their institution and make course corrections as warranted. The guiding principles presented in chapter three, based on the four values of student autonomy, privacy, equity, and educational value provide a structure for this to occur.

The following analysis references the fundamental role of public universities being to foster a liberal education that supports a democratic society. Such an education stimulates critical thinking among its students (Clayton & Halliday, 2017) while preparing them for the working world by helping them become employable, and supports diversity, individuality, and freedom (Pasquerella, 2019) rather than focusing on mere rote memorization of academic material. The mission and values statements of MSU (Michigan State University, 2008) and UT-Austin (University of Texas at Austin, 2021) provide two examples of how providing a liberal education can serve as the foundation for a university's mission and goals .

The speed of change across all sectors of society has increased dramatically in the last several decades. It is no longer adequate for universities to present content for students to ponder in an academic vacuum without seeing how various specializations work in conjunction to impact society. The growth of new information and ease of access to it means that memorization of facts is no longer as important as it was during the early years of higher education (Thelin, 2011). The speed of innovation and technology development means that how people function in society is fundamentally shifting at a rate unimaginable a just a few years ago. The impact of these shifts on society, culture, and the environment require a critical mindset able to process inquiry holistically. This type of analytic mindset is promoted through a liberal education with an additional focus on the ethical implications of decisions made regarding new innovations. The ethical values of student autonomy, privacy, equity, and the educational value that data analytics holds for students align closely with the mission of a liberal education. As discussed in chapter three, the support of these values is critical for the continued prosperity of a democratic society.

Discussion

To begin this discussion, a brief review of the aforementioned values that were discussed in earlier chapters may be helpful to clarify subsequent meanings. For the purposes of this dissertation, student autonomy refers to the degree of which a student is in control of their own educational experiences and outcomes. It is closely related to the concept of individual agency in that it requires self-awareness of one's ability to take ownership and control of their destiny along with the will to do so (Bandura, 2006). It is related to student self-efficacy—the awareness of their ability to make decisions and take action—in that a student needs to have a sense of self-efficacy along with a willingness to act in order to have autonomy (Tilfarlioglu & Doğan, 2011). As was discussed in chapter three, the concept of equity as it applies to education refers to the fair allocation of resources, not necessarily the equal distribution of resources in order to satisfy desires for equality of educational experiences (Stone, 2012). The notion of privacy and how its meaning has changed through the years was explored in depth in chapter three. However, to begin thinking of privacy in the context of student data, it is helpful to consider the degree of control over who has access to student data and how it is used. Perhaps the lesser known—but nonetheless important—value of the four guiding the framework for the forthcoming ethical analysis is education itself. What is meant by educational value is the extent to which students gain an understanding of the data they produce, why it is valued, how it is used, and how they can be agents of influence over the use of personal data.

This chapter presents models for conducting an ethical analysis of university policies and practices around student data analytics based on the four values mentioned above. Many institutional leaders understand the importance of considering ethics when making policy and procedural decisions around the use of student data analytics but are confounded as to how to actually *do* ethics in the decision-making process. The goal of this chapter is to answer this question by explaining how university student data policymakers can analyze current policies and practices through the lenses of the four key values of student autonomy, privacy, equity, and educational value to uncover opportunities to promote each more strategically when developing policies and procedures around student data analytics. It is important to realize that this process is a journey and will require continual analysis by institutional leadership as data analytics capabilities advance and society changes as a result of increasingly sophisticated technology becoming more engrained in everyday life for students, faculty, and staff.

Chapter five provided results from analyzing the policies and procedures related to student data at both Michigan State University and the University of Texas - Austin. The next step in analysis is to identify gaps in addressing the key values through policies and procedures. For example, when analyzing those from MSU, there a few areas where gaps between ethical ideals and institutional procedures appeared. (1) One gap area relates to the issue of equity. The diversity and inclusion statement does not address if or how student data might be leveraged in advancing this goal. Doing so would be a way student data analytics could help improve educational equity. (2) Another area where greater communication could improve equity is in the MSU Privacy Statement section discussing the use of cookies and Google Analytics. Specifically, if MSU provided access to help resources for students who want to control data collected through cookies and Google Analytics but lacked the technical background to be able to do so on their own, it would allow all students the same opportunity to control this type of data sharing rather than the few with the requisite background knowledge. (3) Beyond legal statutes and policies, the ethical use of student data analytics should be addressed. It is possible to follow policy and law to the letter yet still act in a manner that has harmful consequences for students—unintended as they may be. What is mentioned in the MSU Privacy Statement regarding ethical use of student data is a disclaimer in the university privacy statement that the institution and its staff and faculty are not responsible for the actions of those using student data—including the actions of third-party vendors (MSU, 2018).

Another way of approaching this analysis—particularly in the case of educational values—is by looking at institutional mission statements. Thus, a starting point for analysis is to assess how well the policies and procedures around student data reflect the institutional mission statement. The stated mission of an any institution should be the anchor to which all student data policies and procedures are tied. During the course of an ethical analysis of current policies and procedures through the lenses of the four key values, those conducting the analysis should expect to continuously revisit the institutional mission statement to check for alignment between the mission and the policy or practice. If a gap is uncovered, this provides an opportunity for change whether it be a change in the policy or procedure or a change in the institutional mission. UT-Austin and MSU post their mission statements on their websites for anyone to view. Through their mission statements, one can see the role of the institution in promoting a liberal education through statements like "providing outstanding undergraduate, graduate, and professional education to promising, qualified students in order to prepare them to contribute fully to society as globally engaged citizen leaders" (MSU, 2008) and "contributes to the advancement of society through research, creative activity, scholarly inquiry and the development and dissemination of new knowledge, including the commercialization of University discoveries" (UT-Austin, 2021). All policies, procedures and practices of the universities should be aligned to their institutional missions. However, as the capabilities of data analytics technology expands, so too does the demand for it and at a rapid rate. The push to adopt and take advantage of data analytics by universities to inform decision-making requires a careful look back at the mission statements to ensure that the increased frenzy to adopt complex data analytics has not veered institutions off course. The following examines each of the four key values in terms of university student data use—keeping mission statements in mind—and how universities like MSU and UT-Austin can analyze their policies and practices according to their missions and through the lenses of these key values.

Autonomy

When discussing the paternalistic role universities have over their students—whether overtly established or not—students may be put in situations where they do not realize they have autonomy over decisions they make, or decisions made on their behalf.

The students in these situations lack a sense of self-efficacy. Only when a sense of personal self-efficacy exists can a student have autonomy over their educational trajectory and post-graduate life (Baggini & Fosl, 2007). A lack of personal self-efficacy—feeling unable to influence events in their own lives—can lead students to feel a sense of futility and potentially despondency. Causes for this can either stem from students feeling like they truly lack the ability to effectively control things around them or they may feel capable yet don't exercise their autonomy as they don't trust that their efforts will produce results they desire (Bandura, 1982). According to Baggini and Fosl (2007), "it's not simply their right, (…) autonomy makes possible human beings' greatest dignity, fulfillment, and happiness" (p. 185). They posit that autonomy is more than a right, it is a basic human need to live a fulfilling life. The more people feel responsible for their own lives, the greater chance they will be motivated to actually take responsibility.

Universities have an obligation to ensure that students do not get trapped by feelings of futility when it comes to authoring their educational experiences. Self-efficacy "operates through its impact on cognitive, motivational, affective, and decisional processes" (Bandura, 2006). It affects how students see themselves which affects their goals and aspirations along with their motivation to succeed and their persistence in the face of adversity. Self-efficacy has been correlated with increased autonomy, agency (Bandura, 2006), motivation (Kulakow, 2020), and academic achievement (Henri, Morrell, & Scott, 2017; Tilfarlioglu & Doğan, 2011). Beyond their time as students, a personal sense of efficacy can influence the choices they make at critical junctures in their lives (Bandura, 2006) (e.g. whether to accept a challenging job opportunity, run for public office, or make a lifestyle change).

Students who feel empowered will actually exercise that power. The more self-efficacy can be instilled in students, the more students will exhibit autonomy in their actions. Self-efficacy can be developed through different means including engagement in student centered learning environments as opposed to teacher directed learning environments (Kulakow, 2020)—similar to Freire's (2018) call for engaging in problem-posing education in lieu of traditional banking style education. A related strategy for increasing self-efficacy is for students to engage in authentic learning experiences such as practicums, internships, and research projects (Henri et al., 2017). Another influencer on self-efficacy is feedback on student performance. There is a direct correlation between feedback and student self-efficacy—positive feedback can have a short-term positive influence whereas negative feedback can do the opposite. A notable barrier to building self-efficacy among college students is a lack of perceived progress. Perception of progress—or lack thereof—can be difficult to measure though incorporating self-reflection activities throughout a student's educational experience can be helpful, aiding in building self-efficacy and autonomy (Henri et al., 2017). To enhance the development of student self-efficacy and autonomy requires that systems supporting student achievement—such as academic advising and instruction—be responsive to the desires and efforts of students and relinquish a measure of traditional paternalistic behaviors that diminish student self-efficacy and autonomy.

In many cases for students, development of efficacy is influenced by power imbalances. These power differences happen between students and their academic advisors, instructors, admissions counselors, coaches, and parents to name a few. A coach can tell a student that it is their choice whether to attend the team orientation or not but how much self-efficacy does the student feel to make a truly autonomous decision in this case? Related to data analytics, if an academic advisor relayed the results of a predictive analysis showing programs where the student is predicted to find greater academic success yet those results run counter to the interests of the student, what amount of pressure might the student feel to succumb to the data results rather than their own interests and passions? In such instances, the data itself seems to also create a power imbalance.

In order to have autonomy over their educational choices, students must have the necessary information and ability to think rationally (Baggini & Fosl, 2007) about how to best use that information to their advantage. There has been limited empirical research on how autonomy develops in university students but one study trying to answer that question found that students' self-perception of autonomy did not increase during the two year period of the study (Henri et al., 2017). This points to an opportunity for universities like MSU and UT-Austin to direct efforts toward improving autonomy among the student body and with a focus on the use of their data for analytics. This requires a certain amount of transparency from the university in communicating their policies and processes around the use of student data directly to students including opportunities students may have to opt out of certain data collection and use. The implications for students after graduation is that they gain a greater understanding of the value of their data, how they produce it, how it is used by others, and the control they have over it. Ideally, students graduate feeling empowered to become agents of change through policy-making and reform concerning data use in the future as it becomes a greater influence in all aspects of life.

It is not enough for educational institutions to simply respect student autonomy, but they must also actively help it develop. There was little found in the document analysis of MSU and UT-Austin policies and procedures that indicated a conscientious effort to develop student autonomy when it came to their data. In some cases, student autonomy appeared threatened. For instance, the Acceptable Use Policy for University Students at the UT-Austin explains a myriad of data protections, student rights, and responsibilities but it is quite lengthy which can be a barrier for many students in that they simply will not read the entire policy. If students sign a policy they do not fully read, is their consent informed or are they signing off for the sake of efficiently moving on to what they deem as more important endeavors? Informed consent from students is important in helping develop their autonomy as well as communicating succinctly the policies that impact them and the use of their data. UT-Austin's Information Uses and Security Policy includes a multitude of external policy references—including active links to full policy documents from within the document. The inclusion of so many policy links appears to work toward protecting the institution from potential lawsuits rather than actually informing students of their rights and responsibilities. One can assume that most readers of the Information Uses and Security Policy would not click on every related policy link listed or fully understand everything written in those related policies if they did.

Within the MSU Privacy Statement is a statement about students being able to opt-out of Google Analytics. However, in practice, only some students fortunate enough to have background knowledge of how to adjust privacy and cookie settings experience autonomy in this case. Those without the background knowledge or tenacity to hunt down the instructions buried within pages on the Google website are, in a sense, coerced into allowing tracking of their browsing activity.

Privacy

In chapter three, the notion of privacy was reflected upon as being more complex than in the past. No longer is privacy limited to the three traditional principles Nissenbaum (2004) described—protecting against intrusive government, restricting access to confidential information, and curtailing intrusions into personal spaces—but how big data exists and operates in the world must now be considered. To what extent are users aware of the data being collected and how it is being used and possibly shared with others? If consent to share their data is given to one organization, users tend to believe that their data will be used in just that manner and only by those to whom they specifically granted consent. When that does not happen—when their data is shared with others or used in a manner in which permission was not given—then a breach of trust occurs (Waldman, 2018). This breach of trust does not infringe on the three traditional principles Nissenbaum (2004) described but rather raises a newer threat to privacy peculiar to big data use.

Higher education is not immune to these issues of privacy. Student data is collected, stored, and analyzed by many units of a university in order to recruit new students, offer services to help them achieve success while attending the university as well as after graduation, and to run the virtual and physical operations of the university. Students are given a student code of conduct which they sign when enrolling in their institution of choice that most often includes expectations for student use of institutional technology or they may also receive a separate acceptable use of technology agreement to sign. On institutional websites—such as MSU and UT-Austin—there is often a link at the bottom of each page to the university privacy statement or policy. Although not always the case, privacy statements are typically general in scope while the privacy policy explains more of the nuances of how the university handles student data. Explained in detail in chapter five, much of what is explained in the web privacy policies of both MSU and UT-Austin is how student data is safeguarded against unauthorized surveillance and use and how individual identities are protected through deidentification or anonymization of data (IT@UT,2019; MSU, 2018; Tech. at MSU, 2015; Tech. at MSU, 2016). Some confusion may occur for students at MSU when they read within the privacy statement that different university units and websites might have their own privacy policies particular to the unit or website needs. (This occurrence stems from the historical creation of departmental silos when scientific management of universities became popular for boosting efficiencies during the Industrial Revolution; Thelin, 2011; Trujillo, 2014.) There is another disclaimer made by MSU that they are not responsible for any improper use of student data obtained through their website by others (MSU, 2018; Tech. at MSU, 2015; Tech. at MSU, 2016). This is bound to put students a bit on edge regarding how much trust they should put in their institution's efforts toward data privacy. An MSU student may find all the privacy protection efforts laid out by their university hollow if, in the end, the use of the data shared with others is outside the influence of university data use policies. In order to shore up trust between MSU and its students, a policy of sharing data only with outside vendors whose own data privacy policies meet the standards of privacy set by the university should be followed. This should be communicated to students to fortify trust students have in their institution protecting their data privacy.

Students expect that the data collected about them by their university will be used to benefit them and their educational experiences. There is a certain amount of trust students have when it comes to sharing data with their university due to the university being an educational institution. The earlier discussion about the impact of the COVID-19 pandemic on data collection by campuses provides an clear example of the trust students place in their institutions. Part of the condition for living on MSU's campus during the COVID-19 pandemic was that students agreed to register in the COVID-19 Early Detection Program and get tested each week in order to detect the virus quickly and mitigate the likelihood of a large outbreak so classes could continue uninterrupted (COVID-19 Early Detection Program, n.d.). (UT-Austin students were given an option to get tested but no requirements regarding implementation of COVID-19 precautions were required of students returning to campus during the pandemic; Hartzell, 2020.) Students at MSU who chose to live on campus during autumn 2020 willingly gave up data normally protected under HIPAA because they wanted the on-campus learning experience and they trusted MSU to handle their health data responsibly and in the context of detecting COVID-19 and protecting students, staff, and faculty. (Incidentally, the COVID-19 Early Detection Program continued into the spring 2021 term; Davis, 2021.) If the data had been used outside of this context, student faith in their university being truthful and forthcoming when communicating with them about data collection would have been damaged and potentially have hindered the efforts of the university in providing services to students while protecting the university community. Universities must take care to use the data they collect about their students within the context of delivering quality services that benefit students while also serving the institution.

Privacy of student data has gained much attention by leaders at institutions of higher education. The amount of policy documentation presented on how student data privacy is protected at MSU and UT-Austin demonstrates a focus in this area. However, the feeling students may leave with is that privacy policies are for the benefit of the institution—avoiding lawsuits—rather than to inform students. If the information shared with students were actually meant for them to consume and truly understand, policy documents wouldn't be so lengthy, full of external links to additional related policies, and include so much legal jargon. There is an opportunity for institutions of higher education to consider the educational value their work toward student data privacy can have for their students. Institutions should incorporate strategies that lead students to a greater understanding of what privacy means, their right to privacy, and limitations on ensuring privacy in various contexts when developing privacy policies, procedures, and programs.

Equity

Another aim of higher education institutions that espouse the goals of a liberal education is to ensure equity among students. Educational equity can be supported or stifled in many ways. The following discussion looks at barriers and opportunities for equity to be promoted based on the findings from the document analysis in chapter five.

Learning analytics may not be the first opportunity one thinks of for ensuring equity in education but there is potential. In analysis of policies and procedures around the use of student data analytics at UT-Austin, the lack of attention to learning analytics was noticeable. Although a great deal of effort was being placed on student success analytics and predictive analytics—particularly to improve student diversity and inclusion—there was no specific mention of how the university was leveraging learning analytics within courses and across programs. There was no reference to learning analytics on any of the websites and policies posted for public viewing. That is not to say that efforts in learning analytics are not being made at UT-Austin, there was simply no information about such efforts communicated at the time of this study. Much of the efforts from academic advising with student data analytics focused on identifying students falling behind in classes and advisors using that information to intervene before the students fell too far behind in their studies. Some may consider such efforts a form of learning analytics—the definition varies from one institution and context to another. However, with such advanced analytic capabilities as were described on the UT-Austin website (UT-Austin, 2018a; I.S.O., n.d.-a, n.d.-b; Texas One Stop, n.d.), it would seem that course and program level learning analytics would be yet another strategy in the arsenal used by faculty and staff to help identify areas where efforts could be made to redesign parts of the curriculum that pose barriers for certain marginalized student populations. Along those lines and in keeping with the goal of increasing diversity and inclusion in the student body, learning analytics could also be leveraged to identify instructional strategies that promote student inclusion, engagement, and improve performance outcomes.

The paucity of attention to learning analytics presented a similar gap in the use of student data analytics for educational equity at MSU. Although there was discussion regarding efforts toward leveraging learning analytics through collaborations directed mainly through the Hub for Innovation in Teaching and Technology, most of what was shared about the goals for learning analytics through the Hub were discussed at a high level and did not get into specific projects advancing research or practices in the field (Grabill, 2017). Capabilities for basic learning analytics were explained with regard to using D2L Insights by faculty, course designers, and administrators to track student behavior, activity, and performance in the learning management system, Desire2Learn (D2L) (Halick, 2017). It is clear that the use of learning analytics is still in its infancy at each university although there is movement at MSU in this area.

Grabill's (2017) article discussed another threat to educational equity in that some faculty, even when made aware of the capabilities of learning analytics to help them analyze their curriculum and instruction for unintended barriers to student success, might be reluctant to act. The effort required to rewrite curriculum or redesign a course of instruction is significant and some faculty may not be convinced of the value of spending their scarce time addressing issues in their courses. By not attending to issues built into courses that put certain students at a disadvantage in achieving success, inequities in the educational experience not only endure, they compound for the affected students as they progress—or do not progress—through their academic program. This can have lasting effects on their career and earning potential after leaving the university. Although this issue was not specifically address within the documents analyzed in this study, it is a potential barrier to leveraging learning analytics to support educational equity. Incentives and strategies to curtail such a problem should be in place if an institution is intentional about ensuring equitable educational experiences at the course level for students.

Another consideration related to the use of learning analytics is how to message students who are flagged for being at risk for failing. At MSU, D2L Insights was used to calculate whether a student was at risk for failing based on a programmed risk threshold. This is an interesting capability and use of the system as, at the outset, it appears to support educational equity in that students identified as being at risk for failing will assumingly receive extra support to help them succeed in the course. The potential threat to educational equity lies in the treatment of all students identified as at risk by the system the same, with the same outreach messaging, and the same intervention techniques. This would ignore the very personal experience students have when addressing academic challenges they each face. One student may respond positively to being told they are at risk of failing while that message may be enough to cause another student to shut down, quit trying to achieve or drop out altogether. D2L Insights and systems like it do provide great assistance for faculty and advisors working to intervene for students before it is too late for them to improve their performance, but the value of the relationships built between faculty and advisors with the students is still vitally important when communicating with students based on results of the analytics.

Inequity is sometimes unintentionally embedded in student technology use policies. The inequity may not impact learning, but it can have greater impact in other ways. Examples of this occurring is when privacy statements tell students they have control over certain data that is collected by the university by adjusting cookie settings or preferences for Google Analytics. Discussed earlier, such an attempt at supporting student autonomy was seen in the privacy statement at MSU and web privacy policy at UT-Austin. The manner in which these attempts to provide student autonomy threaten equity as a result. Offers for students to adjust their device privacy preferences in order to maintain a desired level of privacy lead to inequities among students based upon their previous knowledge about technology. Specifically regarding cookie preferences and Google Analytics controls, inequity occurs in that those students who know how to adjust cookies and Google Analytics preferences will be able to control the amount of data they have shared with the university and thus limit the potential for harm to them from the possible misuse of their data. Although students are given the option to adjust their cookie settings, in doing so, they will most likely lose functionality in the LMS which will have an impact on their ability to do academic work and participate fully in the online classroom environment. It's one thing to inform students of their rights but only those students who know what cookies and Google Analytics are and how to make adjustments would be able to exercise those rights unless there were support offered and a way to access that support right from the privacy statement itself. Otherwise, inequities will exist between those students who can limit the data collected about them while traversing the LMS and university website and those who cannot. Those who cannot exercise their right to limit such data collection by the university thus have more data collected about them that can be analyzed and used by the university. Although the assumption is that the data would be used in good faith to ultimately support the students and university, the fact that some students have more information collected about them than others simply because they don't have the resources necessary to exercise their right to limit that data collection puts them at a greater risk for misuse of that data.

Presently, universities are answering the call to address racial injustices in the wake of protests around the country stemming from a growing list of cases of police brutality against African-Americans. The University of Texas-Austin's Diversity and Inclusion Action Plan was co-created by faculty, administrators, and students to strategize a commitment to improve diversity and inclusion across campus (UT-Austin, 2018a). Yearly updates report on how student data analytics is being leveraged to achieve the goals laid out in the plan (UT-Austin, 2018b). Similar efforts to use student data analytics to promote student equity were not evident from the available MSU sources analyzed for this study. However, at the time of research and analysis of both cases, the protests demanding change in systems where systemic racism exists had not yet come to a head. It is quite possible that MSU—along with many other universities—will have plans in place or be in the process of developing plans to leverage data analytics to research, evaluate, and improve diversity, equity, and inclusion on their campuses during the 2020-2021 academic year.

Efforts by universities toward supporting equitable education for students will help ensure that all students receive the support they need to succeed—no matter what their personal definition of success may be. The steps taken by universities to safeguard educational equity will have a lasting effect on the lives of every student. By receiving the support needed to be successful in their chosen educational path, students are positioned to embark in careers and other endeavors they will find rewarding. Society as a whole benefits from a more diverse talent pool feeding into private and public employment sectors. There is an educational value inherent in diverse work settings as only by interacting with people of different backgrounds, races, ethnicities, cultures, and experiences can one truly understand the impact of various decisions on others outside of one's own perspective. This enlightenment helps each person be more empathetic and considerate of the particular work they do and the importance of performing it with care toward others who may be impacted by it—not to mention helping create more beneficial products and services for all potential users. This enlightenment is needed in order to improve equity in education and beyond. To move toward a more diverse and equitable future for students and graduates, universities like MSU and UT-Austin need to focus efforts toward engaging diverse students in conversations and actions influencing student data policy and procedures that promote educational equity.

Educational Value

In this section I will discuss the educational value of student engagement with decisions related to student data policies and procedures. There are two perspectives that will be considered: (1) the benefits to the university and (2) the benefits to students. While examining educational value from these perspectives, the mission of each university studied plays a role in demonstrating the importance of this value. To highlight the main tenets that support educational value, UT-Austin's core purpose is "to transform lives for the benefit of society" (UT-Austin, 2021). Included on the UT-Austin webpage with the mission statement is a list of the university's core values:

- Learning - A caring community, all of us students, helping one another grow.
- Discovery - Expanding knowledge and human understanding.
- Freedom - To seek the truth and express it.
- Leadership - The will to excel with integrity and the spirit that nothing is impossible.
- Individual Opportunity - Many options, diverse people and ideas, one university.
- Responsibility - To serve as a catalyst for positive change in Texas and beyond (UT-Austin, 2021, "Core Values" section).

MSU's mission statement specifically states:

Our mission is to advance knowledge and transform lives by:

- Providing outstanding (…) education to promising, qualified students in order to prepare them to contribute fully to society as globally engaged citizen leaders

- Conducting research of the highest caliber that seeks to answer questions and create solutions in order to expand human understanding and make a positive difference, both locally and globally
- Advancing outreach, engagement, and economic development activities that are innovative, research-driven, and lead to a better quality of life for individuals and communities, at home and around the world (MSU, 2008, para. 2).

The following discussion will demonstrate the alignment and importance of developing educational value around student data analytics to the mission and core values of these universities.

The University. It can be a difficult task for universities to think in terms of bringing students into the decision-making process around student data analytics policies and practices when it is quite counter to the traditional paternalistic role institutions have played. Additionally, as mentioned in chapter two, universities are competing more today for students than they have in the past. Growing competition for money to expand facilities, provide amenities, and hire the best and brightest faculty (Bowen, 2013; Carey, 2015; Rosen, 2011; Selingo, 2013) have necessitated streamlining decision-making as much as possible. The chase for student tuition dollars to help make up the void being left as state funding runs drier each year (Bowen, 2013; Rosen, 2011) falls in line with the value of money to oppressive systems. There is never enough money to satiate such systems; there is always a rationale for needing more money (Freire, 2018). Inviting students to participate in policy-making discussions with data and education experts may seem counter to the goals of efficiency and expertise. The value of the input provided by data and education experts over input from students in discussing student data policy matters caters to the paternal role of the university in the assumption that no one would know better than they as to what policy should be and how it should be implemented.

The scenario above demonstrates what Freire called banking education (Nelson et al., 2016). It is very much associated with paternalism in that—in a university setting—university leadership demonstrates through their actions a "we know best" mentality and students should trust university leaders to make decisions in their best interests; it is the leaders who have greater knowledge of the inner workings of the institution and how to care for a large student body. One could argue that institutional leaders at MSU and UT-Austin have more experience, education, and expertise in this regard than students so it is logical for them to set policy so the institution can move forward in its mission efficiently as possible. Another way banking education plays out at a university is in the delivery of education. Faculty and instructors have expertise and knowledge that they impart to students who memorize and report back during examinations. Freire (2018) goes so far as to describe it as "nutritionist banking" when students eat up the knowledge and vocabulary of the instructor who feeds what they themselves deem to be important. This situation exists with the abundance of data procedure and policy information shared via the MSU and UT-Austin websites for students to consume. Each university unit that has information to share—the registrar's office, the OCIO, LMS support units—decides what information is important to share and the best way to present the information to students. Students are in turn expected to ingest the volumes of information with little opportunity to question or discuss with those who created the procedures and policies. This method of instruction and communication stymies critical thinking among students as they are overloaded with information from those imparting knowledge—in this case, unit leaders at MSU and UT-Austin.

Another way to look at how the banking model of education is exhibited through data policies and procedures at MSU and UT-Austin is by considering how those who created the documents were educated for their profession. Extreme high-level specialization—such as data privacy and security specialists, IT specialists, lawyers, etcetera—potentially impedes learners from being able to make linkages as they study a great deal about their own narrow area of the world (Freire, 2018) Evidence of this type of phenomenon can be seen in the university privacy policies at MSU and UT-Austin where the amount of legal jargon used throughout the documents is extensive. Those writing the policies seem to have been so focused on their own purpose and from their own understanding of data privacy that they were seemingly unable to consider the perspectives of the students trying to comprehend the policies. This type of narrow focus benefits oppressive systems as they enjoy the status quo (Freire, 2018). Institutions can continue producing the legal and necessary policy documents to protect themselves from potential lawsuits very efficiently while keeping the risk of having to consider complex questions that may arise from more holistic policy development strategies low.

There are a few important reasons the practice of banking education in higher education is important to reflect upon within the context of the educational value student data analytics can have for students. As stated above, a reliance on banking education methods hinders student critical thinking and imagination. It can also lead to students relying on being told what is right and wrong or truth and fiction rather than the students working through problems to arrive at their own conclusions. Students can feel overwhelmed with information fed to them to the extent that they don't know how to think critically about it so they just accept the information and move on after receiving what they want—a grade that helps them move on to their next course. Should this strategy of moving through their university education prove successful for them, students may treat any situation where they feel overwhelmed with information with the same survival strategy—to have blind faith and simply accept the information without critique in order to gain a benefit and move on.

The unintended consequences of blind faith in student data analytics without oversight of the data sets and algorithms behind the analytic systems can manifest in biased decision-making—especially if the data sets and or algorithms are biased themselves. Historically, bias adversely impacts marginalized groups of students and perpetuates inequity in education, negatively impacting future opportunities for success (Nelson et al., 2016). Unintended consequences also arise when those using data analytics to make decisions receive results without understanding the visualizations provided to them or how the results were derived. They trust the results, sometimes with little or no question because of an assumption that the data is without bias or that the bias within the data is to be expected and there is not much that can be done to correct it—the mechanisms behind the data sets and algorithms are so complex that correcting the bias would be costly and prohibit time-sensitive decisions from being made. Those who do know how to process the analyses and create visualizations for decision-making have the power to create different narratives depending on how they run the analyses and how they present the results to decision-makers. End-users who do not understand how the data is collected, processed, and presented will always be influenced by those who do. They become the oppressed (Nelson et al., 2016). With few exceptions, students are in this position. Hence, universities have an obligation to engage students in learning about how their data is used and the consequences of data misuse so that they can actively participate in shaping policy and processes for the future they want to see.

Both MSU and UT-Austin shared an emphasis in providing information about student data privacy and security policies, the behavioral expectations of students, the rights they have related to accuracy and sharing of their information yet there is no mention of how students may be a part of the policy-making conversations or in the planning process for applications of student data analytics. They also shared a lack of student voice and representation in such conversations. Although it is customary for university administration to develop policies and processes as they deem necessary to deliver education while also protecting the welfare of the campus community, there is an opportunity for universities to educate students about the value and use of their data, how data policy is created, and to critically think through the implications of certain policy decisions by engaging students in the policy-making process regarding university applications of student data analytics. By including students in the policy-making conversations, students stand to gain a deeper appreciation of how their data is used by their university, the opportunity to critically analyze potential consequences of different policy decisions, and to learn first-hand how policy is created and approved. This type of experiential learning helps students see linkages between their classroom learning and current issues directly impacting them. Being included in student data policy and procedure conversations, debates, and development work helps develop students with autonomy and graduates who are more apt to engage in policy decisions at work and in their communities be it locally, nationally, or globally—all of which align back to the mission and values of both MSU and UT-Austin.

The Students. Only by recognizing the causes of oppressive systems can transformative action occur to form a new, more just situation. "However, the oppressed, who have adapted to the structure of domination in which they are immersed, and have become resigned to it, are inhibited from waging the struggle for freedom so long as they feel incapable of running the risks it requires (Freire, 2018, p. 47)." The issues around student autonomy, privacy, and equity may seem so insurmountable due to their complexity and invasiveness in all aspects of life that students may not feel they have any way out of their passive roles and thus accept life as it is.

The pedagogy of the oppressed must be created *with* the oppressed and not *for* them (Freire, 2018). "Authentic education is not carried on by 'A' *for* 'B' or by 'A' *about* 'B,' but rather by 'A' *with* "B,' mediated by the world—a world which impresses and challenges both parties, giving rise to views or opinions about it" (Freire, 2018, p. 93). Any policy, practice, instruction, and action around the use of student data should then be made *with* students' voices at the table, not merely *for* students to read and sign off on or ingest as through a banking education model. It isn't enough for them to gain an understanding of how their data is being used, they must also do something with that knowledge to help shape a new reality. This type of engagement would help MSU and UT-Austin carry out their stated missions of developing future leaders. However, there is no evidence that this is being done at either MSU or UT-Austin with regard to student data analytics procedures and related policies.

Being traditionally in control of policy and data processes, administrators may experience an awakening of their own to the unintended consequences of their policies and actions. In the case of UT-Austin, they experienced an awakening when their practice of using race as a factor for admission to the university was pronounced unconstitutional by the Texas state legislature. To address discoveries made during an awakening, leaders may try to adjust policy but to an end that is merely another form of oppression. For example, the Texas state legislature took the pronouncement of UT-Austin admission practices as unconstitutional a step further in passing the Top 10 Percent Law which mandated that the top 10% of students from every high school in the state must be guaranteed admission to any state funded institution. This included the more elite UT-Austin although for the 2021-2022 academic year the university's threshold for automatic admission will be only the top 6% from each high school graduating class (UT News, 2021). Although it was the state's answer to inequitable admissions practices, an unintended consequence of this law was that those students from lower socioeconomic regions where school resources weren't as plentiful as in higher socioeconomic regions struggled during their first year at UT-Austin due to lack of preparedness for the academic rigors at the university. Many students in that situation suffered from feeling that they didn't belong at the university and many dropped out. Unfortunately, because of that experience, those students may have continued to feel like they don't belong at a university or that they aren't college material after leaving UT-Austin and thus might end their quest for a college degree. This demonstrates how the best of intentions of the Top 10 Percent Law were actually just another form of oppression for some students that legislators were trying to help.

UT-Austin turned to data analytics to try identifying those incoming students who were predicted to struggle and possibly drop out. Those students were enrolled in special leadership programs and smaller classes during their first year in an effort to curtail students feeling like they did not belong and then dropping out. Although the university reported these efforts to be successful, a concerning aspect is that students were unaware of how or why they were chosen for the leadership program or smaller classes (Tough, 2014). The university was acting paternalistically and purposefully excluding students from understanding the mechanisms impacting their university experience. Although the claim would be that the program was to help students, there was a missed opportunity to educate students about how their data was being used by the university and to open up dialog with those students about the value of their data.

In order for true transformation to occur, leadership needs to work alongside students—in equal partnership—to develop new policies and procedures around the use of student data. Leadership at UT-Austin took steps in this direction when the university addressed racial inequities on their campus. Discussed earlier, the establishment of the University Diversity and Inclusion Action Plan (UDAIP) in 2017 included students in a year-long collaboration with faculty and administrators (UT-Austin, 2018a). Since its beginning, updates on specific efforts being made through the UDAIP to increase diversity and inclusion are reported annually (UT-Austin, 2018b). What is not clear is if the collaboration with students has continued beyond 2017.

Following this pedagogy has implications for students beyond the university and after graduation. Outside of academia, the oppressive systems are the market controllers—big business—who gather and use consumer data. Considering this, the duty of higher education is to raise awareness of oppressive actions by the data controllers and their consequences to citizens. University leaders can help students become aware of how their data is used within the university community and become active participants in designing policy and practice around the use of their data at the university so they are best prepared to do the same in their communities, with eyes wide open to oppressive tendencies by market controllers. Thinking back to the earlier discussion about student autonomy, Baggini and Fosl (2007) contended that the more people feel responsible for their own lives, the greater chance they will be motivated to actually take action. Applying that in context here, students who have an understanding of how data is produced, collected and used, the ethical implications of data misuse, and who have actively participated in data policy making at their university will be more likely to engage in public discourse and policy-making with regard to data analytics after graduation.

Recommendations

Higher education leaders may want to analyze how well they are attending to the ethical application of student data analytics but may feel lost as to how to embark on such a process. The following are recommendations for university leaders and policymakers when considering their policies and procedures affecting the use of student data analytics. These recommendations are initial steps and grouped according to the key values for clarity of purpose. The analysis of policies and procedures for each of the key values should be based on the extent to which each value is being attended in accordance with the stated institutional mission.

Autonomy

Higher education institutions like Michigan State University and the University of Texas - Austin can begin assessing the extent to which they promote student autonomy around student data analytics by looking at their existing student data policies and student-facing website pages and documents to see where information about how student data is collected and used is discussed. A starting point should be referencing the three guiding principles of student autonomy—listed below—to see the extent to which those principles are being supported or threatened:

- Processes and activities leveraging student data are designed and implemented in a manner that strategically supports the development of self-efficacy and agency among students—both needed for autonomy to exist.
- Information about how student data is collected and used by the institution is made transparent and communicated in a manner easily understood by students. Additionally, students are informed of their rights pertaining to their data collection, use, and access in such a way as to support their ability to readily act upon those rights if they so choose.

- Consent to collect and use student data is solicited from students consistently and with full transparency of the rationale along with the implications of consent or non-consent. Informed consent is obtained whether the data is used for research or other learning or operational interest at the institution.

Questions to help analyze the extent to which student autonomy around data is supported might include: (1) Where is the collection and use of data discussed? Is the information about how student data is collected and used readily accessible or is it buried within several layers of web links? (2) Once found, is the information easy to comprehend or is there significant use of technological or legal jargon which would hinder obtaining true informed consent if being sought and potentially lead to disparate experiences depending on each student's understanding of the information being communicated? "Often, educators and politicians speak and are not understood because their language is not attuned to the concrete situation of the people they address. Accordingly, their talk is just alienated and alienating rhetoric" (Freire, 2018, p. 96). When presenting students with policies and information they are meant to process and provide their consent to, it is vitally important that the information be written in language the students can easily comprehend. To solicit consent any other way is to solicit false consent and is meaningless beyond meeting legal requirements. (3) Is there mention of students being able to opt out of data collection and analysis? If so, when, where, and is it easy for students to actually exercise their will to opt out? Would opting out cause the students to lose a certain degree of access to educational resources or suffer hardships due to the technology not working as well as it would if they agreed to have their data collected and analyzed—thus enacting a form of coercion to agree to a data practice from which they'd rather opt out? Once these initial investigative questions are answered, university policy makers will be better able to see where adjustments can be made within student data policies and practices to help better instill a sense of self-efficacy and ultimately autonomy in their students.

Privacy

Autonomy and privacy are inextricably linked as autonomy cannot happen without a person having control over their information and hence a sense of personal privacy. By exercising control over who has access to information about them and for what purpose, a person exercises personal autonomy. As the goal of promoting student autonomy is a large part of providing an education for an engaged citizenry, providing for student data privacy is another important factor for universities to ensure.

Once again, using the privacy principles presented in chapter three—listed below—is a helpful place to being assessing how well student data privacy is protected or at risk based on institutional policies and procedures:

- Methods for student data collection and use by the university, along with when and why the data is collected and analyzed, are explained in a concise manner, free of legalese, and written for student consumption.
- Students are informed of when and how they can opt in or out of having their data used by the institution and the immediate implications of their choice so that they are able to give true informed consent if desired.

- The students' right to privacy is of prime importance and is protected by the university through the use of deidentified and anonymous data whenever possible, using identifiable data only when necessary.

The following questions provide guidance when assessing how student data privacy is maintained at a university such as MSU or UT-Austin: (1) How is a commitment to student data privacy communicated? Where and how is privacy of student data explained to students? (2) Are the types of student data collected and their uses described in easily comprehensible terms? To what extent is legal jargon or verbiage that could be confusing to students used in the privacy statements and policies? Are privacy explanations concise while being thorough or are they lengthy to the degree of running the risk of not being read and understood by students? (3) Are student rights and options—including recourse—concerning the collection, storage, and use of their data communicated in a manner easily accessible and understandable to students? (4) How are the contemporary complexities of privacy due to big data analytics—discussed earlier—addressed in university privacy statements and policies to instill trust among students?

After careful analysis of policy documents and university webpages containing any information related to student data privacy, university leaders and policy makers can begin making adjustments as necessary to better meet contemporary privacy expectations and communicate information about student data privacy. Purposeful attention to the complexities that big data bring to student data privacy issues along with improved communication to students will also result in the added benefit of increasing student autonomy.

Equity

The year 2020 presented the global community with many challenges while also raising the collective consciousness about racial injustices that have plagued society and brought a call for action to combat racial injustice. Higher education has an obligation to lead in such efforts by addressing diversity, equity, and inclusion among and for its students. In this case study analysis of MSU and UT-Austin, there were multiple strategies for leveraging student data analytics to improve student equity—a focus of this study—along with diversity and inclusion.

Institutions looking to assess how well they are creating equitable educational experiences can apply the equity principles presented in chapter three—and listed below—to review current policies and procedures:

- The aim of student data analytics is to provide equitable support mechanisms so that all students can achieve their academic goals.
- Data sets, code libraries, and the algorithms behind student data analytic programs undergo formal audits for bias at consistent and contextually appropriate intervals. Programming and analytic team members at all functional levels are empowered to identify and rectify issues that would result in inequity among students.
- Training around methods for interpretation of analytic results that mitigates bias and inequities are required of all faculty and staff who leverage student data analytics.

The following questions and action items can be used as a guide when assessing how student data policies and procedures impact educational equity: (1) Where is student educational equity discussed in the available policy documents and on the university website in relation specifically to student data? (2) What is the process and schedule for auditing data sets and algorithms for bias? (3) Upon reviewing the existing data sets and algorithms used by student data analytics tools for bias—including a review of predictive tools, learning analytics systems, and adaptive courseware—to what extent do they perpetuate inequities? (4) Where are assumptions about students embedded into policies that may place one group of students at an advantage over another. (5) How are faculty and staff trained in the interpretation and use of student data and how to mitigate unintended, inequitable treatment of students?

Educational Value

Universities can promote the educational value of student data analytics by applying Freire's pedagogy of the oppressed in designing strategies for educating and engaging students in university decision-making concerning student data analytics and relevant policies. For students to achieve conscientization they need to be able to reflect upon themselves as producers of data and what that data says about them. Students then need to learn how their data is used by entities in society—including their university—and how they can be beneficiaries of data analytics or victims of it when it is misused. Once students have a clear understanding of how they interact with the different facets of their data world they can become engaged in developing policies and practices around future data use with an eye toward protecting autonomy, privacy, and equity. Following a problem posing model for education is a recommended strategy for achieving this goal.

For the institution interested in evaluating how well they currently provide educational value for their students around student data issues, looking at the educational value principles outlined in chapter three—and listed below—is helpful:

- Student data policy and processes around the use of student data support the educational value for students, promoting the humanization of all students.
- Policy decision-making processes include the student voice as an important contributor to the creation of policy and procedures affecting the collection and use of their data.
- Increasing data literacy through the education of students about the collection and use of their data leverages inquiry-based strategies and ongoing dialog between students and university faculty, staff, and administrators.

The following are questions that can be used to help dig deeper into how educational value is supported or threatened. These questions are based on the characteristics of a problem-posing model of education that promotes the educational value of student data analytics while also bolstering the key values of autonomy, privacy, and equity: (1) How is discussion and self-reflection used to educate students on how their data is collected and used along with related privacy issues? (2) How do students inform policy-makers and practitioners of their views, hopes, and concerns when it comes to the policies and procedures around the use of their data? Are students involved in the work of developing policy and practice around student data and analytics? Are their voices are at the table and being heard in order to inform policy and practice? Is it a collaborative effort between policy-makers, practitioners, and students? (3) Are students developing autonomy as they engage in meaningful work affecting student data analytics policies and practices at their institution? Where can strategies for building student autonomy be integrated with policy making? (4) Is student self-reflection a pre-requisite for involvement in policy-making activities? Self-reflection does not need to occur before students can engage in policy discussion and policy-making work. The process of self-reflection is never-ending and so it is acknowledged that self-reflection and action can occur simultaneously (Freire, 2018).

Conclusion

This study sought to explore how institutions of higher education are writing institutional data policies and procedures that address the ethical complexities of student data analytics in an era of big data in order to protect the institution and its students from potential unintended consequences. A brief history of how data analytics gained importance in higher education from the 1800s to present day provided a look at drivers behind the adoption of data analytics by institutions. A review of the current landscape set the stage for a broad view of how various institutions in higher education are using student data analytics.

A review of the literature discussing ethical issues related to different applications of student data analytics provided the basis for undertaking this case study research. Higher education institutions are under pressure to prove their value to stakeholders so there appears to be a exuberance toward using data analytics for this purpose—in some instances without much thought toward potential unintended consequences. There is a lack of research as to how higher education policies related to student data analytics address the ethical implications of its use. This structured comparative case study sought to uncover answers and identify potential paths for further research and policy action

What was uncovered through the case study analysis were the areas each subject university focused heavily on along with the areas where scant attention was paid. Not surprisingly, a heavy emphasis was placed by both subject universities—Michigan State University and the University of Texas - Austin—on communicating operational IT processes, privacy policies, and laws governing the collection, use, and sharing of student data. Differences were apparent in specialized efforts each university undertook leveraging student data analytics to improve the student educational experience. While MSU posted information about their newly introduced D2L Insights portal for course-level learning analytics and institutional support for learning analytics endeavors through faculty collaborations with the Hub for Innovation in Learning and Technology, UT - Austin promoted their efforts toward diversity and inclusion and their use of predictive analytics to provide advising and extra support for students who may otherwise fall through the cracks of a large university system.

While the enthusiasm for leveraging student data analytics in the manner each university has is necessary to move them forward in providing quality services to their students, when analyzing policies and procedures related to such initiatives, the ethical use of student data analytics is notably absent beyond privacy protections. With the application of big data for student data analytics, it is increasingly imperative to empower students to protect themselves from potential misuse of their data. Because misuse of data can occur even within the confines of what is legal, it is appropriate to address ethical data expectations in policies and procedures. Higher education institutions can lean on the goal of providing a liberal education to fortify a democratic society which entails supporting and advancing the values of student autonomy, privacy, equity, and the educational value of engaging students in policy and procedural discussions and actions. By undertaking the work of analyzing their current policies and procedures according to the four key values and guiding principles presented in this dissertation, institutions will be able to see gaps where they can allocate resources toward making strategic improvements in policies and procedures that support the humanization of all students, faculty, and staff.

References

AtlanticLIVE. (2017, March 22). *The campus experience: Education at a crossroads.* [Video]. YouTube. https://youtu.be/PtlLsAzqlvo

Baggini, J., & Fosl, P. S. (2007). *The ethics toolkit: A compendium of ethical concepts and methods*. Blackwell Publishing.

Baker, R. S. J. d. (2010). *Data mining for education*. International Encyclopedia of Education (3rd ed.). Elsevier. file:///Users/ham.73/Google%20Drive/OSU/Grad%20School/Dissertation%20sources/Data%20Mining%20for%20Education.pdf

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122-147. doi:http://dx.doi.org/10.1037/0003-066X.37.2.122

Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science, 1*(2), 164-180. doi:https://doi.org/10.1111/j.1745-6916.2006.00011.x

Bare, C. E. (1966). Automated data processing in education. *The Educational Forum, 30*(4), 439-445. doi:10.1080/00131726609339741

Becker, S. A., Cummins, M., Davis, A., Freeman, A., Glesinger Hall, C., & Ananthanarayanan, V. (2017). *NMC horizon report: 2017 higher education edition*. https://www.learntechlib.org/p/174879/

Biemiller, L. (2017). Big data for student success still limited to early adopters. *The Chronicle of Higher Education, 63*(2). https://www-chronicle-com.proxy.lib.ohio-state.edu/article/Big-Data-for-Student-Success/239713?cid=rclink

Bloomberg. (2018). Facebook cambridge analytica scandal: 10 questions answered. *Fortune*. http://fortune.com/2018/04/10/facebook-cambridge-analytica-what-happened/

Blumenstyk, G. (2018). Big data is getting bigger. So are the privacy and ethical questions. *The Chronicle of Higher Education* (July 31, 2018). https://www.chronicle.com/article/Big-Data-Is-Getting-Bigger-So/244099

Bowen, W. G. (2013). *Higher education in the digital age*. Ithaka.

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W.W. Norton & Company.

Burbles, N. C., & Callister Jr., T. A. (2000). *Watch it: The risks and promises of information technologies for education*. Westview Press.

Burrell, S. (2017). The past, present and future of big data in higher ed. *The EvoLLLution*. https://evolllution.com/technology/metrics/the-past-present-and-future-of-big-data-in-higher-ed/

Cai, L., & Zhu, Y. (2015). The challenges of data quality and data quality assessment in the big data era. *Data Science Journal, 14*(2), 1-10. doi:http://dx.doi.org/10.5334/dsj-2015-002

Carey, K. (2015). *The end of college: Creating the future of learning and the university of everywhere*. Riverhead Books.

Clayton, M., & Halliday, D. (2017). Big data and the liberal conception of education. *Theory and Research in Education, 15*(3), 290-305. doi:DOI: 10.1177/1477878517734450

College Rank. (2016). The History of College Rankings. http://www.collegerank.net/history-of-college-rankings/

COVID-19 Early Detection Program. (n.d.). About the COVID-19 Early Detection Program. Retrieved February 13, 2021, from https://earlydetection.msu.edu/about

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4 ed.). Sage Publications, Inc.

Curren, R. R. (1995). Justice and the threshold of educational equality. In M. Katz (Ed.), *Philosphy of Education 1994* (pp. 239-248). Philosophy of Education Society.

Cuseo, J. (n.d.). *At-risk prediction instruments, early-alert systems & exit interviews: A proactive-to-reactive continuum of efforts to promote student success*. Shawnee State University. https://www.shawnee.edu/sites/default/files/2019-01/at-risk-early-alert-exit-total-10.pdf

Daniel, B. (2015). Big data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology, 46*(5), 904-920. doi:10.1111/bjet.12230

Davis, J. (2021). MSU's early detection program is back for spring semester - here's what's new. *MSU Today*. https://msutoday.msu.edu/news/2021/EDP-spring-program-update

Diaz, J. (2020). Amazon, TikTok, Facebook, others ordered to explain what they do with user data. *NPR.org*. https://www.npr.org/2020/12/15/946583479/amazon-tiktok-facebook-others-ordered-to-explain-what-they-do-with-user-data

Dudovskiy, J. (n.d.). Exploratory research. https://research-methodology.net/research-methodology/research-design/exploratory-research/

Edirisooriya, G. (2000). Information management in higher education administration: A slow drive on the information superhighway. In L. A. Petrides (Ed.), *Case Studies on Information Technology in Higher Education: Implications for Policy and Practice*. Idea Group Publishing.

Ekowo, M., & Palmer, I. (2016). The Promise and peril of predictive analytics in higher education: A landscape analysis. *New America.*

Ellucian. (2016). *College students and their data: What they expect of institutions*.

Felton, E. (2016). Temple University is spending millions to get more students through college, but is there a cheaper way? *The Hechinger Report*. https://hechingerreport.org/temple-university-spending-millions-get-students-college-cheaper-way/

Freire, P. (2018). *Pedagogy of the oppressed* (50th Anniversary ed. ed.). Bloomsbury Publishing Inc.

George, A. L., & Bennett, A. (2005). *Case studies and theory development in the social sciences*. MIT Press.

Gose, B. (2017). Getting faculty members to embrace student data. *The Chronicle of Higher Education, 64*(6), 5. https://www-chronicle-com.proxy.lib.ohio-state.edu/article/Getting-Faculty-Members-to/241323

Grabill, J. (2017). Vision for learning analytics at MSU. *The Hub for Innovation in Learning and Technology*. https://hub.msu.edu/vision-for-learning-analytics-at-msu/

Halick, S. (2017). Insights portal now available for MSU D2L course analytics. https://tech.msu.edu/news/2017/09/d2l-insights/

Hartzell, J. (2020). Testing update and spring semester. https://president.utexas.edu/messages-speeches-2020/testing-update-and-spring-semester

Henri, D. C., Morrell, L. J., & Scott, G. W. (2017). Student perceptions of their autonomy at university. *Higher Education, 75*(March 2018), 507-516. doi:https://doi.org/10.1007/s10734-017-0152-y

Herold, B. (2017). The case(s) against personalized learning. *Education Week, 37*(12), 4-5. https://www.edweek.org/ew/articles/2017/11/08/the-cases-against-personalized-learning.html#

Hubin, D., Hirsch, D., & Ham, M. (2017). *Ethics and big data in higher education*. [panel discussion]. The Ohio State University. Columbus, OH.

Ikpa, V. W. (2016). Politics, adequacy, and education funding. *Education, 136*(4), 468-472.

Import.io. (2019). Data mining process: The difference between data mining & data harvesting. https://www.import.io/post/the-difference-between-data-mining-data-harvesting/

Information Security Office. (n.d.-a). Acceptable use policy for university students. The University of Texas at Austin. https://security.utexas.edu/policies/aup

Information Security Office. (n.d.-b). Information resources use and security policy. The University of Texas at Austin. https://security.utexas.edu/policies/irusp

IT@UT. (2019). Web privacy policy. *Campus IT Policies.* https://it.utexas.edu/policies/web-privacy

Johnson, J. A. (2014). The ethics of big data in higher education. *International Review of Information Ethics, 21*(21), 3-10. http://www.i-r-i-e.net/inhalt/021/IRIE-021-Johnson.pdf?imm_mid=0c8f35&cmp=em-data-na-na-newsltr_20141217

Kaarbo, J., & Beasley, R. K. (1999). A practical guide to the comparative case study method in political psychology. *Political Psychology, 20*(2), 369-391. https://www.jstor.org/stable/3792081

Khan Academy. What is an algorithm and why should you care? https://www.khanacademy.org/computing/computer-science/algorithms/intro-to-algorithms/v/what-are-algorithms

Klasik, D. (2013). The ACT of enrollment: The college enrollment effects of state-required college entrance exam testing. *Educational Researcher, 42*(3), 151-160. doi:10.3102/0013189X12474065

Kowarski, I. (2019). 10 colleges with the most undergraduates. *Best Colleges.* https://www.usnews.com/education/best-colleges/the-short-list-college/articles/colleges-with-the-most-undergraduates

Kulakow, S. (2020). How autonomy support mediates the relationship between self-efficacy and approaches to learning. *The Journal of Educational Research, 113*(1), 13-25. doi:https://doi.org/10.1080/00220671.2019.1709402

Kumar, V., & Chadha, A. (2011). An empirical study of the applications of data mining techniques in higher education. *International Journal of Advanced Computer Science and Applications, 2*(3), 80-84. doi:10.14569/IJACSA.2011.020314

Kurzweil, M., & Stevens, M. (2018). Setting the table: Responsible use of student data in higher education. *Educause Review*(May/June 2018). https://er.educause.edu/articles/2018/5/setting-the-table-responsible-use-of-student-data-in-higher-education

Latour, B. (2012). Love your monsters: Why we must care for our technologies as we do our children. *The Breakthrough*. https://thebreakthrough.org/index.php/journal/past-issues/issue-2/love-your-monsters

Lestch, C. (2017). Arizona State University expands adaptive learning software from McGraw-Hill Education. *EdScoop*. http://edscoop.com/arizona-state-university-expands-adaptive-learning-software-from-mcgraw-hill-education

Macedo, D. (2018). Introduction to the 50th anniversary edition. In *Pedagogy of the Oppressed*. Bloomsbury Publishing Inc.

Mariani, G. (2018). 10 ways colleges use analytcis to increase student success. *eCampus News*. https://www.ecampusnews.com/2018/08/13/10-ways-colleges-use-analytics-to-increase-student-success/

Michigan State University. (n.d.-a). Colleges and Programs. https://msu.edu/academics/colleges-programs

Michigan State University. (n.d.-b). Notification of student rights under FERPA. https://reg.msu.edu/read/NotificationandDirectoryInformationNotice.pdf

Michigan State University. (n.d.-c). Office of the Registrar. https://reg.msu.edu/

Michigan State University. (2008). MSU mission statement. https://trustees.msu.edu/about/mission.html

Michigan State University. (2018). MSU privacy statement. https://msu.edu/privacy/

Neighborhood Student Success Collaborative (NSSC). (n.d.). Academic advising for exploratory preference and exploring students. Michigan State University. https://nssc.msu.edu/academicadvising/index.html

Nelson, L., Potrac, P., Groom, R., & Maskrey, C. (2016). Paulo Freire: Problem-posing coach education. In L. Nelson, R. Groom, & P. Potrac (Eds.), *Learning in Sports Coaching: Theory and Application*. Routledge.

Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review, 79*(1). https://digitalcommons.law.uw.edu/wlr/vol79/iss1/10

Nissenbaum, H. (2010). *Privacy in context*. Stanford University Press.

Oxman, S., & Wong, W. (2014). *Adaptive learning systems*. http://kenanaonline.com/files/0100/100321/DVx_Adaptive_Learning_White_Paper.pdf

Office of the Registrar. (n.d.-a). About us. Michigan State University. https://reg.msu.edu/aboutus.aspx

Office of the Registrar. (n.d.-b). Michigan State University access to student information. *Academic Programs Catalog.* Michigan State University. https://reg.msu.edu/AcademicPrograms/Text.aspx?Section=112#s542

Office of the Registrar. (n.d.-c). Student privacy rights. Michigan State University. https://reg.msu.edu/ROInfo/Notices/PrivacyGuidelines.aspx

Pasquale, F. (2015a). *The black box society: The secret algorithms that control money and information*. Harvard University Press.

Pasquale, F. (2015b). We're being stigmatized by 'big data' scores we don't even know about. *Los Angeles Times*. http://www.latimes.com/opinion/op-ed/la-oe-0116-pasquale-reputation-repair-digital-history-20150116-story.html

Pasquerella, L. (2019). The purpose of higher education and its future. *Liberal Education, 105*(3-4). https://www.aacu.org/liberaleducation/2019/summer-fall/president

Peters, R. S. (1969). *Ethics and education*. George Allen & Unwin Ltd.

Picciano, A. G. (2012). The evolution of big data and learning analytics in american higher education. *Journal of Asynchronous Learning Networks, 16*(3), 9-20. https://www.researchgate.net/publication/258206917_The_Evolution_of_Big_Data_and_Learning_Analytics_in_American_Higher_Education

Prinsloo, P. (2017). Fleeing from Frankenstein's monster and meeting Kafka on the way: Algorithmic decision-making in higher education. *E-Learning and Digital Media, 14*(3), 138-163. doi:10.1177/2042753017731355

Rajni, J., & Malaya, D. B. (2015). Predictive analytics in a higher education context. *IT Pro*, *July/August*, 24-33.

Reiter, B. (2017). Theory and methodology of exploratory social science research. *Government and International Affairs Faculty Publications, 132*. https://scholarcommons.usf.edu/gia_facpub/132

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

Robinson, R. (2018). Data privacy vs. data protection. [blog] https://blog.ipswitch.com/data-privacy-vs-data-protection

Rosen, A. (2011). *Change.edu: Rebooting for the new talent economy*. Kaplan Publishing.

Rourke, F. E., & Brooks, G. E. (1964). The "Managerial Revolution" in higher education. *Administrative Science Quarterly, 9*(2), 154-181. doi:10.2307/2391232

Scott, S. (2017). Solving the learning data puzzle: The Unizin Common Data Model. Unizin. http://unizin.org/2017/09/13/solving-the-learning-data-puzzle-the-unizin-common-data-model/

Selingo, J. J. (2013). *College (un)bound: The future of higher education and what it means for students*. Amazon Publishing.

Shahar, T. H. B. (2017). Educational justice and big data. *Theory and Research in Education, 15*(3). doi:10.1177/1477878517737155

Siemens, G., & Baker, R. (2012). *Learning analytics and educational data mining: Towards communication and collaboration*. Columbia University.

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist, 57*(10), 1510-1529. doi:10.1177/0002764213479366

Solove, D. J. (2006). A taxonomy of privacy. *University of Pennsylvania Law Review, 154*(3).

Stone, D. (2012). *Policy paradox: The art of political decision making* (Third ed.). W.W. Norton & Company.

Student Success Initiatives. (2017). Data driving UT Austin's increase in graduation rates. The University of Texas at Austin. https://provost.utexas.edu/news/data-driving-increase-graduation-rates

Tampke, D. R. (2013). Developing, implementing, and assessing an early alert system. *Journal of College Student Retention, 14*(4), 523-532. doi:10.2190/CS.14.4.e

Technology at MSU. (n.d.-a). MSU institutional data policy. *Guidelines & Policies.* Michigan State University. https://tech.msu.edu/about/guidelines-policies/msu-institutional-data-policy/

Technology at MSU. (n.d.-b). Student success dashboard: Find out how MSU supports student success with technology. Michigan State University. https://tech.msu.edu/service-catalog/teaching/student-success/

Technology at MSU. (2012). Acceptable use policy for MSU information technology resources. Michigan State University. https://tech.msu.edu/about/guidelines-policies/aup/

Technology at MSU. (2015). Privacy statement. Michigan State University. https://tech.msu.edu/privacy/

Technology at MSU. (2016). Student success dashboard privacy statement. *Academic Technology Service Catalog.* Michigan State University. https://tech.msu.edu/service-catalog/teaching/student-success/privacy/

Texas One Stop. (n.d.). Student records: Personal information. The University of Texas at Austin. https://onestop.utexas.edu/student-records/personal-information/

Thelin, J. (2011). *A history of american higher education* (2nd ed.). The Johns Hopkins University Press.

Tilfarlioglu, F. Y., & Doğan, F. Ş. (2011). Supporting self-efficacy and learner autonomy in relation to academic success in EFL classrooms (A case study). *Theory and Practice in Language Studies, (October)*. doi:DOI: 10.4304/tpls.1.10.1284-1294

Tough, P. (2014). Who gets to graduate? *The New York Times Magazine*. https://www.nytimes.com/2014/05/18/magazine/who-gets-to-graduate.html?_r=0

Trujillo, T. (2014). The modern cult of efficiency: Intermediary organizations and the new scientific management. *Educational Policy, 28*(2), 207-232. doi:10.1177/0895904813513148

University of Texas at Austin. (2018a). University diversity and inclusion action plan (UDIAP). https://diversity.utexas.edu/actionplan/

University of Texas at Austin. (2018b). University diversity and inclusion action plan progress updates - students. https://diversity.utexas.edu/actionplan/progress-updates/students/

University of Texas at Austin. (2020). Appendix C: Chapter 9 - Educational records. *University Catalogs.* https://catalog.utexas.edu/general-information/appendices/appendix-c/educational-records/

University of Texas at Austin. (2021). *Mission & values*. https://www.utexas.edu/about/mission-and-values

University of Texas at Austin Office of the Registrar. (2020). University catalogs. The University of Texas at Austin. https://registrar.utexas.edu/catalogs

U.S. News & World Report, L.P. (2020a). Michigan State University. *Best Colleges Rankings.* https://www.usnews.com/best-colleges/michigan-state-2290

U.S. News & World Report, L.P. (2020b). University of Texas at Austin. *Best Colleges Rankings.* https://www.usnews.com/best-colleges/university-of-texas-3658

UT News. (2021). Top 10 percent law. *UT News*. The University of Texas at Austin. https://news.utexas.edu/key-issues/top-10-percent-law/

Vasudevan, N., & Panthagani, A. (2020). 5 contact tracing lessons higher education can learn from governments. (May 27, 2020). EAB. https://eab.com/insights/expert-insight/business-affairs/contact-tracing-lessons-higher-education/

Velasquez, M., Andre, C., Shanks S.J., T., & Meyer, M. J. (2010). What is ethics? *Issues in Ethics*. https://www.scu.edu/ethics/ethics-resources/ethical-decision-making/what-is-ethics/

Vorhies, W. (2014). How many "V's" in big data? The characteristics that define big data. [Blog] https://www.datasciencecentral.com/profiles/blogs/how-many-v-s-in-big-data-the-characteristics-that-define-big-data

Waldman, A. E. (2018). *Privacy as trust: Information privacy for an informaiton age*. Cambridge University Press.

Weissman, C. G. (2018). How Amazon helped Cambridge Analytica harvest Americans' Facebook data. *Fast Company*. https://www.fastcompany.com/40548348/how-amazon-helped-cambridge-analytica-harvest-americans-facebook-data

Willey, J. (2018). Enterprise analytics: A driver of student-first Cculture. *Educause Review*(July 23, 2018). https://er.educause.edu/blogs/sponsored/2018/7/enterprise-analytics-a-driver-of-student-first-culture

Williamson, B. (2016). Digital education governance: Data visualization, predictive analytics, and 'real-time' policy instruments. *Journal of Education Policy, 31*(2), 123-141. doi:10.1080/02680939.2015.1035758

Zinshteyn, M. (2016, November 1). The colleges are watching. *The Atlantic*. https://www.theatlantic.com/education/archive/2016/11/the-colleges-are-watching/506129/