OTESSA
JOURNAL

Surveillance in the System: Datafication as Critical Change in Global Higher Education

Samantha Szcyrek [iD]
University of Windsor

Bonnie Stewart [iD]
University of Windsor

Correspondence:
Samantha Szcyrek
University of Windsor
Email: szcyrek [at] uwindsor.ca

Abstract
Over recent decades, higher education infrastructures have become increasingly digitized and datafied. The COVID-19 pandemic accelerated adoption of online learning platforms, trading the walls of the classroom for digital systems. Yet the surveillance, privacy, and discrimination issues that such systems raise are minimally understood by those who teach and learn within them. This paper overviews a 2020 pilot survey and 2021-2022 qualitative study of higher education instructors on a global scale. These projects explored the ways in which instructors from various locales and academic status positions understand data and classroom tools using proxy questions surrounding knowledge, practices, experiences, and perspectives. This paper draws on those studies to frame concerns about datafication amplifying issues in higher education. Its premises are twofold: first, if higher education instructors, as knowledge workers, are not knowledgeable about the contexts within which they teach and conduct scholarship, then the construct of shared governance within higher education is inevitably undermined. Secondly, if faculty and academic decision-makers are not intentional about equitable and ethical use of digital platforms within higher education, students' privacy and data is at risk. In this conceptual paper, we outline findings that frame datafication as a critical change within higher education culture.

Keywords: surveillance, data, datafication, higher education, datafied systems

Introduction: Datafication in Higher Education

Datafication refers to both the process and the technical, digital infrastructure through which traces of online engagement are collected and stored at mass scale. Datafication transforms social action into online quantified data, which allows real-time tracking and predictive analysis (Mayer-Schoenberger & Cukier, 2013) of subjects, objects, and practices (Southerton, 2020). "The rendering of social and natural worlds in machine-readable formats" (Williamson, Bayne, & Shay, 2020, para. 2) has marked recent decades with the rise of pervasive digital technologies, quantification, and big data processes (Southerton, 2020). More information is gathered, collected, sorted, and stored about the everyday activities of more people in the world than at any other point in human history (Andrejevic, 2012). Every click and keystroke, often even those we delete, are part of the datafied record created by our digital interactions.

Whenever anyone, anywhere, accesses digital systems, governmental systems, or telecommunications, they effectively open themselves up to datafication. Our perspective as researchers is that this represents a critical change within contemporary culture, with specific implications for higher education. In this paper, the terms "we" and "us" are used to refer to our perspectives as researchers but also to the broader human experience, in an effort to acknowledge the global scale of digital datafication and the fact that it impacts individuals and society regardless of personal buy-in. This does not, however, mean that implications and impacts are the same for all members of this global society, as the paper will explore, nor that we claim to speak for all.

In higher education, datafication has crept into many digital processes and activities that students, faculty, and staff engage in. This transforms social, pedagogical, and administrative actions that may previously have been unseen or untraceable into data that can be monitored, tracked, analyzed, and optimized (Shilova, 2017). This data can in turn be packaged and made visible to instructors and administrators at the senior and systems levels, as well as corporate vendors. In higher education as in the broader culture, datafication owes its rise to the reality that significant numbers of people trust their personal information to corporate platforms (Dijck, 2014), often as the price of access.

However, because education is widely framed as, at least in part, a public good, Williamson, Bayne, and Shay (2020) point out that the surveillance and extraction of data from students in education contexts cannot simply be analyzed as part of what Zuboff (2019) calls 'surveillance capitalism'. Zuboff's concept of surveillance capitalism highlights the increased data collection of individual users in society yet does not encompass the critical change of datafication in higher education systems raising surveillance, privacy, and discrimination implications. Rather, it demands specific analysis because these implications are a present concern in datafied systems currently implemented in higher education. As an example, Pardo and Siemens (2014) acknowledge that the ubiquitous adoption of technology for learning purposes amplifies the capacity to track learners, exposing privacy, security, surveillance, and discriminatory issues.

Yet datafication of education systems often remains invisible to both students and faculty. The role of educational technology vendors and corporate interests in shaping contemporary higher education is still minimally understood and discussed (Williamson, 2015). Datafication matters

not only for students, faculty, staff, and administrators, but also for practical, scholarly, and ethical reasons (Kimmons & Veletsianos, 2022).

Datafication had made significant inroads into higher education even before the COVID-19 pandemic. This is evident in the use of Learning Management Systems (LMSs), Student Information Systems (SISs), library resource systems, Online Program Management systems (OPMs), cheating detection software and systems, proctoring and assessment software, employment matching tools, Customer Relations Management Systems (CRMs). When the response to the pandemic forced institutions all over the world to teach and assess online, with educational technologies situated as a frontline emergency service (Williamson, Eynon, & Potter, 2020), the sector's reliance on datafied tools was only amplified. Due to this, a distinct datafied approach to pedagogy has emerged as a global norm from the educational delivery shift in 2020 (Williamson, Eynon, & Potter, 2020). Kimmons and Veletsianos (2022) state that "understanding the extent of its adoption—the prevalence and patterns—is imperative, since this information may allow estimates not just of expenditures but also of potential harm" (para. 4).

Our position is that this datafied reality and reliance represents a critical change within higher education, which demands significant attention and analysis. When it is addressed at all within academia, datafication is often framed in terms of benefits (Avella et al., 2016). However, risks must also be weighed. Datafication is an extractive (Erickson, 2018) process holding implications for faculty and students. Datafication platforms and processes increase both the capacity and the likelihood of surveillance within higher education systems, without sector-wide transparency or education of academic stakeholders on the topic. The race and gender biases (Noble, 2018; Benjamin, 2019) built into algorithms that sort, rank, and judge students in online classrooms or a proctored exam have been invested with significant power and can create differential and unfavourable outcomes for students.

Conversely, the incursion of corporate datafied systems into higher educational spaces represents a loss of power and knowledge for faculty, with implications for academic shared governance models. Shared governance refers to a transparent process where the university community – such as faculty, staff, students, and administrators – influence policy and procedure decisions by balancing efficiency, effectiveness, equity, and inclusion within operational frameworks such as sector wide legislation and ethics (Youngstown State University, 2017). Authors such as Tsai and Gasevic (2017) have encouraged cross-institutional and cross-sector communications among campus stakeholders, rather than top-down datafication decisions. Yet those of us who teach, even two years into COVID-19 online pivot, are seldom full partners in campus datafication conversations (Raffaghelli & Stewart, 2020). This leaves us ill-prepared, across the profession, to grapple with the predictive data that our classroom tools generate, or with the ethics questions they generate.

Research by Hillman, Bergviken, Rensfeldt, and Ivarsson (2020) suggest that education systems may become increasingly platform-based, decentralized, and marketized in coming years, potentially leading to a system where there is little public governance of the global commercial platforms that public education relies upon. In support of this claim, the Canadian Association of University Teachers (2021) states that shared governance is at risk when higher

education boards are increasingly controlled by corporate appointees, decision-making power becomes concentrated and secretive, and there is little to no input from academic staff and other key stakeholders. As researchers and educators in the higher education sector, we believe this represents a fundamental challenge to the concept of education as a public good, and the two visions have not been reconciled. Our paper therefore focuses on datafication as a critical change within higher education itself, one which we suggest might be addressed by a sector-wide ethics approach. Literature on educational data use identifies student surveillance in virtual learning environments (VLEs) as an ethics issue as early as 2005 (Land & Bayne, 2005) while Ferguson (2012) calls for higher education to address the need for data policy and ethics as tools' reach expanded and concerns about ethics and privacy. A decade later, few inroads have been made.

Our team of researchers has engaged in research efforts to examine higher education instructors' knowledge, practices, experiences, and perspectives regarding online classroom platforms and the data implications of digital tools more generally. Our hope is to inform policy and faculty development responses to datafication. Educators are knowledge workers, a term used to describe individuals with a "high level of education and skills and the use of information technology as an integral part of the informational labour process" (Pyöriä, 2005). As knowledge workers, there has been a shift in the work environment and knowledge about the work environment due to datafication. Our rationale for this research stems from the imminent concern that the sector has not yet "educated educators about data aspects of tools that have come to constitute a large number of its classrooms" (Stewart & Lyons, 2021, p. 64). If educators, as knowledge workers, are not knowledgeable about the conditions and infrastructures in which we teach and conduct scholarship, then the construct of shared governance within higher education is undermined.

The purpose of this investigation was to focus on educator's current data literacies and understandings of online classroom platforms and the data implications of digital tools. Guiding this research were four proxy questions based on instructors' knowledge, practices, experiences, and perspectives (Stewart & Lyons, 2021) to frame concerns about datafication as a critical change in higher education. Our main research question that underpinned this investigation was "What are the perspectives and practices regarding educational technologies and data of university educators who expect to use online teaching tools during the 2020/2021 academic year?" (Stewart & Lyons, 2021, p. 58).

We hope to encourage higher education professionals, as knowledge workers, to consider datafication as something we need to understand. In this paper, we present our findings from the 2020 pilot survey and emergent findings from our 2021 qualitative case study to discuss recommendations for how faculty can remain agential partners in shared academic governance. In doing so, we are not just focusing and addressing on what educators think but attempting to build critical capacity in institutions and the higher education sector.

Surveillance in the System: The Critical Change of Datafication

In contemporary society, surveillance has emerged as a dominant practice of organization, power, and control (Lyon, Haggerty, & Ball, 2012) that can regulate institutional routines and human behaviour through monitoring. Marx (2012) frames surveillance today as "new

surveillance," which "involves the scrutiny of individuals, groups, and contexts using technical means to extract or create information" (p. 22). In higher education, new surveillance is a function of almost all digital, datafied platforms, which together constitute most institutional systems, including registration, finances, communications, course offerings, and grading. This invisible but pervasive digital infrastructure makes surveillance ubiquitous and non-optional for anyone wishing to engage in contemporary academia.

This surveillance environment represents a critical change in the social imaginary that is higher education. In 1940, Sartre (1972) coined the concept of the social imaginary as the "set of values, institutions, laws, and symbols through which people imagine their social whole". The imaginary of the contemporary academy institutionalizes the idea of higher education as a site of public good, while embodying the values of a specific prestige economy (Blackmore & Kandiko, 2011) with status relationships and behaviours. The academic prestige economy has been rooted in cultural, social, and symbolic capital as much as economic (Bourdieu, 1984), and is difficult to reduce to quantifiable terms. Within this prestige economy, faculty and educators gain symbolic and academic capital as knowers and wield that knowledge and capital in the shared governance of the institution. The datafication of digital classroom tools and broader systems within higher education, however, creates a critical change that undermines the academic prestige economy and places decision-making power in mass data, extracted through the surveillant norms of digital platforms. The academic principle of the scientific method is founded in the premise that correlation does not equal causation, but the statistical associations that datafication turns into knowledge are correlational. In the worldview of big data gathered through surveillance and extraction, "numbers can speak for themselves" (Anderson, 2008).

Contemporary higher education decision-making is shaped by datafied tools and the values of quantification and automation that they represent. Nonetheless, a 2020 (Raffaghelli & Stewart) review of 137 papers on data literacy showed little focus on development of critical data literacies or awareness among higher education teachers. The two related research projects outlined in this paper stem from that examination of data literacies, and from the belief that faculty should be encouraged and supported to remain agential participants in higher education governance. Educators representing faculty on academic Senate governing bodies may not recognize the core tenets of this critical change that has been ushered into their domain and systems over the past decade. MacCarthy (2014) states that inadequate knowledge on datafied systems and digital classroom tools can be mitigated, reduced, or eliminated through improved transparency. Similarly, Metcalf (2022) acknowledges pursuing ways of making algorithmic tools more accountable to the public interest through 'algorithmic impact assessment' as a form of practice to help minimize and mitigate risks and harms.

Issues with Datafication

The COVID-19 pandemic saw higher education move towards reliance on datafied systems at an accelerated pace, globally (Bozkurt et al., 2020), particularly in terms of classroom and testing tools. With the rapid shift to emergency remote education (ERE) in March 2020, adaptability may have, in some cases, taken precedence over deep consideration for the ethical and risk implications of hurriedly adopting datafied systems into higher education. Likewise, training educators to use digital, datafied tools took precedence over engaging educators in an exploration of the critical change occuring in their sector or some of the issues it might generate.

Overall, datafication and surveillance work against the public good ideals of the higher education sector. Tracking and surveillance raise ethical concerns, such as inequitable social sorting, classification and profiling, segmentation, and social discrimination. Datafied education systems also risk the exploitation of student data in private markets, refocusing pedagogies to emphasize on what is countable or trackable (Williamson, Bayne, & Shay, 2020), and stigmatization of students through cross-campus data sharing (Benjamin, 2019).

Despite equity being an increasingly visible responsibility for higher education institutions, datafied classroom tools particularly amplify equity issues. The automated and quantified decision-making processes of datafied platforms have been shown to draw on and reinforce biases about class, race, and gender (O'Neil, 2016; Noble, 2018; Eubanks, 2018). Datafication can cause particular harm to racialized and minoritized populations. Facial recognition technologies and some video surveillance systems and online proctoring tools have been shown to disproportionately misidentify racialized people, or fail entirely to recognize dark skin (McIlwain, 2020; Jones, 2020; Stanley, 2022). The automated predictive modelling used in eligibility decisions and personalized learning techniques "includes 'unjustified discrimination' where faulty data or algorithms put people in the wrong category in error or where they are in the right category, but the category is a protected category like gender, age, or ethnicity which are legally impermissible bases for decision making" (MacCarthy, 2014, p. 19). Denial of educational funding for students in "risky" zip codes (O'Neil, 2016) is another way in which datafied systems can fail racialized and marginalized populations, as is digital redlining (Gilliard & Culik, 2016), a term used when digital tools, data analytics, algorithms, and filters are used to reinforce discrimination. Similar to the ways discriminatory – and illegal – loan, mortgage, and insurance policies were used in the mid-twentieth century to 'redline' Black citizens and limit their access to home ownership and attendant wealth, digital redlining occurs through "a set of education policies, investment decisions, and IT practices that actively create and maintain class boundaries through strictures that discriminate against specific groups" (Gilliard & Culik, 2016, para. 13). According to Tufekci (2020), these surveillance, privacy, and discrimination issues necessitate discussion of the datafied systems used within higher education, especially during and post-pandemic. Deepening the transparency and understanding of higher education data collection coul proactively support both the privacy of students and the overall learning environment (Brooks, 2021).

Authors such as Beetham et al. (2022) recognize that surveillance practices existed in education long before digital technologies, referencing Foucault's theoretical framework about surveillance and power. Nonetheless, new digital surveillance practices in higher education include tracking within classroom tools and email systems, and the normalization and implementation of data-based and data-generating digital platforms (Beetham et al., 2022). Cohney et al. (2021) note that some remote learning platforms use a range of tracking techniques to subvert users' expectations, and many platforms interact with third-party advertising and marketing in bidirectional manners. This introduces elements of surveillance that go beyond the bounds of the institution. Additionally, individual educators may not know that free or default licenses do not meet regulatory and normative expectations against surveillance, privacy, and security protections (Cohney et al., 2021).

Learning analytics (LA) are a core tool for collecting and analyzing learning-related data from systems and tools within higher education institutions. LA do offer the potential to increase

understanding of students' learning experiences, potentially resulting in a positive influence on student learning and progression through the analysis of data relating to students and their engagement with learning (Slade & Prinsloo, 2013). Scholes (2016) affirms positive reasons for institutions to use LA, such as predicting risk and supporting students at risk. That said, many LA tools require the input of personally identifiable student information to be effective (Currier, 2021). This required information may pose ethical concerns alongside privacy and security implications and issues because of the lack of consent and decision-making available to those having their data extracted and assessed. Early literature on learning analytics (Siemens, 2013; Slade & Prinsloo, 2013) established ethics and privacy as significant concerns emerging from the adoption of datafied systems, and the collection of LA data and its use raise ethical challenges (Scholes, 2016). Without more investigation into LA issues in relation to justice, diversity, and inclusion, "research limits the potential benefits of LADs generally and risks reinforcing long-standing inequities in education" (Williamson & Kizilcec, 2022, p. 260). As Scholes (2016) states, one of the most significant ethical concerns with LA is the failure to recognize students as individuals, which may impact students' autonomy and agency.

Automation is another significant issue stemming from datafied systems. Humans use agency to cooperate, but when engaging with computational tools implemented for quantitative measurement and metrics such as data analysis, agency is yielded in favour of automated digital systems (Applin, 2019). According to Applin (2019), "using counting, metrics, and implementation of outcomes from extreme data analysis to inform policies for humans is a threat to our well-being" (para. 7) as humans forfeit agency to algorithms and machines for them to operate and in turn provide what we are using the tools for. Automated datafied systems like plagiarism detection have become a widespread classroom tool in the digital age. For students to submit their assignments to receive their grades, they must accept this form of automation and ultimately risk their privacy. Brinkman (2013) acknowledges how the adoption of such services has proven to be controversial due to ethical concerns about students' privacy rights since the computerization and automation of such programs is changing the relationships of trust and responsibility between students, educators, educational institutions, and private corporations. Brooks (2021) states, "the issue of trust that the institution will use the data responsibly stands as the largest factor guiding students' decisions to consent" (para. 5).

Even choosing to opt-out of data sharing can produce significant privacy risks and lead to further issues. Taken together, this collection of issues amplified by datafied systems work to jeopardize mutual trust and relationships between educators and students and represent potential harm to students already most at risk within higher education.

Methods of Inquiry

This paper overviews two connected research projects that aimed to explore faculty understandings of the critical change in higher education represented by datafication. The first, aimed at simply establishing the 'state of the actual' rather than assuming what faculty knowledge, experiences, practices and perspectives might be, was a 2020 pilot survey. The second was an in-depth qualitative comparative case study, that interviewed a small group of participants from the prior year's survey to look at the issue in significantly more depth. The ultimate aim of the case study is to raise awareness about datafication among higher education professionals, and potentially to support faculty development and policy interventions.

Electronic Pilot Survey

The instrument for the pilot survey was a short 5–10-minute online questionnaire of 12 questions, one-third of which were demographic, one-third of which focused on digital platform use in online teaching, and one third of which posed single proxy questions aimed at gauging educators' knowledge, practice, experiences, and perspectives related to data and classroom tools. The instrument was designed as a foundation for the eventual in-depth study.

The pilot survey was intended and designed to capture a broad, international picture of university teaching faculty's surface-level perspectives and knowledge on data and classroom tools during the 2020 online transition due to the pandemic. Survey was chosen as a method for this stage of the research to enable broad and anonymous participation by educators and to ensure that premises for the intended follow-up study were based in a quantifiable reality rather than our own assumptions.

Pilot Survey: Recruitment and Participants

Table 1

Pilot Survey Participant Demographics

Years of Experience Range	N	%
5 Years or less	63	18.6
6-10 Years	82	24.2
11-15 Years	69	20.4
16-20 Years	58	17.1
21+ Years	67	19.8
Gender		
Male	110	32.4
Female	214	63.1
Self-Identify	4	1.2
Prefer not to Answer	11	3.2
Country		
Africa	16	4.7
Australia & New Zealand	7	2.1
Canada	129	38.05
Europe	34	10.0
Mexico	7	2.1
Middle East and Asia	4	1.2
South America	4	1.2
United States	137	40.4
Other	1	0.3
Discipline		
Humanities & Social Sciences	186	54.9

	N	%
Sciences	39	11.5
Engineering	14	4.1
Education	83	24.5
Law	7	2.1
Business	24	7.1
Other	40	11.8

The pilot survey was circulated online for three weeks during August 2020 (Stewart & Lyons, 2021). Participants who were currently teaching or expecting to teach in 2020/2021 in a university setting worldwide was eligible to participate in the pilot survey, whatever their academic status. The survey was completed by 339 participants from 25 countries, all of whom taught in higher education, with a range of experience: 19% had taught for fewer than 5 years, 20% had taught for more than 21 years. All respondents anticipated using digital tools within the 2020-2021 academic year.

Participant demographics showed that 32.5% identified as male, 63% as female. A small group self-identified with an alternate gender designation or selected "prefer not to answer." All respondents to the survey were invited to share contact information for a potential follow-up study. More than 60 of the 339 respondents offered to be contacted later.

Pilot Survey: Data Analysis and Findings

Findings from the summer 2020 pilot survey indicated that the pervasiveness of datafication in higher education does not translate into elevated levels of faculty knowledge or cautionary practice surrounding data and classroom tools (Stewart & Lyons, 2021).

The proxy question for data practice asked respondents how often they read the Terms of Service (TOS) for new educational technologies before bringing them into their classrooms, to gauge educators' habits in relation to classroom tools and data. The knowledge question asked respondents whether they were aware of the country in which the servers for their institutional LMS were housed, since server location can be important in issues of privacy law and the question proxied how informed educators were about the details of the tools they ask students to give up their data to. The experiences question asked respondents to indicate whether they had experienced any kind of privacy breach issues related to technology, and the perspectives proxy asked what respondents think institutions should do with the data from institutionally owned educational technology platforms.

Table 2

How Often the Full Terms of Service (TOS) Privacy Agreements of New Educational Technologies were Read Before Use with Students.

	N	%

90% of the time or more	32	9.4
50-89% of the time	42	12.4
10-49% of the time	63	18.6
Less than 10% of the time	202	59.6

On the practice front, responses indicated that majority of participants read the TOS for new tools less than 10% of the time, while fewer than 10% of participants responded that they read TOS privacy agreements often. The longer educators had been teaching in higher education, the less likely they were to report reading TOS (Stewart & Lyons, 2021), but variations by country were minimal: participants from all countries read the TOS privacy agreement 10% or less of the time.

Table 3

Do You Know Which Country/Countries House the Servers and Data for Your Current Institutional LMS?

	N	%
Yes	116	34.2
No	223	65.8

In terms of knowledge, two-thirds of respondents overall did not know where the institutional LMS servers they taught on were housed, and that number increased to 85% among participants from the United States. The US is the dominant location for educational technology vendors and thus it is US privacy law that educators in other jurisdictions need to be conscious of: that may be a reason for the geographic variance in responses. On the experiences question, a majority of respondents had not experienced digital privacy breaches as far as they were aware, and 90% had not experienced privacy breaches in relation to tool use with students. Finally, responses to that proxy question about what should happen to data were complex. Participants were allowed to choose more than one answer on this question, and of eight options framing different data visibility possibilities, the only option that generated a majority positive response on the data beliefs question was the one that said, "data should not be analyzed without express permission." A consistently larger proportion of the participants were open to the idea of anonymous patterns of engagement being shared within the institution, over identifiable data being shared, and there was greater comfort with student data than faculty data being visible.

Comparative Case Study

After the survey was completed, our PI and team embarked on an in-depth follow-up qualitative case study project, aimed at expanding on the pilot survey proxy categories and investigating how university educators' understandings of data and digital tools are generated and situated across various contexts. Case study methodology was chosen following Yin's (2009) assertion that case study can be a preferred approach when investigating a contemporary phenomenon within a real-life context, when *how* and *why* questions are in focus. Educators' understandings of the critical change that datafication represents in higher education are how and why

questions about complex real-life phenomena, thus the methodology is appropriate as a path to gaining multi-faceted insights regarding the research topic.

Moreover, comparative case study frames participants as individual cases in themselves while also remaining open to commonalities: our analysis employs both a compare and contrast logic to present cases individually, and a processual logic that traces emergent patterns across individuals within the participant group, exploring them as a collective case. Bartlett and Vavrus (2017) posit that the heuristic of comparative case study (CCS) "features an iterative and contingent tracing of relevant actors—both human and non-human—to explore the historical and contemporary processes that have produced a sense of shared place, purpose, or identity with regard to the central phenomenon" (p. 10). The comparative case study approach enables our team to look at both differences and commonalities in educators' experiences with datafication and classroom tools, deepening the snapshot represented by the 2020 pilot survey.

Two specific methods used in our case study approach included interviews and field notes. Interviews with individual participants enable us to investigate the knowledge, practices, experiences, and perspectives of participants in relation to datafication and classroom tools. Questions were semi-structured but open enough to enable thick description (Geertz, 1973) by participants of their own subjective contexts and the ways they constitute their relationships to data and online teaching platforms. Field notes in this case were not those of the researcher: rather, participants were invited to track any institutional or classroom tool-based data notifications and interactions over the course of a month, and to answer a short-written set of questions about those interactions and their perspective on them. The field notes questions paralleled many of those in the interview protocol but invited screenshots and specifics that we hypothesized might generate slightly different perspectives when phrased in writing rather than asked on the spot in an interview.

Individual Interviews: Recruitment and Participants

Table 4

Individual Interview Participant Demographics

Countries	N	%
United States	2	18.2
Mexico	1	9.1
Ireland	1	9.1
Scotland	1	9.1
Saudi Arabia	1	9.1
Canada	5	45.4
Gender	N	%
Male	3	27.3
Female	8	72.7
Role	N	%

Associate Professor	4	36.3
Assistant Professor	2	18.1
Lecturer	1	9.1
Learning Technologist	1	9.1
Coordinator of Teaching and Learning Centre	1	9.1
Program Coordinator	1	9.1
Adjunct Professor	1	9.1

A year after the pilot survey, the project team reached out to 20 of the pilot survey respondents to inquire whether they were still open to participating in a follow-up study. These 20 invitees were selected based on geographic diversity and range of status positions and faculties within higher education.

We hoped to ensure a range of voices and did not want to limit our case study to the dominant North American survey respondent group. 11 of the invitees agreed to be part of the case study project. These participants were based in six countries, including the United States, Mexico, Ireland, Scotland, and Saudi Arabia, with five participants based in Canada.

Of the eleven participants, there were eight women and three men. All three male participants were Associate Professors, in Engineering, Computer Science, and Business faculties. Of the women who participated, one was an Associate Professor (Teaching Stream), two were Assistant Professors, one a Lecturer (UK designation equivalent to an Assistant Professor in North America), one a Learning Technologist, one a Coordinator of a Teaching and Learning Centre, one a Program Coordinator teaching within her own program, and one an Adjunct Professor. The women taught in faculties of Law, Arts, Science, Nursing, and Education as well as in faculty development, teaching courses in History, Linguistics, Drama, Film and Media, Digital Technologies, International Development, Human Development, Law, and Psychology.

Participants all taught at universities, at institutions ranging in size from 2,000 students to 30,000 students, with one institution having 100,000 students across 26 national campuses. Participants had been teaching in higher education from five to 25 years. Only six participants began teaching fully online during the pandemic, while two more had begun in 2018 or 2019. The majority did have some pre-pandemic experience teaching using a LMS as a supplement to face-to-face teaching, though at least one had not done so.

Comparative Case Study: Data Analysis and Findings

do Amaral (2022) specifies that CCS provides effective tools to understanding policy and practice along three different axes of social scientific research: horizontal (spaces), vertical (scales), and transversal (time). According to Stake (2005), a case study refers to the process of inquiry regarding the case and the product of that inquiry, which is in alignment with our research approach as the focus of interests is on global educators' data literacies. We utilized the "collective case approach," a term Stake (2005) uses to describe focusing on multiple cases that are then analyzed for specific and general properties.

Semi-structured interviews with our 11 participants engaged conversations about educators' practices, knowledge, experiences, and perspectives in the context of their work in higher education, prior to and during the pandemic. We collated the interview transcripts with the five Field Notes responses we received from participants and used this collection as our full data set. Then, we used Braun and Clarke's (2006) reflexive version of thematic analysis process: data familiarization, coding, thematic extraction, reviewing themes, and naming themes. This process required us as researchers to be thoughtful and reflective as those are practices that are particularly regarded in quality standards and guidelines (Braun & Clarke, 2022). From the thematic analysis process, two distinct themes emerged: barriers to faculty understanding and agency, with respect to datafication, and beliefs about what should happen with datafication and datafied tools. Each of these included four sub-themes each.

On the horizontal CCS axis, which we used to compare individuals, we examined similarities and differences between participants' beliefs and barriers. Most of our participants, even those responsible for educational technologies on their campuses, indicated that overt conversations about data and its implications were rare. Yet whatever their knowledge of datafied systems and ownership, and their varied feelings about the usefulness of data and classroom analytics, all participants exhibited strong perspectives on what should happen to both student and faculty data collected by institutional tools:

> Claire*: A benefit can be some of the reminders for students to help them. Oh great, there's that assignment. I need to get going on that. From an instructor perspective, getting some sort of notification if someone has disengaged or hasn't accessed the platform. That's helpful to me then to reach out to offer support. As opposed to sort of from a policing perspective.*

> Susana: *I'm really not sure the data should necessarily be collected in the first place, I don't really see the point of it to begin with … I mean I don't wanna be like things were fine in the past, but it's not necessary to collect the amount of data to make things function.*

On the vertical axis, we explored how beliefs and barriers are shaped by influential differences such as networks, institutions, and demographic contexts. One commonality emerged across all interviews: participants consistently constituted their relationship to data and online classroom tools as part of their trust relationships with their institutions. One example of this is participants who work with digital tools more commonly and comfortably tend to have more complex and sometimes critical perspectives on data, because they are networked into those conversations rather than participants who are not. The vertical axis provided opportunity to pull examples of how to build critical capacity in the higher education sector.

> Ilse: *I know from Twitter I'm not the only one, but on campus in my department, I feel like I'm very often the only one's shouting in the desert. And then people go, oh yeah, yeah. But then we're also busy with everything else.*

> Wayne: *I have some ideas, but it's mainly because I'm highly connected from the top people down. I would not think the standard professor would know.*

Finally, the transversal axis gave insight into the big picture of datafication as a critical change in global higher education. An example of this is the trust relationship, which was in many conversations situated not just as an interpersonal employer-employee relationship but within a broader nexus of ethical responsibility related to the idea of education as public good. In interviews and field notes, the perception that institutions have complex, ethics-based responsibilities to students and faculty based in trust and public ideals emerged consistently, as did the conclusion that the sector and institutions have not yet enacted this responsibility sufficiently in relation to data risks and privacy issues. The most significant insight was about the recent and pervasive nature of the shift to datafication and the lack of policy and addressal of educators within institutional/sector responses.

> Cardinal: *So, I have not investigated the university tools because I am choosing to believe that the university has.*

> Wayne: *no discussion about the actual data collection implications on either platform. And I've actually asked the question. I ask the question, how can students and staff opt out of data collection practices in Canvas? But nobody answered me because it's a weird question and I'm a weird Canadian professor.*

Limitations and Delimitations

For the pilot survey, the most significant limitation is the narrow focus posed by the proxy questions regarding reading TOS and the location of LMS servers. Another limitation that emerges from this narrow scope is that participants may know more about the data aspects of their classroom tools, but that was not explored in detail. The third limitation in this study was that the concluded data did not represent a specific discovery, but rather addressed common patterns in the data (Stewart & Lyons, 2021). The delimitations of the survey included a focus on university educators only rather than higher education professionals more broadly, as the survey was shared in part via Twitter and we desired clear terminology regarding eligibility.

There are a few limitations within the qualitative study to address. Firstly, because the research is qualitative rather than quantitative, results are not replicable. Secondly, our participant recruitment and sampling present a limitation due to specific boundaries for inclusion needed, such as being individuals who were currently teaching or expecting to teach in 2020/2021 in a university setting for the pilot survey; and being pilot survey respondents willing to engage in the follow-up interviews. The third limitation was that since participants were in different geographical locations from the researcher and COVID limited both travel and gatherings, interviews could only be conducted online over Zoom. As for delimitations, the choice of qualitative research and methodology stems from the researcher's experience and

comfortability with qualitative over quantitative studies. Another delimitation is the choice to use respondents from the pilot survey as participants for the interviews, as this expanded upon their perspectives and practices highlighted in the survey but limited the opportunity to bring in participants from outside the survey respondent population. Furthermore, framing this study as a conceptual awareness piece rather than a traditional research study is another delimitation of the research. As for delimitations, the choice of qualitative research and methodology stems from the researcher's experience and comfortability with qualitative over quantitative studies. Another delimitation is the choice to use respondents from the pilot survey as participants for the interviews, as this expanded upon their perspectives and practices highlighted in the survey but limited the opportunity to bring in participants from outside the survey respondent population. Furthermore, framing this study as a conceptual awareness piece rather than a traditional research study is another delimitation of the research.

Discussion

Institutions around the world subject students and faculty to being tracked and surveilled by learning platforms during daily studies, without taking responsibility for educating them about this significant shift in the infrastructure of the academy or prioritizing privacy. The landscape of data privacy has been framed as technical and legal, but for higher education to maintain its ties to the concept of public good, it must also focus on the issues and implications of datafication. Through the use of direct participant quotations, our emergent study indicates that educators care about these implications, to the limited extent that they are knowledgeable about them but have not been cultivated as knowers or agents amid this critical change. Faculty knowledge is undermined by the incursion of this invisible layer of surveillance when we are not more critical of the datafied systems we are using in higher education.

Datafication in higher education as critical change calls us to confront issues such as surveillance, privacy, and discriminatory practices within institutions and the datafied classroom tools and systems used within higher education. These faculty within the sector of higher education expressed an underlying theme of social imaginary, highlighting that they are aware of the significance of understanding the datafied systems they extensively use, however this awareness does not translate into their actions when engaging with these datafied classroom tools. When higher education faculty assume the role of knowledge workers yet are unknowledgeable about the contexts within which they teach and conduct scholarship, then the construct of shared governance within higher education is unsurprisingly challenged.

Through our emergent research, it becomes clear that the disconnect between digital systems and the knowledge and ways in which we use them necessitates deeper analysis. Our research suggests that faculty in higher education possess limited knowledge and understanding of data as a critical change in the educational sector because data literacies are not in line with the accelerated pace at which these datafied tools are being used, especially amid the shift in educational delivery during the pandemic (Stewart & Lyons, 2021).

Conclusion

The critical change of datafication in higher education is no longer a concern for the future; it is happening now and warrants our critical attention. Ideas for the future and how they may influence education and pedagogy have a significant impact on educational thinking, policy, and practice in real-time, as ideas and desires for better futures are deeply constitutive of

educational practice, policy, and values (Facer, 2021). However, when ideas are consistently focused on the future of education, present implications and issues are often overlooked and under-addressed. According to Facer (2021), "the future' is both intimately and ubiquitously associated with education and yet this relationship remains poorly conceptualized in mainstream educational thought" (p. 3). Therefore, this paper synthesizes and expands on perspectives from literature regarding datafication's issues in higher education and shares our emergent research about educators' data literacies to explore data as a critical change that is currently happening – and continuously advancing – for both students and faculty.

Literature acknowledges that within recent decades, higher education infrastructures have experienced rapid digitization and datafication. Reliance on datafied systems has been accelerated due to the COVID-19 pandemic, with institutions around the world adopting online learning platforms as digital classroom tools, supplanting or at least supplementing traditional classroom learning environments with digital platforms. Mass datafication represents a critical change in higher education as a sector and brings with it a plethora of surveillance, privacy, and discriminatory issues that amplify and increase issues within the sector. Critically acknowledging datafied changes in the educational sector allow the opportunity to explore new avenues of learning about the risks and impacts of surveillance in the system.

In the digital age, whether we wish to or not, we engage with these datafied classroom tools that raise concerns about issues. Our research shows that this critical change and the issues it raises are only minimally understood by those who teach and learn within higher education. We must acknowledge the pervasive nature of digital systems in educational contexts and recognize the occurrence of datafication as a critical change that is happening with a need to understand it.

Author's Contributions

B.S (Co-author) was PI on both studies mentioned and wrote the methods and findings section in this paper. S.S (First author) was an RA for the 2020 pilot survey and did the background of conceptual research and literature review for this paper. Both authors contributed to the writing of the manuscript.

Open Researcher and Contributor Identifier (ORCID)

Samantha Szcyrek https://orcid.org/0000-0001-5425-0919
Bonnie Stewart https://orcid.org/0000-0001-9576-1037

Funding

Ethics Statement

The authors received approval from University of Windsor REB approval for the two studies.

Conflict of Interest

The authors do not declare any conflict of interest.

Data Availability Statement

Data from the 2020 pilot survey is available as an open data set in Zenodo:
https://doi.org/10.5281/zenodo.4096183

The dataset analyzed in the 2021 qualitative study is not publicly available due to potential identifiability of participants, but is available from the corresponding author upon reasonable request.

References

Andrejevic, M. (2012). Ubiquitous Surveillance. In *Routledge Handbook of Surveillance Studies*. Abingdon, Oxon: Routledge.

Applin, S. A. (2019, April 21). *What we sacrifice for automation*. Fast Company. https://www.fastcompany.com/90336550/how-much-are-we-sacrificing-for-automation

Avella, J., Kebritchi, M., Nunn, S. & Kanai, T. (2016). Learning analytics methods, benefits, and challenges in higher education: A systematic literature review. *Online Learning*, *20*(2), 13-29.

Bartlett, L., & Vavrus, F. (2017). *Rethinking Case Study Research*. Routledge. https://www.researchgate.net/publication/311831037_Rethinking_Case_Study_Research

Beetham, H., Collier, A., Czerniewicz, L., Lamb, B., Lin, Y., Ross, J., Scott, A.-M., & Wilson, A. (2022). Surveillance Practices, Risks and Responses in the Post Pandemic University. *Digital Culture & Education*, *14*(1). https://www.digitalcultureandeducation.com/volume-14-1

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim Code.* Polity.

Blackmore, P. & Kandiko, C. (2011). Motivation in academic life: A prestige economy. *Research in post-compulsory education, 16*(4), 399-411. https://doi.org/10.1080/13596748.2011.626971

Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste*. Harvard University Press.

Bozkurt, A., Jung, I., Xiao, J., Vladimirschi, V., Schuwer, R., Egorov, G., ... & Rodes, V. (2020). A global outlook to the interruption of education due to COVID-19 pandemic: Navigating in a time of uncertainty and crisis. *Asian Journal of Distance Education*, 15(1), 1-126.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2022). Conceptual and design thinking for thematic analysis. *Qualitative Psychology*, *9*(1), 3–26. https://psycnet.apa.org/doiLanding? doi=10.1037%2Fqup0000196

Brinkman, B. (2013). An analysis of student privacy rights in the use of plagiarism detection systems. *Science and Engineering Ethics*, *19*(3), 1255–1266. https://doi.org/10.1007/s11948-012-9370-y

Brooks, C. (2021, September 20). *Privacy Opt-Out May Lead to Inequities*. Inside Higher Ed. https://www.insidehighered.com/views/2021/09/20/should-college-students-be-able-opt-out-data-sharing

Canadian Association of University Teachers. (2021). *Shared governance. Quality education.* https://www.caut.ca/campaigns/shared-

governance#:~:text=For%20centuries%2C%20shared%20governance%20has,in%20their%20decision%2Dmaking%20process.

Cohney, S., Teixeira, R., Kohlbrenner, A., Narayanan, A., Kshirsagar, M., Shvartzshnaider, Y., & Sanfilippo, M. (2021). Virtual Classrooms and Real Harms: Remote Learning at U.S. Universities. *ArXiv:2012.05867 [Cs]*. http://arxiv.org/abs/2012.05867

Cukier, K., & Mayer-Schoenberger, V. (2013). The Rise of Big Data: How It's Changing the Way We Think About the World. *Foreign Affairs*, *92*(3), 28–40. https://www.jstor.org/stable/23526834

Currier, C. (2021). Unresolved Privacy and Ethics Issues Related to Learning Analytics in Higher Education and Academic Librarianship. *Emerging Library & Information Perspectives*, *4*(1), 117–142. https://doi.org/10.5206/elip.v4i1.13463

Dijck, J. van. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, *12*(2), 197–208. https://doi.org/10.24908/ss.v12i2.4776

do Amaral, M. P. (2022). Comparative Case Studies: Methodological Discussion. In S. Benasso, D. Bouillet, T. Neves, & M. Parreira do Amaral (Eds.), *Landscapes of Lifelong Learning Policies across Europe: Comparative Case Studies* (pp. 41–60). Springer International Publishing. https://doi.org/10.1007/978-3-030-96454-2_3

Erickson, K. (2018, July 17). The future of network effects: Tokenization and the end of extraction. *Medium*. Retrieved from https://medium.com/public-market/the-future-of-network-effects-tokenization-and-the-end-of-extraction-a0f895639ffb

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor.* St. Martin's Press.

Facer, K. (2021). Futures in education: towards an ethical practice. *UNESCO Futures of Education Report*. https://unesdoc.unesco.org/ark:/48223/pf0000375792

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, *4*(5/6), 304–317. https://doi.org/10.1504/IJTEL.2012.051816

Geertz, C. (1973). *The Interpretation of Cultures*. Basic Books. https://www.basicbooks.com/titles/clifford-geertz/the-interpretation-of-cultures/9780465093557/

Gilliard, C., & Culik, H. (2016, May 24). *Digital Redlining, Access, and Privacy* [Article]. Common Sense Education. https://www.commonsense.org/education/articles/digital-redlining-access-and-privacy

Hillman, T., Rensfeldt, A. B., & Ivarsson, J. (2020). Brave new platforms: a possible platform future for highly decentralised schooling. *Learning, Media and Technology*, *45*(1), 7–16. https://doi.org/10.1080/17439884.2020.1683748

Jones, W. D. (2020, February 24). *Racial Profiling Goes High Tech with Facial Recognition*. IEEE Spectrum. https://spectrum.ieee.org/do-you-have-the-right-complexion-for-facial-recognition

Kimmons, R., & Veletsianos, G. (2022). Proctoring Software in Higher Ed. *Light + Learning*. https://edtechbooks.org/light_learning_2022/proctoring_software

Land, R. & Bayne, S. (2013). Screen or monitor. In Bayne, S., & Land, R. (Eds.), *Education in Cyberspace* (pp. 165-179). London, UK: Routledge Falmer. https://doi.org/10.4324/9780203391068

Lyon, D., Haggerty, K. D., & Bell, K. (2012). Introducing Surveillance Studies. In *Routledge Handbook of Surveillance Studies*. Abingdon, Oxon: R Jones outledge.

MacCarthy, M. (2014). Student Privacy: Harm and Context. *International Review of Information Ethics*, *21*, 11–24. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3093299

Marx, G. T. (2012). Preface: "Your Papers Please": Personal and Professional Encounters with Surveillance. In *Routledge Handbook of Surveillance Studie*s. Abingdon, Oxon: Routledge.

McIlwain, C. (2020, June 3). *Of course technology perpetuates racism. It was designed that way. | MIT Technology Review*. MIT Technology Review. https://www.technologyreview.com/2020/06/03/1002589/technology-perpetuates-racism-by-design-simulmatics-charlton-mcilwain/?fbclid=IwAR1U-BC1BiN4eab4ROqPU-B9OAzQxROAeVUjgMrj5v-kAFpjUlMOe89c8Zc

Metcalf, J. (2022, March 17). *Raising Algorithmic Impact Assessment*. Medium. https://points.datasociety.net/raising-algorithmic-impact-assessment-1ebc6ba1bbeb

Noble, S.U. (2018). Algorithms of oppression: How search engines reinforce racism. New York, NY: NYU Press.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. London, UK: Allen Lane Penguin Books.

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

Pyöriä, P. (2005). The concept of knowledge work revisited. *Journal of Knowledge Management*, *9*(3), 116–127. https://doi.org/10.1108/13673270510602818

Raffaghelli, J. E., & Stewart, B. (2020). Centering complexity in 'educators' data literacy' to support future practices in faculty development: a systematic review of the literature. *Teaching in Higher Education*, *25*(4), 435–455. https://doi.org/10.1080/13562517.2019.1696301

Sartre, J.-P. (1972). *The Psychology of Imagination*. Routledge.

Shilova, M. (2017, June 15). The concept of datafication: Definitions and examples. *Apiumhub*. https://apiumhub.com/tech-blog-barcelona/datafication-examples/

Siemens, G. (2013). Learning Analytics. *American Behavioral Scientist*, 57(10), 1380-1400. https://doi.org/10.1177/0002764213498851

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

Southerton, C. (2020). Datafication. In L. A. Schintler & C. L. McNeely (Eds.), *Encyclopedia of Big Data* (pp. 1–4). Springer International Publishing. https://doi.org/10.1007/978-3-319-32001-4_332-1

Stake, R. E. (2005). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), The Sage handbook of qualitative research (3rd ed., pp. 443–466). Thousand Oaks, CA: Sage.

Stanley, J. (2022, February 8). *A Scary Demonstration of What Unchecked Face Recognition Could Look Like*. American Civil Liberties Union. https://www.aclu.org/news/

Stewart, B. E., & Lyons, E. (2021). When the Classroom Becomes Datafied: A Baseline for Building Data Ethics Policy and Data Literacies Across Higher Education. *Italian Journal of Educational Technology*, *29*(2), 54–68. https://doi.org/10.17471/2499-4324/1203

Tsai, Y. & Gasevic, D. (2017). Learning analytics in higher education – challenges and policies: A review of eight learning analytics policies. In Proceedings of the 7th International Learning Analytics Knowledge Conference Mar. 2017, pp. 233–242.

Tufekci, Z. (2020). *The Pandemic Is No Excuse to Surveil Students* [Magazine]. The Atlantic. https://www.theatlantic.com/technology/archive/2020/09/pandemic-no-excuse-colleges-surveil-students/616015/

Williamson, B. (2015). Governing software: Networks, databases and algorithmic power in the digital governance of public education." *Learning, Media and Technology* 40 (1): 83–105. doi: 10.1080/17439884.2014.924527

Williamson, B. (2020). A review of datafication and automation in higher education. *University and College Union*. https://www.ucu.org.uk/media/10947/The-automatic-university/pdf/ucus_the-automatic-university_jun20.pdf

Williamson, B., Bayne, S. & Shay, S. (2020). The datafication of teaching in higher education: Critical issues and perspectives. *Teaching in Higher Education*, 25(4), 351-365. https://doi.org/10.1080/13562517.2020.1748811

Williamson, B., Eynon, R., & Potter, J. (2020). Pandemic politics, pedagogies and practices: digital technologies and distance education during the coronavirus emergency. *Learning, Media and Technology, 45*(2), 107–114. https://doi.org/10.1080/17439884.2020.1761641

Williamson, K., & Kizilcec, R. (2022). A Review of Learning Analytics Dashboard Research in Higher Education: Implications for Justice, Equity, Diversity, and Inclusion. *LAK22: 12th International Learning Analytics and Knowledge Conference*, 260–270. https://doi.org/10.1145/3506860.3506900

Youngstown State University. (2017, December 19). *Principles and Practice of Shared Governance*. YSU. https://ysu.edu/provost/principles-practice-shared-governance

Yin, R. K. (2009). Case Study Research: Design and Methods (4th ed.). Thousand Oaks, CA: Sage Publications.

Zuboff, S. (2019). *The age of surveillance capitalism: the fight for a human future at the new frontier of power*. Profile Books.