2022-01

A Grounded Study of Higher Education Leaders’ Perspectives on ‘Big Data’

Harvey, David

UNIVERSITY OF CALGARY

A Grounded Study of Higher Education Leaders' Perspectives on 'Big Data'

by

David James Harvey

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF EDUCATION

GRADUATE PROGRAM IN EDUCATIONAL RESEARCH

CALGARY, ALBERTA

JANUARY, 2022

Acronyms:

AI	Artificial Intelligence
BASIs	Baccalaureate and Applied Studies Institutions
BI	Business Intelligence
CARIs	Comprehensive Academic and Research Institutions
CCCs	Comprehensive Community Colleges
CIRPA	Canadian Institutional Research and Planning Association
CIP	Comprehensive Institutional Plan
CRM	Client Relationship Management
ERP	Enterprise Resource Planning
HE	Higher Education
IB	International Baccalaureate
ICT	Internet Communications Technology
IMAs	Investment Management Agreements
IT	Information Technology
KPIs	Key Performance Indicators
LMS	Learning Management System
ML	Machine Learning
NSSE	National Survey of Student Engagement
PIs	Polytechnic Institutions
REB	Research Ethics Board

SACIs	Specialized and Arts and Cultural Institutions
SIS	Student Information System

Table of Contents

Table of Figures

Tables

Abstract

Higher Education (HE) institutions are facing a series of transformative challenges. From the increasingly pervasive use of Internet Communications Technology (ICT) to relative decreasing public investment and learner demographic changes, HE leaders are seeking new ways to improve learning environments and increase cost-effectiveness. 'Big Data' is an ICT technology that records and analyses massive sets of complex data to reveal previously hidden correlations linked to organizational performance. While 'Big Data' is widely implemented in sectors such as retail and healthcare, adoption in the HE sector is relatively sluggish. The purpose of the study is to develop a conceptual framework that describes and explains the adoption of 'Big Data' technology in the Alberta HE sector. A constructivist approach to grounded theory method is utilized to develop the conceptual framework from seven semi-structured interviews with HE leaders employed in a variety of HE institutions in Alberta. The study concludes that the adoption of 'Big Data' at HE institutions is impeded by IT governance that prioritizes risk mitigation through deductive reasoning. 'Big Data' proposals are not able to meet IT governance requirements. 'Big Data' project proposals are unconvincing because they require data collection and analysis before problems and cost-effective solutions are identified. Despite the challenges, HE leaders in the study demonstrate adaptive organizational learning through small scale 'Big Data' workarounds. Despite the slow pace to date, 'Big Data' adoption is poised to increase as educational technology venders begin integrating capacities in existing products – surmounting current IT governance and risk mitigation challenges. While acknowledging the potential of 'Big Data', the study also encourages a strengthening of HE leaders' ethical understandings, particularly as more advanced predictive and prescriptive analytics emerge and begin to impact learners and other HE stakeholders.

1.0 Chapter One: Research Proposal

Higher Education (HE) is currently experiencing a number of transformational challenges. Changing demographics and the need for more 'lifelong learning' are altering campus constituents (Livingstone & Guile, 2012; Usher, 2018). Neo-liberal political influences encourage greater competition, reducing overall public funding and urges that the remaining public funds should be distributed based upon demonstrated performance (Kirby, 2011). The pervasive growth of Internet Communications Technology (ICT) means greater competition as former geographic monopolies have less and less meaning and also multiplies institutions' capacities to automatically collect, store and analyze data (Calhoun & Kamerschen, 2010; Clark, 2018). The emergent 'Big Data' capacity means that HE institutions now have unprecedented capacity to analyze extremely large data sets to reveal patterns and associations that, if acted upon, could substantially increase learner retention and improve overall cost-effectiveness (Williamson, 2017). For example, the same way that Google can reliable and in real time predict and monitor outbreaks for flu by analyzing a massive amount of internet searches by key phrases and geography, HE institutions can (largely from online behavior) detect when learners likely need a specific intervention in order to succeed academically (Mayer-Schönberger & Cukier, 2013). However, 'Big Data' also raises a number of concerns related to privacy and fairness (Solove, 2004). Additionally, it is unclear if HE leaders are ready and willing to abandon 'knowledge-based' management approaches in favour of decisions informed by 'Big Data'. What is motivating the evolution either towards or away from the adoption of 'Big Data' in HE?

The purpose of the study is to develop a conceptual framework that describes and explains the adoption of 'Big Data' technology in the Alberta HE sector. The study includes seven semi-structured interviews of HE leaders who have leadership experience at a variety of post-secondary institutions in Alberta. The data is coded and analyzed to develop a conceptual framework to examine the phenomenon in future studies.

1.1 Background:

Today, HE leaders confront a myriad of unprecedented and profoundly disruptive forces. Learner demographics are transformed by drastically lower post-war birth rates and swelling enrolment from older domestic as well as international students (Barber et al., 2013). Access to public funding is substantially lower in the face of growing demand from other public sectors such as health care (Doyle & Delaney, 2009). Institutional monopolies are eroding as governments increasingly permit competing publicly-funded programming and ICT provides opportunities for learners regardless of assigned geographic mandates (Farhan, 2016). With the exponential development of ICT and the constantly changing requirements of the 'knowledge economy', learners' demands are magnifying in both scope and complexity (Livingstone & Guile, 2012). Higher tuition rates, greater competition among institutions and increasing skill expectations from the labour market, are also leading to a greater understanding of HE as simply another consumer good with a direct and immediate value for money (Stephenson et al., 2016). Finally, the proliferation of ICT is transforming all aspects of HE, from the design and delivery of programming to how institutions promote themselves and engage with stakeholders in a competitive marketplace (Jameson, 2013).

How should institutional leaders respond given the drastic scope and pace of change in HE? HE leaders managing systems in a state of extreme flux must not only mitigate negative impacts but, more significantly, identify and responsibly exploit new opportunities. As a result, many leaders in HE are investigating the use of 'Big Data' to improve institutional cost-effectiveness and competitiveness. 'Big Data' technology is a set of analytical tools that turn complex raw data into meaningful patterns and relationships (Daniel, 2015). The emergence of 'Big Data' stems from momentous advances in both computational power and the growth of 'data rich' environments. The growing routine use of ICT allows for the recording of an unprecedented level of automated online transactions and behaviours. By analyzing relationships between data sets, 'Big Data' can prescribe specific actions or investments to produce desired results (Picciano, 2012). In HE, 'Big Data' can support learning by determining which interventions or related behaviours have the greatest impact on learner retention and success (Dede et al., 2016). Student recruitment and stakeholder engagement can be enhanced by specific actions and communications that are proven to influence specific groups (Manyika et al., 2011). Financial efficiencies can be discovered through the analysis of heretofore undiscovered relationships between specific investments and improved results such as increased recruitment or lower attrition through enhanced learning environments (Lane, 2016). In addition to analyzing institutional internal processes, more advanced 'Big Data' practices now evaluate real time competitor behavior and recommend actions within a specific marketplace (Olszak & Mach-Król, 2018).

While a technological phenomenon, 'Big Data' is nevertheless governed by human understandings of its intrinsic worth, potential and ethical use (Kitchin, 2014). Is 'Big Data' a tool that offers a new level of understanding or is it a personal infringement that enables greater corporate or state control? Moreover, 'Big Data' raises underlying epistemological questions on the nature of knowledge and how it is defined, created and understood. Just because a relationship can be demonstrated using large amounts of data, should causality be inferred? What are the implications for methodology, bias and other forms of validity assessment? What is the reliability of inferences from large data sets? What tension is there between 'knowledge-based' and 'data-driven' decision making? How mindful are HE leaders of ethical issues regarding the collection of accurate and relevant data and issues of informed consent? What role does a competitive environment, or the lack thereof, impact 'Big Data' adoption by HE institutions?

1.2 Context:

This study examines HE leaders within a single jurisdiction – Alberta, Canada. There are 21 publicly funded post-secondary institutions in Alberta. Public funding is allocated primarily through an annual operating 'Campus Alberta Grant' and distinct capital projects. Campus Alberta operating funds are dedicated to delivering government approved 'credit' and apprenticeship programming. Administrative, faculty and support salaries plus other ongoing costs such as learning equipment for 'credit' and apprenticeship programs are allocated funding through the Campus Alberta grant. The Campus Alberta grant is allocated annually on an 'historical' (see https://www.alberta.ca/publicly-funded-institutions-government-support.aspx) basis that remains relatively static providing enrolment is sustained. Campus Alberta grants are differentiated by the five institution types as defined within the Government of Alberta's Post-secondary Learning Act. First, four Comprehensive Academic and Research Institutions (CARIs) deliver undergraduate and post-graduate programming as well as significant research functions. Second, two Baccalaureate and Applied Studies Institutions (BASIs) deliver university transfer, undergraduate, some post-graduate programming as well as applied research. Third, eleven Comprehensive Community Colleges (CCCs) provide academic and language upgrading, apprenticeship, certificate, diploma, post-diploma, some baccalaureate degrees usually in collaboration with degree granting institutions, post-baccalaureate certificates and applied research. Fourth, two Polytechnic Institutions (PIs) have a similar mandate as CCCs but with a greater emphasis on bachelor degree programming. Fifth, two Specialized Arts and Cultural Institutions (SACIs) offer certificate, diploma, undergraduate, graduate degrees and research in fine arts and culture. Capital projects are funded on an ad hoc basis to support an expansion of instruction, research and administration services. In addition, 'one time' funding opportunities are presented by the government to post-secondary institutions occasionally to address unique post-secondary programming needs.

In addition to public funding, post-secondary institutions in Alberta also generate revenue through tuition fees, research and other fee-based services and donations. The provincial government regulates fees for 'credit' programs through Section (61) of the Post-Secondary Learning Act (Alberta, 2017). Credit programs are degrees, diplomas and certificates approved by the Minister of Advanced Education for funding within the Campus Alberta operational grant. As per Section (2) of the Public Post-Secondary Institutions' Tuition Fees Regulation, specific exceptions to the regulation of tuition fees include: i) courses not approved by the Minister as per Section (61) of the Post-Secondary Learning Act, ii) distance programs for learners outside Alberta, and iii) any differential fees charged to non-Canadians or non-permanent residents (Alberta, 2018). Courses not approved by the Minister are considered "non-credit" and as such as not regulated as long as institutions can demonstrate that they are not funded by the Campus

Alberta grant. In addition, learners taking non-credit courses are not eligible to public student financing.

Historically, while Campus Alberta grants varied significantly from year to year, institutions enjoyed cumulative increases exceeding inflation until 2008. Over the past ten years, Alberta has experienced either cuts or only inflationary increases (Usher, 2018). Given the likelihood of grant revenue remaining flat or even decreasing in the short to medium term, HE leaders must look to alternatives to sustainably grow revenues outside the Campus Alberta grants. These areas principally are i) enrolment in credit tuition-based programs and non-credit courses, ii) international student enrolment, iii) contracts for services, iv) donations and v) reducing overall costs (Barber et al., 2013). Similar to other sectors, HE leaders are considering new opportunities through enhanced technology and decreased regulation to create new sustainable models for institutional development via greater engagement with non-governmental constituent groups.

In addition to the changes in the regulatory environment, technology is impacting HE in Alberta in a number of important new ways. First, former geographic boundaries are eroded by the ability to offer distance programming. As Alberta institutions have increasing latitude to offer similar credit programs at the certificate, diploma and degree programs, institutions inside and outside the province are using technology to compete in areas where public institutions heretofore enjoyed de facto monopolies (Olssen, 2016). Second, the increasingly pervasive use of technology has increased demand for HE (David & Foray, 2002). Given the pace and scope of technological change, training and certification throughout one's career is needed to maintain relevance in the 'knowledge economy' (Livingstone & Guile, 2012). As a result, the fastest growing demographic is learners older than twenty-five (Baer et al., 2015). Particularly among older learners, HE is beginning to be viewed as simply another consumer product. Learners understand like never before that they have choices and it is up to HE institutions to sell the comparative advantages of their products and services (Vrontis et al., 2007). Third, learner expectations of HE are changing as institutions begin a process of integrating technology across all areas of stakeholder interaction – a concept known as 'digital transformation' (Clark, 2018). Digital transformation addresses all the 'touch points' throughout the lifecycle with learners – from prospect, to learner, to alumni, to donor (Norris et al., 2013). Improvements can include providing relevant recruitment information based upon prospective learners' interests and customized programming that addresses individual learning preferences.

It remains unclear how technology and a decrease in relative regulation will impact competition for learners in Alberta. Partial performance-based public funding was discontinued in 2013 and domestic credit learner tuition fees remain frozen at 2015 levels. Nevertheless, tuition-based revenues in non-credit, international learners and selected cost-effective credit programming are vital to Alberta HE institutional sustainability. The perceived urgency among HE leaders to increase competitiveness through the greater use of 'Big Data' and other relatively new technology is uncertain. The implementation of 'Big Data' in HE in general lags behind other sectors such as retail (Daniel, 2015). A greater reliance upon sustained non-governmental revenues in HE may create what (Kirby, 2011) calls a developing 'quasi-market' that rewards cost-effective integration of new technology to create greater value for learners and other HE stakeholders.

1.3 Purpose & Problem Statement:

The purpose of the study is to develop a conceptual framework that describes and explains the adoption of ‘Big Data’ technology in the Alberta HE sector. Factors such as increased competitiveness and demands for increased performance measures are promoting the adoption of ‘Big Data’(Daniel, 2015; Dede et al., 2016). However, inhibiting factors include costs, privacy and fairness concerns and a general lack of confidence in the technology (Boyd & Crawford, 2012; Solove, 2004; Williamson, 2017).

As a result, the following research questions are noted as the focal points for the investigation:

1. What do HE leaders in Alberta consider the main elements for and against in ‘Big Data’ adoption or further development?

2. To what extent are Alberta HE leaders motivated by competitive enrolment and/or cost-effective learning considerations that could be developed by ‘Big Data’ applications?

3. Do Alberta HE leaders believe that meaningful business decisions can or should be influenced or even made by ‘Big Data’ applications?

4. What ethical considerations do Alberta HE leaders integrate in the data collection and analyses processes of ‘Big Data’ initiatives?

5. How able do Alberta HE leaders consider their respective institutions are to implement 'Big Data' successfully from technological and change management perspectives?

1.4 Assumptions

The study is designed with the following assumptions. First, that HE leaders engaged in the study have experience with the applications or potential applications of 'Big Data". Second, that interviewees will discuss the phenomenon in candid and direct ways. Third, that the interviews will yield a 'saturation' point or no new significant data after a reasonable amount of interviews and interview research subjects (Charmaz, 2014).

1.5 Method

The study utilizes a constructivist approach to grounded theory to develop a conceptual framework that describes and explains the adoption of 'Big Data' technology in the Alberta HE sector. Grounded theory is a qualitative research method that uses systematic procedures over time to collect and analyze data for the purposes of theory development or elaboration (Creswell & Gutterman, 2019). Grounded theory methods have particular utility when attempting to develop theory about human action or interaction when existing conceptual frameworks either do not exist or prove inadequate over time (Denzin, 2007). A systematic and iterative process collects and analyzes data (Bryant, 2007). Previously collected data influence subsequent data sets collection in order to continuously root the developing theory with demonstrable evidence (Charmaz, 2000). The study utilizes a constructivist approach which produces findings that are more contextual and flexible unlike positivist or post-positivist approaches that often propose more categorical and conclusive findings (Charmaz, 2000; Glaser & Strauss, 1967; Strauss & Corbin, 1990). The constructivist approach to grounded theory methods, as discussed by Creswell & Gutterman (2019), focuses less on generalizable findings and 'more explanatory, more discursive and more probing on the assumptions and meanings for individuals' (p. 441).

The study includes seven semi-structured interviews conducted before the data collection 'saturation' point was reached (see Figure 2: Interview Saturation Point) and the subjective judgement was made that more data will not provide significant new insight. Interviews provided the verbatim text for the 'theoretical sampling' for coding and analysis (Charmaz, 2014). Interviewees were senior leaders at three types of HE institutions in Alberta. After ethics approval from the University of Calgary and other relevant HE institutions, interviewees were selected after sharing the semi-structured interview questions. Theoretical sampling was collected verbatim and the semi-structured interviews recorded. The data was coded by the researcher as common themes on the phenomenon emerge. The coded data was analyzed and the proposed conceptual framework constructed from the data. Utilizing ‘emergent design’ processes, data collection and theory development occured concurrently (Creswell & Gutterman, 2019). Data analysis informed the ongoing data collection process until the ‘saturation’ point was achieved. The process emphasizes reflexivity with the researcher recording memos to promote an ongoing reflection on behalf of the researcher (Charmaz, 1990). Finally, the emerging theory was validated with the research subjects in order to assure research validity and a means to protect the research subjects from any errors or misunderstanding in the data provided.

1.6 Limitations and Delimitations:

The study is delimited in scope to interviews with HE leaders in Alberta, Canada. As well the study qualitatively examines HE leaders’ perceptions and does not attempt to independently assess the effectiveness of ‘Big Data’ applications. The study is further limited by a relatively small sample size. As a result, findings are not generalizable beyond the context of study sample among HE leaders in Alberta. However, ‘lessons learned’ from selected Alberta post-secondary leaders are documented and analyzed.

1.7 Summary

This chapter outlines the purpose, scope and significance of the study. It discusses the extreme state of flux in HE in general and in Alberta specifically and how the application of 'Big Data' may have a transformative affect in an increasingly competitive environment. The study's limitations and delimitations are discussed as well as an initial conceptual framework. Given that 'Big Data' is a relatively new phenomenon, a qualitative grounded theory approach to the study is adopted to help clarify a working conceptual framework for Alberta HE leaders in the future. In the Chapter 2, a review of the relevant literature further establishes the context and outlines the current and potential impact of 'Big Data' applications in HE.

2.0 Chapter Two: Literature Review

This chapter is a summary or relevant literature related to the research questions. In particular it discusses i) the definition of 'Big Data', ii) how 'Big Data' is currently used in HE, iii) HE leaders perceptions of 'Big Data', iv) data-driven decision making in HE, v) epistemology of 'Big Data', vi) organizational learning / change, vii) ethical considerations for 'Big Data, and viii) competition in HE and how it impacts 'Big Data' adoption and viii) elements for consideration in IT governance specific to the HE sector.

2.1 What is 'Big Data'?

'Big Data' is defined as the set of analytical tools used to translate large and complex data sets into meaningful patterns that represent organizational value (Daniel, 2015). The process of analyzing data is not new or dependent upon technology. However, the rising prevalence of automated data collection and increasing capacity for data storage and processing now creates an unprecedented situation where the value of insights produced by 'Big Data' regularly and significantly exceed costs (Mayer-Schönberger & Cukier, 2013). Traditional data analysis functions by comparing sets of 'structured' data where numbers fit neatly into pre-defined categories (Boyd & Crawford, 2011). Conversely, 'Big Data' records and analyzes vast amounts of unstructured or 'complex' data such as video, audio, diagrams, images or combinations thereof (Basu, 2013). In particular, two concurrent and accelerating ICT innovations serve to enable the development of 'Big Data': "digital transformation" and vastly increased computing power. First, 'digital transformation' is the growing number of interactions among individuals and organizations that are being mediated and recorded digitally through ICT (Clark, 2018). Furthermore, people's passive online behaviour is being recorded and analyzed, providing insight into their 'digital sentiments' (Edwards & Fenwick, 2016). The result is that the amount of data collected historically throughout the world doubles about every two years and almost all data (98 percent) is recorded digitally (Hilbert & López, 2011; Mayer-Schönberger & Cukier, 2013). Data is currently collected, filtered for relevance, and stored in useful formats in 'data warehouses' on the assumption that all data may be valuable at some point in the future (Daniel, 2015). The assumption is that bigger data is more reliable and therefore there is a strong incentive to collect more and more data even if there is not a current purpose (Prinsloo & Slade, 2014). Second, increasing computing power allows for the first time the thorough analyses of large complex data sets (Giacumo & Bremen, 2016). As per 'Moore's Law', computing power is estimated to roughly double every two years (Shalf & Leland, 2015). Because 'Big Data' uses massive data sets and not representative samples, the unprecedented scale identifies new correlations, extracts new insights, creates new value, and ultimately changes organizations (Mayer-Schönberger & Cukier, 2013). Creating the actionable information, including the linking of different data sets is a very complex task performed using algorithms specifically created and modified for the purpose (Daniel, 2015). Once correlations amongst all the collected data are made, a final analysis must be interpreted, distributed and integrated into organizational decision-making processes (Boyd & Crawford, 2012). For its proponents, 'Big Data' is a pivotal technological achievement that marks a fundamentally new way to quantify and understand the world (Anderson, 2008; Mayer-Schönberger & Cukier, 2013; Mazzocchi, 2015)

‘Big Data’ also has significant potential hazards noted by critics. First, by automating research, the understanding and definition of knowledge changes fundamentally (Anderson, 2008; Frické, 2015). Knowledge and research is no longer driven by theory and causality but by massive data sets and the demonstrated correlations (Mayer-Schönberger & Cukier, 2013). Second, the insights produced by ‘Big Data’ are developed by constructed algorithms with inherent biases which need to be considered in the face of uncritical claims of intrinsic objectivity and accuracy (Boyd & Crawford, 2011). Third, just because data may be technologically and legally accessible, it does not mean that the practice is ethical. People have legitimate concerns that the data they provide may be used without their informed consent and against their interests (Patton, 2000; Prinsloo & Slade, 2014; Slade & Prinsloo, 2013; Wintrup, 2017). Finally, many decision makers fundamentally do not understand ‘Big Data’ and how it works, creating a ‘black box society’ where decision are increasingly justified algorithmically (Pasquale, 2015). An uncritical acceptance of ‘Big Data’ may serve to redefine the nature of organizational authority, providing a scenario where key organizational decisions are made with little or no human mediation (Kitchin, 2014).

2.2 ‘Big Data’ in HE

Data reporting has long been an important part of accountability frameworks in HE and is seen as an effective way to measure performance and justify investments in policy, instructors and instructional approaches (Wagner & Ice, 2012). Data collection at an institutional level is also critical to demonstrating the impact of HE national and international ranking systems (Marope et al., 2013). However, with the increasing adoption of ICT, data collection and analysis are escalating in an increasingly data-rich HE environment (Edwards & Fenwick, 2016; Moreira et al., 2017) . Schools are being progressively turned into data production centers because leaders are beginning to link data not just to reporting student and class performance but to potential cost-efficiencies and tangible improvements in the learning environment (Daniel, 2015; Finn, 2016). Others, like the OECD (2013), even assert that 'Big Data' may be the 'foundation on which HE can reinvent both its business model and bring together the evidence to help make decisions about educational outcomes' (p. 910). In short, 'Big Data' is seen by some as a critical tool to lowering costs, increasing enrollment, efficiency, relevance and learner satisfaction in a sector that is often accused of being resistant to change (Lane, 2016; Long & Siemens, 2011; Manyika et al., 2011; Picciano, 2012).

However, despite the potential benefits, the rates of adoption of 'Big Data' in HE lags far behind many sectors such as retail and health care (Picciano, 2012). Some HE leaders see 'Big Data' not as a solution at all but just another cost and implementation challenge in the HE environment (Menon, 2014). Automated data collection processes through Learning Management Systems (LMS) are already largely viewed as desirable, even mundane, at most HE institutions (Williamson, 2017). HE institutions are coming to rely on LMS data for student retention strategies, yet it remains unclear if or how data on individual students are influencing institutional decision-making (Long & Siemens, 2011). Where previously HE was limited to minimal sets of manually collected data such as attendance or test scores, online LMS record plentiful criteria such as how often and for how long specific learners are accessing instructional materials and what learners are making which particular errors (Mayer-Schönberger & Cukier, 2013). The digitization of education (encoding of education practices in software) and the 'datafication' (exponential growth in data points in education) are complementary and mutually reinforcing factors (Dede et al., 2016). Learning software creates more data and the data are used to create better software (Long & Siemens, 2011). As a result, the development of code impacts all aspects of education from policies to pedagogies and yet, to date, the concept draws only fleeting critical attention (Williamson, 2017). Yet while data are collected relatively easily through LMS, analyses to produce prescriptive measures is a much more complicated process. In particular, the data must be considered through a variety of lenses. Specifically, the interrelations between policy, revenue, fund development, accountability requirements and different understandings of 'Big Data' epistemology make the development and acceptance of prescriptive recommendations from the data a very complex and nuanced process (Boyd & Crawford, 2012; Williamson, 2017).

Daniel (2015) describes a number of different types of learner analytics commonly used by HE institutions. First, 'institutional analytics' makes recommendations from data collected from across a variety of divisions. Second, 'information technology analytics' integrates data (often collected manually) from a variety of systems such as the LMS, the Student Information System (SIS), and a constituent group (such as alumni) Client Relationship Management (CRM) System. Third, 'academic/program analytics' combines large data sets with predictive modeling to measure, interpret and report on student learning and make recommendations related to the academic strength and weaknesses of a particular academic program. Finally, 'learning analytics' uses learner data to optimize learning environments by recommending new processes and workflows that enhance individual learner results.

'Big Data' presents unprecedented opportunities for HE institutions to better serve learners. Based upon a better understanding of an individual learner's needs, customized modules and assignments can be constructed and delivered around personalized learning experiences to meet individual needs and capacities (Baer & Campbell, 2011). The subsequent recommendations as a result of data analyses are dependent upon which of the three broad types of data analysis models that are used. First, 'descriptive analytics' simply identify patterns and trends from data such as the frequency of login, page views, assignment completion rates, etc. (Daniel, 2015). Second, 'predictive analytics' seeks to reveal heretofore unrecognized relationships between criteria such as demographics, completion rates and identified potential risk behaviours (Jindal & Borah, 2015). Third, 'prescriptive analytics' assesses and recommends from a number of options, a best course of action such as requiring learners to take courses in a specific order or scheduling learner interventions if specific benchmarks are not achieved at points throughout the academic year (Basu, 2013).

Implementing 'Big Data' is a complex and often institutionally specific process (Laux, 2017). Streamlined and ethical data collection processes and the selection of valid data sources will vary by HE institution (Finn, 2016). HE institutions also need to find ways to link data from typically detached SIS, LMS and CRM systems. Most importantly, processes for integrating 'Big Data' findings into overall decision-making will vary from institution to institution (Williamson, 2017). 'Big Data' can be the platform for exploring scenario modeling simulations where future outcomes may be predicted (Daniel, 2015).

While the potential 'Big Data' benefits are substantial, there remains significant challenges to its implementation in the HE sector. First, the validity of a purely data-driven over knowledge-driven decision-making process is often not accepted as valid (Daniel, 2015; Macfadyen & Dawson, 2012). Second, the costs associated with data collection, data aggregation and analyses are significant and recruitment of the required IT and data science expertise is a significant level of effort (Jones, 2012). Third, technical capacity to organize data, especially unstructured data, in consistent formats (otherwise known as 'data cleaning') can be a substantial challenge (Picciano, 2012; Romero & Ventura, 2013). Fourth, for 'Big Data' to be fully integrated in an HE institutional decision-making process, a co-created governing structure must be established so that all areas of the organization are focused on policies and strategies related to improving the learning environment and its cost-effectiveness (Macfadyen & Dawson, 2010; Wagner & Ice, 2012). Finally, liability associated with potential data breaches and unethical data gathering processes is a significant factor deterring many HE institutions from fully adopting 'Big Data' in their decision making processes (Jones, 2012).

2.3 Leaders' Perceptions of 'Big Data' in Higher Education

'Big Data' is fundamentally altering how leaders across all sectors understand the value of the data collected (and potentially collected) by their organizations (Mayer-Schönberger & Cukier, 2013). Williamson (2017) discusses how specifically in HE there is a new way of thinking of how processes can be 'datafied' and 'digitized' to increase the value to the organization (p. 9). While there is a widespread recognition of the potential of 'Big Data' in HE, there is significant reluctance to abandoning knowledge-based decisions rooted in critical reflection and judgement instead of demonstrated data correlations (Macfadyen & Dawson, 2012). While HE institutions are collecting data at unprecedented levels, most of the efforts are focused on compliance reporting rather than strategic decision-making (Voorhees & Cooper, 2014). Leaders are also generally sceptical of analyses generated by 'Big Data'. As Silver (2012) discusses, a very wide range of potentially contradictory recommendations is often produced from the same data sets if the 'noise' and the 'signal' is not clearly differentiated (p.

162). HE leaders are thus often presented with vast quantities of data but often have little confidence in what it functionally means to their organization much less in whether it provides a meaningful prescription for future strategy (Picciano, 2012).

2.4 Data Driven Decision Making

Data driven decision making is the use of reliable and timely data, in consideration of institutional values, to inform an action or actions related to an institutional policy or procedure (Picciano, 2012). HE institutions have always collected data such as attendance and learner evaluations. However, now as learners interact with software on a constant basis, millions of new data points are automatically recorded in real time (Williamson, 2017). As a result, more than ever before, data are not only available but often instantaneously available through the HE institution's LMS (Knox, 2016). To get a full view of all the data related to a learner, a series of systems like the LMS, SIS and CRM must be seamlessly linked. Yet the reality is that different technologies are very rarely fully integrated and the complete profile of learner or prospective learner is very rarely available (Orlikowski & Iacono, 2001). Likewise, very few HE institutions can produce all of its learner data across all platforms (much of it previously in analog formats) so it can be compared and analyzed (Gregory et al., 2017).

Given the resource restrictions, HE institutions implementing data-based decision making processes are faced with three critical questions: i) what data should be collected, ii) what methods should be used to collect the data, and iii) how can data best inform planning (Menon, 2014). HE institutions should also be clear on the limits of data-based decision making. In his landmark study, Keller (1983) notes that while data and analysis are critical, so is knowledge-based and educated opinions. Likewise, the purpose of data-based decision making processes should be clear to all stakeholders. There is a fundamental difference between collecting data for publishing performance indicators and attempting to prescribe a specific course of action rooted in learner data. Many leaders understand data-based decision making within the context of performance indicators and the need for HE institutions to justify the investment from various constituent groups with measurable results (Neave, 1988). The rationale behind performance data is grounded in neo-liberal ideas of choice in public services and the operation of HE 'quasi markets' (Deem & Brehony, 2005; Kirby, 2011). While performance indicators are measurable results, they are often 'oversold' as prescriptions to meet strategic plan objectives (Taylor, 2014). In fact, some argue that the purpose of data in many HE strategic plans is to create an illusion of rational decision making when the strategies presented are not actually linked to actionable data, resource allocation or valid performance metrics (Voorhees & Cooper, 2014). In this sense, performance indicators lend credence to the idea that managers are making decisions using evidence and facts when the links may be tenuous (Taylor, 2014).

Performance based indicators provide an avenue for public scrutiny, which in turn improve access to public funding (Ziskin et al., 2014). Demonstrated returns on investment in the forms of higher employability, productivity and taxation help to build the business case for HE funding (Psacharopoulos, 2014). Performance indicators also create a competitive environment which help drive efficiency and create a basis where informed consumer decisions can be made (Kirby, 2011). However, distilling performance to simple measures of quantity means that disciplinary competence is no longer necessary to evaluate performance (Bleiklie, 1998). As a result, a 'new managerialism' based upon data-based performance management over subject matter expertise, a concept originating in the corporate sector, is starting to heavily influence how HE institutions are managed (Taylor, 2014)

Data-driven organizations must convincingly fulfill the following three critical roles in order to be successful: i) ensuring data integrity, ii) brokering raw data and transforming them to actionable information, and iii) identifying applicable data and applying them to management solutions (McLaughlin et al., 1998). In addition to the simple data accuracy, data can be compromised by methods used to quantify qualitative measures, a lack of meaningful correlation between results and indicators, and errors of interpretation through illogical inferences (Taylor, 2014). Some characteristics of successful data-driven HE institutions include: i) a strong level of commitment and trust in the data at all levels of the organization, ii) engagement on decision-making with faculty, staff, administrators and students so that it is clear that data is everyone's concern not just the few decision makers, iii) a culture that values both inquiry and evidence, and iv) a firm institutional commitment to ongoing professional development in order to help drive innovation (Voorhees & Cooper, 2014). Yet, despite the growing emphasis upon data-driven decision making, the literature emphasises that data must be integrated with critical reflection, reasoned debate and judiciousness (Keller, 1983; Taylor, 2014).

2.5 Epistemology of 'Big Data'

The logic of 'Big Data' challenges many traditional epistemologies (Dede et al., 2016; Frické, 2015; Mazzocchi, 2015; Williamson, 2017). Epistemology often makes an important distinction between correlation and causation. Theory explaining and predicting cause and effect is central to the concepts of knowing and understanding (Kitchin, 2014). However, 'Big Data' asserts that, provided with enough data, correlation is enough to claim knowledge and solutions can be found without hypothesis (Anderson, 2008). If the correlation can be demonstrated frequently enough, a phenomenon can be predicted without coherent models, unified theories or any explanation at all (Frické, 2015). As Mayer-Schonberger & Cukier (2013) assert, the growing acceptance of 'Big Data' epistemological 'messiness' is embedded in its proven predictive power, overturning beliefs held over centuries on how reality should be comprehended (p. 7). As stated by Anderson (2008), 'With enough data, the numbers speak for themselves.' (p. 4).

However, it is critical to realize that 'Big Data' uses data sets and algorithms that are created by humans and have inherent biases (Crawford, 2013). Just because an inference is proven accurate using a huge set of data, it is still an inference (Kitchin, 2014). How data are analyzed and interpreted is shaped by business plans, objectives and professional values and cultures (Kitchin & Dodge, 2011). The key questions are: who is running the software, what are their agendas and how are inherent biases and objectives shaping the future of education (Boyd & Crawford, 2011; Williamson, 2017)?

Despite the epistemological criticism, 'Big Data' is assuming an increasing role in educational research which is increasingly shaping how we know and understand education through digital data collection and analysis (Cope & Kalantzis, 2016). Some researchers encourage the rapid development of 'actionable data' through 'Big Data' (Voorhees & Cooper, 2014). Compared to the descriptive, predictive and prescriptive power of 'Big Data' some theorists consider traditional data-based decision making as relatively unactionable or 'wallpaper' data (Voorhees & Cooper, 2014). For example, institutional 'fact books' are considered as a priority and have high degrees of validity in many data-driven HE institutions (Wook et al., 2017). However, access to many 'fact books' are restricted to senior management and used not for decision making but primarily for government compliance purposes (Voorhees & Cooper, 2014). If the decision making process is for just a few, then most are alienated from the decisions and will not take ownership (Voorhees & Cooper, 2014). Moreover, by restricting access to the data, most of the HE institutions' stakeholders are alienated from the decisions and feel no ownership or responsibility for success (Dede et al., 2016; Williamson, 2017).

Theorists also point to the significant hazards of an automated 'Big Data' decision making processes. In his Noble prize wining work, Simon (1982) discusses the limits to rationality in organizational decision making and points out that '…a wealth of information creates a poverty of attention' (p. 48). Data-driven means that decisions are informed by data and that data do not replace rational consideration (Picciano, 2012). In this light, there is growing interest for researchers on how algorithms function (Ziewitz, 2016). The political bias of algorithms is noted by many theorists and the widespread implications to the validity of 'Big Data' findings (Crawford, 2013; Gillespie et al., 2014). Others, while mindful of bias, encourage a cautious development of 'Big Data' capacities in HE while remaining critical of data collection and analyses processes (Voorhees & Cooper, 2014)

2.6 Organizational Learning / Change

The adoption of 'Big Data' is a significant organization change, requiring an adaptive capacity that Senge (1997) calls 'organizational learning'. A reluctance within an organization to change or learn can be disabling, particularly within competitive environments (Heifetz et al., 2009). Learning that is required to survive in a competitive environment is what Senge (1997) calls 'adaptive learning'. Senge (1997) makes an important distinction between 'technical learning' where current knowledge and resources may be applied to a particular problem, and 'adaptive learning' where there are no clear historical answers. 'Adaptive learning' requires experimentation and is inherently risky. A further distinction is made for what is called 'generative learning', learning that promotes not only organizational survival but enhances the overall organizational ability to create (Senge, 2006). Ironically, a generative learning environment in competitive organizations can be difficult to create because of a typical emphasis on individual responsibility within the organizational hierarchy over individual or group creativity and innovation. A serious challenge is presented when only those with the most organizational authority are permitted to exhibit adaptive or generative learning behavior because, according to Senge's (1997) 'systems dynamics' theory, leaders tend to focus upon addressing immediate problems and not the long term systemic issues. Generative learning is relatively slow to evolve and develop from shared experiences throughout the organization. As a result, Senge (1997) promotes the re-definition of the relationship between the individual and the organization to a much greater emphasis upon systems thinking, where the ongoing interrelationships between various parts of the system is the priority over snapshot reporting against isolated benchmarks. Senge (2006) notes that ICT is ideally suited to systems analysis across the entire organization and that hypothetical modelling can be a very powerful tool for generative learning within a dynamic system. ICT is thus a central tool in what Senge (1997) calls 'team learning' or the 'fifth discipline' because it can be used to effectively create group comprehension beyond individual perspectives and achieve organizational results based on a shared vision.

Employee engagement, particularly among faculty and administration, is a critical element to any significant HE organizational change (Baer et al., 2015). The collection and interpretation of data in a ‘Big Data’ initiative is dependent upon employees that are purposefully engaged in the research question (Dowd, 2005). The process is more than the selection between two choices and about collaborating with all employees to recognize the options going forward (Picciano, 2012). Leadership must build the bridge between ‘innovation and empirical rigour’ by simultaneously nurturing a culture of inquiry while create high expectations of performance (Lafley et al., 2012). Strategy is the central tool to create a culture of inquiry by generating the need to collect/analyze data and re-imagine processes (Reid & Dold, 2018). A good plan creates many learning opportunities and sets a framework for collaboration among stakeholders and organizational data must be integrated in any sustained organizational change (Senge, 2006).

2.7 Ethical Considerations for ‘Big Data’

Critics raise a number of important issues related to the ethical use of ‘Big Data’, whether one is considering how data are collected or how decisions are ultimately made based upon the data (Patton, 2000; Prinsloo, 2009; Scholes, 2016). Data owners know very little about how personal data are collected yet many share their personal data willingly (Prinsloo & Slade, 2014). Critics note that people often agree to forms of online surveillance and management because of a trade off for greater convenience, efficiency and productivity (Kitchin & Dodge, 2011). Studies show that learners are usually very conservative in allowing HE institutions access to personal information (Ifenthaler & Schumacher, 2016). However, once a learner is enrolled, much of their digital information are collected passively through the institutional LMS and/or SIS (Romero & Ventura, 2013). Students may not know what information are being harvested for what purposes (Prinsloo, 2009)

Instead of the current practices for online data gathering, many theorists propose a moral approach that embraces an 'ethics of care framework' (Peters, 2013). A moral approach is suggested instead of a legal framework because it puts a greater responsibility for authority figures with fiduciary responsibilities (Patton, 2000). Within the framework a number of principles are proposed. First, data gathering processes should not focus on what is effective but on the most appropriate and desirable outcome (Biesta, 2007). Second, learners should be better aware of the data they are sharing and how it may impact them academically. Similarly, HE institutions should recognize that learning does not always occur at an even pace: the data gathered at a static moment in time of a student's academic record may not be representative of their capacity (Stoddart, 2012). Third, data collected need to incorporate temporal constructs that will allow learners to evolve and not have previous poor performance follow them like a 'digital tattoo' throughout their HE institutional record (Mayer-Schönberger, 2009; Olszak & Mach-Król, 2018). Finally, learners need transparency and should know what data are being collected, for what purpose, and how they will be protected from data breaches. Ideally, learners will be given the option for opting in or out of data collection (Prinsloo & Slade, 2014).

In terms of data analytics, critics propose a number of issues related to the ethical use of algorithms. Many critics argue that algorithms are not some disinterested knowledge producing mathematical equation (Boyd & Crawford, 2012; Kitchin, 2014; Pasquale, 2015; Williamson, 2017; Ziewitz, 2016). Algorithms are not separate from social forces but a product of them, influenced by social bias and constantly being redesigned (Beer, 2017). Due to the human development of algorithms, people should be aware of the biases and politics that influence them (Berry, 2014; Solove, 2004). Algorithms can be very powerful and can produce recognitions of normality and abnormality and even influence social comprehensions of 'truth' (Beer, 2017). In fact, as automation increases, it is critical for stakeholders to know how algorithms affect organizational decision making, how they take over choice in governance and how they have a tendency to 'nudge' decision makes in a particular direction (Yeung, 2017)?

Recognizing the inherent bias in the development of algorithms in learning analytics, Prinsloo & Slade (2014) offer guidelines for algorithm development and use. First, following the principles of 'duty of care', learners should have meaningful opportunities to confirm the algorithm derived conclusions. Secondly, similar to data collection processes, learning analytics should be temporal in nature to allow for the non-linear nature of learning (Olszak & Mach-Król, 2018). Finally, learners should be active participants in algorithm development process and not passive suppliers of data (Kruse & Pongsajapan, 2012).

2.8 Competition in Higher Education

In a 2018 survey of 618 HE institutional presidents in the United States, thirteen percent reported that there was a good chance that their institutions could close in the next five years (Mrig & Sanaghan, 2018). Significant challenges in the HE sector includes shrinking enrollments from learners in traditional demographics combined with decreasing public support (Barber et al., 2013). Likewise, governments are less eager to enforce programing monopolies and HE institutions are more free than ever to compete in overlapping non-credit, certificate, diploma, degree and post-degree spaces (Piché & Jones, 2016; Skolnik, 2016). Program delivery through ICT is also eroding previously enjoyed geographic monopolies (Bagley & Portnoi, 2014). The vertical and horizontal integration of the HE sector not only intensifies competition for learners but also for donors with interests in funding the same academic and/or geographic space (Hodson, 2010). Despite increasing costs, frequent tuition caps imposed by governments serve to exacerbate the funding crisis at many HE institutions (Axelrod, 2002). While it is quickly becoming clear that the business model that built the HE sector in the latter half of the 20^{th} century is broken, it is still uncertain what model or models will replace it (Menon, 2014).

Despite the unsustainability of the current HE sector business model, many institutions and governments are addressing their challenges by 'doubling down' on historical tactics with new investment in traditional offerings (Mrig & Sanaghan, 2018). With limited to no funding increases from governments and/or endowments, growth in enrollment and learner tuition is seen one of the few options for HE institutions (Kirby, 2011). However, tuition typically recovers only a portion of the overall per student costs. Non-instructional overheads such as capital costs, learner supports services, financial services, and registration services are typically funded by government grants and/or endowments. Therefore, in the traditional HE sector business model, available overhead resources severely limit HE institutions ability to sustainably increase enrolment over the long term.

In contrast to the 'business as usual' response to the crises in the HE sector, some institutions are taking more strategic approaches. First, instead on focusing on the shrinking learner market direct from secondary school, HE institutions are realizing the potential market size of the 'lifelong learner'. The world of work is changing so rapidly that people require education and training not just at the beginning but throughout their career or careers. It is estimated that sixty-five percent of learners entering primary education will work in a job that does not yet exist (Ritacco & McGowan, 2017). As a result, forward thinking institutions engage with learners throughout their career rather than 'front-loading' all learning at the beginning of their careers. Second, HE institutions are innovating to both improve relevancy and accessibility to learners and lower costs. Some HE institutions are experimenting with Artificial Intelligence (AI) delivering instructing and tutoring services (Bayne, 2015). 'Big Data' is becoming a critical part of not only increasing 'livelong learner recruitment' capacity but also decreasing attrition through learner analytics. Effective data collection and management also helps HE institutions manage other critical relationships with donors, employers and regulators (Marope et al., 2013). While HE institutions have long reported data in compliance with government requirements, there are a growing number of HE institutions (particularly in the United States) that a now using data less for public accountability are more to provide competitive advantages (Burke & Minassians, 2001). HE sector decision-makers are prudent to be cautious in the face of significant costs and sometimes exaggerated claims of 'Big Data' (Booth, 2012; Long & Siemens, 2011; Wagner & Ice, 2012). However, given the power demonstrated in other sectors, it is clear that ethically used 'Big Data' tools can tailor learning environments to individual needs and reduce costs which are both central elements to a sustainable business model in the 21st century HE sector (Picciano, 2012).

2.9 IT Governance in HE

IT governance is the processes used by institutions make decisions and allocate resources to ensure that organizational strategy is achieved and sustained (De Haes & Van Grembergen, 2008; Weill & Ross, 2004). Specifically, in the HE sector, organizations use many and often heterogeneous technologies and processes. It is the governing bodies' role to ensure that new technologies and processes are effectively adopted, integrated and sustained (Bhattacharjya & Chang, 2006). For complex and decentralized organizations such as HE institutions, IT governance is further complicated by many relatively independent actors compared to most other organization types (Fattah & Setyadi, 2021; Isaías Scalabrin et al., 2020). It is also noted that IT plays an increasingly crucial role in many vital functions within HE such as learning systems and research productivity (Wu et al., 2015). Yet, despite its critical role and complexity, empirical studies and analyses specific to IT governance at HE institutions are relatively scarce (Isaías Scalabrin et al., 2020; Jairak et al., 2015; Khouja et al., 2018). Nevertheless, a few common themes emerge from the limited literature that focus directly upon the HE sector. First, while governance structures vary significantly among HE institutions, facilitating organizational goals and maximizing benefits while concurrently mitigating risk are the primary roles of IT governing bodies (Bichsel & Feehan, 2014). Second, despite shared broad objectives, contextual factors vary greatly from institution to institution so specific governance procedures vary widely (Weber et al., 2009). Third, IT decision making must be driven by clear organizational priorities so that competing investment opportunities can be evaluated against their costs and potential benefits (Isaías Scalabrin et al., 2021). Finally, governance models need to confirm their accountability though clear monitoring and control procedures to ensure maximum contribution to organizational goals that are often in a state of flux (Bichsel & Feehan, 2014; Isaías Scalabrin et al., 2021). Yet despite the organizational need to maximize performance, risk mitigation and the effective coordination of numerous projects is a central, if not predominant, theme in IT governance in the HE sector.

2.9 Summary

This chapter examines focal elements influencing HE leaders considering the merits of 'Big Data' technology adoption. 'Big Data' is the ICT tools used to translate large amounts of complex data to organizational insights and recommended future actions. The growth of digital data gathering, through SIS, LMS and CRMs in particular, has exponentially increased HE institutions' capacity to collect and analyze learner and other stakeholder data. Yet, given the experience with other ICT tools, skepticism remains that HE institutions are able to harness the potential of 'Big Data' in a cost-effective manner. Moreover, HE leaders may be reluctant to abandon traditional 'knowledge-based' decision making processes for the data-driven and correlation-based decision making inherent to 'Big Data'. Likewise, HE leader concerns over the ethical collection and use of the stakeholder data collected also inhibits 'Big Data' adoption. However, the increasing pervasiveness of ICT and the progressively more competitive nature of HE may propel a 'systems dynamic' that supports 'Big Data' adoption despite HE leaders' reservations. The next chapter discusses the proposed methodology of the study and how the literature review will be incorporated with a constructivist grounded theory approach to the research problem.

3.0 Chapter Three: Research Design

This chapter describes how the research problem is addressed using a constructivist approach to grounded theory.

3.1 Overview

The purpose of the study is to develop a conceptual framework that describes and explains the adoption of 'Big Data' technology in the Alberta HE sector. Grounded theory is ideally suited for developing theories or explanations of human or social processes, especially when existing concepts do not effectively address the specific research problem (Bryant & Charmaz, 2007; Creswell & Gutterman, 2019; Denzin & Lincoln, 2018). Grounded theory is a qualitative research method that anchors collected data to explanations of human actions and interactions (Charmaz, 2000). Grounded theory methods have particular utility when attempting to develop theory about human action or interaction when existing conceptual frameworks either do not exist or prove inadequate over time (Denzin, 2007). Unlike earlier grounded theory iterations from Glaser and Strauss (1967) and Strauss and Corbin (1990) which assume a positivist or post-positivist epistemology, the constructivist approach to grounded theory focuses upon individuals' perceptions, beliefs, assumptions and ideologies rather than the discovery of universally defined or replicable 'truth' (Creswell & Gutterman, 2019). Likewise, the constructivist approach to grounded theory acknowledges that the values, experiences and priorities of the researcher influence the construction of research conclusions (Bryant & Charmaz, 2007; Charmaz, 2014; Denzin, 2007). Through systematic and rigorous analysis of the data collected, constructivist grounded theory explains the meanings and assumptions individuals have developed within a collective process or processes. Unlike positivist or post-positivist approaches that often propose categorical and conclusive findings, the constructivist approach produces conclusions that are more contextual and flexible (Charmaz, 2000; Glaser & Strauss, 1967; Strauss & Corbin, 1990). Moreover, the constructivist approach to grounded theory methods, as discussed by Creswell & Gutterman (2019), focuses less on generalizable findings and is 'more explanatory, more discursive and more probing on the assumptions and meanings for individuals' (p. 441).

Interviewees are mid to senior leaders at three HE institutions in Alberta. After ethics approval from the University of Calgary Research Ethics Board (REB) and the other relevant HE institutions, interviewees were selected after sharing the semi-structured interview brief and questions. Seven semi-structured interviews were required before the 'saturation' point was reached and the subjective judgement was made that more data would not provide significant new insight. Interviews provided the verbatim text for 'theoretical sampling' from data coding and analysis (Charmaz, 2014). Theoretical sampling was collected in the semi-structured interviews. The data was coded by the researcher as common themes emerged (See Section 7.2 Data Codes). The coded data was analyzed and the proposed conceptual framework constructed from the data. Utilizing 'emergent design' processes, data collection and theory development occured concurrently (Creswell & Gutterman, 2019). Data analysis informed the ongoing data collection process until the 'saturation' point was achieved (Bryant, 2007). The process emphasized reflexivity with the researcher recording memos to promote an ongoing reflection (Charmaz, 1990). Finally, the emerging theory was validated with the research subjects in order to assure research validity and to validate the data provided.

3.2 Research Problem Context

The study examines why, despite many significant potential benefits, HE as a sector is a relative laggard in the adoption of 'Big Data' (Daniel, 2015). HE institutions face unprecedented challenges that include decreasing public funding, escalating competition and significant fluctuation of learner demographics and expectations (Kirby, 2011; Livingstone & Guile, 2012; Usher, 2018). The growing ICT pervasiveness creates a paradox for HE leaders. ICT continues to intensify competition among HE institutions as technology erodes former geographic monopolies. However, ICT's increasing prevalence also potentially multiplies HE institutions' pool of potential leaners. Furthermore, ICT also creates an unprecedented opportunity to automatically collect, store and analyze learner and other stakeholder data. The previously unknown data can be used to discover new ways of increasing value and cost efficiency (Calhoun & Kamerschen, 2010; Clark, 2018; Williamson, 2017). For example, the same way that Google can reliably and in real time predict and monitor traffic flow by analyzing a massive amounts of location data and geography over time, HE institutions can detect when learners need a specific intervention in order to succeed academically from online behavior (Mayer-Schönberger & Cukier, 2013). However, 'Big Data' technologies also raise a number of concerns related to privacy and fairness (Solove, 2004). Additionally, it is unclear if HE leaders are ready and willing to abandon 'knowledge-based' management approaches based upon causality in favour of decisions informed by 'Big Data' technologies' correlation-based reasoning. It remains unclear what is motivating HE leaders' evolution either towards or away from the adoption of 'Big Data'.

3.3 Research Problem and Questions

The purpose of the study is to develop a conceptual framework that describes and explains the adoption of 'Big Data' technology in the Alberta HE sector. The following research questions are key to the investigation: i) what do higher education leaders in Alberta consider the main elements for and against 'Big Data' adoption or further development?; ii) to what extent are Alberta higher education leaders motivated by competitive enrolment and/or cost-effective learning considerations that could be developed by 'Big Data' applications?; iii) do Alberta higher education leaders believe that meaningful business decisions can or should be influenced or even made by 'Big Data' applications; iv) what ethical considerations do Alberta higher education leaders integrate in the data collection and analyses processes of 'Big Data' initiatives? and v) how able do HE leaders consider their respective institutions to implement 'Big Data' successfully from technological and change management perspectives?

3.4 Assumptions

The study is designed with the following assumptions. First, that HE leaders engaged in the study have experience with the applications or potential applications of 'Big Data". Second, that interviewees will discuss the phenomenon in candid and direct ways. Third, that the interviews will yield a 'saturation' point (of no new significant data) after a reasonable number of interviews.

3.5 Role of the Researcher

For over the past twenty-five years, the researcher has held a variety of management responsibilities in Canadian colleges, including domestic and international student recruitment, the design and delivery of continuing education and contract training, international development projects, fund development and alumni engagement. Most recently, the researcher has been responsible for the development and implementation of an institute-wide CRM solution to foster and monitor enhanced engagement with learners, prospective learners and employer stakeholders. The researcher has come to recognize the unprecedented potential of 'Big Data' to improve services to learners and other HE stakeholders but also the risks inherent in the adoption of a costly new technology and the new ethical challenges presented by pervasive ICT. The researcher has also observed a marked reluctance by many HE leaders to forsake 'knowledge-based' decision-making despite prevalent rhetoric affirming the need for 'data-driven' processes and the demonstrated value of the correlations (but not causation) uncovered by 'Big Data.'

3.6 Researcher Ontology

The researcher is situated in a social constructivist philosophy. Reality from a social perspective has both subjective and objective co-existing elements. At one level, reality can be constructed through individual cognitive processes. However, individual perceptions of reality are heavily influenced through social interaction and collective understanding. As a result, individual and communal perceptions of reality are socially and contextually situated and play a central role in the research process, heavily influencing what and how data are collected and analyzed. The researcher's ontology is commensurate with a constructivist approach to grounded theory where the subjective meanings and values of the research participants and the researcher co-create research suppositions rather than the more positivist and post-positivist ontologies of either Strauss and Corbin's (1990) 'systematic' or Glaser's (2001) 'emergent design' grounded theory varieties.

3.7 Researcher Epistemology

The researcher's epistemology is firmly rooted in the constructivist conception that knowledge, particularly in the social sciences, is a human and social construction rather than the objectivist view that there is a single objective 'truth' that can be known. The constructivist concept of multiple valid 'truths' is commensurate with pragmatic or 'bricolage' approaches to research methods (Lincoln et al., 2017). Likewise, the concept of socially constructed knowledge emphasizes the importance of reflexivity from the researcher. Constantly considering other ways to interpret data and presenting conclusions as negotiable constructs rather the final results are central elements to a constructivist epistemology (Bryant, 2007; Charmaz, 2014). Because knowledge is overwhelmingly a social construct, the researcher acknowledges the impact of social context in the research process. Unlike radical subjectivism that proposes that knowledge is determined individually, the constructivist epistemology requires shared interpretations that lead to understanding (Lincoln & Guba, 2013). In short, while the researcher acknowledges a subjective epistemology, subjectivity must be supported by the social context.

3.8 Grounded Theory Methodology

Grounded theory methodology is a series of general principles and strategies rather than a strict or prescriptive process (Atkinson et al., 2003). The flexible guidelines of grounded theory are commensurate with a wide range of differing ontologies and epistemologies (Bryant & Charmaz, 2007; Charmaz, 2014; Denzin, 2007). Despite pragmatic application and a wide variety of individual processes, grounded theory has a number of important elements common to its first iteration in the late 1960s (Glaser & Strauss, 1967). First, data collection and analysis are conducted concurrently and influence each other in a constant reflexive process that categorizes collected data and identifies gaps. Second, new theory development emphasizes inductive reasoning rooted in data rather than the demonstration of a preconceived deduced hypothesis. Third, sampling of research subjects is done in a way to promote new theory development and not for population representation. Although the above elements are applied in many types of research problems and approaches, the purpose of constructing new abstract theoretical explanations for social processes remains the same. Most importantly, grounded theory establishes a rigorous and credible analytical framework based upon qualitative data that has proven valid not just for describing but *explaining* phenomenon (Bryant & Charmaz, 2007; Charmaz, 2014; Denzin, 2007; Lincoln & Guba, 2013).

Criticisms of grounded theory question the method's rigour, often from a positivist or post-positivist ontological perspective. Traditional positivist perspectives that prioritize quantitative data argue that grounded theory lacks rigour and includes incongruous procedures such as combining the data collection and analysis research stages. Underlying much of the criticism from quantitative researchers is a general lack of acceptance that knowledge can be generated from anything other than quantitative data (Charmaz, 2014). Often criticisms of grounded theory derive from other spheres of the ontological and epistemological spectrum. Post-modern and critical theorists assert that grounded theory relies too heavily on the authoritative influence of the researcher in the research process and grounded theory is too often uncritical of the dominant meta-narratives on scientific truth, human nature and world views (Ellis, 1995; Richardson, 1993).

3.9 Constructivist Grounded Theory

Partly in response to the criticism of grounded theory as a whole, a new constructivist approach has developed that encourages researchers and research participants to take an active role in co-creating knowledge and to reflexively contemplate the researcher and research participants’ views and biases in the construction of new knowledge. A constructivist approach acknowledges that ‘reality’ is constructed within the research participants’ and the researcher’s contexts (Clarke, 2005, 2007). Furthermore, the constructivist approach recognizes through reflexivity the researcher’s subjectivity in the collection and interpretation of the data can be acknowledged and governed appropriately (Lincoln & Guba, 2013). The constructivist approach recognizes that research is co-created among the researcher and research participants. The researcher must be aware (as much as possible) of potential sources of conscious and unconscious bias specific to the research context (Charmaz, 2014).

The constructivist approach to grounded theory emphasizes a pragmatic flexibility yet there are characteristics consistent throughout the range of practices. First, data collection and analysis occur simultaneously in an iterative process. Second, an emphasis upon inductive abstract reasoning rooted in data is applied extensively throughout the research process to stimulate new theory development. Third, comparative methods are used to identify common themes and gaps in the data collected. Fourth, perspectives and biases of the researcher are acknowledged as the study is co-constructed with research subjects. Finally, the constructivist approach to grounded theory focuses on explaining the meaning that stakeholders assign within a collective process.

3.10 Preliminary Conceptual Considerations

Grounded theorists are divided on the research utility of a literature review. Early iterations actively discouraged literature reviews for inhibiting new inductive reasoning based only upon open-minded reflection of the collected data (Glaser, 1978; Glaser & Strauss, 1967). Later grounded theory designs considered that researchers should utilize literature and knowledge from the researchers’ professional practice as long as the research findings remain rooted in the collected data and not the researchers’ preconceptions (Strauss & Corbin, 1990). The constructivist approach to grounded theory, however, challenges the positivist and post-positivist ontology that there is one objective theory to be ‘discovered’ (Charmaz, 2000; Lincoln & Guba, 2013). Instead, the researchers’ perspective should be explicitly acknowledged and critical reflection consistently applied to prioritize data-driven inductions over biases derived from previously acquired knowledge (Bryant, 2007). The objective is not to ignore previous knowledge or expertise but instead to critically engage previous knowledge with the collected data (Dey, 1999). Constructivist grounded theorists propose the use of ‘sensitizing topics’ as preliminary ideas to raise initial issues with research participants that can be easily amended or abandoned if proven unwarranted as the data collection and analysis proceeds (Blumer, 1969; Charmaz, 2014).

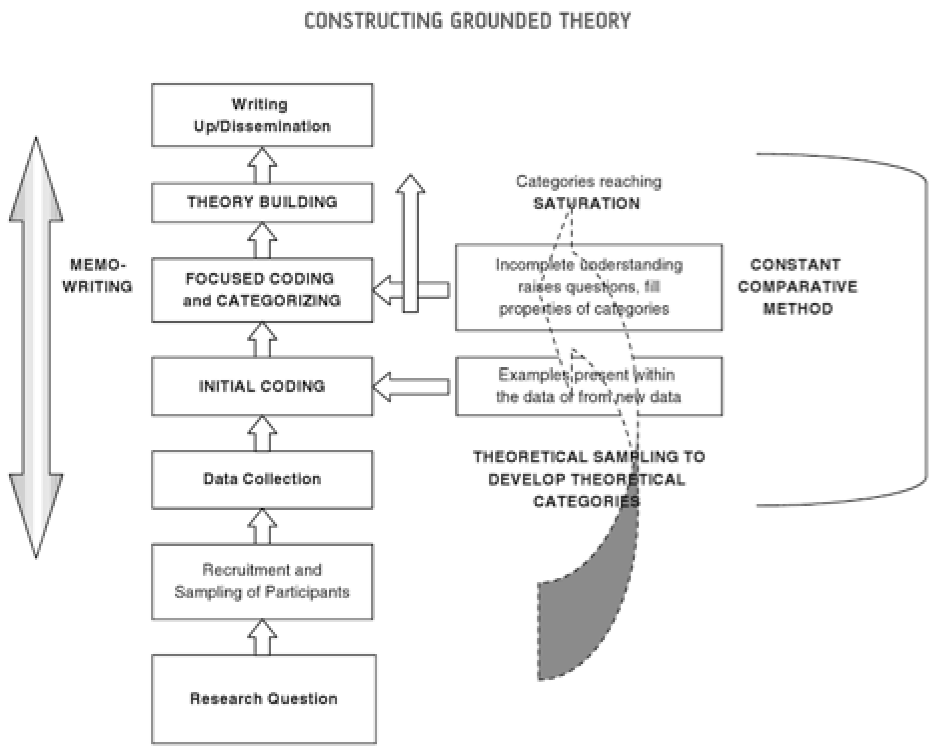
The initially identified 'sensitizing topics' for consideration included the following concepts found in literature review. First, competition for public investment and, especially, for learner tuition fees is increasing and potentially driving the adoption of technologies such as 'Big Data' to increase cost-competitiveness (Kirby, 2011; Williamson, 2017). A leader at an HE institution with a greater funding reliance upon recruiting and retaining learners is presumably more likely to be a 'Big Data' early adopter. Second, an HE leader with an epistemological perspective that prioritizes the more correlative and results-based approach of data-driven decision making over the causality focused methods of knowledge-based decision making is likely more open to 'Big Data' adoption (Anderson, 2008; Daniel, 2015; Frické, 2015; Mayer-Schönberger & Cukier, 2013). Third, an HE leader with a greater overall confidence in ICT to deliver cost-effective solutions instead of costly experiments that fail to live up to early promises is likely to lead to early adoption (Jokonya, 2015; Menon, 2014; Olszak & Mach-Król, 2018). Finally, an HE leader who acknowledges and effectively addresses the important ethical issues related to automated data collection and analysis is more likely to adopt and sustain 'Big Data' practices (Boyd & Crawford, 2012; Kitchin & Dodge, 2011; Patton, 2000; Prinsloo, 2009).

3.11 Research Process

The research project strictly adhered to the requirements of the Research Ethics Board (REB) of the University of Calgary and the relevant ethics requirements of other participating HE institutions. Grounded theory is not a series of prescriptive steps or processes but is instead a collection of principles, potential strategies, and heuristic tools that can be utilized to 'ground' analyses and new theory in the research data (Atkinson et al., 2003). One of the fundamental standards of grounded theory is that analyses and theory development should be conducted concurrent with data collection (Bryant & Charmaz, 2007; Glaser & Strauss, 1967). The data collection was an iterative process and was modified as the abstract theoretical understandings of the phenomenon developed. Consistent with grounded theory, constant reflexivity in the data collection process was maintained in order to follow emerging empirical evidence and promising inductive lines of theoretical reasoning (Bryant, 2007). Likewise, the researcher engaged in the research setting in order to develop a better understanding of the research participants' contexts (Charmaz, 2014). Consistent with grounded theory, the collected data were constructed through observations, interactions and other materials assembled. However, unlike other methods, grounded theory, the study recognized that decisions on the type of data to be collected can change over time as the potential analytic avenues develop (Denzin, 2007). Indeed, instead of reliance upon the description or application of current theories of more recognized researchers, new theory construction through collected data is grounded theory's defining characteristic (Charmaz, 2014).

As per *Figure 1 Constructing Grounded Theory* from Tweed and Charmaz (2011) below, data collection, analysis and theory building are concurrent processes. Constant reflection and adjustment in the data collection tools is required as the theory begins to transpire inductively from the coded data and the researchers' observational memos. Gaps in understanding serve to refine data gathering instruments on an ongoing basis. The constant data comparison yields observed similarities and differences that inductively constructs theory. Once no new data categories relevant to the research problem are identified in the interview process, the 'saturation' point is reached and further interviews are not required for theory development (Charmaz, 2014). Conversely, more data is required when interviews yield surprising results or when constructing particularly complex theory.

Figure 1 Constructing Grounded Theory



3.12 Research Participants

The research was conducted with research participants who are leaders at Alberta HE institutions using a qualitative non-probability or 'purposeful' sampling method (Creswell & Gutterman, 2019). The participants were from HE institutions that represent a cross-section of institution types in Alberta such as CARIs, BASIs, PIs and CCCs. Experience, positive and/or negative, with 'Big Data' technologies was required. Informed consent and formal agreements was sought and granted with each research participant selected from HE leaders who represent a variety of roles with respect to enrolment, budget, ICT integration, and overall decision-making processes. Research participants ranged from manager to vice-president roles. Research participants were recruited individually as needed and as the data was collected and analyzed. Research participants with specific positive and negative experiences were recruited to explore issues as they arose in the analysis. Participants were required to be available and willing to discuss their experiences and perspectives and be prepared to respond to follow up on unanticipated areas of inquiry. Research participants were given the opportunity via email to first correct or amend the verbatim interview record and the later to provide feedback on the initial findings of the study. All the research participants participated in the interview record corrections and about half provided further initial findings feedback through email and telephone interaction. The researcher's network of HE leaders in Alberta provided the starting point for identifying appropriate research participants. A 'snowball' approach identified further research participant candidates with the final selection made based upon: i) experience (positive and negative) with ICT and 'Big Data' in particular, ii) decision making authority at a variety of Alberta HE institution types, and iii) variety of role types within HE institutions (Creswell & Gutterman, 2019).

3.13 Data Collection

Through collecting and analyzing qualitative data that recounts leaders' perspectives, the research project developed a conceptual framework for the adoption of 'Big Data' in Alberta HE. The study used a constructivist approach to grounded theory as outlined by Charmaz (2014) to guide the research process and ensure the conceptual framework developed was rooted in the data collected. Grounded study methods were particularly apt for establishing abstract theoretical explanations of relatively new or unexamined social processes (Creswell & Gutterman, 2019). Individual participant interviews were conducted with academic and IT decision makers at Alberta HE institutions. The number of interviews conducted (seven) was dependent upon the point of 'saturation' in grounded theory – where new data does not yield new data categories under analysis (Charmaz, 2014).

From February to August 2020, semi-structured interviews were conducted to compile 'rich data,' consisting of accurate descriptions of participants' feelings, intentions and actions within the study's context (Charmaz, 2014). Furthermore, data collected and analyses from previous interviews informed subsequent consultations. In the grounded theory method, the researcher is not to be a passive receiver of information but an active participant in the data construction (Charmaz, 2014; Glaser, 1978; Glaser & Strauss, 1967). However, the researcher must remain systematically reflexive throughout the research process and persistently conscious of potential sources of bias (Denzin, 2007). Similarly, the researcher must remain aware of how the research participants' perception of the researcher may affect the data gathering process. A wide variety of criteria such as a perceived insider/outsider status or protagonist/antagonist stance as well as cultural settings and, in some cases, the need to ensure institutional confidentiality could affect the data collected (Charmaz, 2014). Early iterations of grounded theory proposed a strict limiting of prior knowledge of the phenomenon under study, often eschewing literature reviews prior to data collection (Glaser & Strauss, 1967). However, other elements of grounded theory encourage a review of existing relevant research and incorporating reflexivity throughout the research process in order to minimize researcher bias (Dey, 1999).

Interviews used a broad grounded theory approach and started with general open-ended questions that concentrate on the research participants' views and actions related to the adoption (or lack thereof) of 'Big Data' in their HE institution. In these early stages of interviews, the researcher takes a more passive role in the background while exploring the participants' experiences and concerns (Charmaz, 2014). It was important to build rapport and trust with the research participants. Themes emerging for each interview were verified with research participants to ensure accuracy and help to build trust with the researcher. Initially, 'sensitizing topics' from the literature review were used to initialize the discourse (Blumer, 1969). Future interview topics were added, amended and/or discarded as its usefulness to the theory development was assessed (Charmaz, 2014). The priority was to elicit the participants' stories and let the participants describe what is occurring. From the verbatim transcriptions, the interviews were initially coded (see Section 7.1 Data Codes & Memos) with short labels that name and summarize the data and facilitate sorting and analysis (Miles et al., 2014). In some cases, as noted in Chapter Four: Data Collection & Analysis, the verbatim transcription was edited to protect the identities of the research participants and their institutions. Key criteria addressed in the initial coding included: i) what is the data referencing?, ii) what does the data suggest?, iii) from who's point of view?, and iv) what theoretical category does it belong? (Glaser, 1978; Glaser & Strauss, 1967). Concurrent to the initial coding was the development of preliminary analytic notes or 'memos' by the researcher in order to define the relationships between data categories and identify gaps in the data (Charmaz, 2014). Codes and memos guided the development and focus of future interviews. As the interview process progress, the categories begin to emerge within the theory development and the researcher may take a more prominent role to pursue evidence, or lack thereof, to a particular line of theoretical direction. Data collection continued to the point of 'saturation' where the researcher can reasonably conclude the data collected: i) portrays the full range of contexts and participant views under study, ii) reveals the social and theoretical processes occurring beneath the surface, and iii) provides comparisons among the instances that the phenomena is experienced (Charmaz, 2014).

3.14 Preliminary Interview Guide

The preliminary interview guide includes a summary document defining 'Big Data' in HE and the following broad questions for discussions:

General:

1. What is your role at your institution?
2. Describe your experience in incorporating 'Big Data' in decision-making at your institution? Are there specific examples of success or failure?
3. Who was the driving force for 'Big Data' adoption? What was the anticipated benefits and challenges?
4. Are there other ways that you think 'Big Data' can help your institution offer more value to learners and other key stakeholders?
5. Has 'Big Data' influenced decision making? If so, how?
6. What are the factors accelerating or inhibiting 'Big Data' adoptions at your institution?

Competitive Drivers:

1. How big of a role does 'Big Data' have in:
    a. Increasing cost-effectiveness?
    b. Improving services to learners?
    c. Differentiating from competitors for student recruitment?

Ontology / Epistemology:

1. What do you think is the overall comfort-level with 'Big Data' decision-making processes that are based on correlations and inference rather than causality and knowledge?
2. Do you think that 'Big Data' findings are primarily used to drive change independently or more to validate knowledge-based decisions?

Change Management / Organizational Learning:

1. How big a priority is data collection and management at your institution? Do you have enough data and analysis to support good decision-making?
2. Is the integration of 'Big Data' in decision-making valued and rewarded?
3. How do you think decision making could become more effective?
4. Do you think 'Big Data' is cost-effective?

Ethics

1. What are some of the ethical considerations with 'Big Data'?
2. How are learners, potential learners and other stakeholders protected in the data collection and analysis processes?

3.15 Data Analysis

As previously discussed, and in agreement with grounded theory general principles, data analysis occurred concurrent to data collection. Initial coding of interview data was followed by a more focused coding effort as abstract theoretical elements emerge from the data. Coding focuses on the meaning the participants place upon specific activities or events. Often gerunds were used for codes to focus on the activity and in vivo codes were often used to ensure the data was derived from the participants' own understandings (Charmaz, 2014; Glaser, 1978). From the codes developed, broad theoretical directions were proposed that could be validated in the later stages of the data gathering and interviewing process. In the initial coding, the researcher must be cognizant of what the data suggests, from who's point of view, and how does it agree or conflict with other collected data. Grounded theory encourages reflection in the data collection and analysis to become aware of any tension as it emerges (Glaser, 1978). Coding remained simple, precise and focused on the actions of the research participants (Charmaz, 2014). The codes helped the researcher become aware of a phenomenon when it was not explicitly recognized by the research participants. Recognition of phenomenon and theory development was based upon the researcher's inductive reasoning that was supported by the collected data.

3.16 Trustworthiness

Trustworthiness in qualitative research studies is defined in three broad areas: i) credibility, ii) dependability, and iii) transferability (Bloomberg & Volpe, 2016). While the constructivist ontology proposes that the perception of reality is a social construct and that no objective social 'truth' can be known, it remains critical that the qualitative researcher provides a valid portrayal of the participants perceptions (Creswell & Gutterman, 2019). In this light, the research aimed to disclose areas of potential bias, exercise reflexivity throughout the research process, use multiple data sources and to identify and discuss data that are potentially contrary to proposed conceptual framework under development. Furthermore, to build trust with research participants and strengthen the credibility of the study, the interview transcripts were validated by the research participants for validity and accuracy. Intermittent ‘peer debriefing’ with fellow researchers also examined alternative ways of interpreting the data as well as identify possible areas of unconscious bias (Bloomberg & Volpe, 2016). Dependability in the context of qualitative research studies typically refers to the consistency between the data and the findings (Lincoln & Guba, 2000). The study provided detailed descriptions of the data collected so that fellow qualitative researchers could assess dependability independently. Finally, in regards to transferability, qualitative research studies are not typically expected to have generalizable findings (Bloomberg & Volpe, 2016). However, the ‘lessons learned’ in one setting may be useful in another situation in a similar context. As a result, the researcher endeavored to provide as much context as possible so that other researchers may assess the level of the shared experience with the study.

3.17 Ethical Considerations

The research was be approved and adhered strictly to the requirements of the University of Calgary REB. Ethical practices placed the well being and respect of the research participants’ paramount. Informed consent of all research participants was sought and obtained prior to data collection. Informed consent included i) an explanation of the research purpose, processes and any inherent risks, ii) a description of benefits to be reasonably expected, and iii) a clear procedure to ask for further information or withdraw consent at any time (Cohen et al., 2011). Research participants had the explicit option to remain anonymous throughout the research process. Data was provided and recorded in strict confidentiality and research participants had the opportunity to verify statements made during the data collection process. Research participants were mid to senior leaders within HE institutions and any relevant ethics approval of these institutions was officially sought and received as necessary.

3.18 Summary

The purpose of the study is to develop a conceptual framework that describes and explains the adoption of ‘Big Data’ technology in the Alberta HE sector. Grounded theory is an ideally suited research methodology for developing theories or explanations of human or social processes, especially when existing concepts do not effectively address the specific research problem. More a series of principles and strategies than a prescriptive sequence of tasks, grounded theory emphasizes concurrent data collection and analysis and inductive reasoning for theory development. After University of Calgary REB approval was given, seven semi-structured interviews with HE leaders at a variety of Alberta institutions were conducted. Research subjects provided their informed consent and throughout the researcher verified with the research participants the data collected. Research subjects were anonymous and all data was securely stored. Data analysis (see Section 7.1: Data Codes and Memos) was conducted through memo writing and data coding of the interview transcripts. The researcher led initial interviews with ‘sensitizing topics’ as identified in the literature. A reflexive process during data collection and analysis as well as consultation with colleagues was endeavoured to minimise researcher bias in the study.

4.0 Chapter Four: Data Collection and Analysis

> The deductive method is the mode of using knowledge, and the inductive method the mode of acquiring it.
>
> Henry Mayhew (1812-1887)

This chapter describes how and what data was collected and analyzed using a constructivist approach to grounded theory and describes the conceptual framework developed to better understand and resolve the research questions.

4.1 Overview

The purpose of the study is to develop a conceptual framework, grounded in the perspectives of HE leaders in Alberta, that describes and explains the adoption process for 'Big Data' technologies. As noted in Chapter Three: Research Design, the study employs a constructivist approach to grounded theory as detailed by Charmaz (2014). Utilizing the constructivist approach to grounded theory, results focus not upon generalizable findings but explaining how a phenomenon is understood and how that insight is expressed and applied by the research participants (Creswell & Gutterman, 2019). Through the constructed conceptual framework, insight is expected into the study's following secondary research questions:

1. What do higher education leaders in Alberta consider the main elements for and against in 'Big Data' adoption or further development?

2. To what extent are Alberta higher education leaders motivated by competitive enrolment and/or cost-effective learning considerations that could be developed by 'Big Data' applications?

3. Do Alberta higher education leaders believe that meaningful business decisions can or should be influenced or even made by 'Big Data' applications?

4. What ethical considerations do Alberta higher education leaders integrate in the data collection and analyses processes of 'Big Data' initiatives?

5. How able do Alberta higher education leaders consider their respective institutions are to implement 'Big Data' successfully from technological and change management perspectives?

HE leaders across the sector are encouraged to review the constructed conceptual framework and assess its validity to their own experience and context.

4.2 Research Process

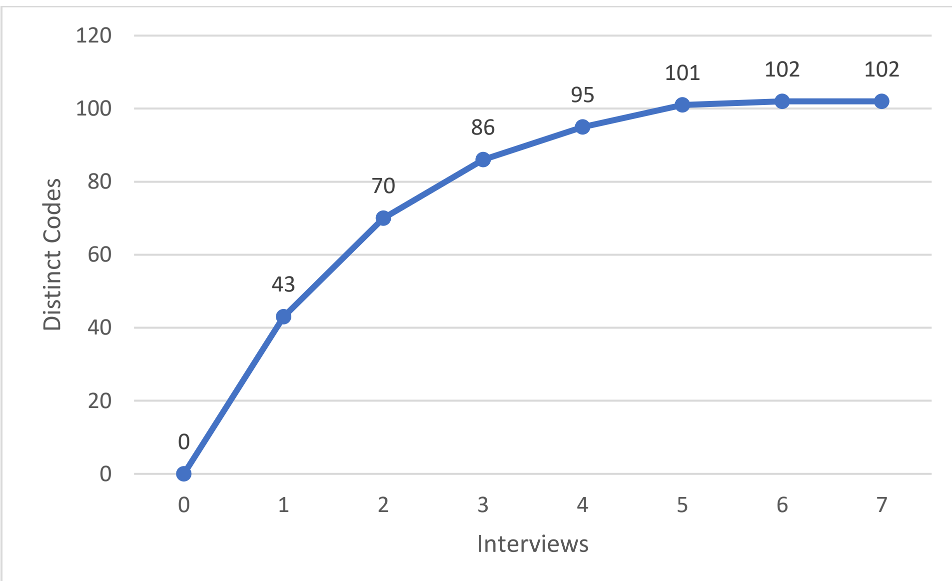
Before commencing primary research, ethics approval was sought and obtained from the University of Calgary REB as well as the REBs at HE institutions in Alberta affiliated with the research participants. A range of institutions are represented that include CARUs, CCCs and PIs in Alberta. As discussed in Chapter Three: Research Design, research participants were recruited in a process that was initialized within the researcher's personal network and expanded using a 'snowball' methodology to identify a diverse range of Alberta HE leaders in various institutional roles that have experience with 'Big Data' (Creswell & Gutterman, 2019). Seven interviews were conducted, producing 124 transcribed pages for analysis. The interview transcripts were shared and reviewed with each respective research participant and edits made as requested. As per the constructivist grounded theory model proposed by Charmaz (2014), 103 distinct codes or 'theoretical samples' were identified in the initial analysis. The initial codes were applied to the semi-structured interview transcripts a total of 723 times. Again, using the Charmaz (2014) constructivist grounded theory method, further analysis yielded eight secondary codes which were distilled to four key themes. The full list of codes developed and assigned are included in Section 7.1: Data Codes and Memos.

The constructivist grounded theory method supports concurrent data gathering or 'theoretical sampling' and data analysis with each process serving to inform the other (Bryant, 2007; Charmaz, 2014; Creswell & Gutterman, 2019; Miles et al., 2014). Emerging themes from the data analysis influence subsequent semi-structured interviews to allow some concepts to be more deeply explored with the research participants. Some themes were anticipated in the research project design and literature review. However, other literature review topics did not materialize to the extent anticipated by the researcher. Reflection on the data is constant throughout the data gathering and data analysis processes. Data is collected, reviewed and compared to how it did or did not correspond with the research participants' and the researcher's experience and potential bias. Using Charmaz (2018) constructivist grounded theory method, initial semi-structured interviews focus upon 'sensitizing topics' discovered during the literature review. Later in the data collection process, commonly identified interview topics are investigated in further depth for theory development. As the conceptual framework emerges from the data and as semi-structured interviews are completed, the researcher reviews the initial findings with each of the research participants in order to, i) socially co-construct the conceptual framework, ii) validate with the researcher's constructivist ontology, and iii) review for bias (Charmaz, 2014).

Throughout the constructivist grounded theory development process, constant reflection upon the collected data offers insight on potential relationships between the data sets as well as opportunities to note data gaps (Charmaz, 1990, 2014). These insights and notes are recorded in observational memos developed throughout the data gathering and analyses processes. The data collection tools are refined to address identified gaps and explore emerging themes and theory. Nine memos are noted from reflection grounded upon the data collected with a view, not just to describe HE leaders' perceptions, but to explain why 'Big Data' technologies are perceived and understood in a particular way. In addition, as the conceptual framework emerges, the researcher reviews topics with the research participants for feedback and facilitate co-creation of theory development. The memos are recorded in Section 7.1 Data Codes and Memos.

As noted by Charmaz (2014), the number of interviews required under the constructivist grounded theory depends upon when the ‘saturation point’ is reached and no new significant insights are realized despite increasing amounts of data. As noted in Chapter Three, the anticipated saturation point was between six and ten interviews. As shown in *Figure 2: Interview Saturation Point*, there is a declining number of new codes as the interviews progressed until the ‘saturation point’ (no new codes) is reached at the seventh interview.

Figure 2: Interview Saturation Point



Of the 102 distinct codes that are assigned in the initial coding phase, four broad themes are characterized. Two of the themes have distinct subthemes identified within them. The themes and subthemes are as per *Table 1: Focused Coding and Categorizing below.*

Table 1: Focused Coding and Categorization

1. Measurable Results	2. Recognized Performance Gaps	3. Viable Potential Solutions	4. Ethics & Legal
	2.1 Collegial / Professional	3.1 Financially Viable	
	2.2 Competition	3.2 Technically Viable	
	2.3 Performance Framework	3.3 Organizational Learning	

Finally, the conceptual framework is constructed from the data gathered and categorized in order to respond to the research questions using the gathered data and the developed theory.

4.3 Conceptual Framework

Upon reflection of the emerging data code categories, the parallels with typical IT governance criteria at HE institutions became increasingly evident. As discussed earlier in section 2.9 IT Governance In HE, IT governance processes guide investment by comparing competing proposals for limited investment funds. Cost-effectiveness is determined by proposals that can convincingly demonstrate the criteria that mirror those that are categorized in the data collection and analysis. The conceptual framework that emerges from the data is one driven by the historical assumptions, objectives and reasoning processes embedded within typical IT governance processes within HE institutions. To mitigate risk, IT governance practices employ deductive reasoning to evaluate potential investment in new technologies. The evaluation determines the course of action with the greatest impact upon the highest priority institutional challenges. Therefore, when applying deductive reasoning, IT proposals most likely to be funded include the following attributes:

1. Provides measurable results: the scale of the problem and subsequent improvement can be gauged and compared with the solution’s costs;

2. Addresses a priority performance gap: the problem is a priority reflected in institutional or government strategic plans;

3. Proposes a viable solution: there is a probable solution that is sustainable from a technological, financial and institutional capacity perspective; and

4. Addresses ethical and legal issues: there are no ethical or legal issues that cannot be mitigated to acceptable levels.

Conversely, technology that relies upon inductive reasoning does not fit governance processes which are designed to limit institutional risk and maximize cost-effectiveness. For example, technology that utilizes inductive reasoning is not able to well define a challenge much less propose a viable solution until a significant investment is made to systematically collect and analyze relevant data sets. An inductive process is holistically driven by data and therefore the 'problem', much less any potential solutions, can only be identified after investments for data collection and analysis. Under normal circumstances, it is extremely unlikely that an IT governance process focused upon risk mitigation and cost-effectiveness would approve an investment prior to a problem and at least one viable solution is proposed. As a result, HE leaders use deductive reasoning within IT governance processes to manage risk through prioritizing problems and assessing the viability or ethical risks of potential solutions before significant investments are made. In short, it is not reasonable to expect HE leaders considering competing proposals that utilize deductive reasoning, to approve significant expenditures when the both the problem and the potential solution are unclear.

Despite significant IT governance challenges to adoption in the HE sector, some small-scale 'Big Data' projects are being implemented. Data on 'Big Data' projects gathered from the research participants are summarized in *Table 2: 'Big Data' Projects Cited by Research Participants*. The projects are noted and analyzed in this chapter to discover common elements and insights into how and why they are approved through standard IT governance processes at HE institutions regardless of the deductive reasoning impediments.

Table 2: 'Big Data' Projects Cited by Research Participants

#	Research Participant	Project Goals
1	1	Modeling number of applicants (planned over five-year cycles) to accept in any program achieve and sustain enrollment at optimum levels considering attrition from applicant to registrant and throughout programming years

#	Research Participant	Project Goals
2	1	Demonstrates evidence to internal stakeholders of improving prospect to registrant yield rates
3	1	Analysis of digital engagement data to predict prospect- and applicant-to-registrant yield rates
4	5	Recording and analysis of applicants' academic background (sequential learners, mature learners, etc.) to inform marketing activities and purchases
5	1	Analysis of applicant data (backgrounds, interests, academic success, etc.) to inform future programming structure (i.e. five-year enrollment plans)
6	1	'Qless' – vendor product that analyzes service demand to reduce registration wait times and more efficiently address peaks and drops in service demand
7	1	CRM tracking and analysis of 'touch points' (clicks, click throughs, etc.) with prospective learners to guide future online engagement investment and predict future enrollment
8	2	Data analysis to produce Fact Book and 'Private Fact Book' for reporting requirements and to identify and populate strategic metrics for senior management dashboards / decision-making
9	2	Analysis of NSSE (National Survey of Student Engagement) learner feedback data for insight on new program programming structure and how to improve the student experience
10	2	Data analysis to produce more student-centric exam scheduling

#	Research Participant	Project Goals
11	4	Analysis of LMS data to identify successful vs. struggling learner behaviours
12	3	Data analysis to predict network hardware failure to reduce down time and repair and maintenance costs
13	3	Data analysis to predict student success from SAT grades and adjust course loads to facilitate learner success
14	3	Analysis of frequent and standard financial functions to automate - Robotic Process Automation
15	5	Data analysis of applicant academic backgrounds (sequential learners, etc.) to inform marketing investment
16	6	Data analysis of faculty LMS adoption rates
17	6	Data analysis of LMS for an early 'at risk' academic alert system
18	7	Data analysis (student survey and enrollment) to inform programing changes and other strategic decisions
19	7	Data analysis to predict enrolment trends and optimize enrolment levels

The projects and how they are approved are examined to identify common characteristics of successful 'Big Data' projects in the HE environment. As stated above, 'Big Data' projects, by their nature, are suited to inductive reasoning processes that are not typically favored by IT governance models in the HE sector. However, none of the research participants indicated that their projects went through an inductive project approval process. Indeed, an inductive project approval process appears a theoretical construct. Instead, the research participants were able to implement their projects despite deductive HE governance processes. Unsurprisingly, the common elements of the noted 'Big Data' projects mirror the typical requirements of the deductive governance process detailed in *Table 1: Focused Coding and Categorization*. As a result, the common elements for project approval comprise the proposed conceptual framework for understanding leaders' perspectives on 'Big Data' adoption in the HE sector as noted below in *Figure 3: Conceptual Framework 'Big Data' Technology Adoption*:

Figure 3: Conceptual Framework 'Big Data' Technology Adoption / Governance

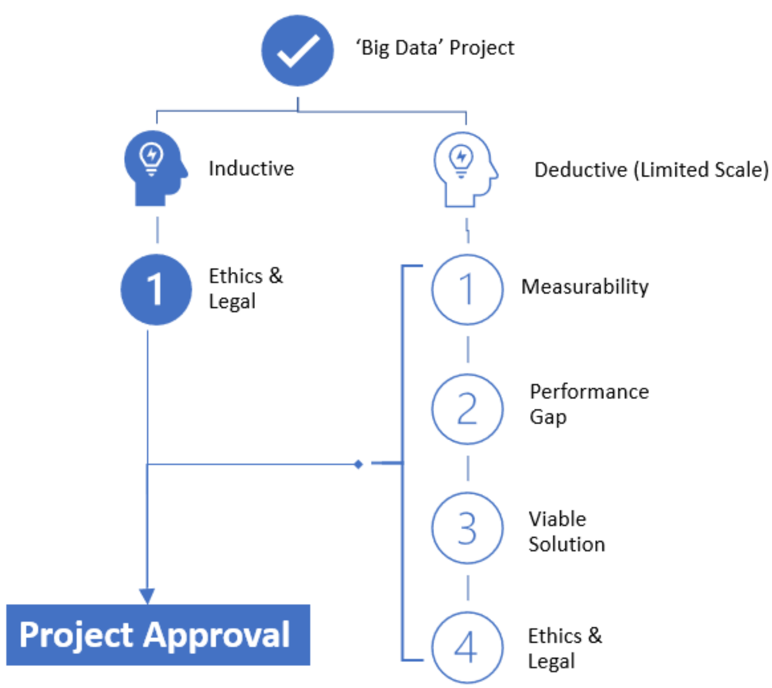


Figure 3: Conceptual Framework ‘Big Data’ Technology Adoption / Governance above shows the criteria typically used in a deductive reasoning process inherent to IT governance in the HE sector and contrasts it with a hypothetical inductively-driven process. The inductive process is labeled ‘hypothetical’ because no example could be found among the research participants’ experience and, as discussed previously, the inductive approach lacks the same risk management rigour in current HE IT governance processes. In the inductive process, investment would be made to collect and analyze data after ethical and legal requirements are fulfilled. Only then could the problem and viable solutions be identified and fully considered. However, as can be seen in *Table 2: 'Big Data' Projects Cited by Research Participants*, approval of some ‘Big Data’ projects does occur in certain circumstances. In cases where an IT project proposal has relatively small investment requirements as a ‘pilot project’, it is possible for ‘Big Data’ proposals to mimic the criteria prerequisites inherent to governance processes driven by deductive reasoning. Once the problem and viable solutions are identified through the ‘pilot project’, the investment needed to scale the ‘pilot project’ can be analyzed using standard deductive reasoning. However, for the ‘pilot projects’ to be initially approved, the following criteria must be convincingly addressed:

1. Measurability

   The project provides measurable results and the solution’s implementation results in a significant and unambiguous impact.

2. Performance Gap

The project addresses a recognized performance gap which typically is identified through i) collegial or professional networks, ii) comparison with other HE institutions; and/or iii) institutional performance frameworks usually represented in strategic plans or comprehensive institutional plans.

3. Viable Solution

   The proposed ‘Big Data’ solution is compelling in terms of i) financial viability, ii) technical capability, and iii) organizational capacity in terms of access to the required implementation skills and innovation adoption.

4. Ethics and Legal

   The solution does not represent ethical or legal risks that cannot be mitigated to acceptable levels.

IT governance processes are driven by deductive reasoning to select the most cost-effective technologies for adoption at an institutional level. Conversely, ‘Big Data’ technologies rely upon inductive reasoning to identify, understand and resolve IT challenges. Despite the governance challenges to ‘Big Data’, there are many examples of small-scale adoption within HE institutions. The common elements proposed and considered for successful adoption comprise a conceptual framework for how HE leaders perceive ‘Big Data’ as they are adopted at the institutional level.

4.3.1 Information Technology Governance

Governance processes serve to select and coordinate the implementation of IT solutions and are a vital element in conceptual frameworks considering technology adoption. The criteria that HE institutions choose to evaluate reveals why or why not certain technologies are adopted. How and why projects are approved or not approved underpins many HE leaders' IT priorities and perceptions towards 'Big Data' technologies. Research participant 1 explains that IT governance processes simultaneously prioritize addressing strategic challenges with mitigating the risks associated with implementing and operating an IT solution.

> We have a tech workplan for our entire office. It's shared with all the managers and the senior leadership so associate registrars here are like directors and then assistant registers and managers. Everybody puts together their wish list, and then based on institutional priorities or government priorities, because sometimes we have to change our system for government initiative, we look at what are sort of the key things that we need to get done. What are the pressure points? What are the solutions to help with those different things? Is it reconfiguring something we already own that just hasn't been configured right? Or is it purchasing something new? We take a very planned approach to any system implementation that we've done.

Prioritization of risk mitigation causes information technology governance to be intuitively drawn towards deductive reasoning processes. A proponent demonstrates one or more performance shortcomings and proposes a viable IT solution. A deductive process proposes a solution with a predictable cost. By contrast an inductive process requires resources to gather and analyze data before a problem is even identified much less a solution. Allocating funding prior to identifying a problem or viable potential solution is inherently risky – especially if the inductive process is competing for funding from proposals that include performance gaps and viable solutions that are identified through deductive reasoning. For many HE leaders, the question quickly becomes how can a technology project that relies upon inductive reasoning be approved through governance processes that prioritizes risk mitigation through deductive reasoning?

The research participants acknowledge the relatively limited scale and slow pace of 'Big Data' technology adoption in HE to date. In particular, research participant 5 notes:

> Post secondary by and large doesn't lack data. We do a lot of data collection just in the nature of our application processes, information we collect, student information, the student backgrounds, where they're applying from, so there is a huge amount of data. I fundamentally admit though, at least in << *research participant's organization's* >> perspective, we've only started to scratch the surface of being able to really make use of the data - to actually turn it into information, which I think really is what I see 'Big Data' being all about.

Despite the abundance of data collected and recorded in higher education in tools such as SIS, LMS, ERP [Enterprise Resource Planning] and CRMs, there is recognition that their data is not remotely being used to its potential. As research participate 7 states:

> Well, the amount of data we have already is pretty staggering. It's just the usage of all that data is hit and miss. We have to get the whole organization up to a certain level and that's a priority for this coming year. But the data is being socialized through a BI [Business Intelligence] tool.

There is concurrence from research participant 1, saying that despite an abundance of stakeholder data available to HE institutions, it is not effectively analyzed and transformed to effectively guide resource allocation or general decision-making.

> We have done a lot of projects. We continue to do a lot of projects that collect data, analyze data. They may not fall, majority of them don't fall with the traditional definition of ''Big Data'' per se but we are slowly getting into the realm of ''Big Data''…

Research participant 7 reflects a widespread view that 'Big Data' technologies and data analysis departments in HE as a whole represent considerable unrealized potential. Although those involved in data science in HE are proud of what they have achieved to date, there is an awareness of much untapped potential for 'Big Data' technologies exists.

> I think we have met the expectations of the institution. Well, I know we have. I don't think we have achieved what I wanted to achieve when the [Business Intelligence and Analytics] department was put together. I want everybody to use the data that we have available.......it's really about creating the users in the different schools and all the academic chairs. That's our strategy this year. I think the institution is quite happy with where we are. I just think we can do so much more. That's really for the Business Intelligence and Analytics department, what's our value proposition? Is it to provide descriptive stats for people to access and it's up to them to access? I've been told that is really our job because we don't control what goes on in academic division. We're not part of it. Or is our value proposition to provide the data and then get everybody up learning curve on how to use the data? Which means we take on a more consultant role. I'm leaning towards a more of a consultant role so we can push change within the organization.

Reflecting a certain degree of caution despite the recognized potential of 'Big Data' among many HE leaders, research participant 3 says that most projects approved are limited in scale and relatively rudimentary in order to limit investment requirements until the technology's value is more widely recognized.

> Because the investment as a capital expenditure is big enough for people, as a new item for them to just use it to prove what they already know. There's an investment factor in how they actually think about this, right. If I'm going to invest this money, there has to be something tangible, something that I really, really need to find out and everybody's saying 'Oh, this tool can actually help me do that'. That's why most of the beginner analytics are proof of concepts kind of projects. If they see value out of maybe they can incrementally increase investment, otherwise they just stop right there.

Research participant 1 reflects on experiences and the relatively rudimentary nature of 'Big Data' technology projects implemented to date.

> We're not using anything sophisticated, and now we are looking at a couple of tools like Power BI, but right now we have used, some of our staff are quite trained, so they're using an SQL server to pull information and then the map, manipulate some additional data and create reports in Tableau, and Excel, so nothing too sophisticated.

There is an appreciation for the potential for 'Big Data' despite the challenges meeting the deductive reasoning requirements of institutional IT governance. Those 'Big Data' projects that are implemented are able to meet the following common challenges of i) measurability, ii) performance gap, iii) viable solution, and iv) ethic and legal constraints.

4.3.2 Measurability

A priority criterion in IT governance processes is the ability to demonstrate measurable results as a return on investment. Research participant 1 affirms, 'anything that we implement, we are always looking at what outputs are that system is going to give us. Not just from how we support and work with students, but how can we measure our success'. The measurability consideration is significant from not only a cost-effectiveness perspective but also because measurability can help influence other departments towards specific action – which has considerable value in distributed decision-making environments common to the HE sector. An example is given by research participant 1 where Deans are encouraged to be more personally active in the recruitment process, increase the number and quality of 'touch points' or interactions with prospective students.

> [The data collected] gives us evidence to show faculties why they need to interact certain ways with students. Then we can say we actually have a better conversion rate with students when we have this many touch points with them. One thing that we introduced a few years ago was the faculty Dean letter. Every Dean now writes a letter to their admitted students. It goes out after our admission package. It goes out by email, and it's just, 'hey, we're so excited to have you come and join us'. We saw a drop in our attrition rate from accept to registered because you still

> always lose a few students even though they pay to accept. We still lost a number of students but we started to see that number drop, which is what we wanted to see.

Emphasizing the importance of measuring changes in performance, research participant 1 underscores the need to report changes in performance that are the direct result of technology investment and implementation.

> Usually, there's something that triggers something for you to look at it to say is this successful or not? Here are some warning signs or here are some really good things. What is the story that you want to tell? It kind of works on both sides. Where we've had things that we've looked at it because it was working really well, we wanted to see how well it was working. And it showed us a whole other lens of information that we hadn't even looked at before. Wow, this is actually a really good story to tell.

One of the most important and challenging functions related to learner recruitment and admissions accurately predicting program 'yield' or the number of qualified prospective learners and program applicants needed by any particular program year in order to fill the planned cohorts with actual registrants. Too few qualified prospective applicants results in wasted capacity and unrealized tuition revenues through unfilled programs while too many represents a redundant use of limited recruitment resources. Research participant 1 notes recent 'Big Data' projects are, based upon historical program enrollment data, accurately predicting the required number of qualified prospects and applicants to fill a program.

> With admissions, it's a lot of data analysis so that we could project what our yield is going to be based on the number of applications we have, the number of offers that we make, and ultimately how many students are actually going to come. Even though they've accepted, there's attrition on those acceptance rates.

Similarly, HE institutions are using historical learner success rates to predict program attrition rates and identify learner behaviour and/or experiences that accurately predict academic success. If learner success rates can be predicted in the recruitment and applicant selection processes, the number of graduating learners, and the corresponding measurable graduation rates and tuition revenues, will be maximized Research participant 5 notes that prospective learner profiles are analyzed and compared with enrolment and academic performance data in order to i) predict the optimum marketing investment level and ii) recommend new programming service levels in response to fluctuating learner demand.

> Getting these realms of different information about applicants’ backgrounds, prior learning that they've had, whether they're coming into the institution directly from high school, from the workplace, with other degrees, understanding the nature or the profile of the applicants that began to inform marketing, to inform trends about where our programming might be going.

In addition to recording and analyzing the selection criteria of prospective applicants and applicants in the recruitment process, data analysis is also influencing ongoing programming decisions. Research participant 7 notes that data collected from CRM and SIS systems (enrolment, academic success) plus data from student satisfaction surveys are affecting programming decisions such as scaling up or down program offerings. Research participant 7 explains the importance of international enrolment and the accompanying tuition differential fees when considering in programming decisions.

> Our department, we look after market research which is survey data and then we look after the student data that we collect through the Office of the Registrar. For ‘Big Data’, let's use the student data, because that is ‘Big Data’. Survey data is not that big in the terms of data points. But with the student data, we analyze our student data to answer many different questions but most of the questions are really about program decisions, as well as strategic decisions per se. We have data on applications. We look at applications based on, is it going up? Is it going down? Is it domestic that's going down or is it international going down and up?

Research participant 7 also highlights the importance of enrollment, learner success and attrition data analysis to quantifiably predict actual registrations at program start dates. With these projections, institutions can plan cohort numbers, optimize enrollment numbers and more fully load graduating cohorts. The predicted enrolment requirements are applied to marketing budgets in programs where more learners are anticipated to meet optimal program loads.

> We look at [registration at attrition levels] at a program level because if we start to see trends that can help predict what is the number of cohorts we should have in a certain program. Let's say we have four cohorts and applications start to go down, we have to take a look at is it still feasible for us to offer four cohorts or do we need to go to three, as an example. You can get a lens to that seven months out or even a year, sometimes. You don't have to suddenly be in a panic situation when you're seeing that there is declining application pool. Because if we don't have waitlists then we run into a summer 'melt' situation, as most post-secondaries do. We have to make sure that we have enough applicants in the pipeline who's also qualified to backfill when you experience a 'melt'. We start to look at our application numbers, confirmed and the waitlists. You start to see patterns that you can actually use to predict what is going to happen in September, today. That's some decisions we have. If one program, I'll use as an example just a couple of days ago, the Dean actually said they are now looking at maybe, eliminating a cohort just based on the data that they see. That's some of the actions. We look at retention numbers. We look at graduation numbers because retention obviously has a big impact on the student who didn't succeed. If the intention was to complete and they don't, there is a revenue issue for the institution because there's two ways for us to earn money. One is to increase the

> number of the intake that we have and the other one is you make sure you retain the students that you have.

In terms of recruitment, more advanced CRM data information systems in the HE sector are able to record the online behaviour of prospective learners, such as email opens, click throughs, etc. For example, a prospective learner may receive a custom follow up email message if the previous message was unopened, if the 'call to action' button was not clicked or if they visited the program web site after a long absence. However, among the research participants interviewed for the study, only a few of the of more advanced automated engagement processes through a CRM are cited. In one case, research participant 7 confirms that their institution uses a CRM to record recruitment and prospective learner behaviour but does not customize automated follow-up actions.

> It just helps us with determining our yield rate in a lot of cases, which is really what we're trying to gauge, so we want to track those things. It won't change our interaction with the student. But it gives us evidence to show faculties why they need to interact certain ways with students..… We analyze that information to create all of our enrollment models. In addition to that we plan out for five years so we kind of want to have a sense what the trends are…'

Research participant 7 observes that gathered data and analysis informs best practices for prospective learner engagement and recruitment generally but there is not a capacity currently to personalize and automate the approaches to individual prospective learner interests.

> We look at open rates, click through rates, time on the click through site when they go through, to see that they just look at it and leave or do they actually engage with the site? We also don't just track email events. We actually track in-person webinars. So, we could see how many touchpoints a student has with us, and then correlate that with…. if a student has 10 meaningful contacts with us, our chances of converting that student are really, really high. But we look at all of those pieces and we track them as best as we can. Obviously, … it's not always possible to do that. But we know we went to a school visit. We had an event and the student came and they came to another event. They went to the faculty event. They attended a webinar. They read everything …that we sent them. We can see all of that.

Over the longer term, research participant 7 says that CRM and SIS data systems record and evaluate the success of recruitment events and initiatives to guide future investment and pinpoint areas for improvement, such as converting applicants to registrants.

> We measure the application uptake first, and then we wait a couple of years to see is there changing trends in the yield rate? If applications spike and yield stays

> low, to us that's actually not really a good recruitment event because we're not ultimately getting the students. But if we see that there is a relationship between applications increasing and yield increasing over time, then that is something we would continue to invest in.

Despite some use of predictive analytics related to future enrollment trends, HE institutional leaders perceive their use of ‘Big Data’ technology to be still at a rudimentary level relative to other sectors. Research participant 7 adds that institutions are planning to expand capacity for analysis in high priority areas for predication of registration and enrollment numbers. The higher value placed on predicting enrolments reflects the significance of measurable tuition revenue and the potential to reduce costs through ‘right sizing’ program offerings based on demonstrated learner demand.

> As of right now obviously our data is mostly descriptive. We don't do advanced analytics on the data yet, because we really need the whole institution to actually use the data that we have - that descriptive data. What I can do, is I can see linkages between different reports and I can then describe what's going to happen based on looking at them. That's what we're trying to bring across to the decision makers which are really all our academic tiers. We need to show the deans, because if they can start to see the linkages between the reports, it is pretty powerful without us having to do predictive analytics. We want to get to predictive analytics, but I want more data in order to do this. We have embarked

> on a new tool at [Research Participant’s Institution] called ‘exports’ which is for student feedback questionnaires, so, cause evaluations, basically. But we can also launch other surveys in tools such as our entrance survey which is a profile of our new students. What motivated them come to [Research Participant’s Institution]? What were the information sources used? How well did they know the programs and the careers they were going to get? And then, by using this tool, we actually have a student ID associated with the results. We always keep our data confidential, but it would allow us to use that data for data modeling and then at the back end, you have a program satisfaction survey when they graduate as well as a grad employment survey so basically you can start to do predictive analytics with a lot of power behind it when you have that kind of data. But we're not there yet. I don't think the institution is there yet. That's why I say we have descriptive data right now that we want them to use before we move to the next phase.

Reinforcing the importance of predicting program attrition, learner success, applications and registrations, research participant 1 notes that program size analysis is the most important and sophisticated application of ‘Big Data’ technologies done through a typical registrar’s office.

> We've also looked at attrition and graduation rates, or retention and graduation rates, so that we have a better sense of, not just the students that are coming in, but how many students are actually going to move through that faculty in a given period of time so that we know how to set those admission targets. So we want a

> faculty to have a population of say 7,000 students, roughly, Faculty of Arts. We know we need to take in x numbers of students every year in order to sustain that population over a long period of time, right, so that's probably the most comprehensive one that we do.

There is some evidence of HE institutions correlating specific investments such as particular marketing channels, events or promotion with particularly successful learner recruitment. However, as research participant 1 observes, a more in-depth analysis of learners' secondary learning experience is seen as a reliable indicator of post-secondary success. Prospect quality is often prioritized over larger numbers of less potentially successful learners.

> We don't really tie [recruitment expenditures] to the event because our admission standards are quite high. It's more about how successful are the students coming in based on the criteria that we're looking at for that point in time. Our retention rates are quite high. So, …we don't see patterns where you can tie it to an event. The success of a student in their post-secondary experience here isn't necessarily tied to that specific recruitment advantage.

Research participant 1 discusses using 'Big Data' technologies to recognize and challenge some assumptions used to identify prospects and applicants with the greatest potential for academic success. For example, academic success at the secondary level is generally correlated to post-secondary academic success. However, data analysis course by course of secondary and post secondary success identifies specific 'average booster' secondary courses in particular jurisdictions that increase academic averages but do not translate to post-secondary success.

> You have students from IB [International Baccalaureate] curriculums, actually, this is the patterns that we see or if a province goes through a curriculum review. And we're looking at specific courses for admission and we say, Okay, we'll accept that. But then we find out that course is not really a good indication of success in university. We're actually doing that right now with all the option courses in Alberta, we're seeing some patterns where they appear to be more GPA booster than they actually are for preparing students for doing well in university. Are those the courses that we really want to take or consider for admission for student who's coming into the kinds of programs that we offer? There isn't really a good relationship between them.

A further example of 'Big Data' supporting measurable learner success is given by research participant 2 where an HE institution uses analytics from SIS data to schedule exams. Data analytics are used to schedule final exams to avoid conflicts and provide sufficient preparation time to learners, thus giving learners a greater opportunity for academic success.

> [The Registrar's Office] have developed a process, a student-centric process, of exam scheduling and that required going through massive amounts of data. Even

> with the 'Big Data' infrastructure that is available to us, they still can't do it instantaneously. It takes them overnight processing. But the improvement they have achieved is the process that used to take them a week. They have narrowed it down to eight hours. So massive improvement compared to a couple of weeks it used to take them before. What the process has done is ensure that students don't have conflicts. Students don't have multiple exams on the same day, or multiple exams at the same time. Also, multiple exams that are of the same level, so, a variety of metrics that they've looked at. That had a direct impact in student success in their grades, as well as their experiences. … I have a feeling it has an impact on student retention as well. If you don't fail courses, if you pass courses, and it's not because you're not studying, but because of exam scheduling, that has an impact on people's perception of what they do at the institution.

Among HE leaders there is potential for translating the mass number of data sets within institutional LMS to discover correlations between measured academic success and learning online behavior and/or instructional design structure or strategies. Research participant 4 described the initial thinking of one such attempt.

> The driver [of the 'Big Data' project] was largely the desire of programs to have a better sense of which students might start to struggle in a program, and would benefit from early intervention. We started to look at ways of tapping the data that existed in the learning management system. And that's things like assignment completion, and even going into things like, most learning management systems

> will let you look at how often, and even whether, a student has looked at a particular element of the course. You can start to see what the successful students' behaviors are, what the struggling students' behaviors are in terms of how long they're spending on things, whether they're accessing all of the potential resources, things like that. And then, trying to come up with a system that will allow early intervention for students who were demonstrating behaviors associated with eventually doing poorly in the course.

However, as research participant 4 later discusses, the project does not live up to expectations as developing meaningful predictive and prescriptive results remains challenging and changes to current practices to measurably enhance learning are not clearly identified.

> Well it's certainly true that you can have a look and see whether students are exhibiting these kinds of behaviors associated with eventually performing poorly. For us it started to get to the point that, even if you knew these things, what difference would it make? What would you do about it? It is certainly possible to see some of those things developing across a larger group of students. At this point our experience was that you didn't really know more, and you weren't really able to do more than an instructor would have been able to do on their own - just keeping a close eye on their students.

Research participant 6 also portrays the perception among HE leaders that transitioning from descriptive results from LMS data to more sophisticated processes yielding predictive and/or prescriptive results is not yet among key institutional priorities. Investment in higher levels of sophistication of ‘Big Data’ technologies is seen as i) costly, ii) with unclear benefits and iii) best addressed at the sectoral level by educational technology vendors instead of individual HE institutions. However, some small scale and *ad hoc* predictive innovations using ‘Big Data’ technologies are occurring without resources specifically dedicated through an IT governance process.

> It's very clear at the start of [the LMS data analysis] initiative, with our management and Deans and Directors, that everything we were doing was historical. We would take all the information over the previous year and give a report on the utilization and adoption that had already occurred. We had no skill sets … at that time to do any forward looking in terms of predictive or prescriptive analytics. In 2018, we started to look at moving in that direction. We had a good foundation underneath us in terms of historical analytics to say how well did we do and we want to move into that space. How do we get out of that reactive and starting to move into the proactive space? Another initiative we spun up with the same kind of informal little group that we had was, how could we possibly move into that area of predictive analytics and what might be a way to do it? And what we determined through some conversations around campus and with some of our leaders was that a better approach or a better area to do that in was early alert systems for at-risk students. And so, our early alert system at our institution is very manual. It's very voluntary. Students can either ask for help or instructors can suggest that the students need help. In all cases it's voluntary and

> paper-based and manual and conversations, and data is not really used that much. It's also very late. Sometimes it's after the course has ended and they got a bad grade. This isn't an early alert system. This is simply a 'students are already at-risk system' and have either failed, or are at-risk of failing, and it's very far into the semester or maybe even after. It's a very reactive thing. Some of our research into this space indicated that it wouldn't be that difficult to develop an early alert system because other institutes have already done this. There's several in the states that have already gone down this path. There's actually tools and third party service providers that will give you this for a fee or cost. But, of course, we didn't have a budget, and there was no institutional drive at that time to fund or dedicate resources.

Further emphasizing the informal and organic growth of data gathering and analytics capacity in HE, research participant 6 describes an individual initiative that began experimenting with what was deemed promising technology prior to formal investigation through IT governance structures.

> …we didn't have capacity to spin up on a dedicated team or resources for any organizational change into or move into data science as a field or as a team. And so, myself and a couple others - I just tapped her on the shoulder and said 'are you willing to put in maybe a little bit of time here and there in between stuff or off

> the side of your desk, so to speak, and just start to look at how this might be useful?’ Like, let's just play with no expectations.

Research participant 7 affirms that data analytics investments to date are small in scale and relatively low in complexity due to decision makers’ marginal risk tolerance and doubts regarding measurable returns on investment.

> The ‘Big Data’ you're talking about here, we don't do that because you're talking really ‘Big Data’. You're basically talking about machine learning, artificial intelligence. We're not there yet. I don't know if we ever will be because we are not dealing with that kind of ‘Big Data’. I would say what we call ‘Big Data’, I’d say is small data. I would say, post secondaries for the most part, and especially right now because we are all under scrutiny for cost and expenditures, there will be no desire or appetite to spend on that unless there is a tangible outcome. So, for me, the tangible outcome is if we're in a position to actually use the findings. Which means that you need high data literacy and digital literacy in order to be able to deal with the data. I don't think we're there yet. I’m not just talking right now, not just talking [Research Participant’s Institution], I'm talking post-secondary, period. Because you have to have the deans, and you have to have the program coordinators, academic chairs or whatever term that is used. You have to make sure that they are able to actually use the findings and the data that you are going to get. And as I said, we are still working with descriptive data and we're trying to then move them to down the path of data literacy, where you can now

> look at this data and look at all the reports as companion pieces and then start to form a picture of what is going on and what do we need to do. So, that’s the project for this year. So, I would say right now, for post secondaries, … they don't understand the ‘Big Data’ and AI [Artificial Intelligence] and machine learning. For them it's an abstract concept. If they approved that they have to have confidence that the lower ranks know how to deal with this. I'm not saying that as a negative. They have a bigger fish to fry than AI and ML [machine learning]. I just can't see them move in that direction unless there is a tangible outcome. Over the years and objective, they say by doing this, we're going to achieve x y z. And if you can’t do that, then I don't think there's an appetite to sign on for what could be a million dollars or more.

Related to recruitment, enrolment and retention, research participant 7 makes the point that many HE institutions are only now beginning to understand how ‘Big Data’ can improve the effectiveness of recruitment and retention efforts by analyzing CRM, SIS and, to a lesser extent, LMS data to reveal heretofore hidden challenges. Recent trends, such as increasing international learner demand, tends to obscure longstanding challenges to domestic recruitment and retention, but recent events are creating a fresh urgency for HE leaders.

> Post secondaries are pretty happy with [increasing international enrollment]. But then you have a crisis as you have today. So, as long as you have an awareness, that's as far as we can take it. But I said a year ago that we really need to be paying attention to this because we have a great reputation and all the data points

> through surveys, but it's not really translating into applications. Why is that? Is it a product offering issue? Is it people just want a university for the sake of having a university degree? It doesn't really matter. But, this is a trend we've seen. The curve is kind of being masked because of the increase in the international applications. But then you end up in unforeseen circumstances as of right now, with the pandemic. Where international funnel starts to dry up because they don't know if they can get here. They don't know if they want to spend $20,000 for a year of online education as an example. So, if you don't have a pipeline of domestic to fill those spots that was normally filled by international, you're going to have an intake problem.

In addition to directly quantifiable impacts such as recruitment and retention, HE leaders also are increasingly recognizing proxy indicators from criteria such as change in student and employee satisfaction resulting from an IT project. Research participant 1 discusses an IT project that analyzed CRM learner data to prioritize services to those that need them most and to customize services to each learner's specific needs. The result is both more satisfied learners measured through survey data as well as increased staff engagement defined by staff survey and absentee data.

> We implemented a system called Qless A few years ago, which is our queue monitoring system downstairs. And we used to have horrible lines, like the actual joke was, you know, you could watch one of the 'Lord of the Rings' series before you actually got service here, which is really bad. Now it's down to, you know, a

> Pepsi commercial, just for comparison. … that metric is more about staff engagement in their job, right? When they're always getting students, who are upset and angry with them. Not really good. That translates into increased absences increased sick time, right? When, in that particular measure, what we did is we looked at having implemented this queuing system and looking at not just our response time, …wait times are significantly lower, call times are significantly lower because it's multifaceted. … Our savings or cost benefit is actually in less sick days. Less absenteeism. …that's a huge benefit. We have people who are here more, they're happier in their job and students aren't as upset necessarily after a long week to get a question like they need. In the CRM, it's direct tuition dollars. The CRM, we can see, here's where we're, we're now able to track these things that we never were able to track before. We can look at the cost-benefit of all of these different activities that we're doing. And this is how much these activities … generate revenue for the institution. And when you factor that against the cost of the tool, okay, we recruit one student that pays it.

Supporting the importance of measuring student or client satisfaction in HE generally, research participant 2 discusses how 'Big Data' provide HE leaders with insights needed to quantify and categorize the types of learners at a particular institution. Correlating CRM and SIS data with anonymous student satisfaction survey data, institutions now better understand individual learner priorities, challenges and objectives. For example, if childcare services are a particular barrier for lower academically performing learners and if a particular service or intervention has a measurable impact on retention and/or satisfaction.

The first example was analyzing data, reporting data for decision making for institutional leadership. The ‘Big Data’ work that we're looking at… is looking at our survey responses. Essentially, like most higher eds in Canada and the US, we participate in NSSE [National Survey of Student Engagement] service. Over the past many years NSSE has been used just to look at responses to individual questions. We haven't really had an opportunity to correlate or see if there's correlation between those responses. So one of the tasks that we took on this last year, we presented this at the CIRPA [Canadian Institutional Research and Planning Association] conference as well. is trying to understand if there was correlation between the responses and, if there are negative responses, are they from the same people? Would we be able to glean any insight based on the response and the correlation between the responses that we saw? And what we found is, some of the findings were eye-opening. What we found that we could categorize the students based on responses, not just knowing the Student IDs, we didn't look at the student IDs, we basically wrote an unsupervised classification model using AI. That is one of the ‘Big Data’ technologies that we used. We haven't really used Hadoop or Spark to store the data. We're able to store it just with a traditional database. But that allows to cluster our students, our respondents, into certain categories. So, the insights that we found out, we have to take it back to the leadership to see if those findings could guide program development for student success in improving student experience.

Measurability of IT project results is a clear priority for HE leaders. There are some exceptions, such as when an IT project is required for legal or security compliance. However, the IT governance process invokes a cost-benefit calculation and comparison with other proposed IT projects. Those projects that can convincingly quantify results better than competing proposals are most likely to be funded. Projects, like those driven by 'Big Data' and require some investment to collect and analyze data prior to building a case for full implementation, are shown to potentially garner some limited support through 'pilot' projects.

4.4 Recognized Performance Gaps

Consistent with the deductive reasoning inherent to IT governance processes, a common theme among the research participants is that technology adoption is driven, in part, by the ability of a proposed IT solution to convincingly addresses one or more recognized performance gaps. The interviews with research participants demonstrate three ways that performance gaps are commonly understood. First, a priority is identified by an HE institution in the existing strategic plan or Comprehensive Institutional Plan (CIP). Second, best practises that are shared and recognized within professional networks can lead to perceived deficiencies at HE leaders' institutions. Third, to a lesser degree, competition among HE institutions can drive a level of recognition that a specific type of performance must be improved at an institutional level. Research participants explained how 'Big Data' projects address recognized performance gaps using the following three broad methods.

4.4.1 Performance Frameworks

A fundamental component of IT governance at HE institutions is the demonstration of how a project will address a need or needs identified in an institution-based or provincial strategic plan. Research participant 1 concisely notes.

> We have very specific goals that we're trying to achieve as an institution. It's outlined in our strategic plan and in our academic research plans. Everything that we do in the registrar area is to try and support the institution and achieving those goals.

Performance frameworks are constantly monitored by HE leadership, says research participant 2. There is significant demand for current data on institutional performance against specific metrics. Successes are constantly gauged and resource and/or strategy adjustments are made whenever required.

> One of the examples is collecting data from our [ERP and SIS] system for student retention, student performance, as well as our HR data. This is primarily used for our Fact Book, which is available publicly. However, behind the scene, we have what's called a private Fact Book that is available only to leadership. The idea is that we use that same data to identify certain metrics for institutional strategies

> and those help our strategy heads gauge the performance of those strategies and make tweaks and pull levers as needed.

As research participant 6 notes, measurement towards precise strategic objectives encourages leaders within HE to pitch innovative solutions. Often a key component to initial IT innovation is the ability to convincingly propose a potential solution while concurrently keeping incremental costs and other risks low. Using already budgeted resources in new pilot programs to assess cost effectiveness is a method often used to test and evaluate new IT approaches in the typically risk adverse HE environment.

> I found a couple of individuals and a number of people in different departments that were, like myself, interested in [the 'Big Data] space. We started poking around and started to realize that there were certain things that we could possibly do that align with our strategic plan for the institute. We went to our VP academic and said you have within your strategic plan, in your education plan, you have some goals of moving faculty and student adoption of the LMS as a goal for the institute. That was just a strategic vision to say, we know that education is going in this direction and we want to move our faculty and our students to adopting technology and specifically the LMS as a strategic goal over the next couple of years. We can use the data from the data hub, some of the datasets, to actually measure that adoption, however we wanted to define it. That conversation kicked off our first foray into using data to support a strategical initiative. We set up

> some processes and some rules and, … over the next couple years we defined some criteria, like KPIs [Key Performance Indicators], and said here's how we're going to measure it. Here's the data points we're going to use. Here's how we're going to measure these and extract these, because you have to be very clear and transparent as to the process, not just the data.

Research participant 4 provides another example where a ‘Big Data’ project is implemented as an ad hoc small-scale activity, using already budgeted resources. As a result, the project is implemented more as a result of individual HE leader initiative and not driven or authorized by a conventional IT governance process.

> Well, we, got lucky in being able to hire some students with those skills and even lucky here that we had, in particular, one faculty member who - just because of their experience and background at a fairly high level in data analysis and computer programming - was able to bring that lens to the work. Otherwise it actually would have been a big problem. To even know how to manipulate the databases, get the information we needed. It wouldn’t have been insurmountable. But it would have required institutional resources that were very unlikely to be made available. We got kind of lucky.

During the data collection, the Province of Alberta is beginning to design three-year ‘Investment Management Agreements’ (IMAs) to serve as sector-wide performance funding frameworks. The details of IMA agreements with each HE institution in the province have not yet been determined. At the point of data collection, the pending agreements do not appear to significantly impact institutional IT project planning. Performance below the agreed upon benchmarks will impact as much as 15 percent of the province’s funding for the institution with the proportion increasing in subsequent years. However, as of mid 2020, the exact performance indicators and the associated performance measures are not yet announced. Later in 2020 it was announced that the new funding frameworks would not take effect in the 2020/1 academic year due to the worldwide COVID 19 pandemic (Government of Alberta, 2020). Research participant 1 describes how HE leaders are aware of the pending funding framework but, because the exact measures have not yet been announced, the impact on IT planning is minimal to date.

> [The provincial government] have identified metrics but we don't know specifically which ones are going to be applied to all the individual sectors and what they're talking about, or what they even mean in the greater context, right? So yes, they're going to look at, you know, enrollment, for example. Okay, but what's the measure? What's the outcome that you want? We don't have enough context yet to really know. But they've already pulled like all of the elements that you look at. Those are things we’re pretty much all pulling except for a couple, which we're not quite sure what.

There is some concern among HE leaders that, given the way they are currently constructed, IMAs may spur unintended and potentially counter-productive results. Research participant 5 reasons that performance frameworks and performance funding could hinder innovation if the measures do not accurately reflect the desired outcome. For example, if measures are focused exclusively on reducing expenditures, there can be insufficient recognition for important criteria such as increasing efficiency, relevancy, revenues, not to mention learning itself.

> I think [IMAs] has the possibility to [encourage innovation] but also it is possible to do quite the opposite, particularly when some of the measures are expenditure limits. You will not spend any more for this, irrespective of the benefit you might gain from the expenditures. So, I think having a somewhat draconian view that expenses are the only thing that matters at a top level. It is probably the result given the economic state of the province but it's very regressive thinking. There are no awards for innovation, other than the story that well, because we'll post these KPIs and force you to think differently. I'm not sure that is necessarily the case. So, I'd like to see more focus or some focus placed on what are you doing to do to enable your improvements and not constraints. And I am completely supportive of KPIs for government as it is really public money, whether it be tuition or government funding. It's coming from the public. We do need to be accountable. Those measures need to be ones that enable us to, not just the status quo. So, there isn't really a lot in what I've seen in the key performance indicators that will suggest that is being looked at as an opportunity.

Research participant 4 agrees that it is too early to know the likely impact of the IMAs but recognizes that the objective is to incentivize HE institutions in a way that the previous static institutional funding formula did not.

> Based on my experience, so much will depend on what's actually in those. But you tend to get what you measure and what you fund. And so, I think that they're very likely to incentivize whatever they seek to incentivize.

In terms of getting 'Big Data' projects approved through institutional IT governance processes, all the research participants recognize the overriding importance of the priorities articulated in institutional strategic plans. There is no consensus on if and how IMAs will intensify or potentially misdirect IT innovation at the institutional level going forward. In any case, to receive initial approval, IT projects need to justify costs by convincingly demonstrating the tangible results related to performance frameworks established by the institutions and/or the province.

4.4.2 Professional Networks

Many of the research participants indicated that they are influenced by what IT solutions other professionals in their network are using to enhance services. A professional expectation for an optimal learner or client experience is most cited as the main motivating drive. Research participant 1 discusses how innovation within professional networks is tracked and sometimes emulated if the institutional contexts are similar and the solutions match their own institutional strategic priorities.

> We do look at other institutions to say, what are they doing and how are they doing these things because someone's doing something really good. Usually you want to try and see how you may adapt it. Sometimes we adapt it to your area. But it wouldn't be to say, we're going to do that on top of what they're doing. Like we're going to outrun them on this. It's more of how can this help us achieve our goals? And is this the right way for our institutions? So, when we're looking at our data, and we're looking at the decisions we need to make, that's the approach that we take.

Professional influences are noteworthy when HE leaders consider relatively new technologies like 'Big Data'. There are widespread expectations that 'Big Data' technologies eventually will have a great impact upon HE institutions, but HE leaders are largely unsure how and what specifically will adjust. Many HE leaders do not want to be perceived as laggards in adopting new and potentially influential technology. HE leaders realize how academic programming is being influenced by increased demand for data science based on increasing demands from learners and employers. Furthermore, data science is also heavily influencing the research agenda at HE institutions, translating to a widespread recognition of potential impact of 'Big Data' technology in HE. Research participant 3 highlights the position of many HE leaders that recognize the importance of the technology but are unsure how to best manage its inherent risks.

> I think the institution is looking at this as if there's writing on the wall that this ‘Big Data’ analytics and data science is actually coming out very strong and you need to do something about it. So the next step is to look at it and say what can we do in order to design and provide courses to our students so that we can, not only keep the students that we have, but we can actually attract more. It's more of a self motivating kind of idea - actually bringing in these technologies for teaching. That's the main thing really; teaching and learning. That's the focus. The other thing is, they do look at other universities, but not from a competitive standpoint, but more of we don't want to be lagging behind in terms of knowledge of these technologies, in terms of courses that we offer around these technologies and so forth. That's why you will see a department like computer science for example having a new program around ‘Big Data’ analytics and data science. It's not about competitiveness. It's really more about being in the forefront of actually providing this kind of experience.

Despite a relative lack of experience with data collection and analysis technologies relative to other sectors, research participant 2 says that HE leaders often feel professional pressure to ensure their management decisions are data informed. The impetus for enhanced for data informed decision making originates more from expectations within professional networks rather than, what is more typical in other sectors, a direct means to improve metrics such as revenues or costs. However, there is the perception that HE professionals’ accountability to performance metrics is increasing as those indicators become more clearly and constantly available on institutional decision makers’ dashboards.

> For us, I don't think it's [metrics for success] as clear. However, having said that, there has been research on this by institutional researchers across North American and elsewhere as well that in order for institutions to survive and thrive, we have to be data informed. Even though one of the directions that we got is that we have to build a data-informed institution. Even though it does not spell out ‘Big Data’, essentially, that's where it's leading us to. I think as we do more and more dashboards and more and more look at data, we’ll hit a roadblock at some point. It's just that we haven't hit that roadblock block yet. For-profit industries hit that roadblock a little faster.

While HE leaders generally agree on the likely significant future impact of ‘Big Data’ technologies on the sector, the specifics of the evolution remain largely unresolved. Furthermore, while they are under some professional pressure to incorporate data analysis within institutional decision-making, there lacks a consensus on how an HE organization should become more ‘data-driven’. Despite the ambiguity, professional expectations do play a more significant role than factors such as competition in the HE sector.

4.4.3 Competition

Perceptions of competition as a motivating factor for 'Big Data' technology adoption among the research participants are more nuanced than expected from the literature review in Chapter 2. One explanation is that, because HE in Canada is primarily publicly-funded as opposed to more heavily tuition-funded institutions covered within the literature review, Alberta HE leaders may perceive some competitive pressure much less frequently than other jurisdictions. Research participant 1 says that competitive motivation is not prevalent. Pointing out that, despite the relatively high tuition rates for international students, their institution remains focused upon recruiting international students that are academically prepared for success above all. As a result, the research participant's institution does not use overseas agents to boost the numbers of prospective learners and applicants.

> We're not looking for high volume of international students. We're looking for a good blend and mix. And so that really creates a very different environment when you're working with agents.

Research participant 1 also does not consider HE driven by competition but instead more by internal goals in institutional strategic plans and measures within those plans.

> I don't feel we're in a competition. So, if you look holistically, you know, we're not. We're competing with ourselves is the way that I look at it… .

Research participant 2 notes that competition in the HE sector is mostly internalized within the institution.

> Perhaps there's a competition within ourselves amongst ourselves when we're just challenging ourselves to do better. That's the competition I can think of that is more evident right now.

Research participant 2 argues that HE leaders in Alberta do not perceive a competitive environment related to learner recruitment or retention. Instead, HE leaders recognize a professional need to continually improve and better serve prospective learners, applicants and current learners as well as other clients. However, competitive pressure is acknowledged related to recognized academic standards or research capacities, rather than enrolment. Nevertheless, the presence of competitive pressure, or lack thereof, is not seen as a significant driver or inhibitor of ‘Big Data’ adoption.

> I think competition exists in a variety of areas. It might not be as evident in student recruitment and student retention; I think we are fortunate to be one of the older universities. We do well in recruitment and retention. However, I think there is still competition. Whether nationally or locally, there's some form of competition. I think the institution leadership, as well as the institution

> community, understands that we need to continuously improve. We can't just relax and expect to keep up the standard where it needs to be for recruitment and retention, not just for students, but also staff and our academics. So, the competition is a bit more obvious I think in the research realm, publications, grants, and things like that. But I don't think that that is the reason behind… a slow pace of 'Big Data' adoption. There is adoption happening at the institution. Obviously, it's just that it's not the pace that we see in for-profit sector. Primarily because in the for-profit sector, perhaps they see a little bit more direct relationship between what they see in the data and the bottom line.

Reinforcing the perspective that HE institutions work more collegially than competitively, research participant 6 affirms that IT departments from various HE institutions tend to collaborate to address common problems in the sector. Professional groups are the impetus for sharing information and cooperation among institutions on common problems. There is a desire to be recognized amongst peers in different areas of strength such as research capacity, relevance to labour market, graduate employment and/or collaboration with employers.

> From what I've seen, and I've been in my organization for about 10 years now, we really tend to lean towards being as collaborative as possible. We're involved in a number of different networks and associations and groups to work with our partners and other post-secondaries, …involved in many of the boards and provincial activities as well as the national ones. Some good examples of that is

our applied research group is actively involved in research projects across Canada with many partners and so, highly collaborative. And at the same time, I'm not sure if competition is the right word, but we're competitive with ourselves, in terms of continuous improvement. And so, [given our institutional mandate], this separates us a little bit in terms of the metrics we're measuring and how we define success. But we did find it really clearly in terms of its graduate outcomes. You know, graduation rates and its employment. So, … are our students getting jobs? Are our graduates employed and are they employed in fields that they actually trained for? If you come out of electrical engineering but then go get a job in basket weaving or something, then it's hard to see how that education led to the transferability of skills. We really monitor those things carefully. Are we doing something that is producing good graduates? Are they graduating and is it worth it? … And so, we strive to monitor those KPIs, envision ourselves in terms of continuous improvement. I would say we're highly competitive in that sense. And the third piece would be where we're highly aligned with industry. Almost all of our programs have industry partnerships, sponsorships. We have program advisory councils or boards and they're almost always industry driven. And so, we're very hand-in-glove with our partners from all of our different faculties to work with industry and say, okay, are you hiring our grads? Are they working? What are you looking for? How is this going in that feedback loop of continuously improving and growing with all the changes and trends around us? So that's another thing, we're highly competitive and making sure that we're doing that really well.

Research participant 5 agrees that the HE sector in Alberta is not particularly competitive. However, some exceptions are noted in areas of general programming and geographic overlap.

> I think by nature of our geography, we're a loosely competitive industry. Just use Calgary and Edmonton as two examples [two Alberta post-secondary institutions] have very complimentary programs that offer many of the same programs- some unique and different at times. But if you look at our student draw, they come from either the north, or the south. It's very rare that we'll see a northern student come to take a program [in the south] that is available [in the north]. … We are generally non-competitive in that space. I think the same would apply to [two other Alberta higher education institutions], although they have a lot of common programs but they have some unique ones. I think from a student draw perspective, generally speaking, we are not overly competitive.

Nevertheless, research participant 5 acknowledges the current trend of increasing competitiveness that will drive greater adoption of data analytics to increase cost-efficiency in areas such as recruitment in some markets.

> The data sets are there. What we've lacked is probably the foresight to make better use of them. We use them on a highly transactional basis. We want to

> make sure that we have the seats in the program filled. It is that measure of supply and demand and how we size that. But in light of where postsecondary is going, which certainly has been blown up due to the COVID situation, the competitive nature of what post secondary is going to be. If you're now competing in the online space with everybody else for a program, how do you take the data that's available to allow you to do a better job of understanding the marketplace and how you supply the markets.

The competitive nature of HE in Alberta is likely to intensify with the eventual adoption of the IMA funding framework, adds research participant 5. In line with the probable IMA funding priorities, Alberta institutions are likely to compete in a much more direct manner as recruitment continues from a limited domestic learner base into comparable types of programming dedicated to a limited number of ‘high demand’ employment opportunities.

> When you get into a funding-based model that has key metrics associated with it, particularly around the number of students you serve, the number of graduates, graduation rate employment rates, I think we're going to see a much different competitive set. You have as your funding, is absolutely tied now to your performance. If the pie is only so big, and the bigger piece goes to the better institution, you bet there will be competition.

Competition is also heavily impacted by geography and the increasing reach of online program offerings says research participant 5. As institutions are increasingly offering the same types of programs in the same geographic regions and as online program learning becomes more and more prevalent, HE leaders are cognizant that increasing competitiveness can be expected in the future.

> So, I think, generally speaking, we are not overly competitive just by nature. Now if this was a GTA [Greater Toronto Area] or region where they've got multiple colleges, they do compete for students because in the space of 20 kilometers you've got two huge institutions that offer the same product. …Not the case so much in Alberta, I think. Where the competition lies is I think it's more of a future consideration. If the move to online becomes the norm, or more of the norm in post-secondary, I might choose to do an online program, but now I'm not bound by geography, but by reputation. I may choose the institution I think that graduating from an institution that will give a better education experience.

There are conflicting views among Alberta HE leaders about if the post-secondary system is designed to promote or inhibit competition among the institutions. Research participant 4 views the system as designed for Alberta HE institutions not to compete but instead to occupy niche roles within the provincial system in terms of sectors, programming level and/or geography. However, there is an important recognition that, despite the system intent, HE leaders often behave competitively.

> My own perspective has been that institutions aren't really in competition. Especially, well I don't see them that way because I tend to think of the institution as part of a larger provincial system and not so much competing for students, but trying to find the right mix of programming that will meet local needs best without spreading too thin or getting outside of it sort of - no institution can be all things to all people. I tend to view each institution has an important node within a provincial network and an institute that has to focus on local catchment primarily, but should also offer niche programming that might not be found elsewhere. And so, I haven't tended to view them - I just haven't viewed education as a competitive undertaking. That said, my experience working with Ecampus Alberta was that the downfall of that organization, despite its tremendous potential, was that too many players, and especially too many of the very senior players, presidents and vice presidents, within the Alberta system viewed their institution very much, as in competition with each other. To me that was actually that competitive attitude really prevented the system from being as efficient as it could have been.

Research participant 7 says that there is some competition in particular geographic and sector areas where there is greater commonality, such as business programming in high population densities. However, HE leaders tend to prioritize collegial behaviour over competitiveness.

> I would say it's very collegial. We obviously work with people from other institutions. The guy from [another HE institution] and myself we have many conversations throughout the year and we don't really have much of a problem sharing our own viewpoint of what's going on. But I think we also have to accept that, let's say in business programs as an example. We all offer business programs. Even in Calgary, [Another Institution] , [Another Institution], [Another Institution] , and [Another Institution]   and we all offer business programs and we are going after to the same students. You want the best qualified students because then you also have the highest retention rates. Because they have the highest likelihood of actually being able to complete the program. I would say it really depends. If we look at our health and public safety programs, I don't think we are really competing with anybody. But if you're looking at our business programs, especially our business diploma and business degree, we are absolutely competing with every single post-secondary in Alberta. I would say that's competition. If we didn't have competition, we probably wouldn't market ourselves. I think, we are collegial but there is certainly competition for students. We all have a marketing budget. We all market. We all buy billboards. We all buy transit. We all go with radio once in a while and that's to try and fill the pipeline.

Research participant 7 agrees that the HE system is designed purposely by the provincial government to reduce or even eliminate competition, at least among the publicly funded institutions. However, in some cases, there is motivation to compete in certain high demand sectors, programming levels and geographic areas.

> I think the government is looking at ways [to eliminate] competition because you want to cut down on the cost of post-secondaries. If you carve up the market and say you got to do this, you look after that, then maybe the competition is going to go away. But I would say it's there now.

Research participant 2 sums up reflections of the research participants as overall sentiment that, despite isolated competitive instances, Alberta HE leaders' first instincts are to collegially solve sector-wide IT challenges instead of seeking competitive edges through autonomous action.

> We are actually using 'Big Data' technologies in ensuring our students are successful. So, it's not a competition, per se in with other institution's recruitment. But we are in competition to make sure our students do better in terms of success rate……

The research participants reflect an overall collegial nature of HE, where leaders are inclined first to want to solve problems on a sectoral basis. While there are isolated instances of relative competitive markets, the motivating factors for IT adoption are most associated with addressing recognized performance gaps in institutional strategic priorities and meeting expectations of professional networks. Once measurable performance gaps are identified, HE leaders still need to propose a viable solution for a proposal to make the convincing case for funding and for the adoption of the technology.

4.5 Viable Solutions Identified

In addition to showing that there is a measurable performance gap, the deductive reasoning processes within IT governance requires that a viable solution be at least proposed and assessed. Decision makers need to gauge the cost of the potential solution and likelihood of success. The assessment is broken down to three broad criteria. First, is the benefit worth the initial and sustained financial cost of the proposed IT solution? Second, particularly applicable in experimental circumstances, is the proposed solution technically viable? Third, does the organization have or can it obtain and sustain the knowledge and skills to operationalize the proposed solution?

4.5.1 Technology Viability

Research participant 2 notes that there is a growing awareness among HE leaders that ‘Big Data’ technologies can be used to better understand interaction and engagement with multiple external and institutional stakeholders. Initial emphasis has been on potential and current learners but other stakeholders such as alumni, donors, sponsors, employers and regulators are also areas for potential engagement. Stakeholder engagement could be improved through a fuller understanding of cross-institutional relationships – not just department by department relationships with single stakeholders. Significant potential is recognized for the application of 'Big Data' technologies utilizing CRM data to uncover how HE institutions can form deeper stakeholder relationships.

> Based on my relationship management understanding there is a push on that side as well. Fundraising, I think traditionally has been seen as more of a relationship-based industry and more of an art than a science. But I think fundraisers are also recognizing that there's a lot of insights to be found in unstructured data and they are looking at that now. They are looking at emails and responses to emails, what people are saying about the institution, outside of the institution, outsider of their traditional approaches. So, there is some momentum at the institution in adopting 'Big Data' in identifying donors to better stewarding donors, understanding their needs and things like that.

Research participant 4 notes the potential of 'Big Data' technologies to direct resources to instructional approaches that are proven to be most effective for learning. Knowing what parts of a course are the most and the least engaging to learners demonstrates what content and strategies are most effective. However, while realizing the value, HE leaders like research participant 4 are unsure if such insight can be technically demonstrated.

> Yeah, this gets again to the…even if you knew, what would you do about it? So, there were a lot of areas where it might have been possible to do something, if we have … the people power to address some things. One area that I think there is quite a bit of potential for the academic side, for curriculum development, was to do a much deeper dive across a variety, even an institutions entire online course base and now that because people use learning management systems so much now, that its basically everybody's courses across the institution and try to get an understanding of what kinds of content were most impactful. Now you'd be using a bit of a proxy measure for that. Because you'd be looking at what did students go back to. But you could also do an analysis of the kinds of instructional strategies that seem to lead to the greatest success for groups of students. And I think that that was entirely unrealized - that's more of an institutional and individual teaching and learning kind of intervention. But I do think that that could have a really significant impact on decisions of where to put resources. What's worth taking the time for? What elements of a course seem to make the most difference? And, we didn't get to that level. It would take quite a bit of work, but I think it could actually yield real gains.

Research participant 4 further highlights skepticism that 'Big Data' technologies as they exist today can provide actionable data to improve learning curricula. There is confidence in the 'business side' of HE, such as learner recruitment and stakeholder engagement, but there remains suspicion that the collection and analysis of learner data may not lead to results that will enhance learning.

> [Technology viability] is actually my greatest concern. I think that some of the 'Big Data' collection, analysis, has a lot of potential in figuring out sort of consumer behavior in students. Courses that will be most appealing, things like that. On the marketing side, the business side, I think there can be likely some really very important gain. Very similar to what other businesses might gain from better understanding their clients. But when it came to the area that is more my expertise and where I work more, which is the actual teaching and learning and assisting students to be successful, it was still really challenging to know - no matter how much you collected, no matter how finely you analyzed it - what exactly you would do to make a difference. You know, in other words, even if we knew this, what difference would it make on the learning side. I remain a little bit skeptical.

Research participant 4 again emphasizes that 'Big Data' clearly adds value for non-instructional elements of HE, such as enrollment management. However, the technology has not convinced some leaders of its potential impact upon learning.

> I think that if you are trying to figure out how many sections of a course to offer and what modes to offer them in, some things like that - they're not quite same as market research, but trying to figure out how to best allocate a course mix, there is quite a lot to be learned. … When it sort of comes to the idea of actual learning

> analytics, right now I'm not yet certain what we would do differently - no matter how much we knew.

Research participant 4 notes that even if 'Big Data' were able to determine every learner's 'preferences', it would not likely be economically feasible to meet each of those preferences within individually designed courses.

> I think that really just to put it simply on the business and administration side, I really see a lot of value. On the learning side right now, it reminds me a lot of instructional design, where you're always trying to figure out individual differences and how you can best meet the needs of individuals with various different learning styles or, that's a bit out of vogue now but - different preferences. But no matter how much you know about all of those things, it's really not economical to try to meet all those preferences. And so no matter how much data you collect and how much you know about individuals, it's just really tough to do anything with that, and that's been a challenge of instructional design for a long time.

Some experiences in applying 'Big Data' to learning, in particular, have left some HE leaders underwhelmed. As research participant 4 confirms, descriptive analytics developed from LMS data simply confirmed what instructors already heavily suspected. Confirmation of what is presumed does provide value. However, given the costs and effort required to implement 'Big

Data' technologies, HE leader's expectations in terms of insight are often greater than what is yielded.

> I don't think there was a lot that was really unexpected. Some things that were confirmed, and this is where some of my - not skepticism but just limits on what I think is possible here. We … confirmed a lot of things that intuitively, you kind of already know. Students who spend more time on task tend to do better in the course. More interactive elements in the courses tend to lead to greater student performance. More instructor engagement with students, communications, tended to yield greater student performance, things like that. So, it confirmed and it gave us a way to collect data that showed those things to be valid. But they weren't things that that anyone experienced as an instructor or student really wouldn't already have known.

Skepticism that data analytics can improve the overall effectiveness and efficiency of learning through technology is reflected by research participant 4. The lack of confidence related to improving learning may be related to the relative lack of recognized measurable and meaningful indicators. Although ubiquitous in HE, grades are not universally accepted as the sole indicator of learning and there lacks a commonly accepted rubric to measure learner engagement. Without the accepted measures for learning and engagement, how can 'Big Data' technologies convincingly demonstrate effectiveness? Furthermore, research participant 4 also reveals the philosophic challenges of automation applied to a constructivist perspective on learning. If people learn through shared experiences, can those experiences be shared or simulated in automated environments?

> I think one thing is that, … what I learned over the years and that is, a lot of topics like this, and this is just one of many, that have kind of deal with the rise of information technologies and becoming part of every sector of the economy. There's often been the sort of underlying question: when is technology finally going to result in more efficient educational systems - meaning less inputs and greater outputs? I guess at this point, I do think that there still remains a lot of potential for technology and ‘Big Data’ analysis to create efficiencies at the organizational and administrative business side of the institution. I have yet to see anything in either data analysis or educational technologies delivery mechanisms and schemes built around them that I think will have much impact on efficiency. I have seen some tools that seemed to help with the effectiveness of education with flexibility and you could make some arguments that those things contribute to efficiency because they mean that people don't have to completely leave the workforce for training and things like that. I think those are important kinds of efficiency gains. But I remain very skeptical of the idea that technologies will allow people to become sort of autodidacts in a meaningful full sense. I'm open to being shown to be wrong there. But after sort of 25 years, watching this unfold, I haven't seen it yet that makes me think that that's going to happen

A more confident tone on the potential for 'Big Data' related to learning is struck by research participant 5. However, there remains concern over the availability of the specialized human resources needed to effectively build and utilize the technology. However, the increased development of education technology and information systems through AI automation may serve as a solution and drive increasing adoption in the HE sector.

> We've got to start building the algorithms, and potentially using artificial intelligence to layer into that. We know the challenge with 'Big Data' is that it takes big resources to manipulate the data, if you're doing it in a human way. It not only depends on the analytic capability of the individual but the number of individuals who can handle to actually work with the data that you're going to turn into useful information. I think the advent of artificial intelligence opportunities to allow the programming to start working with the data to give us better information for decision making

Reflecting concerns over internal capacity to adopt more advanced IT solutions, research participant 6 discusses a project to drive LMS adoption. Even getting accurate and reliable data on IT project adoption at the institutional level can be a challenge and it reflects a concern that IT adoption is not at the necessary level as more and more programs are delivered online.

> Here's how we're going to clean it up because data is really messy and here's how we're going to present the report. We did that over a couple of years and showed

> the impact of some of the initiatives that other departments and areas were working on in terms of driving adoption, like their training, their supports, their social or network interactions of trying to drive technology use, and said here's where you were in 2017 and here's where you are in 2018 and here's where you are in 2019. The outcome of that was, there was a lot of work going on in terms of training and support to move the needle on Brightspace adoption and then we use the data, as defined, to actually show that we were moving the needle institutionally. Our adoption as we defined it went from about 54% to about 70% over the course of two years, which from all of the research that I've done, which was mostly just environment scans and literature reviews, that was pretty significant. Moving adoption in any measurable way at an institute, if you can hit double digits like 10 percent is significant and close to 20 was something of a feather in our cap really.

Following up in these comments, research participant 5 identifies 'Big Data' scalability as an important issue across the HE sector. Given that HE institutions are facing similar challenges in terms of recruitment, retention, learner success and stakeholder engagement, the solution is best addressed not by individual institutions but by a sector-wide approach adopted by educational technology organizations. By integrating 'Big Data' capacity into current SIS, ERP, LMS and CRM products, institutions can adopt the benefits without the substantial risk of innovating solutions only for their HE institution.

> We use an Ellucian product, Banner, as many post-secondaries do. … I think the database, designers, in this case Ellucian Banner, what has surprised me is they have not brought to the table that type of capability. They still see themselves, I think, as the designer of the database, the architect of the database. That's yesterday's generation of capability. How do you work with your partners in this case [research participant's institution] to understand what they need and help them design or build in the algorithms and artificial intelligence to make use of the thing. I think we're at the forefront of post secondary on that space.

Scalability is also highlighted as a significant viability consideration by research participant 6. Recently at the research participant's institution, IT governance transitioned from the department level to an institute-wide approach. The increased scale provided greater efficiency across the institution. The same principle may be applied to education technology vendor organizations that address sector-wide issues on a much more efficient basis than individual institutions all attempting to independently resolve the same or similar IT challenges.

> Our [approach to IT governance] has really bubbled to light through that organizational change of moving from decentralized to centralized. Let's say areas like … finance or facilities or IT or employee services - this has happened in a number of different departments. One of the big things for us has been, we've been able to look at the data of things that have been previously managed independently and autonomously in a decentralized model and bring them

> together and say - one example is, we had multiple departments managing their own infrastructure so they had their own licensing and their own management consoles and their own, I'm not sure if you're familiar with, like VMware vCenter, and Vspheres and things like that, and so while it was great in the decentralized model because it was rapidly scalable and adaptable and within that business unit. When we brought them all together, we realized we were way over provisioned. We had way more management councils and all these things and way more than we need to actually run our environment. So, we were able to consolidate and find some significant efficiencies. We only actually need one of these and 10 of these, not 60. And there's absolutely no impact to operations. We can actually run it better, because of scalability and efficiencies that we found with the vendor. There were some key insights into how we were using data to support decision making and that's really easy when you bubble it up to the top to say, hey, we've done some digging and here's some data that actually supports, it's not going to cost you money it's going to save us save us money. So, all it took was a bit of that time in terms of gathering the data and bringing it together and saying, this is interesting. Let's move into that more and pull out more data and work with a vendor.

Specifically referencing ‘Big Data’ technologies to support learning, research participant 5 notes that HE institutions have a tremendous amount of learner data that has been passively collected and recorded in institutional LMS. However, there is no clear way to transform the data to planning insight or other useful information. Referring to the preceding comment from research participant 5, a sector-wide approach may be the best method to resolve the problem of generating practical learning insight from LMS data rather than relying on individual institutions because of the high cost and relatively smaller scale benefit.

> The other areas that we have huge data sets is in our learning management system. We use the D2L/Brightspace product. Again, we're just starting to get into what is the fundamental opportunity there from being able to engage participants in the program to be able to look at to their performance with quizzes, assignments, whatever the case may be. We look at trends to highlight where, for example if a number of students are doing poorly in one area is a very quick indicator to the faculty member that something is not being learned so what has to be done to address that? What are the trends around? If you start to map the background of the individual to the performance in the classroom we can start to yield, much better understanding about how we implement learning. What's the learning methodology and pedagogy performance? So, what I'm getting at is that we understand 'Big Data' is there and what we need to do is develop our capability to truly leverage that and really help the people in our organization and those, whether they be a support staff from the recruiting side or a faculty member, how do we take that information or the data and turn it into useful information?

Reinforcing the idea that scalability is an important consideration in the adoption of ‘Big Data’ technologies, research participant 6 observes that effort and expertise required for distinct data gathering and analysis solutions are often not viable for individual institutions only. The requests of a growing number of institutions for greater access to their raw data and the need for greater analysis capacities from educational technology vendors are also important considerations as institutions begin to look more towards vendors for sector-wide solutions to be applied at the institutional level.

> Within the LMS at the institute that I work for, it became a rather interesting time because at the time they only allowed a small little widget insight into the back end of all the data in the LMS. The institution, through this vendor, didn't actually have access to the raw data - like the full back end, all of the data for your institute. When this little tiny widget that gave us a few little snippets that they thought were relevant and it was, in my opinion, not very useful and not satisfactory. There was there was quite a push from ourselves and a number of other institutes to open up access to all of the data. It's, arguably, our data. It's our users, our faculty, our staff, using these tools and we would like to move into this space of looking at data for insights, or for actionable supporting evidence for decision making. And so, they did. So, around 2016, maybe, 2017, they opened up access to the back end through a secure portal that was called the ‘data hub’. And so they basically said okay here's your data hub, here's how you access it, and here's all the data sets that are within that space as well as API access to the raw data if you want to move in that direction for integrations, or API

> connections. And then, to be honest we were like this is great, thank you very much and then we went 'oh crap, now what?'. We had no plan. We had no data governance. We had no experts in data science, then we opened up this massive window into this space and we just went oh, okay, maybe we didn't think this through very well. What do we do now?'

The technical viability of 'Big Data' is a question for some HE leaders, particularly as it relates to measurably enhancing learning. There is more agreement about potentially effective 'Big Data' applications for functions such as recruitment and stakeholder engagement. Solution scalability remains a concern but education technology organizations may be best positioned to scale 'Big Data' across the sector rather than individual HE institutions developing their own IT solutions.

4.5.2 Financial Viability

Ability to access to financial resources is a substantial inhibitor to 'Big Data' technology adoption says research participant 1. HE leaders need to consider not only start-up costs such as training and change management but also ongoing and recurring costs such as IT licensing fees. HE institutions must assess return on investment against all costs to determine if the proposed technology is financially sustainable and represents value for money.

> Money is a quick one [inhibitor to IT projects]. Some of them are quite expensive to do. And staff resources, sometimes it's the knowledge of the staff. You may have ideas to do it, but the staff just either don't have the expertise or knowledge of how to action that. And so that's where it's getting the right folks around the table who can help you facilitate that. Money's probably the biggest one though… The initial outlay and the ongoing maintenance. Usually, if you get the initial outlay it is the most expensive thing when you look at implementation costs, and because you're factoring in your training costs, your change management costs, all of those pieces are part of that. When you actually look at what the annual licenses are for a lot of things, it's actually not that expensive. Yeah, that's what the companies get you on. Right? And so, we kind of have to look at both pieces.

Scaling 'Big Data' solutions across the HE sector would likely improve viability says research participant 5. In Alberta, the HE institutions face very similar challenges related to data collection and analysis. There may be an opportunity to approach individual institutional situations at a provincial system-wide perspective and urge educational technology partners to develop and implement solutions to province-wide issues.

> I think part of the solution going forward I believe has to be more partnerships in the post-secondary sector. Even though we are part of the Campus Alberta system we are in effect independent entities. We can collaborate as we choose. There really hasn't been a leadership mindset or initiative. I don't want to put too much

> on government because sometimes they can get in the way. But I think you need to enable partnerships. So, if the whole move towards 'Big Data' and analytics is expensive for each institution, that expense can be leveraged across the institutions to find common solutions so that there's that level of partnership I think needs to happen. We also need to get much more established partnerships in the IT world – Microsoft, Cisco, IBM. That's their business to be effective in 'Big Data' sets. So, I realize we are asking them to be partners and vendors at the same time. But analytics and 'Big Data' is what's going to allow these organizations to survive so they're going be good at it. This can help us develop those tools so that we can become good at it and use these tools.

From a research perspective, HE leaders may tend to initially judge 'Big Data' technology solutions as relatively low-cost because in a research environment many tools are easily accessible, even free, says research participant 2.

> As far as the cost of adopting the technology, I don't think there's too many issues largely because there is a vast amount of freely available tools however storage can be expensive, but it's going down in price as well.

However, use of the same research data tools on a commercial basis (even in an education technology scenario) is often much more costly compared to research purposes. The difference in expectations and actual cost may be a significant factor inhibiting 'Big Data' adoption. Research participant 3 reiterates that HE is used to 'Big Data' technologies for research applications but the higher price tag for educational technology applications can be viewed as very costly by comparison.

> If you're looking at 'Big Data' analytics in the research community for example, it's really, really worth it because a lot of the work, a lot of development happens internally. So, it's not costing the university as much…. In the industry, it's quite expensive because you're not using 'open source' packages for example, tools. You actually need to go to a vendor and have a relationship with that vendor in order to maintain your infrastructure with 'Big Data' tools and so forth. So, it's very, very expensive in that sense. In the institution, it's actually looked at again from a research perspective, it's absolutely fine because they know they're not going to spend a lot of money on tooling here because research can actually handle that. In the end you buy some tools but again, most of it is open source platforms. So actually, no problem there. But if you start looking at academics,…the registrar's office for example, if they want to have their own tool and so they have to [collaborate] with the business intelligence team in IT. For that you need to go to a vendor and actually buy a platform and buy the maintenance, the license software... It is it is expensive but I think the institution is still hesitant to go in that direction because of the cost and because of the running cost of maintaining this platform. It's really expensive so, in the times we're in right now, they are more dependent on researchers to do whatever they

> need to do, but they're not willing to invest much in this technology for other areas.

HE leaders are relatively familiar with 'Big Data' from a research perspective if not for institutional management processes. However, as 'Big Data' costs are often discounted for research purposes, fees for institutional uses may seem prohibitive. Likewise, HE leaders are aware that many institutions within the provincial system face similar data challenges and opportunities. A scaled province-wide approach to common data challenges would increase financial viability so many HE leaders encourage the HE system and educational technology vendors to provide sector-wide solutions rather than institutions investing in unique solutions.

4.5.3 Organizational Learning

In addition to technical and financial feasibility, HE leaders express concern over the required institutional capacity to initialize, maximize the utility and sustain 'Big Data' projects. Concerns range from HE leaders requiring evidence of causation over data correlation offered by 'Big Data' as a legitimate basis for institutional resource allocation and decision-making. Other organizational concerns include a lack or potential lack of skilled human resources needed for project implementation. HE leaders also are cognizant that an IT investment proposal driven by inductive reasoning is not likely to be funded without a convincing description of the problem and potential solution. Nevertheless, some HE leaders are confident in 'Big Data' technologies' long term potential in the HE sector and are discovering innovative ways to dedicate initial resources and evaluate pilot projects.

Research participant 1 gives an example where a 'Big Data' initiative, after the collection and analysis of the data, provided results that were not anticipated or projected in the IT governance process. Despite the unexpected results, the project was still deemed to have value because after data set analysis, the problem and subsequent potential solutions became better understood. The process shows how HE organizations, at times, learn to consider adjusting the goals and measures inherent in deductive reasoning-based IT governance structures when warranted by the data collected and analyzed.

> We actually have a program that we thought would result in a certain type of behavior and a certain type of outcomes from a student experience and student success perspective, and actually wasn't what we thought was going to happen. And so, it was working closely with an area that was running this program. We had to make some decisions - institutionally not me, individually - but some decisions had to be made about whether it was viable to continue with that program.

Expanding on the theme of unexpected results, research participant 1 relates how HE institutions can learn to shift from the proposed project hypothesis to another, based on the data collected and analyzed. However, the need within the IT governance structure to demonstrate at least a probable causal relationship explaining the relationship between project inputs and outputs is still expected. Although IT governance structures in HE are inherently driven by deductive reasoning, HE organizations are learning to adopt a more iterative approach and examine assumptions and re-evaluate project goals and measures as defined by the ongoing data analyses.

> It was obvious over time that what the expectation was wasn't actually happening, and so we started doing a review of the program and the process and all the different elements. And through that review, and looking at some very specific data around the kinds of students that were in that program, it actually showed other things that supported that this wasn't really a good direction to go in. So usually, there's something that triggers something for you to look at it to say, is this, is this successful or not? Here's some warning signs, or here's some really good things. What's the story that you want to tell? So, it kind of works on both sides, where we've had things that we've looked at it because it was working really well, we wanted to see how well it was working. And it showed us, you know, a whole other lens of information that we hadn't even looked at before. Wow, this is actually a really good story to tell.

Research participant 2 emphasizes the organizational significance put upon recognizing causation over data correlation. Data correlation is helpful in discovering causation but there is often a reticence among HE leaders to focus strategy or resource allocation justifications solely upon data correlation.

> I think one of the biggest challenges is for data science is distinguishing causality and correlation because they're used interchangeably. When you show correlation, most likely people think it is causation. I think decision makers want to see causation because it's an easier lever to pull. If x is doing y we just change x, or we do more of x. However, I think once we're able to explain that there is correlation in the data, we need to further dig deeper into the data to see if there's causation. Showing the correlation helps gain support to find causation if there is anything, but you're absolutely right, I think showing causation gains more support.

Often researchers themselves, HE leaders are generally more aware than other sectors of the distinction between correlation and causation for institutional decision making and perhaps, as a result, causation is a relatively greater priority in HE says research participant 2. Correlation is still useful to direct further research based upon causal insights. As a result, correlation is often used to direct further research to determine causation where correlation is enough to justify strategic planning and resource allocation.

> I think higher education is unique in the sense that when we explain the differences between causation and correlation, the leadership is also mainly made

> up of academics as well. So, in their research, they understand. We have an easier task distinguishing between the two. Perhaps outside of higher ed, that distinction would be difficult to make.

Research participant 3 agrees that organizationally there is a reluctance to trust the insights from 'Big Data' technologies because of: i) the priority put upon causation for decision making, ii) the distributed decision making structure common in the HE sector, and iii) the increasingly complex and adaptive nature of HE institutions. Even though correlation may be enough to convince some HE leaders to pursue a given strategy or allocate resources in a particular way, institutional resistance is expected based upon historically accepted decision-making processes. Furthermore, without causation, the risk for negative and unintended consequences is often perceived to be higher in complex and adaptive organizations like post-secondary institutions.

> This is one of the main things that probably made the acceptance or adoption of 'Big Data' analytics and machine learning, which is the proof that this model and what it creates actually works. People don't understand the concept of 'I gave you a lot of data and you didn't tell me what I need to do in a situation'. There is a lot of reluctance actually there. I'll tell you why, because over the years, these institutions have been around for like forty or fifty years and they went through each and every gyration of economics, social and technology changes. They built this knowledge inside the institution and it's very hard not to rely on, especially in

> tough times that they like the ones where we're going through. I'm saying that there is a reluctance, and for a good reason for institutions like [research participant’s institution] to actually accept a platform to tell them exactly what to do, strategically at least, for the institution. The efforts still matter but the decision I would say is made in siloed areas, like the education team, that can look at very specific things as I said earlier. They would look at stuff that you cannot do with the current knowledge and resources. If somebody tells them ‘all you can use this tool. It helps you do this and that.’ Then you can go there. But at a strategic level or institutional level is still very hard to actually go there.

Research participant 2 argues the opposite, stating that because the HE sector is not driven by profit and has a better understanding of research, there is less of a need for causation. But when it comes to making resource allocation decisions, are HE leaders as willing to commit to decisions without causation or is their acceptance of data correlation restricted to research activities?

> I think in the higher eds the need to see causation is a little lower - primarily because we're not driven by profits. And they think in research-based organizations and other non research-based organizations as well as mainly in higher eds the understanding is there that there's not always causation. Research does not always lead to what we are after. It could give us something surprising. People, I think, are a little bit more in tune to the fact that research at the end of

> the day - 'Big Data' analytics is a form of research so it could lead to something that we were not expecting. It could lead to not finding any causation.

Research participant 4 further discusses the need in institutional IT governance for both a hypothesis and a projected measurable result and return on investment. There is not much value for insight on a performance indicator in the HE sector if there are no technically or financially viable options available to resolve or ameliorate the problem.

> I think that what it would take [for greater 'Big Data' adoption is] a strong argument for what you could do differently if you knew something more. I think you have a really hard time - and I don't blame any administrator for this - if your desire is just to understand better what the correlations are. I'm not sure what the return is on that. If you had a pretty strong argument for if we knew x, we could do y, and you could really demonstrate that at least there was a reasonable likelihood that you could make interventions or change practice or change curriculum to increase student success. If you knew something, I think that would probably be enough. And that's just because education tends to work that way. But, I think it's really fair to say, well, we hope that if we understood this better that then we might be able to come up with strategies.

Research participant 4 reinforces the idea that, from an HE organizational point of view, IT projects without a proposed solution to a problem is a virtual 'non-starter' in terms of getting funding approved.

> Or at the very least, I guess you could call it hypothesis, but, you'd at least have to have some notion of what you could do differently. If you knew something.

Research participant 5 says that organizational cultures that are attuned to minimizing risk affect HE institutions' IT decision-making processes far more than correlation versus causation philosophic perspectives. The desire to avoid risk is a significant impediment to innovation across the sector but institutional priorities and understandings related to technology hazards are evolving.

> I think it speaks to a cultural inhibitor that is present in post secondary and I'm speaking in general terms but post secondary has never been accused of going too fast in areas that are really forward leaning. That's a cultural issue imbedded in the nature that we still are largely government funded. We have controls. We're not rewarded for being risk takers. I believe that is shifting quite quickly, as I mentioned earlier. I think COVID will force a different level of thinking around that to some degree. Hopefully we can find a way to generate that cultural shift in the organization, as we've proven we can do by going online. So, it's not that we can't. It's just that we just don't seem to be motivated. Culturally in many ways

> we are typically hierarchical and rule bound and hierarchical organizations typically have a cultural perspective counter to the innovative and creative mindsets. I think leadership and post secondary has to start seeking those type of mindsets and learning to take more measured engaged risks, and that can be done within the IT space for sure. Whereas before, I'm speaking to my perspective - post secondary, they [IT] were not seen as a business partner to enable things to happen. As long as our Banner system kept running the networks were running, PC passwords were protected and we were able to get password changes when we forgot our passwords, that was really seen as the nature of IT, in my perspective. I believe moving into the area with data sets, the opportunities for artificial intelligence, and those types of things, IT leadership has to, and it's not just IT, executive leadership has to embrace their role as an enabler of really driving future capabilities of a fast moving organization, and seeking out those tools to allow us to be not just a supplier of tools they are really servicing the needs of the organization and they are critically needed.

Continuing the theme of complex and adaptive organizations, research participant 5 says that HE leaders tend to infer causation to support institutional decisions based on correlation. Among HE leaders, there is an increasingly prevalent need to demonstrate the data backing a given decision. However, HE institutions are complex and adaptive organizations and there is a growing realization that unintended consequences are not necessarily undesirable. HE leaders are encouraged to incorporate all available data in decision-making and resource allocation and, over a longer term, arrive at a level of causal understanding rooted in intended and unintended consequences of a particular IT project.

> I would say from [the research participant's institutional] perspective we are moving towards more data or evidence informed decision making. That's the first step, recognizing data rather than opinion or suspicion is the means by which decisions are made. We are still low on the curve. Overall, I think we still probably are more correlation level - that we see something that we get drilled into the understanding of the causation. The correlation piece really is what I consider very linear decision making. You see something happen therefore it's probably what you draw from your past experiences why that happened. I think the data world, I hope, will start us to get into a more system level understanding about what causes something to happen. You talk about unintended consequences of decision making by not understanding the true nature of what caused that situation. If you look just too high level of data, you tend to be talking about unintended consequences being the bad outcome. I prefer to use the term 'unanticipated consequences' because you can get very good things come out of an unexpected situation. But if you don't understand when the good thing happens, you can't leverage that to grow it. On the same side, if there's something intended happens if you don't understand why, you may make the wrong decisions about what caused it and may result in further unintended consequences.

Furthermore, in complex and adaptive environments research participant 5 says data driven decision-making is becoming an increasing focus for timely decisions that optimize HE institutional performance based upon specific meaningful and measurements - not assumptions of causality.

> I think there is an opportunity for post secondary to perform a lot better than it has in the past. It will be rooted in our ability to make use of data, 'Big Data', and the analytics associated with that. So, we're making more correlation/causation based decisions, but making them more quickly.

The HE sector as a whole is at an initial point of the adoption curve related to 'Big Data' says research participant 6. As a result, a significant distinction is not always drawn between insights or decisions based on correlation or causation.

> In terms of the maturity level of data analytics, I would suggest that we, and myself included, are at a maturity level zero or a one on a scale of five, as you move through most maturity models. We're at the beginning. I'm not even in a place where I would feel comfortable reporting on the data that I have access to and make the statement that these variables are correlative and these ones are actually causative. I wouldn't be. I'm not at that place, and outside of maybe a data scientist, that'd be kind of tricky. Right now, we're at the exploratory stage of saying well here's the data. Here's what we think it's telling us. Here's the insights

> we think we're pulling out of it but we're being very transparent in that whole thing. We're all in the same boat in the decision-making process. It's probably time for us to start having conversations around, well maybe these are correlative and maybe we have to find some causation and I haven't had those conversations yet.

Among HE leaders there is a growing awareness that 'Big Data' technology can sometimes provide a much deeper understanding than what was available in the past of what factors are impacting institutions and their stakeholders. Research participant 5 says that a more holistic understanding is particularly important as HE institutions operate in increasingly complex and adaptive environments. 'Big Data' technology may also be a good check to make sure decisions are data-driven and not based upon casual observation or unchecked bias:

> I may be getting beyond the topic of this conversation, but it really speaks to the types of leaders that we need in an organization that understand that complexity. It is complex, and that decisions can't necessarily be drawn from a simple conclusion of an outcome based on what you see happening. Hopefully, the use of data and, speaking to the correlation causation understanding, will give a better understanding of what is that how is the system weighting something that's happening. If you see a, I'll give you an arbitrary example, if you see a decline in a program, you might think that it's because students don't see a future in it. But it might be that students are hearing the program is lousy. I'm not sure how you can

> collect the data but this is an example that to make decisions you need to understand the true causes underneath it and they're not necessarily visible and to really understand you need to be able to collect the right data to support the right analysis.

In addition to understanding the potential for 'Big Data' technologies, organizational learning can be challenged by the ability to recruit the required skills sets to effectively implement data analytics in the HE sector. Research participant 2 highlights the difficulty in recruiting and retaining the needed data science skills sets to implement 'Big Data' technologies:

> I think a bigger issue is attracting talent. So even if you have the infrastructure, people who know how to do this stuff, they're not cheap and we're not as competitive as the for-profit industry. I know of one department that has gone through five different data scientists in four years. They just don't stay. One of those reasons is it's difficult to keep them excited, just because of the slower adoption, but we're just not being able to pay the kind of salaries that the for-profit industry is paying.

Conversely some HE leaders like research participant 4 say recognition of the importance of data analytics is growing at institutions and, consequently, access to the required resources and skills is improving considerably.

> At the time that we did this at [the institution] I would say 'yes' [the institution had sufficient access to required skills]. I would say that if, if you'd asked me the same question, even just five years previously, I would have said 'no'. IT functions and our relationship with our vendor and just the sophistication of that part of the institution had expanded considerably. And I didn't have any concerns but even five years ago, I would have said there wouldn't be support and there wouldn't be a sophistication level to help us.

Research participant 6 indicates that HE leaders are improvising due a lack of dedicated resources for 'Big Data' technologies to provide access to predictive analytics. Reallocation of budgeted resources to pilot 'Big Data' technology solutions is occurring, recognizing of the importance of the technology despite the lack of initially budgeted resources through IT governance processes.

> We thought, well, maybe we can use some of the data in the LMS to build this around some of the tools and precedents that already exist. We just fortuitously had access to a data scientist within our applied research group that I just reached out and said 'hey I know you're working on other things and you have no connection to what we're doing, but if you're interested and have a little bit of time, and I'd love you to even just take a little doing and tell us if we're on-track or way off based and, thankfully, he was just another one of those people that was

> keen to be engaged and help so he came on board. We actually developed our own process of building a model using the data sets out of the data hub and, almost fumbling but we had the data scientist who was really good, our way through training a model to actually get to the point where we had prototyped an early-alert predictive algorithm that worked in our environment. We were scheduled to present the paper for the IBM SAS convention down in the States this March but it was canceled. I'll be presenting on that work this summer virtually now at our vendor's national conference event on how we use data to drive that. That was the example of how we moved into our first foray into predictive analytics.

Although some initial forays in 'Big Data' are at individual leaders' initiative and access to data science skills sets is a crucial consideration for HE institutional adoption, research participant 3 affirms that if 'Big Data' technologies are to be much more widely used in the long term, executive support (from the President and Provost in particular) is needed with technical backing from IT departments.

> In most institutions that I've studied it is the Provost's office and the President's office that are the business stakeholders for this kind of activities. And then there's usually institutions within the university or departments, actually, education, for example, are the ones that were taking most of the load as far as driving this kind of analysis. They need to work with the IT department as well

> as hiring some data scientists to create these models. It's machine learning models and the prediction algorithm as well. So, it's all within the institution.

In addition to executive, management and IT support, research participant 5 also notes the important role instructional staff have in 'Big Data' adoption. Faculty, especially, must feel the positive impact on learning from the adoption of 'Big Data' if technology is to be successful in the HE sector over the long term.

> One thing about culture, I'm not sure if I made this as clear as I could about the cultural elements in post secondary. Just by nature, I'm not picking on faculty but, they're used to a certain mode of operation. And so, if using LMS information and data, to be used we have to make sure that faculty understand or brought along as part of the journey. Again, you can collect all the data you want but you can't turn it into useful information. So, the use of 'Big Data' really is going to be driven by the ability to affect corporate change management and leadership, to really understand the value that helps faculty to do a better job without increasing your workload.

The capacity for organizational learning is an important factor influencing the adoption of 'Big Data' in the HE sector. While many HE leaders have experience with 'Big Data' in research, it is uncertain if institutional priorities and resource allocation could be influenced by data correlation or if demonstrated causation is necessary. Despite the need for a clear problem and proposed potential solution in deductive reasoning-driven IT governance processes, HE leaders' recognition of the potential for 'Big Data' is growing, especially in complex and adaptive environments. Recognizing 'Big Data' technologies' potential value, HE institutions are building technical capacity and initiating small-scale pilot projects using existing resources to test and demonstrate 'Big Data' effectiveness. Despite the challenge of deductive reasoning driven IT governance process, support for 'Big Data' is being championed and further cultivated with institutional executive, faculty and IT departments.

4.6 Ethics and Legal

From the literature review, ethics and legal concerns among HE leaders regarding the implementation of 'Big Data' technologies were anticipated at the outset of data collection. However, while research participants are aware of ethical and legal issues, existing institutional research methods to mitigate these risks are largely deemed as sufficient when applied to 'Big Data' technology solutions. Research participant 2 relays that there are not many unaddressed concerns at the institutional level and existing ethical and legal checks and balances are currently adequate. The research participant's viewpoint is potentially a reflection that most 'Big Data' projects cited produce descriptive level results only. More advanced predictive or prescriptive analytics and integration of automated AI or Machine Learning (ML) decision-making have much greater potential impact upon interaction with learners or other stakeholders and therefore have much larger possible ethical or legal impacts.

> I think that fortunately, it's very interesting the concerns that are being raised are by the analysts themselves. They're asking whether or not we're supposed to do this. So just because we have access to the data, what is the ethical use of that data? We are actually actively looking at and asking around whether or not there is some sort of an ethical framework for 'Big Data' analytics. So far, we haven't had a whole lot of issues raised. We haven't had any issues raised. There's only one instance where there was a survey thing. Your institution's part of it as well. But there was a government survey done that did not allow opting out of survey questions. Those kind of ethical dilemmas exist. Fortunately, it wasn't as big of an issue, students still had options to opt out of the survey completely. But perhaps this is a function of the maturity of the organization in 'Big Data' analytics. We haven't done a whole lot yet. So hence, not a lot of issues have come up. I'm not familiar with them. Perhaps this is just because I'm disconnected from this area, perhaps more on the medical side or on the Faculty of Medicine side, there are more issues raised, but we worked very closely with research ethics or REB. So, they do a very good job at telling us what can and cannot do.

Reflecting a common perspective among HE leaders, research participant 4 demonstrates that, at the current utilization sophistication levels, data analytics can be effectively managed within existing ethical and legal frameworks at the individual HE institutions:

> I have heard some people argue about that. [ethical concerns about data collection and analysis]. But, within the groups that have done this there wasn't any real concerns about it because the data being collected wasn't any different than what we would have done with any instructor, watching over their class.

Reinforcing the view that existing frameworks are currently sufficient, research participant 4 says that data gathering and analytics that are currently being applied at the descriptive level are comparable to other ethical research processes.

> We didn't run into any concerns. The data that was being collected wasn't different than what the system collected anyways. So, we weren't inventing new things to collect. We were just looking at ways to analyze it. The one issue that did come up that was important was when you get to the point of analyzing the data you do have to be just a little bit careful about who has access before it's anonymized. So in our case, we did have some student interns working on a couple of projects, and we had to build in some safeguards where a faculty member did some cleaning of the data before we let students get involved with analysis. And that was not a new issue we've faced that challenge with some of the research projects previously where it was just information that you wouldn't normally have students have access to - more about other students’ performance, things like that.

Noting similar experiences to date, research participant 7 states that legal or ethical issues have not been a particular challenge to date and the institution is able to use existing legal and ethical IT and research protocols to meet risk mitigation or compliance requirements:

> We, in market research, we adhere to the market research industry association guidelines when it comes to research and how you treat response data. Everything is always confidential. And in terms of the data from our office of the register there is a student ID associated with it. We would never divulge any personal information as for one particular student. Everything is locked down in aggregate. And even when it comes to course evaluations, if there's less than six responses, we will not release the results because we don't want any instructor to be able to identify who the students may be based on the comments or whatever they get. Data is secure. The use of personal data has never been an issue.

Noting the importance of ethical and legal considerations but recognizing that HE institutions are generally at the more rudimentary level of 'Big Data' application, research participant 5 says that HE is just now developing a greater understanding of the ethical implications of enhanced digital data gathering and analytics. Privacy and data breaches continue to be a priority concern with all technologies including 'Big Data':

> The whole aspect of ethics and privacy has really come to the top of the agenda in the last, certainly last few months here. I mean if you think about, even the

> COVID tracing app, is that a violation of somebody's privacy that you're tracking contact with somebody else? So, is that privacy waived by the public good? So fascinating challenges and ethics. Have we thought about it? I think we started but I'm not sure we necessarily understand it, to be honest. Our data sets that we collect really, well let me back off a little bit, we are very careful about now. Student marks, for example, we do know that there is an element of ownership and privacy around student marks. When I was at university, the marks for a course were posted on the board. You can see everybody else's marks. You would not get away with that today. There has been a movement away from that. We do consider, we were not calling it 'Big Data' yet, but it certainly is part of our framework that we have to be very careful when we're protecting the privacy of information of individuals, students or employees. So, the awareness is there. I think we've got to be really careful about how we use the data in a manner that protects the ethics and privacy of individuals, while still achieving the outcomes of program…. We're at the ground floor on this.

Acknowledging the multitude of ethical and legal issues, such as potential unconscious bias or discrimination incorporated in algorithms developed directly by data scientists or indirectly through AI or ML, research participant 3 notes some potential challenges and that researchers must be wary of bias as well as ensure access to sufficient data for sound decision-making.

> I think initially, when this machine learning and artificial intelligence became well known, people thought that you can use the same model to solve many similar problems. It turns out there's biases in the models that you're using for certain problems. To give you an example, for security applications even within a university environment the initial seed of images that the modeler used to model face detection and use it to determine if this person is a security risk. Most of these initial averages of detection errors were very high for African-American versus white people. We call this a bias in the model. It is not intended to be that way, it just happened to be that way. The result is that the detection process, created an issue because there was a lot of errors in the data model. If people are working on this and these models in particular they're really very cognizant of the models that they create like model biases for example. You can see today even with COVID-19, analysts are trying to create a 'Big Data' analytic model to look at the curve and see how much data they need. That's why you keep hearing people talking about you need more and more data. So one of the issues is - do you have enough data? That's the main issue with this approach. Do you have enough data so that you trust the outcome or the insights that you get out of your 'Big Data' analytics experiments?

Recognizing ethical and legal issues that are emerging particularly as AI and ML is increasingly being applied to 'Big Data' technologies, research participant 6 notes ethical and legal concerns are not widespread at HE institutions currently. However, transparency and security in data collection and analysis are likely to become significant if not dominant issues in the near future:

> I'm sure anyone who's even dipped their toe in 'Big Data' is aware of rights - it's privacy, it's access to data, it's all the dangers inherent with almost that fear of relinquishing control or that intelligence to an algorithm or a machine. I personally didn't encounter any significant or very harsh opposition to the idea as in you can't do this because it's wrong sort of thing. But I think that was in part due to the way we approached it and that was keeping transparency at the forefront. We started with engaging our faculty groups to say, with the adoption initiative, it was at the beginning to say listen to what students are telling us in in crazy numbers, like 98% of students, want more learning accessibility in the LMS. This isn't a 'us' thing. This is a student thing. They want this and so we're trying to move in that direction, and we think it's a good idea supported by the research. Having those kinds of conversations about how you're going to transparently use data is really, really important. It helps diminish those barriers and have the productive conversations around well how are you going to do this securely? How are you going to do it confidentially? How are you going to do this reliably and in a way that is able to be validated? How is it going to be for cyber security and access to information and all of those things that, that are good things to have? It all boils down to data ethics. You have to abide by the law. You have to abide by cybersecurity. But also to abide by the culture and the environment that you work within. Ethics isn't just about the law. It's about operating within

> your societal, and cultural context. Building those relationships and having that transparency into the process and asking for insight and feedback and thoughts really goes a long way to preventing it from getting to that oppositional environment.

Research participant 6 later continues to affirm that legal and ethical considerations will dominate once institutions begin using predictive and prescriptive analytics and learners are more directly impacted by understandings developed through 'Big Data' technologies.

> Yeah, and we haven't gotten there yet. So just to be clear, we did a pilot prototype in a very closed sandbox, where we were using data from the previous year to train the algorithm and then see if the algorithm accurately predicted which students went on at risk from last year. We did not deploy this out into a live production environment so we actually didn't engage in students and faculty and say hey the algorithm said you're at risk so you need help. We haven't gone down that path and I'm sure when we even propose the idea, there's going to be a lot of conversations around data ethics.

As research participant 6 further notes, when legal and ethics considerations associated with 'Big Data' become more prominently understood, IT governance will become much more active in risk mitigation and legal compliance.

> Exactly. I mean we had I think every institute went through the same thing when there was a few of those over the last few years where every post-secondary was like okay, we cannot be the next headline. What do we need to do to make sure?

Contrary to early expectations from the literature review, legal and ethical concerns do not significantly inhibit 'Big Data' technology adoption in the HE sector. Some research participants noted potential concerns as 'Big Data' continues to develop but there is a consensus that current legal and ethical practices for general research practices meet the required standards to protect learners and other stakeholders in data collection and analyses processes. However, HE sector utilization of data analytics to date focus at the descriptive level and not much work has been done to produce or apply predictive or prescriptive analytics. If and when HE institutions use 'Big Data' to allocate resources or constrain stakeholders' choices based upon demonstrated likely results, the legal and ethical environments become much more complex. While legal and ethical concerns do not currently inhibit 'Big Data' adoption, they are likely to be significant issues as the technology grows in use and sophistication.

4.7 Conclusion

Using a constructivist approach to grounded theory as outlined by Charmaz (2014), a conceptual framework is developed by examining the perspectives of HE leaders to explain the adoption rate of 'Big Data' technologies in the sector. After receiving relevant REB approvals, the 'saturation-point' for data collection of seven semi-structured HE leader interviews is reached. Data analysis conducted concurrent with data collection yields 102 distinct data codes which are applied to the verbatim interview transcripts a total of 723 times. Further analysis produces eight secondary codes and nine memos recorded as the researcher reflects upon the collected data. The emerging themes, grounded in the collected data, form the conceptual framework to explain HE leaders' perceptions towards 'Big Data' and their adoption rate in the sector. It is discovered that the deductive reasoning processes that IT governance overwhelmingly uses to mitigate investment risk and maximize cost-effectiveness effectively discourage 'Big Data' adoption which depends upon inductive reasoning approaches. Ideally, as per *Figure 3: Conceptual Framework 'Big Data' Technology Adoption / Governance*, a proposal for the adoption of 'Big Data' technologies would receive funding for data collection and analyses if legal and ethical concerns could be managed. Once completed, the IT problem and potential solutions would be identified and assessed. However, current IT governance driven by deductive reasoning demands that a significant problem and feasible solutions be identified prior to significant investment. Constrained by the need for risk mitigation, 'Big Data' projects that are approved are those that are best able to address the deductive-reasoning requirements inherent to existing IT governance processes. These criteria are demonstration of: i) measurable results, ii) identified performance gaps through collegial networks, competition and/or institutional performance networks, iii) solution viability in terms costs, technical operation and organizational capacity to innovate, and iv) manageable ethical and legal risks. As noted in this chapter, some 'Big Data' project proposals can meet the requirements inherent to deductive reasoning inherent to IT governance. Other smaller scale 'Big Data' projects are piloted using already budgeted resources and thus can, at least temporarily, side-step IT governance processes.

As 'Big Data' become more widely understood in the sector, many HE leaders view adoption as best addressed at the system level – not at institutional levels. As the data challenges and implementation costs are relatively similar across institutions Big Data' solutions are almost certainly best addressed at a sector level by system-wide or educational technology organizations by further integrating data analytics capacity in existing SIS, CRM, LMS, and ERP tools. Further reflection on the conceptual framework and how it responds to the research questions is provided in Chapter 5: Study Results.

5.0 Chapter 5: Study Results

The purpose of this study is to construct a conceptual framework, grounded in the perspectives of HE leaders in Alberta, that describes and explains how and to what extent 'Big Data' technologies are being adopted in the sector. As noted in Chapter Three: Research Design, the study employs a constructivist approach to grounded theory as detailed by Charmaz (2014). Given a current lack of theory that effectively describes or explains the relatively slow 'Big Data' adoption rate in HE, grounded theory is ideally suited to the research questions. Grounded theory explains human or social processes from data collected with research participants. The results of the study focus not upon generalizable findings but explaining how a phenomenon is understood, expressed and applied by the research participants (Creswell & Gutterman, 2019). Through the conceptual framework, the insights provided are related to each of the study's secondary research questions, as follows:

1. What do higher education leaders in Alberta consider the main elements for and against in 'Big Data' adoption or further development?

2. To what extent are Alberta higher education leaders motivated by competitive enrolment and/or cost-effective learning considerations that could be developed by 'Big Data' applications?

3. Do Alberta higher education leaders believe that meaningful business decisions can or should be influenced or even made by 'Big Data' applications?

4. What ethical considerations do Alberta higher education leaders integrate into the data collection and analyses processes of 'Big Data' initiatives? and

5. How prepared do Alberta higher education leaders consider their respective institutions are to implement 'Big Data' successfully from technological and change management persepctives?

The study demonstrates that Alberta HE leaders recognize the significant potential of 'Big Data' and the increasing impact of 'digital transformation' overall within the HE sector. These leaders are conscious that digital technology is increasingly the principal medium through which HE organizations interact with learners, potential learners and other stakeholder groups. As new technologies emerge, HE leaders seek different approaches to improve cost-effectiveness and recognize that they operate in 'data rich' environments. SIS, LMS and CRM information systems automatically collect vast amounts of data with each interaction with learners, prospective learners and other stakeholders. The question for HE leaders is how to best use their collected data to improve the quality and cost-efficiency of services to stakeholders? HE leaders are particularly convinced of the potential impact of 'Big Data' technology on the 'business side' of institutional operations, such as learner recruitment and donor engagement. What is much less clear is how these technologies can improve individual learning processes. Furthermore, while HE leaders recognize the potential and increasing impact of 'Big Data', most readily agree that their organizations are in early stages of adoption. The broad question at the root of this study's conceptual framework is, what is impeding adoption among HE leaders?. The study will examine the broad questions as the sub-questions are addressed sequentially in the next section.

5.1 Study's Questions

- *Question 1: 'What do HE leaders in Alberta consider the main elements for and against in 'Big Data' adoption or further development?'*

On the surface, the environment appears ripe for rapid 'Big Data' adoption by HE institutions. While HE leaders recognize that institutions operate in a data-rich environment, they also realize that institutional data is an underutilized resource and that 'Big Data' technology adoption lags behind other sectors. Research participants recognize that HE institutions are only now beginning to translate their data stored in SIS, LMS and CRM systems into information that can improve institutional performance. Despite recognizing the importance of becoming data-driven organizations, HE leaders agree that the pace of transforming available data to 'actionable information' remains relatively slow.

The proposed conceptual framework as identified in Chapter 4 Data Collection and Analysis describes and explains how and why HE lags behind many other sectors in 'Big Data' adoption. In summary, IT governance structures at HE institutions use deductive reasoning to recognize and mitigate risk. A successful proposal for an IT project must address the following broad criteria: i) measurability, ii) performance gap, iii) viable solution and iv) ethics and legal compliance. All 'Big Data' projects identified in the study are relatively small in scale but do address each of the conceptual framework elements. Particularly noteworthy is the emphasis from all the research participants on the 'business side' for 'Big Data' projects. For example, project proposals related to student recruitment, service level planning or donor engagement are recognized as better able to meet the required criteria for IT governance approval. Conversely, ‘Big Data’ project ideas related to student learning do not address the required deductive criteria in the identified conceptual framework. While it may be argued that benefits to learning are measurable through demonstrated student achievement and retention, it remains unclear among HE leaders what are the performance gaps and potentially improved new instructional or learning methods. Moreover, without a proposed solution contained within a project proposal, the potential for funding is very remote. As research participants noted, even if a learning performance gap is identified, if there is often no clearly understood solution that would measurably improve performance. Thus, with respect to learning and instruction, HE leaders are not convinced that ‘Big Data’ technology can discover a performance gap. Even if the performance gap exists, in most cases it is unclear to the HE leaders interviewed how ‘Big Data’ could provide viable courses of action to improve learning.

Central to the recognition of ‘Big Data’ technology’s value to the business of HE, is the growing importance that HE leaders place on non-governmental revenue sources. Domestic and especially international learner tuition together with other revenues such as donations and sponsorships are increasingly motivating factors among HE leaders as provincial revenues proportionately decrease over time. The study found that technology solutions that can increase non-government revenue are a high priority among the research participants. In addition to the recognized current value on the business side, ‘Big Data’ is also widely recognized among HE leaders as having a significant future impact upon the sector. Research participants judge that they and their colleagues are only beginning to explore ‘Big Data’ technology’s potential in HE’s data-rich environment. Given the sector-wide expectation of substantial future impact, research participants relate how 'Big Data' may be widely supported within the HE sector despite not being well understood. However, while 'Big Data' technology is recognized as an important trend, it often is obstructed when project proposals cannot meet the requirements of the IT governance process. For example, at the proposal stage, a 'Big Data' project often cannot identify the specific problem that will be addressed (much less a solution) until the project is underway and the data is collected and analyzed. Instead, the research participants describe various workarounds such as smaller scale 'Big Data' projects that use existing budgeted resources or other types of informal collaboration designed to allow HE leaders to pilot 'Big Data' initiatives while avoiding some institutional IT governance processes.

In addition to the disadvantages inherent to a technology driven by inductive reasoning in a deductive governance ecosystem as described above, another related major impediment to the adoption of 'Big Data' is the issue of scale. A typical 'Big Data' project is a major undertaking and investment for any HE institution. Moreover, at the initial proposal stage, 'Big Data' projects represent a major risk with an often imprecise institutional benefit. However, paradoxically, the risk associated with 'Big Data' projects can be proportionately reduced by increasing scale. Compared with the high risk associated with 'Big Data' projects at a single institution, the incremental cost of adding additional institutions is relatively modest. Therefore, HE leaders propose two approaches that can increase 'Big Data' technology adoption by increasing scale. First, the provincial government can promote a more collaborative approach among the all the publicly funded HE institutions in the province to promote and fund 'Big Data' adoption. However, it should be noted that, given the many different IT systems employed by HE institutions in Alberta, the technical viability of this option is questionable. Second, educational technology vendors could integrate 'Big Data' technologies with their current SIS, LMS and CRM products. The approach allows education technology vendors to develop and offer sector-wide solutions at a manageable cost to HE institutions.

In terms of 'Big Data' adoption and organizational learning, HE leaders demonstrate evidence of what Senge (2006) characterizes as 'adaptive' learning. Because 'Big Data' technology is relatively new, there is no established processes for adoption and each institution is left to develop its own 'blueprint' through formal and informal collaboration among engaged groups. Examples are provided where 'Big Data' projects are initiated by HE leaders who see the potential applications of the technology. However, as discussed previously, many of the projects cited are small and informal and, therefore, do not require additional budget resources and additional IT governance approval. HE leaders see these small scale 'pilot projects' as adaptive ways to test 'Big Data', while avoiding the incompatibility of such projects with current IT governance processes.

In summary, while HE leaders agree 'Big Data' will eventually have a significant influence on how HE institutions are managed, adoption is significantly hindered by the conflict between the inductive thinking that drives the technology and the deductive IT governance processes which authorize implementation. In short, 'Big Data' project proposals are not able to meet all the requirements needed by current IT governance processes to mitigate risk. Costs and undefined project benefits remain the major unresolved risks in a typical 'Big Data' proposal. Despite these challenges, HE leaders demonstrate adaptive organizational learning by implementing relatively modest pilot projects through either the formal IT governance process or through informal group collaboration. The approved pilot programs are typically related to the business side (as opposed to learning) because the potential benefits are more easily measured.

Question #2: To what extent are Alberta higher education leaders motivated by competitive enrolment and/or cost-effective learning considerations that could be developed by 'Big Data' applications?

In the study, HE leaders consistently opine that competition with other HE institutions is either a non-factor or a low-level motivator. Instead, there is substantial evidence of a more collegial approach. HE leaders demonstrate a concern for ensuring their institution is perceived among their peers as an innovator. Conversely, demonstrating to peers an ability to compete for a limited pool of learners is not a top priority. While the research participants themselves did not view HE as a particularly competitive sector, in a few instances it was noted that higher level managers are seen as more competitive than the rest of management or faculty. Some HE leaders acknowledge a degree of competition in certain limited circumstances such as for recruiting business learners for similar programs within the same geographic area. Yet, overall, despite some limited forms of competition, management approaches noted in the study are considered by HE leaders to be far more collegial than competitive in nature.

HE leaders' views of the provincial government's role in promoting collaboration and/or competition among HE institutions are noteworthy. The impact of the Province of Alberta's IMAs with each post-secondary institution is not yet fully understood by HE leaders. HE leaders are aware that decreasing IMA performance indicators will eventually negatively impact provincial funding. However, how performance will be measured is still largely unknown. Furthermore, there is a widespread concern among HE leaders that IMAs could easily become counter-productive if measures have a single focus (expenditures, for example), at the expense of other critical management functions such as efficiency or relevancy to stakeholders. HE leaders also recognize the tension inherent in the province's approach between efficiency and concurrently ensuring a minimum level of service across a wide number of sectors and geographic areas. Emphasizing the point, some research participants suggest that the province's objective is to, somewhat paradoxically, increase HE institutional efficiency while at the same time as decreasing competition amongst the same HE institutions. While it is clearly not sustainable to fund numerous public institutions competing in the same marketplace, the IMAs are intended to act as a competitive proxy to encourage greater efficiency, more relevancy to the labour market and improved learning. It remains unclear to HE leaders in Alberta how the province's proposed tools like the IMAs will translate to their policy goals.

Given the perceptions and expectations of HE leaders regarding competition among post secondary institutions, it is not surprising that there is only low level 'generative' organizational learning as defined by Senge (2006) demonstrated in the study related to 'Big Data' technology. Institutional improvements are most often measured towards self-defined strategic goals and not directly assessed against the performance of other post-secondary institutions. There is evidence that the collegial environment drives innovation as HE leaders want to be effective in areas deemed worthwhile by their peers. However, competition is typically understood as each HE institution's capacity to facilitate individual learner success overall and not by efficiency and effectiveness comparisons with similar HE institutions. Therefore, performance is most often not understood and measured against competing HE institutions but instead gauged against self-determined milestones and professional expectations. While competition among post-secondaries exists, it is not a prominent driving factor for 'Big Data' technology adoption among HE leaders. As a result, risk-taking and generative organization learning is not prioritized in HE as much as other more competitive sectors.

In summary, competition among institutions is a consideration in some cases, but is usually not a major factor motivating the HE leaders in the study. Annual targets are generally self-determined within institutional strategic plans. HE leaders are influenced broadly by collegial expectations in the sector, but direct comparisons with competing institutions are infrequent and restricted to the business side of particular programs in specific geographic areas. There is an expectation among HE leaders that the incoming IMAs developed with the Province of Alberta may develop a pseudo-competitive environment to encourage greater cost-effectiveness. However, many details remain unclear as the IMAs are just now being implemented and HE leaders' expectations for provincial budgetary support has not yet changed as a result. It is clear that competition does not drive generative learning with respect to 'Big Data' technology in HE institutions. With respect to IT governance, HR leaders are motivated primarily by risk mitigation over potentially developing a competitive advantage. The deductive-thinking embedded in IT governance processes serve to keep 'Big Data' technology on the margins until project risks can be sufficiently abated.

Question #3: Do Alberta higher education leaders have faith that meaningful business decisions can or should be influenced or even made by 'Big Data' applications?

Alberta HE leaders in the study clearly recognize the potential influence of 'Big Data' technology for improving institutional decision making. As discussed in Question #1, the research participants recognize that the potential is just now beginning to be better understood. One of the main challenges that emerges, as highlighted in the conceptual framework, is that typical IT governance processes in HE heavily deter an inductive approach and systems thinking. The deductive approach values a proposed viable solution to a distinct and measurable problem. Conversely, the inductive approach first requires an analysis to understand the inter-relationships of the collected data before even the problem, much less a proposed solution, can be identified. As a result, governance models based on deductive reasoning typically rely upon data snapshots to identify measurable performance gaps rather than systems thinking that focuses more on identifying how different data sets interrelate over time.

HE leaders express satisfaction in the 'Big Data' projects identified in the study, and they are particularly willing to make management decisions related to the business side, such as recruitment and stakeholder engagement, based on the results of these projects. There are no examples in the study of HE leaders changing practices related to instruction or learning based upon a 'Big Data' project, and there are no examples of large-scale 'Big Data' project that have gone through an IT governance process based upon inductive reasoning. 'Big Data' projects do heavily influence many Alberta HE leaders in specific business areas of institutional management. However, current 'Big Data' projects are relatively small and tend to address an already identified and measurable problem and include a potential viable solution. Barriers in the IT governance process, limit large-scale 'Big Data' projects based upon inductive reasoning.

In an inductively driven process, data collection and analysis are completed prior to identifying the performance gap and potential solutions. The inductive governance model then further challenges HE leaders' typical decision-making processes because it can be unclear if the discovered data set inter-relationships are causal or correlated. However, because no example is found, the study cannot make an assessment on HE leaders' confidence in large scale 'Big Data' findings. Nevertheless, the absence of such large scale 'Big Data' projects indicates that only when the projects can mitigate the inherent risks at the proposal stage, will HE leaders be significantly challenged to consider correlation versus causation issues in their management decision making.

In summary, HE leaders are confident of making business side decisions based upon 'Big Data' that meet the requirements of a deductive IT governance process. However, the number of cases meeting the requirements are relatively few. Projects focus on the business side of HE institutional operations and no changes in practices related to learning were identified by using 'Big Data'. Because all the 'Big Data' projects are relatively small in scale, HE leaders are not yet presented with significant epistemological questions of data correlation versus causation. Decision making is still made largely on the basis of data 'snap shots' and not upon continuous inter-relationships of data sets. HE leaders agree that 'Big Data' technologies will play an expanding role in HE institutional management going forward. There are some examples of organizational adaptive learning where small groups are collaborating on 'Big Data' projects despite the restrictions put in place based upon institutional IT governance process that prioritize deductive reasoning. Yet, the number of business decisions currently being made in the HE sector using 'Big Data' technology is small and limited in scope.

Question #4: What ethical considerations do Alberta higher education leaders integrate in the data collection and analyses processes of 'Big Data' initiatives?

The study shows that Alberta HE leaders are aware of many of the ethical considerations associated with 'Big Data', particularly as they relate to data collection. The consistent view is that although the risks are significant, current institutional research practices effectively address ethical concerns related to 'Big Data' used to date. Mitigation methods include ensuring data security, requiring informed consent for provision of data, and using existing, anonymized and/or aggregate data sets. However, as many HE leaders also recognize, institutions are only now beginning to explore 'Big Data'. The projects identified in the study are almost exclusively restricted to 'descriptive' analytics that define and characterize an identified phenomenon through the collected data. Other types of analytics such as 'predictive' or 'prescriptive' have greater potential to encounter new and more complex ethical dilemmas. For example, predictive analytics could calculate expected academic performance based upon previous online activity, such as number and duration of LMS log-ins. Those learners with greater success likelihoods may be prioritized for course or program admissions. Likewise, to better ensure academic success, prescriptive analytics could be used to require some learners to take additional programming based upon correlations identified in learners' SIS, LMS and/or CRM records. The requirement may not be based on a learner's academic success to date, but upon other data points in a leaner's record that correlate to demonstrated academic risk. However, basing learner recommendations or requirements on anything other than academic records entails a level of digital surveillance that most learners are not aware of and would require informed consent.

In addition, the validity of algorithms used to identify and prioritize correlations are subject to conscious and unconscious bias and human errors. The vulnerability inherent to predictive and prescriptive analytics is that institutions may require learners to undertake specific activities or restrict program access for some learners based upon outdated data or a flawed correlation. Furthermore, the biases in data analytics are often insidious due to a common misconception that decisions based ‘Big Data’ technology are impartial because they are automated and grounded in empirical data.

Although HE institutions are not yet implementing ‘Big Data’ technology at an advanced level, it is concerning that the common understanding among HE leaders participating in the study is that current research ethical and legal protections are sufficient going forward. As institutions integrate more advanced ‘Big Data’, a more thorough understanding of digital surveillance, informed consent and the source of bias and ‘truth’ within data analytics is needed among HE leaders. In addition, the concepts of ‘duty of care’ as noted by Prinsloo & Slade (2014) as well as engaging learners and other stakeholders to actively participate in analytic development and validation as discussed by Kruse and Pongsajapan (2012) should be better understood as ‘Big Data’ technology is further developed in the HE sector. Without HE leaders that are more more fully engaged in the development and implementation of quickly evolving ‘Big Data’ technology will rapidly and inevitably lead to ‘black box’ management processes - where decisions are increasingly automated with little or no human mediation or understanding of algorithms and their potential flaws.

In summary, HE leaders participating in the study demonstrate an appropriate understanding of the ethical requirements for the preliminary stages of ‘Big Data’ adoption. However, a new level of ethical considerations loom for HE leaders as analytics move beyond the descriptive to predictive and/or prescriptive in nature. HE leaders need to fortify their ethical understandings in relation to learners and other stakeholders as ‘Big Data’ technology grows and is increasingly integrated to HE operations. Without a more robust ethical understanding decisions as HE institutions gradually turn to outsourcing more advanced ‘Big Data’ to education technology vendors and key decisions are justified by algorithms that are not fully understood by HE leaders.

Question #5: How able do Alberta higher education leaders consider their respective institutions are to implement ‘Big Data’ successfully from technological and change management perspectives?

As many research participants noted, they consider their institutions to be in the early stages of ‘Big Data’ technology adoption. While some noted barriers related to accessing and retaining the required IT expertise, the main inhibiting factor to larger scale adoption is a perceived high cost coupled with ambiguous benefits of ‘Big Data’ technology. All the ‘Big Data’ projects identified by the study are relatively small in scale and originate with clearly defined IT problems and specifically recommended solutions. An instance of a truly inductive approach where data is collected and analyzed prior to identifying the IT problem and potential solutions was not found in the study. Instead, either the ‘Big Data’ project was implemented with already allocated resources, thus avoiding the IT governance process all together, or the approval conformed to the deductive approach of current IT governance systems (see *Figure 3: Conceptual Framework 'Big Data' Technology Adoption/Governance*). There is little evidence found that overall access to financial or human resources is a significant inhibiting force for 'Big Data' adoption. Instead, the challenge is that 'Big Data' projects are unable to clearly identify the problem, much less the resources required for a solution, until the project is underway with the data collected and analyzed. In short, the required resources for 'Big Data' project implementation may be available within current HE institutional budgets but the resources needed are not fully known at the time the project is under funding consideration. As a result, because they are unable to mitigate project risk and prioritize investment, IT governance bodies are very often unable to seriously consider most 'Big Data' project ideas.

A summary of the research results is provided below in Table 3: Summary of Study Results:

Table 3: Summary Study Results

Questions:	**Results:**
Question 1: 'What do HE leaders in Alberta consider the main elements for and against in 'Big Data' adoption or further development?'	• Leaders recognize the HE sector is a data rich environment conducive to 'Big Data' technologies • Leaders recognize significant potential benefits, particularly on the 'business side' • Adoption is significantly because 'Big Data' technology is driven by inductive reasoning whereas

	typical IT governance in HE prioritizes risk mitigation through deductive reasoning • Smaller scale 'Big Data' projects, particularly on the 'business side' of HE have been implemented largely on a pilot basis in order to reduce risk • In future, implementation risks may be significantly reduced if educational technology vendors development and offer 'Big Data' solutions sector-wide
Question #2: To what extent are Alberta higher education leaders motivated by competitive enrolment and/or cost-effective learning considerations that could be developed by 'Big Data' applications?	• Leaders typically view competition with other HE institutions is either a non-factor or a low-level motivator • Some leaders acknowledges competitive factors in selected programs in some geographic areas but leaders demonstrate collegial more than competitive instincts • The upcoming IMAs may create a more pseudo-competitive environment in the future but there is no evidence yet • Risk mitigation remains the major motivating factor in IT governance and developing competitive advantages is not expressed as a key priority
Question #3: Do Alberta higher education leaders have faith that meaningful business decisions	• Leaders' experience with more advanced predictive analytics is limited to date but there is evidence that

can or should be influenced or even made by ‘Big Data’ applications?	‘Bid Data’ influenced resource allocation in non-academic areas • The scale and scope of ‘Big Data’ adoption levels remains too low to fully assess leader’s faith in the technology
Question #4: What ethical considerations do Alberta higher education leaders integrate in the data collection and analyses processes of ‘Big Data’ initiatives?	• Leaders demonstrate an appropriate understanding of the ethical requirements for the preliminary stages of ‘Big Data’ adoption • As ‘Big Data’ incorporates more predicative and prescriptive analytics, leaders need to fortify their ethical understandings ensure fairness and minimize bias towards learners and other stakeholders
Question #5: How able do Alberta higher education leaders consider their respective institutions are to implement ‘Big Data’ successfully from technological and change management perspectives?	• Accessing appropriate IT expertise is not a major inhibiting factor for ‘Big Data’ adoption • Leaders are prepared to invest in ‘Big Data’ as long as the risk can be appropriately mitigated through existing IT governance structures but most large scale projects are not able to meet the requirements

As discussed in *Question #3: Do Alberta higher education leaders have faith that meaningful business decisions can or should be influenced by 'Big Data' applications?,* a likely solution to the governance and risk management challenges is for educational technology vendors to increase scalability across the HE sector. Integrating 'Big Data' technology to SIS, LMS and CRM data products sector-wide would provide HE institutions with access to 'Big Data' technology at an incremental and predictable cost. However, in the near future, there are at least two substantial challenges inherent to the scaling solution. First, as discussed above in *Question #4: What ethical considerations do Alberta higher education leaders integrate in the data collection and analyses processes of 'Big Data' initiatives?,* outsourcing of 'Big Data' capacities to educational technology vendors could lead to the establishment of 'black box' management processes and decisions made on the sole basis of vendor-provided technology and algorithms. Therefore, although educational technology vendors have an important role in resolving 'Big Data' adoption issues, HE leaders still need a strong understanding of how the technology can be implemented ethically and effectively. Second, as 'Big Data' technology evolves, HE leaders will be increasingly challenged by inductive decision-making processes. As data and data analyses capacity becomes increasingly available and automated, management decisions will be made more frequently on the basis on data correlation and less upon HE leaders' knowledge or experience. Most HE leaders have yet to practice inductive decision-making in a meaningful way but, as more sophisticated predictive and prescriptive analytics are adopted, HE management will be challenged to adjust. HE leaders in the study express a willingness to adopt a more systems thinking approach. However, risk mitigation methods to date require the application of deductive thinking to IT governance processes. As scalable 'Big Data' technology solutions become more available, there will be more opportunities to test if HE leaders can translate their willingness to test systems thinking on truly data-driven management approaches at HE institutions.

In summary, HE institutions are not currently well prepared for large scale ‘Big Data’ adoption. To date, IT governance processes are rooted in deductive thinking and prioritize risk mitigation over large scale ‘Big Data’ technology adoption. However, new approaches from educational technology vendors may drastically increase adoption by integrating ‘Big Data’ technology with current SIS, LMS and CRM products. Increasing adoption through education technology vendors will create some unfamiliar challenges to HE leaders. A greater utilization of inductive systems thinking and an enhanced ethical framework to manage decision-making in an increasingly automated environment are two capacities that will grow in importance.

5.2 Conclusion

The purpose of the study is to develop a conceptual framework that describes and explains the adoption of ‘Big Data’ technology in the Alberta HE sector. Using grounded theory, an analysis of semi-structured interviews of seven HE leaders in Alberta reveals insights that distinguish the adoption of ‘Big Data’ from other technology types. Among HE leaders, there is strong support for ‘Big Data’ technology and a recognition that the sector has only ‘scratched the surface’ of its potential impact. Adoption of ‘Big Data’ in the HE sector is slow relative to other sectors because risk mitigation inherent to IT governance processes utilizes deductive thinking and requires proposals to address key criteria that include the following: i) measurability, ii) performance gap, iii) viable solution and iv) ethics and legal compliance. Conversely, ‘Big Data’ technology projects are focused upon uncovering the relationships between data sets in the overall system. Problems, much less solutions, cannot be identified until the project starts and the data is collected and analyzed. 'Big Data' projects in the study are all small-scale with limited objectives. As a result, 'Big Data' projects in the study were able to either meet or avoid IT governance approval requirements because they are: i) too small to be reviewed, ii) being implemented with existing resources and/or iii) focused upon a specific and predetermined objective. Despite the modest rate of adoption, the existence of even small scale 'Big Data' projects clearly demonstrates adaptive organizational learning within HE institutions. HE leaders recognize the future value of 'Big Data' technology and use workarounds for implementation despite barriers inherent to IT governance processes. Conversely, despite the recent development of IMAs with the provincial government, the study did not find evidence that HE leaders consider their institutions in competition. The absence of competition may explain a lack on generative organizational learning in HE compared to other sectors. HE leaders are not incentivized to assume the risks inherent to inductive reasoning and systems thinking in IT governance processes that are necessary to facilitate more rapid adoption of 'Big Data'.

Despite the relatively slow pace of 'Big Data' adoption in the HE sector, the study reveals a viable pathway forward. Unlike individual HE institutions, education technology vendors have access to a much larger scale of potential technology users. Far greater scalability combined with ownership of current SIS, LMS and CRM products makes it far more likely that vendors will develop and integrate 'Big Data' with current and widely used software. This scenario addresses the IT governance barrier as educational technology vendors assume most of the 'Big Data' development investment and risks. However, as 'Big Data' adoption becomes more rapid and comprehensive, the study demonstrates that HE leaders will be quickly presented with two new priorities. First, HE leaders will need to significantly augment current ethical frameworks to accommodate more advanced ‘Big Data’ technology functions such as predictive and prescriptive analytics. Specifically, by outsourcing ‘Big Data’ capacity HE leaders will play less direct roles in developing and implementing the technology. The conditions will be ripe for a management ‘black box’ – where key decisions, such as resource allocation or learner supports, are automatically made without HE leaders fully understanding how or why. As a result, HE leaders need to increase their capacity to meet the impending and increasingly complex ethical responsibilities for data gathering, storing and usage. Second, the expansion of ‘Big Data’ technology will further test HE leaders’ commitment to ‘data-driven’ management approaches. Specifically, ‘Big Data’ technology will, based upon data correlation, automatically prescribe management approaches that will not necessarily agree with causal or knowledge-based management methods. Due to the relatively low adoption levels currently in the sector, the underlying epistemological challenge has not yet materialized for most HE leaders. However, increasing adoption of ‘Big Data’ will dramatically accelerate truly data-driven management approaches and HE leaders need to be better philosophically prepared for the greater automation of decision-making.

In summary, ‘Big Data’ technology offers both promises and pitfalls for HE leaders. The study found that HE leaders in Alberta, despite a relatively rudimentary level of adoption currently, believe ‘Big Data’ technology will heavily influence the sector in the future. As demonstrated by the proposed conceptual framework, the deductive nature and focus upon risk mitigation of current IT governance structures significantly inhibit ‘Big Data’ adoption in the HE sector. However, there are potential lower risk pathways now forming that may soon drastically increase the adoption of 'Big Data'. In particular, education technology vendors are well placed to mitigate risk for institutions by integrating 'Big Data' technology in existing software and scaling their solutions across the sector. In this scenario, the scale and complexity of 'Big Data' utilization will increase drastically. HE leaders need to be better prepared for the newly emerging ethical and philosophical issues that greater adoption of 'Big Data' technology will quickly bring to the forefront the HE institutional management.

6.0 References

Anderson, C. (2008). The End Of Theory. In (Vol. 16, pp. 108): Conde Nast Publications, Inc.

Atkinson, P., Silverman, D., & Delamont, S. (2003). *Key themes in qualitative research: Continuities and changes*. Rowan and Littlefield.

Axelrod, P. (2002). *Values in conflict: the university, the marketplace, and the trials of liberal education*. McGill-Queen's University Press.

Baer, L., & Campbell, J. (2011). Game Changers. *EDUCAUSE Review*.

Baer, L., Duin, A. H., & Bushway, D. (2015). Change agent leadership. *Planning for Higher Education*, *43*(3), 1.

Bagley, S. S., & Portnoi, L. M. (2014). Setting the Stage: Global Competition in Higher Education. *New Directions for Higher Education*, *2014*(168), 5-11. https://doi.org/10.1002/he.20109

Barber, M., Donnelly, K., & Rizvi, S. (2013). *An Avalance is Coming: Higher Education and the Revolution Ahead*. http://www.studynet2.herts.ac.uk/intranet/lti.nsf/0/684431DD8106AF1680257B560052BCCC/$FILE/avalanche-is-coming_Mar2013_10432.pdf

Basu, A. (2013). Five pillars of prescritive analytics success. *Analytics*, (March/April Issue), 8-12.

Bayne, S. (2015). Teacherbot: interventions in automated teaching. *Teaching in Higher Education*, *20*(4), 455-467. https://doi.org/10.1080/13562517.2015.1020783

Beer, D. (2017). The social power of algorithms. *Information, Communication & Society*, *20*(1), 1-13. https://doi.org/10.1080/1369118X.2016.1216147

Berry, D. M. (2014). *Critical theory and the digital*. Bloomsbury Academic.

Bhattacharjya, J., & Chang, V. (2006, 2006). Evolving IT governance practices for IT and business alignment: A case study in an Australian institution. Facing the Information Society, Retrieved from https://espace.curtin.edu.au/bitstream/handle/20.500.11937/42026/20128_downloaded_stream_116.pdf?sequence=2&isAllowed=y

Bichsel, J., & Feehan, P. (2014). *Getting Your Ducks in a Row: IT Governance, Risk and Compliance Programs in Higher Education*. EDUCAUSE. Retrieved December 20 from https://library.educause.edu/resources/2014/3/getting-your-ducks-in-a-row-it-governance-risk-and-compliance-programs

Biesta, G. (2007). WHY "WHAT WORKS" WON'T WORK: EVIDENCE-BASED PRACTICE AND THE DEMOCRATIC DEFICIT IN EDUCATIONAL RESEARCH. *Educational Theory*, *57*(1), 1-22. https://doi.org/10.1111/j.1741-5446.2006.00241.x

Bleiklie, I. (1998). Justifying the Evaluative State: New Public Management ideals in higher education. *European Journal of Education*, *33*(3), 316;299;-316.

Bloomberg, L. D., & Volpe, M. (2016). *Completing Your Qualitative Dissertation* (3rd ed.). SAGE.

Blumer, H. (1969). *Symbolic Interactionism: Perspective and method*. Prentice Hall.

Booth, M. (2012). Learning analytics: The new black [http://www.educause.edu/ero/article/learning-analytics-new-black]. *EDUCAUSE Review*.

Boyd, D., & Crawford, K. (2011, September, 2011). *Six Provocations for Big Data* A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1926431

Boyd, D., & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, *15*(5), 662-679. https://doi.org/10.1080/1369118x.2012.678878

Bryant, A. (2007). Introduction: Grounded Theory Research: Methods and Practices. In A. Bryant & K. Charmaz (Eds.), *The Sage handbook of grounded theory* (pp. 1-28). Sage Publications. https://doi.org/10.4135/9781848607941

Bryant, A., & Charmaz, K. (2007). *The SAGE Handbook of Grounded Theory* (1 ed.). SAGE Publications, Limited.

Burke, J., & Minassians, H. (2001). *Linking State Resources to Campus Results: From Fad to Trend - The Fifth Annual Survey*. S. U. o. N. York.

Calhoun, J., & Kamerschen, D. R. (2010). The impact of governing structure on the pricing behavior and market structure of public institutions of higher education in the U.S. *International Review of Economics*, *57*(3), 317-333. https://doi.org/10.1007/s12232-010-0089-2

Charmaz, K. (1990). 'Discovering' chronic illness: Using grounded theory. *Social Science & Medicine*, *30*(11), 1161-1172. https://doi.org/10.1016/0277-9536(90)90256-R

Charmaz, K. (2000). Grounded Theory: Objectivist and constructivist methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 509-535). SAGE.

Charmaz, K. (2014). *Constructing grounded theory: A practical guide through qualitative analysis* (2nd ed.). SAGE.

Charmaz, K., Thornburg, R., & Keane, E. (2018). Evolving Grounded Theory and Social Justice Inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE Handbook of Qualitative Research* (5th ed., pp. 411-443). SAGE.

Clark, E. (2018). Digital Transformation: What is it? , *2018*(July 30, 2018). https://er.educause.edu/articles/2018/5/digital-transformation-what-is-it

Clarke, A. E. (2005). *Situational analysis: grounded theory after the post-modern turn*. Sage Publications.

Clarke, A. E. (2007). Grounded theory: Conflicts, debates and situational analysis. In W. Outhwaite & S. P. Turner (Eds.), *Handbook of social science methodology* (pp. 838-885). Sage.

Cohen, L., Manion, L., & Morrison, K. (2011). The ethics fo educational and social research. In L. Cohen & K. Morison (Eds.), *Research methods in education* (pp. 75-104). Routledge.

Cope, B., & Kalantzis, M. (2016). Big Data Comes to School:Implications for Learning, Assessment, and Research. *AERA Open*, *2*(2), 2332858416641907. https://doi.org/10.1177/2332858416641907

Crawford, K. (2013). The hidden biases in Big Data. *Harvard Business Review*.

Creswell, J. W., & Gutterman, T. C. (2019). *Educational Research: Planning, Conducting, and Evaluating Quantiative and Qualitative ResearchNe* (6th ed.). Pearson.

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

David, P. A., & Foray, D. (2002). An introduction to the economy of the knowledge society. *International Social Science Journal*, *54*(1), 9.

De Haes, S., & Van Grembergen, W. (2008). An exploratory study into the design of an IT governance minimum baseline through Delphi research. *Communications of the Association for Information Systems*, *1*(22), 24.

Dede, C., Ho, A., & Mitros, P. (2016). Big Data Analysis in Higher Education: Promises and Pitfalls. *EDUCAUSE Review*, *51*(5), 1.

Deem, R., & Brehony, K. J. (2005). Management as ideology: the case of 'new managerialism' in higher education. *Oxford Review of Education*, *31*(2), 217-235. https://doi.org/10.1080/03054980500117827

Denzin, N. K. (2007). Grounded theory and the politics of interpretation. In A. Bryant & K. Charmaz (Eds.), *The Sage Handbook of grounded theory* (pp. 454-471). SAGE.

Denzin, N. K., & Lincoln, Y. S. (2018). *The SAGE handbook of qualitative research* (Fifth ed.). Sage.

Dey, I. (1999). *Grounding grounded theory*. Academic Press.

Dowd, A. (2005). *Data Don't Drive: Building a Practioner-Driven Culture of Inquiry to Assess Community College Performance*.

Doyle, W. R., & Delaney, J. A. (2009). Higher Education Funding: The New Normal. *Change: The Magazine of Higher Learning*, *41*(4), 60-62. https://doi.org/10.3200/CHNG.41.4.60-62

Edwards, R., & Fenwick, T. (2016). Digital analytics in professional work and learning. *Studies in Continuing Education*, *38*(2), 213-227. https://doi.org/10.1080/0158037X.2015.1074894

Ellis, C. (1995). EMOTIONAL AND ETHICAL QUAGMIRES IN RETURNING TO THE FIELD. *Journal of Contemporary Ethnography*, *24*(1), 68-98. https://doi.org/10.1177/089124195024001003

Farhan, B. Y. (2016). Competitive Behaviour in Publicly Funded Academic Institutions. *Interchange*, *47*(4), 357-373. https://doi.org/10.1007/s10780-016-9283-9

Fattah, A., & Setyadi, R. (2021). Determinants Effectiveness Information Governance in Higher Education Institutes (HEI) using partial least squares structural equation modeling approach (PLS0SEM). *Journal of Physics: Conference Serices*, *1807*(1), 012007.

Finn, M. (2016). Atmospheres of progress in a data-based school. *Cultural Geographies*, *23*(1), 29-49. https://doi.org/10.1177/1474474015575473

Frické, M. (2015). Big data and its epistemology. *Journal of the Association for Information Science and Technology*, *66*(4), 651-661. https://doi.org/10.1002/asi.23212

Giacumo, L. A., & Bremen, J. (2016). EMERGING EVIDENCE ON THE USE OF BIG DATA AND ANALYTICS IN WORKPLACE LEARNING: A Systematic Literature Review. *Quarterly Review of Distance Education, 17*(4), 21.

Gillespie, T., Boczkowski, P. J., & Foot, K. A. (2014). *Media technologies: essays on communication, materiality, and society*. The MIT Press. https://doi.org/10.7551/mitpress/9780262525374.001.0001

Glaser, B. (1978). *Theoretical sensitivity: advances in the methodology of grounded theory*. Sociology Press.

Glaser, B. (2001). *The grounded theory perspective: conceptualization contrasted with description*. Sociology Press.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine. Doubleday.

Government of Alberta. (2020). *Implementing Budget 2020*. Retrieved August 20, 2020 from https://www.alberta.ca/implementing-budget-2020.aspx

Government of Alberta. (2017). *Post-Secondary Learning Act*. Edmonton: Alberta Queen's Printer Retrieved from http://www.qp.alberta.ca/documents/Acts/p19p5.pdf

Government of Alberta. (2018). *Post-Secondary Learning Act: Public Post-Secondary Institutions' Tuition Fees Regulation*. Edmonton: Alberta Queen's Printers

Gregory, K., McMillan, C., & Daniel, J. (2017). Introduction. In K. Gregory & C. McMillan (Eds.), *Digital Sociologies* (pp. pp. xvii-xxx). Policy Press.

Heifetz, R. A., Grashow, A., & Linsky, M. (2009). *The practice of adaptive leadership: Tools and tactics for changing your organization and the world*. Harvard Business Press.

Hilbert, M., & López, P. (2011). The World's Technological Capacity to Store, Communicate, and Compute Information. *Science*, *332*(6025), 60-65. https://doi.org/10.1126/science.1200970

Hodson, J. B. (2010). Leading the way: The role of presidents and academic deans in fundraising. *New Directions for Higher Education*, *2010*(149), 39-49. https://doi.org/10.1002/he.379

Ifenthaler, D., & Schumacher, C. (2016). Student perceptions of privacy principles for learning analytics. *Educational Technology Research and Development*, *64*(5), 923-938. https://doi.org/10.1007/s11423-016-9477-y

Isaías Scalabrin, B., Rui Dinis, S., & Ruben, P. (2021). Information Technology Governance for Higher Education Institutions: A Multi-Country Study. *Informatics (Basel)*, *8*(26), 26. https://doi.org/10.3390/informatics8020026

Isaías Scalabrin, B., Rui Dinis, S., Ruben, P., & Irineu Manoel de, S. (2020). Effective IT governance Mechanisms in Higher Education Institutions: An empirical study. *RISTI : Revista Ibérica de Sistemas e Tecnologias de Informação*(E25), 412-423.

Jairak, K., Praneetpolgrang, P., & Subsermsri, P. (2015). Information technology governance practices based on sufficiency economy philopsopy in the Thai university sector. *Inf. Technol. People*(28), 195-223.

Jameson, J. (2013). e-Leadership in higher education: The fifth "age" of educational technology research. *British Journal of Educational Technology*, *44*(6), 889-915. https://doi.org/10.1111/bjet.12103

Jindal, R., & Borah, M. D. (2015). Predictive Analytics in a Higher Education Context. *IT Professional*, *17*(4), 24-33. https://doi.org/10.1109/MITP.2015.68

Jokonya, O. (2015, 2015). Towards a Conceptual Framework for Big Data Adoption in Organizations.

Jones, S. J. (2012). Technology Review: the possibilities of learning analytics to improve learner-centered decision-making. *The Community College Enterprise*, *18*(1), 89.

Keller, G. (1983). *Academic strategy: the management revolution in American higher education*. Johns Hopkins University Press.

Kirby, D. (2011). Strategies for widening access in a quasi-market higher education environment: recent developments in Canada. *Higher Education*, *62*(3), 267-278. https://doi.org/10.1007/s10734-010-9386-7

Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*(1), 205395171452848. https://doi.org/10.1177/2053951714528481

Kitchin, R., & Dodge, M. (2011). *Code/Space: software and everyday life*. MIT Press.

Knox, J. (2016). Posthumanism and the MOOC: opening the subject of digital education. *Studies in Philosophy and Education*, *35*(3), 305-320. https://doi.org/10.1007/s11217-016-9516-5

Khouja, M., Rodriguez, I. B., Ben Halima, Y., & Moalla, S. (2018). IT Governance in Higher Education Institutions. *Int. J. Hum. Cap. Inf. Technol. Prof.*(9), 52-67.

Kruse, A., & Pongsajapan, R. (2012). Student-centred learning analytics. *CNALS Thought Papers*. Retrieved November 10, 2018, from https://cndls.georgetown.edu/m/documents/thoughtpaper-krusepongsajapan.pdf

Lafley, A. G., Martin, R. L., Rivkin, J. W., & Siggelkow, N. (2012). Bringing science to the art of strategy. In (Vol. 90, pp. 56). United States: Harvard Business Review.

Lane, J. (2016). BIG DATA: THE ROLE OF EDUCATION AND TRAINING. *Journal of Policy Analysis and Management*, *35*(3), 722-724. https://doi.org/10.1002/pam.21922

Laux, C., Li, Na L., Seliger, C., & Springer, J. (2017). Impating Big Data analytics in higher education through Six Sigma techniques. *International Journal of Productivity and Performance Management*, *66*(5), 662-679.

Lincoln, Y. S., & Guba, E. G. (2000). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 163-188). Sage.

Lincoln, Y. S., & Guba, E. G. (2013). *The constructivist credo*. Left Coast Press, Inc. https://doi.org/10.4324/9781315418810

Lincoln, Y. S., Lynham, A., & Guba, E. G. (2017). Paradigmatic controversies, contradictions, and emerging confluences, revisited. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE Handbook of Qualitative Research* (5th ed., pp. 108 - 150). SAGE.

Livingstone, D. W., & Guile, D. (2012). *The knowledge economy and lifelong learning: a critical reader / edited by D.W. Livingstone and David Guile* (Vol. 4). Sense Publishers.

Long, P., & Siemens, G. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*(September/October), 31-41.

Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & Education*, *54*(2), 588-599. https://doi.org/10.1016/j.compedu.2009.09.008

Macfadyen, L. P., & Dawson, S. (2012). Numbers Are Not Enough. Why e-Learning Analytics Failed to Inform an Institutional Strategic Plan. *Journal of Educational Technology & Society*, *15*(3), 149-163.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., . . . McKinsey Global, I. (2011). *Big Data: The next frontier for innovation, competition, and Productivity*. McKinsey Global Institute.

Marope, P. T. M., Wells, P. J., & Hazelkorn, E. (Eds.). (2013). *Rankings and accountability in higher education; Uses and abuses*. Unesco.

Mayer-Schönberger, V. (2009). *Delete: the virtue of forgetting in the digital age*. Princeton University Press.

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

Mazzocchi, F. (2015). Could Big Data be the end of theory in science? *EMBO reports*, *16*(10), 1250-1255. https://doi.org/10.15252/embr.201541001

McLaughlin, G., Howard, R., Balkan, L., & Blythe, E. (1998). *People, Processes, and Managing Data* (Resources in Institutional Research, Issue.

Weber, K., Otto, B., & Österle, H. (2009). One Size Does Not Fit All---A Contingency Approach to Data Governance. *ACM journal of data and information quality*, *1*(1), 1-27. https://doi.org/10.1145/1515693.1515696

Menon, E. (2014). Introduction. In E. Menon, D. Terkla, & P. Gibbs (Eds.), *Using Data to Improve Higher Education: Research, Policy and Practice* (pp. 1-5). Sense Publishers.

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A methods sourcebook*. SAGE.

Moreira, F., Ferreira, M. J., & Cardoso, A. (2017). Higher Education Disruption Through IoT and Big Data: A Conceptual Approach. In *Learning and Collaboration Technologies.*

Novel Learning Ecosystems (pp. 389-405). https://doi.org/10.1007/978-3-319-58509-3_31

Mrig, A., & Sanaghan, P. (2018). *The future of higher education: will higher education seize the future or fall victim to it?*

Neave, G. (1988). On the cultivation of quality, efficiency and enterprise: An overview of recent trends in higher education in Western Europe, 1986- 1988. *European Journal of Education*, *23*(1-2), 2<7-23;27;-23. https://doi.org/10.2307/1502961

Norris, D., Brodnick, R., Lefrere, P., Gilmour, J., & Baer, L. (2013). Transforming in an age of disruptive change: Part 1: Back to the future, zooming to the present. *Planning for Higher Education*, *41*(2), 18.

OECD. (2013). *OECD Report: The State of Higher Education 2013*. O. H. E. P. Publications. www.oecd.org/education/imhe/thestateofhighereducation2013.htm

Olssen, M. (2016). Neoliberal competition in higher education today: research, accountability and impact. *British Journal of Sociology of Education*, *37*(1), 129-148. https://doi.org/10.1080/01425692.2015.1100530

Olszak, C., & Mach-Król, M. (2018). A Conceptual Framework for Assessing an Organization's Readiness to Adopt Big Data. *Sustainability*, *10*(10), 3734. https://doi.org/10.3390/su10103734

Orlikowski, W. J., & Iacono, C. S. (2001). Research Commentary: Desperately Seeking the "IT" in IT Research--A Call to Theorizing the IT Artifact. *Information Systems Research*, *12*(2), 121. https://doi.org/10.1287/isre.12.2.121.9700

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

Patton, J. W. (2000). Protecting privacy in public? Surveillance technologies and the value of public places. *Ethics and Information Technology*, *2*(3), 181-187. https://doi.org/10.1023/A:1010057606781

Peters, M. A. (2013). Managerialism and the neoliberal university: prospects for new forms of "open management" in higher education. *Contemporary Readings in Law and Social Justice*, *5*(1), 11.

Picciano, A. G. (2012). The evolution of big data and learning analytics in American higher education. *Journal of Asynchronous Learning Networks*, *16*(3), 9.

Piché, P. G., & Jones, G. A. (2016). Institutional Diversity in Ontario's University Sector: A Policy Debate Analysis. *The Canadian Journal of Higher Education*, *46*(3), 1.

Prinsloo, P. (2009). *Ethics and learning analytics as a Faustian pact: Between Orwell, Huxley, Kafka and the deep blue sea* LASI13 conference,, Pretoria, South Africa. http://www.up.ac.za/telematic/sahela2013/day2-9h30-sahela2013 paul_UNISA_ethics_and_learning_analytics_as_a_fuastian_pact.pdf

Prinsloo, P., & Slade, S. (2014). Student Privacy and Institutional Accountability in an Age of Surveillance. In E. Menon, D. Terkla, & P. Gibbs (Eds.), *Using Data to Improve Higher Education: Research, Policy and Practice* (pp. 187-214). Sense Publishers.

Psacharopoulos, G. (2014). The Return on Investment in Higher Education: Methods, Data and Policy Implications. In E. Menon, D. Terkla, & P. Gibbs (Eds.), *Using Data to Improve Higher Education: Research, Policy and Practice* (pp. 121-149). Sense Publishing.

Reid, W. M., & Dold, C. J. (2018). Burns, Senge, and the Study of Leadership. *Open Journal of Leadership*, *Vol.07No.01*, 28, Article 83458. https://doi.org/10.4236/ojl.2018.71006

Richardson, L. (1993). Interupting discursive spaces: Consequenses for the sociologial self. In N. K. Denzin (Ed.), *Studies in Symbolic Interaction* (Vol. 14, pp. 77-84). JAI Press.

Ritacco, A., & McGowan, H. (2017, December 5, 2018). Preparing students to lose their jobs (and faculty to keep theirs). https://www.academicimpressions.com/blog/preparing-students-to-lose-their-jobs-and-faculty-to-keep-theirs/

Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *3*(1), 12-27. https://doi.org/10.1002/widm.1075

Scholes, V. (2016). The ethics of using learning analytics to categorize students on risk. *Educational Technology Research and Development*, *64*(5), 939-955. https://doi.org/10.1007/s11423-016-9458-1

Senge, P. M. (1997). THE FIFTH DISCIPLINE. *Measuring Business Excellence*, *1*(3), 46-51. https://doi.org/10.1108/eb025496

Senge, P. M. (2006). *The fifth discipline: the art and practice of the learning organization* (Rev. and updat ed.). Crown Business, Currency.

Shalf, J. M., & Leland, R. (2015). Computing beyond Moore's Law. *Computer*, *48*(12), 14-23. https://doi.org/10.1109/MC.2015.374

Silver, N. (2012). *The signal and the noise: why most predictions fail-- but some don't*. Penguin Press.

Simon, H. A. (1982). *Models of bounded rationality*. MIT Press.

Skolnik, M. L. (2016). How do quality assurance systems accommodate the differences between academic and applied higher education? [journal article]. *Higher Education*, *71*(3), 361-378. https://doi.org/10.1007/s10734-015-9908-4

Slade, S., & Prinsloo, P. (2013). Learning Analytics: Ethical Issues and Dilemmas. *American Behavioral Scientist*, *57*(10), 1510-1529. https://doi.org/10.1177/0002764213479366

Solove, D. J. (2004). *The digital person: technology and privacy in the information age*. New York University Press.

Stephenson, A. L., Heckert, A., & Yerger, D. B. (2016). College choice and the university brand: exploring the consumer decision framework. *Higher Education*, *71*(4), 489-503. https://doi.org/10.1007/s10734-015-9919-1

Stoddart, E. (2012). A surveillance of care: Evaluating surveillance ethically. In K. Ball, K. D. Haggerty, & D. Lyons (Eds.), *Routledge handbook of surveillance studies*. Routledge.

Strauss, A. L., & Corbin, J. M. (1990). *Basics of qualitative research: grounded theory procedures and techniques*. Sage.

Taylor, J. (2014). Informing or Distracting? Guiding or Driving? The Use of Performance Indicators in Higher Education. In E. Menon, D. Terkla, & P. Gibbs (Eds.), *Using Data to Improve Higher Education: Research, Policy and Practice* (pp. 7-25). Sense Publishing.

Tweed, R., & Charmaz, K. (2011). Grounded theory for mental health practitioners. In A. Thompson & D. Harper (Eds.), *Qualitative research methods in mental health and psyco-theratpy: A guide for students and practioners* (pp. 131-146). Wiley-Blackwell.

Usher, A. (2018). *The State of Post-Secondary Education in Canada*.

Voorhees, R., & Cooper, J. (2014). Opportunities and Barriers to Effective Planning in Higher Education: Data Sources and Techniques. In E. Menon, D. Terkla, & P. Gibbs (Eds.), *Using Data to Improve Higher Education: Research, Policy and Practice* (pp. 25-39). Sense Publishers.

Vrontis, D., Thrassou, A., & Melanthiou, Y. (2007). A contemporary higher education student-choice model for developed countries. *Journal of Business Research*, *60*(9), 979-989. https://doi.org/10.1016/j.jbusres.2007.01.023

Wagner, E., & Ice, P. (2012). Data changes everything: Delivering on the promise of learning analytics in higher education. *EDUCAUSE Review*(July/August), 32-42.

Williamson, B. (2017). *Big data in education: the digital future of learning, policy and practice*. SAGE Publications.

Wintrup, J. (2017). Higher Education's Panopticon? Learning Analytics, Ethics and Student Engagement. *Higher Education Policy*, *30*(1), 87. https://doi.org/10.1057/s41307-016-0030-8

Wook, M., Yusof, Z. M., & Mohd Zakree Ahmad, N. (2017). Educational data mining acceptance among undergraduate students. *Education and Information Technologies*, *22*(3), 1195. https://doi.org/10.1007/s10639-016-9485-x

Wu, S., Straub, D., & Liang, T. P. (2015). How Information Technology Governance Mechanisms and Strategic Alignment Influence Organizational Performance: Insights from a Matched Survey of Business and IT Managers. *MIS Quarterly*, *39*(2), 497-518.

Yeung, K. (2017). 'Hypernudge': Big Data as a mode of regulation by design. *Information, Communication & Society*, *20*(1), 118-136. https://doi.org/10.1080/1369118X.2016.1186713

Ziewitz, M. (2016). Governing Algorithms: Myth, Mess, and Methods. *Science, Technology, & Human Values*, *41*(1), 3-16. https://doi.org/10.1177/0162243915608948

Ziskin, M. B., Hossler, D., Raboum, K., Cekic, O., & Hwang, Y. (2014). *Outcomes-Based Funding: Current Status, Promises Practices and Emerging Trends*.

7.0 Appendices:

7.1 Primary Data Codes

#	Name	1	2	3	4	5	6	7	SUM
1	Admissions	4				2			6
2	Enrollment models	2		2		1	1	2	8
3	Predictive planning	7	3	5	2	5		3	25
4	Sophistication	2	3		1	5	3	3	17
5	HR Capacity	4	2	2	1	1	4		14
6	Attrition	1		2					3
7	Retention	2	2				1	1	6
8	Student Success	3	1	5	2	1	3		15
9	CRM	4							4
10	SIS	2	1			1		1	5
11	External information	1	1			1		2	5
12	Measure success	9	5	2	5	3	4	2	30
13	Access new markets	2			2	1		1	6
14	Connections	1	2			1	1		5
15	Recruitment strategies	2				2		1	5
16	Patterns / Correlations	6	3	6	4	4	4	3	30
17	Advantage	1							1
18	Recruitment	5	1			1		2	9
19	Better/good indication	4		1	1				6
20	Efficiency	4	1	3	3	1	1	1	14
21	Yield / conversion	6	1					1	8
22	Relationships	1	1			1			3
23	Invest	2			3	1	2	1	9
24	Non-competitive	3	1	1		2		1	8
25	Inaccurate prediction	1				1			2

26	Program review	2				1		1	4
27	New lens, perspective	1	2			3	1	2	9
28	Strategic goals	2	2		1	1	3	2	11
29	Investment agreements - too early	1		1	1	2			5
30	Tech adoption	6	2	1	4	5	5	4	27
31	Institutional priorities	1			1		3	2	7
32	transparency	1				1	4		6
33	Data quality	1		1	1		1	1	5
34	evidence for resource allocation	2	3	5	4	5	1	3	23
35	User focus	2				1	1	1	5
36	Organizational learning	1	3	2	4	9	9	6	34
37	Continuous improvement	2	1	2	1	6	3		15
38	Value for $$	1	4	1	6	1		2	15
39	Client satisfaction	1	2		1	1	1		6
40	Tracking touch points	2	1	1					4
41	Lack of initial investment resources	1			1	1	3	1	7
42	Maintenance fees	1			1				2
43	Change management	1	1			2	1		5
44	Training	1					1	1	3
45	Unstructured data		2						2
46	Data collection		3	1		3	4	3	14
47	Data analysis		4	3	1	4	6	8	26
48	ERP		1						1
49	Institutional strategies		1			1	1		3
50	LMS		1	4		2	6		13
51	Learner Engagement Survey		1					2	3
52	Improving student experience		2		1		2		5
53	Prescriptive results		2	4	2	3	2		13
54	AI/ML		1		1	1		1	4
55	Insight		4	4	3	5	5	4	25
56	Unsupervised classification model		1						1
57	Unexpected results		1						1

58	Project proposals- deductive vs inductive rationale	4	3	1				8
59	Hazy understanding	1		1	1		2	5
60	Profit motive (lack of)	2		1				3
61	Some form of competition	1		2	4		1	8
62	Competition in research	1						1
63	Fundraising	1						1
64	Relationship based industry	1						1
65	Art vs science	1		1	1		1	4
66	Innovation	1		1	3	2	3	10
67	Exam Scheduling	1						1
68	Ethical Use	1	2		1	2	2	8
69	HR turnover	1						1
70	Slower adoption	1		1	3		1	6
71	Learner behaviors		1		1	4	1	7
72	Mixed results		1				1	2
73	Descriptive results		1				2	3
74	Predictive results		1	4	1	2	1	9
75	Lack of prescriptive results		6			1	1	8
76	Unrealized potential		4	2	2	2	3	13
77	Measuring specific LMS content impact on learner success		2		1	2		5
78	Measuring instructional strategies impact and learner success		2			2		4
79	Proxy measures		1	1	1	1	1	5
80	Lack of unexpected results		1	1				2
81	Confirmed inferences		1	1	1			3
82	consumer vs learner behavior		2		1			3
83	Perceive competitiveness at senior levels		1	1	1			3
84	Provincial focus on what is measured and funded		1		1			2
85	Administrative improvements but not enhanced learning		1					1
86	Sceptical of tech/autonomous learning potential		1					1
87	More data for 'big data' to work more effectively			1			1	2
88	Driven by executive			1	1			2
89	Machine learning			2			1	3

90	Algorithms	1	1	3		5
91	Bias in models	1			1	2
92	Need more data to increase confidence levels	1			1	2
93	Data security / ethics	1		2	2	5
94	Open source for research inexpensive but commercial tools costly	1				1
95	Pilot project initial investment	2		1		3
96	Risk adverse / not early adopters		2		1	3
97	IT integrated to operations - not just a service provider		1		1	2
98	Need more sector wide IT partnerships		1			1
99	Need for causal understanding		1		1	2
100	Leadership in a complex environment		2			2
101	Impact of technology on organizational culture		1	3	6	10
102	Informal skill development for unknown applications			2		2
103	IT Project governance models					0
						723

7.2 Secondary Data Codes

1	**Measurable Results**	
	1	Admissions
	2	Enrollment models
	6	Attrition
	7	Retention
	8	Student Success
	12	Measure success
	18	Recruitment
	21	Yield / conversion
	41	Tracking touch points
	65	Fundraising
	71	HR turnover
	86	Provincial focus on what is measured and funded
2	**Recognized Performance Gaps**	
	1.1	**Collegial / Professional**
	13	Access new markets
	14	Connections
	15	Recruitment strategies
	20	Efficiency
	25	Non-competitive
	27	Program review
	29	Strategic goals
	40	Client satisfaction
	53	Learner Engagement Survey
	54	Improving student experience
	57	Insight
	68	Innovation
	69	Exam Scheduling
	73	Learner behaviors
	78	Unrealized potential
	90	Driven by executive
	102	Leadership in a complex environment
	1.2	**Competition**
	17	Advantage
	62	Profit motive (lack of)
	63	Some form of competition
	64	Competition in research
	84	consumer vs learner behavior
	85	Perceive competitiveness at senior levels

1.3 Performance Framework

28 New lens, perspective
30 Investment agreements - too early
32 Institutional priorites
38 Continuous improvement
51 Institutional strategies

3 Viable Potential Solutions

3.1 Financially Viable

3 Predictive planning
23 Invest
35 evidence for resource allocation
39 Value for $$
43 Lack of initial investment resources
44 Maintenance fees
96 Open source for research inexpensive but commercial tools costly
97 Pilot project initial investment

3.2 Technically Viable / HR

3 Predictive planning
4 Sophistication
5 HR Capacity
9 CRM
10 SIS
16 Patterns / Correlations
22 Relationships
26 Inaccurate prediction
31 Tech adoption
34 Data quality
36 User focus
47 Unstructured data
48 Data collection
49 Data analysis
50 ERP
52 LMS
55 Prescriptive results
56 AI/ML
58 Unsupervised classification model
74 Mixed results
75 Descriptive results
76 Predictive results
77 Lack of prescriptive results

79	Measuring specific LMS content impact on learner success
80	Measuring instructional strategies impact and learner success
81	Proxy measures
82	Lack of unexpected results
83	Confirmed inferences
87	Administrative improvements but not enhanced learning
88	Sceptical of tech/autonomous learning potential
89	Need to gather more data for 'big data' to work more effectively
91	Machine learning
92	Algorithms
94	Need more data to increase confidence levels
3.3	**Organizational Learning**
37	Organizational learning
46	Training
56	AI/ML
57	Insight
59	Unexpected results
60	Project proposals- deductive vs inductive rationale
61	Hazy understanding
66	Relationship based industry
67	Art vs science
72	Slower adoption
98	Risk adviers / not early adopters
99	IT intergrated to operations - not just a service provider
100	Need more sector wide IT partnerships
101	Need for causal understanding
103	Impact of technology on organizational culture
104	Informal skill development for unknown applications

4 Ethical & Legal

11	External information
33	transparency
48	Data collection (duplicate with tech viability)
49	Data analysis (duplicate)
55	Prescriptive results (duplicate)
70	Ethical Use
93	Bias in models
95	Data security / ethics

7.3 Memos

1. Code 16 Patterns/correlations - although not seen as a zero-sum competition, still important to leaders to measure success

2. Code 19 Better/good indication - institutions admit based on early marks - correlation - indicator of probable success

3. Code 20 Efficiency - when will tech lead to greater efficiency - lower costs?

4. Code 35 Organizations learning – Many organizations do not have skills sets for predictive or prescriptive analytics

5. Code 38 Value for $$ - value on 'business side but lack of prescriptive results for learning' - a business innovation not a learning innovation

6. Code 58 Project proposals, deductive vs inductive rationale - leadership needs a hypothesis to allocate resources - hypothesis linked to strategic goals/plans

7. Code 88 Driven by executive - IT needs to be a full exec member for programming- not just a service provider

8. Code 97 IT integrated to operations not just a service provider – does IT provide solutions or just data/analysis and different departments craft the tools they need – need a provider vs data

9. Code 100 Leadership in a complex environment - environments are so complex - the correlation based knowledge may be more reliable - causality more or an illusion in complex environments and is perhaps more biased?