This chapter has been accepted for publication in "Technology-Enhanced Learning and the Virtual University". Available at https://doi.org/10.1007/978-981-19-9438-8_29-1

Chapter 29: Using institutional data to drive quality, improvement and innovation

Sarah Dart (corresponding author)
Learning & Teaching Unit and Faculty of Engineering, Queensland University of Technology, Brisbane, Australia
sarah.dart@qut.edu.au

Samuel Cunningham
Learning & Teaching Unit and Faculty of Engineering, Queensland University of Technology, Brisbane, Australia
sam.cunningham@qut.edu.au

Abstract

Although higher education has historically been a data-rich sector, the move to fully online learning and teaching dramatically increases the quantity and scope of data that institutions can collect. This is because students produce a significant amount of trace data in virtual learning environments, which can be readily captured and dynamically analysed. When coupled with the more traditional university data, analytic insights from this comprehensive data set can be harnessed, presenting an enormous opportunity for driving evidence-based improvement. This chapter presents the various data sources that can underpin a data-driven strategy within a virtual higher education institution, and key considerations for leveraging this data to extract value. These strategies include strategic alignment, data management, integration into decision-making, and institutional culture. A series of contemporary examples are then used to illustrate how institutional data can be translated into action to successfully drive quality and innovation in a range of virtual learning and teaching situations.

Keywords

Data, analytics, quantitative, qualitative, evidence-based decision-making, quality enhancement, continuous improvement

Introduction

Although higher education has historically been a data-rich sector, the move to fully online learning and teaching dramatically increases the quantity and scope of data that institutions can collect. This is because students produce a significant amount of trace data in virtual learning environments (Leitner et al., 2017), which can be readily captured and dynamically analysed (Pardo et al., 2019). When coupled with the more traditional university data, analytic insights from this comprehensive data set can be harnessed, presenting an enormous opportunity for driving evidence-based improvement at all levels of virtual higher education institutions.

In addition to the shift toward online learning, the emergence of "big data" and the political environment are contributing to institutional adoption of data-driven strategies (Daniel, 2015; Ferguson, 2012; Knight et al., 2016). With regard to big data, tools for extracting, aggregating, storing, and managing large data sets have become more accessible, while visualisation tools have improved greatly (Ferguson, 2012). This has led to institutions implementing increasingly sophisticated algorithms and approaches for a range of applications including to identify at-risk students for intervention, customise communications to nudge student behaviour, personalise learning pathways, and provide dashboards that enable students to reflect on how they benchmark against their cohort (Liu et al., 2017). In terms of the political environment, institutions are under pressure to increase student participation while improving educational quality within a reduced funding envelope (Knight et al., 2016; Small et al., 2021). These developments are also coinciding with the diversification of students' backgrounds through widening participation trends (Dart & Spratt, 2021; Small et al., 2021). This context has meant finding highly scalable and efficient approaches to teaching and learning support that cater to student variability has become a key priority (Pardo et al., 2019). Given data can be used to implement this personalisation efficiently at scale, data-driven approaches have become increasingly prevalent (Dart & Spratt, 2021).

While leveraging data offers great promise, there are numerous challenges that must be navigated for benefits of a data-driven strategy to be realised in practice. For example, it is important that data collection is aligned to strategic priorities, and that analytical insights are closely integrated into decision-making practices to maximise impact (Fritz, 2011). Additionally, weak data governance, ethical concerns, poor institutional culture, and limited data capabilities can serve as barriers to implementation (Colvin et al., 2015). Where these issues are not overcome there is a significant risk that data-driven insights remain dormant, without ever being acted upon (Liu et al., 2017; Pistilli et al., 2014).

This chapter initially outlines the data sources that can underpin a virtual university's data-driven strategy, and then goes on to discuss key considerations for implementing such a strategy in practice. Finally, the chapter draws on contemporary examples from the literature to illustrate how institutional data can be used to successfully drive quality, improvement, and innovation within a virtual higher education landscape.

Data sources

An exceptionally wide variety of data sources can guide decision-making and developments within higher education institutions. Figure 1 summarises specific types according to whether the data originates from students, the institution, or external entities. Although classified separately in Figure 1, it should be noted that these varied sources can (and should) be considered in combination to facilitate deeper insight. For example, graduate outcomes can be assessed from a socio-demographic background perspective, while student and staff perceptions of a learning and teaching experience can be compared. The specific data types falling under each classification are discussed in more detail below.

Figure 1 – Data sources and types

Student data

Student data is fundamental to any data-driven strategy in the higher education context. There are several forms of data that students contribute through their relationship with an institution. This includes static data initially provided at enrolment, as well as the dynamic data progressively generated through students' interactions with the institution.

Socio-demographic background

Universities capture socio-demographic data from students at enrolment. This includes basic information like gender, age, cultural background, citizenship, and language background, as well as measures more specific to the sector like first-generation student status, tertiary admission score, fee-paying type, prior study outcomes, and admission pathway (Campbell, 2007; Patfield et al., 2021). Attributes like socio-economic status and regional background are often derived by connecting measures like high school attended or postcode at enrolment to other data sets, such as those generated from census records (Patfield et al., 2021). Collection of socio-demographic data enables analysis of sub-cohort participation rates, with the representation of students from traditionally disadvantaged backgrounds often of interest (Patfield et al., 2021). Socio-demographic data is also vital for analysing further data sets through an equity lens.

Study attributes

Fundamentally, universities hold data about students' study choices. This includes department, program, and course enrolment, part-time or full-time study status, level of study, majors or minors selected, and university standing (such as probation or semester honours) (Campbell, 2007). The timing and frequency of changes to these attributes (such as withdrawing from a subject or taking a leave of absence) can also be valuable, as these events can be lead indicators for outcomes such as attrition (Pistilli et al., 2014). Students gradually develop an educational history within the institution, which feeds into the number of credit points achieved and failed. Credit points can also be used to determine students' year of study within their program.

Subject and educator perceptions

Student Evaluation of Teaching (SET) surveys, also known as course evaluations, are widely used to assess student satisfaction with subjects and teaching (Goos & Salomons, 2017). SET surveys are delivered toward the end of each teaching period, and typically consist of both quantitative and qualitative questions about a specific subject and the associated teaching team (Cunningham-Nelson et al., 2020). Quantitative responses (often measured on Likert scales) enable comparisons to be readily drawn across educators, courses, departments, and institutions (Goos & Salomons, 2017), and trends can be tracked over time (Cunningham-Nelson et al., 2020). Qualitative free-text responses provide rich and specific feedback that can guide enhancements in teaching, curriculum design and assessment (Cunningham-Nelson et al., 2020). Although employed relatively infrequently, student perceptions can also be gathered through focus groups and interviews. These mechanisms allow greater depth and elaboration of student responses but are much more time-consuming and resource-intensive to implement. Therefore, these data collection methods are usually only considered when there is a clear purpose that goes beyond general evaluation, such as a major subject redesign or collecting evidence for an accreditation process.

Program perceptions

Surveys can be conducted at a program level to elicit broader student perceptions data. This is appropriate when evaluating central student support services or assessing the development of skills scaffolded across a program. Program-level surveys can be run internally by an institution or via an external body. Examples of the latter include the Student Experience Survey in Australia, the National Student Survey in the United Kingdom, and the National Survey of Student Engagement in the United States (Whiteley, 2016). These national surveys involve students from many institutions, with this reach and focus on quantitative questioning enabling benchmarking of key performance indicators across the respective sectors (Whiteley, 2016).

Online engagement

The virtual learning environment makes capturing student behaviour within the learning management system (LMS) highly accessible, given student-tracking capabilities are typically included as a software feature (Ferguson, 2012). Log data can provide insight into what students have engaged with on the LMS (such as assessment, discussion boards, and videos), as well as frequency and duration of access (Campbell, 2007). Other data that can be drawn out includes attendance at live online teaching sessions, the timing of an assessment submission (including whether a student has submitted at all), whether students have viewed their assessment feedback (Dart & Spratt, 2021), the content of and interaction patterns for discussion boards (Leitner et al., 2017), and video viewing patterns such as tendency to pause and rewind (Mirriahi et al., 2018). An emergent area for online engagement data is collaboration tools (such as Microsoft Teams and Google Docs) which may be used to record individual contributions within a virtual team environment. Log data as well as the content students progressively develop within these systems may be drawn into analysis, however it is key for these systems to be a part of the university's learning and teaching ecosystem for the data to be readily accessible to a data-driven strategy.

Engagement with extra-curricular activities and support services

Data can be collected on student engagement with extra-curricular activities, such as participation or leadership within student clubs or societies. Data may also be captured around student interactions with support services, such as accessing a drop-in help service or using resources from the library. A virtual university context makes gathering this type of data much more straightforward than an equivalent face-to-face environment, as online systems tend to automatically have these data-tracking capabilities.

Learning outcomes

Institutions collect a range of data relating to student learning outcomes. In addition to the actual artefacts that students submit, grades are recorded on individual assessment items. When criteria-based grading is utilised, grades may even be broken down into achievement against named benchmarks (Ragupathi & Lee, 2020). Qualitative feedback may also be provided to students – usually this is in a written form, but it can also be communicated via audio or video recordings (McCarthy, 2015). The outcomes of individual assessment items feed into students' final subject grade, and in turn, their grade point average (GPA) (Campbell, 2007).

Progression, retention, and completion outcomes

Higher education outcomes are measured in terms of progression, retention, and completion (Australian Government Department of Education and Training, 2018). Progression measures the pass rate for individual subjects. Retention measures the percentage of a student cohort who study in one year, and then continue study or graduate in the following year (Dart, 2019). This is the opposite of attrition, which captures those within the cohort who choose to withdraw from their studies without graduating. Completion measures the percentage of students within a cohort who graduate with degrees (Edwards & McMillan, 2015). As students must pass individual subjects to advance through their degrees, progression underpins completion and is a vital factor influencing retention (Crosling et al., 2009). It is crucial to note that while progression data becomes available at the end of each teaching period, retention and completion outcomes lag significantly. For retention, outcomes of a given starting cohort cannot be fully finalised until the following year has concluded (although only small changes are likely as the year draws to a close) (Dart, 2019). Similarly, students can take an extended period of time to complete a degree, particularly when they are studying in a part-time capacity or take a leave of absence (Edwards & McMillan, 2015).

Institution data

Institution data largely relates to staff in terms of their demographic backgrounds, perceptions of students' learning and teaching experiences, and engagement with professional development. However, two other aspects of institution data to consider are curriculum and finance.

Staff demographic background

Staff demographic attributes include department, school, and subject alignment, appointment type (such as full-time or casual), role focus (such as teaching-intensive or research-intensive), and level (such as lecturer or professor). Like the student socio-demographic data, recording staff backgrounds can provide valuable insight when connected to other data sets. For example, staff perceptions of learner engagement may differ across departments or experience level. Understanding which staff are performing well and which are struggling enables more nuanced actions to be taken.

Staff perceptions

Staff can provide valuable perspectives on learning and teaching designs, including insight into resourcing, technology, and student engagement. Gathering these perceptions in a systematic manner is particularly important where casual staff are employed in front-line teaching roles, given they usually have few avenues for sharing their feedback but commonly facilitate smaller classes where richer student-educator interactions transpire (Hemming & Power, 2021). Methods for gathering staff perceptions mirror the student perceptions data and thus centre mostly on surveying toward the end of teaching periods, with focus groups and interviews used sparingly. Surveying of staff may include both quantitative and qualitative questioning, and when interpreted alongside staff demographic and curriculum information (discussed below), results can be contextualised. Surveys may be conducted internally or externally such as the Faculty Survey of Student Engagement that is run nationally in the United States (Center for Postsecondary Research, 2020).

Professional development engagement

Given professional development is strongly tied to learning and teaching capability (Inamorato et al., 2019), collecting data on staff development can be valuable. Many institutions provide internal professional development through central learning and teaching units (Inamorato et al., 2019) that can form a foundation for systematic data capture. These central units may deliver ad-hoc development as well as operate more formal programs such as accredited qualifications in tertiary education or educational fellowship schemes (Greer et al., 2021). Data types that can be considered include workshop attendance, engagement with professional development modules, and recognised learning and teaching expertise against specific benchmarks. As an extension of this, scholarship of learning and teaching outputs and educational research publications can be tracked (Inamorato et al., 2019).

Curriculum information

Curriculum provides another source of institutional data, which can be key for contextualising information from other sources. One aspect of curriculum data is mapping where specific skills are developed across a program, including the level to which these skills are demonstrated. Curriculum data can also include learning outcomes, subject prerequisites, assessment methods, class types (such as lecture or tutorial), credit points, and synchronous contact time.

Financial information

The financial impact of decision-making is critical in higher education institutions, especially given the political pressures forcing increased operational efficiency (Goldstein & Katz, 2005; Leitner et al., 2017). Therefore, financial data needs to be represented in decision-making to ensure viability of actions from an economic perspective. Aspects to consider include student fees, and the costs of staffing, learning materials, and technology (Twigg, 2003). Investment in staffing and infrastructure is also required to operationalise a data-driven strategy, which should be considered when making decisions about which approaches are pursued (Corrin et al., 2019).

External data

External stakeholders, including university graduates and industry, provide another source of data that can strongly influence learning and teaching practices. These stakeholders provide data relating to employment outcomes, graduate competencies, and accreditation.

Graduate employment outcomes

Employment is one aspect of employability, and represents a key data set for understanding the success of graduating students (Small et al., 2021). Employment outcomes are gathered from alumni, usually through large scale surveys such as the Gradate Outcomes Survey in Australia (Whiteley, 2016) and the National Survey of College Graduates in the United States (United States Census Bureau, 2019). Data of interest includes employment or further study status, usual hours worked, employment type (such as voluntary, casual or continuing), typical work activities, salary, industry, and overall well-being. This data may be collected shortly after students have graduated, and after alumni have developed experience in the workforce in order to support a longitudinal understanding (Whiteley, 2016).

Post-program perceptions of alumni

Past students can provide insight into the skills and capabilities developed through a program, and how they perceive this to relate to their subsequent employment. This insight can be used to pinpoint gaps or weaknesses, as well as strengths in curriculum and teaching approaches. This data may be collected via surveying where a wide sample is desirable, or through interviews and focus groups where the richness and depth of response is prioritised.

Industry and employer perceptions

Industry can provide valuable perspectives on the skills and capabilities expected of graduates when entering their respective fields. The data can help to identify curriculum gaps, while ensuring programs maintain currency with industry developments. Employers of an institution's graduates can also provide feedback on whether graduates are meeting expectations. Industry and employer perceptions data may be collected through surveying (such as the Employer Satisfaction Survey in Australia (Whiteley, 2016)), or through focus groups or interviews with industry representatives.

Accreditation outcomes

Many programs undergo regular accreditation cycles, an evidence-based process for evaluating whether programs meet stated benchmarks (Borrego & Henderson, 2014). Here institutions typically put forward a range of evidence for consideration by an external panel of peers and industry experts (Borrego & Henderson, 2014). Alongside the judgement of whether a program should be accredited, institutions receive feedback, which serves as a further data point for informing curriculum and teaching approaches.

Strategies for leveraging institutional data

The previous section has highlighted the significant amount of data that virtual higher education institutions can consider feeding into an analytics strategy. However, channelling this in a way that maximises value presents as an immense challenge. Key factors in this process are presented below, comprising alignment, management, integration, and culture.

Strategic alignment

Given the vast data institutions have at their fingertips, developing a data-driven program that aligns data collection and analytical activities to core priorities is critical for impact (Pistilli et al., 2014). A necessary starting point for this involves clarifying the overarching objectives for a data-driven program, and where it fits within institutional goals (Fritz, 2011). Where objectives are unclear, a needs assessment should be completed (Spector & Yuen, 2016) to establish the vital context that underpins decisions about what data is necessary for the program's activities, as well as what should be measured to monitor the program's performance and demonstrate its value.

Logic models can be considered as a framework for connecting how a program operates to a range of measures that collectively demonstrate program success (Figure 2) (McLaughlin & Jordan, 2015; Spector & Yuen, 2016). This technique first considers the resources and activities that underpin a program. Explicitly articulating these elements is useful for identifying fundamental data requirements and infrastructure necessary for a program to function. This includes data sources and types (such as those summarised in Figure 1), specific hardware and software, and staffing. Unsurprisingly, these requirements vary considerably according to the specific objectives of the program. The second stage of logic models considers the different types of results produced, which are characterised as either outputs, outcomes, or impacts. Outputs are immediate short-term results, outcomes capture intermediate results, and impacts represent the high-level attributes changed over the long-term (McLaughlin & Jordan, 2015). For example, if a program's objective is enhancing students' employability (the impact), an output could be the number of students the program reaches, while an outcome could be improvements in students' professional skills theorised to have a relationship with employability. Where possible, programs should seek to utilise existing data sets rather than generating new data (Pistilli et al., 2014).

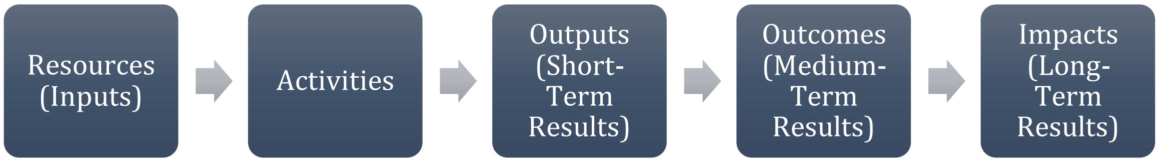

Figure 2 – Logic model structure for connecting how a program operates to its results; adapted from McLaughlin and Jordan (2015)

Often the objectives of a data-driven program are very high-level, such as to improve student retention or enhance the employability of graduates. It is important to evaluate programs at this level to maintain an ongoing understanding of performance and to evidence impact. However, measuring these high-level objectives directly can take an enormous amount of time and be extremely costly (Spector & Yuen, 2016). For example, retention outcomes for a newly enrolled student cohort cannot be determined until the following year, while comprehensively measuring employability requires resource-intensive longitudinal tracking of graduates in the workforce. Additionally, there may be confounding factors that influence these high-level attributes (such as the state of the employment market), thus making it difficult to isolate the program's contribution. To address these challenges, data collected on short- to medium-term measures of program success (that is, outputs and outcomes) can act as indicative proxies (Pistilli et al., 2014). Where success can be shown across the results chain, a stronger claim can be made about the program's influence. Identifying key measures across the results chain should be prioritised upfront in program design to ensure processes are put in place to capture the required data in a systematic manner. This is particularly important for the impact measures that must be gauged in a consistent manner over a long period of time to be highly informative and persuasive (Spector & Yuen, 2016).

Engaging in a continuous improvement cycle is recommended for enhancing alignment between activities and strategic objectives over time (Borrego & Henderson, 2014; Spector & Yuen, 2016). One popular framework for facilitating continuous improvement is "Plan, Do, Check, Action" (Bond, 1999). Planning involves designing enhancement initiatives, which are then piloted in the doing phase. The influence on performance indicators is evaluated in the checking phase. The final stage of the framework involves broadly applying actions deemed successful, before repeating the cycle (Bond, 1999). Over progressive iterations, improved alignment between activities undertaken and strategic aims is expected.

Selecting key performance indicators is advised for guiding continuous improvement processes to overcome the potentially overwhelming amount of data available (Bond, 1999). It is important to select indicators from across a logic model's results chain, given shorter-term measures can contribute to more responsive monitoring and improvement actions but longer-term measures capture overall impact. Moreover, indicators should be drawn from a variety of sources and types (including both quantitative and qualitative data). This is because consideration of multiple sources works to increase the external validity of conclusions made about a program, while most data collection mechanisms are not specifically tailored to evaluating a program so only become valuable when interpreted alongside other data (Smith, 2008). The "Four Quadrant" approach proposed by Smith (2008) argues self-reflection (which may refer to an individual or institution), student learning, peer review, and student experience should be triangulated when holistically evaluating the impact of a learning and teaching intervention (Figure 3). Examples of the how these recommendations can be fulfilled include internal quality assurance mechanisms that facilitate self-reflection, using learning outcomes from assessment to measure student learning, considering SET survey results for student experience, and leveraging the peer review that occurs through accreditation.

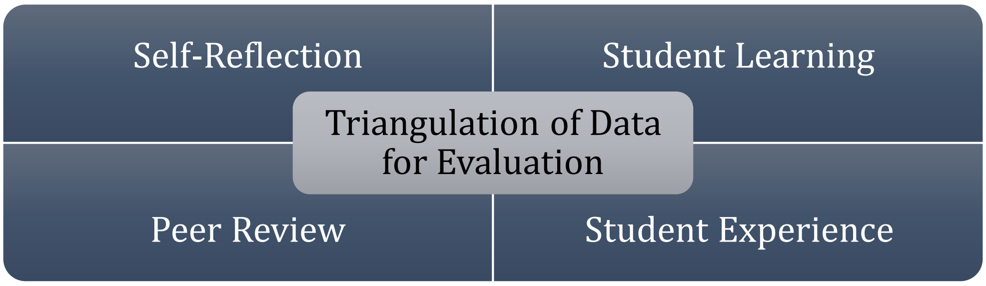

Figure 3 – "Four Quadrant" approach to triangulating data when evaluating learning and teaching interventions; adapted from Smith (2008)

Data management

Governance is concerned with the structure and processes for distributing accountability and control in higher education institutions (Elouazizi, 2014). This control is typically applied at an institutional level for key data assets (such as student and curriculum data). Often extensive information technology systems are used to bring this data together to centrally report an agreed source of truth (Goldstein & Katz, 2005). Enabling access to this repository is crucial for empowering end-users to ask questions of the data that are relevant to their contexts, which can rapidly grow the reach and impact of analytics efforts. However, access needs to be balanced against security and privacy concerns by giving due consideration to data management controls (Colvin et al., 2015). This is particularly critical for sensitive data sets, such as those containing extensive student information which is individually identifiable.

Beyond provisioning access, ease of manipulation and interrogation of data for end-users is critical for extracting value (Colvin et al., 2015). This includes how easily data sets can be connected (even to data sets that fall outside central systems), as this is a precursor to many of the more sophisticated analytical approaches (Barber & Sharkey, 2012; Fritz, 2011; Pistilli et al., 2014). Consequently, developing an agreed set of operational definitions for how variables are measured (including underlying assumptions) (Patfield et al., 2021; Spector & Yuen, 2016) is vital for ensuring data accuracy and transparency in interpretation, and enables data sets (that each adopt the definitions) to be jointly analysed (Colvin et al., 2015). For example, defining how a student's socio-economic status is determined is critical for interpreting equity data (Patfield et al., 2021), while understanding how retention outcomes are calculated is necessary for fair comparisons, given retention may be considered at a program, institution, or sector level.

Pistilli et al. (2014, p. 85) note that "absent a strong foundation of good data, any analytics effort will likely fail." Consequently, an ongoing challenge for data-driven strategies is maintaining clean data (Barber & Sharkey, 2012). For manually derived data (such as from surveys), simple steps at data capture can reduce this burden substantially. For example, where there is a finite group of categorical options, drop-down or checkbox menus can limit selection to the permissible list. Similarly, if individuals must enter numbers that sum to a certain value (such as 100%), validating the inputs is strongly recommended. To reduce incidence of missing data, forced responses can be applied. Where missing data is highly problematic, ambient data collected in a non-intrusive manner should be considered as an alternative (Pistilli et al., 2014). For example, log data from a virtual learning environment is likely a preferable substitute to surveying students directly about their engagement.

Ethical protocols must underpin how data are collected, shared, and utilised (Colvin et al., 2015; Corrin et al., 2019; Pargman & McGrath, 2021). Fundamental to this is the implicit expectation that analytical developments generate benefit without causing harm (Corrin et al., 2019). However, this creates questions about the level of intrusion permissible for data collection efforts (Pistilli et al., 2014), informed consent, data transparency and validity (Pargman & McGrath, 2021), and institutional responsibility in intervening on the basis of analysis (Corrin et al., 2019). Managing these ethical concerns is crucial to obtaining buy-in for data-driven strategies from all stakeholders, including students, and this becomes increasingly salient as approaches to using data mature (Colvin et al., 2015; West et al., 2020).

Integration into practices, processes, and policies

A goal of data-driven strategies involves shifting decision-making practices to emphasise use of the evidence-base, rather than pure instinct. This is because data supports identification of new opportunities, enables exploration of alternative options, and counters group thinking, ultimately leading to better outcomes (Bonabeau, 2003). Core to achieving this is the integration of data into decision-making processes, such that insights are acted upon in an appropriate and timely manner (Liu et al., 2017; Pistilli et al., 2014). This is not just applicable for strategic decision-makers at upper management levels of institutions, but also educators who frequently make decisions about how to best support students (Borrego & Henderson, 2014; Colvin et al., 2015).

Policy is a key lever for influencing adoption of evidence-based practices (Borrego & Henderson, 2014). This operates on the premise individuals and teams will "collect and analyze their own evidence to evaluate and improve their ability to meet stated goals" (Borrego & Henderson, 2014, p. 235). Policies may signal specific sources of evidence that need to be considered in a decision-making process. Examples of where this is commonly employed at an institutional level include program accreditation, curriculum design, and academic integrity management. At an individual level, academic promotion policies can substantially influence data usage in learning and teaching contexts, such as when SET scores are nominated as a key indicator of educator quality (Heffernan, 2021). However, setting narrow definitions for measuring quality risks perverse consequences. For example, SET scores have been shown to disadvantage educators from specific minority backgrounds (Heffernan, 2021) and can mask other factors at play. Consequently, triangulation of multiple indicators should be encouraged (Figure 3).

The extent to which data drives decision-making within an institution depends on the approaches employed. Goldstein and Katz (2005, p. 60) proposed a five-stage hierarchy for classifying institutions' data usage maturity. Immature institutions relied on extracting and reporting transaction-level data (Stage 1), which then evolved to basic analysis to monitor operational performance (Stage 2) and support "what-if" planning decisions (Stage 3). As institutions further developed, they pursued predictive modelling (Stage 4), and eventually embedded automatic triggers into business processes, such as when a key metric fell outside a desired range (Stage 5). From this hierarchy it can be observed that increasing the sophistication of data-driven approaches creates opportunities for more informed, relevant, and timely decision-making.

In terms of educators, an area of interest is how technological tools can improve decision-making related to teaching and student support (Ferguson, 2012). This is particularly relevant to virtual environments for two key reasons. First, educators lack traditional student engagement cues when teaching online, such as visually observing whether a student is actively participating and on-track, or overwhelmed and absent (Ferguson, 2012). Second, virtual environments generate a substantial amount of trace data related to engagement and performance (Leitner et al., 2017). Learning analytics has sought to find ways of unifying this data with reporting and analysis that supports educators in observing what would otherwise go unseen (Liu et al., 2017). However, benefits will only be realised when educators integrate these tools into routine practices by deploying actions based on the evidence (Liu et al., 2017). This may be at the macro-level, where teaching modifications are made for an overall cohort, or at the micro-level, where adjustments are made for individual students. Ensuring analytical developments are responding to the real needs of educators in their teaching is core to achieving this goal (Liu et al., 2017).

Institutional culture and capability

Institutional culture plays a key role in the adoption of evidence-based practices, and the extent to which individuals embrace measurement and improvement as an ongoing process (Borrego & Henderson, 2014; Pistilli et al., 2014). The commitment of those in leadership positions to data-driven practices contributes to motivating this culture (Goldstein & Katz, 2005) due to the strong influence over data availability, including brokering systematic collection mechanisms and mediating the relationship between data owners and end-users (Liu et al., 2017). This is particularly critical in the virtual university where selection of software tools greatly influences the type of data available. Leadership is also required in ensuring quality in data sets and analysis tools (Colvin et al., 2015). The interplay between data capture, sharing, and utilisation shapes desire within the institution to engage, and sets the tone for analytic efforts growing or fading. In fact, Colvin et al. (2015) found formation of a shared vision that effectively responded to institutional needs to be a core factor in ensuring sustainable analytics implementations. Additionally, leadership establishes the risk appetite. Where individuals feel supported to experiment with new approaches to using data, one can expect innovations to emerge. This contrasts against a culture where data is tightly regulated and where evaluation is used as a "hammer" to exert control (Pistilli et al., 2014).

Strategic capability is crucial within a data-driven environment (Colvin et al., 2015). This is because specific expertise is required to guide decisions around what data is collected to align with needs, the development of strategies for managing this data appropriately, and how developments can be diffused through the institution for maximum impact. Consequently, a commitment to resourcing a highly-skilled workforce with strong capabilities in technology, data, analysis, and evaluation is required (Goldstein & Katz, 2005). In fact, Goldstein and Katz (2005) emphasise that it is the skills of staff (especially in understanding and manipulating data) that limit what can be achieved when embarking on analytical approaches, rather than the technology. However, it is worth noting that a strong understanding of the educational context is also important in guiding developments. Professional development, secondments, and supporting staff to complete postgraduate study have been recommended for capacity building (Colvin et al., 2015).

Individuals can be encouraged to evolve toward data-driven practices through a range of mechanisms. One such path involves building awareness of the opportunities and advantages of data-driven approaches, such as through planned communication strategies or word of mouth. As individuals progressively choose to adopt these approaches, a "tipping point" is reached where it becomes a normal part of practice anchored in the culture (Borrego & Henderson, 2014). Professional development can also be used to influence adoption decisions through raising capacity and confidence to engage with analytics (Colvin et al., 2015). Providing these developmental opportunities coupled with ongoing support is particularly important for ensuring data is embraced on mass by those in a wide range of roles, not just those who are enthusiastic about data and technology (Liu et al., 2017). Training can also minimise data misuse, such as where messiness in underlying data sets leads to misinterpretation and inappropriate action (Liu et al., 2017). Colvin et al. (2015) further highlight the importance of empowering students in analytical developments, especially for those interventions that rely on students exhibiting their agency as part of self-regulated learning processes. Here students require sufficient skills in understanding and interpreting data-driven insights to take responsibility in driving their learning effectively.

Motivating educators to approach their teaching as a scholarly activity is another avenue for increasing adoption of evidence-based approaches (Borrego & Henderson, 2014). This involves shifting educators' to be consistently thinking about the effectiveness of their teaching methods and how they could improve. Means for this exist on a spectrum ranging from personal reflection to formal educational research, with data integral to the process regardless of the pathway (Borrego & Henderson, 2014). Higher Education Academy fellowships, which facilitate reflection against a Professional Standards Framework, have been used to strategically support this shift (Greer et al., 2021). Rewarding learning and teaching quality (including educational research) through academic promotion processes has also been recommended as a lever for stimulating cultural change (Borrego & Streveler, 2015).

Fostering communities of practice can complement scholarly teaching endeavours. This involves creating spaces to connect those across an institution with interest in specific topics (Liu et al., 2017), such as using a particular software tool or deriving meaning from a specific data set. These forums facilitate practice-sharing, contributing to the development of increasingly mature evidence-based practices and a culture of continuous improvement. This approach can also create change by linking data experts with end-users to collaboratively work on solving context-specific problems.

Contemporary data-driven approaches for advancing educational practices

Pistilli et al. (2014, p. 80) emphasise "institutions cannot simply collect and report on data… [They] must take specific actions to enhance student success." As institutions have incorporated increasingly sophisticated data-driven strategies, a range of these actions have emerged including "prediction, intervention, recommendation, personalization and reflection" (Leitner et al., 2017, p. 6). Contemporary examples of these actions applicable to the virtual university context are discussed below.

Predictive modelling of student outcomes

Predictive modelling represents one of the most popular data-driven approaches employed by higher education institutions (Corrin et al., 2019; Goldstein & Katz, 2005; Leitner et al., 2017). This relies on data from past cohorts to make predictions about those that follow, with subject-based performance and retention the most common outcomes to predict (Corrin et al., 2019). Predictive models can methodically flag at-risk students, with educators or support staff often used to contact these students directly with tailored support interventions (Barber & Sharkey, 2012; Colvin et al., 2015). Predictive models can also enhance understanding of key risk factors (Dart, 2019), enabling more strategic actions to be taken at the cohort level (such as changing program admission requirements, subject prerequisites, or assessment approaches).

A wide range of statistical techniques have been employed for predictive models including regression, decision trees, support vector machines, and neural networks (Corrin et al., 2019; Kotsiantis et al., 2010; Marbouti et al., 2016). These techniques differ greatly in terms of complexity and computational cost, with suitability highly dependent on underlying data and model purpose. Predictive modelling of at-risk students may be conducted at a subject level (e.g. Marbouti et al. (2016)) as well as within or across programs (e.g. Campbell (2007)). While high-level models can draw in larger numbers of observations to improve predictive power, this comes at a cost of being limited to considering students' demographics, educational backgrounds, and high-level LMS engagement as these are the only variables common across all students (Liu et al., 2017). Conversely, subject-level models can consider very specific variables, such as performance in assessment items and engagement with individual activities, which have been shown to substantially improve model accuracy (Dart, 2019; Kotsiantis et al., 2010; Marbouti et al., 2016). It is worth noting that a virtual learning and teaching landscape dramatically increases the number of indicators available for predictive models, given engagement and performance data is often captured by online tools in a highly structured manner that minimises missing data (Ferguson, 2012).

There are several ethical issues that need to be considered for predictive modelling. In particular, Liu et al. (2017, p. 146) highlight that making predictions based only on students' background characteristics "risks limiting our view of students' ability to their past performance and, at worst, perpetuates stereotypes." Corrin et al. (2019) emphasise ethics when selecting variables for inclusion in modelling, while expert oversight is critical when developing and interpreting results to mitigate against validity concerns. Missing data and data quality issues can compromise predictive modelling efforts (Barber & Sharkey, 2012). Similarly, where data lies in disparate systems, a significant amount of work is required to unify sources (Barber & Sharkey, 2012). This is becoming a greater challenge as educators rely on a wider variety of online learning tools and students look to sources outside the LMS (such as YouTube), which means trace data is not always recorded in institutional systems (Dart et al., 2020; Pistilli et al., 2014).

Personalised guidance and feedback through email nudging

Personalised email communications have emerged as a scalable method for nudging students toward certain actions, with data-driven insights used to tailor messaging (Dart & Spratt, 2021). Data underpinning this approach typically includes students' demographic and educational backgrounds, engagement (especially with online systems), and assessment performance (Liu et al., 2017). The approach is well-suited to a virtual university context given the extent of the systematically collected data available (especially relating to learner engagement with online tools). Mail merging (Dart & Spratt, 2021) or specialist software (Lim et al., 2019; Liu et al., 2017; Pardo et al., 2018) can be used for disseminating emails. However, the latter is preferable as it brings together data, analysis, and the email interventions in a unified workflow (Liu et al., 2017).

Early developments in the personalised email space tended to focus on students deemed likely to fail a subject or withdraw from their program (Liu et al., 2017). This meant predictive modelling was often a precursor step necessary to identify target students. One of the most well-known examples is Course Signals (Arnold & Pistilli, 2012) where LMS engagement, performance, and background data were used to generate a risk rating for each student. Educators developed personalised emails for those most at-risk to guide these students toward supports and remedial actions.

More recently there has been a move away from the deficit approach that focuses on at-risk students, to instead consider how personalised messaging can be embedded as a support mechanism in routine learning and teaching practices (Dart & Spratt, 2021). In these situations, statistical methods are not employed to calculate risk. Instead, educators' intrinsic understanding of behaviours that demonstrate effective learning practices guide what data is consulted and the associated messaging. This centres the relationship between the educator and student, and firmly contextualises recommended actions within the learning environment (Liu et al., 2017). Specific examples of where this has been employed include to aid student transitions to university by "fostering feelings of belonging, supporting effective engagement, and easing navigation of university systems and processes" (Dart & Spratt, 2021, p. 10), providing feedback to students showing different levels of interaction with learning resources (Pardo et al., 2019), and prompting engagement for online students showing signs of disengagement (Lawrence et al., 2019). Personalised email interventions have been shown to translate to increased perceptions of support (Lawrence et al., 2019), enhanced course satisfaction (Dart & Spratt, 2021; Pardo et al., 2019), and improved grade performance (Dart & Spratt, 2021; Lim et al., 2019).

Adaptive learning pathways

Online learning provides substantially more opportunities for personalising experiences than traditional face-to-face environments, enabling more efficient and effective learning (Vie et al., 2017). This is because the systematically captured engagement and performance data can be readily used to customise learning to individuals' needs, which is becoming increasingly important as cohort diversity increases with widening participation (Small et al., 2021). In simplistic implementations, students are asked multiple-choice or numerical response questions, with feedback given according to chosen answers. This feedback can include whether students have answered the question correctly, clarifying reasoning, and guidance on other resources to engage with for further support. More sophisticated implementations respond to each students' needs by dynamically changing the content presented according to digital trace data (Corrin et al., 2019; Khosravi et al., 2017). Systems incorporating this more advanced approach are often referred to as adaptive learning platforms (Corrin et al., 2019).

Decisions about how the learning pathway is adapted can be made using "if-then" rules. This draws upon the expertise of educators in understanding how learning progresses within their context. For example, students may be given a diagnostic activity at the beginning of a course. According to their performance, they can be directed to content that addresses the specific gaps identified in their assumed knowledge (Vie et al., 2017). Alternatively, variations on a problem may be given until the student is verified to possess robust understanding (Corrin et al., 2019). Here the real-time feedback educators receive from the system on students' progress is important for driving critical reflection that leads to content enhancements and teaching adjustments in corresponding synchronous class components (Prusty et al., 2011). Learning pathway decisions may also be based on models generated from other users' data (Vie et al., 2017). An example of an adaptive learning system using this approach is the Recommendation in Peer-Learning Environments (RiPLE) system of Khosravi et al. (2017). This was designed to improve the relevance of crowd-sourced questions by suggesting content corresponding to individuals' interests and knowledge gaps. It provides students with rich, relevant, and timely peer-generated feedback to further drive learning.

Textual analysis of student-written passages

Data in the form of numbers and categories represents structured data, while unstructured data comes in formats such as text-based documents, video, audio, and images (Daniel, 2015). Most institutions rely heavily on structured data for their analytics as this is far easier to analyse, but with developments in analysis and visualisation tools, mining the more complex unstructured data for meaningful patterns is on the rise (Daniel, 2015). Examples of natural language processing in a learning analytics context include assessing students' written language (Peña-Ayala et al., 2017) and conceptual understanding (Cunningham-Nelson et al., 2018) to provide immediate feedback, exploring connections between ideas in discussion forum or social media settings (Peña-Ayala et al., 2017), and extracting trends from institutional evaluation data containing a large number of qualitative comments (Cunningham-Nelson et al., 2019). Again, application of analysis techniques is well-suited to the virtual university environment where this type of data is automatically collected as part of routine practices.

One common technique for analysing large samples of qualitative comments is sentiment analysis. This involves labelling data as either positive or negative (Medhat et al., 2014), which can be on a binary scale (positive/negative) or on continuous scale (such as -5 to +5). This enables comments to be grouped based on positive and negative sentiment, as well as sorted from most to least positive for further consideration. Another method for analysing textual data is thematic analysis. Traditionally, this has been performed manually by individuals or teams defining common codes, tagging where these are evident in the data, and then following a verification process (Braun & Clarke, 2012). While effective, this thematic analysis is not time efficient and not easily replicable. A more recent alternative approach is topic modelling, which can be used to automatically extract key terms and themes from textual data (Murakami et al., 2017). This can identify recurring patterns in large quantities of text, and in turn provide a summary of areas that may need to be addressed with action.

Cunningham-Nelson et al. (2019, 2020) provide examples of how sentiment analysis and topic modelling can be combined to generate visualisations of SET data, such as in Figure 4. The visualisations include automatically generated key topics or themes on the horizontal axis and a sentiment count on the vertical axis. This allows educators to easily identify topics commonly mentioned, and whether students are discussing those in a positive or negative way, supporting educators to prioritise responsive actions. The visualisations also allow for comparisons to be made between the same subject across teaching periods, highlighting trends in student perceptions over time.

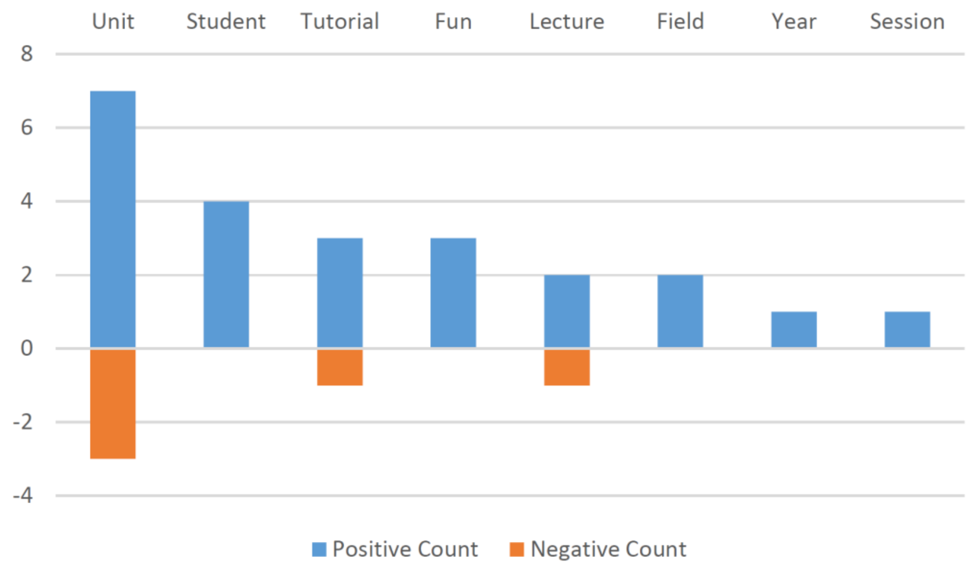

Figure 4 – Example of sentiment and topic visualisation for a subject

Benchmarking against peers

Benchmarking involves drawing comparisons between an entity and a broader sample from the population. Benchmarking has historically been conducted across institutions (Tasopoulou & Tsiotras, 2017), especially through large-scale national surveys measuring student experience and graduate employment outcomes (Whiteley, 2016). These surveys tend to have a strong emphasis on quantitative questioning to facilitate simple numerical comparisons across the sector. Engaging in this type of metric benchmarking process enables institutions to systematically compare themselves to similar entities for the purpose of identifying relative strengths and weaknesses (Tasopoulou & Tsiotras, 2017). This insight can be used within continuous improvement processes to drive quality enhancement (Bond, 1999). However, it is important to set reasonable expectations for benchmarking indictors, as unrealistic goals can have a harmful impact on motivation (Tasopoulou & Tsiotras, 2017).

A more recent innovation with regard to benchmarking involves student-facing dashboards that compare a student to their cohort. These provide students real-time feedback on their engagement and performance throughout the semester, and can in turn support self-regulated learning (Fritz, 2011). This can also reduce the burden on educators to monitor and act, which is particularly advantageous in large classes where instructor time is stretched (Pardo et al., 2019). An example of where this has been implemented is Course Signals, where the risk rating developed from predictive modelling was displayed as a traffic light signal on each students' LMS home page (Arnold & Pistilli, 2012). Similarly, Fritz (2011) implemented a "Check My Activity" tool that compared students' grade and LMS activity to their peers anonymously. However, in these student-facing systems due consideration needs to be given to ethics, especially around the types of data used (West et al., 2020) and the influence of negative judgments in demotivating students (Corrin et al., 2019). For example, Arnold and Pistilli (2012) noted a handful of students felt "demoralized" by negative feedback, while signals that were not regularly updated risked giving false indications. Poorly designed dashboard systems can also be confusing for students, causing unintended stress and damage (Corrin et al., 2019). The significance of on-boarding students has also been highlighted, with Fritz (2011) stating students needed to be proactively introduced to their tool to understand how and why it would be useful.

Conclusion and future directions

This chapter has presented the various data sources and types that can underpin a data-driven strategy within a virtual higher education institution, and key considerations for leveraging this to extract value. These strategies included strategic alignment, data management, integration into decision-making, and institutional culture. A series of contemporary examples were then used to illustrate how institutional data can be translated into action to successfully drive quality, improvement, and innovation in a range of learning and teaching situations.

Contextual factors continue to disrupt the higher education landscape including through growing student diversity, changes in funding structures, greater external accountability expectations, and evolutions in technology (Ferguson, 2012; Knight et al., 2016). The latter has been particularly spurred by the COVID-19 public health crisis, which has forced universities to redefine how teaching is delivered and learning is undertaken in fully online modes. This recent shift has meant the extent of the data opportunities in virtual learning environments continues to grow rapidly. Thus, as we move beyond this period of peak disruption, we would expect institutions to be increasingly looking to data-driven approaches to guide their decision-making when responding to external pressures, while implementing the more sophisticated and innovative advances made possible within a virtual environment.

References

Arnold, K. E., & Pistilli, M. D. (2012). *Course signals at Purdue: Using learning analytics to increase student success.* 2nd International Conference on Learning Analytics and Knowledge, Vancouver, Canada. https://doi.org/10.1145/2330601.2330666

Australian Government Department of Education and Training. (2018). *Higher education standards panel final report - Improving retention, completion and success in higher education*. https://docs.education.gov.au/system/files/doc/other/final_report_for_publishing.pdf

Barber, R., & Sharkey, M. (2012). *Course correction: Using analytics to predict course success.* 2nd International Conference on Learning Analytics and Knowledge, Vancouver, Canada. https://dl.acm.org/doi/abs/10.1145/2330601.2330664

Bonabeau, E. (2003). Don't trust your gut. *Harvard Business Review*, *81*(5), 116-123. https://hbr.org/2003/05/dont-trust-your-gut

Bond, T. (1999). The role of performance measurement in continuous improvement. *International Journal of Operations & Production Management*, *19*(12), 1318-1334. https://doi.org/10.1108/01443579910294291

Borrego, M., & Henderson, C. (2014). Increasing the use of evidence-based teaching in STEM higher education: A comparison of eight change strategies. *Journal of Engineering Education*, *103*(2), 220-252. https://doi.org/10.1002/jee.20040

Borrego, M., & Streveler, R. A. (2015). Preparing engineering educators for engineering education research. In A. Johri & B. M. Olds (Eds.), *Cambridge Handbook of Engineering Education Research* (pp. 457-474). Cambridge University Press.

Braun, V., & Clarke, V. (2012). Thematic analysis. In *APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological.* (pp. 57-71). American Psychological Association. https://doi.org/10.1037/13620-004

Campbell, J. P. (2007). *Utilizing student data within the course management system to determine undergraduate student academic success: An exploratory study* Purdue University]. West Lafayette, Indiana. https://docs.lib.purdue.edu/dissertations/AAI3287222/

Center for Postsecondary Research. (2020). *Faculty survey of student engagement*. Retrieved February 15 from https://nsse.indiana.edu/fsse/about-fsse/index.html

Colvin, C., Rodgers, T., Wade, A., Dawson, S., Gasevic, D., Shum, S. B., Nelson, K. J., Alexander, S., Lockyer, L., & Kennedy, G. (2015). *Student retention and learning analytics: A snapshot of Australian practices and a framework for advancement. Final report*. https://research.usc.edu.au/discovery/fulldisplay/alma99449564202621/61USC_INST:ResearchRepository

Corrin, L., Kennedy, G., French, S., Shum, S. B., Kitto, K., Pardo, A., West, D., Mirriahi, N., & Colvin, C. (2019). *The ethics of learning analytics in Australian higher education: A discussion paper*. https://melbourne-cshe.unimelb.edu.au/research/research-projects/edutech/the-ethical-use-of-learning-analytics

Crosling, G., Heagney, M., & Thomas, L. (2009). Improving student retention in higher education: Improving teaching and learning. *Australian Universities' Review*, *51*(2), 9-18. https://search.informit.org/doi/abs/10.3316/ielapa.159225407205474

Cunningham-Nelson, S., Baktashmotlagh, M., & Boles, W. (2019). Visualizing student opinion through text analysis. *IEEE Transactions on Education*, *62*(4), 305-311. https://doi.org/10.1109/TE.2019.2924385

Cunningham-Nelson, S., Laundon, M., & Cathcart, A. (2020). Beyond satisfaction scores: visualising student comments for whole-of-course evaluation. *Assessment & Evaluation in Higher Education*, 1-16. https://doi.org/10.1080/02602938.2020.1805409

Cunningham-Nelson, S., Mukherjee, M., Goncher, A., & Boles, W. (2018). Text analysis in education: a review of selected software packages with an application for analysing students' conceptual understanding. *Australasian Journal of Engineering Education*, *23*(1), 25-39. https://doi.org/10.1080/22054952.2018.1502914

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

Dart, S. (2019). *Developing predictive models of student success in undergraduate engineering mathematics courses.* 30th Australasian Association for Engineering Education Annual Conference, Brisbane, Queensland. https://aaee.net.au/wp-content/uploads/2020/07/AAEE2019_Annual_Conference_paper_38.pdf

Dart, S., Pickering, E., & Dawes, L. (2020). Worked example videos for blended learning in undergraduate engineering. *Advances in Engineering Education*, *8*(2), 1-22. https://advances.asee.org/worked-example-videos-for-blended-learning-in-undergraduate-engineering/

Dart, S., & Spratt, B. (2021). Personalised emails in first-year mathematics: Exploring a scalable strategy for improving student experiences and outcomes. *Student Success*, *12*(1), 1-12. https://doi.org/10.5204/ssj.1543

Edwards, D., & McMillan, J. (2015). *Completing university in a growing sector: Is equity an issue?* https://research.acer.edu.au/higher_education/43/

Elouazizi, N. (2014). Critical factors in data governance for learning analytics. *Journal of Learning Analytics*, *1*(3), 211-222. https://doi.org/10.18608/jla.2014.13.25

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

Fritz, J. (2011). Classroom walls that talk: Using online course activity data of successful students to raise self-awareness of underperforming peers. *The Internet and Higher Education*, *14*(2), 89-97. https://doi.org/10.1016/j.iheduc.2010.07.007

Goldstein, P. J., & Katz, R. N. (2005). *Academic analytics: The uses of management information and technology in higher education*. https://www.educause.edu/ir/library/pdf/ers0508/rs/ers0508w.pdf

Goos, M., & Salomons, A. (2017). Measuring teaching quality in higher education: assessing selection bias in course evaluations. *Research in Higher Education*, *58*(4), 341-364. https://doi.org/10.1007/s11162-016-9429-8

Greer, D. A., Cathcart, A., & Swalwell, G. C. (2021). Compounding the impact of teaching development programs in China and Hong Kong SAR: using the Professional Standards Framework to deepen learning and improve teaching self-efficacy. *International Journal for Academic Development*, 1-15. https://doi.org/10.1080/1360144X.2021.1878032

Heffernan, T. (2021). Sexism, racism, prejudice, and bias: a literature review and synthesis of research surrounding student evaluations of courses and teaching. *Assessment & Evaluation in Higher Education*, 1-11. https://doi.org/10.1080/02602938.2021.1888075

Hemming, A., & Power, M. (2021). Student 'voice' and higher education assessment: is it all about the money? *Journal of University Teaching & Learning Practice*, *18*(1), 1-16. https://ro.uow.edu.au/jutlp/vol18/iss1/6/

Inamorato, A., Gaušas, S., Mackevičiūtė, R., Jotautytė, A., & Martinaitis, Ž. (2019). *Innovating professional development in higher education: Case studies*. http://doi.org/10.2760/712385

Khosravi, H., Cooper, K., & Kitto, K. (2017). RiPLE: Recommendation in peer-learning environments based on knowledge gaps and interests. *Journal of Educational Data Mining*, *9*(1), 42-67. https://doi.org/10.5281/zenodo.3554627

Knight, D. B., Brozina, C., & Novoselich, B. (2016). An investigation of first-year engineering student and instructor perspectives of learning analytics approaches. *Journal of Learning Analytics*, *3*(3), 215-238. http://dx.doi.org/10.18608/jla.2016.33.11

Kotsiantis, S., Pierrakeas, C., & Pintelas, P. (2010). Predicting students' performance in distance learning using machine learning techniques. *Applied Artificial Intelligence*, *18*(5), 411-426. https://doi.org/10.1080/08839510490442058

Lawrence, J., Brown, A., Redmond, P., & Basson, M. (2019). Engaging the disengaged: Exploring the use of course-specific learning analytics and nudging to enhance online student engagement. *Student Success*, *10*(2), 47-59. http://doi.org/10.5204/ssj.v10i2.1295

Leitner, P., Khalil, M., & Ebner, M. (2017). Learning analytics in higher education—a literature review. In A. Peña-Ayala (Ed.), *Learning Analytics: Fundaments, Applications, and Trends: A View of the Current State of the Art to Enhance e-Learning* (pp. 1-23). Springer. https://doi.org/10.1007/978-3-319-52977-6_1

Lim, L.-A., Gentili, S., Pardo, A., Kovanović, V., Whitelock-Wainwright, A., Gašević, D., & Dawson, S. (2019). What changes, and for whom? A study of the impact of learning analytics-based process feedback in a large course. *Learning and Instruction*, *2019*, 101202. https://doi.org/10.1016/j.learninstruc.2019.04.003

Liu, D. Y.-T., Bartimote-Aufflick, K., Pardo, A., & Bridgeman, A. J. (2017). Data-driven personalization of student learning support in higher education. In A. Peña-Ayala (Ed.), *Learning Analytics: Fundaments, Applications, and Trends: A View of the Current State of the Art to Enhance e-Learning* (pp. 143-170). Springer.

Marbouti, F., Diefes-Dux, H. A., & Madhavan, K. (2016). Models for early prediction of at-risk students in a course using standards-based grading. *Computers & Education*, *103*, 1-15. https://doi.org/10.1016/j.compedu.2016.09.005

McCarthy, J. (2015). Evaluating written, audio and video feedback in higher education summative assessment tasks. *Issues in Educational Research*, *25*(2), 153-169. http://www.iier.org.au/iier25/mccarthy.html

McLaughlin, J. A., & Jordan, G. B. (2015). Using logic models. In *Handbook of practical program evaluation* (Fourth edition. ed., pp. 62-87). John Wiley & Sons.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093-1113. https://doi.org/10.1016/j.asej.2014.04.011

Mirriahi, N., Jovanovic, J., Dawson, S., Gasevic, D., & Pardo, A. (2018). Identifying engagement patterns with video annotation activities: A case study in professional development. *Australasian Journal of Educational Technology*, *34*(1), 57-72. https://doi.org/10.14742/ajet.3207

Murakami, A., Thompson, P., Hunston, S., & Vajn, D. (2017). 'What is this corpus about?': using topic modelling to explore a specialised corpus. *Corpora*, *12*(2), 243-277. https://doi.org/10.3366/cor.2017.0118

Pardo, A., Bartimote, K., Shum, S. B., Dawson, S., Gao, J., Gašević, D., Leichtweis, S., Liu, D., Martínez-Maldonado, R., & Mirriahi, N. (2018). Ontask: Delivering data-informed, personalized learning support actions. *Journal of Learning Analytics*, *5*(3), 235–249-235–249. http://dx.doi.org/10.18608/jla.2018.53.15

Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., & Mirriahi, N. (2019). Using learning analytics to scale the provision of personalised feedback. *British Journal of Educational Technology*, *50*(1), 128-138. https://doi.org/10.1111/bjet.12592

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

Patfield, S., Gore, J., & Weaver, N. (2021). On 'being first': the case for first-generation status in Australian higher education equity policy. *The Australian Educational Researcher*. https://doi.org/10.1007/s13384-020-00428-2

Peña-Ayala, A., Cárdenas-Robledo, L. A., & Sossa, H. (2017). A landscape of learning analytics: An exercise to highlight the nature of an emergent field. In A. Peña-Ayala (Ed.), *Learning Analytics: Fundaments, Applications, and Trends: A View of the Current State of the Art to Enhance e-Learning* (pp. 65-112). Springer. https://doi.org/10.1007/978-3-319-52977-6_1

Pistilli, M. D., Willis, J. E., & Campbell, J. P. (2014). Analytics through an institutional lens: Definition, theory, design, and impact. In J. A. Larusson & B. White (Eds.), *Learning analytics: From research to practice* (pp. 39-60). Springer.

Prusty, G. B., Russell, C., Ford, R., Ben-Naim, D., Ho, S., Vrcelj, Z., Marcus, N., McCarthy, T., Goldfinch, T., Ojeda, R., Gardner, A., Molyneaux, T., & Hadgraft, R. (2011). *Adaptive tutorials to target threshold concepts in mechanics-a community of practice approach.* 22nd Australasian Association for Engineering Education Annual Conference, Fremantle, Western Australia. https://aaee.net.au/wp-content/uploads/2018/10/AAEE2011-Prusty_Russell_Ford_et_al-Adaptive_tutorials_to_target_Threshold_Concepts_in_Mechanics.pdf

Ragupathi, K., & Lee, A. (2020). Beyond fairness and consistency in grading: The role of rubrics in higher education. In C. S. Sanger & N. W. Gleason (Eds.), *Diversity and inclusion in global higher education* (pp. 73-95). Palgrave Macmillan.

Small, L., McPhail, R., & Shaw, A. (2021). Graduate employability: the higher education landscape in Australia. *Higher Education Research & Development*, 1-15. https://doi.org/10.1080/07294360.2021.1877623

Smith, C. (2008). Building effectiveness in teaching through targeted evaluation and response: connecting evaluation to teaching improvement in higher education. *Assessment & Evaluation in Higher Education*, *33*(5), 517-533. https://doi.org/10.1080/02602930701698942

Spector, J. M., & Yuen, A. (2016). *Educational technology program and project evaluation*. Routledge.

Tasopoulou, K., & Tsiotras, G. (2017). Benchmarking towards excellence in higher education. *Benchmarking: An International Journal*, *24*(3), 617-634. https://doi.org/10.1108/BIJ-03-2016-0036

Twigg, C. (2003). *Improving learning and reducing costs: Lessons Learned from Round I of the Pew Grant Program in Course Redesign*. https://www.thencat.org/PCR/Rd1Lessons.pdf

United States Census Bureau. (2019). *What is the National Survey of College Graduates?* Retrieved February 15 from https://www.census.gov/programs-surveys/nscg/about.html

Vie, J.-J., Popineau, F., Bruillard, É., & Bourda, Y. (2017). A review of recent advances in adaptive assessment. In A. Peña-Ayala (Ed.), *Learning Analytics: Fundaments, Applications, and Trends: A View of the Current State of the Art to Enhance e-Learning* (pp. 113-142). Springer. https://doi.org/10.1007/978-3-319-52977-6_1

West, D., Luzeckyj, A., Searle, B., Toohey, D., Vanderlelie, J., & Bell, K. R. (2020). Perspectives from the stakeholder: Students' views regarding learning analytics and data collection. *Australasian Journal of Educational Technology*, *36*(6). https://doi.org/10.14742/ajet.5957

Whiteley, S. (2016). Creating a coherent performance indicator framework for the higher education student lifecycle in Australia. In R. M. O. Pritchard, A. Pausits, & J. Williams (Eds.), *Positioning higher education institutions* (pp. 143-160). Brill Sense.