MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

DATA ENABLED STUDENT SUPPORT

Ethical use of student analytics–interim guidelines

BACKGROUND

For some time, Massey University has collected and analysed student data to support student wellbeing and achievement. The use of data analytics to meet student retention and completion targets [1] is now strongly encouraged by the Tertiary Education Commission (TEC) and is becoming increasingly necessary to meet pastoral care outcomes and performance required by the Ministry of Education. [2]

As our approach to proactive student support expands, we need to further develop our relevant policy and principles to establish a clear university-wide framework for the ethical use of student data.

The development of this framework needs to be undertaken in partnership with Māori and in consultation with stakeholders including Pacific, students, international students, disabled persons, and staff and should be undertaken as a priority.

In the interim, principles for the use of student analytics are needed now to ensure the Data-Enabled Student Support (DESS) programme initiative (DESS programme) can be undertaken in a controlled, safe, and ethical way taking into account Māori data governance protocols in terms of access and analysis.

[1] As set out in Massey's investment plan with TEC.
[2] The Education (Pastoral Care of Tertiary and International Learners) Code of Practice 2021.

PURPOSE

The purpose of this document is to outline those interim ethical considerations and to ensure that an appropriate amount of transparency and accountability is in place for the analysis and use of data related to students' wellbeing, achievement and learning experiences at Massey University via the DESS programme. Given these guidelines are interim, all references to the use of student analytics in this document for current purposes is therefore confined to use within the scope of that programme.

RELATED POLICIES

These interim guidelines have been developed with reference to and in support of the following policies held within the Massey University Policy Library:

- Information Security Classification Policy and Framework
- Data Management Policy
- Information and Records Management Policy
- Policy on Staff Conduct
- Privacy Policy
- Privacy Statement

ACKNOWLEDGEMENTS

These guidelines acknowledge Māori Data Sovereignty and that as taonga, Māori data is subject to Māori Data Governance.

These guidelines also acknowledge the need to protect an individual's interests covered by the Privacy Act 2020, the Bill of Rights 1990, the Human Rights Act 1993 and as applicable the European Union (EU) General Data Protection Regulations (GDPR) or as applicable other global privacy legislation New Zealand becomes subject to over time. These guidelines are therefore intended as an interim measure to guide the ethical use of student data until such time as higher-level governance and sovereignty decisions have been created in partnership, appropriately consulted, and implemented across the university.

These interim guidelines draw extensively on guidance provided by TEC and the Policy on Ethical use of Student Data for Learning Analytics for The Open University but within the legislative context applicable to New Zealand as noted above.

DEFINITIONS

Data vs analytics: data in these guidelines refers to qualitative and quantitative information that is collected; analytics refers to the careful and complete analysis of data using a model, usually performed by a computer

Institutional analytics: the analysis of aggregated data to help monitor institutional success

Student analytics: an overarching term for both learning analytics and learner analytics

Learning analytics: a subset of student analytics which refers specifically to the measurement, collection, aggregated analysis, and reporting of data about learner outcomes for the purpose of understanding and optimising learning and the environments in which it occurs. Learning analytics use data from multiple sources to analyse student learning behaviours and outcomes and provides opportunity where required to improve course and programme structure and design and teaching approaches, leading to more effective teaching and learning experiences

Learner analytics: a subset of student analytics which refers specifically to information about an individual student or group of people engaged in activities where learning is an output. Learner analytics involves using student data to better understand the pastoral and education needs of students. Student data includes previous education records, course/ programme characteristics, age, gender, ethnicity, nationality, course attendance, level of engagement with the university and its programmes, outstanding fees status, enrolment status and academic results

Māori Data Governance refers to the principles, structures, accountability mechanisms, legal instruments, and policies through which Māori exercise control over Māori data – Te Mana Raraunga

Māori Data Sovereignty: refers to the inherent rights and interests that Māori have in relation to the collection, ownership, and application of Māori data – Te Mana Raraunga, Māori Data Sovereignty Network

Predictive analytics: is the use of current and historical data patterns, statistical algorithms, and machine learning techniques to determine the likelihood of performance and future outcomes

Interventions: an intervention derived from learning analytics that may be used to inform changes to teaching and learning design. An intervention derived from learner analytics may include the provision of support and advice by university staff to one or more students. Interventions (other than those resulting from predictive analytics) may also impact eligibility to enrol at the university or in a course/ programme of study.

OVERSIGHT

Initiatives related to the more systemic use of student analytics stems from the work of Pūrehuroatanga, 'a coordinated and centralised body of work which aims to remove institutional barriers to success as well as providing targeted, and proactive and data-driven support for those learners who need it.' The intent is to use:

- Learning and institutional analytics at a macro level to aggregate information about the student learning experience at an institutional level, and to inform strategic priorities that will improve student retention and progression. This aggregation includes the deidentification of the data;
- Learner analytics at a micro level to drive or support learner specific short, medium, and long-term interventions.

THE PRINCIPLES

The principles are split into seven categories and provide guidance on not just the act of intervening based on student analytics but also on information collection, communication and intended usage. A summary of the principles is provided below followed by further explanation and rational for each principle.

Summary

Strategic use

Principle 1 – good governance practices are used to ensure student analytics are conducted according to ethical consideration and align with core university values, strategy, policy, and responsibilities.
Principle 2 - the use of student analytics should be purposeful.

Transparency

Principle 3 - we will be transparent about the purpose of the analytics, and how the data is collected, used, and shared, including its retention.

Lawful Purpose/Consent

Principle 4 - the authority to use learner and learning analytics is based on lawful purpose/legitimate interest (in context of EU GDPR) grounds and informed consent for any specific categories of sensitive data requiring that.

Communication and Engagement

Principle 5 – students are well informed about the university's approach to student analytics.
Principle 6 - stakeholders, including staff, students, and technology providers (where appropriate) should be involved in designing, developing, and reviewing the use of analytics and its governance.

Ethical and Appropriate Use

Principle 7 – students are real and diverse individuals rather than data or information.
Principle 8 – those producing and using student data for analytics are adequately trained, experienced, and supervised.
Principle 9 – use of student analytics will be systematic and coordinated.

Reliability and Validity of Data

Principle 10 - the reliability and validity of data is essential to the effectiveness of learning analytics.

Māori Cultural Values

Principle 11 – respecting and upholding the mana and dignity of the people who share their data.

THE PRINCIPLES EXPLAINED

Strategic use

Principle 1 – good governance practices are used to ensure student analytics are conducted according to ethical consideration and align with core university values, strategy, policy, and responsibilities.

The university is committed to proactively supporting students as part of achieving its broader strategic goals and compliance requirements.

This principle requires us to understand how institutional strategy can combine with ethical considerations to enable more effective use to be made of data about individual students and their interactions with the university. Massey University's largely open entry pathways bring a greater responsibility for ensuring that students have the information, advice, guidance, and support needed to make appropriate choices and to complete their study successfully. Our approach to supporting learning and teaching is outlined in Paerangi our learning and teaching strategy and specific activities supporting student achievement are outlined in our University Learner Success Plans. Student analytics are an important tool for delivering on these commitments in a focused and impactful way.

The New Zealand Government has been clear that the support of student learning and wellbeing is a shared responsibility between government, tertiary education providers, learners, whānau and the wider community. The Government has also rightly identified the importance of learner wellbeing in enabling students to achieve their aspirations in education and beyond. Therefore, it is important that where data indicates that there is potential for action to be taken which might better support students in achieving their study goals or in reaching their potential, the university has a responsibility to act on this. For example, if there is evidence that a student is not engaging with essential learning activities, we should consider making an appropriate intervention.

This responsibility needs to be balanced against considerations for effective use of resources. It is likely that we will identify many more potentially helpful interventions than can be reasonably resourced. In the context of resource availability to deliver student support, priorities will need to be established for targeting particular curriculum areas and/or priority groups. Once priorities are established, consideration will be given to ways of improving student outcomes.

Principle 2 - the use of student analytics should be purposeful

The purposes of student analytics at Massey University is to identify aspects of a student's experience that may enable the university to match support services more closely to need and to understand how we may do this effectively. Student analytics also support the further development of our curriculum, policies and business processes including the delivery of advising and support services.

Any initiative developed that utilises student analytics should be clear and explicit about how it contributes to this purpose.

Transparency

Principle 3 - we will be transparent about the purpose of the analytics, and how the data is collected, used, and shared, including its retention.

Learner and learning data and the results of its analysis (analytics) will be used to shape students' potential learning journey including student wellbeing with the expressed aim of enhancing student achievement. Transparency includes making clear and accessible the data sources, the purposes of the analytics, the processes involved in producing the analytics, who has access to the analytics, the period of retention, and how to interpret the data (see also Principle 5 – 6 Communications and Engagement).

All data captured during the university's interaction with the student has the potential to provide evidence for student analytics. Data will, however, only be used for student analytics where there is likely to be an expected benefit (which will be evaluated) to students' achievement and where the university has a valid lawful basis for using the data for that purpose.

Lawful Purpose/Consent

Principle 4 - the authority to use learner and learning analytics is based on lawful purpose/legitimate interest (in context of EU GDPR) grounds and informed consent for any specific categories of sensitive data requiring that.

Students give their general consent during admission and enrolment to their data being used in accordance with the university's Privacy Statement . This statement notes that profiling, analytics, or tracking activities (including by automated means) will be undertaken for a number of listed purposes and has the additional section giving more specific information about the process and purposes for which their data may be used for DESS programme.

The privacy statement therefore enables the university to process most student data within the scope of the DEES initiative upon the basis of lawful basis/legitimate interest, but:

- an additional informed and express consent will still be required for including certain categories of sensitive data covered by the Privacy Health Information Code, the EU/UK GDPR and where applicable other global privacy legislation; and
- future enhancements to the DESS programme must be both covered by the listed purposes (i.e.; lawful purpose/legitimate interest grounds) disclosed in the privacy statement and meet the transparency, communication and engagement requirements prescribed in these guidelines (Principle 3, 5, 6).

Students should be given the option to withdraw their consent or opt-out of any profiling, analytics, or tracking activities (including by automated means) via the process set out in the Privacy Statement so the programme must always be designed to accommodate this. This follows advice from the Privacy Commission regarding analytics activity in general and is a requirement under the EU/GDPR regarding any sensitive data included with the informed and express consent of an individual.

It may not however be possible for the:

- student to opt-out of all learners and learning analytics activity (i.e.; if such activity is required under New Zealand Law); and/or
- for the university to delete all data as a result of a withdrawn consent or opt-out option notification, due to Public Records Act requirements to keep the data for minimum retention periods; and where this the case, the reasons must be clearly explained to the student.

Requests to students to participate in educational research associated with the DDES programme will follow existing university practices including required Ethics Committee approvals under the Code of Responsible Research Conduct.

Communication and Engagement

Principle 5 – students are well informed about the university's approach to student analytics.

The techniques used in student analytics are based on standard statistical methods, but can also involve the development of complex models, the full working of which will only be apparent to those familiar with the data and with the statistical methods employed. It is understood that users may want to know more about how the models are produced, the information being captured and the outcomes which they then deploy.

Care should be taken when communicating directly with students based on their analytics. Students may want to understand why they have been selected for an intervention and, in some cases, may want to challenge the basis for their selection.

It is important that there is a clear plan to communicate with students the purpose and approach in the use of their information in student analytics to ensure they feel confident that their data is used responsibly, using best practice.

Students will also be assured that data is being used to enhance support services and is specifically focused on improving successful student outcomes.

The Privacy Statement will provide layered levels of transparency depending on the information needs of the student as follows:

- Students are notified of the broad uses and purposes of profiling, analytics, or tracking activities (including by automated means) undertaken by the university;
- A link to these Interim Guidelines is provided for those wishing to access further information;
- A specific section related to the DDES programme is included; ;
- The DDES section includes a link to a high-level overview of the DESS end-to-end process;
- The end-to-end process overview provides a link through to FAQ's and a contact address for students wishing to access further information about the predictive analytics process deployed at the university; and
- The FAQ's provide a link to the methodology used for developing the predictive analytics model for the DESS programme including a table identifying the data sources used in the model.

Principle 6 - stakeholders, including staff, students, and technology providers (where appropriate) should be involved in designing, developing, and reviewing the use of analytics and its governance.

Wherever possible we will share our interpretation of data with students, in particular students will have access to the data they need to make informed decisions about their study. They will also be actively involved in helping the university to design and shape interventions that will support them. This not only aids in enhancing any informed consent requirements but also in improving the quality of data interpretation and our understanding of what forms of intervention and support are most appropriate. Both the university and students have a responsibility to ensure that student personal data is current and accurate.

Ethical and Appropriate Use

Principle 7 – students are real and diverse individuals rather than data or information.

Students should not be wholly defined by their visible data or our interpretation of that data. Analysis based on the characteristics of individual students at the start of their study must not be used to limit the university's or the students' expectations of what they can achieve.

We acknowledge that data and the algorithms created from it are not neutral, and thus commit to working to avoid creating or perpetuating inequity and no form of profiling or data analytics can be engaged with that would give rise to a claim of discrimination under the Bill of Rights Act 1990 or Human Rights Act 1993.

We understand the limitations of data. Hence:

- Students will not be defined only by their data
- Where predictive analytics are used, it is accepted that there will always be individuals whose learning behaviours do not follow the typical pattern. While it is recognised that some insights from analytics may be directed more at some students than others, we do not propose a deficit model targeted at only supporting students at risk of failure
- No decisions relating to individual student wellbeing, achievement and learning experiences will be taken solely based on learner data. The university recognises that while data is a tool to inform human decision-making, further human input and analysis is always required including review of the raw data where predictive analysis has been used
- Predictive analytics will not be used for determining a student's suitability, eligibility, or qualifications to be:
  - enrolled or progress through their course or qualification (including to post-graduate level);
  - included in a selected entry programme, oversee exchange programme, or articulation process;
  - excluded from their course or qualification once enrolled; or
  - given an award, scholarship, honour, or other benefit.

Principle 8 – those producing and using student data for analytics are adequately trained, experienced, and supervised.

Data used for student analytics typically falls into one of two categories: that captured at registration or at later points as a result of the student supplying information to the university and that are derived from ways in which the student engages with university systems as a result of their ongoing study.

Any interpretation or manipulation of data to extract meaning is to be based on sound technique and is subject to expert peer review and, where necessary, advice and mentoring by those more experienced in techniques of qualitative and quantitative data analysis. The DDES programme meets these requirements by aiming to use the most appropriate models and by ensuring that members of staff using the data or information are best placed to do so. The DESS model has also been peer reviewed upon creation and will continue to be reviewed if significant changes to the model are proposed.

Access to individual student data and analytics will be restricted to those identified as having a legitimate need to view them within the scope of their role at the university. Requests for access to the student data for the purpose of any analytical use outside the scope of the DESS must be approved by the Data Custodian in consultation with the steering group as necessary, and depending on the nature of the request will be assessed by any applicable governance processes such as ethics, sovereignty, privacy, and security.

The implementation of the DESS programme at Massey University requires broad acceptance of the values and benefits of a proactive, data driven approach to student support and the development of appropriate skills across the organisation. Skills in the collection and use of data including privacy obligations must be further developed as appropriate to the role. Some group members will need a higher level of technical expertise; others will need more focus on interpretative skills; and all staff will need adequate awareness and understanding of ethical considerations.

Principle 9 – use of student analytics will be systematic and coordinated.

Student analytics can be applied to individual students as well as to defined groups of students (as a result of identifying a student via combinations of characteristics and/or study behaviours), and to whole cohorts of students. These guidelines apply in all cases. To make these guidelines of practical use, interventions will be delivered within a broader framework where each intervention will have a written brief which clearly outlines:

- The institutional, wellbeing or student achievement issue being addressed
- The responsibilities of different staff
- The data being used
- The resulting report or evaluation.

Reliability and Validity of Data

Principle 10 - the reliability and validity of data is essential to the effectiveness of learning analytics.

People making decisions, as well as those impacted by decisions, expect data to be complete and accurate. To ensure that outputs from learning analytics are both reliable and valid, the university needs to ensure that data collected and analysed is both accurate and represents the issue being measured. Datasets should be kept as up to date as possible, with opportunities for the data to be refreshed to replace existing data to keep it relevant.

Predictive analytics involves statistical calculation; therefore, it is also important to ensure that data sets are complete and sufficient to enable robust calculations to be made. Further, the models used to analyse, interpret, and communicate learning analytics to stakeholders (academic staff, support staff, advisors, students, etc) should be sound, largely free from algorithmic bias; transparent and where possible clearly understood by end users.

Alongside analysis of data and the development of interventions, work will continue ensuring that data used for student analytics, including that drawn from the learning management system Stream, are reliable across the university.

Māori Cultural Values

Principle 11 – respecting and upholding the mana and dignity of the people who share their data.

Those using data for any analytical purpose must understand that Māori students, their whānau, hapū or iwi may be affected by the use of data analytics, and are therefore responsible for ensuring that the creation, interpretation, and use of those data sets are respectful and uphold the mana and dignity of those students who share their data.

Reliable data must be used in a cultural context to support Māori students to succeed. Data analysis that emphasises a deficit view of Māori results in individual and collective harm and should be actively avoided.

Māori have inherent rights and interests in relation to the collection, ownership, and application of Māori data therefore those performing data analytics will authentically engage with Māori representation and end users to inform appropriate practices.

CURRENT AND POTENTIAL SOURCES OF DATA

All information collected by the university is potentially available for the purposes of student analytics, providing its use is consistent with the guidelines and scope outlined.

Descriptive analytics are based on static data derived from several sources, such as course evaluations, student exit surveys, student information systems, Stream activity and Customer Relationship Management (CRM) interactions. The reporting may include both quantitative and qualitative data, and the nature of this data is post hoc and often summative.

Predictive analytics may extend data from the same sources but focuses on trying to measure actual learning. The data may come from assessments and assignments, log files, simulations designed to capture the learning process and/or direct observation.

Whereas descriptive and predictive analytics involve asking questions and receiving answers, data mining is an exploratory process of discovery. Researchers develop and apply algorithms and models to data sets about the learning environment to reveal previously unseen patterns of information.

The DESS programme uses descriptive analytics, predictive analytics, and data mining. The data that may be used by the DESS programme for this purpose, subject to further evaluation of some data sets expected benefits of inclusion and lawful basis for collection and use, includes:

- Student management system:
  - Pre-enrolment questionnaire
  - Admission questions
  - Enrolment questions
  - Progression data
  - Fee payment data
  - Career Readiness questions
  - Grades
- Library access
  - Headcount
  - Borrow count

- Learning management system
  - Mediasite
  - Stream Reports (Activity logs; assessment results)
  - eReserve
- Office 365
- Customer Relationship Management system
- Records of interactions between students and staff
- Attendance at events or workshops
- Geographical and location data such as WiFi logons, IP addresses, access cards
- RAPID data warehouse
- Akari (the curriculum management)
- Student Surveys and Evaluations.

IN SCOPE

The current scope of the DESS programme includes data captured by Massey University as part of its interaction with students and as individual or combined data sets for use in student analytics. With permission from the student, this includes:

- personal information provided at the time of admission and enrolment including ethnic origin of students from jurisdictions where this information usage is not subject to express consent requirements or has otherwise been obtained during or after enrolment; student information collected during their continued education at the university such as a student's study record held by the university including information on assignments, assessments, exams, and quizzes;
- sensitive information collected during their continued education that the university has a valid lawful purpose to use (e.g.; outstanding fee status for students from jurisdictions where this information usage is not subject to express consent requirements);
- details of contacts/interactions between the student and the university; for example, the library gathers information generated by student use of a subscription service
- interactive content generated by students; for example: completing diagnostic tests, student responses to surveys and research
- system-generated data such as the date and frequency of accessing the student learning management systems (Stream)
- clearly understood data derived by the university from other data, for instance, whether a student falls into a priority learner category
- data from the categories above that is held or generated internally in combination with data provided by third parties may be used by the university to tailor support, where:
  - there is agreement to do so from the third party concerned; and
  - the university can clearly identify where the data came from (provenance); how it has transformed e.g., data lineage; understands how the information has been collected and the purpose of collection
  - the collection of such data from the third party rather than the individual to whom the information relates, satisfies an exception to Principle 2 of the Privacy Act 2020.
- miscellaneous sources of data specifically approved by the Steering Group, following evaluation of its expected benefits of inclusion and lawful basis for collection and use.

OUT OF SCOPE

The university does not intend to use the following types of data in its analytical models for the DESS programme. This list is subject to review:

- Data on student complaints
- Data that identifies individuals created on external sites, e.g.; social networking sites not owned by the university, third party sites where there is no permission to employ shared information, etc
- Data relating to prospects, enquirers, and informal learners rather than enrolled (or admitted) students or alumni
- Sensitive information related to ethnicity (if express consent is required but not obtained), religious belief, disability and other health matters, sexual life, political opinions, offences (including alleged offences), criminal proceedings, outcomes and sentences and membership of trade unions.[3] Should any of these sensitive data items become required for learning analytics (whether by legislative requirement or as a result of expected benefit evaluation), informed consent will have to be obtained by a suitable means, or any alternative lawful basis relied upon for its' inclusion will need to be validated. Any combinations of data or derived data that may contravene an individual's right to respect for their private and family life will not be used.
- Analytics used for academic research purposes fall outside these interim guidelines and are subject to other policies such as the Code of Responsible Research Conduct.

[3]Also, being Special Category data under the EU GDPR that would require informed and express consent for such use.

POLICY DEVELOPMENT

A policy that provides governance to all areas of student analytics must be formed in partnership with Māori and in consultation with stakeholders including Pasifika, students, international students, disabled persons, consider issues of data sovereignty, privacy, ethics, and human rights considerations and be informed by a review of existing practice in other higher education institutions and research literature.

The policy will govern the ongoing application and further development of these interim guidelines for the ethical use of data analytics for the Data-Enabled Student Support (DESS) Programme.