Policy on Ethical use of Student Data for Learning Analytics

1 Introduction

1.1 Rationale

1.1.1 The Open University has collected and analysed student data as a means of

providing information relating to student support and retention for many years. The changing landscape of higher education has seen the rapid expansion of uses to which student data is put. The Open University, in common with many other higher education institutions, is now looking at its use of **learning analytics**.

1.1.2 In the context of the Open University, learning analytics is the use of raw and analysed student data to proactively identify interventions which aim to support students in achieving their study goals. Such interventions may be designed to support individual students and/or the entire cohort. This is in line with the University's vision of "…delivering a step-change in how effectively we help students achieve their study goals" and to deliver "…a study experience that maximizes students' chances of success in achieving their study goals whilst maintaining academic standards."[1]

1.1.3 The use of a learning analytics approach to drive student support within the University is relatively new. There is a need to establish guiding principles which help provide a clear framework for the ethical application of learning analytics.

1.2 Problem statement

1.2.1 All data captured as a result of the University's interaction with the student has the potential to provide evidence for learning analytics. Data will, however, only be used for learning analytics where there is likely to be an expected benefit (which will be evaluated) to students' learning.

1.2.2 The techniques used in learning analytics are based on standard statistical methods, but typically 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 likely, however, that users will want to understand how the models produce the outcomes which they then deploy. Students will want to understand why they have been selected for an intervention and, in some cases, may want to challenge the basis for their selection. A potential conflict exists therefore between creating models which provide the most reliable outcomes and those which work in ways that can be made transparent to users and subjects. The University has agreed a communication strategy to explain how individual subjects are affected by specific applications of learning analytics, in order to off-set this complexity.

1.2.3 In order to make the new policy of practical use, staff will need support to facilitate their understanding of how the policy applies in practice. This will be provided separately in the form of guidance notes aimed at different staff roles, providing practical examples of the uses to which data may be put and the ethical

---

[1] The Strategic Plan 2012-2015
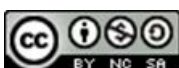
considerations. Access to student data is determined by staff permission levels and most staff will have permissions to access a wide range of student data. This policy does not in itself attempt to define how different categories of staff will engage with data and assumes that existing policy and guidance applies (for example, the Data Protection Policy states that "All personal information is held in secure computer and manual files, and we will only transfer data within the University on a 'need-to-know' basis so that we can support our academic and other services to you.")

1.2.4   Learning 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 (as a result of amending the assessment regime on a module following observed behaviours and/or results, for example). The policy and principles created apply in all cases.

1.2.5   The use of student data in new ways will necessarily impact on current practice. As such, a number of existing policies relating and referring to potential uses of student data require review.

1.2.6   The Policy is informed by significant consultation with key stakeholders and review of existing practice in other higher education institutions and detailed in the literature. A review of literature of external practice suggests that policy development in response to new uses of student data is relatively scarce. As a result, any new policy developed within the University has few exemplars and should be subject to review and amendment in 12 months' time.

1.2.7   The success of any learning analytics approach is dependent on a reasonably complete and up to date dataset. As part of the stakeholder consultation, it has been established that there should be a responsibility on the student to ensure that their own held data reflects current circumstances. This is detailed in Principle 4 of the Student Charter[2]. Further guidance will be provided to students to ensure that this is well understood. Both the University and students have a responsibility to ensure that student personal data is current and accurate.

1.2.8   Any use made by the University of data regarding individual students must be compliant with the Data Protection Act. More specifically, data must be handled (processed) in accordance with the data protection principles. In order to assure best practice and to reduce the risk of non-compliance with the Data Protection Act, a formal Data Protection Assessment has been conducted. The risks identified and the resulting actions required to eliminate or mitigate these risks have been recorded with the University's Data Protection Coordinator.

1.3   Definitions

The following definitions are intended to provide clarity about terms used throughout the Policy.

a)   **Learning analytics** has been defined as the collection and analysis of data generated during the learning process in order to improve the quality of learning and teaching[3].

b)   The term **student** refers to individuals registered to study on a module or qualification. This does not include enquirers or informal learners.

c)   The term **cohort** may refer to, for example, all students linked to a qualification or students registered on a specific module-presentation.

---

2 The Student Charter

3 Siemens, Dawson, & Lynch, 2013

d) An **intervention** derived from a learning analytics approach may refer to information, advice and guidance (IAG) directed from University staff to one or more students. Learning analytics at a module or qualification level may be used to inform changes to teaching and learning design.

e) **Data** used for learning 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 (typically labelled as *Student characteristic data*), and that derived from ways in which the student engages with University systems as a result of their ongoing study (typically summarised as *Study behaviour data*). Access to this data is governed by existing policies such as the Data Protection Policy.

f) **Sensitive data**: The Data Protection Act defines sensitive data and generally requires an individual's permission to collect and use sensitive data for a specified purpose. The items of sensitive data that the University currently collects are:

- racial or ethnic origin
- religious or similar beliefs
- disability and other health matters
- sexual life.

The items of sensitive data that the University does not currently collect are:

- political opinions
- offences (including alleged offences)
- criminal proceedings, outcomes and sentences
- membership of trade unions.

g) In the context of this policy, **informed consent** refers to the process whereby the student is made aware of the purposes to which some or all of their data may be used for learning analytics and provides consent. Informed consent applies at the point of reservation or registration on to a module or qualification. Requests to students to participate in educational research will follow existing university practices including applying to the Student Research Project Panel (SRPP)[4] and Human Research Ethics Committee (HREC)[5].

2   Scope

The OU's approach of using learning analytics stems from the vision for the Analytics for Student Success programme "*To use and apply information strategically (through specified indicators) to retain students and progress them to complete their study goals*". This is supported by the mission statement, which aims at achieving this:

- at a **macro** level to aggregate information about the student learning experience at an institutional level to inform strategic priorities that will improve student retention and progression
- at a **micro** level to use analytics to drive short, medium and long-term *interventions*.

[4] Student Research Project Panel
[5] Human Research Ethics Committee

In the future, the use of learning analytics may be extended to personalised learning paths, adaptive learning, personalised feedback, visualisations of study journey, intelligent e-tutoring, intelligent peer support, etc. Furthermore, new technological innovations might allow for more targeted, measured approaches.

2.1   In scope

2.1.1   Categories of data captured by the University as part of its interaction with students and potentially available as individual or combined data sets for use in learning analytics:

- personal information provided by the student at registration

- the student's study record held by the University

- sensitive information that the University has consent to use. Currently consent for the use of ethnic origin and disability to identify students who may be in need of additional support is obtained at registration when students agree to the Data Protection Policy

- details of contacts between the enquirer or student and the University

- interactive content generated by enquirers or students; for example: completing diagnostic tests, student responses to surveys and research (subject to existing SRPP and HREC restrictions and approvals), etc.

- system-generated data such as the date and frequency of accessing VLE pages;

- data derived by the University from other data, for instance, whether a student falls into a widening participation category

- data 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. For example, the Library gathers information generated by student use of a subscription service. Student data supplied to third parties is subject to existing policies such as the Data Protection Policy

- anonymised data from external sites, e.g. social networking sites not owned by the University, where this is used to generate information on the cohort rather than the individual student. For example, where this forms part of an activity within a module

- miscellaneous sources of data, for example, forum posts could be anonymised and analysed to shape module design.

2.1.2   A set of guidance notes will be produced for students, Associate Lecturers and other staff setting out the purpose and boundaries of the University's learning analytics approach. These guidance notes may be updated as the scope changes in the future. The notes will include, where feasible, examples of ethical and unethical practice in order to make the principles on which the policy is built more meaningful. The guidance notes will be used as part of planned communications to stakeholder groups.

2.2   Out of scope

2.2.1   In its adoption of a learning analytics approach to provide student support, the University does not intend to use the following types of data. 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 enquirers and informal learners rather than registered (or reserved) students or alumni.

- Sensitive information on religious belief and sexual life will not be used as part of the analytical models. Should any other sensitive data items be required for learning analytics, consent will have to be obtained by a suitable means, such as through changes to the Data Protection Policy. 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.

2.2.2   There is a large number of related existing University policies which consider the use and storage of student data. These have been considered as part of the development of this policy. The final version of the policy will be provided to other relevant policy owners requesting them to review and amend their own policies as required.

2.2.3   Information relating to individual Associate Lecturers, except in so far as tutor details may be linked to identified students (to alert specific tutors of issues, or as part of a tutor's own use of the Student Support Tool).

2.3   Ethical issues relating to the use of student data for academic research

2.3.1   Applications to use student data for the purposes of research will need to be made in accordance with the standard processes in place currently, such as Student Research Project Panel (SRPP) and the Human Research Ethics Committee (HREC). Bodies considering applications for research using learning analytics should assess if the projects comply with this policy. The bodies, within the remit of their own terms of reference, may approve research proposals that test the boundaries of this policy. If the outcomes of that research may then be applied to operationally targeting individuals or groups of students, further alignment with this policy will be required.

3   Oversight on the ethical use of data for learning analytics

Consideration needs to be given to a formal process or owner for oversight on the ethical use of data for learning analytics. The policy owner will need to appoint a named contact and set up a formal approval process to ensure compliance with this policy.

4 Policy statement

This policy aims to set out how the University will use student data in an ethical way in order to shape the student support provided. The document and accompanying guidelines are not regulatory in nature but are intended to inform and guide the ethical use of student data.

The policy is based around eight key principles discussed in more detail in the policy statement below.

- **Principle 1**: Learning analytics is an ethical practice that should align with core organisational principles, such as open entry to undergraduate level study.
- **Principle 2**: The OU has a responsibility to all stakeholders to use and extract meaning from student data for the benefit of students where feasible.
- **Principle 3**: Students should not be wholly defined by their visible data or our interpretation of that data.
- **Principle 4**: The purpose and the boundaries regarding the use of learning analytics should be well defined and visible.
- **Principle 5**: The University is transparent regarding data collection, and will provide students with the opportunity to update their own data and consent agreements at regular intervals.
- **Principle 6**: Students should be engaged as active agents in the implementation of learning analytics (e.g. informed consent, personalised learning paths, interventions).
- **Principle 7**: Modelling and interventions based on analysis of data should be sound and free from bias.
- **Principle 8**: Adoption of learning analytics within the OU requires broad acceptance of the values and benefits (organisational culture) and the development of appropriate skills across the organisation.

Each of the above principles is linked to particular aspects of learning analytics.

4.1 Aligning the use of student data to core University values

Principles 1 and 2 aim to reflect key tenets that the University should adopt a responsible approach to using learning analytics to enhance the chances of student success. At the same time, the University has a commitment to proactively support particular cohorts in line with broader strategic values and the approach taken should complement that. These principles support the Student Charter which states that "We treat each other with dignity and respect: As members of a welcoming and inclusive community we … respect confidentiality and only use information we have been given for the legitimate purposes that we have been given it."[6]

4.1.1 Principle 1: Learning analytics is an ethical practice that should align with core organisational principles, such as open entry to undergraduate level study.

- Interpretation of data derived from learning analytics must be carried out in accordance with principles for ethical practice, which include the need for clarity

---

[6] The Student Charter

about the purpose of using such data and for engagement of stakeholders, of which students are the most significant group.

- This principle enables us to understand how core institutional values can combine with ethical considerations to enable more effective use to be made of data about individual students and their interactions with the University. The University's open entry mission is a defining organisational principle which ensures a vibrant and diverse student community. Learning analytics provides us with the ability to understand and respond in a more targeted way to diverse student needs based on the data we collect as a part of our interaction with the student.

- As a University with an open entry admissions policy, we have a higher level of responsibility to ensure that students have the information, advice, guidance and support needed to make appropriate choices and to complete their study successfully. This is reflected in our stated vision of "…delivering a step-change in how effectively we help students achieve their study goals" and to deliver "…a study experience that maximizes students' chances of success in achieving their study goals whilst maintaining academic standards." (The Strategic Plan 2012-2015). In addition, the University has a defined approach to providing support for particular groups in line with the Widening Access and Success Strategy. Learning analytics offers the potential to enable more effective support of these groups.

- Transparency (T), responsibility (R) and effectiveness (E) are key features of our approach to ethical practice and are referred to where relevant below.

4.1.2 **Principle 2: The OU has a responsibility to all stakeholders to use and extract meaning from student data for the benefit of students where feasible.**

- 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 consideration of 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 available to deliver student support, priorities will need to be established – in terms of targeting particular curriculum areas (key modules, etc) and/or priority groups (for example, WASS or potentially vulnerable groups). Once priorities are established, consideration may be given to ways of achieving maximum impact.

4.2 Purposes and boundaries

Principles 3 and 4 make clearer why the University has adopted learning analytics as one of many means of providing effective and targeted student support whilst recognising that students, as real and diverse individuals, rather than data or information, drive appropriate student support.

4.2.1 **Principle 3: 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.

- Learning analytics will generate data and insights which enable us to provide targeted and specific support to student groups with shared characteristics and/or behaviours. For example, these may be students who fall within our widening access priority groups.

- Predictive analytics reflect what has happened in the past, not the future. In their calculation of error rates, it is accepted that there will always be individuals whose behaviours do not follow the typical pattern.

- We should guard against stereotyping. Students who do not fall within any priority group may encounter difficulties during their study which become apparent as a result of learning analytics data, and subsequently benefit from targeted interventions.

- Caution needs to be exercised in the interpretation of data for a variety of reasons and guidance provided to staff will aim to support this. For example, individual members of staff may not have access to the full data set that is available to the University and may have an incomplete view of the student and their experience.

- The primary purpose of our approach to learning analytics is to support students in achieving their study goals: thus consideration can be given to forms of additional communication or support offered in line with our 'open' mission to support the diversity of student need. Services and support that are available potentially to all students will continue to be provided.

4.2.2 **Principle 4: The purpose and the boundaries regarding the use of learning analytics should be well defined and visible.**

- The primary purpose of learning analytics activity must be to identify ways of effectively supporting students to achieve their declared study goals.

- The purposes of learning analytics at the Open University are to:

  a) identify aspects of a student's record and engagement with learning activities which may enable us to match services more closely to need and to understand how we may do this effectively (E)

  b) support the further development of our curriculum, policies and business processes including the delivery of Information, Advice and Guidance services to enquirers and students (E)

  c) improve the likelihood that a student will achieve his or her stated learning outcomes. (R/E)

4.3   Engaging students in the use of their data

Principles 5 and 6 reflect the shared responsibility of both the student and the University for student learning.

4.3.1 **Principle 5: The University is transparent regarding data collection, and will provide students with the opportunity to update their own data and consent agreements at regular intervals.**

- Accurate and representative student data should not be the sole responsibility of the University. Students also have a duty to accurately maintain their personal data and to inform the University of any changes which might impact on their studies. The University must provide students with the opportunity to periodically update their records.

- This principle highlights the importance of a clear plan to communicate with students about our approach to learning analytics:

  a) in order that students understand our approach and feel reassured that data is used responsibly (T)

  b) where we can share our interpretation of data with students we will do so unless there are good pedagogical reasons to do otherwise. It may, for instance, become good practice to highlight for students those avoidable factors such as late registration which analytics identify as increasing the likelihood that a student will fail to complete a module (T/R)

  c) to support students in making informed decisions about their studies (R/E)

  d) to reassure that this is about enhancing support services in order to improve qualification completion which will benefit all students of the University. (T/E)

- Data may not be reliable, nor reflect a student's current status, if it is not up to date. Students must have opportunities to check and update their personal data, including their consent to its use, at clearly defined intervals.

4.3.2 **Principle 6: Students should be engaged as active agents in the implementation of learning analytics (e.g. informed consent, personalised learning paths, interventions).**

- Students should be actively involved in helping the University to design and shape interventions that will support them.

- It is essential to engage students with the development of our approach to learning analytics, because this is likely to enable us to:

  a) ensure that students understand their responsibility for keeping personal information up to date and can give informed consent (R/E)

  b) achieve a more accurate interpretation of data relating to student behaviours (E)

  c) improve our understanding of what forms of intervention and support are most appropriate (E)

  d) know how to communicate with students in general and individually about our approach (T/E/R)

  e) understand how to tailor a student's learning journey to meet their needs, potentially as a personalised learning path (R/E)

  f) produce outcomes that students will find useful and be able to respond positively to; this might include a decision to continue or discontinue with their studies. (E)

4.4 Ensuring that data is used wisely

The final principles which support the policy relate to the need to ensure that any interpretation or manipulation of data to extract meaning is based on sound technique which is subject to expert peer review and, if necessary, through advice and mentoring by those more experienced in techniques of quantitative data analysis. The principles address this 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.

4.4.1 **Principle 7: Modelling and interventions based on analysis of data should be sound and free from bias.**

- A set of best practise principles must be established for the development, maintenance, interpretation and review of the statistical models used for learning analytics to enable periodic internal audits.

- All information collected by the University is potentially available for the purposes of learning analytics, providing its use is consistent with these principles. Information selected for use in learning analytics should demonstrate its value in delivering the agreed outcomes. Analytical models should aim to be transparent, such that their method of working can be described to staff and students, they are based on standard statistical techniques, and they can be tested and audited to provide assurances that they use data which meet quality criteria, correctly apply that data and produce results that reach an agreed level of accuracy.

4.4.2 **Principle 8: Adoption of learning analytics within the OU requires broad acceptance of the values and benefits (organisational culture) and the development of appropriate skills across the organisation.**

- Regular communications to staff will ensure that they understand the approach: the underlying values linked to the University's mission and strategy; the anticipated benefits for students; the limitations of data and its interpretation; and guidelines for ethical practice.

- Skills in the collection and use of data must to be further developed as appropriate to role. Some groups will need a higher level of technical expertise; others will need more focus on interpretative skills; and all staff will need some basic awareness and understanding of ethical considerations.

5 Implementation

The final timeline for implementation is dependent on ensuring that key stakeholders are informed of the need to ensure appropriate use of student data. A communications strategy has been developed which considers each of the key stakeholder groups, focusing on specific needs and responsibilities.

In addition to the creation of new policy, a number of systems changes are likely to be recommended. These are not within the scope of the policy *per se*, but have emerged as a result of consultation. It is hoped that any systems changes will be relatively minor and should not prevent the adoption of the new policy.

6 Schedule

This Policy was approved at LTSSC on 16th July 2014. It was implemented with regard to the use of the learning analytics approach to providing student support from October 2014,

subject to the approval of a clear communications plan and any required systems developments.

7   Review

Aspects of this policy will be reviewed every two years with a more extensive review to align with wider Teaching and Learning Policy reviews, to ensure fitness for purpose; any significant changes will be taken through the appropriate governance structure.

8   Executive owner

PVC - Learning and Teaching