Use of Analytics in South African Higher Education
Saide Working Paper 2024

The Siyaphumelela 3.0 initiative, funded by the Kresge Foundation, supports 20 South African universities to investigate the use of student data analytics to improve student success. In order to support this initiative *Saide* initiated a project to explore the ethical use of data within a South African context. This included the development of a draft narrative framework for the ethical use of data authored by Professor Paul Prinsloo (Unisa)[1].

This framework includes seven guiding principles that support the ethical use of data to support student success:

1. **The moral relational duty of learning analytics**:
   Learning analytics should be viewed as relational and emancipatory, rather than a process that perpetuates injustice and inequality.

2. **Defining student success in the nexus of student, institution and macro-societal agencies and context**:
   "How can we collaboratively craft a shared understanding of what engagement means for them [students], for us, and then determine what criteria are most useful and descriptive of where they are coming from, where they are and where they want to go?" (p. 24).

3. **Understanding data as framed and framing**
   Data never provide the complete or full picture and that the data we do have of our students are glimpses and often proxies of complex, layered lives" (p. 25).

4. **Student data sovereignty**:
   "Student data are not something separate from students' identities, their histories, their beings … data are an integral, albeit informational part of students' being." (p. 25).

5. **Accountability**:
   "[H]igher education institutions are also accountable to students … [and] learning analytics allows institutions to grow in their understanding of the complexities of learning" (p. 28).

6. **Transparency**:
   "[T]ransparency, as a principle in the ethical collection, analysis and use of student data underpins the other principles … aiming to institutionalise ethical data collection, analysis and use" (p. 32).

7. **Co-responsibility**:
   "[E]mphasises the interdependency between institutions and students in facilitating effective and appropriate learning experiences" (p. 33).

This working paper provides background information for institutions that wish to consider the implications of collecting and using student data in the context of analytics for student success.

---

[1] Prinsloo, P. (2017). Guidelines on the Ethical Use of Student Data: A Draft Narrative Framework. Siyaphumelela conference 27 June 2017, Johannesburg.

For clarity, the following definitions of analytics and different types of analytics, will form the basis of usage in this document.

Analytics: Statistical techniques are applied to data, usually to answer a particular question, and the results are used for modelling, interpretation, decision support, etc. Data sets are usually large (e.g. all students taking a particular module; success rate of all students in the institution).

Academic Analytics: 'Academic analytics (AA) is the improvement of organizational processes, workflows, resource allocation, and institutional measurement through the use of learner, academic, and institutional data'[2]. Academic Analytics uses aggregate data from different university systems such as the student information system and these data are collected by an institutional research office for decision-support within the university (for instance, for the Executive) and external reporting (e.g. to DHET through HEMIS).

Learner Analytics: The analysis and interpretation of student data in order to support student success effectively. The data on students – biographical, demographic, academic, learning and institutional – on the university system or collected through a survey instrument are used in order to intervene as early as possible if the profile shows any risk. The analytics can be personalised, help us to learn from the data and help us to predict certain outcomes. Learner analytics can help us to decide on the best course of action for a particular student or group of students. Raw data from different sources combine to form derived variables for the purpose of analysis and reporting.

Learning Analytics: 'Learning analytics (LA) is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs'[3]. LA is usually dynamic, formative and personalized. One purpose is to improve the learning process. Data on student engagement with their learning and their emerging success are gathered through formative marks data, clickstream data, library usage or any other engagement data in order to intervene as early as possible if trends emerge that might lead to a student not fulfilling his or her academic potential. The results can be used to improve teaching and/ or to evaluate and improve student retention and success interventions. Raw data from different sources combine to form derived variables for the purpose of analysis and reporting.

The introduction of learning analytics to support student success is complex and multifaceted. The New America's "Predictive Analysis in Higher Education" [4] and Jisc's "Code of practice for learning analytics" [5] discuss the introduction and ethical use of data for learning analytics from two different contexts. The current document provides an integrated overview of these two documents aligned

---

[2] Long, P. and Siemens, G. (2011). Penetrating the Fog: Analytics in Learning and Education. *Educause Review*. https://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-and-education.

[3] Siemens, G. (2011). 1st International Conference on Learning Analytics and Knowledge, Banff, Alberta, February 27–March 1, 2011, http://tekri.athabascau.ca/analytics/.
Lang, C., Siemens, G., Wise, A. and Gašević, D. (2017). (eds). *The Handbook of Learning Analytics*. ISBN: 978-0-9952408-0-3. DOI: 10.18608/hla17 https://solaresearch.org/hla-17/. *Journal of Learning Analytics: Supporting and Promoting Learning Analytics Research*. A journal started in 2014 by the Society for Learning Analytics Research (SoLAR). http://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/3908.

[4] Ekowo, M. and I. Palmer (2017) Predictive analysis in Higher Education. New America (https://na-production.s3.amazonaws.com/documents/Predictive-Analytics-GuidingPractices.pdf)

[5] Jisc (2016) Code of practice for learning analytics – June 2015. (https://www.jisc.ac.uk/sites/default/files/jd0040_code_of_practice_for_learning_analytics_190515_v1.pdf)

to the guiding principles. In addition, it attempts to identify the role players, processes and responsibilities to ensure that the introduction of student learning analytics is carried out in a responsible manner and includes an understanding of the ethical and logistical matters. It is understood that South African institutions are legally bound to adhere to the terms of the Protection of Personal Information Act (POPI) that is not discussed here. However, as a baseline, POPI requires that the minimum amount of information possible to the performance of our legal reporting duties and our duty to support student success be requested from students and that it be used only for the purpose for which it was collected. Ethics clearance for any research project using student data should require approval from the Registrar of the university over and above clearance from an Ethics Committee.

Planning

Before embarking on the journey to integrate the use of learning analytics to support student success, it is advisable to set goals, that may be part of the institutional strategic plan. The development of goals/vision should clearly articulate the purpose and outcomes associated with the use of learning analytics. In addition, the development of such a vision statement, or set of goals, should include key stakeholders (for example, institutional management, faculty leadership, institutional researchers, academic members and students). A detailed plan could include processes and management roles related to the following:
- Communications and transparency,
- Student consent,
- Data collection and use,
- Interventions,
- Mitigation of advert impacts, and
- Professional Development.

Communication/transparency

One way to be transparent in the collection and use of data for learning analytics is to communicate all aspects of associated goals including the potential benefits associated with learning analytics, and the associated processes, policies and potential impacts (both positive and negative).

Student consent

The "Code of practice for learning analytics"[4] discussed a number of criteria that should be considered in order for students to make informed decisions so that an institution could make use of their data to support their success. These include:
- Institutional objectives for the use of learning analytics, the data required to achieve the objectives, data excluded for analyses and reasons why such data are excluded, metrics used in the analyses, how the data are interpreted, the algorithms[6] used during data analysis, and who has access to the data.
- Students should be asked for their consent for institutional use of their data to support their success. Consent could be during the registration period or some time thereafter. It is generally included in the contract that the student signs with the university. Students are entitled to opt out (refuse consent). However, as South African institutions are legally bound

---

[6] *Algorithm*: A model containing a set of criteria/ parameters (both numerical and non-numerical) with defined steps/ rules that, when applied to data sets, calculates a result, such as a prediction of success.

to report certain student data to the South African government entities, students may not opt out of the reporting of such data. But, data reported to authorities should be clearly stated.

- Any adverse consequences of giving consent should be made clear to students and the consequences if a student opts out should be clearly explained.

Data collection and use

As data are core to the use of learning analytics, it is necessary to ensure that quality data are accurately recorded, made available to those authorised to view the information, is anonymised and securely stored. However, it is understood that, owing to legal reporting, some student data may not be anonymised. Each of these issues is discussed in more depth.

The collection of data should be accurate and comprehensive. In order to support students in a timeous manner it is necessary to collect accurate data consistently to identify trends and outlying data points when they occur. However, data should be collected from a variety of sources that might be both quantitative and qualitative in nature. This is important as students are not only described by their performance data but also through their well-being. To ensure the validity of data, institutions should understand inaccuracies in the data, account for missing data, and preclude spurious correlations. In addition, the accurate and comprehensive data collection, and the interpretation of data, should be sensitive and appropriate.

Who should have access to learning analytics data needs to be defined. Students need to understand who has access to their data. Policy should make it clear that students are the owners of the data and have consented to selective access to:

- Those who advise them on their academic progress and well-being;
- Faculty members who support them in improving their performance;
- Institutional and Faculty administrative staff (anonymised data) for reporting and high-level analyses; and
- Institutional members responsible for reporting student data to authorities.

Those who have access to anonymised data should not be able to identify individuals from the metadata or re-identify individuals through the aggregating of multiple data sources, or reverse engineer the anonymisation process. Furthermore, institutions need to protect student data especially when data are shared to, or collected by, outside contracted third parties.

Appropriate controls and back-up processes need to be in place for the "retention and institutional stewardship of all data for and generated by learning analytics."

Interventions

Acknowledging the lasting, inter-generational effects of colonialism and apartheid, institutions collect, analyse and use student data with the aim of addressing these effects and tensions between ensuring quality, sustainability and success (Prinsloo, 2017, p. 9). One of the mechanisms to address past inequalities is to identify and develop interventions to overcome such inter-generational effects.

Jisc[5] (p.3) advised that institutions "should specify under which circumstances they believe they should intervene when learning analytics suggests that a student could benefit from additional

support". The scope and outcomes of such an intervention needs to be made clear to the student who might decline the advice.

All interventions, and their potential impacts on student success, should be documented and made available to all students so that they can take control of their own learning.

Through the use of learning analytics an institution might identify a number of interventions that could increase student success. Therefore, funds need to be allocated to interventions identified by institutional leaders, faculty members and students.

Mitigation of adverse impacts

"Data [can] never provide the complete or full picture and that the data we do have of our students are glimpses and often proxies of complex, layered lives" (Prinsloo, 2017, p. 25). Institutions need to ensure that "trends, norms, categorisation or any labelling of students do not bias staff, student or institutional perceptions and behaviours towards them, reinforce discriminatory attitudes or increase social power differentials."[2] (p.3). Therefore, interventions and approaches to data analyses should be reviewed to allow students to make decisions based on the analytics results of their data that in no way could be used to impact on their academic progress or well-being. Review of all interventions and their impact on students is necessary to identify, and correct, unintended consequences of analytics.

Professional Development

The successful implementation of learning analytics to support student success requires that the institution provides professional development to institutional researchers to minimise biases, guarantee privacy and protection of data; academic staff members in the use of data to improve their teaching and learning designs and interventions; students in their use of their data to improve their performance; and development of institutional skills in the evaluation of interventions that make use of both qualitative and quantitative methodologies.

Role players

The introduction of learning analytics to support student success includes many role players. However, specific individuals should be tasked with specific responsibilities to support the ethical use of data:

- Collections of data for learning analytics,
- Anonymisation of the data,
- Analytic processes,
- Intervention, and
- Retention and stewardship of data.