



## Evidence for Enhancement: Improving the Student Experience

# Learning Analytics: Data creation and collection

**Heather Gibson and Alan Clarke**  
**The Open University**

### About this paper

This paper has been written for institutional managers and academics who are using, or wish to use, learning analytics to support the enhancement of the student experience. The aim of the paper is to help inform conversations with learning analytics experts in their institutions about some of the issues and challenges that are emerging from the learning analytics research field that may impact on institutional activities.

An overarching trend is the need to increase capacity for institutional staff and students to engage with ethics, design, understanding and using learning analytics. Where this has previously been the concern of a relatively small number of experts, it is becoming increasingly important that a broader community is equipped to participate in the conversation.

This is one part of a longer discussion paper<sup>1</sup> based on an adaptation of Clow's 2012 cycle of learning analytics. The main paper includes four key sections:

- **data creation and collection (this paper)**
- working with and understanding data
- using data to enhance the student experience
- implementing learning analytics in institutions.

The main paper also includes a series of 'hot topics', which we have made available as separate factsheets.

---

<sup>1</sup> [www.enhancementthemes.ac.uk/docs/ethemes/evidence-for-enhancement/learning-analytics-discussion-paper.pdf](http://www.enhancementthemes.ac.uk/docs/ethemes/evidence-for-enhancement/learning-analytics-discussion-paper.pdf).

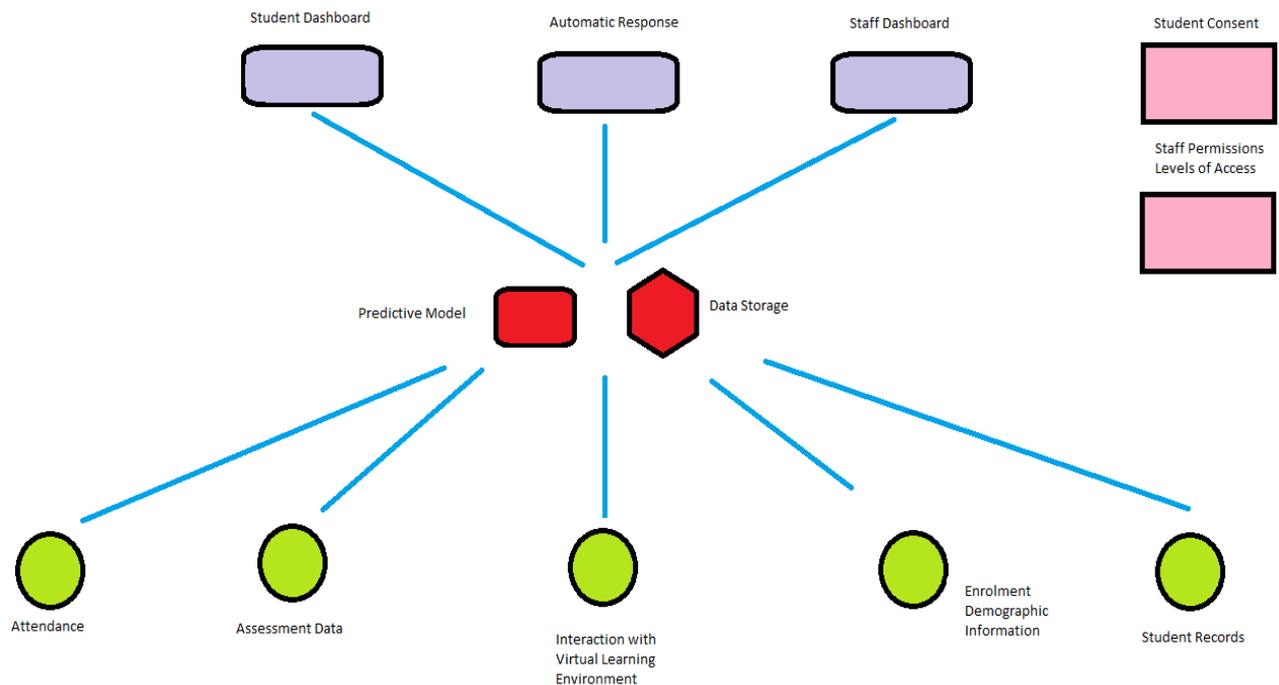
## Data creation and collection

When learners interact with their institutional systems, their activity, communication and assessment is captured. Figure 1 below, based on a diagram created by Jisc, summarises a typical learning analytics system. The green shapes denote some of the data that is captured. These include:

- attendance data - at lectures, online tutorials, library usage
- assessment data - assignment scores, submission rates, dates of submission
- interactions with any VLE - pages accessed, how often these are accessed, repeated/return access, time of access, downloads, discussion forum use
- demographic information - age, ethnicity, gender, previous educational qualifications etc
- student records - modules studied, how fees are covered, location.

The model has two overarching aspects (identified as pink boxes) which deal with ethical issues. These are:

- students consent to the use of their data
- staff access to the data is controlled and managed so that student data is protected.



**Figure 1: representation of how learning analytics can be structured (adapted from Jisc)**

It is important to recognise that the potential of learning analytics comes with the need to consider the ethics of using personal data. As Sclater (2017, p 203) points out, the consequences for the student can be considerable: the algorithms used to create and present learning analytics data will influence the institution's and the students' (and potentially employers') perceptions of student success. At a societal level, the public's relationship with data and how it is used by large organisations is contentious. In his keynote

presentation for LAK18, Selwyn (2018) makes the important point that for many people outside the field, 'the idea of learning analytics is an uneasy and sometimes controversial proposition', and that cultures of suspicion about data/technology in society have emerged that can be articulated through the messages: technology and data are not used for societal good, and the benefits of technology will not be equally shared across society. Perhaps the most important step for institutions to consider when implementing learning analytics is to work with all stakeholders to ensure that they know that the use of learning analytics data will be beneficial and ethical.

## Ethics: it's not just privacy

Learning analytics involves collecting a great deal of data of all kinds from individual learners, including personal (and often sensitive) data as well as evidence of their engagement and performance. How institutions use that data responsibly, and how the rights of the students are protected in that use, is an area of ongoing concern. On a practical level, if ethical concerns are not addressed, or perceived not to be addressed, they can inhibit the use of learning analytics in an institution, as the risks for institutional managers may appear too high (see Sclater (2016), Drachsler & Greller (2016)). As Gasevic et al (2016) note: 'It is well recognized that these (ethical) issues lie at the very heart of the field and that great care must be taken in order to assure trust building with stakeholders that are involved in and affected by the use of learning analytics.'

### Good review

Drachsler & Greller (2016) offer a thorough consideration of ethical and privacy issues and what can be done to address both. This paper also articulates the DELICATE Framework (see Figure 2, below).

Look out for: Sharon Slade (The Open University, UK) and Paul Prinsloo (University of South Africa).

Researchers in the learning analytics field agree that there is a need for more studies examining ethics and learning analytics (Ferguson & Clow, 2017). Viberg et al (2018) reviewed 252 papers covering learning analytics since 2011, finding that only 18 per cent of these mentioned ethics in relation to the research itself and that there were very few articles that considered ethics systematically. Similarly, Gasevic et al (2016), in the introduction to a special edition of the Journal of Learning Analytics on ethics, stated that more research was required. This is clearly an issue for the field to consider, and the Learning Analytics Community Exchange (LACE) project has an ongoing sub-strand of work looking at this: Ethics and Privacy in Online Learning (EP4LA). Among the work of this strand is the DELICATE framework (see Figure 2, page 5).

## Issues for institutions

Slade and Prinsloo (2013) considered whether existing university policies covering the use of student information had kept pace with the development of learning analytics, concluding that in general they had not. Privacy is exercising staff in higher education institutions because of the recent introduction of General Data Protection Regulation (GDPR). To help institutions address GDPR, Jisc has provided information and advice to help institutions respond to the challenges.<sup>2</sup> Sclater also addresses some of the common questions

---

<sup>2</sup> [www.jisc.ac.uk/guides/preparing-for-the-general-data-protection-regulation-gdpr](http://www.jisc.ac.uk/guides/preparing-for-the-general-data-protection-regulation-gdpr).

institutions may ask.<sup>3</sup> In summary, with regard to GDPR, institutions are encouraged to clearly explain to students *what* data is collected, *how* it is collected and *what* it is used for. In particular, institutions should articulate whether there is a lawful basis for collecting and processing personal data, that is, for the purposes of supporting students to succeed and to operate effectively. The Open University has developed a Student Privacy Notice<sup>4</sup> for this purpose and students are referred to this when they register on a course.

To help provide practical assistance for institutions to help develop policies to support ethical use of learning analytics, Drachsler and Greller (2016) developed a framework for institutions to use. This could be used to initiate and maintain the internal discussions within the institution that are needed in order to develop policy. The framework is called DELICATE, and Figure 2 presents it in more detail.

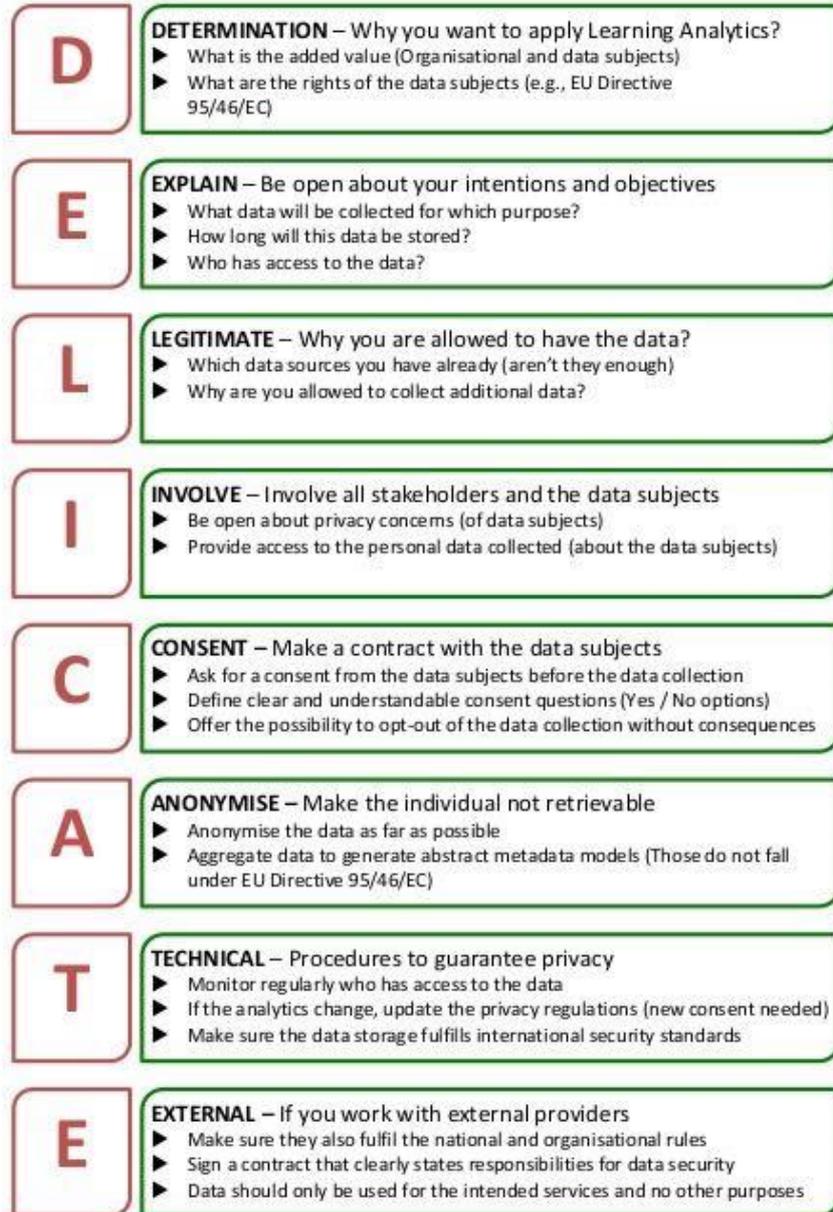
---

<sup>3</sup> [analytics.jiscinvolve.org/wp/2018/06/01/gdpr-and-learning-analytics-frequently-asked-questions](https://analytics.jiscinvolve.org/wp/2018/06/01/gdpr-and-learning-analytics-frequently-asked-questions).

<sup>4</sup> <https://help.open.ac.uk/documents/policies/privacy-notice/files/47/student-privacy-notice.pdf> (132KB).



The DELICATE Checklist  
to implement trusted  
Learning Analytics



Drachsler, H. & Greller, W. (2016). Privacy and Analytics – it's a DELICATE issue. A Checklist to establish trusted Learning Analytics. 6th Learning Analytics and Knowledge Conference 2016, April 25-29, 2016, Edinburgh, UK.

LACE Project is supported by the European Commission Seventh Framework Programme under grant 619424.



Figure 2: the DELICATE Framework

This framework offers a series of prompts for institutions to use when considering work to develop a learning analytics ethics policy.

Another framework that might be useful to institutions is provided by Ferguson et al (2016), and identifies 21 learning analytical challenges related to ethics.

1	Use data to benefit learners
2	Provide accurate and timely data
3	Ensure accuracy and validity of analysed results
4	Offer opportunities to correct data and analysis
5	Ensure results are comprehensible to end users
6	Present data/results in a way that supports learning
7	Gain informed consent
8	Safeguard individuals' interests and rights
9	Provide additional safeguards for vulnerable individuals
10	Publish mechanisms for complaint and correction of errors
11	Share insights and findings across the digital divides
12	Comply with the law
13	Ensure that data collection, usage and involvement of third parties is transparent
14	Integrate data from different sources with care
15	Manage and care for data responsibly
16	Consider how, and to whom, data is accessible
17	Ensure data is held securely
18	Limit time for which data is held before destruction and for which consent is valid
19	Clarify ownership of data
20	Anonymise and de-identify individuals
21	Provide additional safeguards for sensitive data

Jisc offers a Code of Practice for learning analytics which covers many of the areas of DELICATE, as well as the challenges set out by Ferguson et al (2016). It emphasises privacy, consent, responsibility, validity, access, use and legality and sets out expectations for each.<sup>5</sup>

These frameworks, and the Jisc Code of Practice, provide a set of pointers to help institutions initiate and maintain the internal discussions necessary to ensure that learning analytics activity is carried out ethically. They can also act as a series of ethical touchstones, or act as an 'arbitrator' for different types of staff and students who may have different perspectives on what learning analytics should be used for.<sup>6</sup> However, it should be noted that any framework or policy will be a political construct, with values, agenda and messages determined by those who create it. This argues for increased involvement of all stakeholders throughout the institutional community in the creation and development of their learning analytics ethical frameworks.

There are valid reasons why policy development may not involve the entire institutional community and a policy may be developed centrally. However, policy implementation should be complemented by raising awareness and encouraging its use throughout the institution. As Gunn et al (2017) note: 'Policies and acceptable use guidelines need to be written and synergies between policy and practice encouraged'. One method of doing this can involve the construction of an institutional policy that is then used to inform other institutional

---

<sup>5</sup> [www.jisc.ac.uk/guides/code-of-practice-for-learning-analytics](http://www.jisc.ac.uk/guides/code-of-practice-for-learning-analytics).

<sup>6</sup> [ict-innovatie.uva.nl/2013/09/13/towards-a-uniform-code-of-ethics-and-practices-for-learning-analytics](http://ict-innovatie.uva.nl/2013/09/13/towards-a-uniform-code-of-ethics-and-practices-for-learning-analytics).

policies, processes and practices. An example of an institutional policy that aims to do this is provided by the Open University. The original policy<sup>7</sup> is based on eight principles:

- '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 (for example, 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.'

These principles are then carried into University policy and practices, for example, principle 8 has generated activity within the University to increased data capacity and capability in staff, and the student privacy notice clearly sets out how and why the University collects student data (principles 2 and 5).

## Learning analytics and ethics: looking deeper

Selwyn (2018) argues that education, technology, and learning analytics are political in nature: they are not value neutral, because they have been designed to produce particular societal and political effects. He encourages learning analytics researchers to consider some key questions:

- 'What is it you are actually doing?
- Why are you doing it?
- What are the key values, ideas, agendas and ideologies built into the design of the learning analytics/data you use?'

Higher education is also a political system. As Selwyn (2018) points out, from the early years to university, education has become more dependent on data and in tandem the use of data in education has become more contested. He argues that learning analytics has become part of the ongoing debate about what education is about. Is it just about learning? Should it focus on the individual good rather than the societal good? And so on.

At institutional level, Prinsloo and Slade (2016) note that there are intrinsic power imbalances between institutions and students and that there are dangers that students' vulnerabilities can be exacerbated even if the use of learning analytics is being used to address issues of equity and equality. The authors suggest that ethical debates should 'go

---

<sup>7</sup> [help.open.ac.uk/documents/policies/ethical-use-of-student-data/files/22/ethical-use-of-student-data-policy.pdf](https://help.open.ac.uk/documents/policies/ethical-use-of-student-data/files/22/ethical-use-of-student-data-policy.pdf) (152KB).

beyond a simple 'rights' or 'privacy' perspective and explore the notion and scope of learner agency through vulnerability as a lens' (Prinsloo & Slade, 2016, p 166). This 'discursive-disclosure' approach allows what they call 'surveillance' to be located in 'the context of what is being done, by who, and for what purpose and then investigates alternative approaches to satisfy the need that initially resulted in the surveillance'. Like the questions raised by Selwyn (2018) above, the purpose for which data is being collected and used is articulated as a key concept, but this set of questions also considers who is collecting the data and encourages thinking around alternatives.

Both these references serve to emphasise that learning analytics operates in active social systems, and without careful examination of how these impact on the design and use of learning analytics, human biases (intentional or unintentional and from all stakeholders) can be inbuilt and exacerbated. The involvement of other disciplines in learning analytics such as political science, philosophy, educational research etc, will serve to hold a mirror to the discipline and help it develop a robust and ethical foundation.

The main message emerging from this brief discussion is that ethics is an ongoing concern for both the field and institutions, particularly in light of recent privacy concerns in other sectors and the introduction of GDPR legislation. However, it is clear that both the field and institutions are developing research, policy and processes to address ethics. More work could be done specifically around linking the field's work with institutional activity and concerns about student agency: for example, Prinsloo and Slade (2016) consider the issues of student agency and the potential for unintentional exacerbation of disadvantage through learning analytic work. Engaging with ethical concerns could stimulate discussion about the use of learning analytics in institutions - that is, the conversations that could arise as a result of considering ethics have the potential to unite different stakeholders and foster a sense of ownership. It is in the interests of everybody using and impacted by learning analytics that ethical issues are addressed.

As Gasevic et al (2016, p 2) note:

'We would like to take a different perspective to this and encourage the community to see ethics and privacy as enablers rather than barriers. It is natural that learning analytics stakeholders have serious questions related to a number of issues such as protection of their privacy, ownership and sharing of data, and ethics of the use of learning analytics. We would also like to posit that learning analytics can be only widely used once the critical factors are addressed, and thus, these are indeed enablers rather than barriers for adoption of learning analytics.'

## References

- Clow, D. (2012). The learning analytics cycle: closing the loop effectively. *Second International Conference on Learning Analytics and Knowledge*, (p. 134).
- Drachsler, H., & Greller, W. (2016). Privacy and Analytics – it's a DELICATE issue. A Checklist for Trusted Learning Analytics. *Sixth Learning Analytics and Knowledge Conference*. Edinburgh.
- Ferguson, R., & Clow, D. (2017). Where is the evidence? A call to action for learning analytics. *LAK '17: Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 56–65). New York: ACM International Conference Proceeding Series.
- Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., & Vuorikari, R. (2016). Research evidence on the use of learning analytics - implications for education policy. In J. Vuorikari, & J. Castaño Muñoz, *Joint Research Centre Science for Policy Report*.
- Ferguson, R., Hoel, T., Scheffel, M., & Drachsler, H. (2016). Guest Editorial: Ethics and Privacy. *Journal of Learning Analytics*, 3(1).
- Gasevic, D., Dawson, S., & Jovanovic, J. (2016). Ethics and Privacy as Enablers of Learning Analytics. *Journal of Learning Analytics*(28), 1-4.
- Gunn, C., McDonald, J., Donald, C., Milne, J., & Blumenstein, M. (2017). *Building an evidence base for teaching and learning design using learning analytics*. Wellington: Ako Aotearoa – The National Centre for Tertiary Teaching Excellence.
- Prinsloo, P., & Slade, S. (2016). Student vulnerability, agency, and learning analytics: an exploration. *Journal of Learning Analytics*, 3(1), 159–182.
- Slater, N. (2016). Developing a Code of Practice for Learning Analytics. *Journal of Learning Analytics*, 3(1), 16-42.
- Slater, N. (2017). *Learning Analytics Explained*. New York: Routledge.
- Selwyn, N. (2018). The Promises and Problems of Learning Analytics. *Eighth International Learning Analytics and Knowledge Conference*.
- Slade, S., & Prinsloo, P. (2013). Learning analytics: ethical issues and dilemmas. *American Behavioral Scientist*, 57(10), 1509–1528.
- Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The Current Landscape of Learning Analytics in Higher Education. *Computers in Human Behavior*.