Dimensions of Trust and Transparency in the Learning Analytics process

# Guiding Principles

1. **Transparency** - Share information on who, what, where and why data is collected and analysed
2. **Communication** - Open informal (no jargon) - website
3. **Rights & Responsibilities** - What we do as an educational institution. What you can do as a student data stakeholder. Maybe what either party cannot do??
4. **Governance** - processes? Where can i find out more, who can I speak to for action/remediation, what structures and processes exist for data protection within the university? HREC, Research Hub, etc etc.
5. **Legal vs ethical obligations** - ‘do no evil’..

# Learning Activity

This process involves setting up the LA process, i.e., deciding upon objectives, what data sources to use etc. In doing so, the institutional policies on LA are a foundation, codified in a Code of Practice outlining principles for purpose of data collection, responsibilities and roles, transparency and consent policies, questions of validity of analytics, access to results and data, policies on interventions, principles of minimising adverse impacts, stewardship of data, etc.

Learning Design aspects - outline to students how the learning will take place, the design of the activity and the data that will inform assessment, feedback and progress through the learning process.

Data decision - what are the data sources for the learning activity? What is the justification for the collection of the data? What sources of data will be used? If an upstream system is to be used, does this use constitute a secondary purpose? If feedback is not to be provided to students due to a legitimate educational reason, outline this at the beginning.

One of the first questions to ask when setting up a LA session will be when to involve the student(s) in this particular session.

Trust questions from a student perspective:

* Do you know what you’re doing in this learning activity?
* Have you done it before?
* What data are you using and why?
* Is my data safe?

[Scope for defining a model]

# Data Collection

Need for de-identification? Never share personal identifying information (PII).

Privacy Impact Assessment (PIA) - process done ahead of system design/data collection

No: What conditions? What storage?

* User controlled data capture - e.g., Personal Learning Record Store, could be stored in the Cloud - and contributed to the group mix upon request.
* Class or course setting, where there is established trust, and the data will not be shared and deleted after end of the course.

Yes: De-identification needed.

1. What do you think is possible in terms of anonymization?
2. Will there, in the future, be a need for adding other datasets to the analysis?
3. Risk assessment approach
4. Access to the right set of skills?
5. Organisational measures needed, i.e., do the institution need to build capacity. Should more institutions collaborate?

What consequences will the different options here have for storage solution and management structure - and how to communicate this to the stakeholders?

Data minimization -> adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed.

Data quality -> input data can be innately sparse, uncertain and incomplete.

Trust questions from a student perspective:

* Where is the data coming from?
* How are you collecting the data?
* Will it be used for any other purpose?
* Who will have access to it (is it just for this activity/class/module/institution)?
* Can I opt out?

# Storage and Processing of Data

How can manipulations of input data be detected?

Who has access?

What actions are taken to protect my data in terms of security?

Trust questions from a student perspective:

* What kinds of information are you keeping as part of this process?
* Is there anonymisation being used?
* Can my data be re-identified?
* How long is the data stored for?

# Analysis

What algorithms are in place to process the data? Can you describe the algorithm? Can any risks be identified in the analysis? Do we need to communicate these to students? If not, why not?

Trust questions from a student perspective:

* Is the analysis valid?
* What decisions will flow from the analysis?
* How is the analysis conducted?
* When is the analysis happening? Is is real-time or after the fact?

# Visualisation

Who will see the results? How will access be controlled? How is this communicated?

Do the student have the right to know (and see) all the visualisations that are done (in terms of dashboards), even if the dashboard has a different target group than the student herself?

Trust questions from a student perspective:

* Does the visualisation obscure the underlying data in any way?
* Is the presentation model valid for the analysis being visualised?

# Feedback Actions

Trust questions from a student perspective:

* How will I know if an action involving me has happened?
* Can I appeal a decision based on this analysis?
* How is the action and the data involved in the decision communicated to me?