Do researchers take student privacy seriously?

Dr. Brid Lane
What About Education?

Academia is very quickly catching up with the commercial world

Using learning analytics (LA) to track and trace students and their online activity...

...within college-provided platforms.
A Data Controller must...

1. **Obtain** and process the information fairly
2. **Keep** it only for one or more specified and lawful purposes
3. **Process** it only in ways compatible with the purposes for which it was given to you initially
4. **Keep it safe and secure**
5. **Keep it accurate and up-to-date**
6. **Ensure** that it is **adequate, relevant and not excessive**
7. **Retain it no longer than is necessary** for the specified purpose or purposes
8. **Give a copy of his/her personal data** to any individual, on request.

---

Siemens (2010) defines Learning Analytics as “the use of intelligent, **learner-produced data**, and analysis models to discover information and social connections, and to predict and advise on learning.

  - Protection of students’ identities
  - Prevention of data misuse
  - Preserve confidential user information
  - Protect the identities of the users at all times

- Individualised vs. aggregated collective data
Are we as educators doing enough (or doing anything at all) on ethical grounds to inform students that they are being tracked online

Which of the following are we doing?

- Do students know what educators mean by tracking them online?
- Do they know what PII (personally identifiable information) we are collecting?
- Do they know why we are doing it?
- Are the reasons why we are doing it reasonable?
- Are students given reasonable opportunity to give explicit consent or to opt-out?
Educational institutions should -

“ensure that learning analytics is carried out responsibly, appropriately and effectively, addressing the key legal, ethical and logistical issues which are likely to arise”

(JISC, 2015).

Methodology

• Searched for publications from academic peer-reviewed journals (key search focus was the EBSCO database)
• Search key term “learning analytics” and variations thereof (metrics, learning metrics, learning prediction).
• Inclusion / exclusion criteria
  ▪ Excluded articles that were not empirical e.g. literature reviews, framework constructions, opinion pieces, etc.
  ▪ Excluded articles that addressed multiple research cases e.g. Clow, (2013) who addressed a very large number of studies but none in any detail, meaning it was not feasible to expect coverage of ethics / privacy therein.
  ▪ Articles on MOOCs were omitted as their massiveness and the nature of student interaction therein is different to groups that are more traditional.
• Limitation - the search could not be exhaustive. Articles ranged from the years 2012 to 2016 inclusive, with one article from 2010 and one from 2017
Each article was perused with the view to identifying what (if anything) was said about participant ethics and privacy.
Findings
Of 108 empirical articles, based on the in/exclusion criteria

<table>
<thead>
<tr>
<th>Journal</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computers in Human Behaviour</td>
<td>12</td>
</tr>
<tr>
<td>Computers &amp; Education</td>
<td>8</td>
</tr>
<tr>
<td>Journal of Learning Analytics</td>
<td>6</td>
</tr>
<tr>
<td>Internet &amp; Higher Education</td>
<td>4</td>
</tr>
<tr>
<td>Educational Technology &amp; Society</td>
<td>3</td>
</tr>
<tr>
<td>British Journal of Educational Technologies</td>
<td>3</td>
</tr>
<tr>
<td>Online Learning</td>
<td>2</td>
</tr>
</tbody>
</table>
Findings

• Of 52 articles, 39 did not mention the ethical process undertaken with respondents, or privacy of their data.

• …. 3/4 did not deem privacy important enough to discuss in their methodology.

• While this is not to suggest that ethical considerations were not taken, it is a surprisingly high proportion.

Findings

The remaining 13 varied in their coverage.

• At the lowest level → those who merely mentioned privacy as being important but didn’t actually say what they did.

• The most impressive → Dyckhoff, et al., (2012) -
  • Drew on data protection law
  • Recognised that research involves a trade-off of “data privacy” versus “pedagogical useful indicators”.
  • Curtailed some aspects of their research because of privacy considerations.

• Others carried out tasks such as anonymising the data but it was not always clear if the students had been asked for their permission to be part of the study.
Conclusion

• Ultimately, personal data is becoming the life force of any aspect of education that moves online.
• There is a large-scale task involved in ensuring that we as educators are sufficiently responsive to and cognisant of student rights to data privacy.
• We do not want a situation where students are overly willing to give up privacy, and educators overly willing to take it from them.
• But where is metaphorical line and how do we know if we’ve crossed it?

THANKS!

Any questions?

You can find me at
@pennybridged & 
brid.lane@dbs.ie

“I’m not just a dot
on a big scatter plot
or a cell in Excel
or a simplified thought
I’m a human in action—
irreducible fraction—and
day-to-day data can’t say
what I’ve got.”
References