Using student data: Moving beyond data and privacy protection to student data sovereignty as basis for an ethics of care

Keynote at the 13th International Conference on E-Learning (ICEL), Cape Peninsula University of Technology (CPUT), Cape Town, South Africa

Paul Prinsloo
University of South Africa (Unisa)
@14prinsp
ACKNOWLEDGEMENT

I do not own the copyright of any of the images in this presentation. I therefore acknowledge the original copyright and licensing regime of every image used.

This presentation (excluding the images) is licensed under a Creative Commons Attribution 4.0 International License.
Let us talk, for a moment, about your data...
Who knows what about you?

How does, what they know, impact on your chances of:
• Getting car and household insurance
• Access to a medical aid
• Getting employment/promotion?

Where did they get your data?

Who do they share the data with?
• How much control do you really have about what data are collected, analysed, and used/
• What are your options to alter interpretations of data and definitions/categories?
• Can you opt out of personalised data collection, analysis and use and have your data deleted? Or has it become impossible?
Last thing I remember, I was Running for the door 
I had to find the passage back 
To the place I was before 
"Relax, " said the night man, 
"We are programmed to receive. 
You can check-out any time you like, 
But you can never leave!"

_Eagles – Hotel California_
It is not only a question of not being able to leave, but recognizing that “data are agentic”, data do things (Gullion, 2018, p. 72)
Can we ignore the fact that data collection, analysis and use were and continue to be *political acts* and serve declared and hidden assumptions of those who harvest and analyse the data and the masters they serve?

When we talk about (student) data at a conference hosted on the *African* continent, dare we ignore...
How colonialism...

• Objectified humans as mere data points and information in the global, colonial imaginary
  • Stole the dignity and lives of millions based on arbitrary criteria and beliefs about meritocracy supported by asymmetries of power
  • Extracted value in exchange for bare survival

• Controlled the movement of millions based on arbitrary criteria such as race, cultural grouping and risk of subversion?
When we talk about (student) data at a conference hosted in South Africa, dare we ignore...
How data were used during Apartheid to classify humans according to those worthy of humanity and dignity and those who were, somehow, less human, less worthy, and of lesser merit?
How do we recognise that students’ data are not, per se, indicators of their potential, merit or engagement, but rather entangled in the intergenerational impact of the skewed allocation of value and resources based on race, gender and culture?
Break the hold of digital colonialism

Michael Kwet  29 Jun 2018 00:00

Source credit: https://mg.co.za/article/2018-06-29-00-break-the-hold-of-digital-colonialism
“In South Africa, the US is colonising the digital economy, which will mean, increasingly over time, the whole of society. US-based Big Tech products are spreading, threatening to decimate local industries. At the same time, Western doctrines of what a digital society should look like — the models of how to construct a digital society — have captivated local audiences. South Africans are facing a very powerful threat, and they must change course quickly if they are to avert an insidious new phenomenon, digital colonialism” (emphasis added)

Michael Kwet (2018)

Source credit: https://mg.co.za/article/2018-06-29-00-break-the-hold-of-digital-colonialism
Data colonialism through accumulation by dispossession: New metaphors for daily data

Jim Thatcher
University of Washington—Tacoma, USA

David O’Sullivan
University of California, Berkeley, USA

Dillon Mahmoudi
Portland State University, USA

• The utopian imaginaries of digital frontierism
• “the asymmetry of this data capture process is a means of capitalist ‘accumulation by dispossession’ that colonises and commodifies everyday life in ways previously impossible” (p. 990)
• “Information about daily life is quantified and privatised, not in the hands of those who generated it, but of those who created the application... they that are dispossessed of the right to control those data” (p. 996)

How do we understand the collection, analysis and use of student data through the lens of **data colonialism** where students are dispossessed of their data, where they are surveilled, measured and classified according to imported zombie categories?

Image credit: https://en.wikipedia.org/wiki/British_Empire
To what extent does learning analytics produce narratives of dispossession, dismemberment and alienation?
• How much **control** do students have to determine what data institutions harvest; to challenge the meaning of *their* data and *our* categories of analysis?

• Can students (re)define/*alter* interpretations of data and definitions/categories? Can they offer us counter-narratives to our understanding of *their* learning and *their* life-worlds?

• Can students opt out of *personalised* data collection, analysis and use and have their data **deleted**?
Overview of the presentation

• Understanding the **macro-context** of collecting, analysing student data
• Learning analytics as a particular form of collecting, analysing and using student data
• Counter-narratives/concerns and issues
• What about the **law**? Whose law?
• **Data sovereignty** as a basis for an ethics of care
• Pointers for consideration
• (In)conclusion
The Unholy Trinity of Big Data

Different sources of data:
• Directed
• Automated
• Gifted

Different role-players with different interests
• Individuals
• Corporates
• Governments
• Higher education
• Data brokers
• Fusion centers

The fragility of consent

Different methods/types of surveillance, harvesting and analysis

Adapted & refined from Prinsloo, P. (2014). A brave new world. Presentation at SAAIR, 16-18 October
http://www.slideshare.net/prinsp/a-brave-new-world-student-surveillance-in-higher-education
CORPORATE SURVEILLANCE IN EVERYDAY LIFE

Report: How thousands of companies monitor, analyze, and influence the lives of billions. Who are the main players in today’s digital tracking? What can they infer from our purchases, phone calls, web searches, and Facebook likes? How do online platforms, tech companies, and data brokers collect, trade, and make use of personal data?

By Wolfie Christl, Cracked Labs, June 2017.

### Predicting Personal Attributes from Facebook Likes

<table>
<thead>
<tr>
<th>Predicted Attribute</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethnicity</td>
<td>95%</td>
</tr>
<tr>
<td>Gender</td>
<td>93%</td>
</tr>
<tr>
<td>Sexual Orientation (male)</td>
<td>88%</td>
</tr>
<tr>
<td>Political Views</td>
<td>85%</td>
</tr>
<tr>
<td>Religion</td>
<td>82%</td>
</tr>
<tr>
<td>Sexual Orientation (female)</td>
<td>75%</td>
</tr>
<tr>
<td>Nicotine Usage</td>
<td>73%</td>
</tr>
<tr>
<td>Alcohol Usage</td>
<td>70%</td>
</tr>
<tr>
<td>Relationship</td>
<td>67%</td>
</tr>
<tr>
<td>Drug Usage</td>
<td>65%</td>
</tr>
<tr>
<td>Parents Divorced</td>
<td>60%</td>
</tr>
</tbody>
</table>

Predicting personal attributes from Facebook Likes. Source: Kosinski et al 2013.

### Predicting Character Traits from Phone Call Records and App Usage

<table>
<thead>
<tr>
<th>Trait</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotional Stability</td>
<td>72%</td>
</tr>
<tr>
<td>Extra-VERSION</td>
<td>76%</td>
</tr>
<tr>
<td>Openness</td>
<td>69%</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>75%</td>
</tr>
<tr>
<td>Agree-Ableness</td>
<td>70%</td>
</tr>
</tbody>
</table>

### Recognizing Emotions from the Rhythm of Keyboard Typing Patterns

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONFIDENCE</td>
<td>83%</td>
</tr>
<tr>
<td>HESITANCE</td>
<td>82%</td>
</tr>
<tr>
<td>NERVOUS-NESS</td>
<td>83%</td>
</tr>
<tr>
<td>RELAXATION</td>
<td>77%</td>
</tr>
<tr>
<td>SADNESS</td>
<td>88%</td>
</tr>
<tr>
<td>TIRED</td>
<td>84%</td>
</tr>
</tbody>
</table>

### Source

We Want Privacy, but Can’t Stop Sharing

By KATE MURPHY  OCT. 4, 2014

... digital promiscuity

Our bodies are “leaking” (both passively and actively) “flows of data to an overlaying surveillant assemblage where they are subject to analysis”

Our data are not something separate from our identities, our histories, our \textit{beings}. Our data are an integral, albeit informational part of our being. Data are therefore not something we own and can give away. \textbf{We don’t own our data but we are, increasingly, constituted by our data.}

Big Data's Coming Of Age In Higher Education

Guest Post Written By

Bridget Burns, Executive Director, University Innovation Alliance

Source credit: https://www.forbes.com/sites/schoolboard/2016/01/29/big-datas-coming-of-age-in-higher-education/#5384ce8c1c41
The world of (student) data

- Individuals
- Corporates
- Governments
- Data brokers
- Fusion centers

(Higher) Education

- Directed
- Automated
- Gifted

Academic Analytics

Learning Analytics

The world of (student) data
New Study on “the Marketplace for Student Data”

Source credit: https://fpf.org/2018/06/06/new-study-on-the-marketplace-for-student-data/

By Sara Collins, Stacey Gray, Tyler Park, and Amelia Vance
“there is very little transparency as to the overall marketplace for student data. These types of lists about students are evidently available for purchase from data brokers, but the commercial marketplace is opaque from the perspective of students and parents.”

Source credit: https://fpf.org/2018/06/06/new-study-on-the-marketplace-for-student-data/
“There are no safeguards in place to protect anonymity. There is no public transparency or accountability of its purposes or use.”
Black people in UK 21 times more likely to have university applications investigated, figures show

‘This shocking practice highlights just how pervasive institutional racism is across the higher education sector,’ says Labour’s shadow education secretary

Eleanor Busby Education Correspondent | Monday 23 April 2018 19:30 | 28 comments
“Learning institutions and corporations make little use of the data learners "throw off" (sic) in the process of accessing learning materials, interacting with educators and peers, and creating new content” (emphasis added)
“Learning analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs”.

Source credit: https://tekri.athabascau.ca/analytics/
Student data as the ‘new black”, as oil, as a resource to be mined

The Predictive Learning Analytics Revolution: Leveraging Learning Data for Student Success

Tuesday, October 6, 2015

Author(s): Sakinah Alhadad, Kimberly Arnold, Josh Baron, Ilana Bayer, Christopher Brooks, Russ Little, Rosemary Rocchio, Shady Shehata, John Whitmer

Sources(s): EDUCAUSE Center for Analysis and Research (ECAR)

Collection(s): ECAR Working Group

Academic Information Systems, Analytics, Data Administration and Management, Data Mining, Learning Analytics, Learning Management Systems (LMS), Student Information Systems, Student Learning Support, Student Success

Page credit: https://library.educause.edu/resources/2015/10/the-predictive-learning-analytics-revolution-leveraging-learning-data-for-student-success
Higher Education and the Revolution of Learning Analytics

2016

Page credit: https://icde.memberclicks.net/assets/RESOURCES/anne_la_report%20cc%20licence.pdf
Universities are tracking their students. Is it clever or creepy?

Learning analytics are becoming increasingly popular for improving learning and cutting drop-out rates - but critics question the impact on privacy

Source credit: https://www.theguardian.com/higher-education-network/2016/aug/03/learning-analytics-universities-data-track-students
The Higher Ed Learning Revolution: Tracking Each Student’s Every Move

January 11, 2017 - 6:23 AM ET

ERIC WESTERVELT
How (and Why) Ed-Tech Companies Are Tracking Students' Feelings

By Benjamin Herold | Education Week  June 12, 2018

Source credit:
https://mobile.edweek.org/c.jsp?cid=25919761&bcid=25919761&rssid=25919751&item=http%3a%2f%2fapi.edweek.org%2fv1%2few%3fuuid%3dC08929D8-6E6F-11E8-BE8B-7F0EB4743667
At this Chinese school, Big Brother was watching students — and charting every smile or frown

Using the latest artificial intelligence software, the devices tracked students’ behavior and read their facial expressions, grouping each face into one of seven emotions: anger, fear, disgust, surprise, happiness, sadness and what was labeled as neutral.
Concerns, resistance and counter-narratives

Learning Analytics
Ethical Issues and Dilemmas
Sharon Slade, Paul Prinsloo
First Published March 4, 2013 | Research Article

Abstract

The field of learning analytics has the potential to enable higher education institutions to increase their understanding of their students' learning needs and to use that understanding to positively influence student learning and progression. Analysis of data relating to students and their engagement with their learning is the foundation of this process. There is an inherent assumption linked to learning analytics that knowledge of a learner’s behavior is advantageous for the individual, instructor, and educational provider. It seems intuitively obvious that a greater understanding of a student cohort and the learning designs and interventions they best respond

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...
Stanford Calls for Responsible Use of Student Data in Higher Ed

6 October 2016

By Krysten Crawford  Oct 6, 2016
From High School to Harvard, Students Urge for Clarity on Privacy Rights

By Jenny Abamu  Jun 13, 2017
Privacy

When Learning Analytics Violate Student Privacy

The rise of analytics in higher education raises questions about the responsible use of student data. Here are some of the issues and how institutions are dealing with them.

By David Raths | 05/02/18

Source credit: https://campustechnology.com/articles/2018/05/02/when-learning-analytics-violate-student-privacy.aspx
Setting the Table: Responsible Use of Student Data in Higher Education

by Martin Kurzweil and Mitchell Stevens  Monday, May 7, 2018  In Print  PDF

Source credit:

Credit: Steve McCracken, © 2018
South Africa: Protection of Personal Information Act, 2013

Guide to the General Data Protection Regulation (GDPR)

Introduction
The Guide to the GDPR explains the provisions of the GDPR to help organisations comply with its requirements. It is for those who have day-to-day responsibility for data protection.

This is a living document and we are working to expand it in key areas. It includes links to relevant sections of the GDPR itself, to other ICO guidance and to guidance produced by the EU’s Article 29 Working Party. The Working Party includes representatives of the data protection authorities from each EU member state, and the ICO is the UK’s representative.

Alongside the Guide to the GDPR, we have produced a number of tools to help organisations to prepare for the GDPR:

- GDPR: 12 steps to take now
- Data protection self assessment toolkit

Legislation (permanently) lags behind

A View from Vivek Wadhwa

Laws and Ethics Can’t Keep Pace with Technology

Codes we live by, laws we follow, and computers that move too fast to care.

April 15, 2014

Source credit: https://www.technologyreview.com/s/526401/laws-and.ethics.cant.keep.pace.with.technology/
Legislation has to be enforced

• Once the data have been shared/collected, the ‘new’ owners of those data are difficult/impossible to trace/confront and summoned
• The ”black box’ of algorithms
• The cost/expertise required to enforce legislation
Legislation is never...

“...when a just decision has been reached, there is still much ethical work to be done”

(Noddings, 1999)
Student data sovereignty as a basis for an ethics of care

### Sources of data

**Directed** - A digital form of surveillance wherein the “gaze of the technology is focused on a person or place *by a human operator*”

**Automated** - “an inherent, automatic function of the device or system and include traces …”

**Gifted/volunteered** - In exchange for services and/or digital promiscuity

<table>
<thead>
<tr>
<th>The right of privacy</th>
<th>Privacy as a human right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student personal data</td>
<td></td>
</tr>
<tr>
<td>The legitimate use of data: aggregated and/or personalised</td>
<td></td>
</tr>
<tr>
<td>The fiduciary role of (higher) education</td>
<td></td>
</tr>
</tbody>
</table>
We have to recognize that... (1)

Data are *not* neutral, raw, objective and pre-analytic but framed “technically, economically, ethically, temporally, spatially and philosophically. Data do not exist independently of the ideas, instruments, practices, contexts and knowledges used to generate, process and analyse them”

(Kitchen, 2014, p. 2)
“... the ontologies and the epistemologies of those conquered were replaced by the narratives, ontologies and epistemologies of the colonisers. Indigenous peoples lost their right to narrate their own beliefs in a language that would make sense to them. They lost their right to self-define. Indigenous peoples beliefs and narratives were demonized and rated as primitive. They had to accept the beliefs and narratives of the colonisers as their own. The colonisers’ account was the only account that counted”

(Poole, 2016; p.58 ; emphasis added)

To what extent have our analyses and our dashboards become voice-overs, ventriloquist dolls to speak in our voices, our categories, and our understanding of their learning?
Central to the act of colonizing was to define the colonised as data-objects, who were classified into typologies defined by the coloniser – ‘we call people like you x’.

Once the category and the metrics were established, the coloniser used these metrics to validate, normalise and perpetuate these categories and types.

Sounds familiar?
We have to recognize that our data are broken (2)

Embracing Imperfection in Learning Analytics

Kirsty Kitto  
Connected Intelligence Centre,  
University of Technology Sydney  
Sydney, NSW  
kirsty.kitto@uts.edu.au

Simón Buckingham Shum  
Connected Intelligence Centre,  
University of Technology Sydney  
Sydney, NSW  
simon.buckinghamshum@uts.edu.au

Andrew Gibson  
Connected Intelligence Centre,  
University of Technology Sydney  
Sydney, NSW  
andrew.gibson@uts.edu.au

ABSTRACT

Learning Analytics (LA) sits at the confluence of many contributing disciplines, which brings the risk of hidden assumptions inherited from those fields. Here, we consider a hidden assumption derived from computer science, namely, that improving computational accuracy in classification is always a worthy goal. We demonstrate that this assumption is unlikely to hold in some important educational contexts, and argue that embracing computational “imperfection” can improve outcomes for those scenarios. Specifically, we show that learner-facing approaches aimed at “learning how to learn” require more holistic validation strategies. We consider what information must be provided in order to reasonably evaluate algorithmic tools in LA, to facilitate transparency and realistic performance comparisons.

if a genuine dialogue between disciplines is emerging, one that sees the establishment of new common ground in a move towards transdisciplinarity. Such a dialogue requires an ongoing and critical examination of the hidden assumptions that are imported into LA from other fields; are some fields getting an “easier” treatment because they came with tightly defined methodologies? Here, we will call into question the suitability of appropriating validation criteria from computer science for computational models underpinning student-facing feedback, for particular forms of learning.

We think it timely, with the rising popularity of data science, to call attention to some of the problems that can infiltrate a field like LA if we do not pay careful attention to our underlying assumptions. Sometimes it is all too easy to evaluate a methodology, using a set of field specific criteria, while leaving out the learning. Therefore, the purpose of this paper is to ensure that the field of LA maintains

Research as a form of colonialism (3)

“Research is one of the ways in which the underlying code of imperialism and colonialism is both regulated and realized”

“It galls us that Western researchers and intellectuals can assume to know all that is possible to know about us, on the basis of their brief encounters with some of us”

(Smith, 2012, in Gullion, 2018, p. 830)
Destabilising our Zombie Categories (4)

“categories from the past that we continue to use even though they have outlived their usefulness and even though they mask a different reality”

(Plummer, 2011, p. 195)
(a) Indicators name things

“To claim ‘naming rights’, indigenous peoples need to replace indicators that have been constructed according to hegemonic categories and motivated by Global North normative assumptions with indicators that reflect their own local understandings of their social world” (Morphy, 2016, p. 104)

(b) Indicators compare and rank

“Encapsulated indigenous minorities within settler states constantly find themselves being compared, as a ‘population’, with the ‘mainstream population’—and found wanting. [...] They have ‘gaps’ that need to be ‘closed’, and improvement is defined in terms of the indicators that measure the gaps” (Morphy, 2016, p. 105)
(c) Indicators simplify complex phenomena

“...categorisation is used as a tool of simplification with respect to complex phenomena such as the ‘family’ and the ‘household’” (Morphy, 2016; p. 106).

To what extent do our categories of analysis **dismember** and **dislocate** our students, **alienate** them from what **they** know and how **they** see the world and their learning journeys?
Gullion (2018) does not propose that we stop using categories, but we need to recognize the ontologies and epistemologies that informed and normalised our categories of analysis.
“Epistemic freedom speaks to cognitive justice. Epistemic freedom is fundamentally about the right to think, theorise, interpret the world, develop own methodologies and write from where one is located and unencumbered by Eurocentrism” (Ndlovu-Gathsheni, 2018, p. 3)
We have to move the centre

“I am concerned with moving the centre in two two senses at least. One is the need to move the centre from its assumed location in the West to a multiplicity of spheres in all the cultures of the world. ... The second sense is even more important [...]. Within nearly all nations today the centre is located in the dominant social stratum, a male bourgeois minority [...]. Moving the centre in the two senses... will contribute to freeing of the world of cultures from the restrictive walls of nationalism, class, race and gender” (emphasis added)

(Ngugi wa Thiong’o, 1993, in Ndlovu-Gatsheni, 2018, p. 4)
Principle 1: Equal partners

“If you have come to help us, you can go home. If you have come to accompany us, please come. We can talk”


Principle 2: Data are political

Data collection, analysis and use are *political acts* and serve declared and hidden assumptions of those who harvest and analyse the data and the masters they serve?

**Principle 3: Student data sovereignty**

Student data is not something separate from students’ identities, their histories, their *beings*. We must accept that data are an integral, albeit informational part of students being. Data is therefore not something a student owns but rather is. Students do not own their data but are constituted by their data.

**Students do NOT “throw off” data** (LAK11)

**Principle 4: Student data sovereignty (cont.)**

- Students have a right to control what personalised data is collected from them, for what purposes, by whom, and how it will be stored, governed.
- Students have the right to access the data we have of them, to know who accessed their data, and how their data was used.
- Students have a right to know what the rationale/criteria are for how we categorise them, our ‘regimes of truth’ and to engage with us making sense of their data.
- We need to ask them what categories and analyses will help them to make sense of their learning.
- Students have a right to supported and accessible recourse.
Principle 5: Let’s talk

Student data is an invitation to start a conversation
(In)conclusions

Without student data sovereignty, our care and fiduciary duty towards our students can very easily result in perpetuating bias and stereotypes, pathogenic vulnerability and normalising asymmetrical power relationships and new forms of colonialism...
You call me a misfit, a risk, a dropout and stop-out
Your research indicates that ‘students like me’ may not make it
You ask me questions regarding my financial status, where I live,
how many dependents I have, and I know that once I tell you,
I will become a number on a spreadsheet
I will be color-coded
I will become part of a structural equation model that re-affirms that
People like me
Don’t belong here
Somehow I don’t fit in you spreadsheet
But I want you to know that I am so much more
I am so much more than how you define me
I am so much more than my home address
(the one I lied about to get access to funding or to get a place in residence)
I am also a brother, a sister, a mother, a dependent, a carer
I don’t fit in your spreadsheets
I am not a dropout, I am a refugee, a migrant
I am in exile
Talk to me
THANK YOU

Paul Prinsloo

Research Professor in Open Distance Learning (ODL)
College of Economic and Management Sciences, Samuel Pauw Building, Office 5-21, P.O. Box 392
Unisa, 0003, Republic of South Africa
T: +27 (0) 12 433 4719 (office)
prinsp@unisa.ac.za
Skype: paul.prinsloo59
Personal blog:
http://opendistanceteachingandlearning.wordpress.com
Twitter profile: @14prinsp