



SoLAR Webinar, 6<sup>th</sup> August, 2019

# Learning Analytics as Educational Knowledge Infrastructure

Simon Buckingham Shum

University of Technology Sydney  
Professor of Learning Informatics  
Director, Connected Intelligence Centre  
[@sbuckshum](https://twitter.com/sbuckshum) • <http://cic.uts.edu.au>  
<http://Simon.BuckinghamShum.net>





Deep acknowledgements to  
the team whose joint work has  
shaped my thinking...

<https://cic.uts.edu.au/about/people>



Simon Knight

Lecturer



Ágnes Sándor

Honorary Associate



Shibani Antonette

Lecturer



Radhika Mogarkar

Applications  
Developer



Andrew Gibson

Research Fellow  
(Reflective Writing  
Analytics)



Kirsty Kitto

Senior Lecturer  
(Data Science)



Carlos Prieto

PhD Student  
(Learning Analytics)



Gabrielle Gardiner

Senior Manager



Sophie Abel

PhD Student



Roberto Maldonado  
Martinez

Research Fellow  
(Educational Data  
Science)



Vanessa Echeverria  
Barzola

PhD Student



Ming Liu

Research Fellow  
(Educational Data  
Science)

# An extraordinary sensor, modelling and prediction infrastructure



# “infrastructure”

educational  
infrastructure  
we can trust?

ethically and scientifically

People are literally on the streets protesting against AI in education



We need trust-building conversations for an informed dialogue. A luddite rebellion won't help anyone...



7:15 PM - 26 Apr 2019

<https://twitter.com/AGavrielatos/status/1121704316069236739>

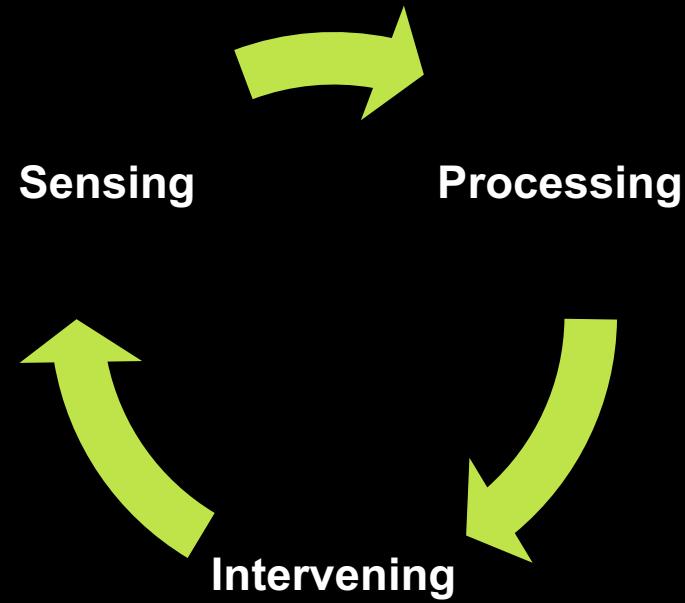
Proposition for today:

We're now in a transitional phase — we're laying foundations for the next

educational  
knowledge  
infrastructure



Machines see the world through  
**computational models**  
analysing new forms and volumes of digital data



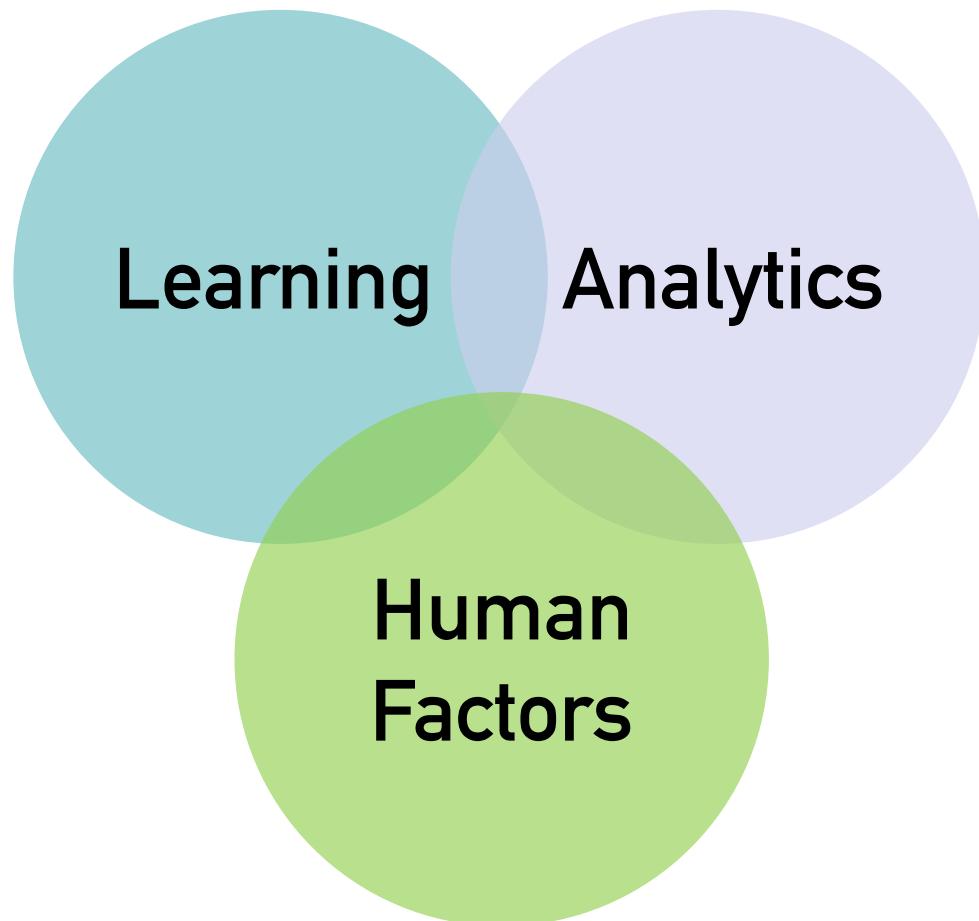


If we are tracking **learner activity**  
through the lens of data / analytics / AI  
we better have a very good idea...

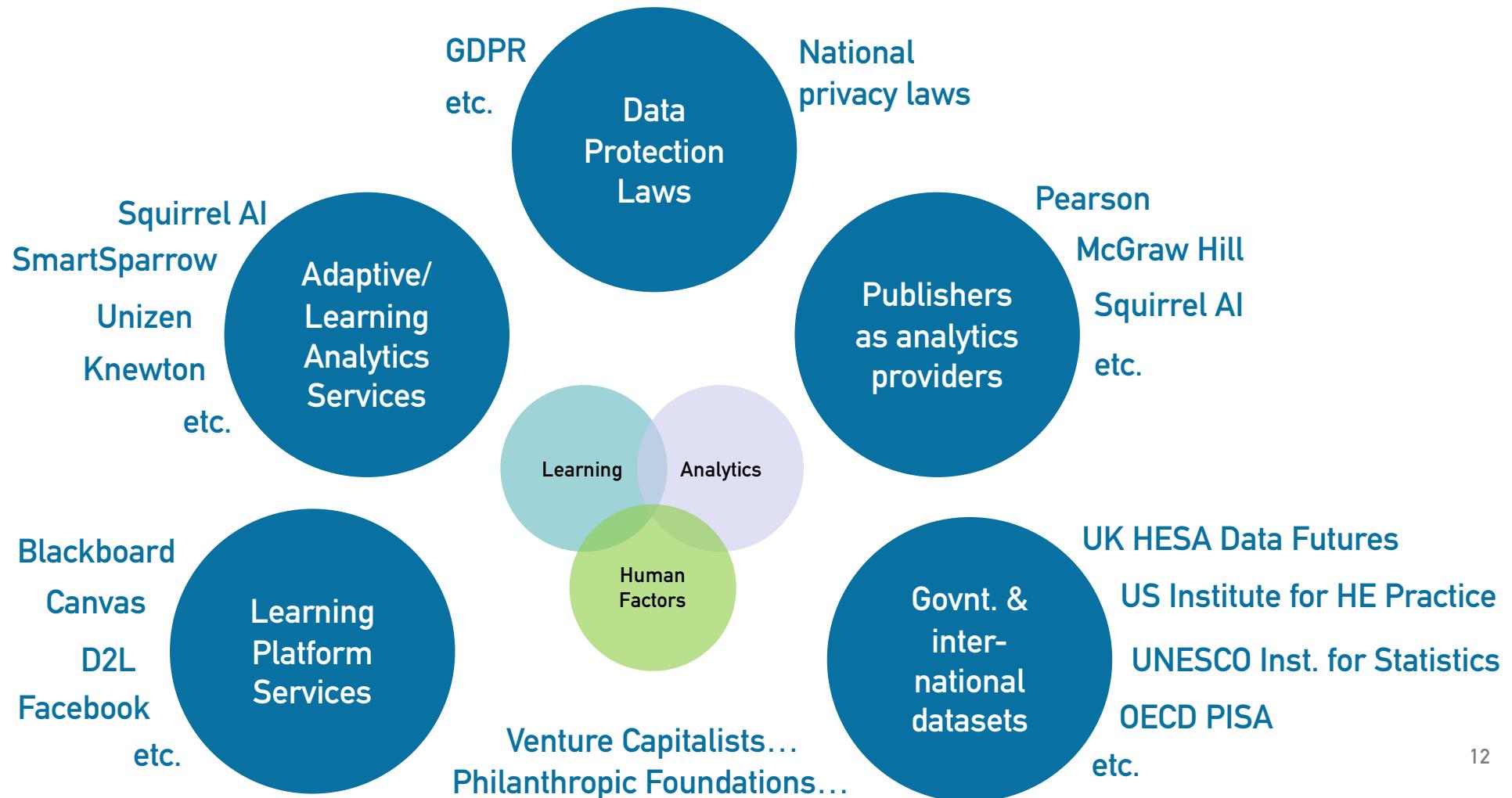
**how those lenses are cut**  
**who cut them, for whom**  
**how they distort the view**

We need  
**wholistic (systemic, human-centred)**  
lenses to design, monitor and evaluate  
data-intensive educational infrastructure.

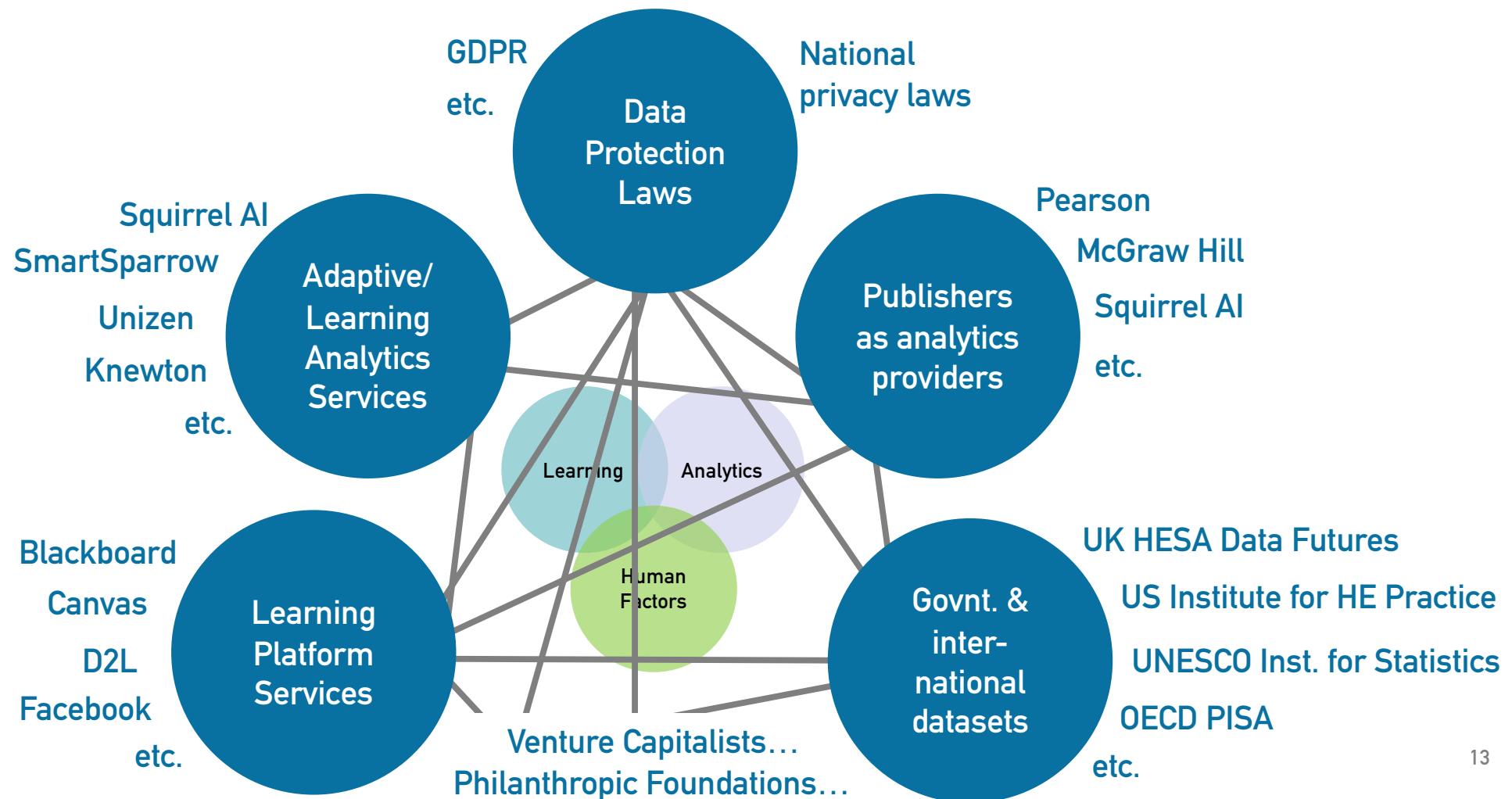
# Learning Analytics: A Human-Centred Design Discipline



# A rapidly changing educational data/analytics ecosystem...



# A rapidly changing educational data/analytics ecosystem...

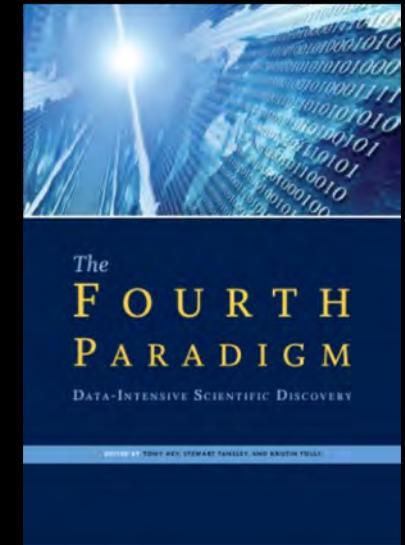


Expand from...

# “The Fourth Paradigm”

a Computer Science vision of how research is building on the  
Empirical, Theoretical and Computational paradigms  
moving into a Data-Intensive paradigm

<https://www.microsoft.com/en-us/research/publication/fourth-paradigm-data-intensive-scientific-discovery>



To see the wider systems...

# “Knowledge Infrastructures”

a critical lens on how **human+technical systems** in science  
interoperate to construct, share, contest and sanction knowledge

<http://hdl.handle.net/2027.42/97552>

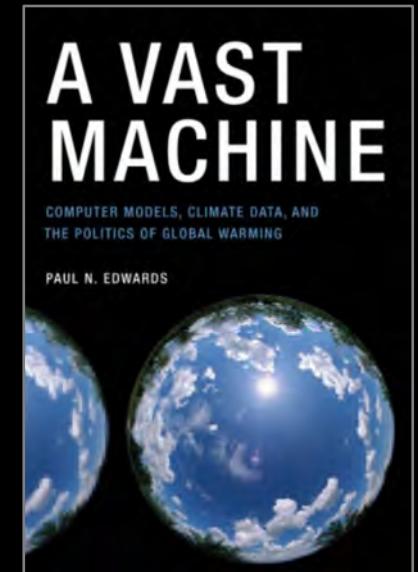


e.g. Paul Edwards on...

# climate science

How do **global data, models, visualisations, science and politics** combine to produce knowledge about the past, present and future, and how do they handle uncertainty?

<https://mitpress.mit.edu/books/vast-machine>



That's what a knowledge infrastructure looks like after nearly 200 years' evolution

“Computer Models, Climate Data, and the Politics of Global Warming”

“Computer Models, Learning Data, and the Politics of Education” ...??

# “Knowledge Infrastructures”

“robust networks of people, artifacts, and institutions that generate, share, and maintain specific knowledge about the human and natural worlds.”

Routine, well-functioning knowledge systems include the world weather forecast infrastructure, the Centers for Disease Control, or the Intergovernmental Panel on Climate Change — individuals, organizations, routines, shared norms, and practices.

# “Knowledge Infrastructures”

“**Infrastructures are not systems**, in the sense of fully coherent, deliberately engineered, end-to-end processes.

...**ecologies or complex adaptive systems** [...] made to **interoperate** by means of **standards**, **socket layers**, **social practices**, **norms**, and **individual behaviors**.“

# “Knowledge Infrastructures”

“**Infrastructures are not systems**, in the sense of fully coherent, deliberately engineered, end-to-end processes.

...ecologies or complex adaptive systems [...] made to **interoperate** by means of **standards**, socket layers, social practices, norms, and individual behaviors.”



I think we can see the  
educational ecosystem  
here

# “Knowledge Infrastructures”

“I intend the notion of knowledge infrastructure to signal parallels with other infrastructures [...] Yet this is no mere analogy [...]”

Get rid of infrastructure and you are left with claims you can't back up, facts you can't verify, comprehension you can't share, and data you can't trust.” (p.19)

# “Knowledge Infrastructures”

...perform 3 key functions...

1

Monitoring

2

Modelling

3

Memory

# Knowledge Infrastructure concepts

## metadata friction

“People long ago observed climate and weather for their own reasons, within the knowledge frameworks of their times.

You would like to use what they observed — not as they used it, but in new ways, with more precise, more powerful tools.

[...]

So you dig into the history of data. You fight metadata friction, the difficulty of recovering contextual knowledge about old records.”

(p.xvii)

# Knowledge Infrastructure concepts

## metadata friction

“People long ago observed climate and weather for their own reasons, within the knowledge frameworks of their times.

You would like to use what they observed — not as they used it, but in new ways, with more precise, more powerful tools.

[...]

So you dig into the history of data. You fight metadata friction, the difficulty of recovering contextual knowledge about old records.”

(p.xvii)

cf. Reanalysis of educational data  
(your own and others') using  
computational methods

# Knowledge Infrastructure concepts

Models, models, models...

“Everything we know about the world’s climate — past, present, and future — we know through models.” (p.xiv)

“I’m not talking about the difference between “raw” and “cooked” data. I mean this literally. Today, no collection of signals or observations [...] becomes global in time and space without first passing through a series of data models.” (p.xiii)

# Knowledge Infrastructure concepts

Models, models, models...

Machines 'see'  
learners only through  
models

"Everything we know about the world's climate — past, present, and future — we know through models." (p.xiv)

Today, no collection of signals or observations [...] becomes global in time and space without first passing through a series of data models." (p.xiii)

"Raw data is an oxymoron"  
(Geof Bowker)

# Knowledge Infrastructure concepts

## infrastructural inversion

“The climate knowledge infrastructure never disappears from view, because it functions by infrastructural inversion: continual self-interrogation, examining and reexamining its own past. The black box of climate history is never closed.”

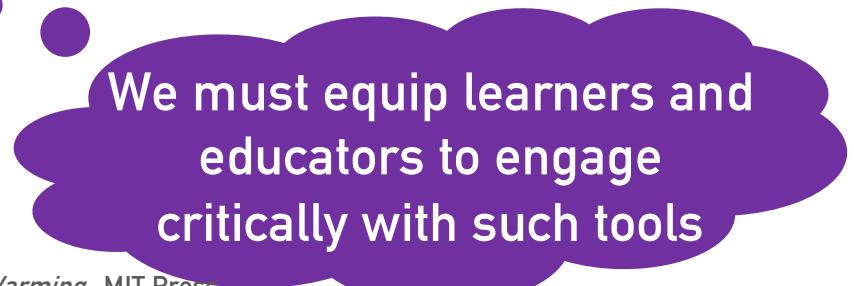
# Knowledge Infrastructure concepts

## infrastructural inversion

“The climate knowledge infrastructure never disappears from view, because it functions by infrastructural inversion: continual self-interrogation, examining and reexamining its own past. The black box of climate history is never closed.”



We must keep lifting the lid  
on learning analytics  
infrastructures

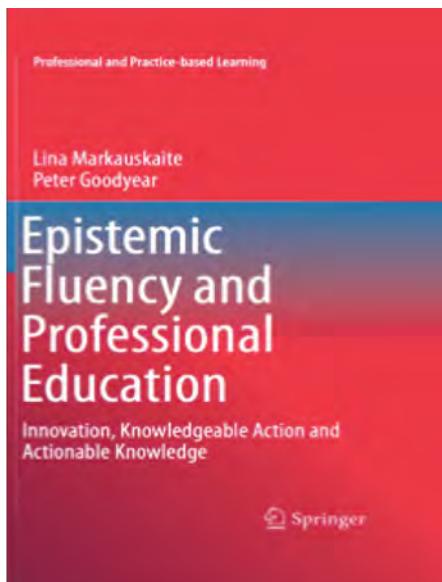


We must equip learners and  
educators to engage  
critically with such tools

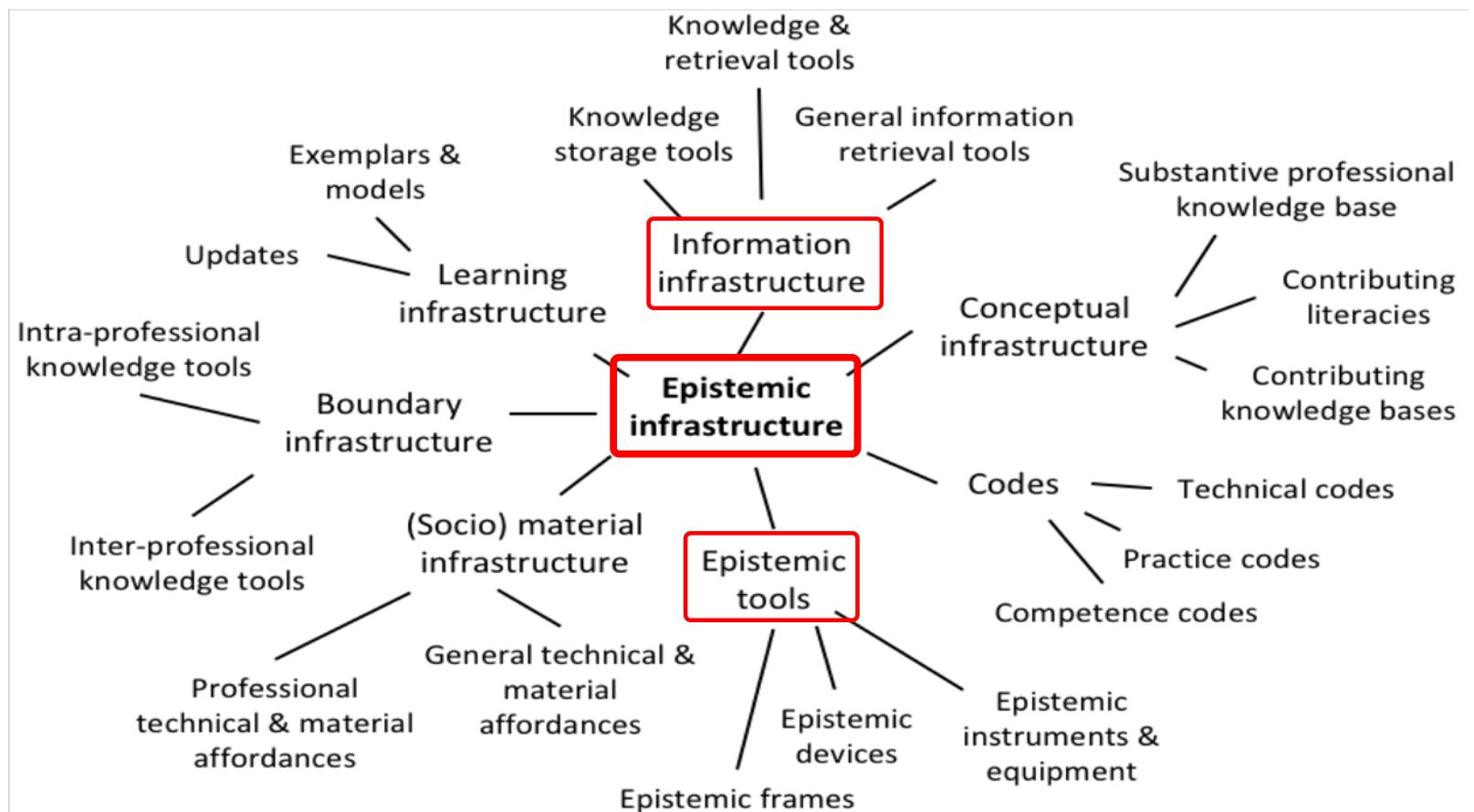
Paul Edwards (2010). *A Vast Machine. Computer Models, Climate Data, and the Politics of Global Warming.* MIT Press

# *Epistemic Infrastructure* taxonomy for professional knowledge

## Partic contributions at the “Micro-KI” level: how professionals construct their EI



Markauskaite, L. & Goodyear, P. (2017). *Epistemic Fluency and Professional Education: Innovation, Knowledgeable Action and Actionable Knowledge* (Springer, 2017), p.376



In what senses might Learning Analytics constitute,  
or at least contribute to, an emerging

knowledge  
infrastructure  
?

# In what senses might Learning Analytics constitute, or at least contribute to, an emerging KI?

1

KI concepts seem  
to apply to critical  
perspectives on LA

2

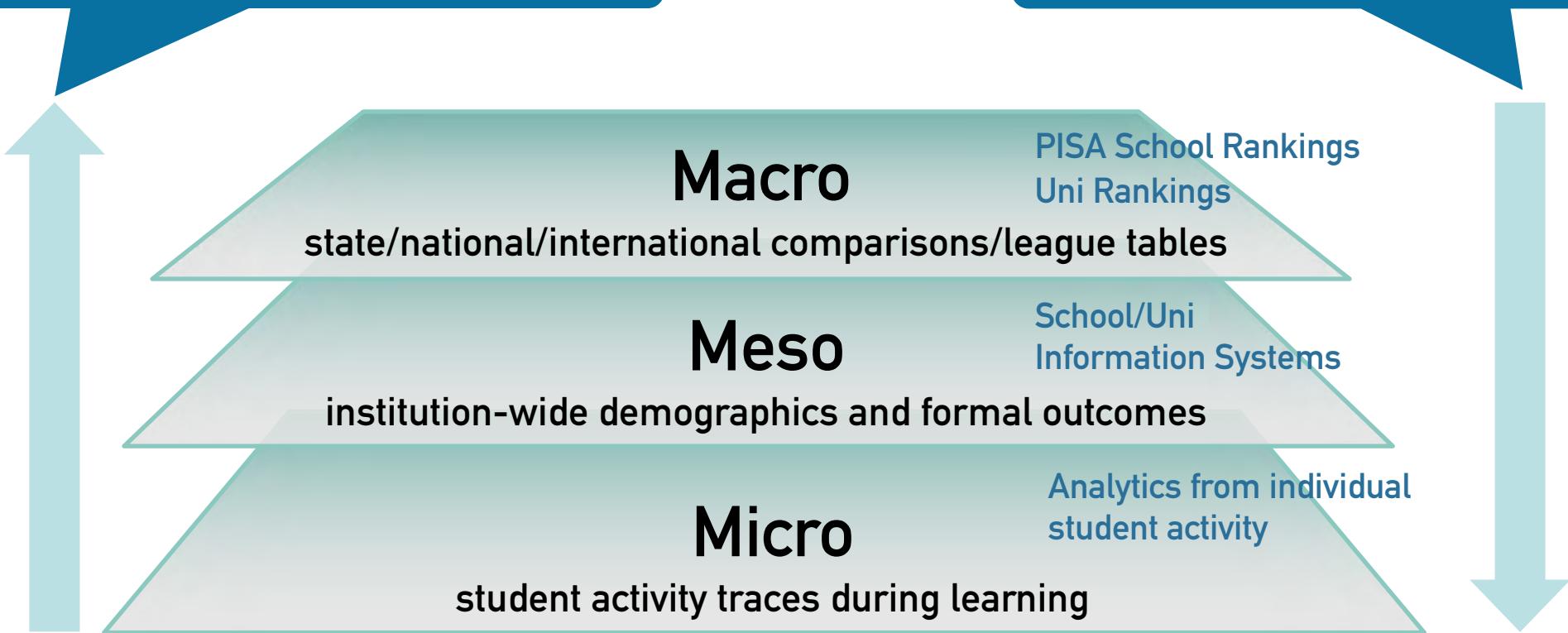
LA is starting to display  
KI properties at different  
levels of the system

LA is only 10 years old, and there's much to do.  
But knowing what functioning KIs look like could help us prioritise.

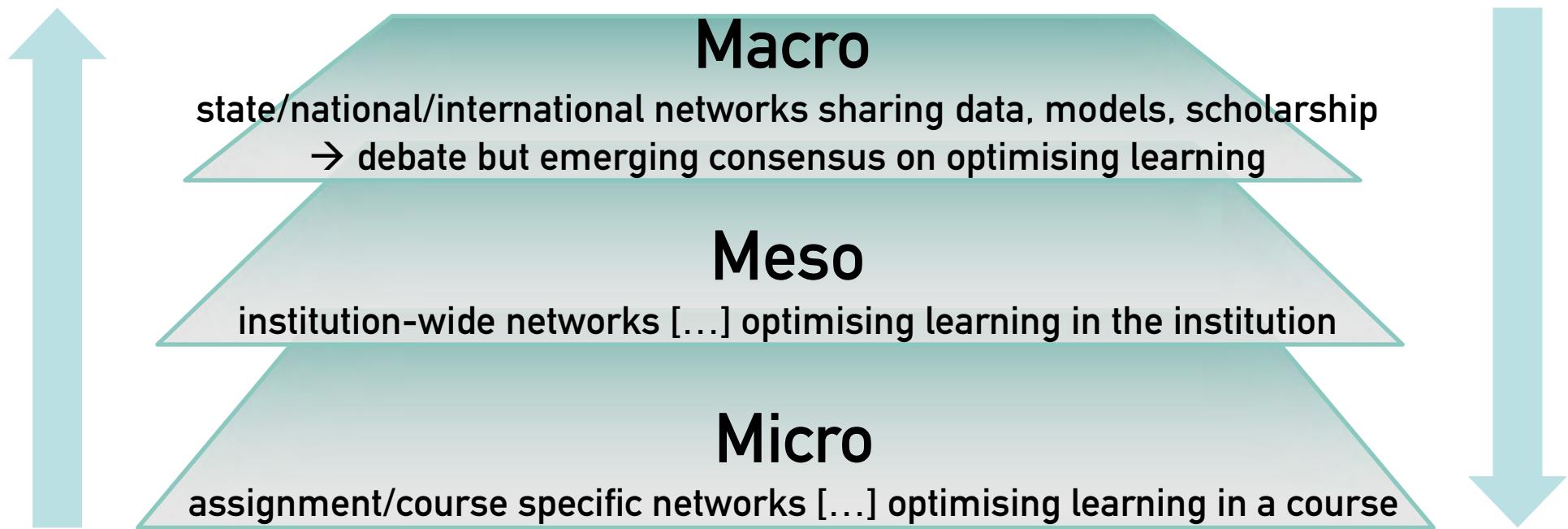
# Macro/Meso/Micro Learning Analytics

Aggregation of user traces enriches meso + macro analytics with finer-grained process data

Breadth + depth from macro + meso levels add power to micro analytics



# Macro/Meso/Micro Educational KI



**Trusted data sources \***

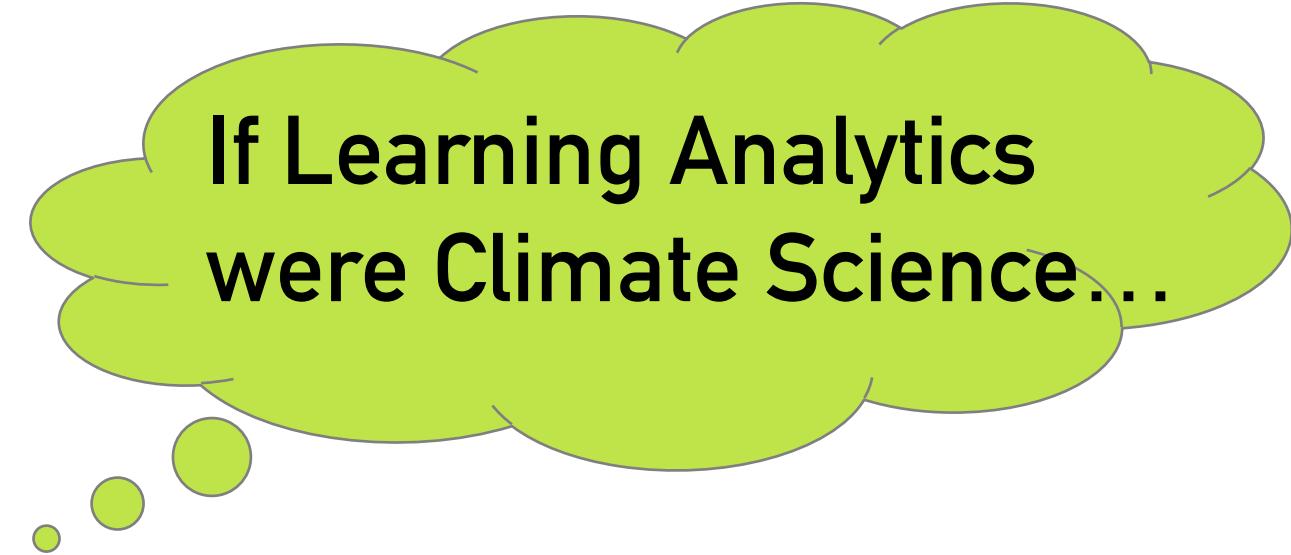
**Validated models \***

**Interoperable data flows and models \***

**Established research methodologies \***

**Government policy held accountable to international scientific consensus \***

\* all under rigorous scholarly review and debate



# If the challenge is to build education's KI, what are the practical implications for LA?

## Macro

state/national/international networks sharing data, models, scholarship  
→ debate but emerging consensus on optimising learning

## Meso

institution-wide networks [...] optimising learning in the institution

## Micro

assignment/course specific networks [...] optimising learning in a course

**Accountability: ground models in educ. research + learning sciences**

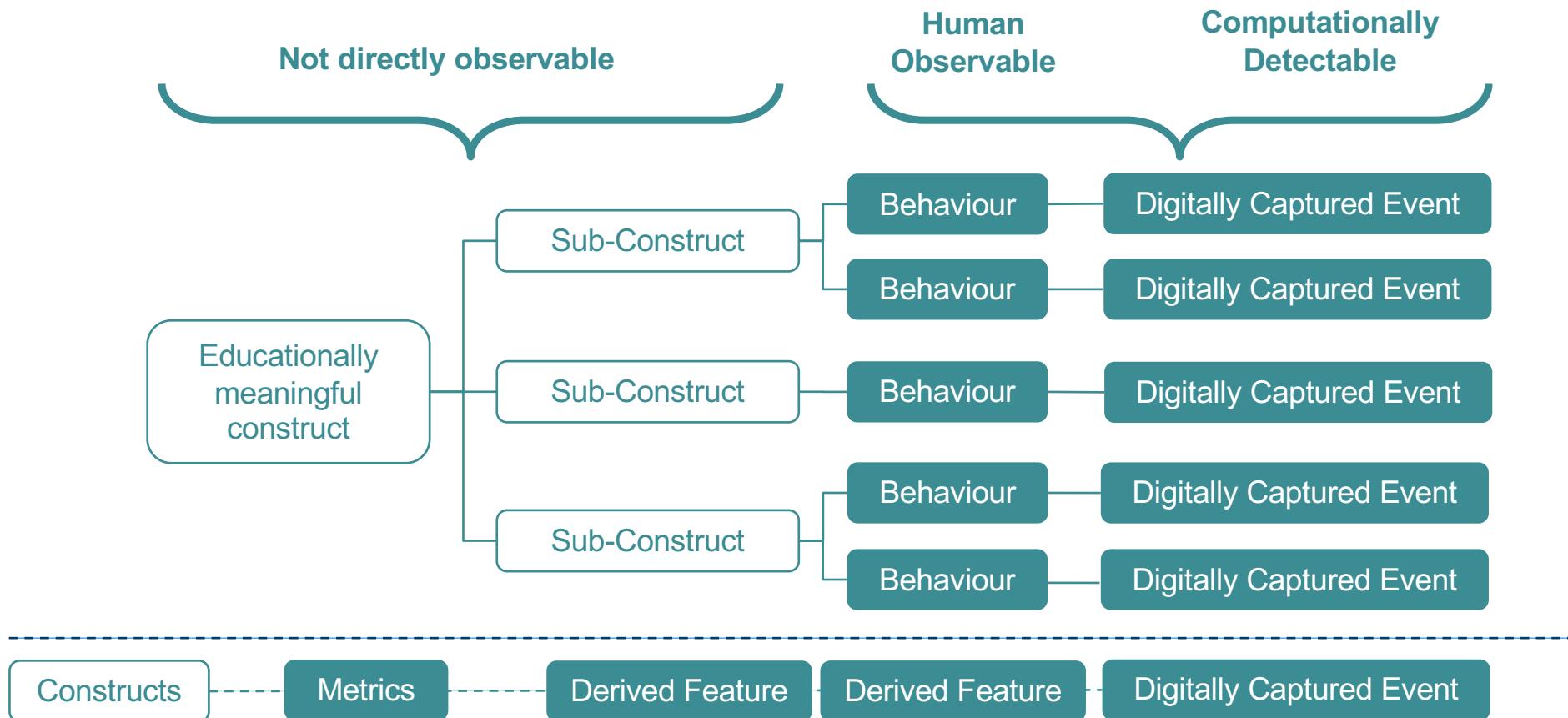
**Impact policy + practice: make the evidence base accessible**

**Share models (and data?)** climate data ≠ Learner data

## Macro

state/national/international networks sharing data, models, scholarship  
→ debate but emerging consensus on optimising learning

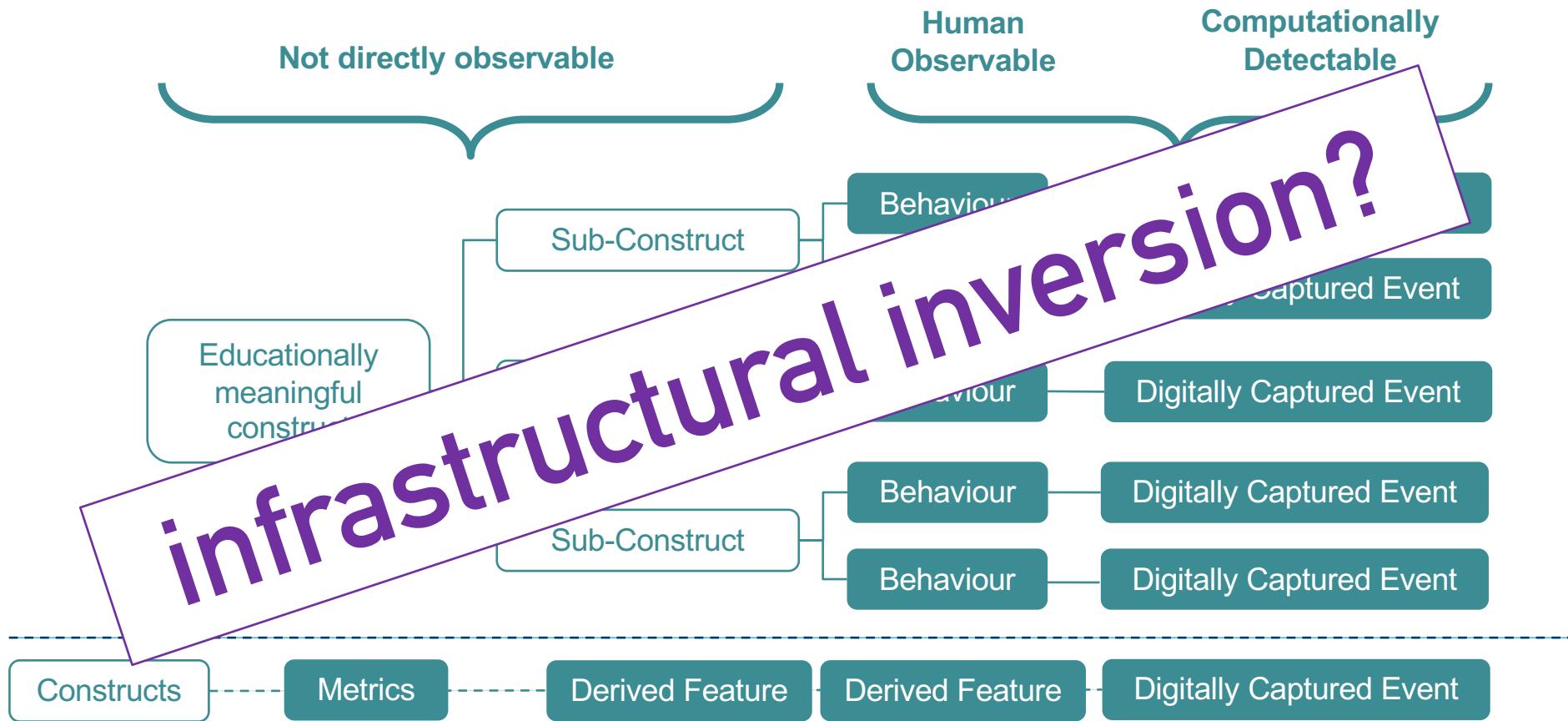
# Ground models in learning sciences + educ. research



Adapted from: Wise, A., Knight, S., Buckingham Shum, S. (In Press) Collaborative Learning Analytics. In: Cress, U., Rosé C., Wise A., & Oshima, J. (Eds.) *International Handbook of Computer-Supported Collaborative Learning*. Springer

See also: Buckingham Shum, S. (2016). *Envisioning C21 Learning Analytics*. Keynote Address, LASI-Asia, Seoul. <https://cic.uts.edu.au/lasi-asia-keynote2016>

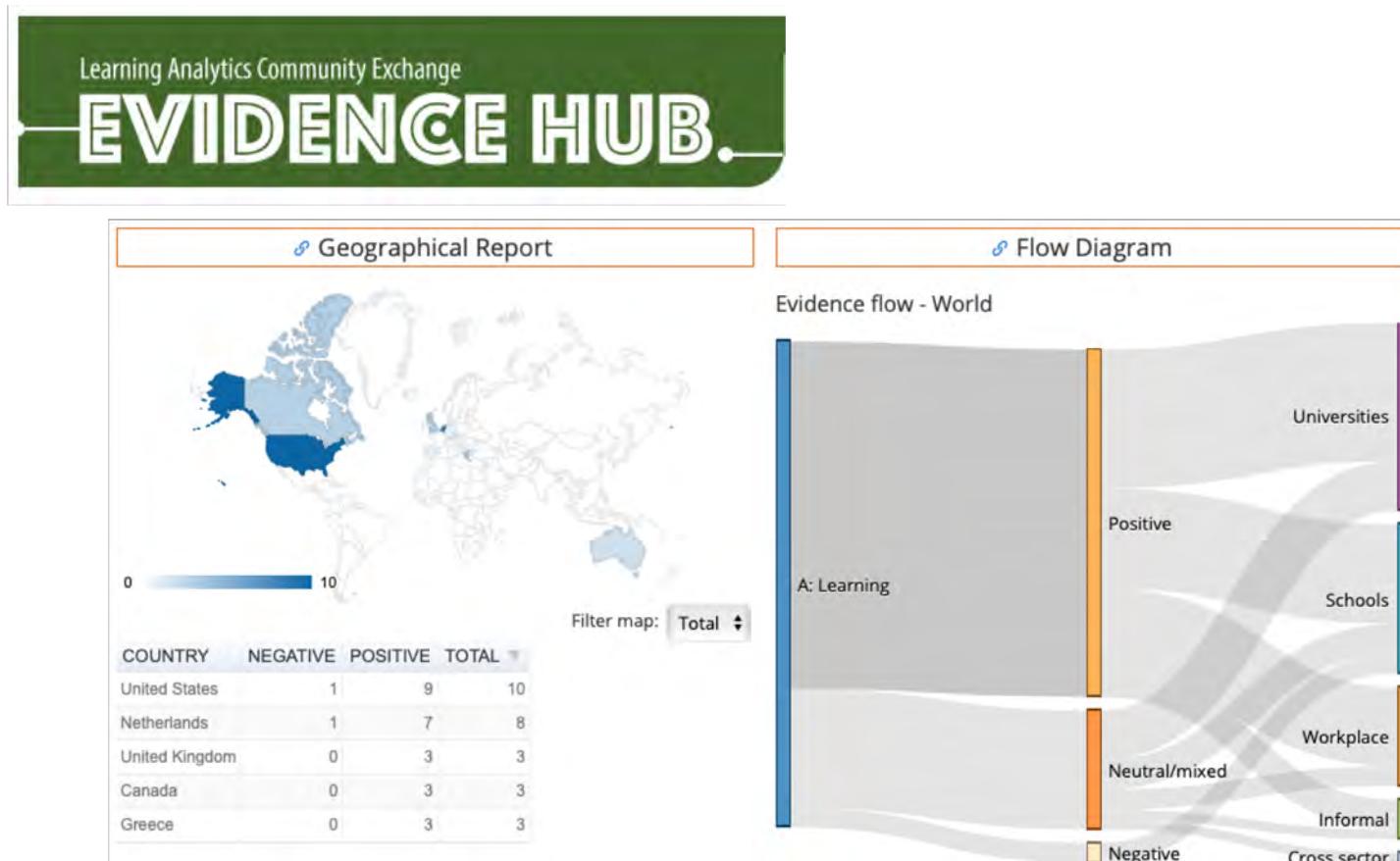
# Ground models in learning sciences + educ. research



Adapted from: Wise, A., Knight, S., Buckingham Shum, S. (In Press) Collaborative Learning Analytics. In: Cress, U., Rosé C., Wise A., & Oshima, J. (Eds.) *International Handbook of Computer-Supported Collaborative Learning*. Springer

See also: Buckingham Shum, S. (2016). *Envisioning C21 Learning Analytics*. Keynote Address, LASI-Asia, Seoul. <https://cic.uts.edu.au/lasi-asia-keynote2016>

# Impact policy + practice: make the evidence base accessible



<http://evidence.laceproject.eu>

In principle, as variation reduces (e.g. timescale, geography, methodology), so do the KI challenges. So MACRO to MESO should help simplify the KI.

But institutions still have long histories

Institutional data and knowledge are still notoriously slippery to curate

And institutionalized teaching practices slow to change

“data management”

“knowledge management”

“progressive pedagogy”

“authentic assessment”

Meso

institution-wide networks [...] optimising learning in the institution

**Nonetheless, it's at the MESO + MICRO layers  
where LA can really add to KI**

Enable data flows  
Tune analytics for the institution's specific needs  
Co-design with stakeholders

## Meso

institution-wide networks [...] optimising learning in the institution

## Micro

assignment/course specific networks [...] optimising learning in a course

# Envisioning the learning ecosystem beyond the LMS, in the wild

**“How are we going to deliver LA over that type of complexity?”**

Kirsty Kitto: Designing Learning Analytics Ecosystems (LASI 2019)  
<https://www.beyondlms.org/blog/LASIworkshop>



Kitto, K., O'Hara, J., Philips, M., Gardiner, G., Ghodrati, M. & Buckingham Shum, S. (2019) The Connected University: Connectedness Learning Across a Lifetime. In Ruth Bridgstock and Neil Tippett (Eds.), *Higher Education and the Future of Graduate Employability: A Connectedness Learning Approach*. <https://doi.org/10.4337/9781788972611>

# Towards LA data flows over an emergent ecosystem: LA-API infrastructure designed for huge diversity in data + analytics

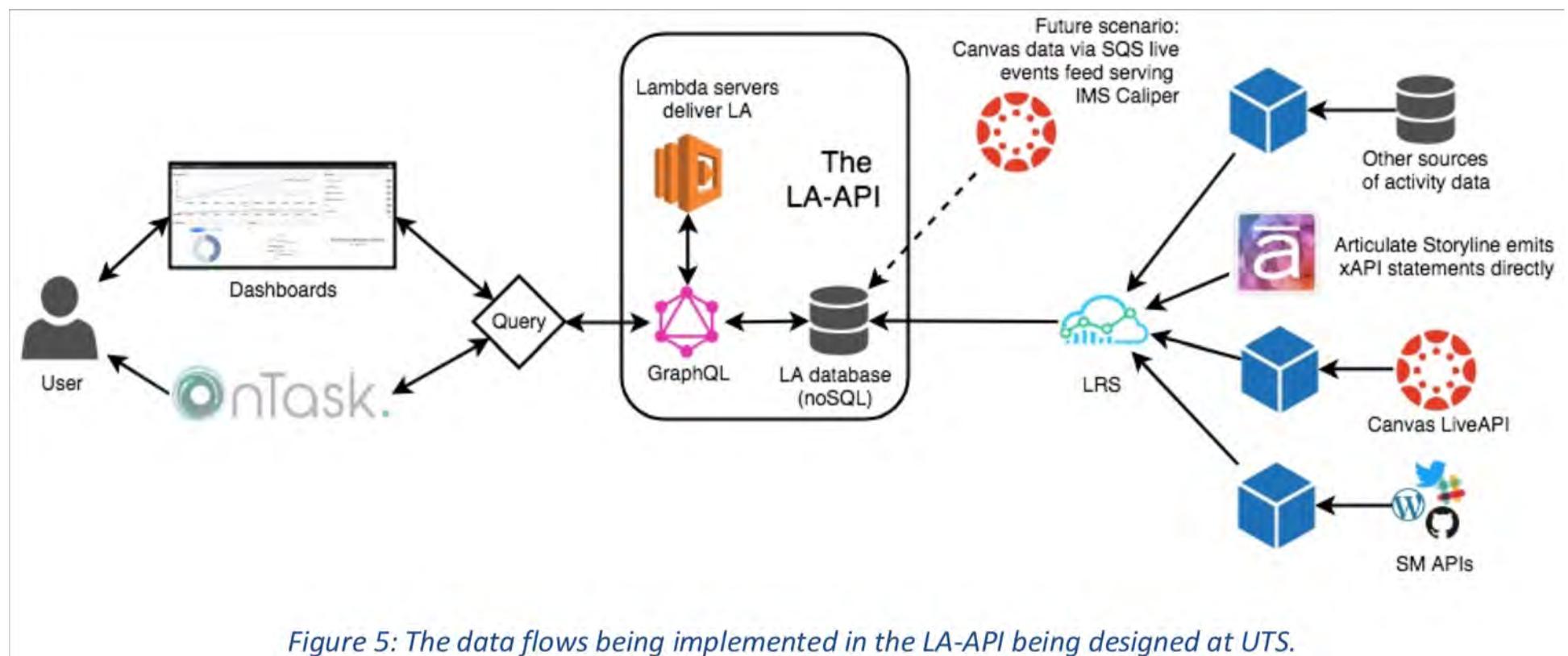


Figure 5: The data flows being implemented in the LA-API being designed at UTS.

Kirsty Kitto, Zak Waters, Simon Buckingham Shum, Mandy Lupton, Shane Dawson, George Siemens (2018): *Learning Analytics Beyond the LMS: Enabling Connected Learning via Open Source Analytics in “the wild”*. Final Report, Office for Learning and Teaching, Australian Government: Canberra. <http://www.beyondlms.org>

# **generalisable models without sacrificing context-sensitivity**

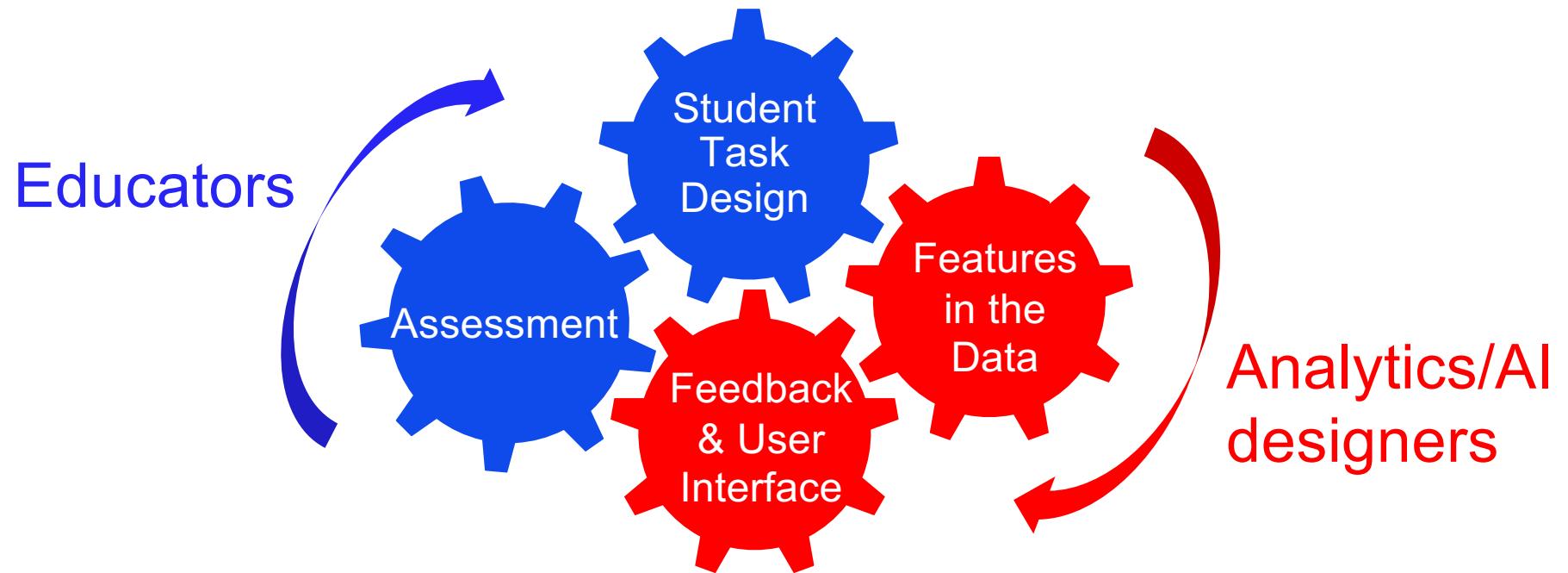
## **Meso**

institution-wide networks [...] optimising learning in the institution

## **Micro**

assignment/course specific networks [...] optimising learning in a course

# Framework @UTS for educators to **co-design** Analytics/AI → **augment teaching practice**



Shibani, A., Knight, S. and Buckingham Shum, S. (2019). Contextualizable Learning Analytics Design: A Generic Model, and Writing Analytics Evaluations. *Proc. 9<sup>th</sup> International Conference on Learning Analytics & Knowledge (LAK19)*. ACM Press, NY, pp. 210-219. DOI: <https://doi.org/10.1145/3303772.3303785>. Eprint: <https://tinyurl.com/lak19clad>

# AcaWriter feedback tuned for Civil Law



**NOTE:** Computers don't understand writing like humans. So, AcaWriter may highlight rhetorically good sentences that actually make no sense, or leave un-highlighted a sentence that you feel is really good. It's fine to disagree with the feedback — but it's also your job to check your facts!

## Analytical Report

## Feedback

## Examples

The analytical report highlights salient rhetorical moves AcaWriter identified in your essay for reflection. For more specific feedback, go to the Feedback tab.

### Rhetorical Moves

- S Summarises or signals the authors goals    P Perspective or stance    E Emphasis to highlight key ideas
- N Novel improvements in ideas    C Contrasting idea, tension or critical insight
- B Background information and previous work    S Surprising or unexpected finding
- Q Question or gap in previous knowledge    T Trend or tendency related to ideas

Technology is an enabler in providing greater access to justice through its ability to connect people with legal needs to legal assistance, information, and advice. T With the increasing popularity of internet-enabled hand held devices and laptop computers, there is a tendency to assume that even the socio-economically vulnerable in our society have access to technology and the skills to use online services with confidence. This is not necessarily the case.

Examples of the application of technology to provide legal information and assistance include case studies, guides and virtual legal advice clinics. S C The 2012 Review does not address the role of courts in serving the legal needs of the community. The court system is not regarded as a part of the wider legal assistance services. C This omission questions the role of the court in facilitating access to its services, including dispute resolution and trials. It also identified uses of technology to expand the delivery of services, many of which are transferable to an online court. These services include e-access for remote communities, availability outside of business hours, interactive processes and virtual appearances. S This essay will discuss uses of technology to expand the delivery of services, many of which are transferable to an online court.

# Building UTS trust with an “AcaWriter micro-KI”

- A pedagogically robust writing exercise was rated significantly more useful with the addition of AcaWriter
- Students who used AcaWriter made significantly more academic rhetorical moves in their revised essays
- A significantly higher proportion of AcaWriter users improved their drafts (many students degraded them across drafts)
- Students who used AcaWriter produced higher graded submissions if they engaged deeply with AcaWriter’s feedback

Shibani, A., Knight, S. and Buckingham Shum, S. (2019). Contextualizable Learning Analytics Design: A Generic Model, and Writing Analytics Evaluations. *Proc. 9<sup>th</sup> International Conference on Learning Analytics & Knowledge (LAK19)*. ACM Press, NY, pp. 210-219. DOI: <https://doi.org/10.1145/3303772.3303785>. Open Access Eprint: <https://tinyurl.com/lak19clad>

Shibani, A. (2019, In Prep). *Augmenting Pedagogic Writing Practice with Contextualizable Learning Analytics*. Doctoral Dissertation, Connected Intelligence Centre, University of Technology Sydney

# Building the AcaWriter micro-KI → educator trust

“Overall, since we’ve been working with CIC around written communication over the course of the last four of five semesters, **we have seen marked improvement in students’ written communication.** Overall their individual assignment pass-rate is going up... We are seeing improvements in the number of students who are either **meeting or exceeding the expectations** around written communication”

# Building the AcaWriter micro-KI → student trust

“It's like having a tutor or another person check and **give constructive feedback on your work.**”

# Building the AcaWriter micro-KI → student trust

“When you’re editing your own writing, **you automatically think that your work sounds good** and that all your ideas and views have been clearly conveyed. This exercise was useful in the sense that it **indicated areas where I needed to be more explicit, which on my own I would not have noticed.**”

# Building the AcaWriter micro-KI → student trust

“I think what is being taught is something I was already aware of. However, by being forced to actually identify **ways of arguing, along with the types of words used to do so**, it has broadened my perspective. I think I will be **more aware of the way I am writing now.**”

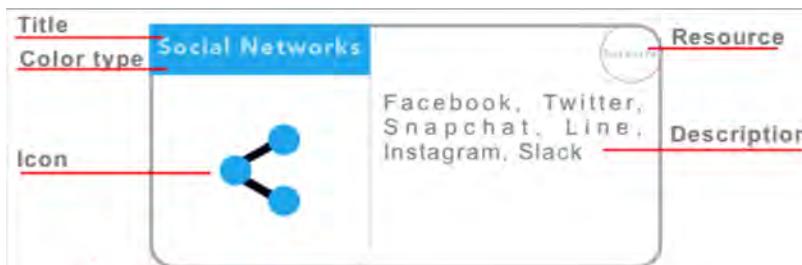
# co-design techniques

educators trust analytics when they can see  
that they're really shaping the design

# Learning Analytics Deck for co-design

<http://ladeck.utscic.edu.au>

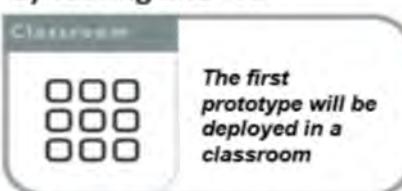
Carlos Prieto's PhD: 'Playing cards' to help stakeholder communication as they design a new kind of analytics tool



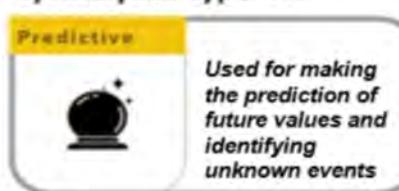
a) Learning objective -LO



b) Testing Site -TS



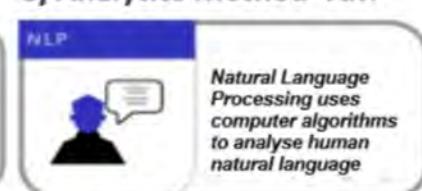
c) Analytics type -AT



d) Data source -DS



e) Analytics Method -AM



f) Privacy - PR



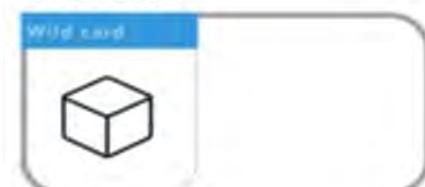
g) User interface -UI



h) Developer tools - Dev



Wild card



Resources

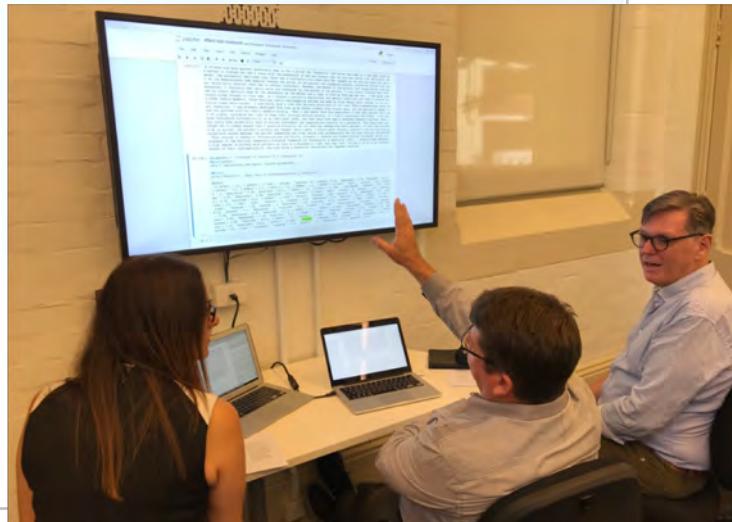


# Co-design with educators to tune writing analytics

jupyter Affect test notebook Last Checkpoint: 11 hours ago (autosaved)

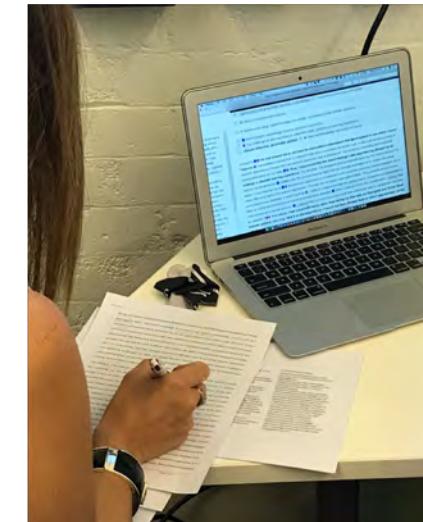
In [103]: `checkAffect(paras[1],4.5)`

```
26 words matched out of 4035 total words in the text - 0.644361833952912 percent
ability      >> 4.85
developed    >> 4.5
male        >> 4.5
taken       >> 4.52
medication   >> 4.56
failed     >> 5.5
learnt      >> 4.8
improved    >> 4.61
disrespectful >> 4.65
learning    >> 4.8
issue       >> 4.55
improve     >> 4.61
rush        >> 6.55
intriguing  >> 5.18
excellent   >> 5.15
impressed   >> 4.82
admiration  >> 5.52
skills      >> 4.94
discovered  >> 5.7
learning    >> 4.8
disrespectful >> 4.65
gave        >> 4.57
positive    >> 5.5
motivated   >> 5.09
```



Goal: calibrate the parser detecting affect in reflective writing, working through sample texts

Rapid prototyping with a Jupyter notebook to agree on thresholds



# More on LA + KI...

(in particular on LA's relationship to the learning sciences)



The slide features a red circular logo for 'ICLS LONDON 2018' on the left. To its right, text reads 'Keynote Address, International Conference of the Learning Sciences London Festival of Learning, June 2018'. On the far right is a black square logo for 'UTS' with a white stylized figure. The main title 'Transitioning Education's Knowledge Infrastructure' is centered in large blue text. Below it, in smaller text, is 'Shaping Design or Shouting from the Touchline?'. At the bottom left is the 'UTS: CIC' logo, and at the bottom right is contact information for Simon Buckingham Shum.

Keynote Address, International Conference of the Learning Sciences  
London Festival of Learning, June 2018

UTS

**Transitioning Education's  
Knowledge Infrastructure**

Shaping Design or Shouting from the Touchline?

**UTS: CIC**  
CONNECTED INTELLIGENCE CENTRE

Simon Buckingham Shum  
@sbuckshum • <http://utscic.edu.au>

<http://simon.buckinghamshum.net/2018/06/icls2018-keynote>

# More on Human-Centred AIED & Learning Analytics...

Collections of insider accounts from teams who are building these infrastructures: how do they engage with issues of epistemology, pedagogy, politics, ethics...?

Human-Centred Learning Analytics. *Journal of Learning Analytics*, 6(2), pp. 1–94 (Eds.) Simon Buckingham Shum, Rebecca Ferguson, & Roberto Martinez-Maldonado

Learning Analytics and AI: Politics, Pedagogy and Practices. *British Journal of Educational Technology (50th Anniversary Special Issue)*, (Eds.) Simon Buckingham Shum & Rose Luckin. (late 2019)

What's the Problem with Learning Analytics? *Journal of Learning Analytics*. Invited Commentaries on Neil Selwyn's LAK18 Keynote Talk, from Carolyn Rosé, Rebecca Ferguson, Paul Prinsloo & Alfred Essa (late 2019)

## Human-Centred Analytics/AI in Education

By sbs on May 4th, 2019 | [Edit](#)

A heads-up that three collections will hit the streets this year focused on how we can design so that human needs and values are well and truly centre-stage in educational tools powered by data, analytics and AI. It will be good to have detailed ‘insider accounts’ from researcher/developers who are reflecting deeply on how values are baked into their design practices and the infrastructures they are building, and how different stakeholders can engage meaningfully in shaping design. I’m excited about the papers shaping up for these volumes, so watch out for their releases mid- and end-2019...

JOURNAL OF  
LEARNING ANALYTICS

BJET | British Journal of  
Educational Technology

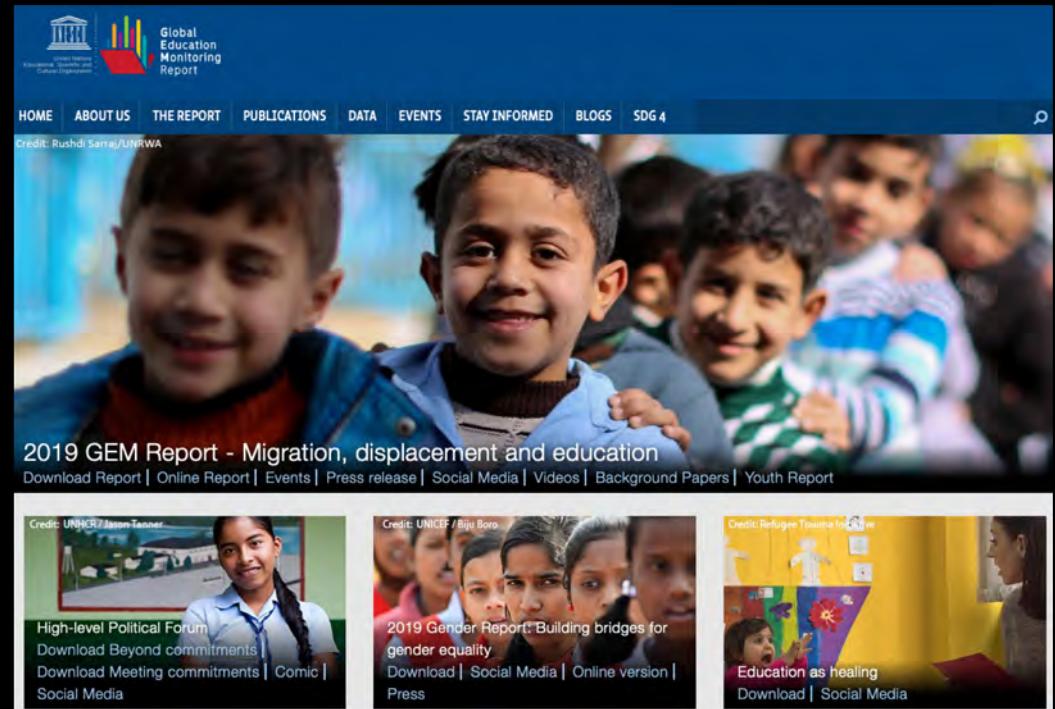
# Reflections on the future educational KI

Are we aspiring for an  
“Intergovernmental Panel on Learning?”

Is part of this already in place?...

UNESCO Global Education Monitoring Report <https://en.unesco.org/gem-report>

A conventional form of educational KI



# Reflections on the future educational KI

Commercial platforms and their R&D programs are 'vertical Knowledge Infrastructures' at national and increasingly international scales

Knowledge about learners from proprietary platforms, primarily ITS (but expanding beyond no doubt)

All the usual questions and concerns around multinational platforms, data ownership, commercial products in education...

The image contains two side-by-side screenshots of educational company websites. The top screenshot is for Squirrel AI Learning, featuring a white header with the logo '松鼠AI·智适应' and 'Squirrel AI Learning'. Below the header, the main title 'What is Squirrel AI Learning?' is displayed in large blue text. A subtext explains they are the first pure-play AI-powered adaptive education provider in China, offering personalized K-12 after-school tutoring at an affordable price. The bottom screenshot is for Pearson, showing a navigation bar with 'PreK-12 Education', 'Higher Education', 'Industry & Professional', 'About Us', and 'Search'. A subpage for 'Data, Analytics & Adaptive Learning' is shown, with a heading 'Looking at the big picture' and a subtext 'helps us personalize a learning path for every student'. To the right is a photo of a teacher and a student looking at a tablet. Below the photo, a section titled 'Data, Analytics, & Adaptive Learning' discusses the importance of data in education.

# Conclusion

We know how a mature, functioning Knowledge Infrastructure operates, and the influence it can have on science, policy and practice (not that this is straightforward)

Insights into KI structure and dynamics should help the LA community focus its efforts to invent an educational KI that can be sustained, and trusted

Your feedback welcomed!

@sbuckshum • [Simon.BuckinghamShum@uts.edu.au](mailto:Simon.BuckinghamShum@uts.edu.au) • [Simon.BuckinghamShum.net](http://Simon.BuckinghamShum.net)