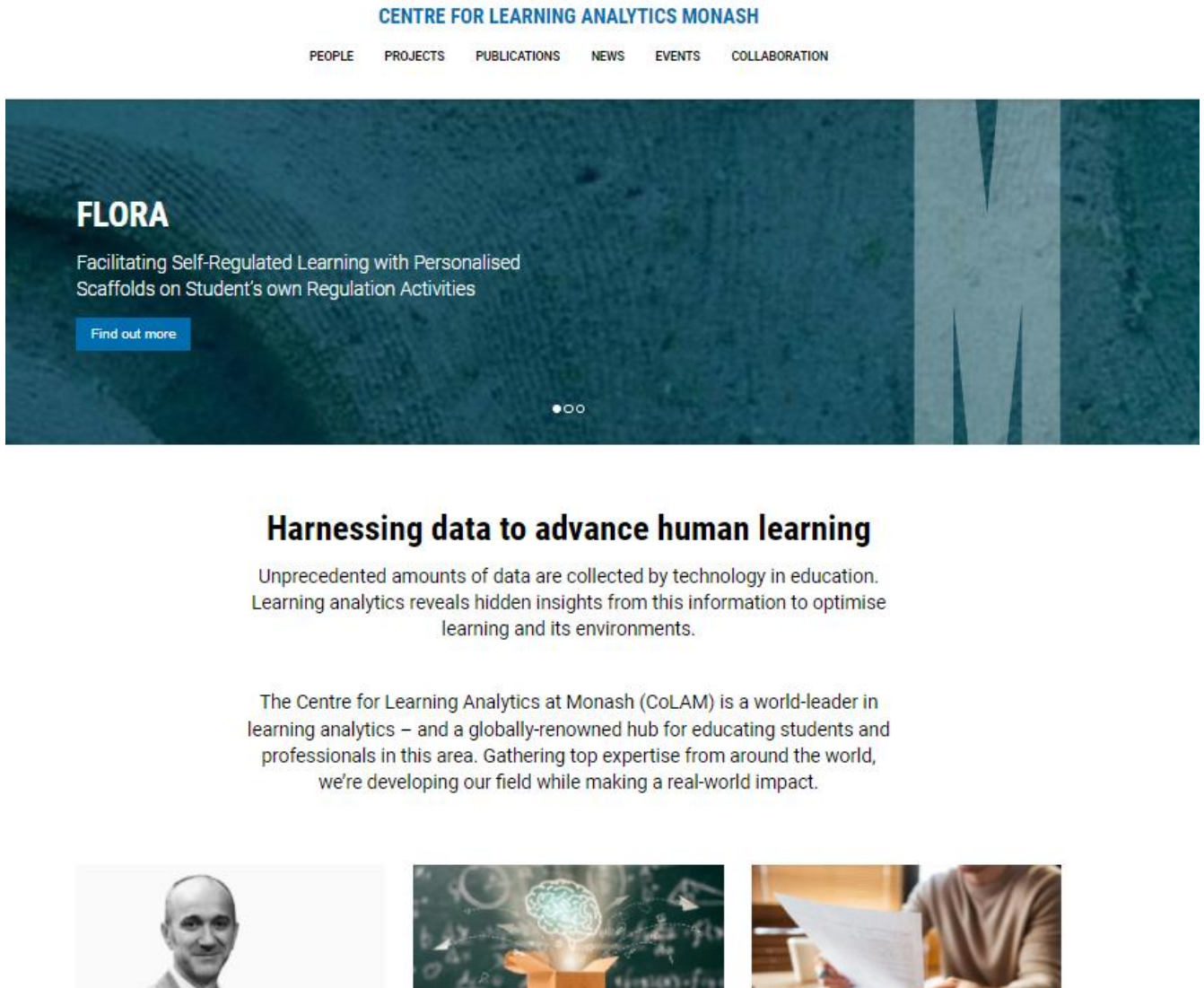
MONASH University | CLAM Centre for Learning Analytics Monash

Directions for dealing with ugly truths about learning analytics

Dragan Gašević
@dgasevic

Jan 11th, 2023
Faculty of Education
University of Hong Kong



Quick facts

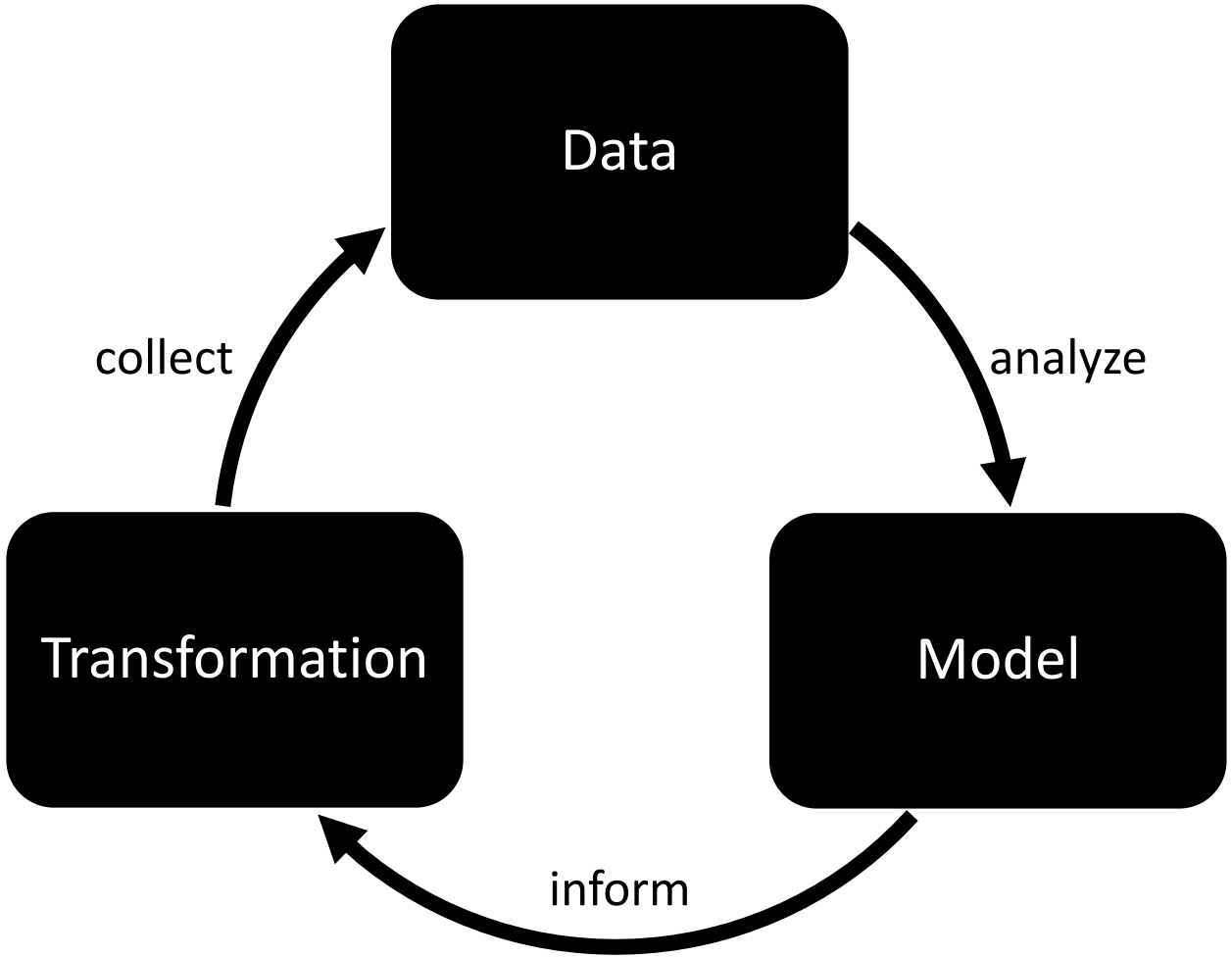
- Largest learning analytics institute in the world
- Over 70 members and affiliates
- Analytics in schools, higher ed, and workplace
- Over 100 partners from all over

https://www.monash.edu/colam

Objective of learning analytics

Using data to establish and enhance feedback loops

Closing the loop



Gašević, D., Tsai, Y-S., Dawson, S., & Pardo, A. (2019). How do we start? An approach to learning analytics adoption in higher education. *International Journal of Information and Learning Technology, 36*(4), 342-353.

Today's talk

Recongizing what we have learned to inform future work

Key takeaways

Data are not flawless

Key takeaways

Applications of AI should be informed by education

Key takeaways

Impact can’t be made by sweeping complexity under the rug

Disclaimer

This will be
a condensed list of ugly troughs

DATA – MODEL – TRANSFORMATION – FINAL REMARKS

Dragan's list of *ugly* truth about data

D1 Datasets are relatively small

D2 Data points are not sampled sufficiently frequently

D3 Measurement validity and reliability not frequently considered

Ugly truth

D1

Datasets are relatively small

Siemens, G. (2013). Learning analytics: The emergence of a discipline. American Behavioral Scientist, 57(10), 1380-1400.

Learning analytics

MOOC platform

MOOC: *Flipped classroom*
12 offerings in 3 years
Selected 292 retakers
177,845 unique events

Raw trace data

15:06:06	/learn/announce
15:07:34	/learn/content
15:10:22	/learn/announce
15:11:01	/learn/content
15:12:27	/learn/content?type=detail&id=1002579286
17:49:58	/info
17:51:44	/learn/announce
17:51:46	/learn/content
17:52:02	/learn/content?type=detail&id=1002579307
17:52:38	/learn/content?type=detail&id=1002579307&cid=1002813724
17:56:32	/learn/content?type=detail&id=1002579307&cid=1002813725
20:44:19	/info

.....

20:44:30	/learn/announce
20:44:32	/learn/content
20:44:34	/learn/content?type=detail&id=1002579275
20:44:41	/learn/score
20:44:41	/learn/custom?id=1002062038
20:44:42	/learn/announce
20:44:44	/learn/content?type=detail&id=1002579275&cid=1002813499
20:44:45	/learn/content?type=detail&id=1002579275&cid=1002813500
11:53:47	/info
11:53:50	/learn/announce
11:53:52	/learn/content?type=detail&id=1002579275&cid=1002813500
10:05:40	/learn/content
10:05:45	/learn/content?type=detail&id=10

Fan, Y., Jovanović, J., Saint, J., Jiang, Y., Wang, Q., & Gašević, D. (2022). Revealing the regulation of learning strategies of MOOC retakers: A learning analytic study. *Computers & Education*, *178*, 104404.

Learning analytics

MOOC platform

MOOC: *Flipped classroom*
12 offerings in 3 years
Selected 292 retakers
177,845 unique events

Raw trace data

15:06:06	/learn/announce
15:07:34	/learn/content
15:10:22	/learn/announce
15:11:01	/learn/content
15:12:27	/learn/content?type=detail&id=1002579286
17:49:58	/info
17:51:44	/learn/announce
17:51:46	/learn/content
17:52:02	/learn/content?type=detail&id=1002579307
17:52:38	/learn/content?type=detail&id=1002579307&cid=1002813724
17:56:32	/learn/content?type=detail&id=1002579307&cid=1002813725
20:44:19	/info

.....

20:44:30	/learn/announce
20:44:32	/learn/content
20:44:34	/learn/content?type=detail&id=1002579275
20:44:41	/learn/score
20:44:41	/learn/custom?id=1002062038
20:44:42	/learn/announce
20:44:44	/learn/content?type=detail&id=1002579275&cid=1002813499
20:44:45	/learn/content?type=detail&id=1002579275&cid=1002813500
11:53:47	/info
11:53:50	/learn/announce
11:53:52	/learn/content?type=detail&id=1002579275&cid=1002813500
10:05:40	/learn/content
10:05:45	/learn/content?type=detail&id=10

609 events/learner for seven weeks,
87 events/learner/week,
many of which are just logins

Fan, Y., Jovanović, J., Saint, J., Jiang, Y., Wang, Q., & Gašević, D. (2022). Revealing the regulation of learning strategies of MOOC retakers: A learning analytic study. *Computers & Education*, *178*, 104404.

Learning analytics

Why not combine data and problem solved?!

Course designs are diverse

Table 5: Blended courses dataset: Course module usages

	ACCT	BIOL 1	BIOL 2	COMM	COMP	ECON	GRAP	MARK	MATH
Assignment	X	X		X	X	X		X	X
Book	X		X			X			
Chat								X	
Course Logins	X	X	X	X	X	X	X	X	X
Feedback			X						
Forum	X	X	X	X	X	X	X	X	X
Gallery	X								
Map			X						
Quiz		X	X		X	X			
Resource	X	X	X	X	X	X	X	X	X
Turnitin	X			X	X	X		X	X
Virtual Classroom			X						

Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. The Internet and Higher Education, 28, 68-84.

Direction

Data collection approaches to capture context

Gašević, D., Greiff, S., Shaffer, D. W. (2022). Towards Strengthening Links between Learning Analytics and Assessment: Challenges and Potentials of a Promising New Bond. *Computers in Human Behavior, 134*, 107304.

Learning analytics

D2

Data points are not sampled sufficiently frequently

Gašević, D., Tsai, Y-S., Dawson, S., & Pardo, A. (2019). How do we start? An approach to learning analytics adoption in higher education. *International Journal of Information and Learning Technology, 36*(4), 342-353.

Learning analytics

Do counts of clicks count for learning?

Jovanović, J., Gašević, D., Pardo, A., Dawson, S., & Whitelock-Wainwright, A. (2019, March). Introducing meaning to clicks: Towards traced-measures of self-efficacy and cognitive load. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 511-520).

Learning analytics

What happens between clicks?

Fan, Y., Lim, L., van der Graaf, J., Kilgour, J., Raković, M., Moore, J., ... & Gašević, D. (2022). Improving the measurement of self-regulated learning using multi-channel data. *Metacognition and Learning*, 1-31.

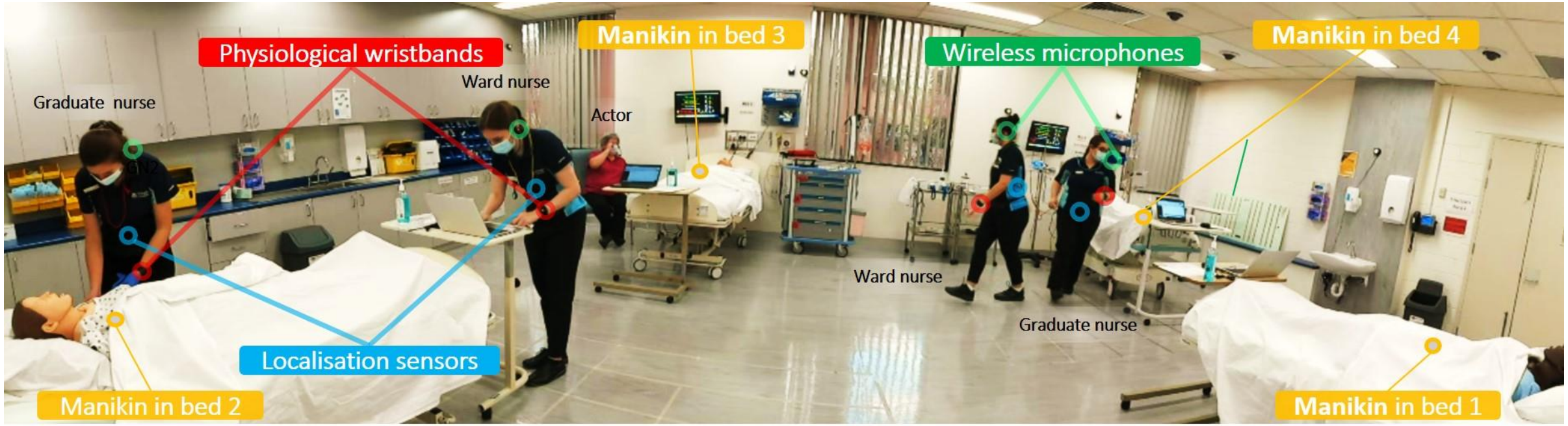
Direction

Using multichannel data

Mouse moves, keystrokes, and eye tracking

Fan, Y., Lim, L., van der Graaf, J., Kilgour, J., Raković, M., Moore, J., ... & Gašević, D. (2022). Improving the measurement of self-regulated learning using multi-channel data. *Metacognition and Learning*, 1-31.

Multimodal sensing technologies for teamwork



Echeverria, V., Martinez-Maldonado, R., Yan, L., Zhao, L., Gašević, D., Fernandez-Nieto, G., Buckingham Shum, S. (2023). HuCETA: A Framework for Human-Centred Embodied Teamwork Analytics. *IEEE Pervasive Computing*, in press.

Learning analytics

D3

Measurements are not necessarily valid and reliable

Winne, P. H. (2020). Construct and consequential validity for learning analytics based on trace data. *Computers in Human Behavior, 112*, 106457.

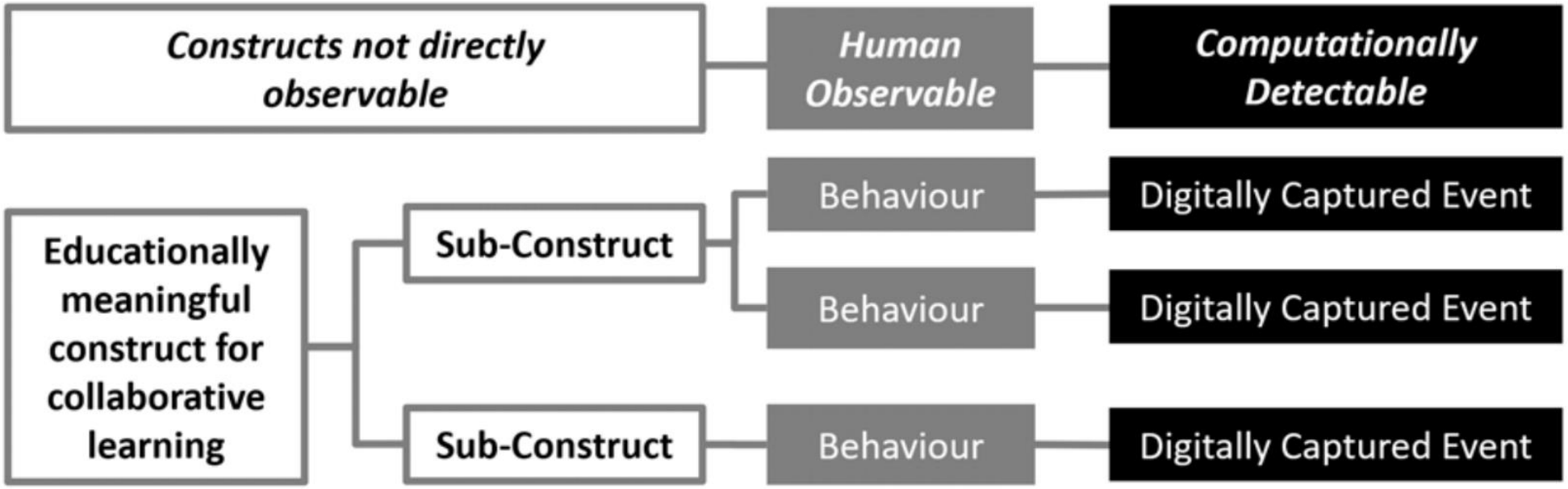
Learning analytics

What's the meaning of digital traces?

Jovanović, J., Gašević, D., Pardo, A., Dawson, S., & Whitelock-Wainwright, A. (2019). Introducing meaning to clicks: Towards traced-measures of self-efficacy and cognitive load. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 511-520).

Direction

From data to constructs



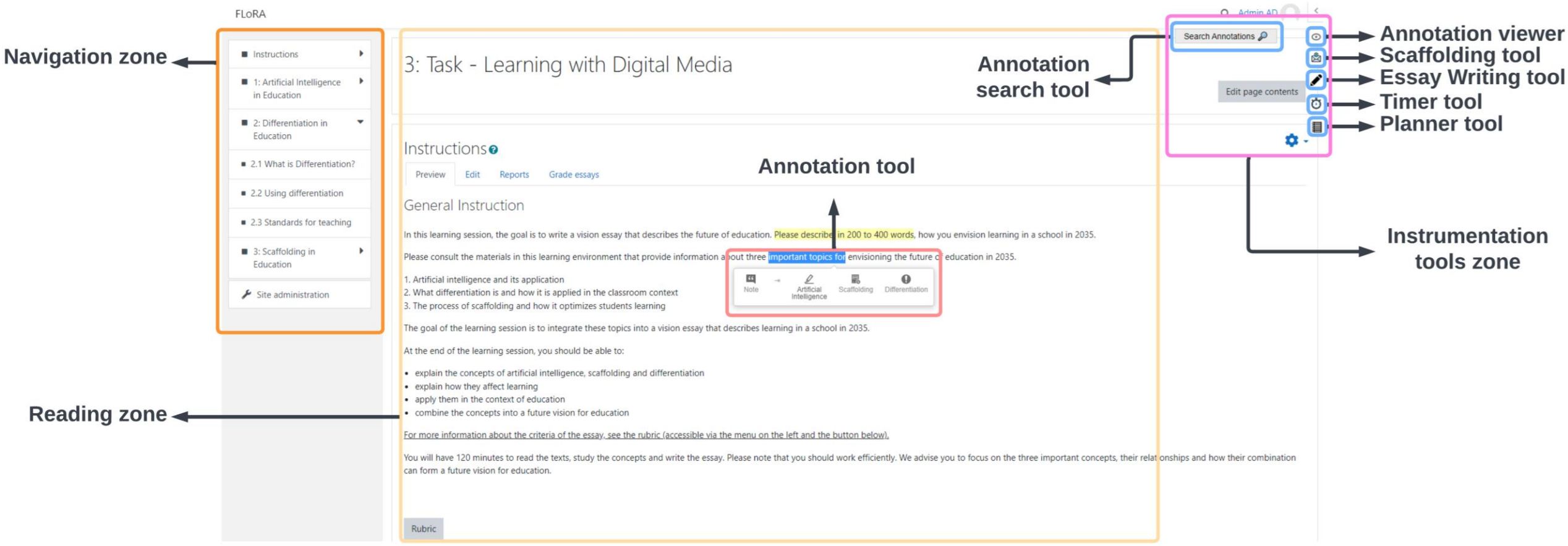
Martinez-Maldonado, R., Gaševic, D., Echeverria, V., Fernandez Nieto, G., Swiecki, Z., & Buckingham Shum, S. (2021). What Do You Mean by Collaboration Analytics? A Conceptual Model. *Journal of Learning Analytics, 8*(1), 126-153.

Direction

Pedagogically-valuable instrumentation tools

van der Graaf, J., Lim, L., Fan, Y., Kilgour, J., Moore, J., Bannert, M., ... & Molenaar, I. (2021). Do instrumentation tools capture self-regulated learning?. In *Proceedings of the 11th international learning analytics and knowledge conference* (pp. 438-448).

Instrumentation tools

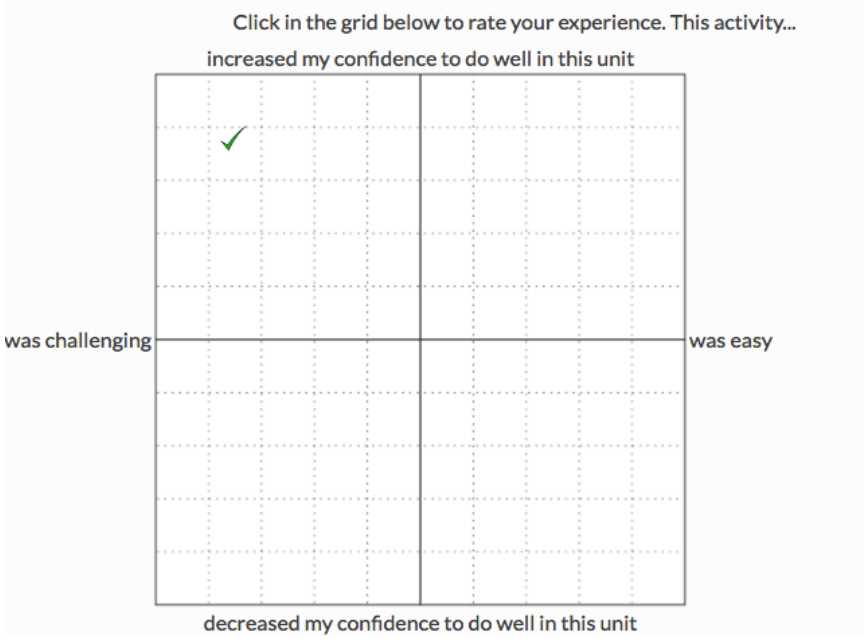


moodle + hypothes.is + extensions

https://floraproject.org

Introduction meaning to clicks

Just-in-time self-reporting



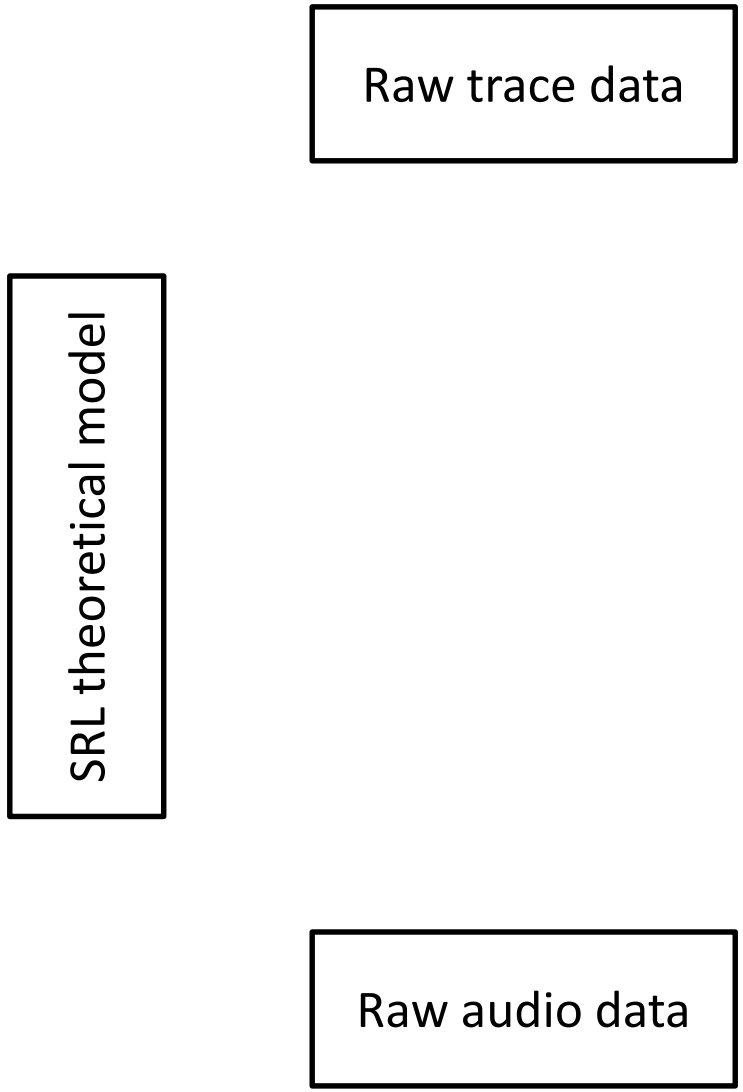
Jovanović, J., Gašević, D., Pardo, A., Dawson, S., & Whitelock-Wainwright, A. (2019). Introducing meaning to clicks: Towards traced-measures of self-efficacy and cognitive load. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 511-520).

Direction

Using other data sources as reference points for measurement validation

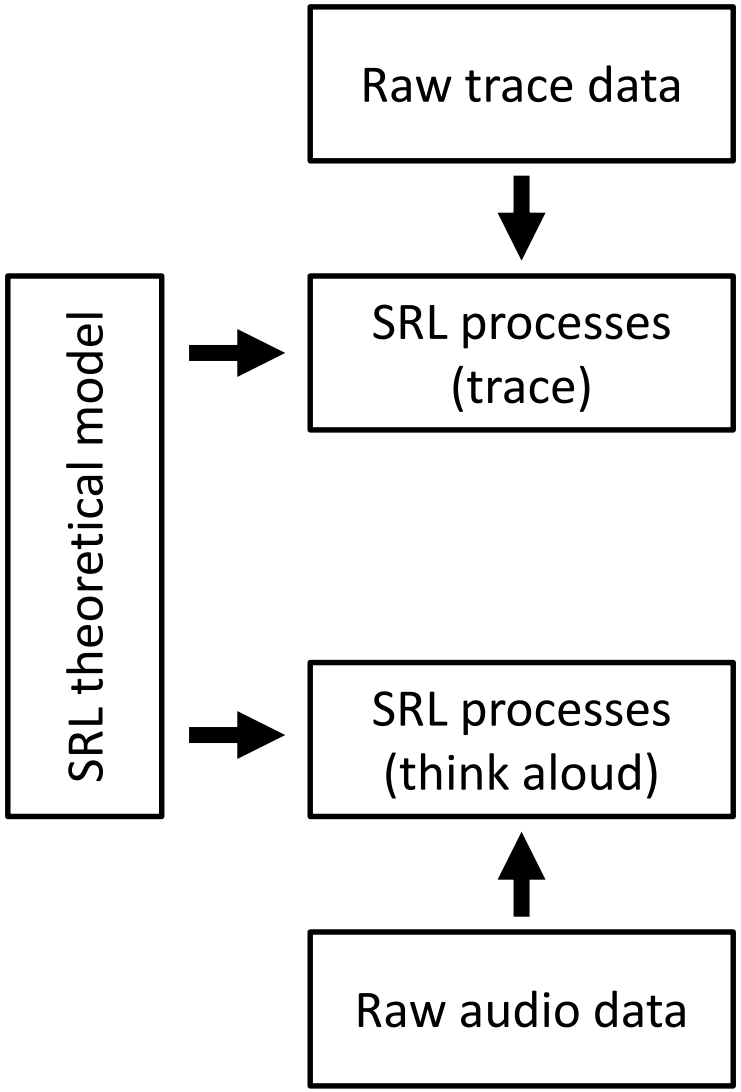
Fan, Y., van der Graaf, J., Lim, L., Raković, M., Singh, S., Kilgour, J., ... & Gašević, D. (2022). Towards investigating the validity of measurement of self-regulated learning based on trace data. *Metacognition and Learning*, in press.

Theory- and data-driven approach



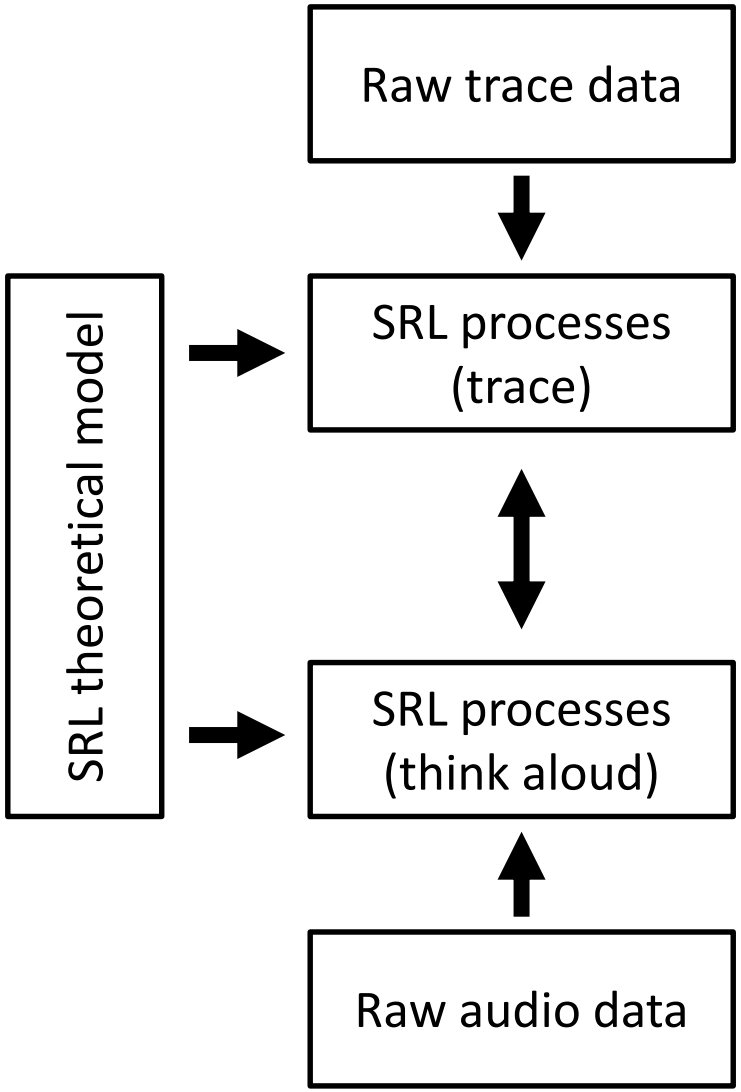
Fan, Y., van der Graaf, J., Lim, L., Raković, M., Singh, S., Kilgour, J., ... & Gašević, D. (2023). Towards investigating the validity of measurement of self-regulated learning based on trace data. *Metacognition and Learning*, in press.

Theory- and data-driven approach



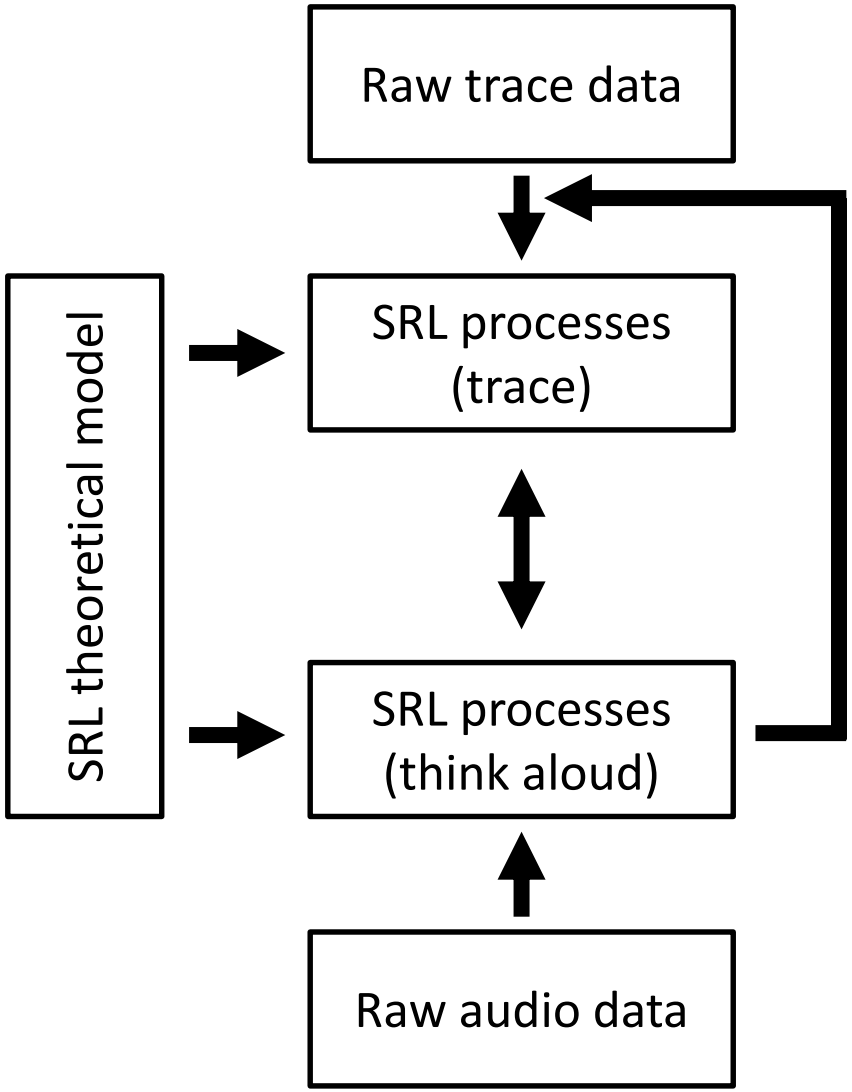
Fan, Y., van der Graaf, J., Lim, L., Raković, M., Singh, S., Kilgour, J., ... & Gašević, D. (2023). Towards investigating the validity of measurement of self-regulated learning based on trace data. *Metacognition and Learning*, in press.

Theory- and data-driven approach



Fan, Y., van der Graaf, J., Lim, L., Raković, M., Singh, S., Kilgour, J., ... & Gašević, D. (2023). Towards investigating the validity of measurement of self-regulated learning based on trace data. *Metacognition and Learning*, in press.

Theory- and data-driven approach



Fan, Y., van der Graaf, J., Lim, L., Raković, M., Singh, S., Kilgour, J., ... & Gašević, D. (2023). Towards investigating the validity of measurement of self-regulated learning based on trace data. *Metacognition and Learning*, in press.

COLAM

DATA– **MODEL** – TRANSFORMATION –
FINAL REMARKS

Dragan's list of *ugly* truth about models

- M1 Generalizability of models is limited
- M2 Obsession with predictive accuracy, bias rarely measured
- M3 Insufficient focus on causality – mostly correlational
- M4 Explainability is essential for education

Learning context

Generalizability of models is limited

Gašević, D., Dawson, S., Rogers, T., Gašević, D. (2016). Learning analytics should not promote one size fits all: The effects of course-specific technology use in predicting academic success. *The Internet and Higher Education, 28*, 68–84.

What shapes generalizability?

Instructional conditions shape learning analytics results

Gašević, D., Dawson, S., Rogers, T., Gašević, D. (2016). Learning analytics should not promote one size fits all: The effects of course-specific technology use in predicting academic success. *The Internet and Higher Education, 28*, 68–84.

What shapes generalizability?

Students matter the most in learning analytics

Jovanović, J., Saqr, M., Joksimović, S., & Gašević, D. (2021). Students matter the most in learning analytics: The effects of internal and instructional conditions in predicting academic success. *Computers & Education, 172*, 104251.

Direction

Analytics of learning strategies

COLAM

Direction

Analytics of learning strategies

Unsupervised machine learning

+

Sequence mining

Process mining

Network analysis

Key findings (1/2)

Analytics of learning strategies

Regulation of strategies is consistent with relevant theory

Gašević, D., Jovanović, J., Pardo, A., & Dawson, S. (2017). Detecting Learning Strategies with Analytics: Links with Self-reported Measures and Academic Performance. *Journal of Learning Analytics*, *4*(2), 113–128.

Key findings (2/2)

Analytics of learning strategies

Strategies are predictive of academic performance

Fincham, O. E., Gašević, D., Jovanovic, J. M., & Pardo, A. (2019). From Study Tactics to Learning Strategies: An Analytical Method for Extracting Interpretable Representations. *IEEE Transactions on Learning Technologies*, *12*(1), 59–72. https://doi.org/10.1109/TLT.2018.2823317

Ultimate goal

Models of *individual* learners

An idiographic approach
Identify learning signatures of individual learners

Malmberg, J., Saqr, M., Järvenoja, H., & Järvelä, S. (2022). How the monitoring events of individual students are associated with phases of regulation: A network analysis approach. *Journal of Learning Analytics, 9*(1), 77-92.

Ugly truth

Obsession with predictive accuracy, bias rarely measured

Gardner, J., Brooks, C., & Baker, R. (2019). Evaluating the fairness of predictive student models through slicing analysis. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 225-234).

Direction

Protective Attributes	Relevant Studies				# Papers
	Included		Not included	Unknown	
	Detection	Enhancement			
Gender/Sex	[1, 6, 7, 33, 38, 59, 78] [13, 32, 61, 69, 77]	[15, 31, 40–45] [47, 49, 55, 59, 69, 72]	[21, 25, 56] [46, 60, 79]	[14, 34, 63, 64, 66, 68, 71]	37
Race/Ethnicity	[6, 13, 32, 38, 77, 78]	[5, 9, 31, 41–45]	[46, 56, 79]	[14, 17, 34, 68]	21
Geographic/Region/Country	[6, 33]	[15]	[26, 27, 52, 53, 62]		8
Age/Year-of-Birth	[1, 6, 33, 59]	[15, 59]		[71]	6
Disability	[6, 33, 59]	[59]	[56]		4
First-language background	[61]			[17, 64, 66]	4
Income	[78]	[49, 55]	[3]		4
Parents' education level	[32, 61]			[34, 71]	4
Others	[6, 19, 33, 59, 61, 69, 77, 78]	[49, 59, 69]	[46, 56]	[14, 71]	15
# Papers	13	17	12	9	

Li, L., Sha, L., Li, Y., Rakovic, M., Rong, J., Joksimovic, S., Selwyn, N., Gašević, D., Chen, G. (2023). Moral Machines or Tyranny of the Majority? A Systematic Review on Predictive Bias in Education. In *Proceedings of the 13th International Conference on Learning Analytics and Knowledge* (in press).

Direction

Types	Fairness Metric	Educational Tasks							# Papers
		Admission	Recommendation	Forum	Assessment	Performance	Dropout	Other tasks	
Group	Equal accuracy				[46]	[6, 15, 33] [17, 78]	[14, 17, 34, 61, 77]	[3, 52, 53, 62, 79]	16
	ABROCA		[44, 45]	[63, 64, 66]	[21]	[32, 59, 64]	[25, 61, 64]		10
	Statistical parity	[49]	[7, 72]		[68]	[17, 31, 33, 71]	[17, 69]		9
	Disparate impact	[9]	[7, 26, 27]		[68]	[17, 59]	[17, 69]	[60]	9
	Equal Opportunities				[68]	[6, 17, 33] [71, 78]	[17, 61, 69]	[47]	9
	Equalized Odds		[41, 45]		[13, 68]	[17, 42, 43]	[17, 69]		8
	Others				[5, 46, 56]	[59, 71, 78]			6
Individual	Consistency		[48]		[18]	[17, 31]	[17]		4
	Counterfactual fairness					[38]			1
Other metrics						[17, 55]	[17]	[19, 40]	4
	# Papers	2	8	3	7	14	8	9	

Li, L., Sha, L., Li, Y., Rakovic, M., Rong, J., Joksimovic, S., Selwyn, N., Gašević, D., Chen, G. (2023). Moral Machines or Tyranny of the Majority? A Systematic Review on Predictive Bias in Education. In *Proceedings of the 13th International Conference on Learning Analytics and Knowledge* (in press).

Direction

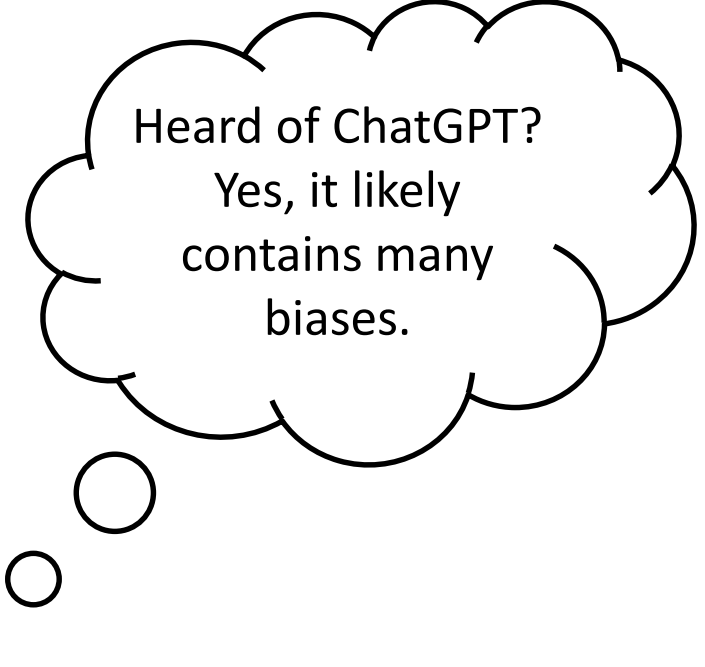
Dataset balancing approaches

Pre-processing approaches

Sha, L., Raković, M., Das, A., Gašević, D., & Chen, G. (2022). Leveraging Class Balancing Techniques to Alleviate Algorithmic Bias for Predictive Tasks in Education. *IEEE Transactions on Learning Technologies, 15*(4), 481-492.

COLAM

Direction



Bias in pretrained language models

Pre-processing approaches

Sha, L., Li, Y., Gasevic, D., & Chen, G. (2022). Bigger Data or Fairer Data? Augmenting BERT via Active Sampling for Educational Text Classification. In *Proceedings of the 29th International Conference on Computational Linguistics* (pp. 1275-1285).

COLAM

Fundamental challenge

Can bias in data ever by removed?

Ugly truth

Insufficient focus on causality, mostly correlational

Weidlich, J., Gašević, D., & Drachsler, H. (2022). Causal Inference and Bias in Learning Analytics: A Primer on Pitfalls Using Directed Acyclic Graphs. *Journal of Learning Analytics, 9*(3), 183-199.

Direction

Methods that can model causality and bias

Weidlich, J., Gašević, D., & Drachsler, H. (2022). Causal Inference and Bias in Learning Analytics: A Primer on Pitfalls Using Directed Acyclic Graphs. *Journal of Learning Analytics, 9*(3), 183-199.

COLAM

Direction

Supplanting models with qualitative insights

Quantitative ethnography

Andres, J. M., Hutt, S., Ocumpaugh, J., Baker, R. S., Nasiar, N., & Porter, C. (2021). How Anxiety Affects Affect: A Quantitative Ethnographic Investigation Using Affect Detectors and Data-Targeted Interviews. In *Proceedings of the 3rd International Conference on Quantitative Ethnography* (pp. 268-283).

Ugly truth

Explainability is essential for education

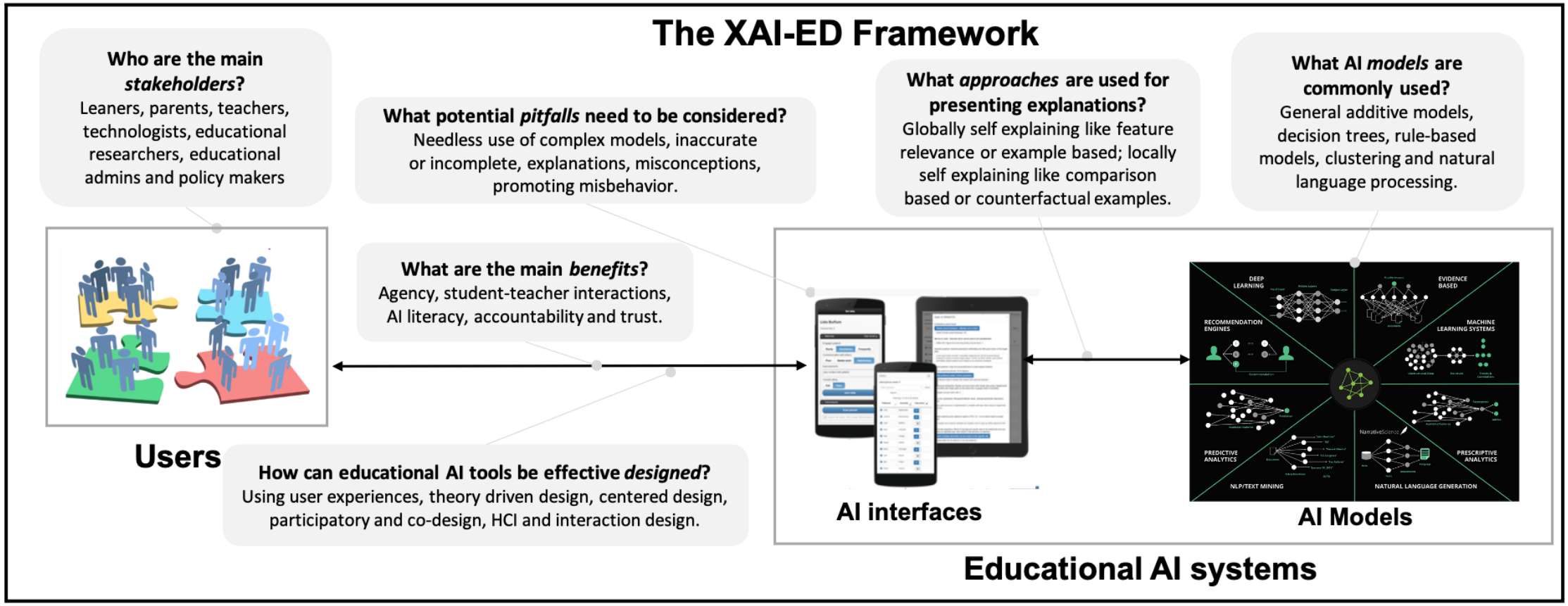
Deep learning is not a prevalent technology

Sha, L., Raković, M., Lin, J., Guan, Q., Whitelock-Wainwright, A., Gašević, D., & Chen, G. (2023). Is the Latest the Greatest? A Comparative Study of Automatic Approaches for Classifying Educational Forum Posts. *IEEE Transactions on Learning Technologies*, in press.

Determining when explainability is and isn't critical

Baker, R. S., Gašević, D., & Karumbaiah, S. (2021). Four paradigms in learning analytics: Why paradigm convergence matters. *Computers and Education: Artificial Intelligence*, *2*, 100021.

Direction



Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y. S., Kay, J., ... & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *3*, 100074.

DATA– MODEL – **TRANSFORMATION** –
FINAL REMARKS

Dragan's list of *ugly* truth about transformation

T1 Needs analysis insufficiently performed

T2 Analytics is not an impartial black-box to outsource accountability

T3 Limited understanding analytics affects stakeholders

T4 Social and organizational complexity shape adoption and impact

Ugly truth

T1

Needs analysis insufficiently performed

Bodily, R., & Verbert, K. (2017). Review of research on student-facing learning analytics dashboards and educational recommender systems. *IEEE Transactions on Learning Technologies, 10*(4), 405-418.

Learning analytics

Dashboards can be harmful

Matcha, W., Ahmad Uzir, N., Gašević, D., Pardo, A. (2020). A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective. *IEEE Transactions on Learning Technologies, 13*(2), 226 - 245.

Direction

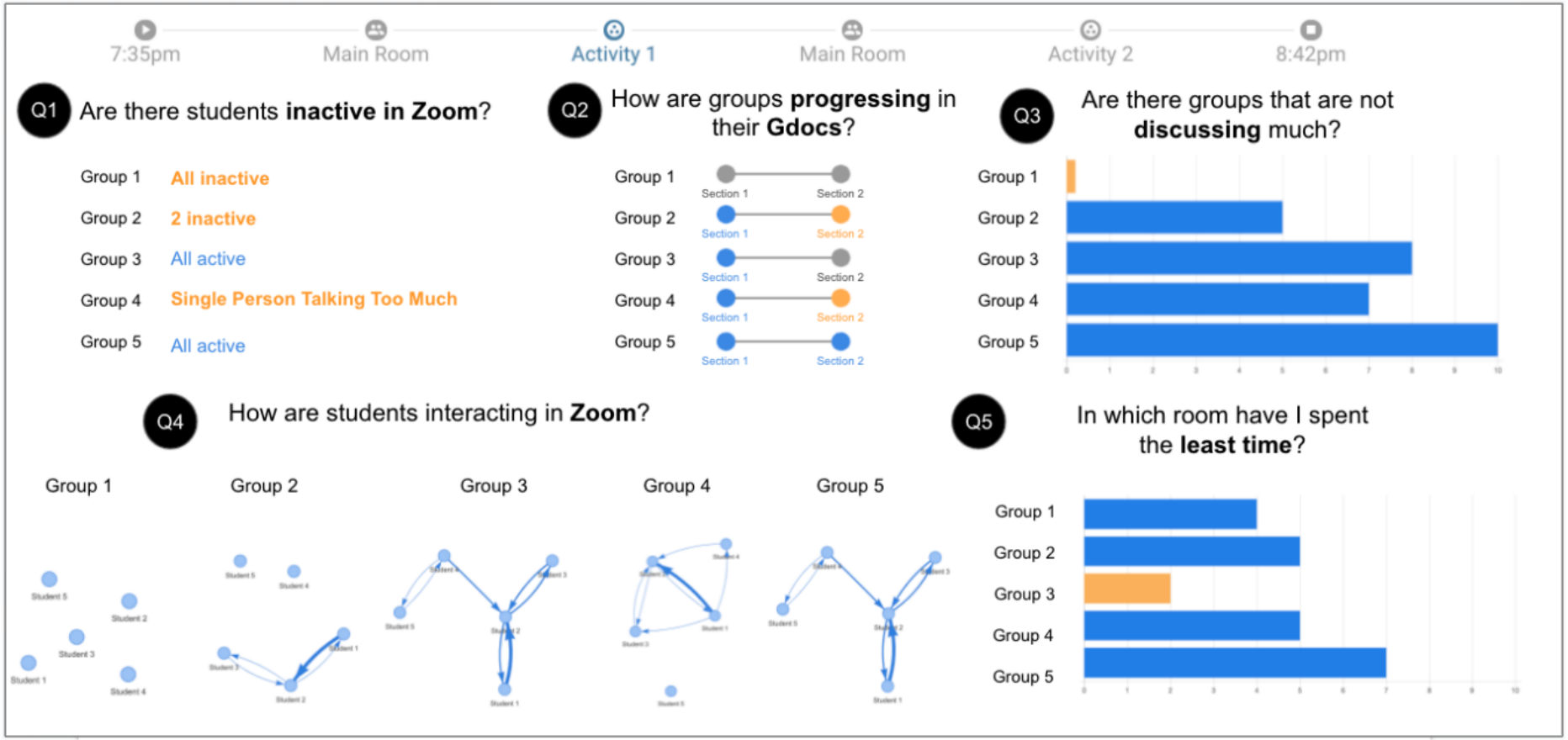
Participatory and co-design is gaining momentum

Sarmiento, J. P., & Wise, A. F. (2022). Participatory and Co-Design of Learning Analytics: An Initial Review of the Literature. In *Proceedings of the 12th International Learning Analytics and Knowledge Conference* (pp. 535-541).

ZoomSense

https://zoomsense.io

Question-driven data storytelling



Pozdniakov, S., Martinez-Maldonado, R., Tsai, Y. S., Cukurova, M., Bartindale, T., Chen, P., ... & Gasevic, D. (2022). The Question-driven Dashboard: How Can We Design Analytics Interfaces Aligned to Teachers' Inquiry?. In *Proceedings of the 12th International Learning Analytics and Knowledge Conference* (pp. 175-185).

Ugly truth

T2

Analytics is not an impartial black-box to outsource accountability

Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., ... & Gašević, D. (2022). Assessment in the age of artificial intelligence. *Computers and Education: Artificial Intelligence*, *3*, 100075.

Dilemma

Who controls decision-making – humans vs machines?

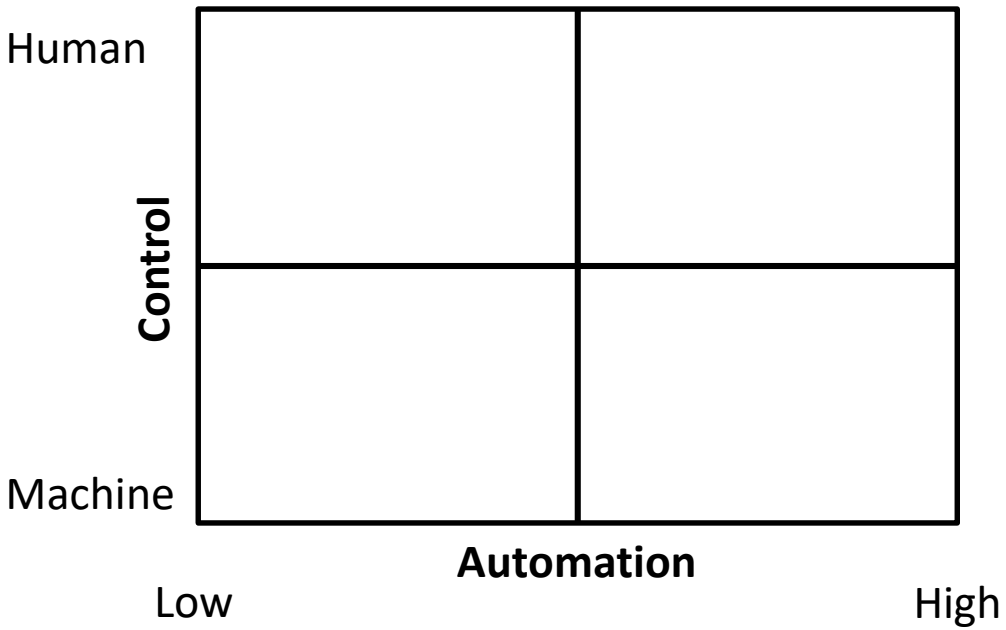
Shneiderman, B. (2022). *Human-Centered Artificial Intelligence*. Oxford University Press

Direction

Analytics in the loop
(human is already in the loop!)

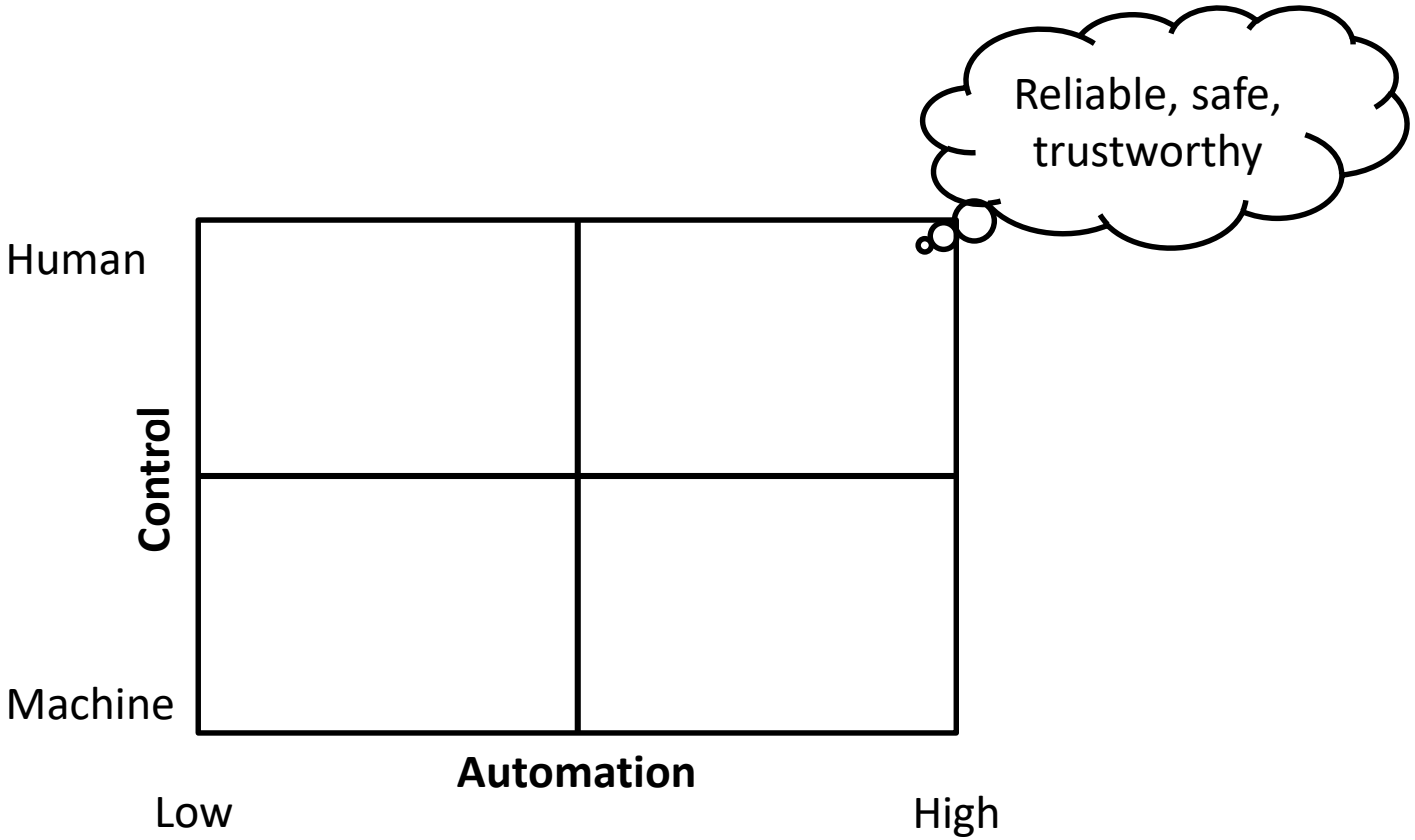
Shneiderman, B. (2022). *Human-Centered Artificial Intelligence*. Oxford University Press

Direction



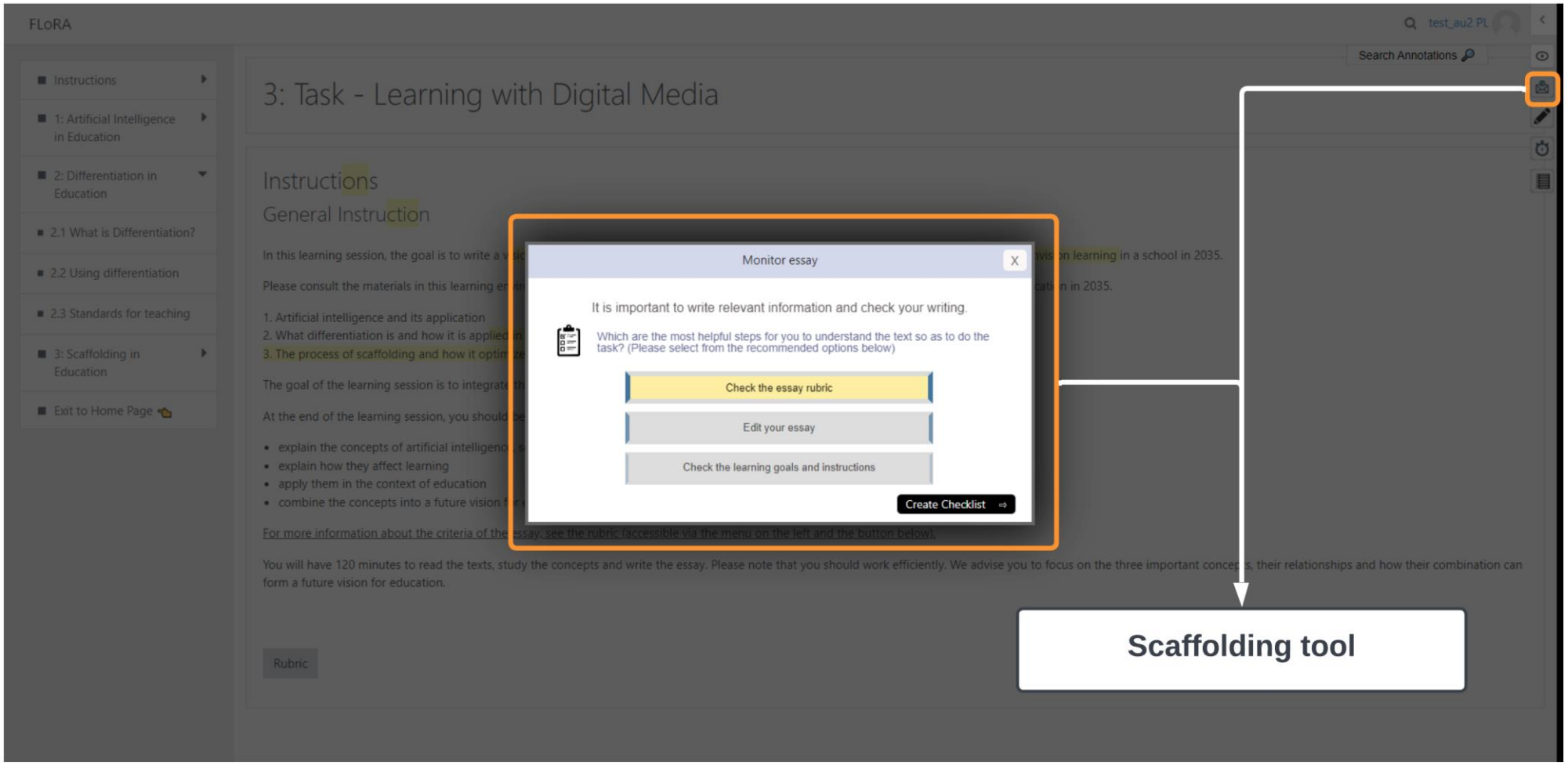
Shneiderman, B. (2022). *Human-Centered Artificial Intelligence*. Oxford University Press

Direction



Shneiderman, B. (2022). *Human-Centered Artificial Intelligence*. Oxford University Press

Analytics-based personalized scaffolding



Srivastava, N., Fan, Y., Rakovic, M., Singh, S., Jovanovic, J., Van Der Graaf, J., ... & Gasevic, D. (2022). Effects of Internal and External Conditions on Strategies of Self-regulated Learning: A Learning Analytics Study. In *Proceedings of 12th International Learning Analytics and Knowledge Conference* (pp. 392-403).

Direction

Novel human-centred design methods

Shneiderman, B. (2022). *Human-Centered Artificial Intelligence*. Oxford University Press

Ugly truth

T3

Limited understanding analytics affects stakeholders

Fan, Y., Li, T., Tsai, Y-S., Rakovic, M., Singh, S., Li, X.,… Gašević, D. (2022). How learners perceive and benefit from personalised SRL scaffoldings: a qualitative study. *Journal of Computer Assisted* learning, submitted.

Responsibility

How do stakeholders sensemake data about diversity, equity, and inclusion?

Williamson, K., & Kizilcec, R. F. (2021). Learning Analytics Dashboard Research Has Neglected Diversity, Equity and Inclusion. In *Proceedings of the Eighth ACM Conference on Learning@ Scale* (pp. 287-290).

Responsibility

Who takes responsibility for actions taken based on analytics?

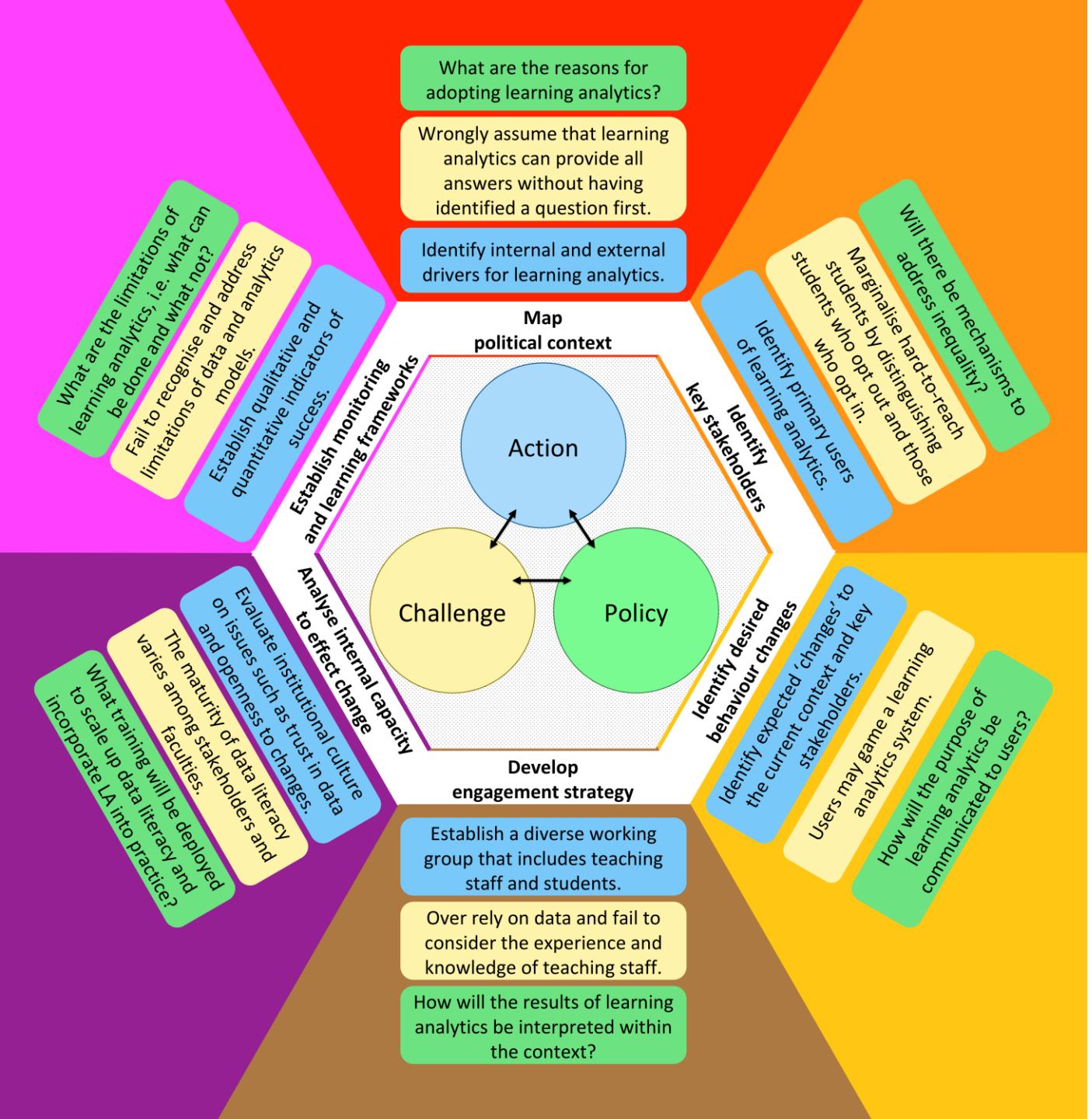
Williamson, K., & Kizilcec, R. F. (2021). Learning Analytics Dashboard Research Has Neglected Diversity, Equity and Inclusion. In *Proceedings of the Eighth ACM Conference on Learning@ Scale* (pp. 287-290).

Ugly truth

T4

Social and organizational complexity shape adoption and impact

Macfadyen, L. P., Dawson, S., Pardo, A., & Gašević, D. (2014). Embracing big data in complex educational systems: The learning analytics imperative and the policy challenge. Research & Practice in Assessment, 9, 17-28.



Learning Technologist of the Year 2019
Winner

ALT

SHEILA framework

Formation of institutional strategies and policies for learning analytics

https://sheilaproject.eu/

T4

Direction

Adaptive complex systems and leadership

Tsai, Y. S., Poquet, O., Gašević, D., Dawson, S., & Pardo, A. (2019). Complexity leadership in learning analytics: Drivers, challenges and opportunities. *British Journal of Educational Technology*, *50*(6), 2839-2854.

COLAM

DATA– MODEL – TRANSFORMATION –

FINAL REMARKS

We need to embrace and address limitations of data we are using

AI can offer many good things, but educational needs come first

Impact can be achieved by accepting social-technical complexity

MONASH University | CLAM Centre for Learning Analytics Monash

Directions for dealing with ugly truths about learning analytics

Dragan Gašević
@dgasevic

Jan 11th, 2023
Faculty of Education
University of Hong Kong