

HUDK 5053: FEATURE ENGINEERING STUDIO

5/2/19 1:40 PM

Today

- Photo release
- Intro to Methodology Mapping
- Methodology Mapping Exercise
- Mystery Unit Work

Next Week

- Tuesday: Workshop on Mystery Unit
- Thursday: Mystery Unit Due 5:00pm
 - ❖ (52% Chose Github)

Mapping Methodologies In Learning Analytics

Purpose

- Introduce the breadth of methodology in the field of learning analytics
- Demonstrate that it is a work in progress
- Have you participate in a research project



Yoav Bergner, NYU

Geraldine Grey, ITB

Document, organize & critique the use of
methods within learning analytics and
educational data science



Learning
(science of)

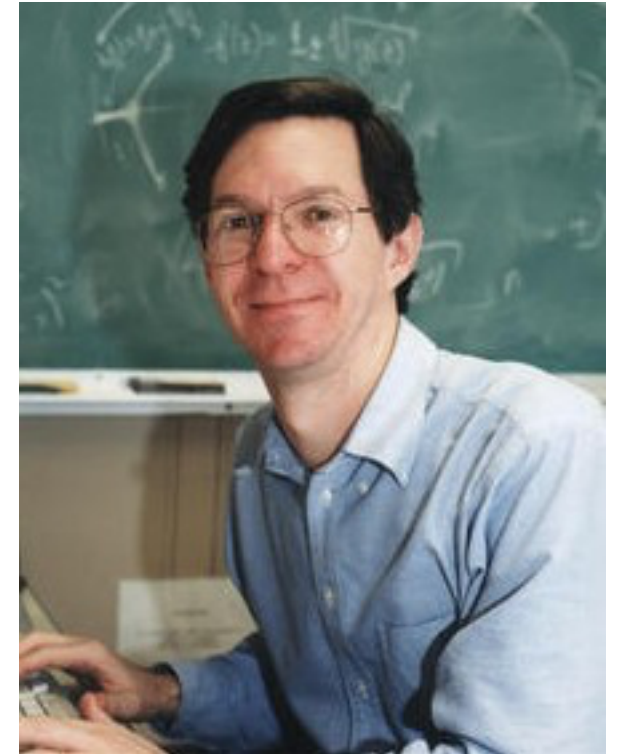
Analytics
(methods of)

© 2009 M. Stone
Victoria Falls Bridge (Copyright)

Victoria Falls Bridge (Construction)

The Sokal Affair

- Sokal, A. D. (1996). Transgressing the boundaries: Toward a transformative hermeneutics of quantum gravity. *Social text*, (46/47), 217-252.
- The paper was a parody/hoax designed, at least in part, to embarrass its target. It succeeded.
- Sokal also wanted to make his own mark on the “two cultures” problem (Snow, 1959) and the need for two-way understanding.



Science Wars. As part of the campaign against “political correctness,” the history and theory of science studies is increasingly subject to intense political scrutiny. In this special issue edited by Andrew Ross, many of the leading figures in the social and cultural study of science respond to recent debates in the field. Contributors to this issue include Sandra Harding, Steve Fuller, Emily Martin, Hilary Rose, Langdon Winner, Dorothy Nelkin, George Levine, Sharon Traweek, Sarah Franklin, Ruth Hubbard, Richard Levins, Joel Kovel, Stanley Aronowitz, Andrew Ross, Les Levidow, and Alan Sokal. **Spring/Summer 1996**

SOCIAL

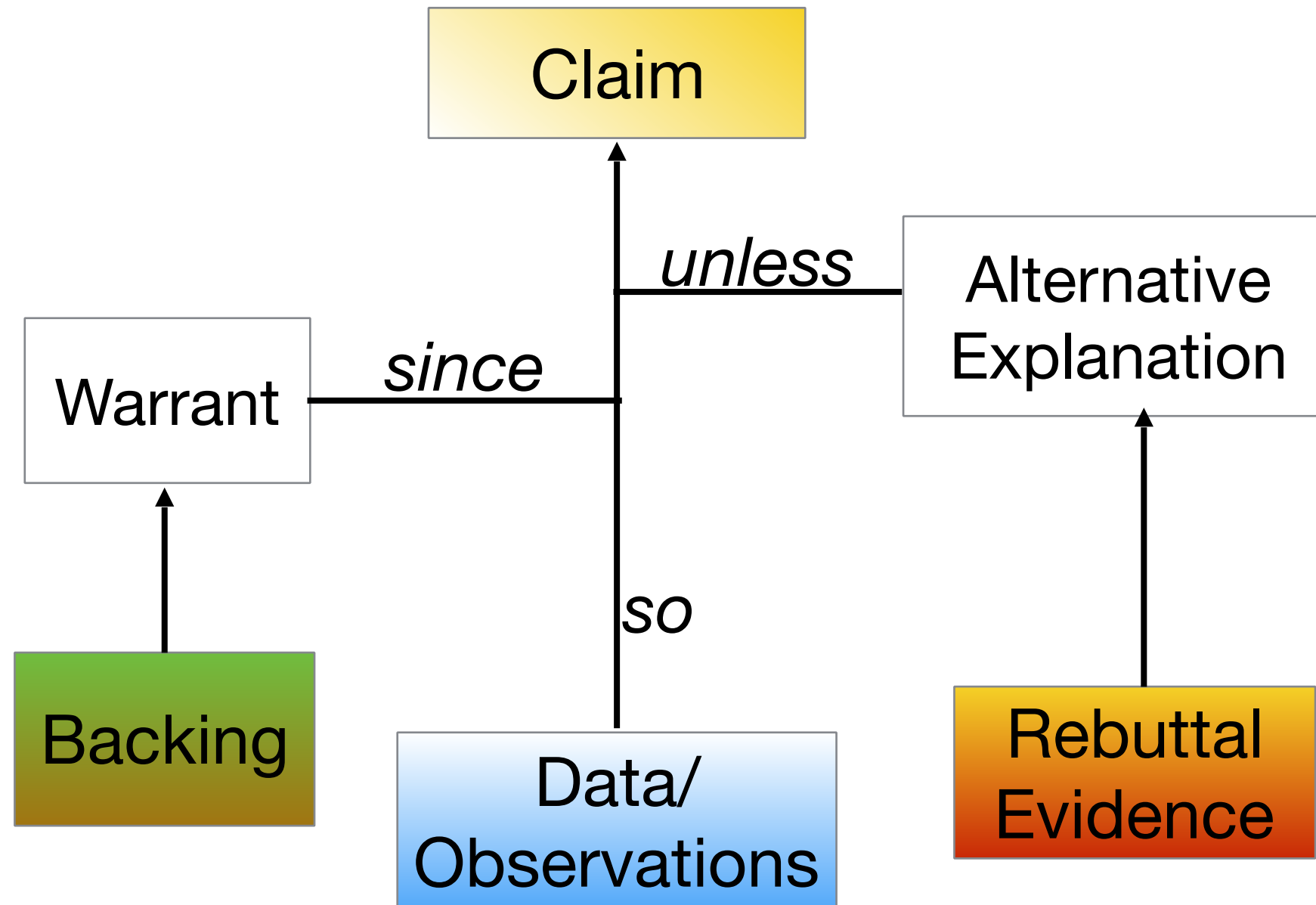
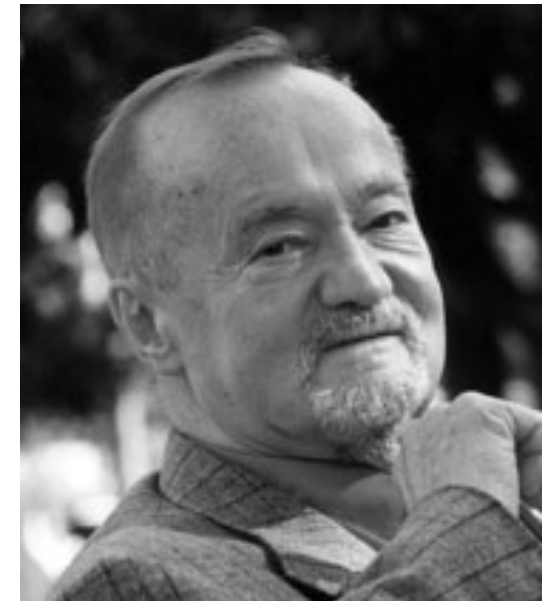
46-47

TEXT



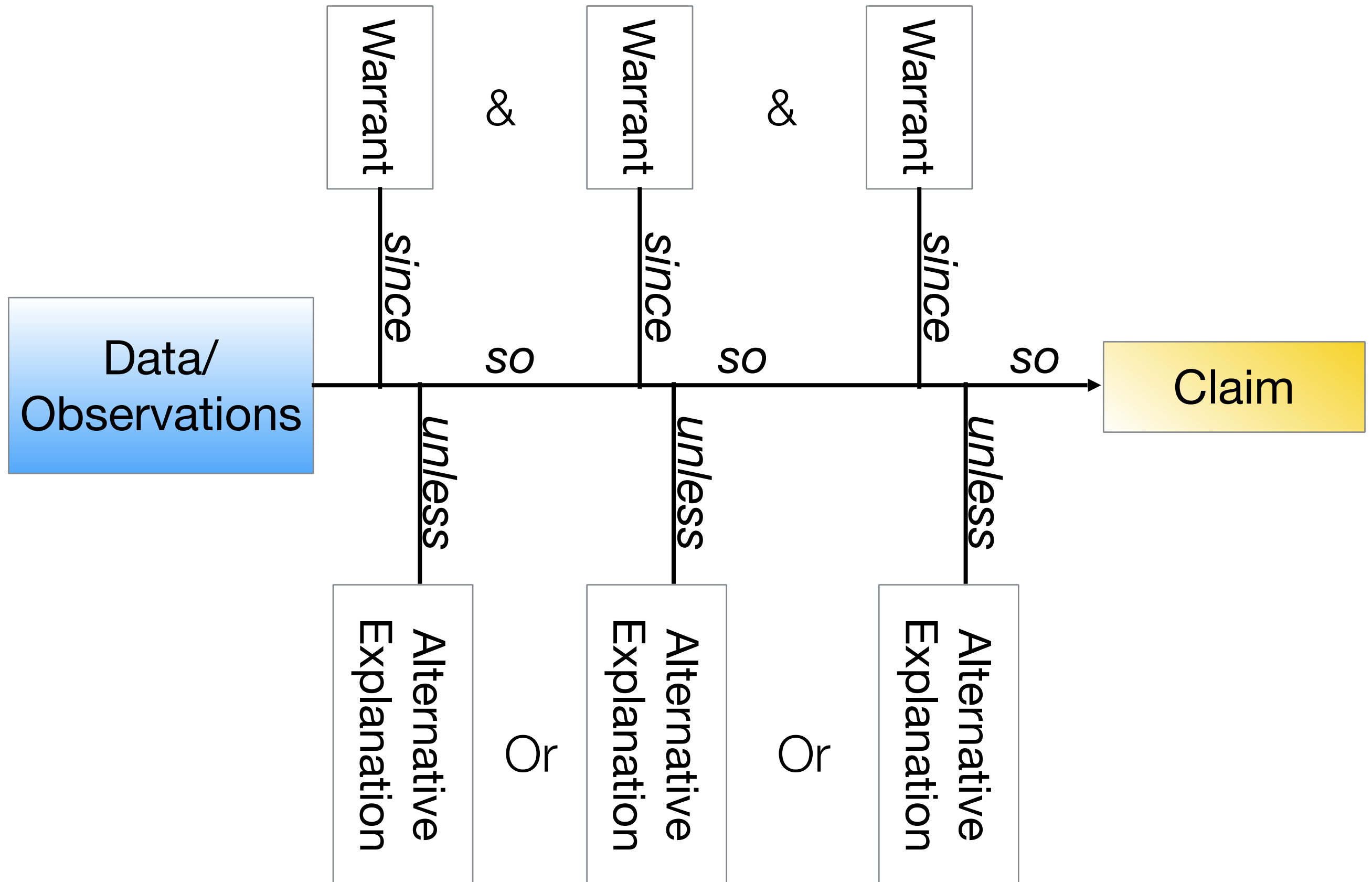
Sokal's Bridge Project

Papers are arguments



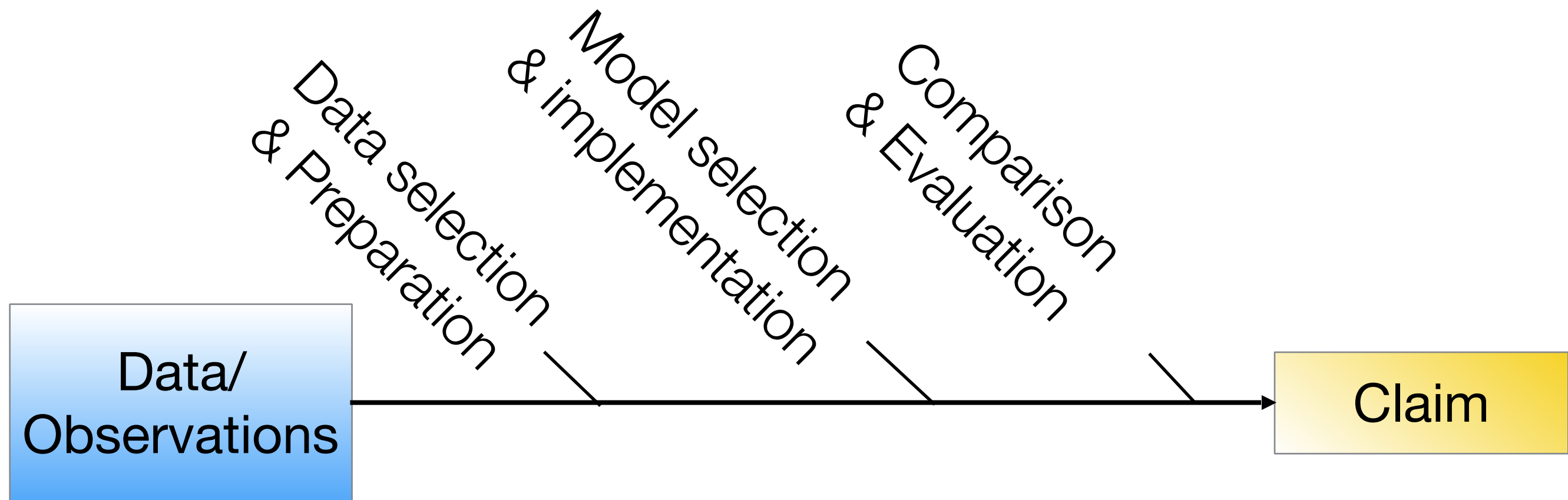
Toulmin (1958) Argument Model

Many arguments involve chains of sub-claims



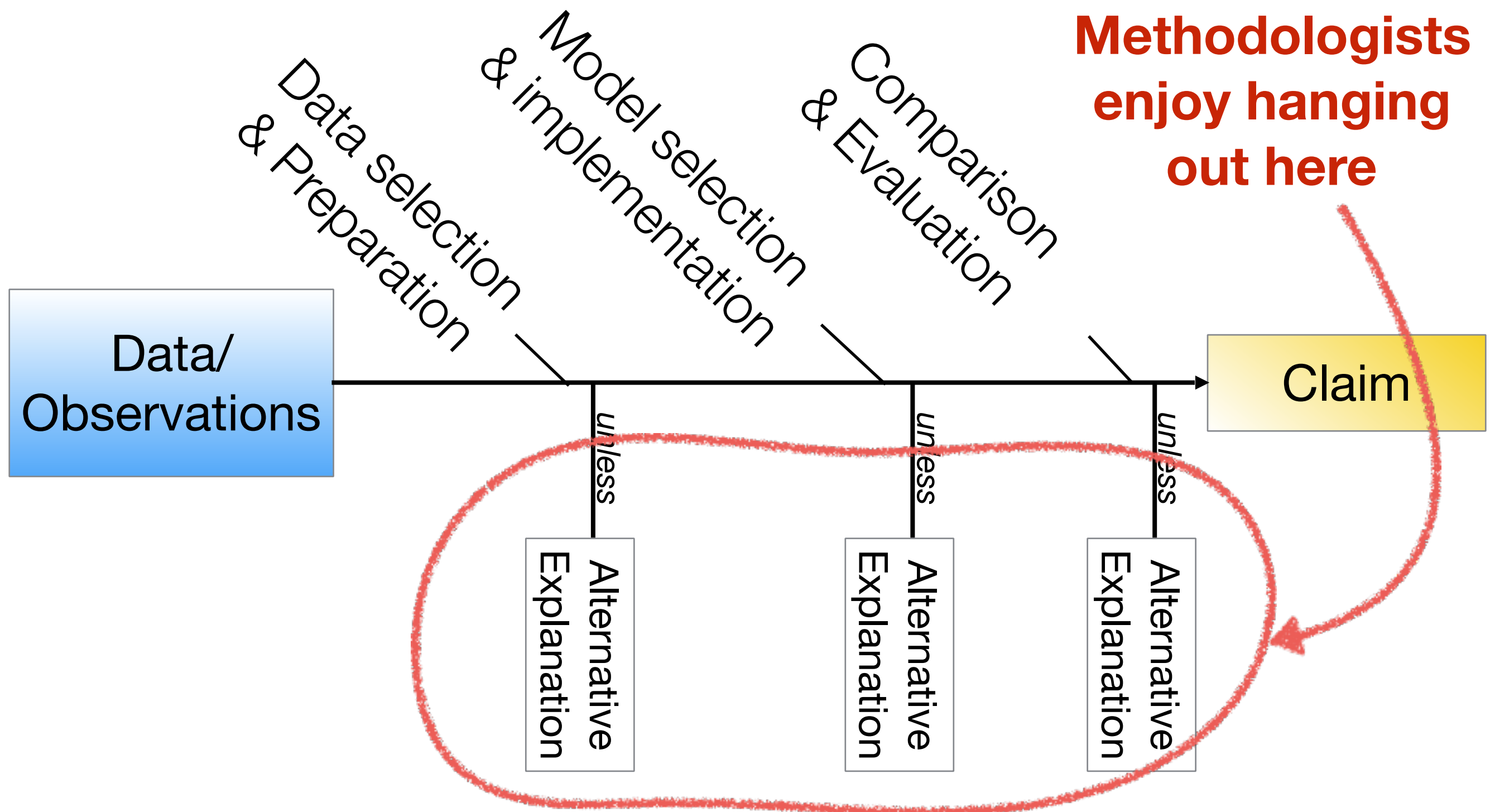
A learning analytics paper is typically a chain

For example



A learning analytics paper is typically a chain

For example



Quantifying uncertainty

- The methods we use in learning analytics are subject to random and systematic error
- If we do not make explicit efforts to quantify uncertainty, we chain together weak links
- Sensitivity analysis, cf. Kennedy's (2002) 10th commandment of applied econometrics:
 - *Thou shalt confess in the presence of sensitivity*
(Corollary: Thou shalt anticipate criticism)

Laudable example

“The choice of a particular time-on-task estimation strategy plays an important role in the overall model fit and subsequent model interpretation.”

(LAK15 best paper award)



Kovanović, V., Gašević, D., Dawson, S., Joksimović, S., Baker, R. S. J. D., & Hatala, M. (2015). Penetrating the black box of time-on-task estimation. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge - LAK '15* (pp. 184–193). New York, New York, USA: ACM Press.

And a cautionary tale

- Kizilcec, Piech, and Schneider (LAK 2013) Deconstructing Disengagement: Analyzing Learner Subpopulations in Massive Open Online Courses
- One of the most cited LAK papers (400+), especially for the claim that MOOC learners can be categorized as **completing, auditing, disengaging, or sampling**.
- There are some really great things about this paper! And it's findings are probably true
- But there are also some methodological disasters.

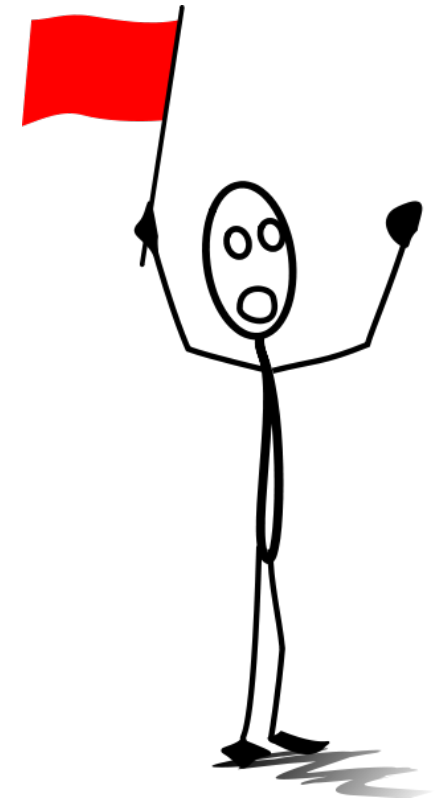
A chain of methods

- **Step 1:** In each assessment period, all participants are labeled
 - “on track” (did the assessment on time),
 - “behind” (turned in the assessment late),
 - “auditing” (didn’t do the assessment but engaged by watching a video or doing a quiz), or
 - “out” (didn’t participate in the course at all).
- e.g., [T, T, T, T, T, B, A, A, A].

A chain of methods

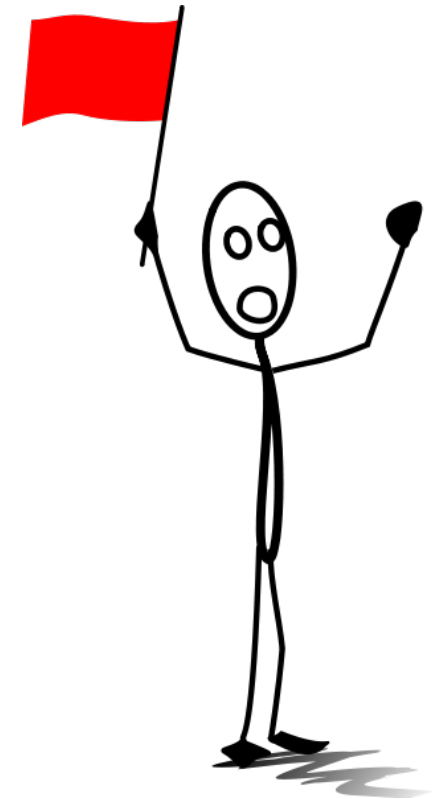
- **Step 2:** calculate the similarity between engagement descriptions for two students:
 - assign numerical values to each label
 - on track = 3, behind = 2, auditing = 1, out = 0
 - and compute the L1 norm of the list of numbers.
- **Step 3:** k-means clustering (repeated 100 times)

A chain of methods



- MOOC data
- **Step 1:** [T, T, T, B, A, A], [A, A, O, B, B, B]
- **Step 2:** L1 norm of [3, 3, 3, 2, 1, 1], [1, 1, 0, 2, 2, 2] (= 9)
- **Step 3:** k-means clustering (repeated 100 times)
- Claim: **completing, auditing, disengaging, or sampling**

A chain of methods



- MOOC data
- **Step 1:** [T, T, T, B, A, A], [A, A, O, B, B, B]
- **Step 2:** L1 norm of [3, 3, 3, 2, 1, 1], [1, 1, 0, 2, 2, 2] (= 9)
- **Step 3:** k-means clustering (repeated 100 times)
- Claim: **completing, auditing, disengaging, or sampling**

If you don't lock your front door, and no one breaks into your house, it doesn't mean you shouldn't have locked your door.

–Roger Peng (paraphrasing)

Methodology Mapping

- Catalogue the methods being used in LA
- Looking for miner-grained data

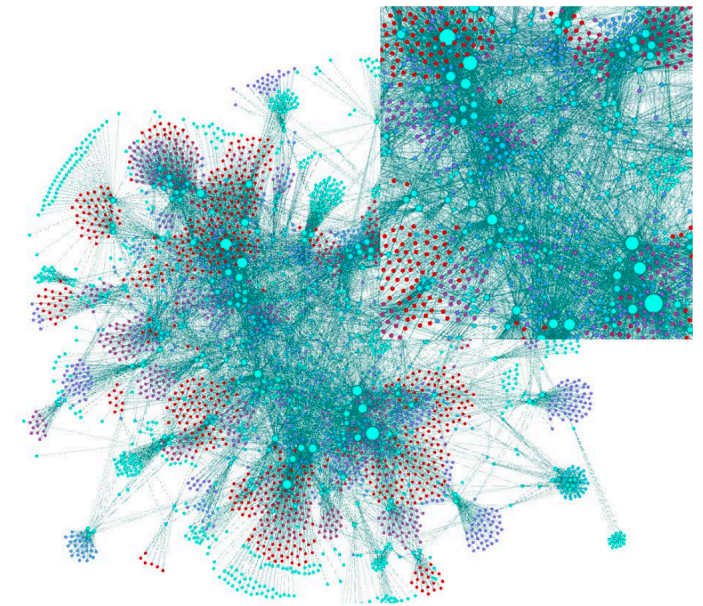


Figure 1b: All LAK conference citations. The network has been configured to highlight the clustering. Nodes sized by degree centrality

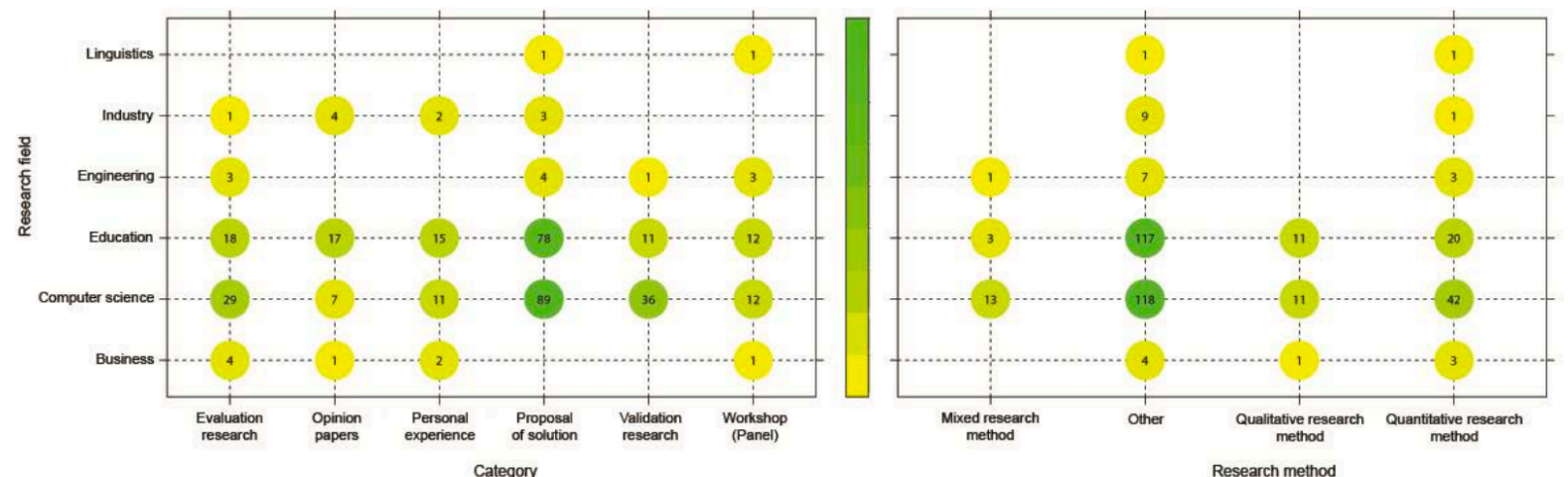
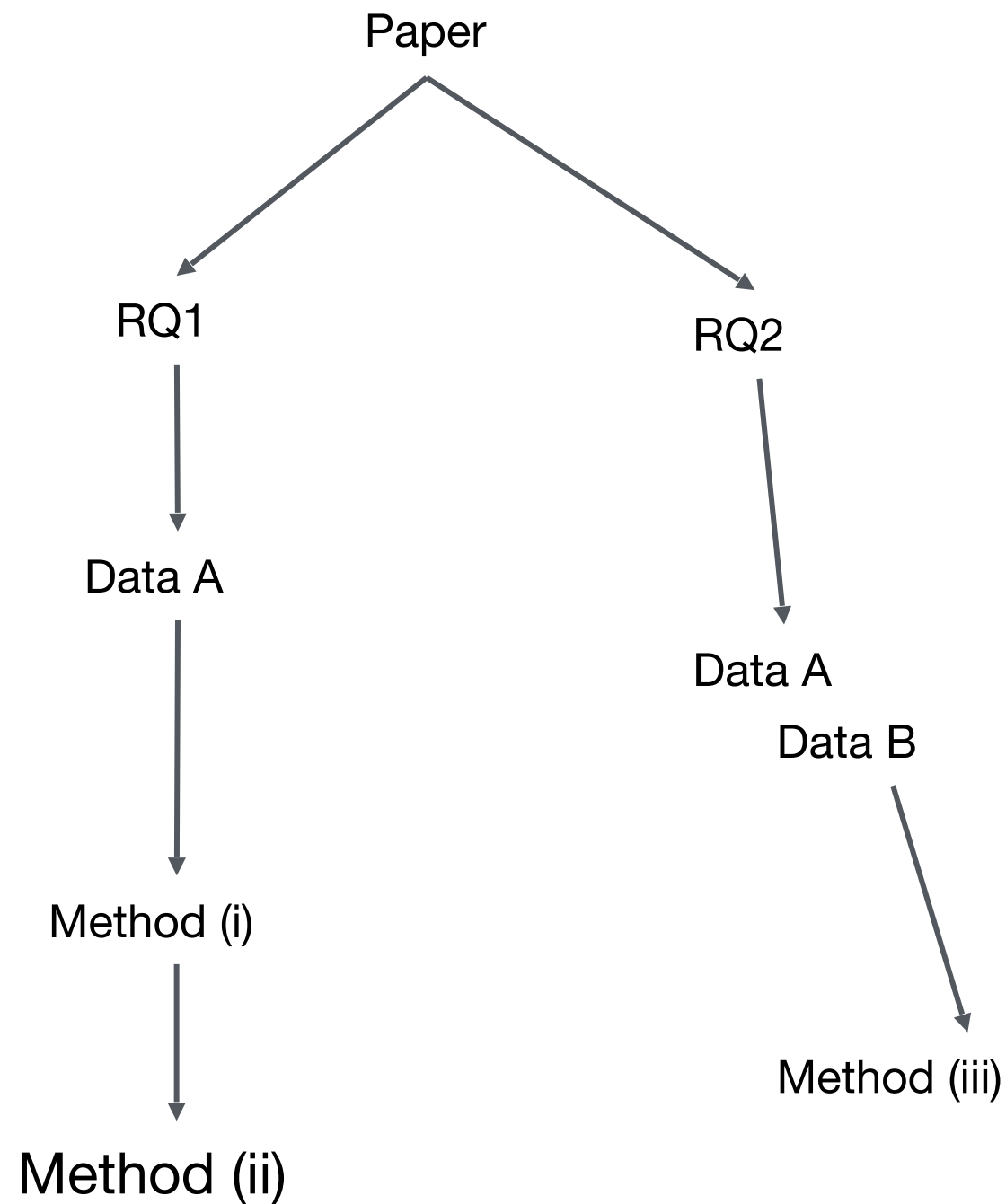


Figure 5b: Cross-tabulation of the authors' home disciplines (i.e., numbers in the circles represent the number of authors) with both research methods reported in the conference papers and types of conference papers.

Methodology Mapping



Methodology Mapping

- Open a random paper from the LAK18 folder
- Read the Abstract
 - Identify research question(s)
 - Open Airtable
 - Input a RQ code (18LAuthor first initials_RQ#, EG - 18LJBTO_1)
 - Input the research question under RQ
 - Choose a method(s) (or add one under the methods tab if it doesn't exist)
 - Choose a data form (or add one under the data tab if it doesn't exist)
 - Do you think the assumptions of the method are met?
 - Do you think limitations with the method are mentioned?
 - Input the author info under the "Papers" tab

The RAP System: Automatic Feedback of Oral Presentation Skills Using Multimodal Analysis and Low-Cost Sensors

ABSTRACT

Developing communication skills in higher education students could be a challenge to professors due to the time needed to provide formative feedback. This work presents RAP, a scalable system to provide automatic feedback to entry-level students to develop basic oral presentation skills. The system improves the state-of-the-art by analyzing posture, gaze, volume, filled pauses and the slides of the presenters through data captured by very low-cost sensors. The system also provides an off-line feedback report with multimodal recordings of their performance. An initial evaluation of the system indicates that the system's feedback highly agrees with human feedback and that students considered that feedback useful to develop their oral presentation skills.