

Lessons from Autonomous Personalization

Georgetown University, Thursday January 18th 2018

Charles Lang
Visiting Assistant Professor
Teachers College Columbia University

charles.lang@tc.columbia.edu

Slides available at:
<https://github.com/charles-lang/georgetown>

Structure of Talk

- Personalization algorithms
- Three lessons: goals, models and implementation
- Applied to teaching
- Masters Degree

Autonomous Personalization



Xue Ji

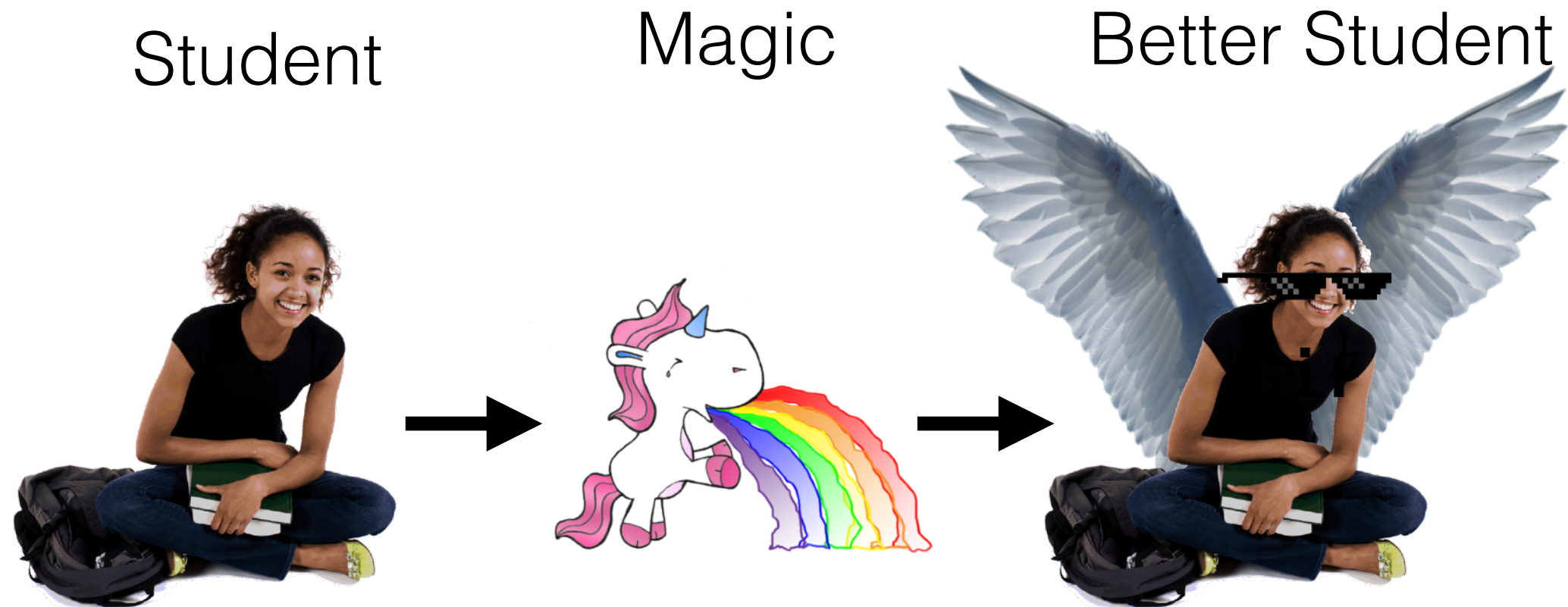
The success of
education depends
on adapting
teaching to
individual
differences among
learners

329 BC, Xue Ji

“Uniformity is the curse of American schools...Individual instruction is the new ideal.”

Charles Eliot, 1899

Best Definition



Definitionally True

If we provide what the student needs, when they need it, they will learn better

Vocabulary

Individualization: learning goals are the same for all students, but students can progress through the material at different speeds

Differentiation: learning goals are the same for all students, but the method or approach of instruction varies according to the preferences of each student

Personalization: learning goals and content as well as the method and pace may all vary (so personalization encompasses differentiation and individualization)

ed.gov (2010)

Autonomous

We trust the machine to do some of the
decision making

Fuel for machines = variation

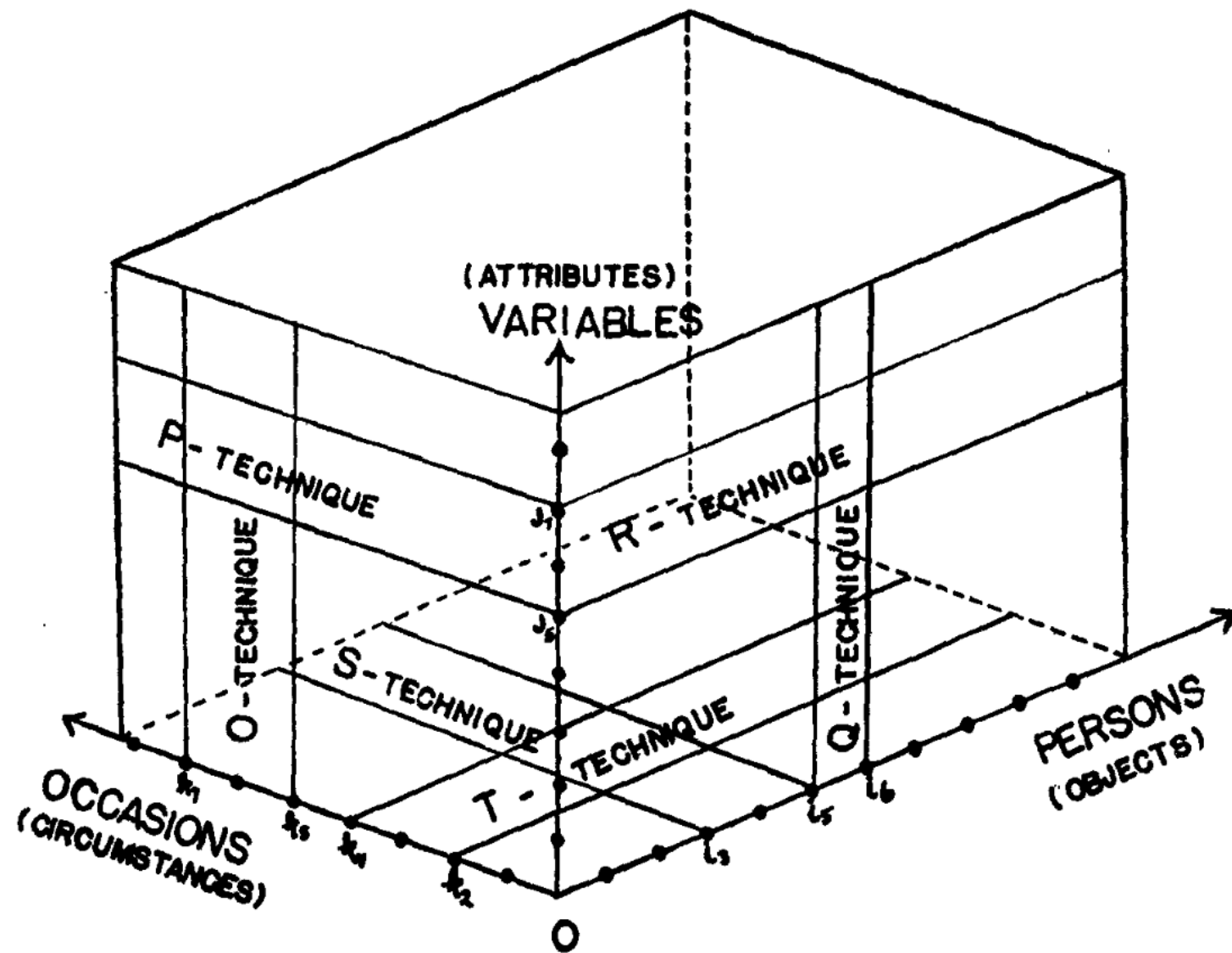


FIG. 1. THE COVARIATION CHART

Cattell, 1952

Bayesian Student

$$P(K|D) \propto P(D|K).P(K)$$

behavior

context

knowledge

(posterior)

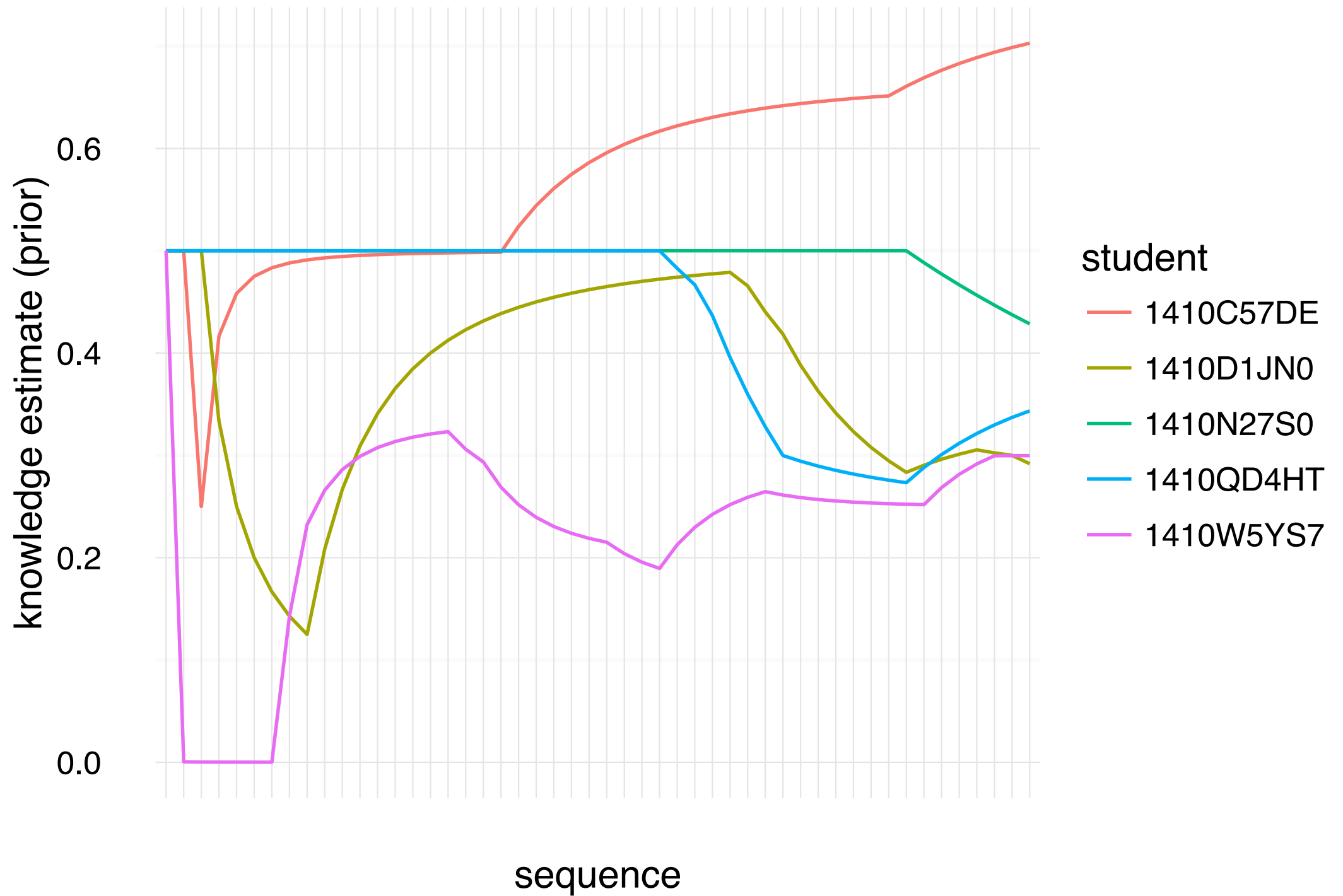
(likelihood)

(prior)



Results

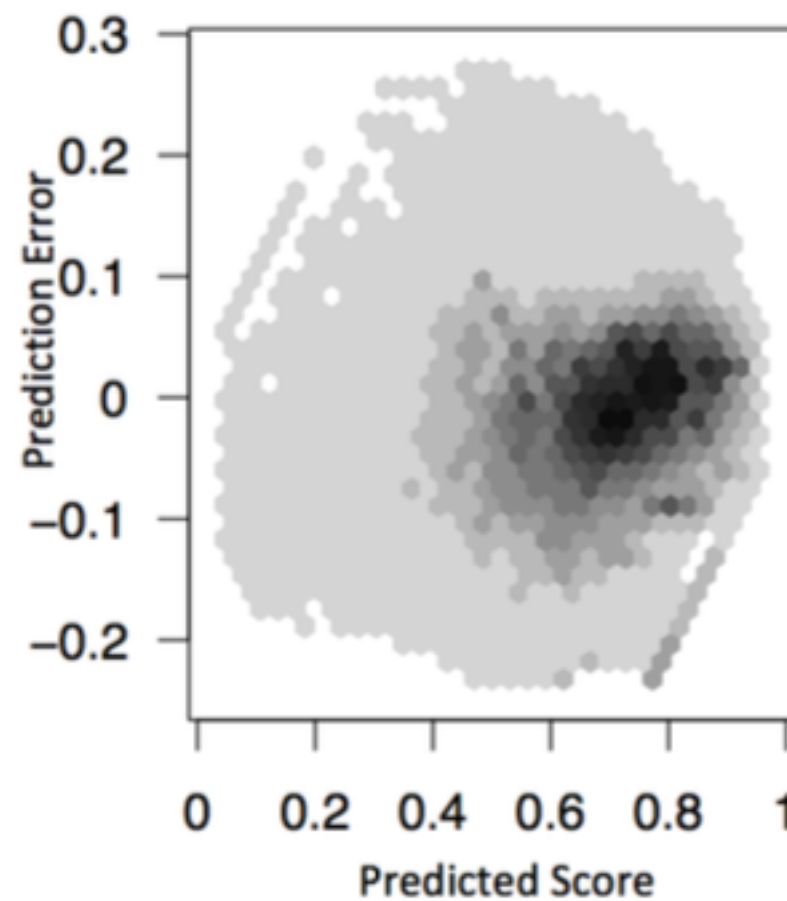
Real Students



Real Students

Model	Overall	Per Student	Per Skill
BKT	0.28	0.39	0.28
Bayes	0.29	0.25**	0.31

**
 $p < .01$



Why?

- Being able to predict student behavior is useful
- Data collection parsimony is *more* useful

Success?

- Objective setting

Lesson 1
Lesson 2
Lesson 3
- Models of learning
- Data location

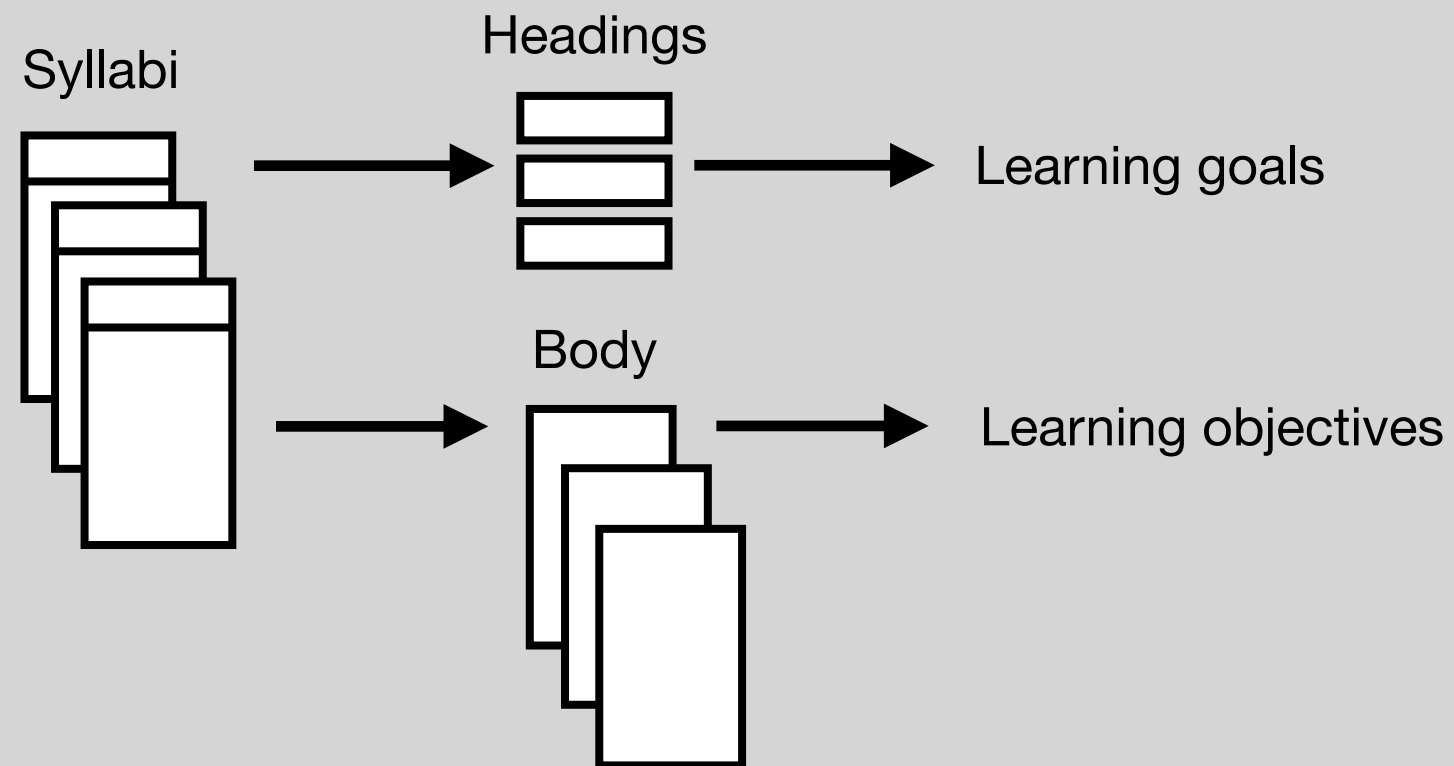
Lesson 1: Objective
setting is non-trivial

Objective Setting is Non-Trivial

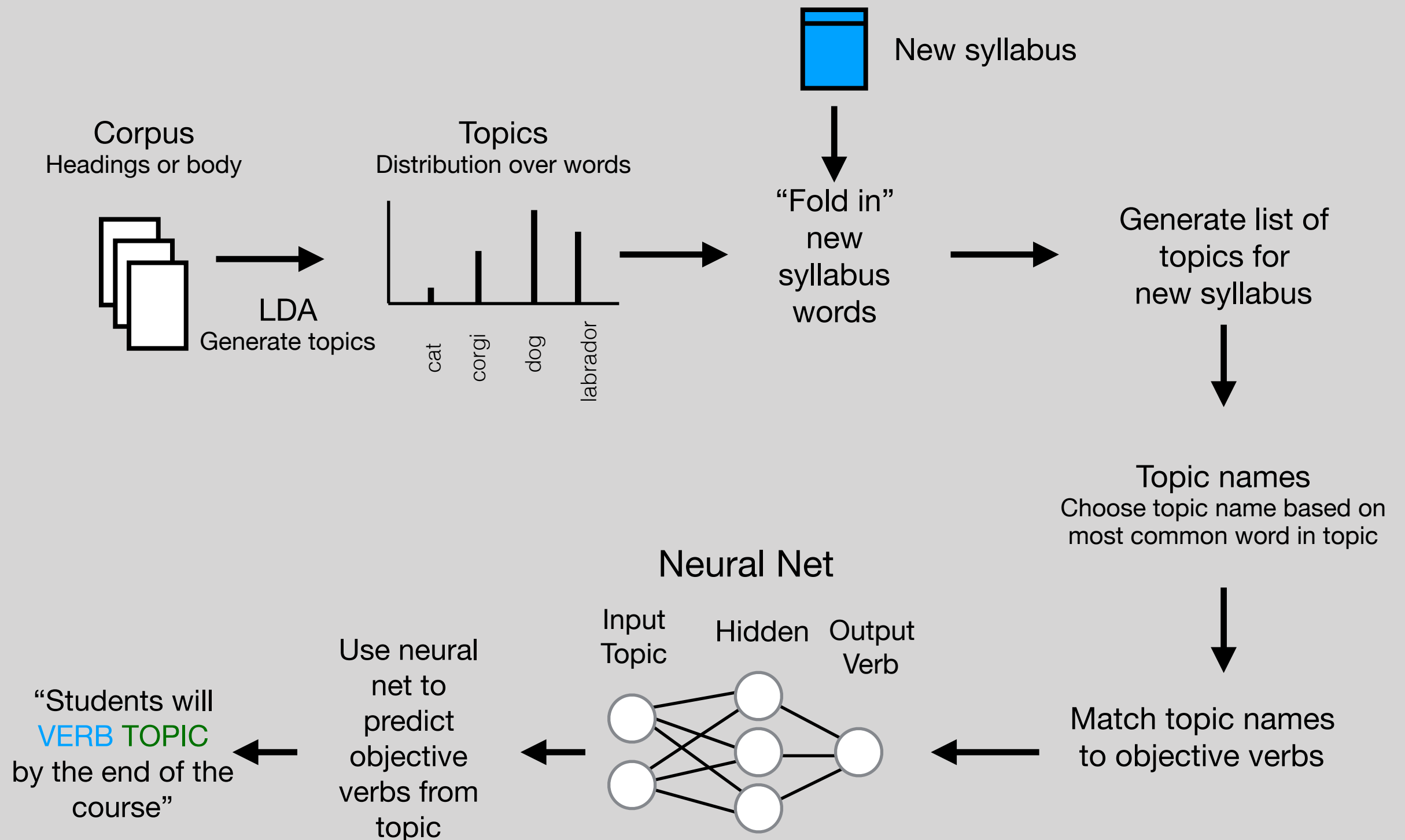
- **Model requires that a desired behavior is defined**
- Time consuming
- Conceptually difficult
- Requires training

Syllabus Project

Can we auto-generate learning goals and objectives from syllabi?



Syllabus Project



Objectives

- **Not to prove The Answer**
- Provide suggestions
- Reduce friction
- Reduce time spent

Lesson 2: The need for (at least) three models of learning

General Framework

- **Narrative Model**: Dinner party version of your personalization model or theory
- **Operational Model**: What you count and how you count it
- **Validation Model**: Convincing yourself there is a connection between your narrative and operational models

Swimming Experience

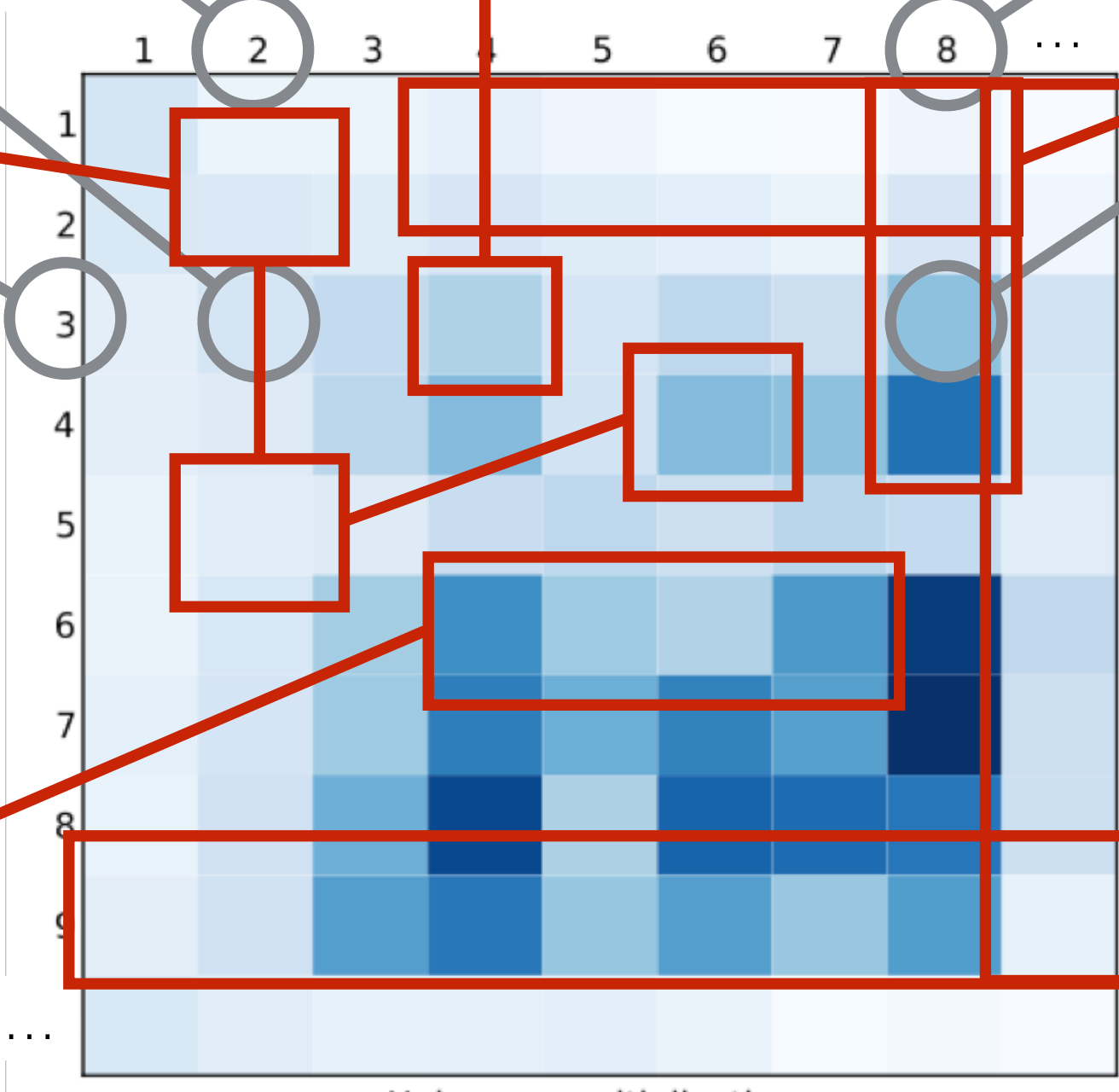
Experimental Psych

Classroom

Context

Social Psych

Ed Stat,
Economics
Bad
Spell "cat"



Behavior

Psychometrics/
Cog Tutor

Qual/Big
Data

Pragmatic Validity

- Relax pursuit of generalizability to other populations
- Concentrate on internal validity and prediction of future events

Grapentine, T. (1995)

Narrative Model

Trigger

Action

Target

< average
usage

email
reminders

increase
deadline
adherence

Operational Model

Count

logins

read emails

submission rate
submission time

Variation

across
students

within
student

within
student

Validation Model

Threshold

below mean
login

Acceptable
number?

Success?

Why
trigger?

Did action
occur?

Did it work?

Unintended
consequences?
Auto-iterate?
Narrative
assumptions

Lesson 3:
Implementation is
everything

Implementation

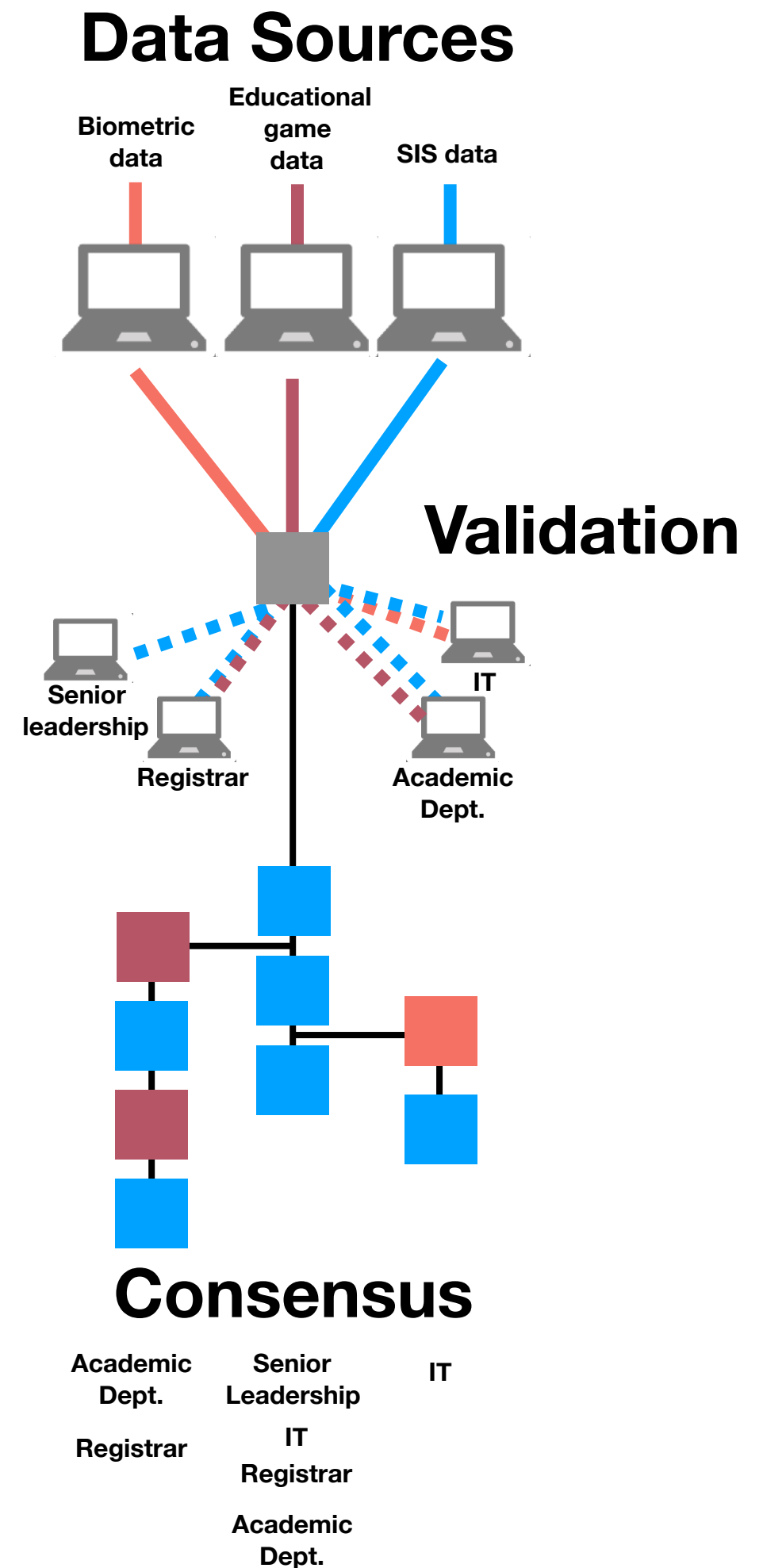
- Data sources are controlled by different groups within an institution
- Negotiation between groups is a major stumbling block to utilizing data

BUT

- Job is too big for any one group, there is more work than can possibly be achieved by one person, unit, department, university, government...

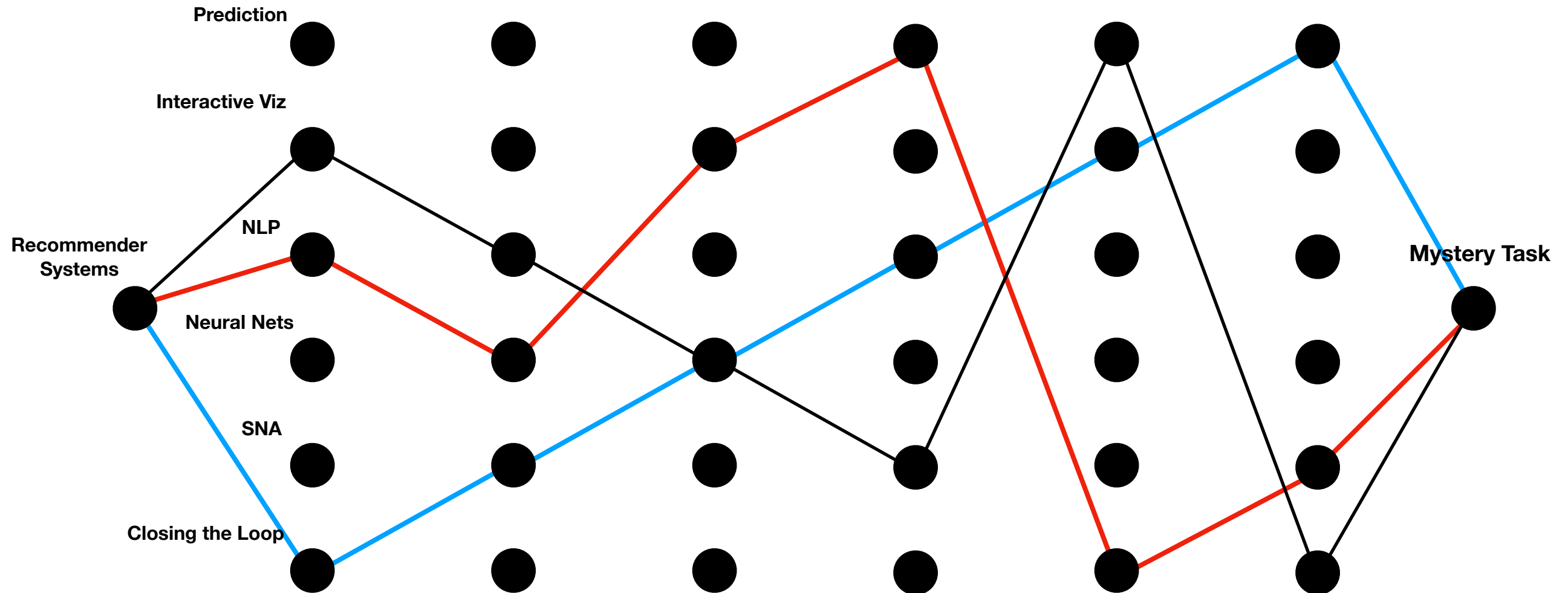
Blockchain Project

Can we use a distributed data base (Blockchain) to build consensus about what data to collect?



Impact on Teaching

LA: Process & Theory



<https://github.com/la-process-and-theory/syllabus>

Why?

- Content isn't necessary cumulative
- Learning is more effective if you can follow your own interest
- Students are better arbiters of what they need to know than instructor
- Promotes self-motivated learning
- More time getting instructor feedback
- Get to test the tool you build in unit 1

Narrative Model

Trigger

Recommend

Action

Choose
content

Target

Motivation

Operational Model

Count

Ratings
(Int/Dif)

Choice
patterns

Commits
Additions

Affect?

Variation

across
students

across
students

across
students/time

Differential
use/utility by
subgroup?

Validation Model

Threshold

Similarity

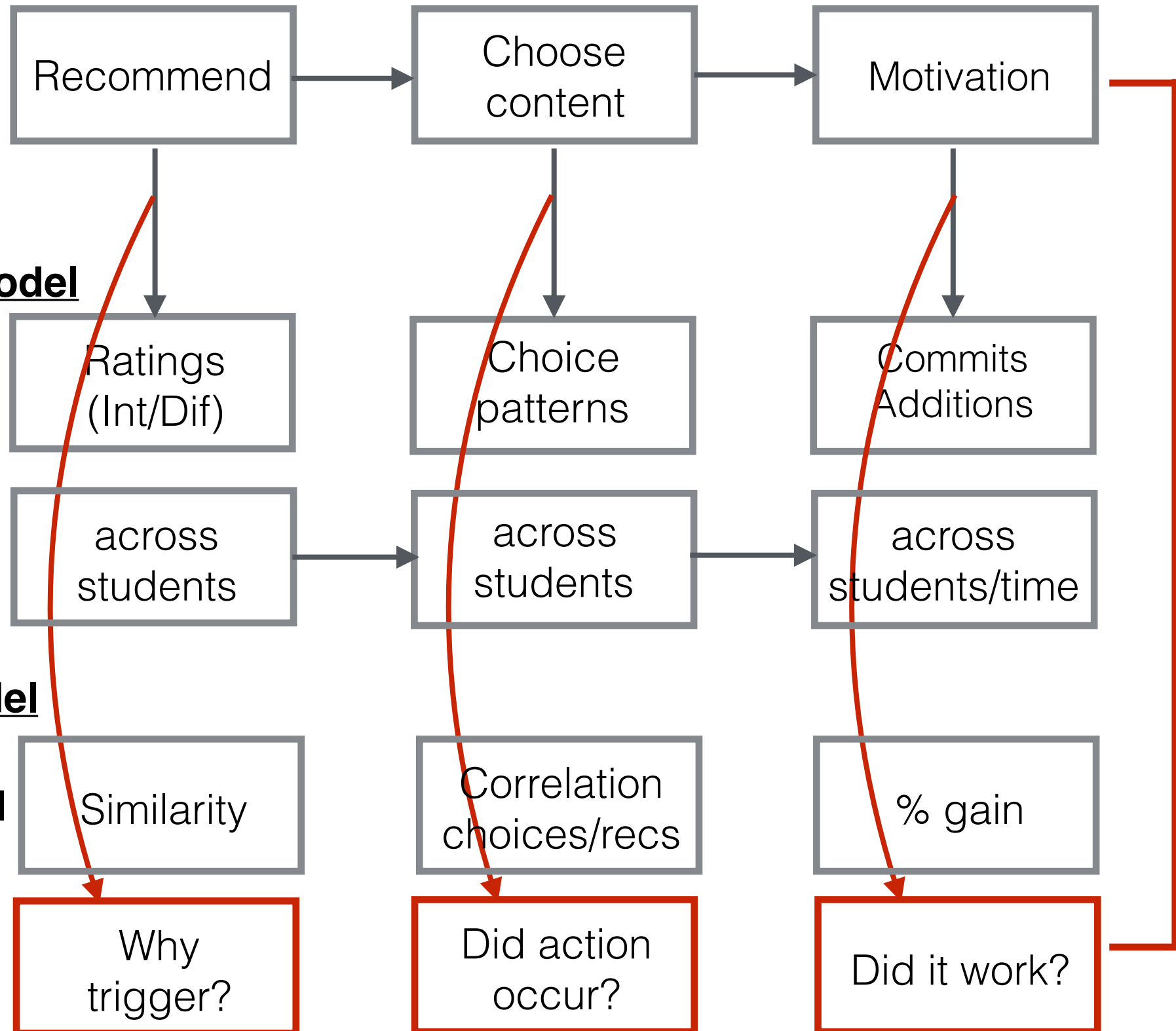
Correlation
choices/recs

% gain

Why
trigger?

Did action
occur?

Did it work?

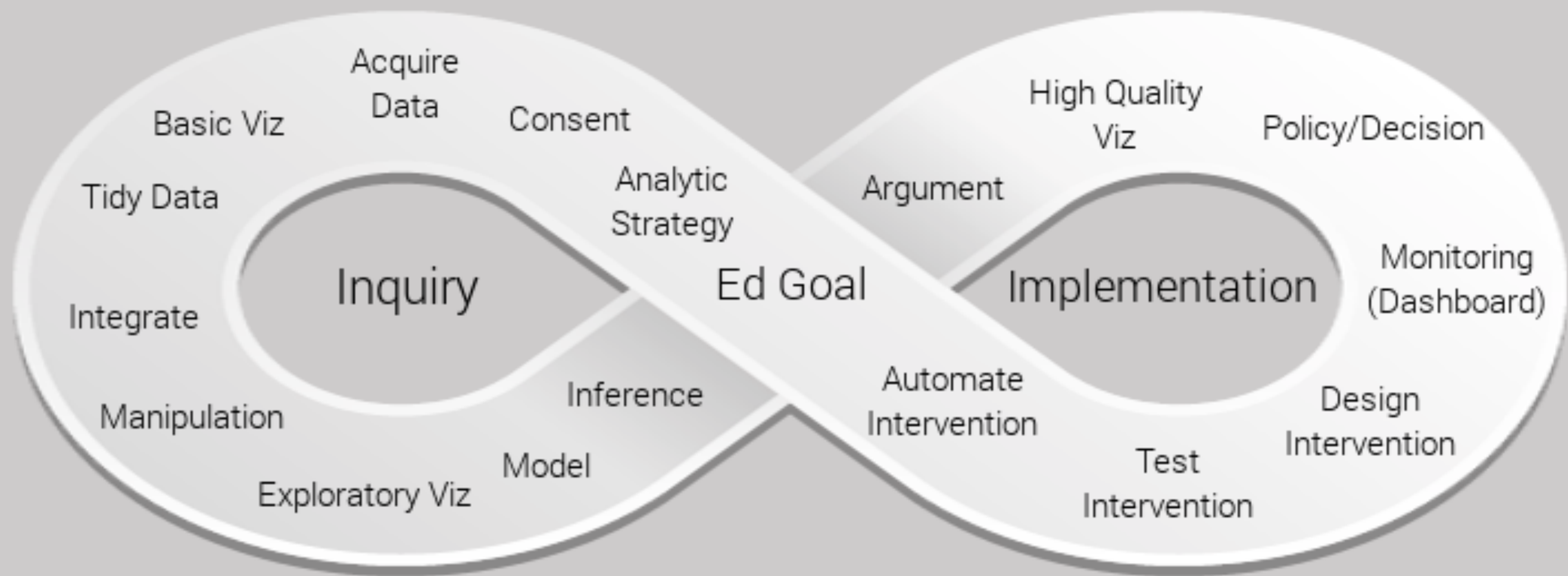


M.Sc in Learning Analytics

M.Sc in LA

- 18-24 months
- 3 core course methods sequence: Intro, Intermediate, job readiness
- Required courses in sociology, cognition & learning, data visualization, statistics
- Placed with employers for group based capstone projects

Ed Data Science Cycle



Example Capstone Projects



- Which modalities within platform correlate with highest engagement?
- Interpreting clustering of students within personalized learning platform
- College math readiness personalized tutor evaluation
- Patterns within admissions system usage among applicants

Thanks!

Questions?