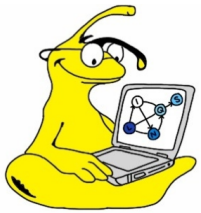
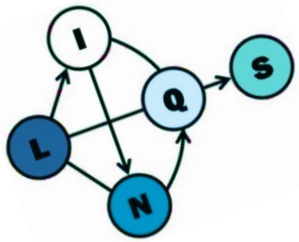
CSLS Panel

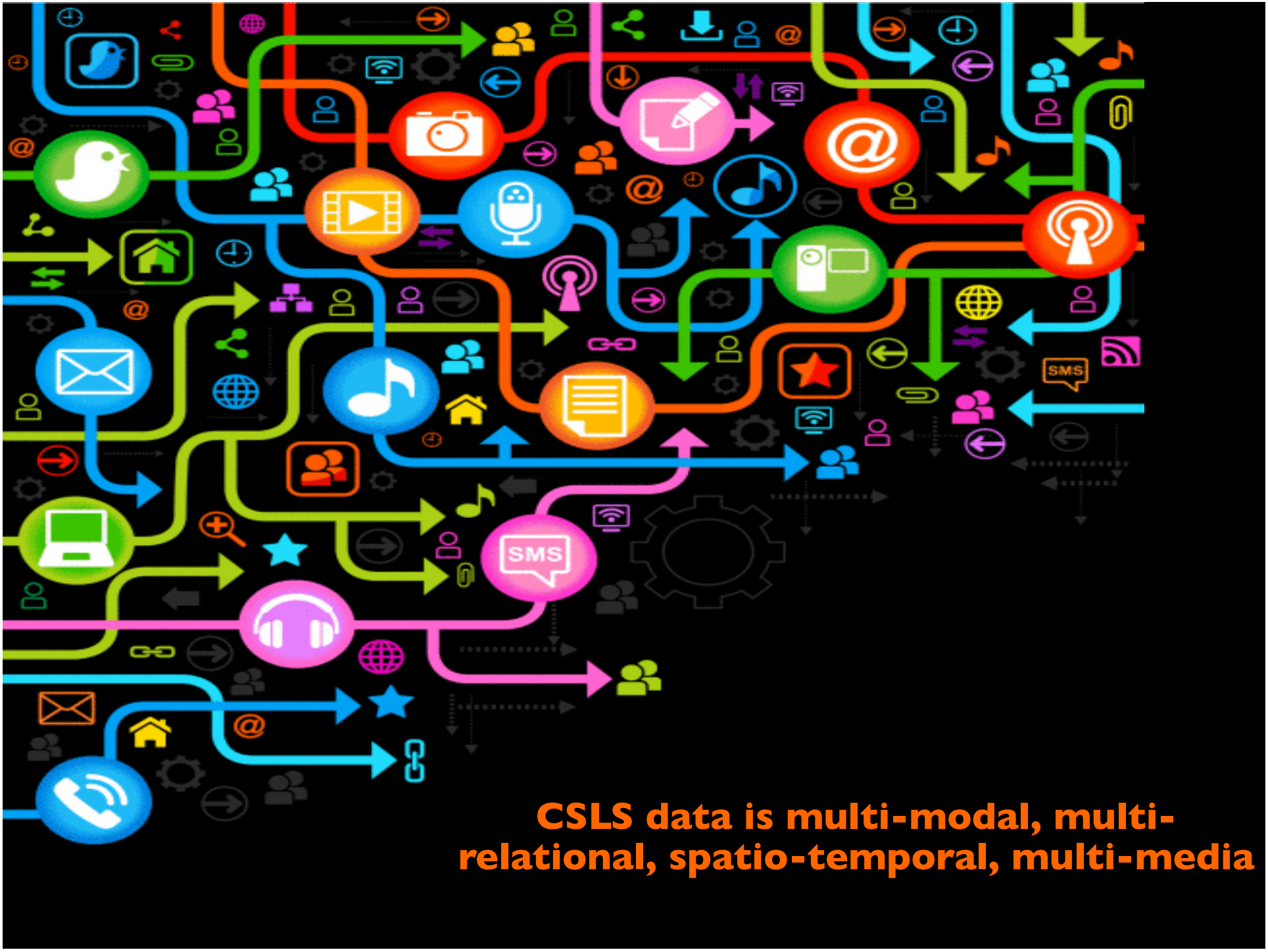
Lise Getoor
Computer Science Department
University of California, Santa Cruz
@lgetoor

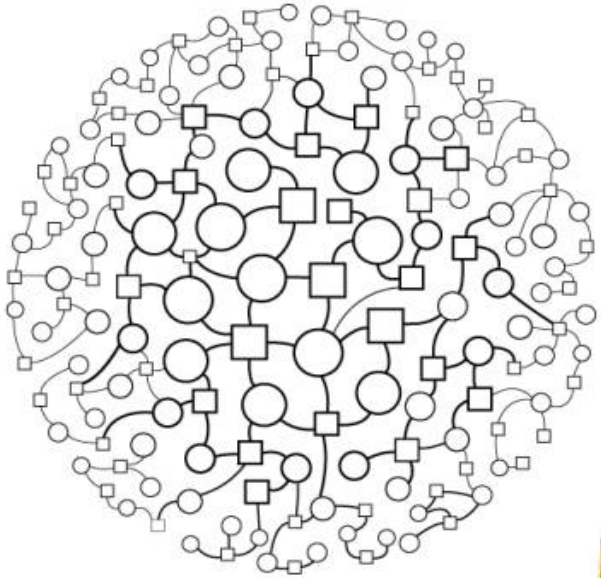


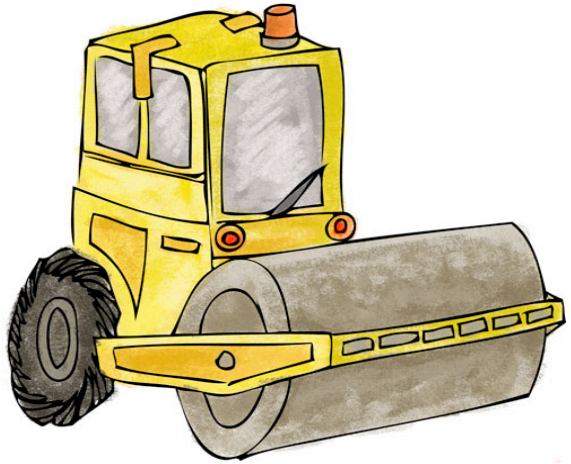
August 30, 2016

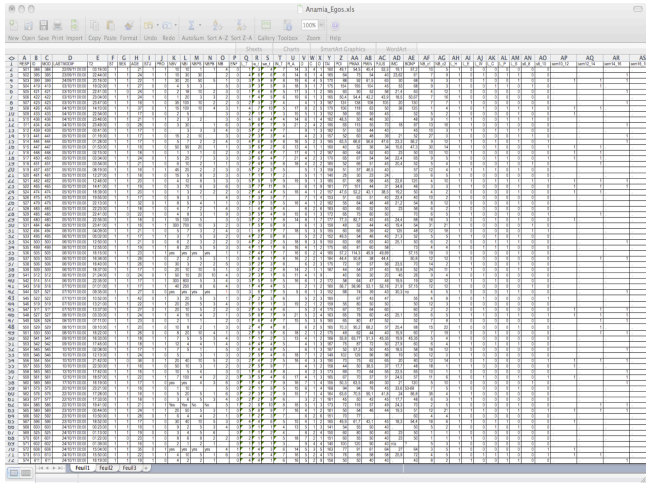


ML


SMS
SMS
CSLS data is multi-modal, multi-relational, spatio-temporal, multi-media






Anamia_Egos.xls

NEED: ML methods that take into account *relational context*

CSLS ML Patterns

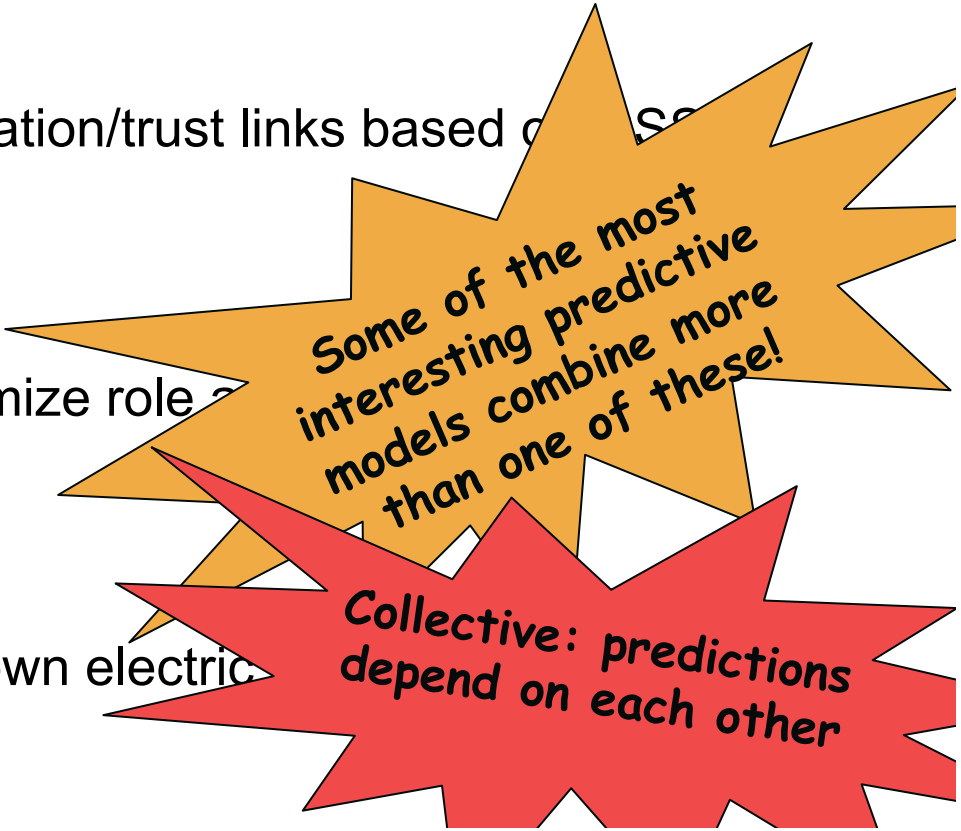
Predictive
Model Type
Intervention

x

Individual
Dyads
Social Context
Community
Society

Predictive Models

- Individual
  - Personalized recommendation
    - e.g., use CSS information to improve user-level predictions
- Dyad
  - Link recommendation
    - e.g., recommend collaboration/trust links based o
- Community
  - Ecosystem optimization
    - e.g., team formation, optimize role
- Society
  - Predict tipping points
    - e.g., when will everyone own electric



Interventions

- Support what/if analysis: if an intervention/treatment was made, what would be the outcome
  - Benefits from a causal model

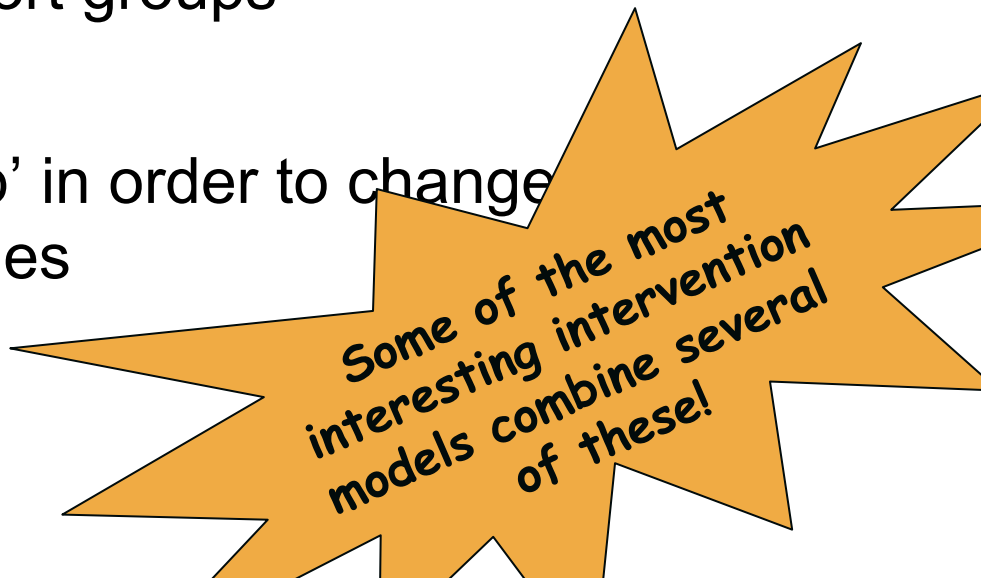
- Possible interventions:
  - New: information, social interaction, group membership

Optimize: engagement, health, cost

- Challenges in CSLS: social interference
  - Outcomes of treatments may not be independent
    - E.g., get out the vote, HIV treatment

Models that Support Interventions

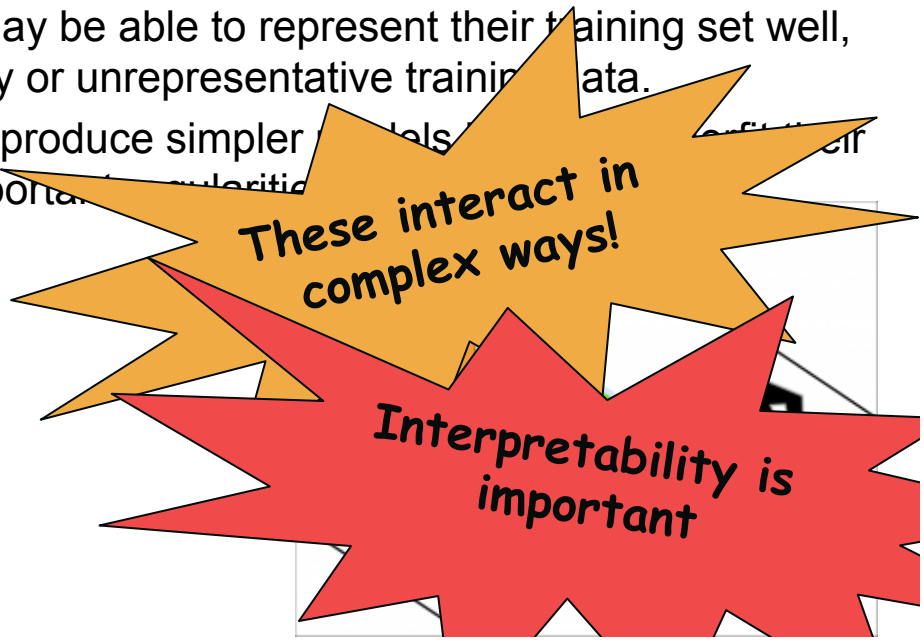
- Individual
  - Treatment recommendation
- Dyads
  - Which links to make to maximize information flow
- Community
  - Which members to recruit to optimize group performance, e.g. support groups
- Society
  - Which individuals to ‘flip’ in order to change population level outcomes

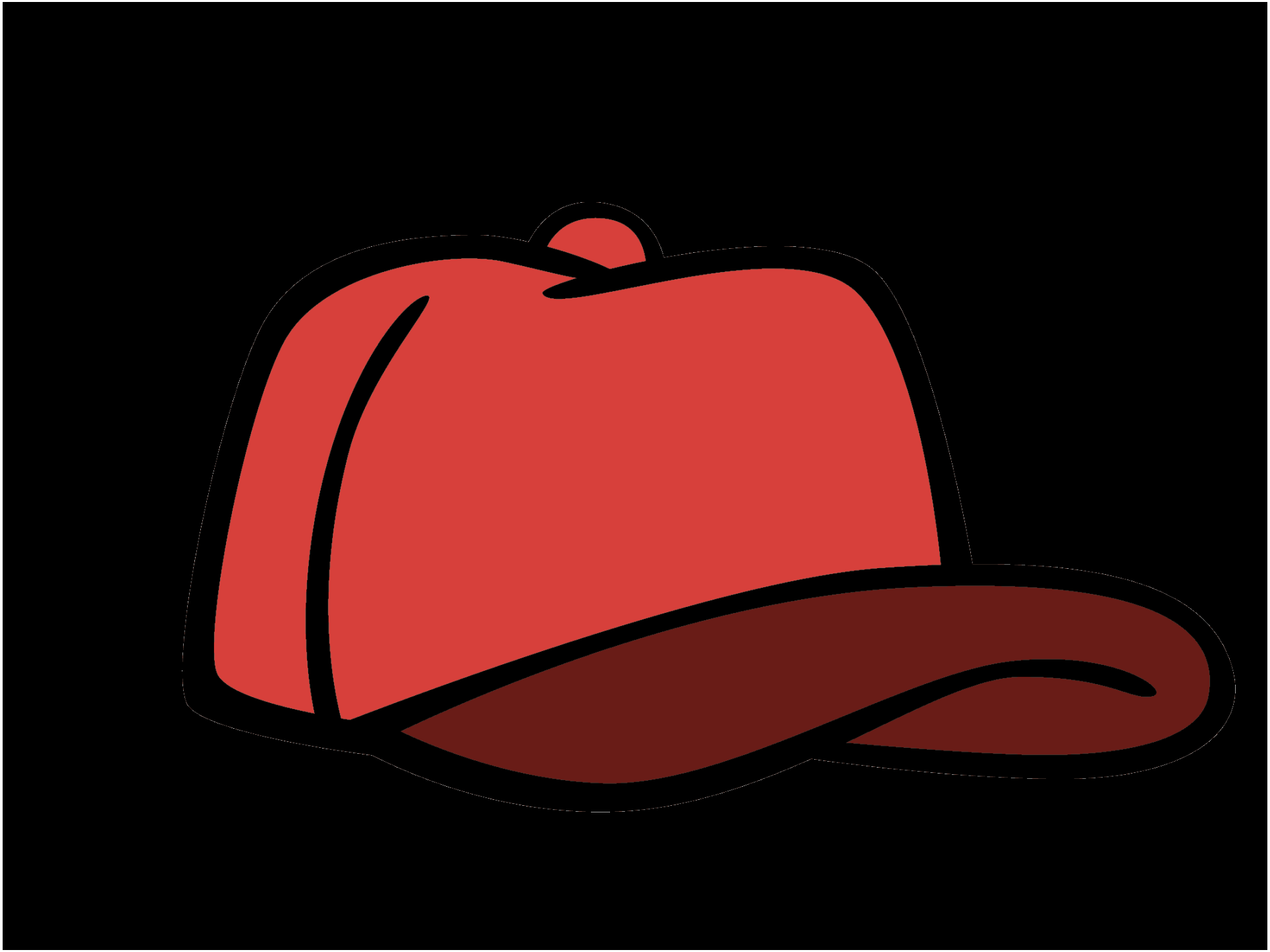


bias

Bias: Three Interpretations

- Sampling Bias
  - Subset of the population overrepresented in data
- Statistical Bias
  - Bias-Variance Tradeoff:
    - Want model that accurately captures the regularities in its training data and also generalizes well to unseen data
    - impossible to do both simultaneously
    - High-variance learning methods may be able to represent their training set well, but are at risk of overfitting to noisy or unrepresentative training data.
    - Algorithms with high bias typically produce simpler models that may underfit their training data, failing to capture important regularities
- Ethical Bias
  - Fairness, equity
- CSLS Bias?





ethics

social engineering (noun) :
a discipline in social science that refers to efforts influence particular attitudes and social behaviors on a large scale, whether by governments, media or private groups in order to produce desired outcomes in a target population.

Challenges

- Who defines social good?
- Dimensions
  - Personal good
  - Collective good
  - Societal good
- Aligning incentives to maximize benefit

CSLS

Research Opportunities #1

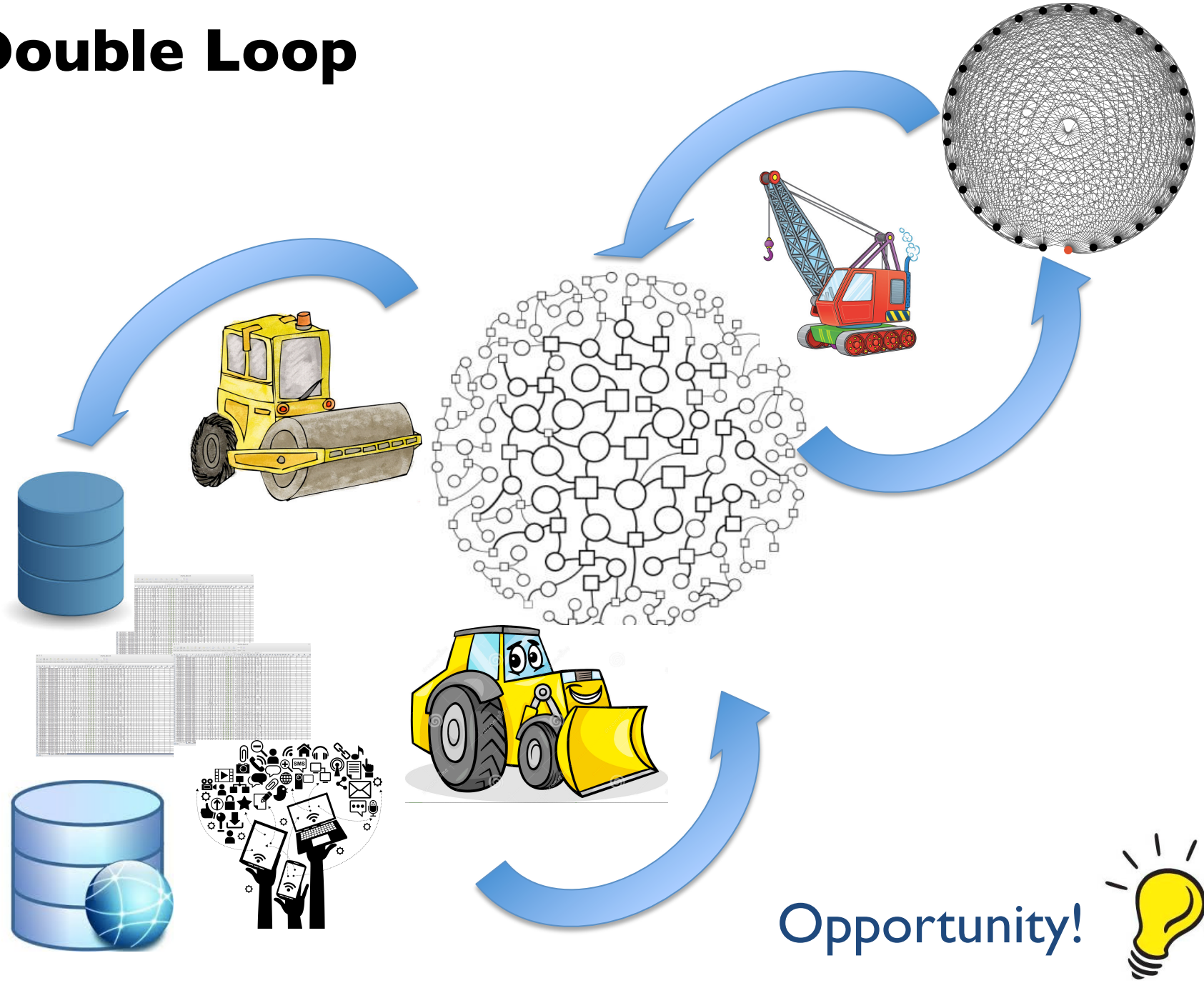
- ML
  - Most ML methods designed for independent, identifiably distributed data (IID) represented as a table
  - Often need causal models, not just predictive models
  - Models need to be able to handle social interference
- Social Science
  - How to mix psychological with sociological models? Social and geographic? Social and political?
  - How to mix big data (ML), small data (ethnographic) and social science theory?
  - How does culture affect data? Affect context?

Research Opportunities #2

- Human in the loop
  - Human in the data
  - Human in the system
- Communities and societies in the loop
  - Human***s***
  - Develop new systems engineering science for understanding how people and data and systems interact
  - And how to improve systems over time

Research Opportunities #3

- There ***are*** common abstractions that go across a broad number of the most compelling cyber-social systems domains including heath, education, smart cites, citizen engagement and more

- Three Cs:
  - Context
  - Comparative modeling
  - Collective reasoning

Double Loop