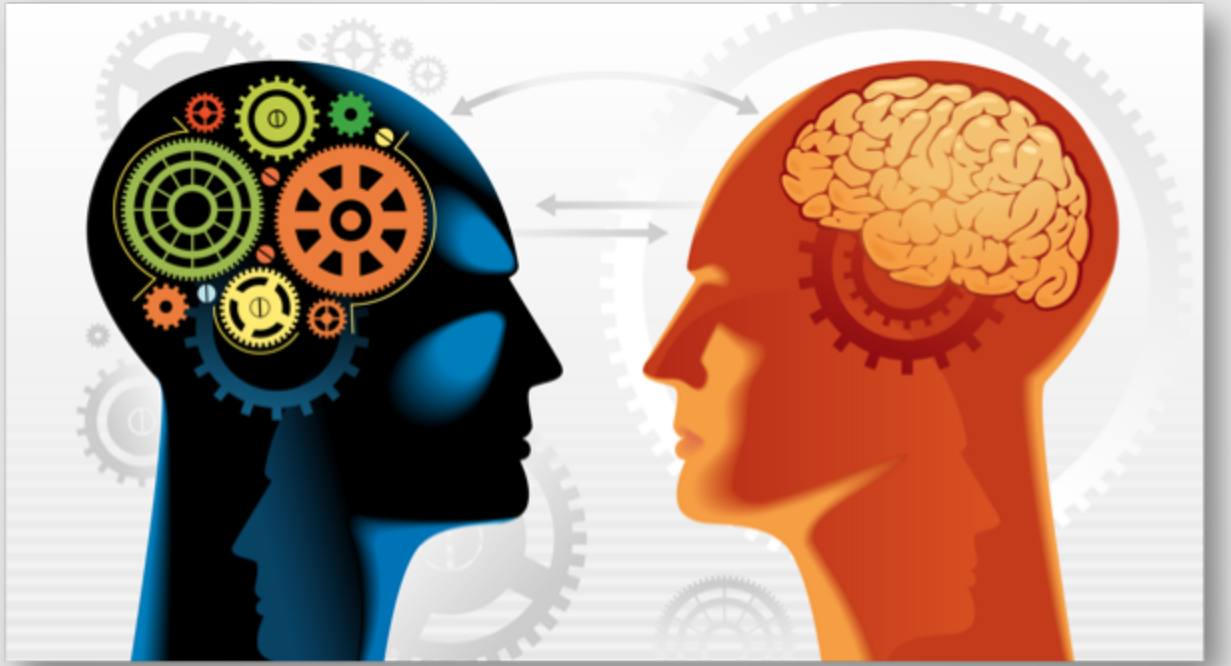
# SIG 4: Prediction & Intervention

Agoritsa, Jeff, & Fei

# Schedule

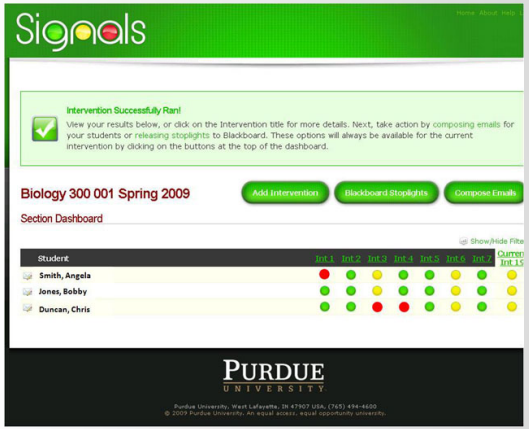
- (20 min) Machine Learning
- (25 min) Macro-Meso-Micro discussion & comments in KF
- (15 min) split-group discussion on teacher/student point of view
- (15 min) break
- (25 min) discussion following split-group
- (35 min) activity: Agoritsa’s showcase / hands-on activity



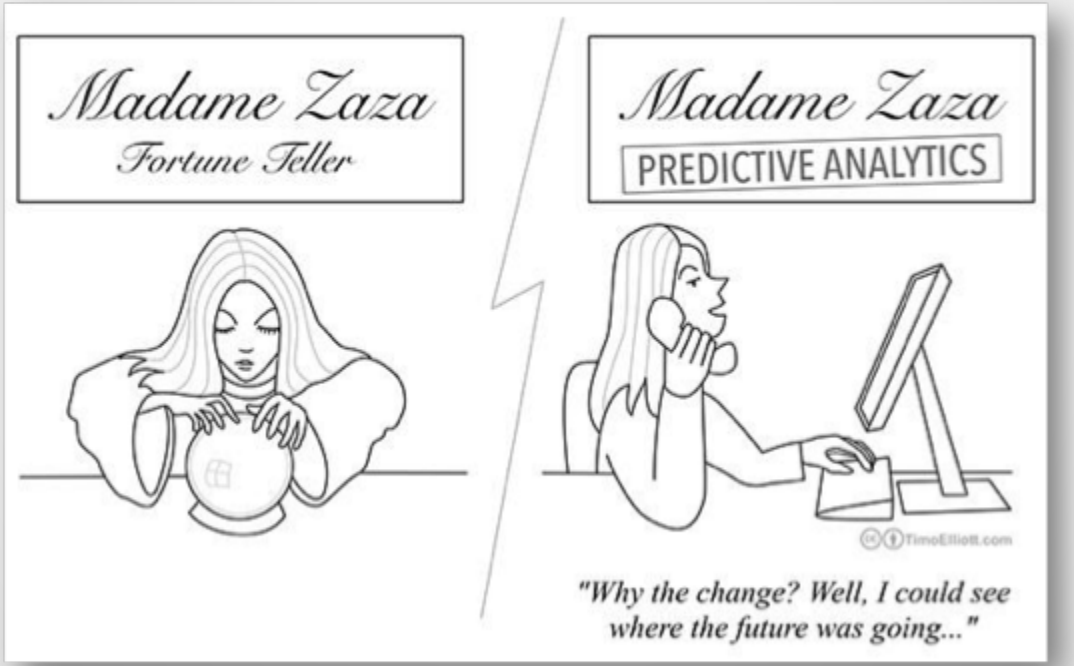
Machine Learning

# It’s not new

- Guest speakers: George Karypis & Jeff Grann
- Signals: Applying academic analytics. Educause Quarterly, 33, 1-10. (Week 6)
- The open university: Improving Retention by Identifying and Supporting "At-Risk" Students. EDUCAUSE Review Online. (Week 6)
- Learner modeling: A review of recent advances in learner and skill modeling in intelligent learning environments. User Modeling and User-Adapted Interaction, 22(1-2), 9–38. (Week 8)
- Bayesian knowledge tracing model: Properties of the bayesian knowledge tracing model. Journal of Educational Data Mining, 5(2), 1-10. (Week 8)



# Predicting the future…



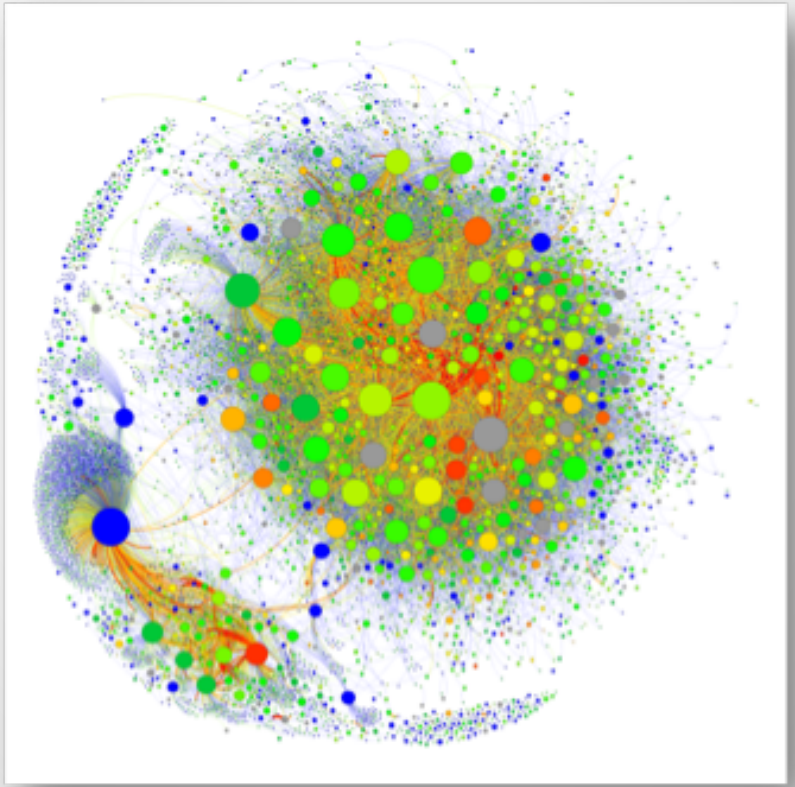
# A short intro video

https://vimeo.com/41995910

# Predictive analytics?

Relations

Algorithms



# What is machine learning

Machine learning is a scientific discipline that explores the construction and study of algorithms that can learn from data. Such algorithms operate by building a model from example inputs and using that to make predictions or decisions, rather than following strictly static program instructions. Machine learning is closely related to and often overlaps with computational statistics; a discipline that also specializes in prediction-making.

From: http://en.wikipedia.org/wiki/Machine_learning

# Machine learning: types

- Supervised learning:
  - Example inputs and desired outputs
  - The goal is to learn a general rule that maps inputs to outputs
  - E.g. Pardos' article, Kloft's article, Baker's article, DeBoer's article
- Unsupervised learning:
  - Leaving it on its own to find structure in its input (discovering hidden patterns, etc.)
- Semi-supervised learning
- Reinforcement learning:
  - Interacts with a dynamic environment
  - E.g. driving a vehicle, playing a game, teaching

# Machine learning workflow

Training set

Unsupervised

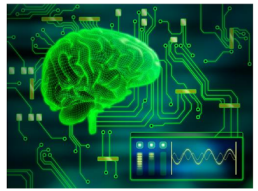
Supervised

Feature extraction


Machine learning algorithm

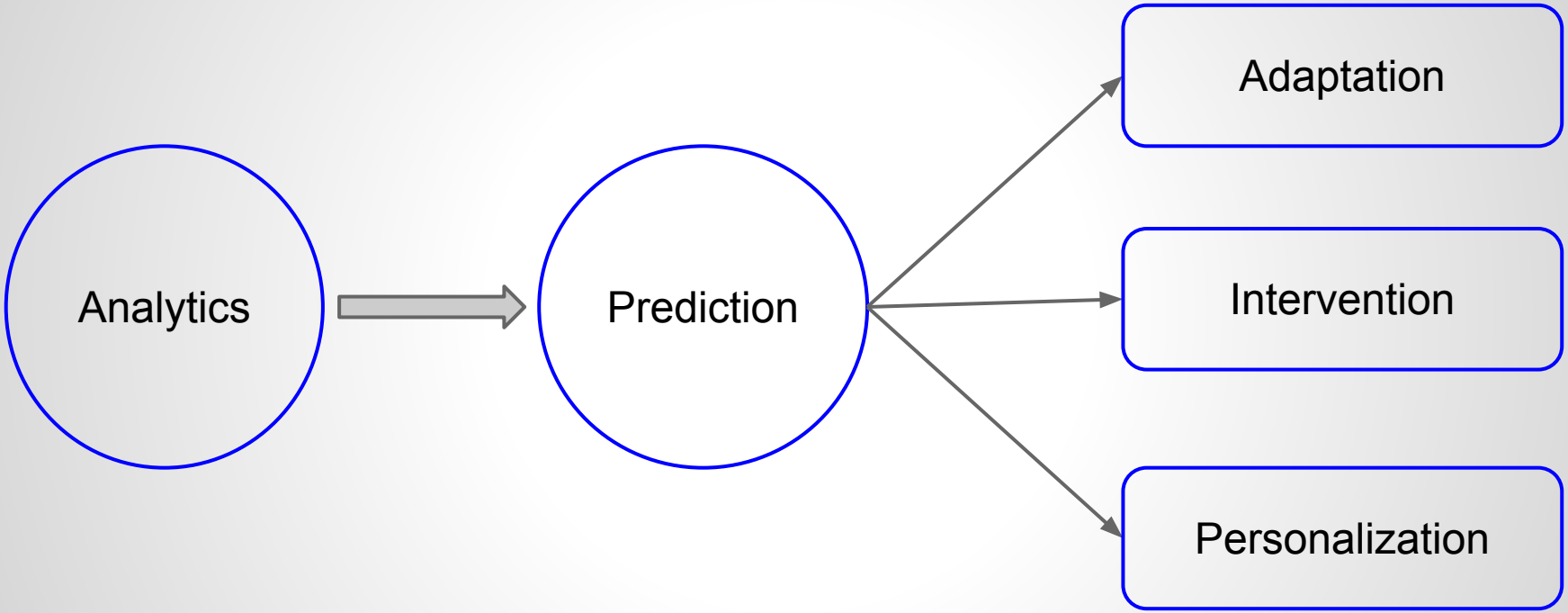
Grouping of objects

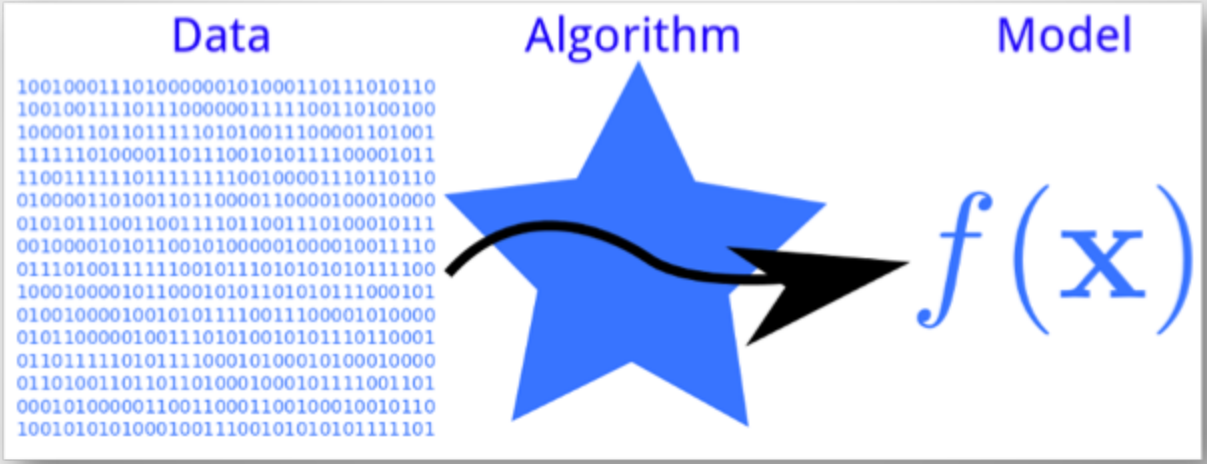

Predictive model



New Data

Annotated data

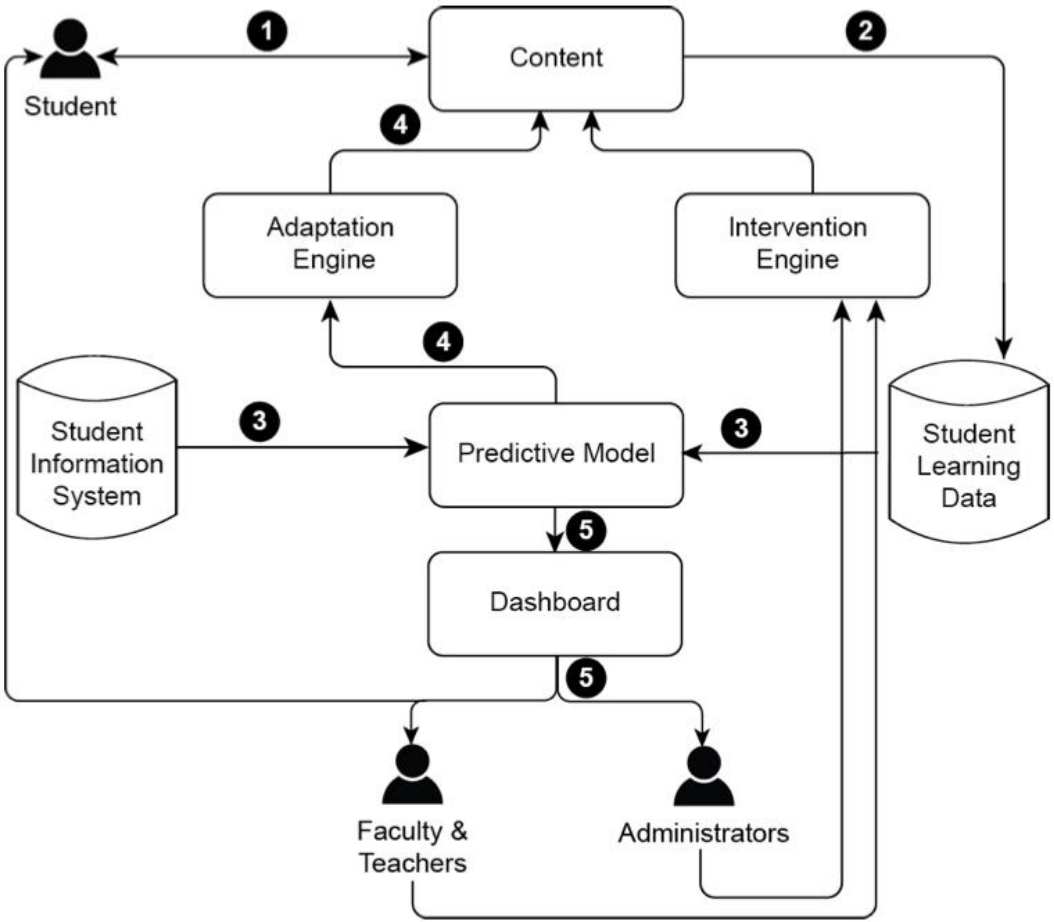

Analytics
Prediction
Adaptation
Intervention
Personalization


Data
Algorithm
Model
10010001110100000010100011011101011
10010011110111000000111110011010010
10000110110111110101001110000110100
11111101000011011100101011110000101
11001111110111111110010000111011011
01000011010011011000011000010001000
01010111001100111101100111010001011
00100001010110010100000100001001111
01110100111111001011101010101011110
10001000010110001010110101011100010
01001000010010101111001110000101000
01011000000100111010100101011101100
01101111101011110001010001010001000
01101001101101101000100010111100110
00010100000110011000110010001001011
10010101010001001110010101010111110
$f(\mathbf{x})$

# Advantages

- Iterative and automation
- Thomas H. Davenport wrote in *The Wall Street Journal* that with rapidly changing, growing volumes of data, "... you need fast-moving modeling streams to keep up." And you can do that with machine learning. He says, "Humans can typically create one or two good models a week; machine learning can create thousands of models a week."

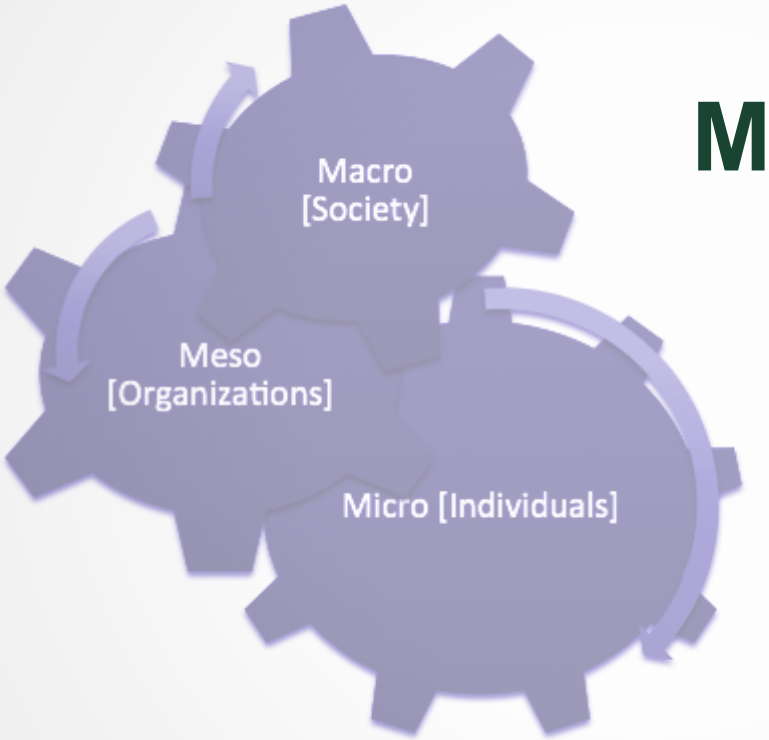
# The Components and Data Flow Through a Typical Adaptive Learning System



# Machine learning vs. Data mining

- Data mining discovers previously unknown patterns and knowledge: more exploratory
- Machine learning is used to reproduce known patterns and knowledge, automatically apply that to other data, and then automatically apply those results to decision making and actions
    - “explores the construction and study of algorithms that can learn from data”


MACHINE
LEARNING
POWER TO THE DATA




# Macro-Meso-Micro discussion

# Macro-Level Prediction and Intervention Analytics (Regional, State, National)

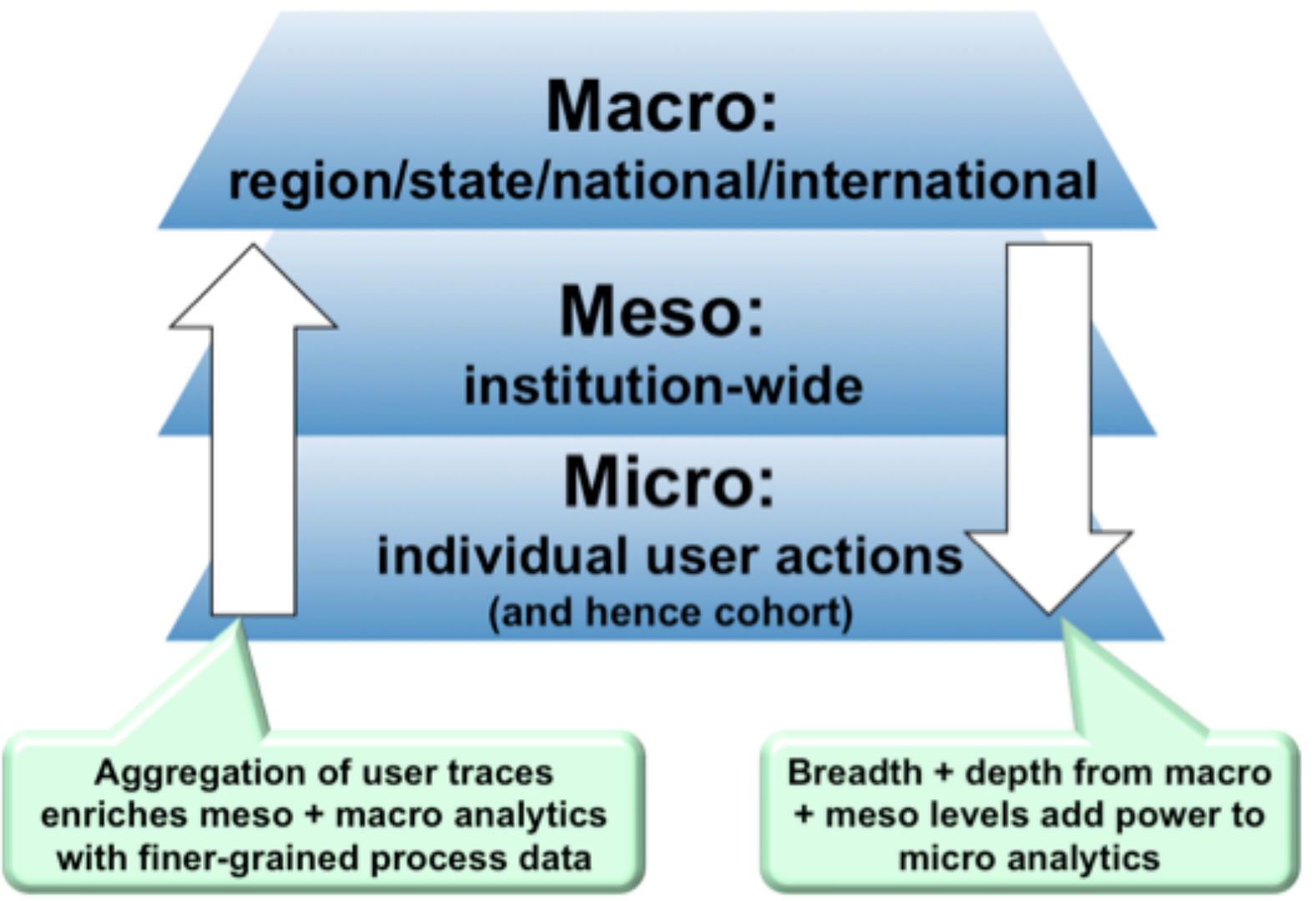
- Collect information from broad cross-institutional resources
- Large databases drawn from regional, state, and national sources show prediction and intervention models and results over long period of time.
- Prediction and intervention data leads to high-level policy decision-making and resource allocation models

# Meso Level Prediction and Intervention Analytics Institution Wide (Universities, School Districts)

- Patterns of learner information collected from institutional data sources create inferences where educators can predict behavior
- Intervention strategies can be implemented and measured at an institutional level to create different student outcomes
- Prediction and intervention data permit institutions to prioritize and optimize resource allocation

# Micro Level Prediction and Intervention Analytics Individual Users (Teachers and Students)

- Feedback and evaluation can be immediate for the student and teacher
- Interventions can be implemented in real-time
- Interventions can disrupt predictive trends


Macro:
region/state/national/international
Meso:
institution-wide
Micro:
individual user actions
(and hence cohort)
Aggregation of user traces enriches meso + macro analytics with finer-grained process data
Breadth + depth from macro + meso levels add power to micro analytics

# Using Predictive Analytics and Intervention to Increase High School Graduation Levels from 81% to 90%

- What three metrics or predictors should the State of Minnesota focus on to improve high school graduation rates?

- How can predictive analytics and intervention strategies be fine toned at the school system level to improve high school graduation rates?

- What real time predictive analytics can be used to provide meaningful intervention at the classroom level? Discuss possible intervention strategies.

# 15 min break.

# Split-group discussion

Two groups:

- Prediction and intervention for learners
- prediction and intervention for teachers

# Try to answer to questions like:

- How well will a student perform on an exam? / Does a student need help?
- What grade will a student get in the end of a course?
- Will a class of students like a specific type of exercises? Which teaching style is better for every student?
- What will be the dropout rate after the first year?
- Which classes should a student take?
- What will be the average grade of the class?
- What sequence of classes is the most effective for a specific student?
- Which student actions indicate satisfaction, engagement, learning progress, etc.?
- What features of an online learning environment lead to better learning?
- Does completing an assignment on one day or more affect performance?
- detect student behaviors as: when they are gaming the system,
  engaging in off-task behavior, or
  failing to answer a question correctly despite having a skill