## Modeling and Simulation

# Agent-Based Modeling & Applications

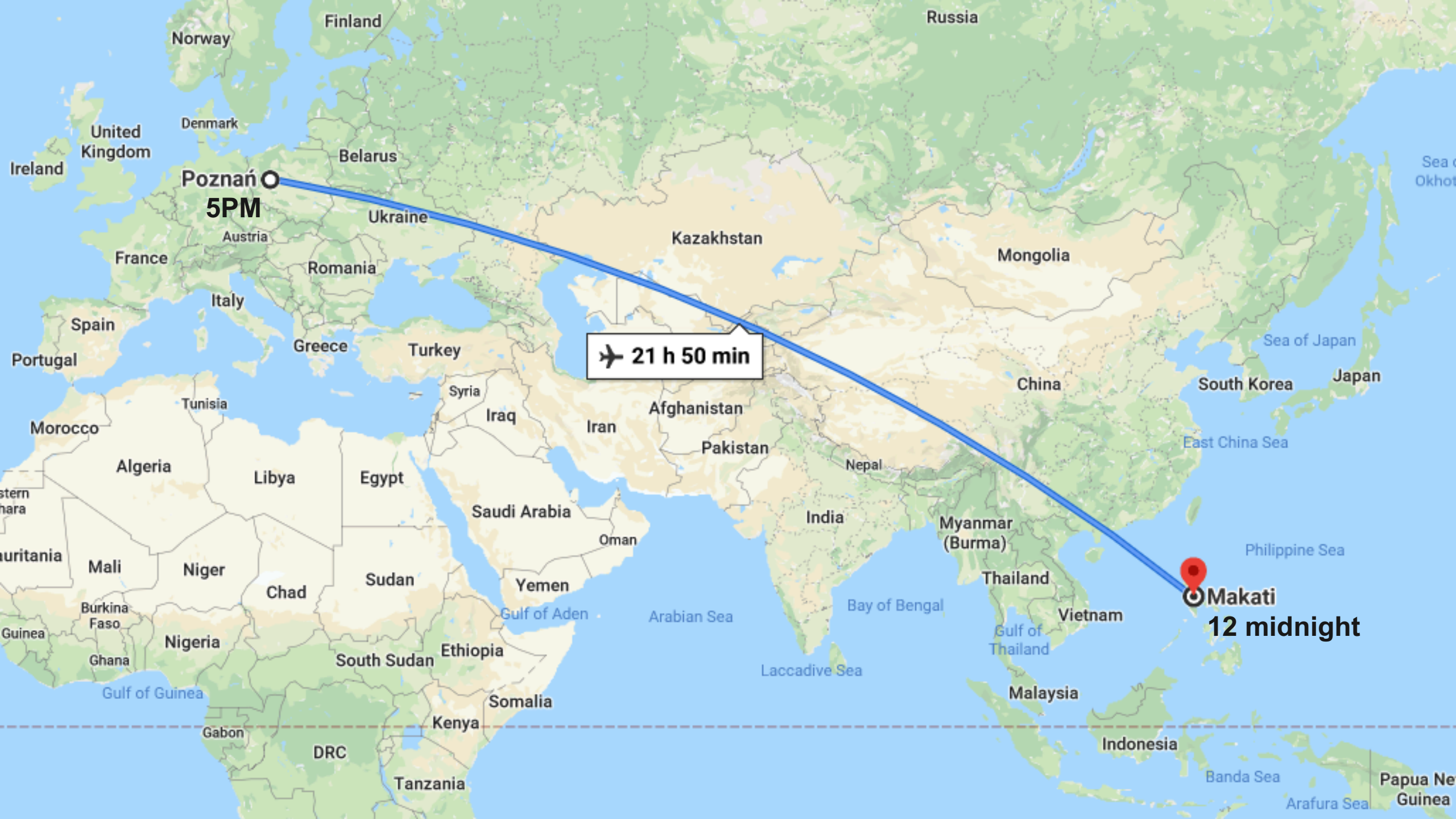
Erika Fille T. Legara, Ph.D.

eflegara

26 March 2020, 5:00pm

WiMLDS

Women in Machine Learning & Data Science


Poznań
5PM
21 h 50 min
Makati
12 midnight
Finland
Russia
Norway
Denmark
United Kingdom
Ireland
Belarus
Ukraine
Austria
France
Romania
Kazakhstan
Mongolia
Italy
Spain
Portugal
Greece
Turkey
Syria
Iraq
Iran
Afghanistan
Pakistan
China
South Korea
Japan
Sea of Japan
East China Sea
Tunisia
Morocco
Algeria
Libya
Egypt
Nepal
India
Saudi Arabia
Oman
Myanmar (Burma)
Philippine Sea
Thailand
Mali
Niger
Chad
Sudan
Yemen
Gulf of Aden
Arabian Sea
Bay of Bengal
Vietnam
Burkina Faso
Guinea
Nigeria
Ghana
Gulf of Thailand
South Sudan
Ethiopia
Laccadive Sea
Malaysia
Gulf of Guinea
Somalia
Kenya
Gabon
DRC
Indonesia
Tanzania
Banda Sea
Arafura Sea

I'm a physicist[1] by training, a *complex systems scientist* and educator by profession. I was based in Singapore for 5.5 years, working at a government think tank[2]. We built models[3] and simulations for partners and clients; this allowed me to work with a lot of data[4]. I moved back to the Philippines to join a business school, AIM. I also am part of AIM's **Analytics Computing, and Complex Systems Laboratory (ACCeSs)**.

[1] Ph.D. in Physics from the University of the Philippines (Diliman Campus).
[2] Scientist at the Institute of High Performance Computing, Agency for Science, Technology, and Research, Singapore.
[3] Statistical models, mathematical and numerical models, machine learning and neural network models, agent-based models
[4] Social network data, transport data, weather data, geographic/GIS data, marketing data, and other industry data.

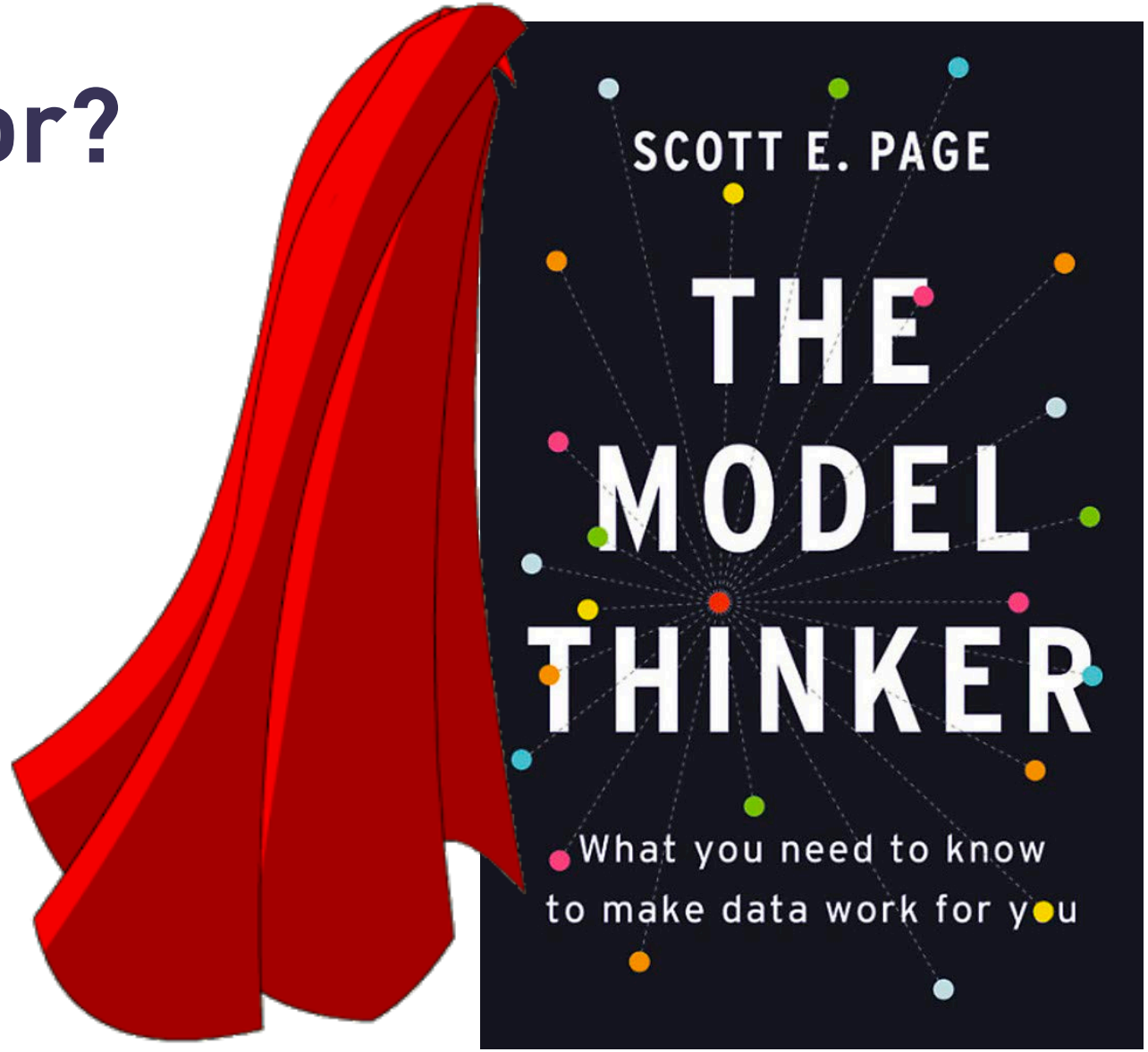
# What is a model?

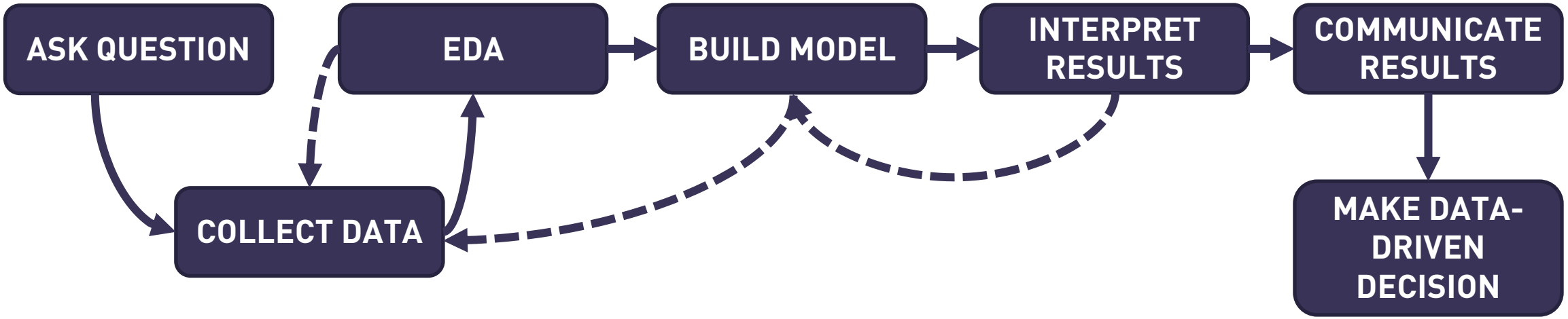
# What is it not?

# What are models for?

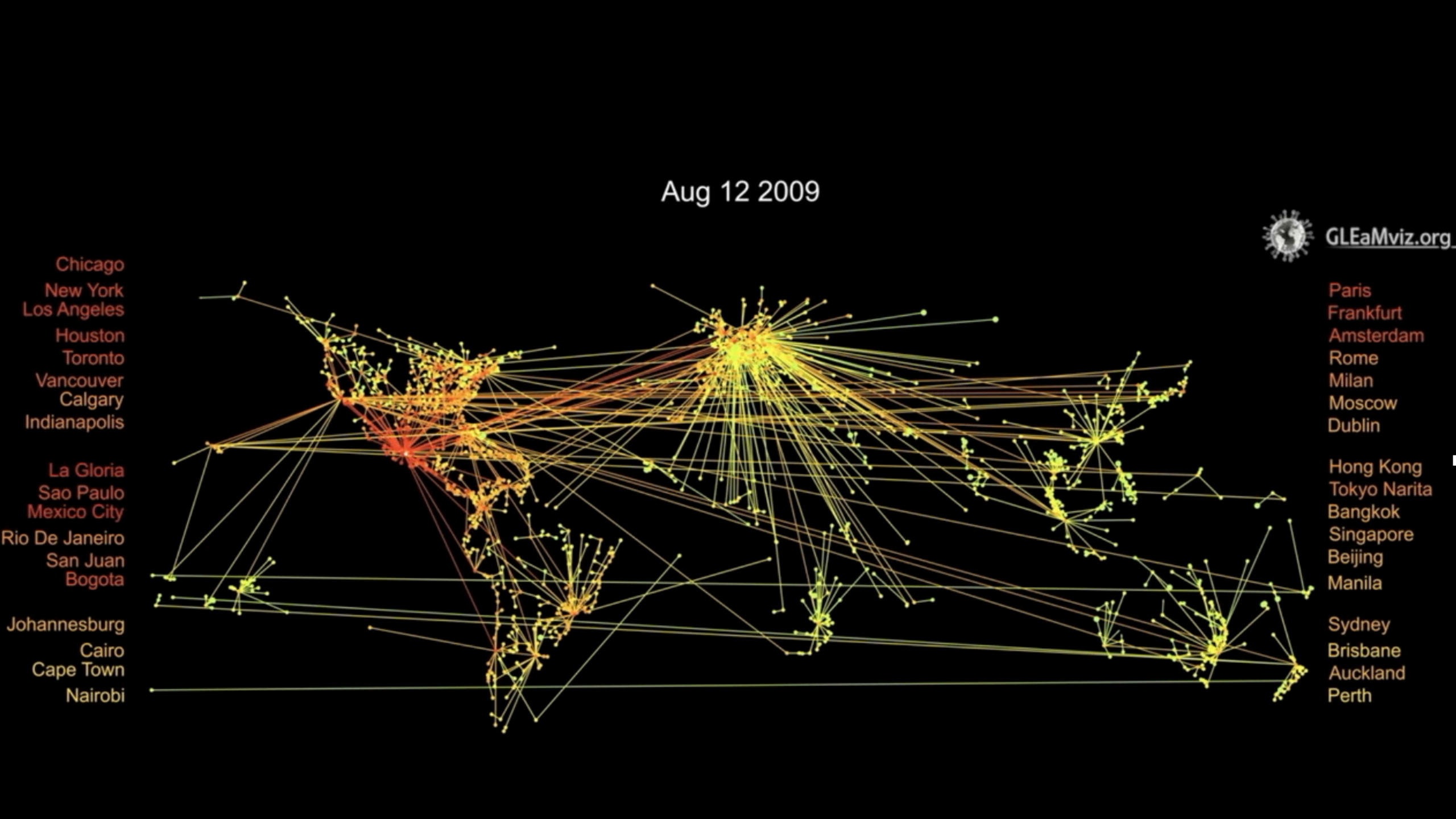
# What are models for?

**R** eason
**E** xplain
**D** esign
**C** ommunicate
**A** ct
**P** redict
**E** xplore



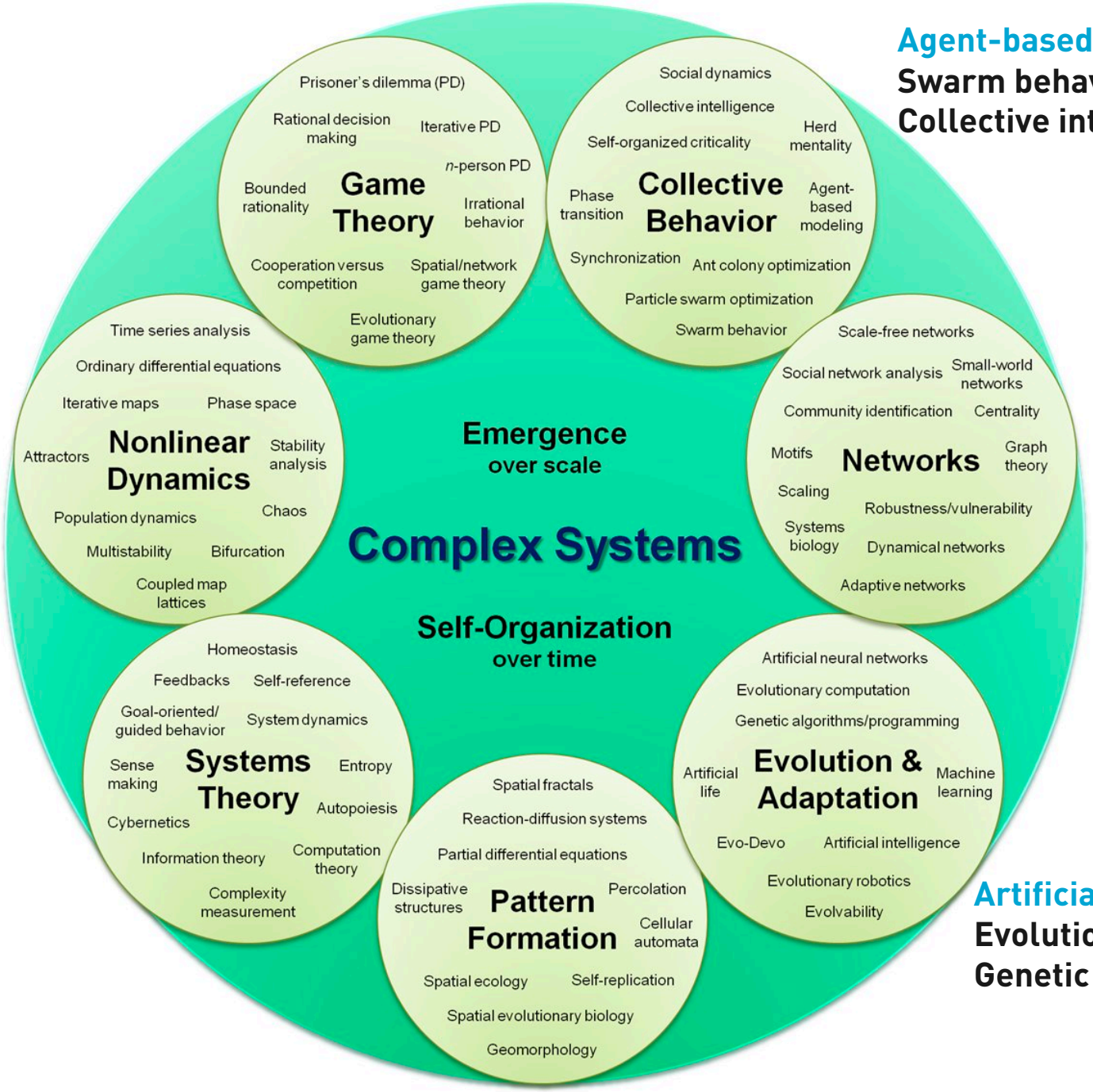

ASK QUESTION
EDA
BUILD MODEL
INTERPRET RESULTS
COMMUNICATE RESULTS
COLLECT DATA
MAKE DATA-DRIVEN DECISION

WiMLDS
Women in Machine Learning & Data Science


Aug 12 2009
GLEaMviz.org
Chicago
New York
Los Angeles
Houston
Toronto
Vancouver
Calgary
Indianapolis
La Gloria
Sao Paulo
Mexico City
Rio De Janeiro
San Juan
Bogota
Johannesburg
Cairo
Cape Town
Nairobi
Paris
Frankfurt
Amsterdam
Rome
Milan
Moscow
Dublin
Hong Kong
Tokyo Narita
Bangkok
Singapore
Beijing
Manila
Sydney
Brisbane
Auckland
Perth

**BUILD MODEL**

Source: Hiroki Sayama




**Agent-based modeling**
**Swarm behavior**
**Collective intelligence**

**Social network analysis**
**Robustness/vulnerability**
**Information spreading**
**Dynamics networks**

**Artificial Intelligence / ML**
**Evolutionary robotics**
**Genetic algorithms**

WiMLDS
Women in Machine Learning & Data Science

# Agent-based modeling

- “Bottom-Up” approach
- Simulate actions and interactions of autonomous agents
- No "representative agent" assumption
- Agents have diverse, dynamic, and interdependent behavior

# Agent-based modeling

- “Bottom-Up” approach
- Simulate actions and interactions of autonomous agents
- No "representative agent" assumption
- Agents have diverse, dynamic, and interdependent behavior

## Agent Attributes

Personality
Characteristics
Status

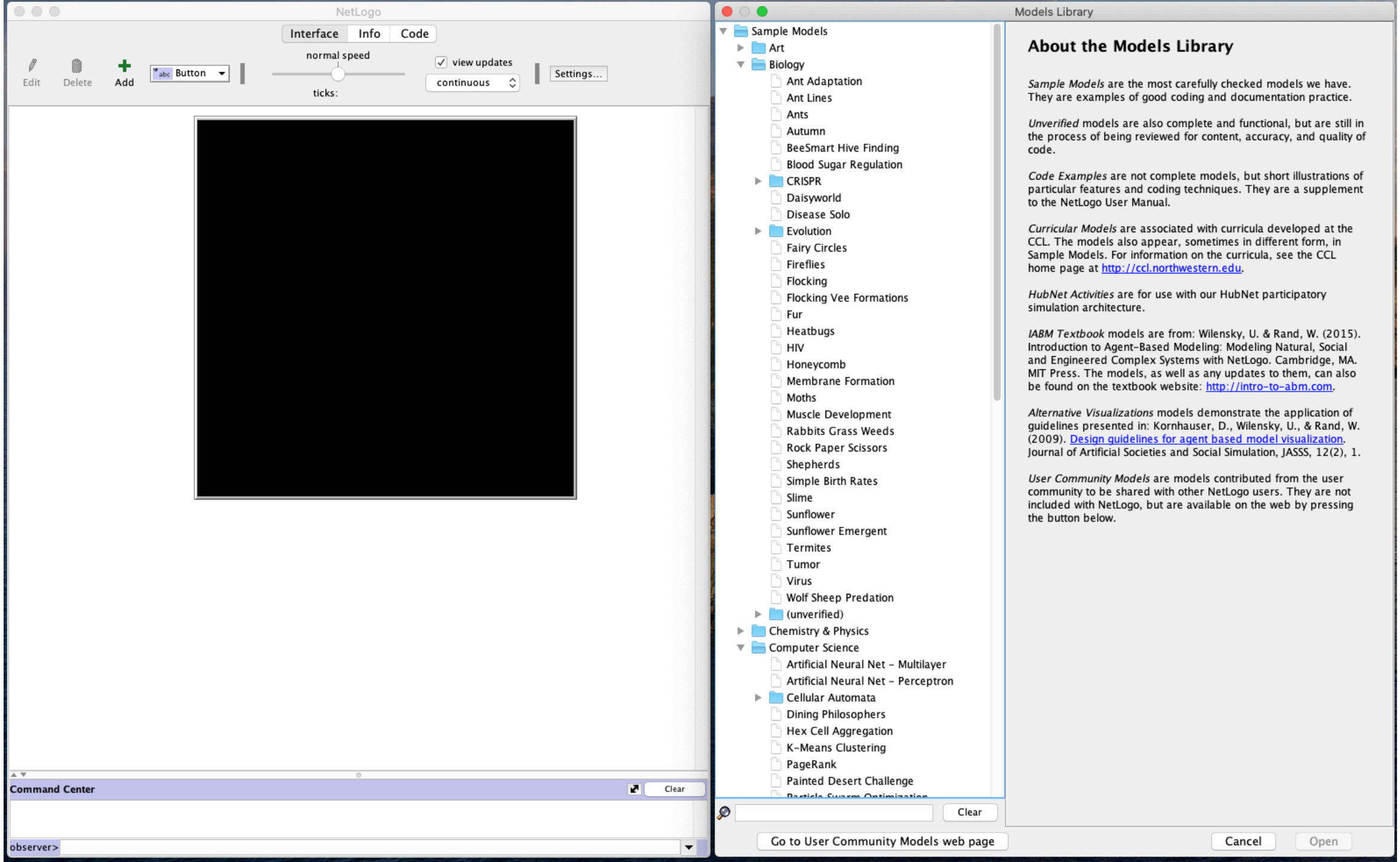
## Rules/Behavior

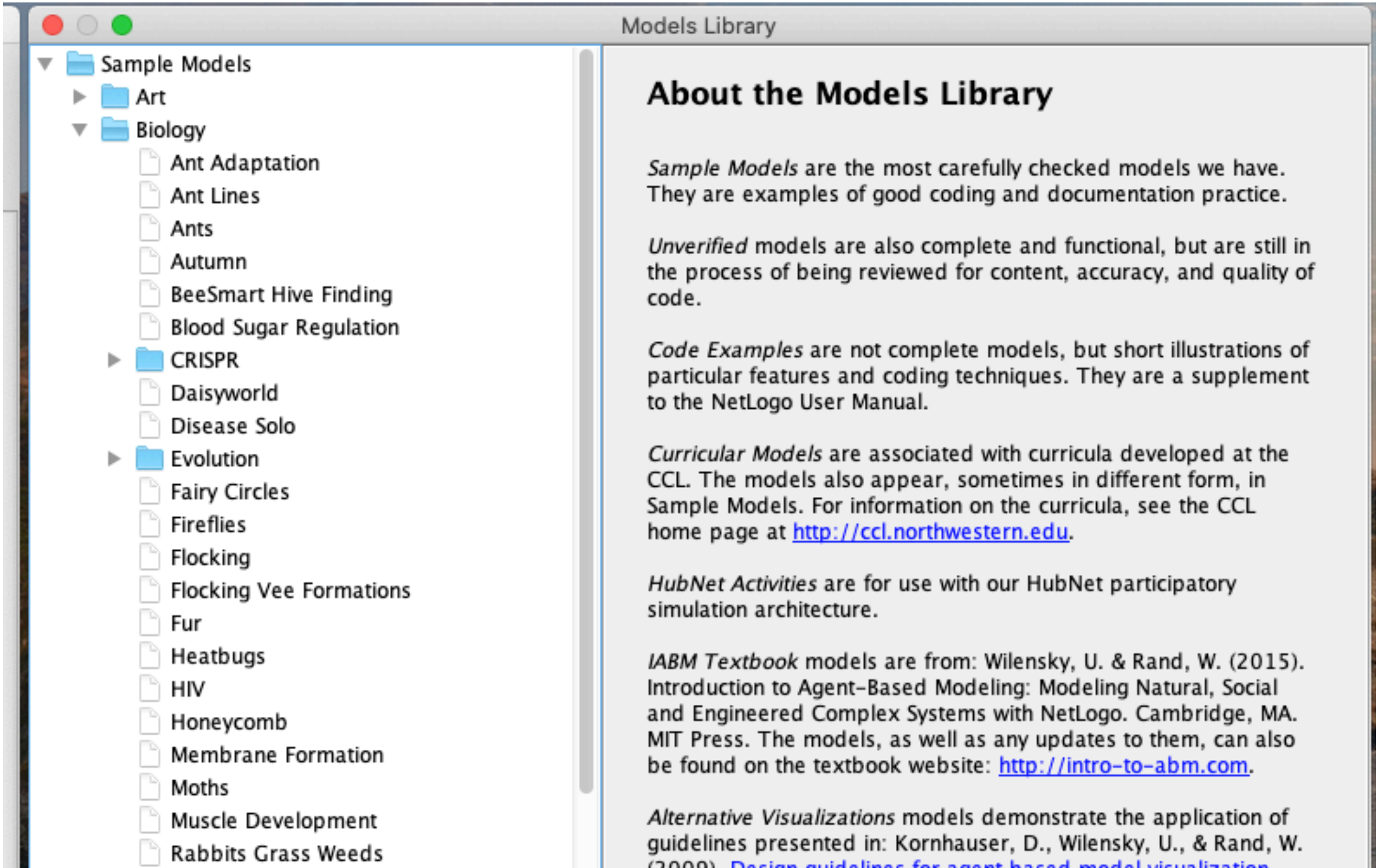
Behavior towards environment
Interaction with other agents

# Software for Agent-based Modeling

https://ccl.northwestern.edu/netlogo/


NetLogo is a multi-agent programmable modelling environment. It is used by tens of thousands of students, teachers and researchers worldwide.


NetLogo
Interface
Info
Code
Edit
Delete
Add
Button
normal speed
ticks:
view updates
continuous
Settings...
Command Center
Clear
observer>
Models Library
Sample Models
Art
Biology
Ant Adaptation
Ant Lines
Ants
Autumn
BeeSmart Hive Finding
Blood Sugar Regulation
CRISPR
Daisyworld
Disease Solo
Evolution
Fairy Circles
Fireflies
Flocking
Flocking Vee Formations
Fur
Heatbugs
HIV
Honeycomb
Membrane Formation
Moths
Muscle Development
Rabbits Grass Weeds
Rock Paper Scissors
Shepherds
Simple Birth Rates
Slime
Sunflower
Sunflower Emergent
Termites
Tumor
Virus
Wolf Sheep Predation
(unverified)
Chemistry & Physics
Computer Science
Artificial Neural Net - Multilayer
Artificial Neural Net - Perceptron
Cellular Automata
Dining Philosophers
Hex Cell Aggregation
K-Means Clustering
PageRank
Painted Desert Challenge
Clear
About the Models Library
Sample Models are the most carefully checked models we have. They are examples of good coding and documentation practice.
Unverified models are also complete and functional, but are still in the process of being reviewed for content, accuracy, and quality of code.
Code Examples are not complete models, but short illustrations of particular features and coding techniques. They are a supplement to the NetLogo User Manual.
Curricular Models are associated with curricula developed at the CCL. The models also appear, sometimes in different form, in Sample Models. For information on the curricula, see the CCL home page at http://ccl.northwestern.edu.
HubNet Activities are for use with our HubNet participatory simulation architecture.
IABM Textbook models are from: Wilensky, U. & Rand, W. (2015). Introduction to Agent-Based Modeling: Modeling Natural, Social and Engineered Complex Systems with NetLogo. Cambridge, MA. MIT Press. The models, as well as any updates to them, can also be found on the textbook website: http://intro-to-abm.com.
Alternative Visualizations models demonstrate the application of guidelines presented in: Kornhauser, D., Wilensky, U., & Rand, W. (2009). Design guidelines for agent based model visualization. Journal of Artificial Societies and Social Simulation, JASSS, 12(2), 1.
User Community Models are models contributed from the user community to be shared with other NetLogo users. They are not included with NetLogo, but are available on the web by pressing the button below.
Go to User Community Models web page
Cancel
Open


Models Library
Sample Models
Art
Biology
Ant Adaptation
Ant Lines
Ants
Autumn
BeeSmart Hive Finding
Blood Sugar Regulation
CRISPR
Daisyworld
Disease Solo
Evolution
Fairy Circles
Fireflies
Flocking
Flocking Vee Formations
Fur
Heatbugs
HIV
Honeycomb
Membrane Formation
Moths
Muscle Development
Rabbits Grass Weeds
About the Models Library
Sample Models are the most carefully checked models we have. They are examples of good coding and documentation practice.
Unverified models are also complete and functional, but are still in the process of being reviewed for content, accuracy, and quality of code.
Code Examples are not complete models, but short illustrations of particular features and coding techniques. They are a supplement to the NetLogo User Manual.
Curricular Models are associated with curricula developed at the CCL. The models also appear, sometimes in different form, in Sample Models. For information on the curricula, see the CCL home page at http://ccl.northwestern.edu.
HubNet Activities are for use with our HubNet participatory simulation architecture.
IABM Textbook models are from: Wilensky, U. & Rand, W. (2015). Introduction to Agent-Based Modeling: Modeling Natural, Social and Engineered Complex Systems with NetLogo. Cambridge, MA. MIT Press. The models, as well as any updates to them, can also be found on the textbook website: http://intro-to-abm.com.
Alternative Visualizations models demonstrate the application of guidelines presented in: Kornhauser, D., Wilensky, U., & Rand, W.

# Agent-based modeling

## Spread of the HIV

This model **simulates the spread of the human immunodeficiency virus (HIV)**, via sexual transmission, through a small isolated human population. It illustrates the effects of certain sexual practices across a population.

## Virus on a Network

This model demonstrates the **spread of a virus through a network**. Each node may be in one of three states: susceptible, infected, or resistant. In the academic literature such a model is sometimes referred to as an SIR model for epidemics.

# Agent-based modeling

## Spread of the HIV

```
globals [

turtles-own [
  infected?           ;; If true, the person is infected.  It may be known or unknown.
  known?              ;; If true, the infection is known (and infected? must also be true).
  infection-length    ;; How long the person has been infected.
  coupled?            ;; If true, the person is in a sexually active couple.
  couple-length       ;; How long the person has been in a couple.
  ;; the next four values are controlled by sliders
  commitment          ;; How long the person will stay in a couple-relationship.
  coupling-tendency   ;; How likely the person is to join a couple.
  condom-use          ;; The percent chance a person uses protection.
  test-frequency      ;; Number of times a person will get tested per year.
  partner             ;; The person that is our current partner in a couple.
]
```

initial-people 300

setup | go

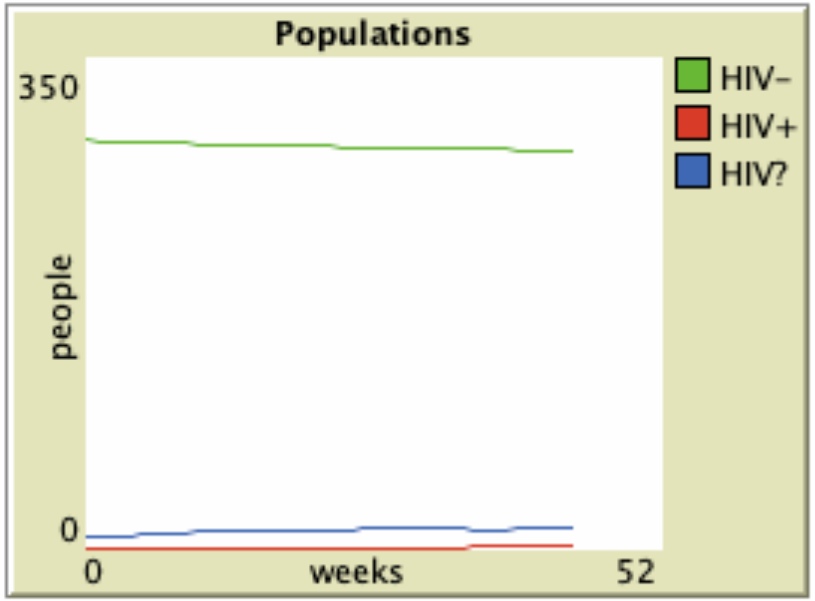
% infected 5.67

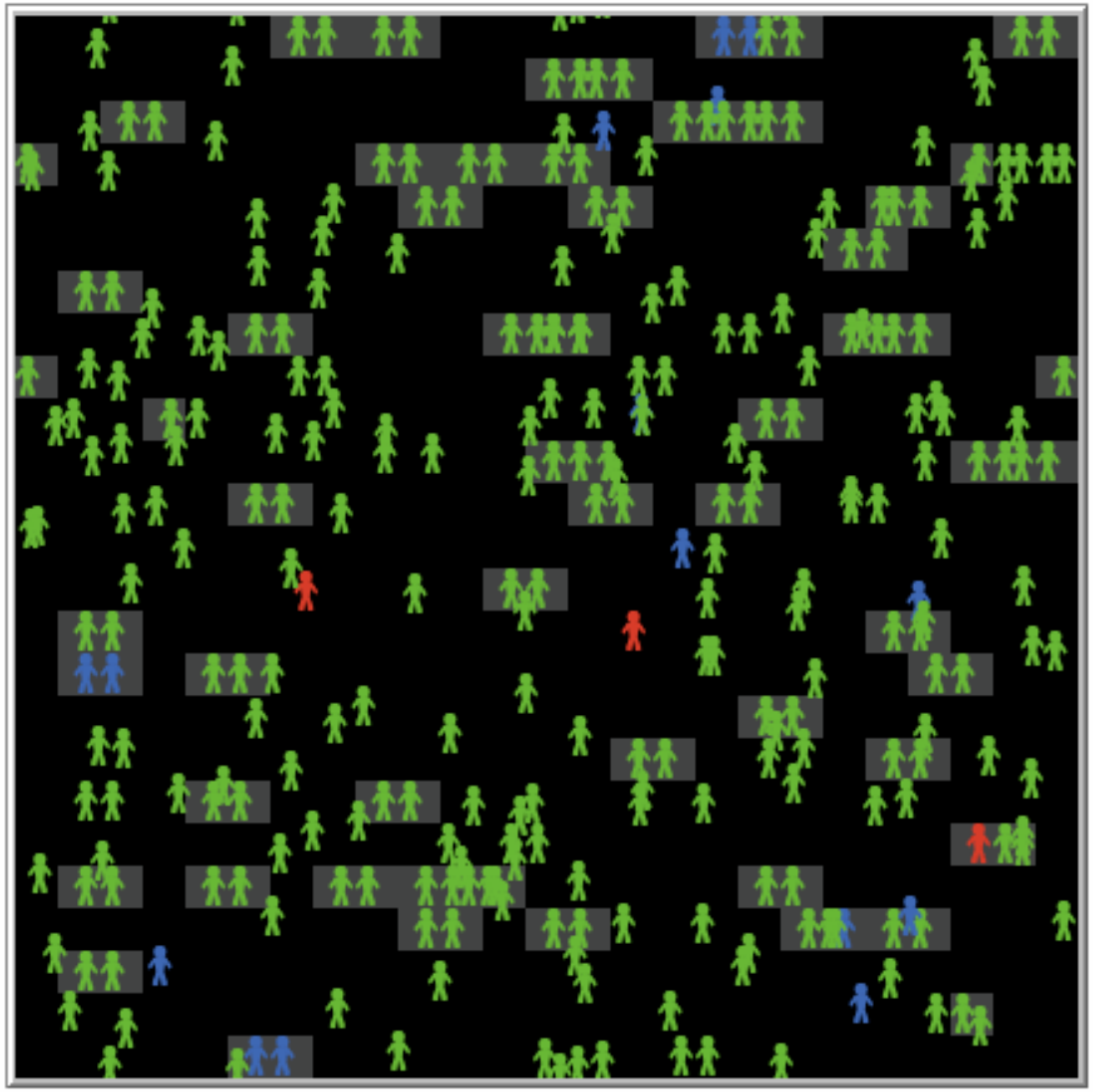
average-coupling-tendency 6

average-commitment 20 weeks

average-condom-use 4

average-test-frequency 0.09 times/year



# Agent-based modeling

## Virus on a Network

```
turtles-own
[
  infected?           ;; if true, the turtle is infectious
  resistant?          ;; if true, the turtle can't be infected
  virus-check-timer   ;; number of ticks since this turtle's last virus-check
]

to setup
  clear-all
  setup-nodes
  setup-spatially-clustered-network
  ask n-of initial-outbreak-size turtles
    [ become-infected ]
  ask links [ set color white ]
  reset-ticks
end
```

number-of-nodes 150

average-node-degree 6

initial-outbreak-size 3

setup

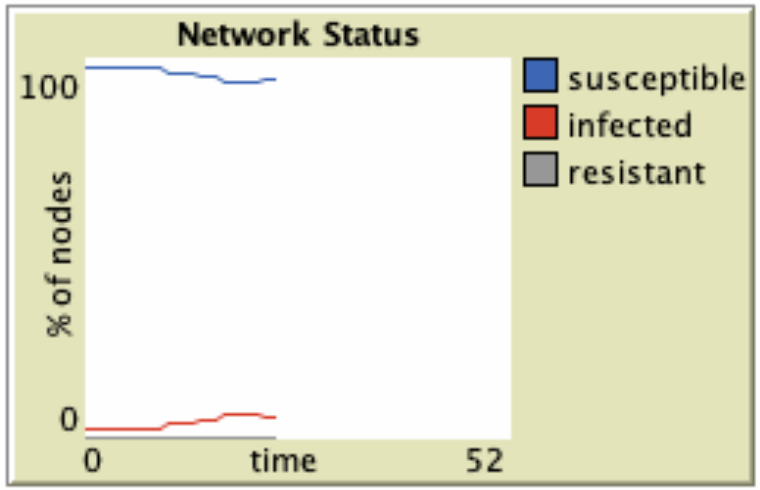
go

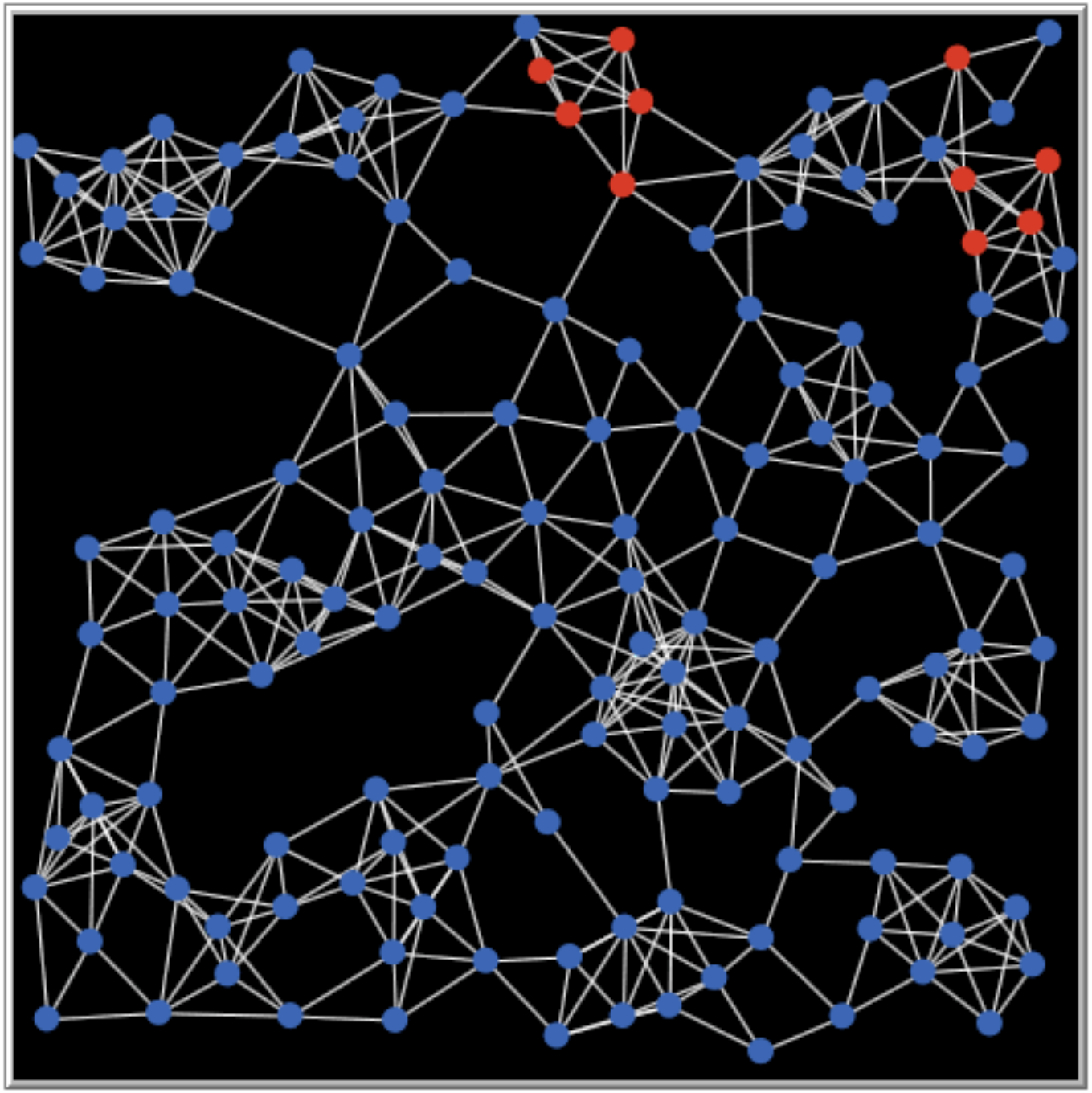
virus-spread-chance 2.2 %

virus-check-frequency 1 ticks

recovery-chance 3.0 %

gain-resistance-chance 5 %



# Network Science

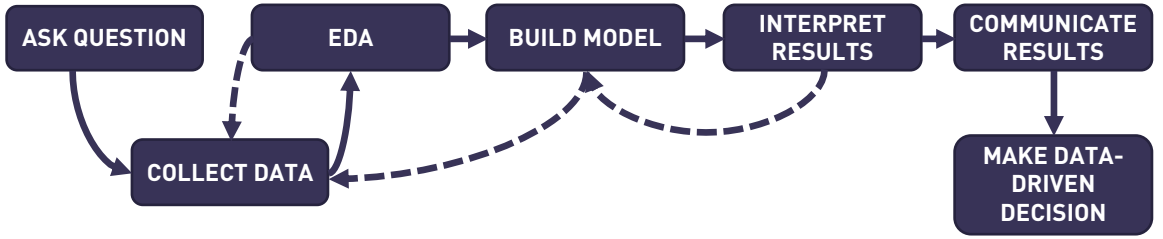

In this repository, I share some of the Jupyter notebooks I use in the Network Science class I offer under the Asian Institute of Management's MSc. in Data Science program where Network Science is a core data science course.

## Table of Contents

1. Intro to Network Analysis
2. Intro to Network Models
3. Exploring Network Vulnerability and Robustness with Centrality Measures
4. Community Detection in Complex Networks
5. Target and Attack in Complex Networks
6. Application: Information Flow
7. Application: Social Distancing (An Elementary Model)

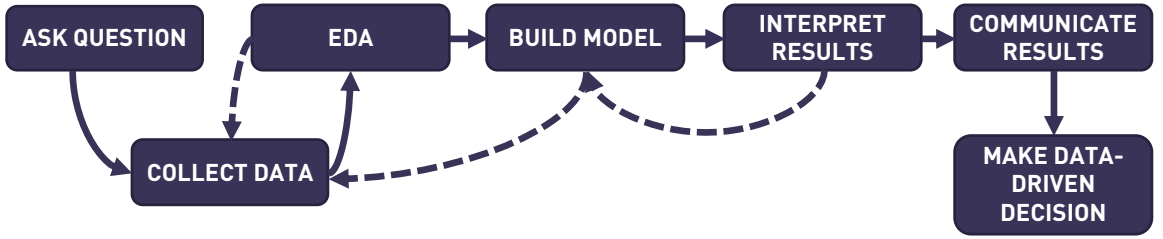
# A Use Case: A multi-sectorial data-driven project

**SINGAPORE SMART FARE CARD**




- What are the critical origin-destination (OD) pairs?
- How often are people unable to board trains?
- How do we probe route choices?
- How can we model various scenarios such as **travel delays** resulting from overcrowding of platforms?
- Can we predict the efficacy of congestion-reducing strategies?

# SMART FARE CARD




Journey ID
Card ID
Passenger Type
Boarding Station
Alighting Station
Tap In Date
Tap In Time
Tap Out Date
Tap Out Time
Transport Mode

- What are the critical origin-destination (OD) pairs?
- How often are people unable to board trains?
- How do we probe route choices?
- How can we model various scenarios such as **travel delays** resulting from overcrowding of platforms?
- Can we predict the efficacy of congestion-reducing strategies?

WiMLDS
Women in Machine Learning & Data Science

# What are the critical origin-destination (OD) pairs?

Journey ID

Card ID

Passenger Type

Boarding Station

Alighting Station

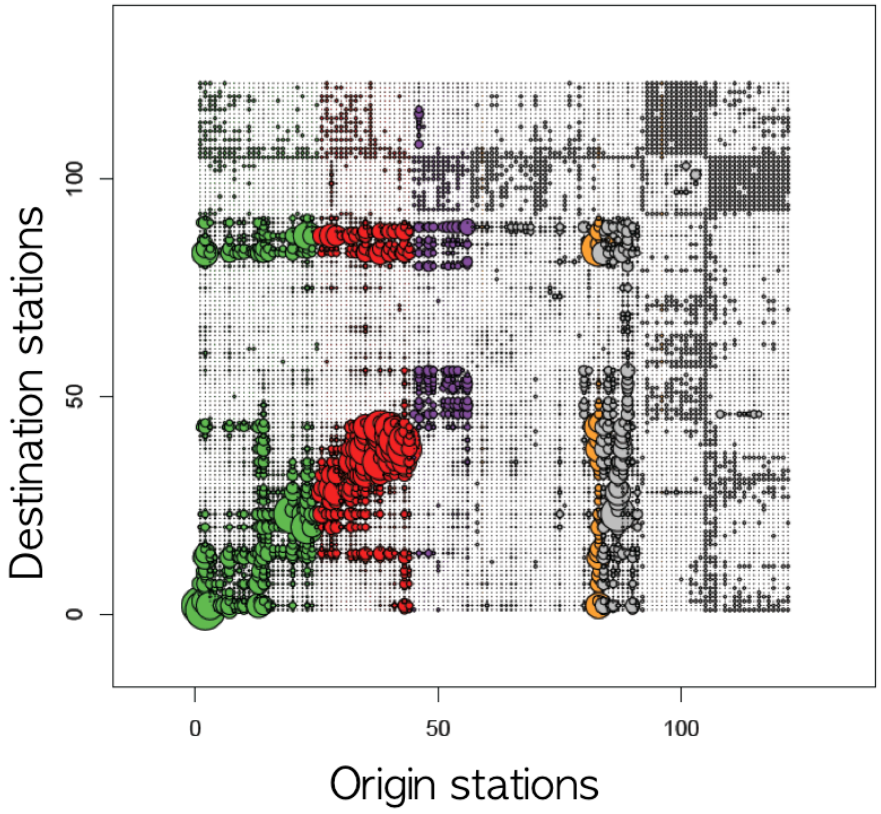
Tap In Date

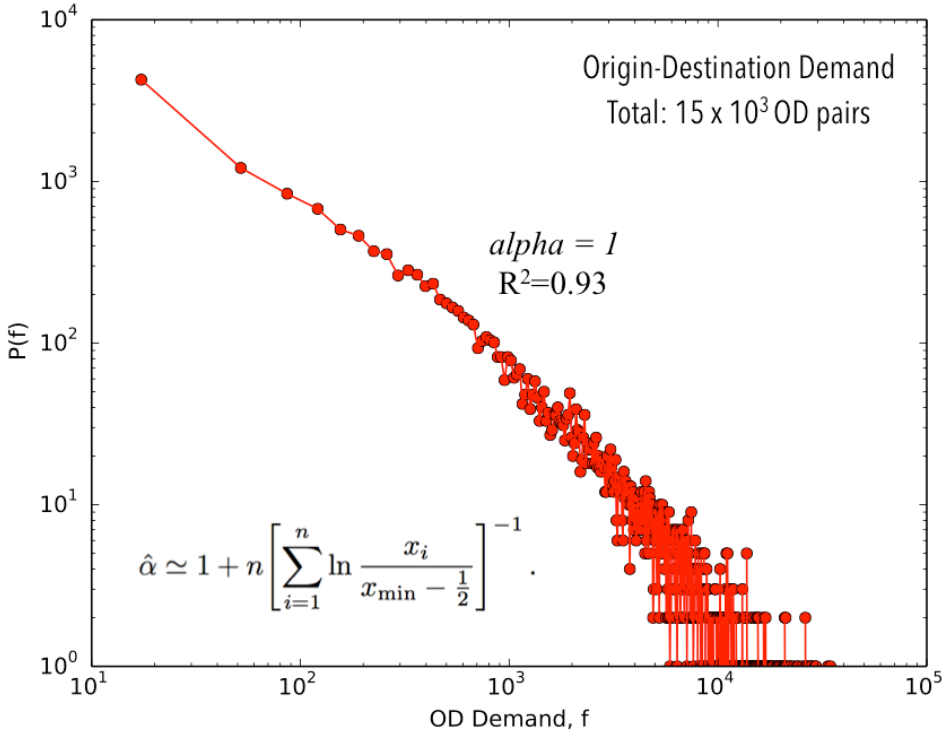
Tap In Time

Tap Out Date

Tap Out Time

Transport Mode




WiMLDS
Women in Machine Learning & Data Science

EF Legara, KK Lee, GG Hung, and C Monteorla, "Mechanism-based model of a mass rapid transit system: A perspective," Int. J. Mod. Phys. Conf. Ser. 36, 1560011, 2015.

# Develop a **transport simulation tool** to explore "what-if" scenarios

Keywords: Scheduling System, Asset Management, Capacity Monitoring, Robustness/Resilience of System, Scenario Modeling

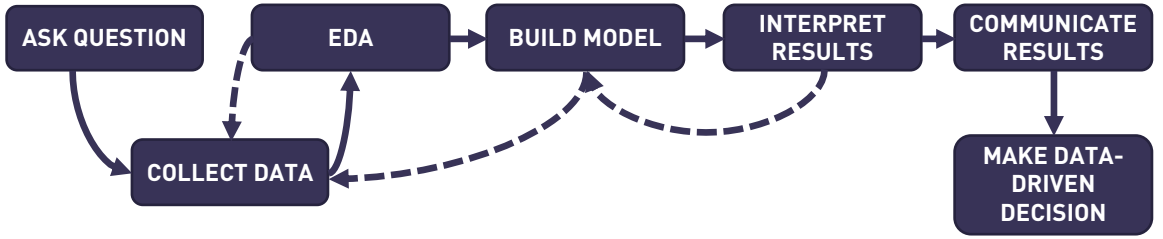


Institute of High Performance Computing
A*STAR

N Othman, **EF Legara**, V Selvam, and **C Monterola**, "A Data-Driven Agent-Based Model of Congestion and Scaling Dynamics of Rapid Transit Systems," J of Comp. Sci. 10, pp. 338-350 (2015).
N Othman, **EF Legara**, V Selvam, and **C Monterola**, "Simulating Congestion Dynamics of Train Rapid Transit using Smart Card Data," Procedia Computer Science 29, pp. 1610–1620 (2014).
**EF Legara**, **C Monterola**, KK Lee, GG Hung, "Critical capacity, travel time delays and travel time distribution of rapid mass transit systems," *Physica A* **406**, pp. 100-106 (2014).

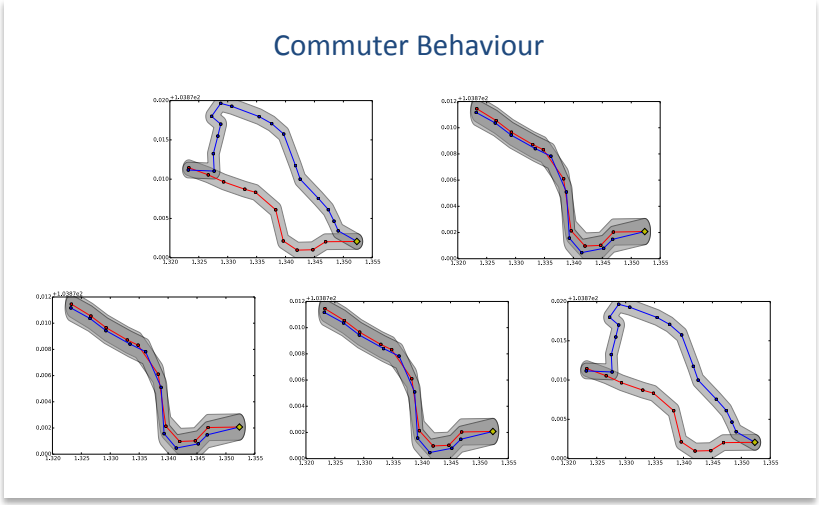
WiMLDS
Women in Machine Learning & Data Science

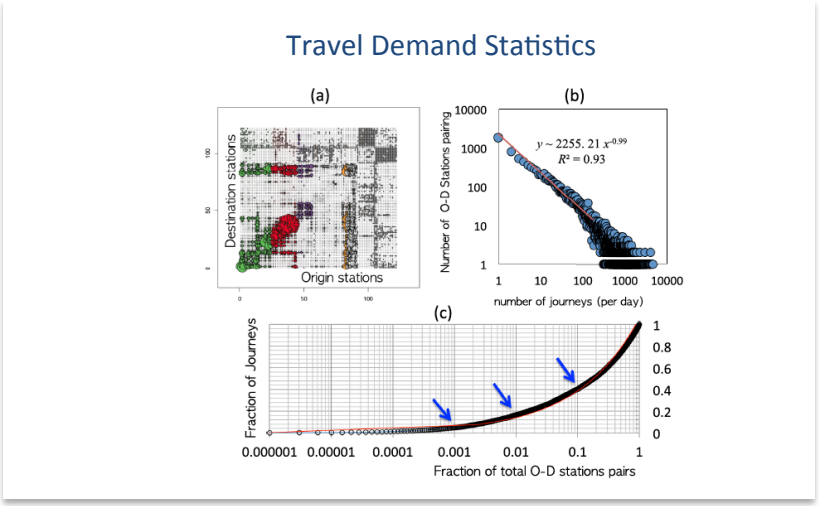
**SINGAPORE SMART FARE CARD**

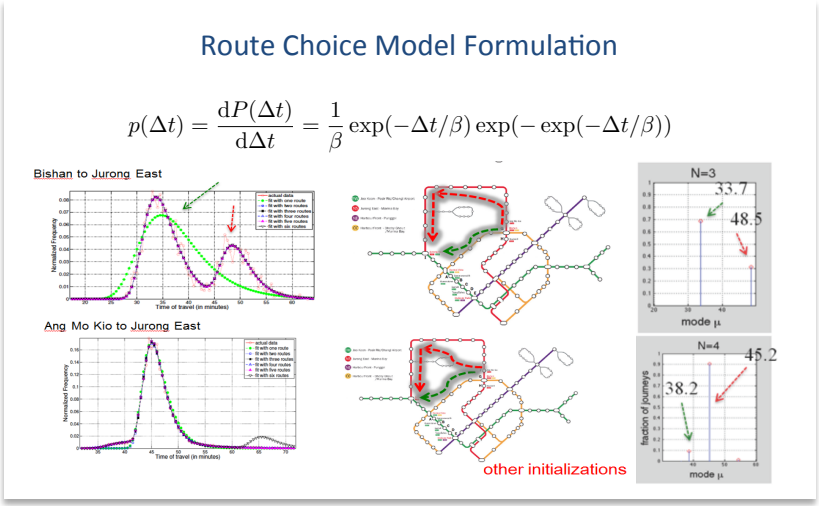



- What are the critical origin-destination (OD) pairs?
- How often are people unable to board trains?
- How do we probe route choices?
- How can we model various scenarios such as **travel delays** resulting from overcrowding of platforms?
- Can we predict the efficacy of congestion-reducing strategies?

# Data-Driven Transport Research





**C Monterola**, **EF Legara**, P Di, KK Lee, GG Hung, "Non-invasive procedure to probe the route choices of commuters in rail transit systems Procedia Computer Science 80, 2016, Pages 2387–2391. Proceedings of International Conference on Computational Science (2016).

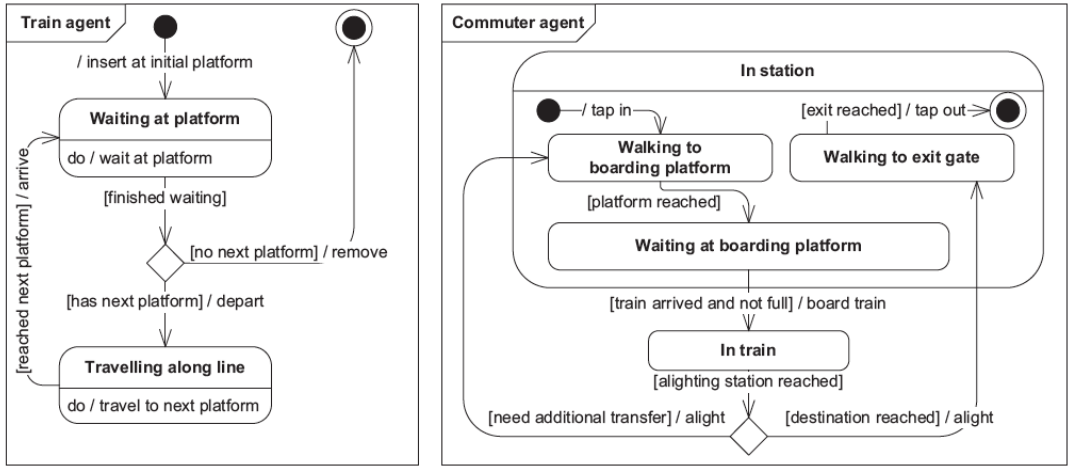
**EF Legara**, N Othman, GG Hung, and **C Monterola**, "Absence of a dominant transit route drives asymmetry in the route paths of commuter," 2014 European Conference on Complex Systems, IMT Lucca, Italy, 12-26 September 2014.

**EF Legara**, **C Monterola**, KK Lee, GG Hung, "Critical capacity, travel time delays and travel time distribution of rapid mass transit systems," *Physica A* **406**, pp. 100-106 (2014).

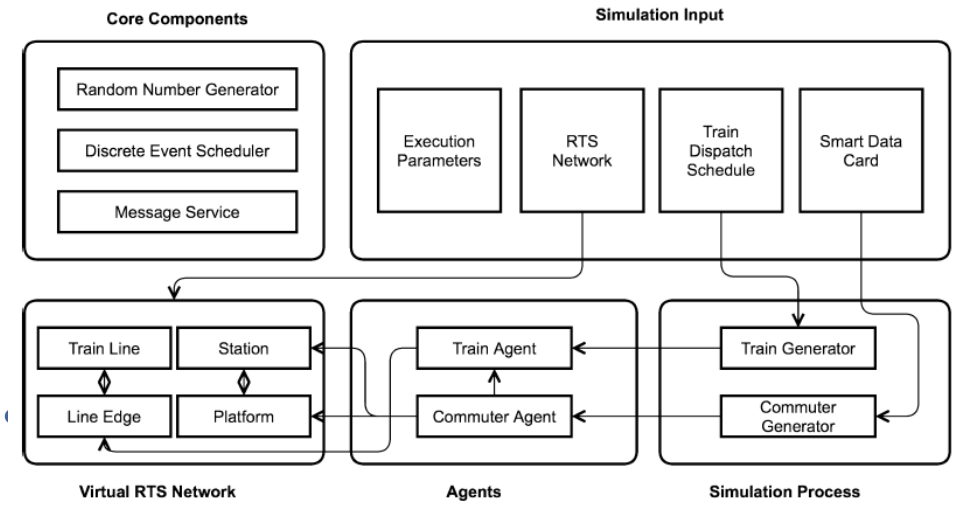

WiMLDS
Women in Machine Learning & Data Science

# Data-Driven Transport Research

## Agent-Based Modeling




## Simulation Architecture



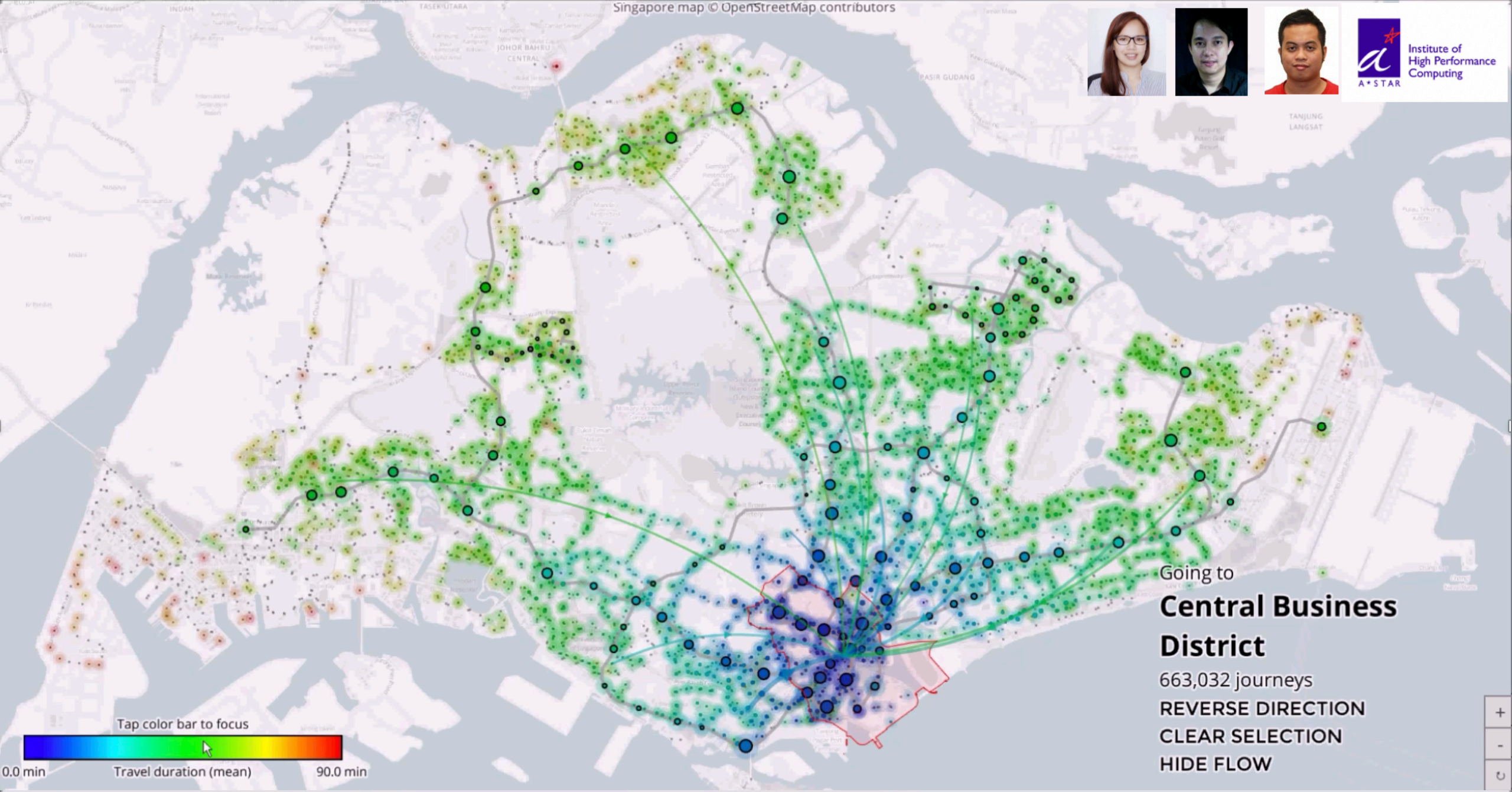

## Model Input and Parameters

| Input | Description |
|---|---|
| Random engine | • Mersenne Twister (MT19937) implementation. |
| Random seed | • Re-use enables reproducibility of simulation. |
| Rail network | • Directed graph connecting stations via train lines.<br>• Vertices are stations.<br>• Edges are annotated with travel duration and train line. |
| Dwell times | • Waiting time at each platform.<br>• Values are longer for interchanges.<br>• Currently uniformly distributed. |
| Train capacities | • Maximum allowable passenger load per train.<br>• Trains on different lines may have differing values. |
| Train schedules | • At least one schedule for each launching station.<br>• Mean &standard deviation for time to next launch given per schedule.<br>• Additional trains are inserted during peak periods. |
| Walking times | • Mean duration given for each walkway.<br>• One walkway for each gate-to-platform combination.<br>• One walkway for each platform-to-platform transfer.<br>• Variable reduction due to symmetry and common station layouts.<br>• Currently normally distributed with standard deviation = 1/3 mean. |
| Safety distance | • Minimum headway between trains. |
| Travel demand | • Temporal demand given for each origin–destination combination.<br>• Commuters are sampled using Monte Carlo method on empirical data. |
| Population scaling | • Ratio of simulated commuter population over empirical population. |
| Route choices | • Pre-computed probability distribution of commuter route choices.<br>• Up to three routes for each origin–destination combination.<br>• Each route is assigned a probability.<br>• Probabilities must sum to unity for each origin–destination combination. |


N Othman, **EF Legara**, V Selvam, and **C Monterola**, "A Data-Driven Agent-Based Model of Congestion and Scaling Dynamics of Rapid Transit Systems," J of Comp. Sci. 10, pp. 338-350 (2015).

N Othman, **EF Legara**, V Selvam, and **C Monterola**, "Simulating Congestion Dynamics of Train Rapid Transit using Smart Card Data," Procedia Computer Science 29, pp. 1610–1620 (2014).

**EF Legara**, **C Monterola**, KK Lee, GG Hung, "Critical capacity, travel time delays and travel time distribution of rapid mass transit systems," *Physica A* **406**, pp. 100-106 (2014).

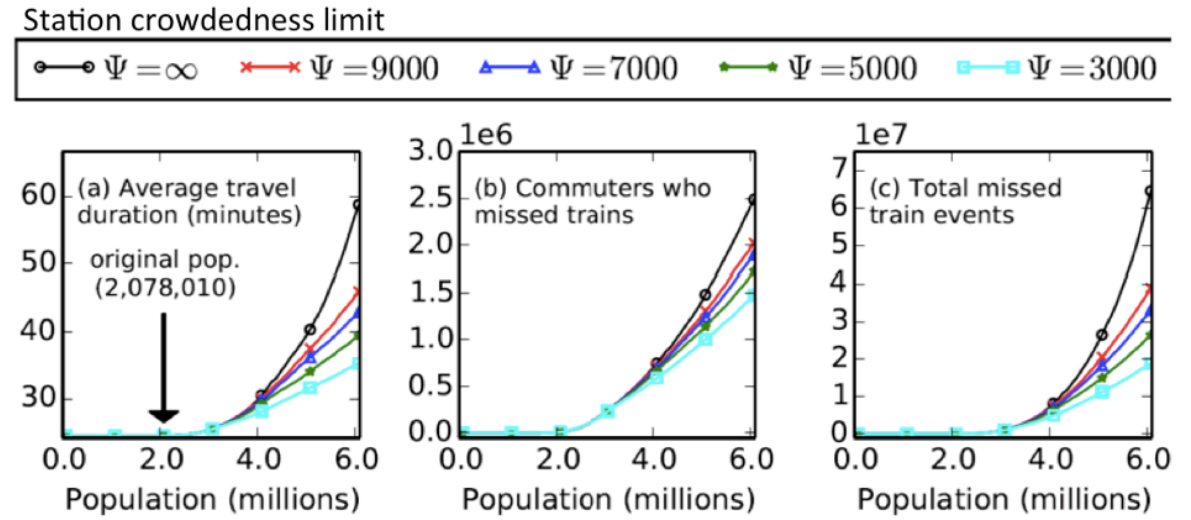

Singapore map © OpenStreetMap contributors
Institute of High Performance Computing
A*STAR
Going to
Central Business District
663,032 journeys
REVERSE DIRECTION
CLEAR SELECTION
HIDE FLOW
Tap color bar to focus
0.0 min
Travel duration (mean)
90.0 min
SEARCH
ANALYSE
SIMULATE
VALIDATE
?

# Scenario Modeling

N Othman, **EF Legara**, V Selvam, and **C Monterola**, "A Data-Driven Agent-Based Model of Congestion and Scaling Dynamics of Rapid Transit Systems," J of Comp. Sci. 10, pp. 338-350 (2015).

N Othman, **EF Legara**, V Selvam, and **C Monterola**, "Simulating Congestion Dynamics of Train Rapid Transit using Smart Card Data," Procedia Computer Science 29, pp. 1610–1620 (2014).

*Can the transit system take in **more commuters? How much more?***




Population Increase

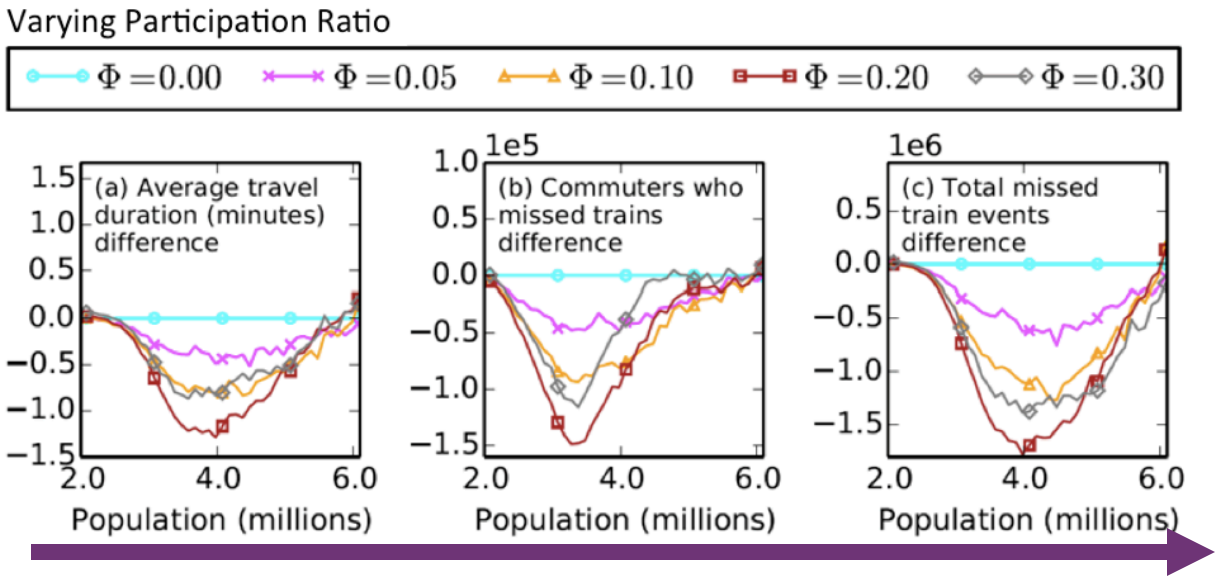
WiMLDS
Women in Machine Learning & Data Science

# Scenario Modeling

N Othman, **EF Legara**, V Selvam, and **C Monterola**, "A Data-Driven Agent-Based Model of Congestion and Scaling Dynamics of Rapid Transit Systems," J of Comp. Sci. 10, pp. 338-350 (2015).

N Othman, **EF Legara**, V Selvam, and **C Monterola**, "Simulating Congestion Dynamics of Train Rapid Transit using Smart Card Data," Procedia Computer Science 29, pp. 1610–1620 (2014).

*How effective are **congestion reducing strategies**? How many participants do we want?*




WiMLDS
Women in Machine Learning & Data Science

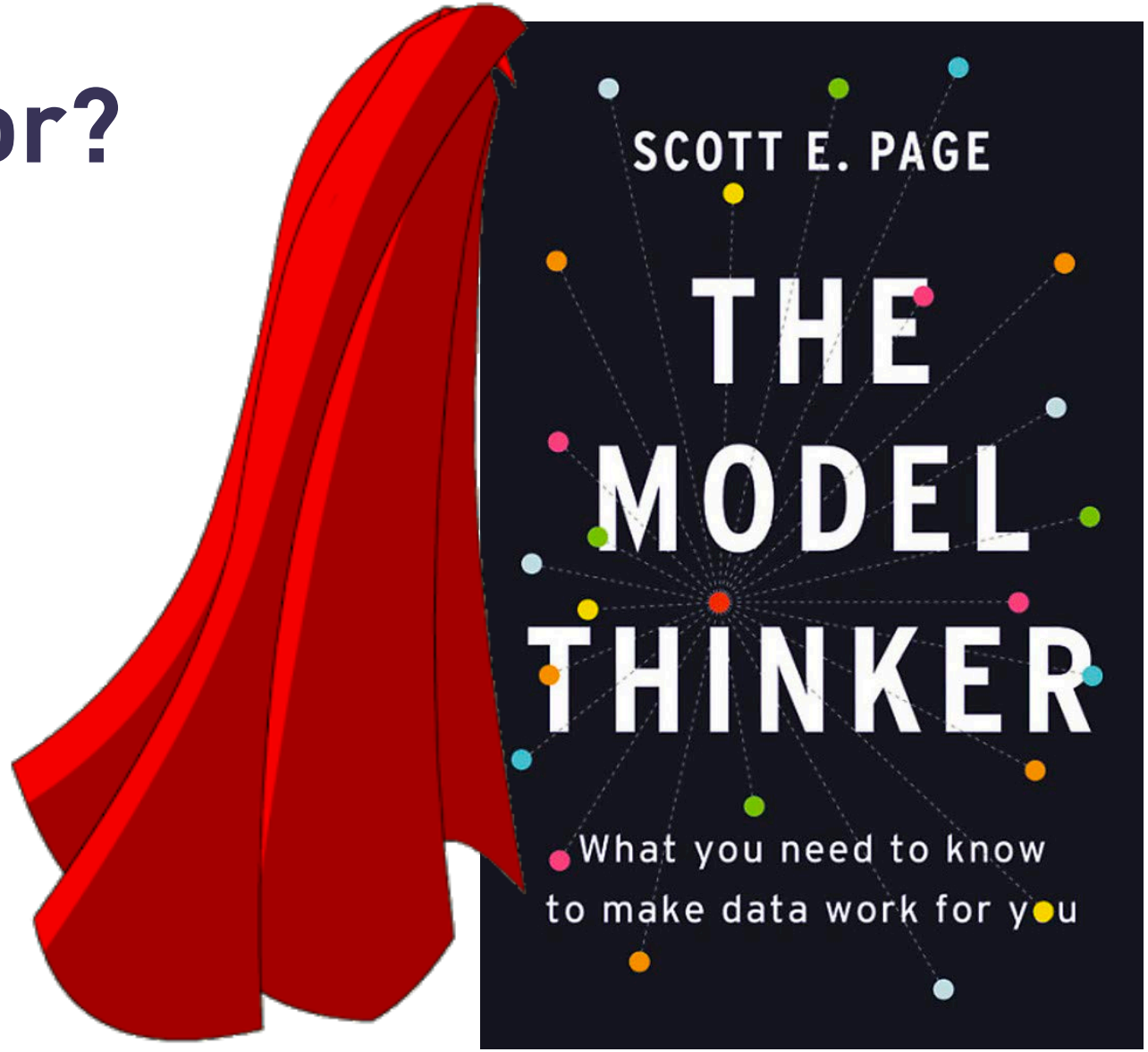
# What is a model?

# What is it not?

# What are models for?

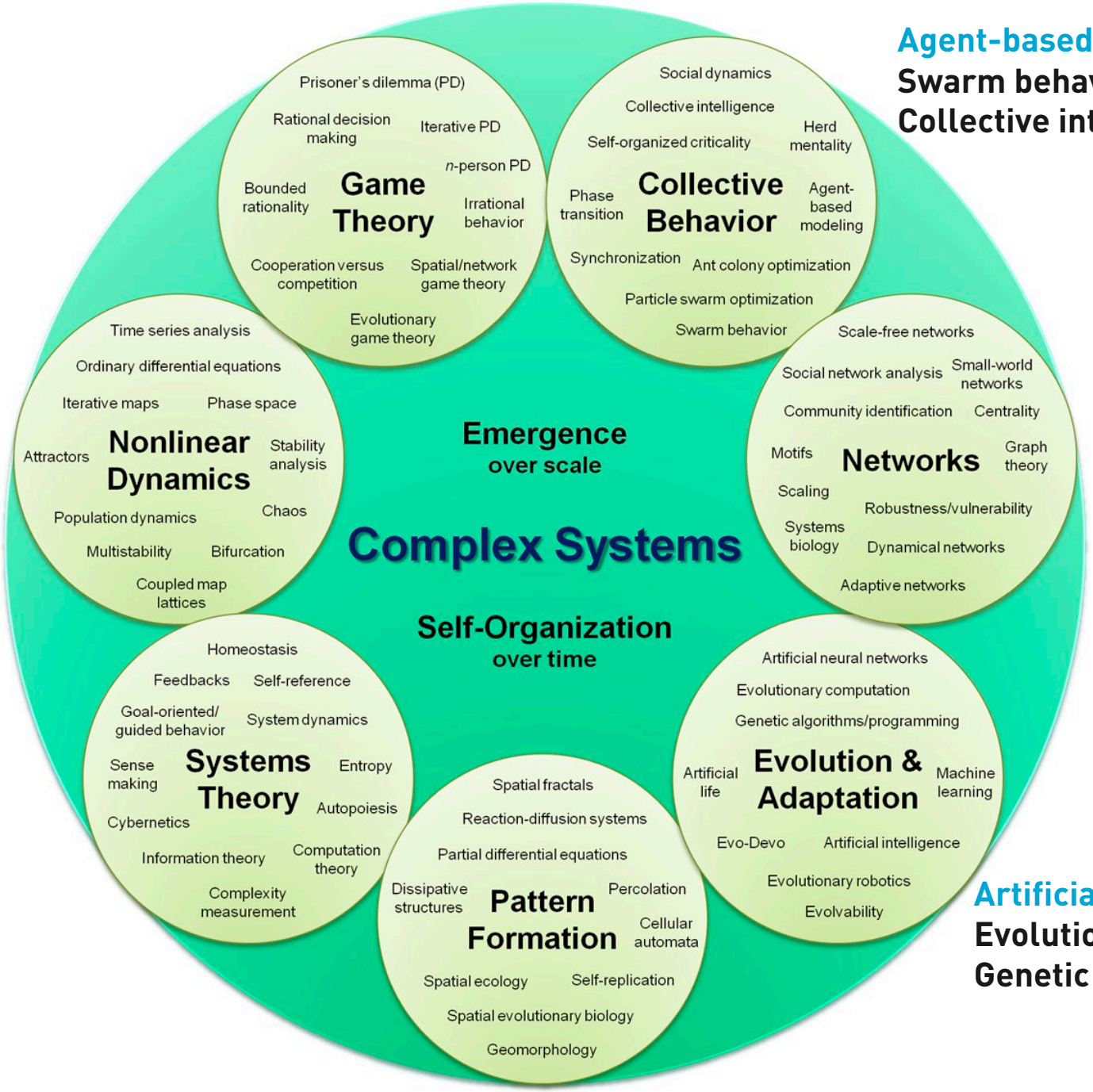
# What are models for?

**R** eason
**E** xplain
**D** esign
**C** ommunicate
**A** ct
**P** redict
**E** xplore



**BUILD MODEL**

Source: Hiroki Sayama




**Agent-based modeling**
**Swarm behavior**
**Collective intelligence**

**Social network analysis**
**Robustness/vulnerability**
**Information spreading**
**Dynamics networks**

**Artificial Intelligence / ML**
**Evolutionary robotics**
**Genetic algorithms**

WiMLDS
Women in Machine Learning & Data Science

## Modeling and Simulation

# Agent-Based Modeling & Applications

Erika Fille T. Legara, Ph.D.

eflegara

26 March 2020, 5:00pm

WiMLDS

Women in Machine Learning & Data Science