ICICLE
DEMOCRATIZING AI

Community-based Systems Modeling and Conversational Artificial Intelligence

Ayaz Hyder, PhD

Assistant Professor in the College of Public Health

Core Faculty in Translational Data Analytics Institute

Joint work with Harsh Panday, Amad Hussain, Erika Goetz, Huan Sun, and Eric Fosler-Lussier

THE OHIO STATE UNIVERSITY
COLLEGE OF PUBLIC HEALTH

TRANSLATIONAL DATA ANALYTICS
INSTITUTE

A conversation…

Community member: “I overheard the employees at the Colony Food Mart on College Corner Pike say they are closing the store by the end of the year.”

Public Health Commissioner: “Oh no, that’s terrible news. Let me see if I can talk them out of closing. Too many people rely on that store. I am not sure how they will be affected by the store closing.”

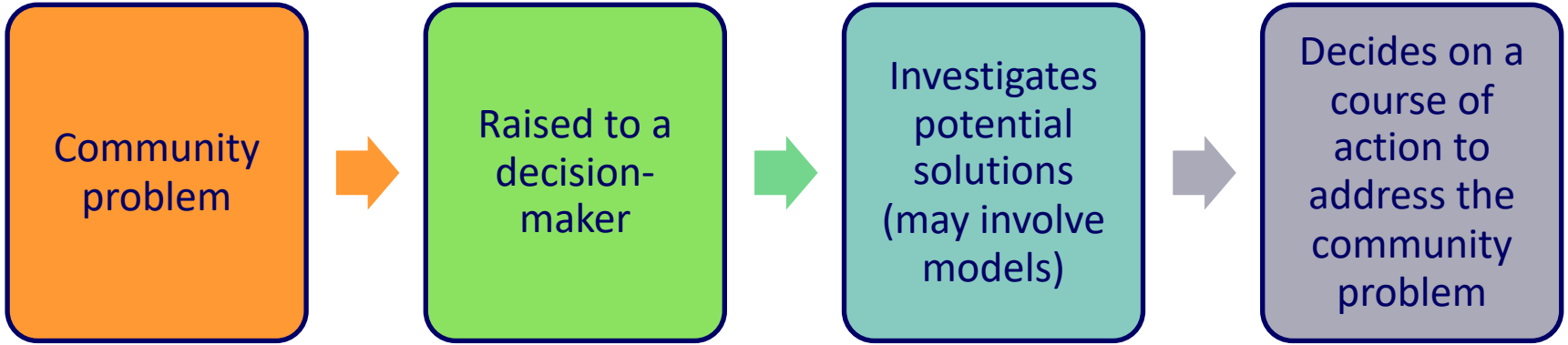
Another conversation…

Public Health Commissioner (goes to the ICICLE webpage and types into a text box): “What will happen to the level of food insecurity if the Colony Food Mart on College Corner Pike closes down? ”

ICICLE Conversational AI agent: “Thank you for your query. I’ll contact you soon with some answers.”

Then what happens…

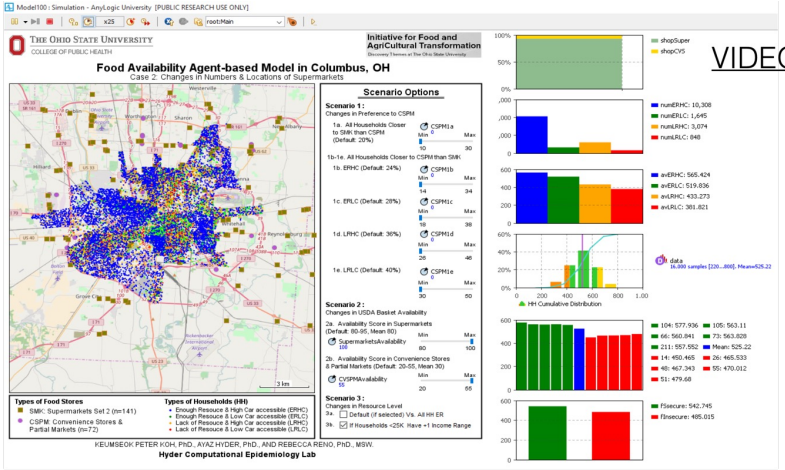
Well… let's take a step back and put into context what is being asked of the ICICLE Conversational AI agent.



"If all models are wrong but some are useful", then **communities** are critical for identifying which models (and data analytics tools) are useful.

Community-Based System Dynamics/Modeling

- CBSD is a <u>participatory modeling approach.</u>
- CBSD is increasingly being used to develop, test, implement, and evaluate <u>simulation</u> models in partnership with community stakeholders from the **<u>BEGINNING</u>** of the modeling process.
- CBSD has its limits when it comes to <u>model literacy</u> and <u>translation to practice</u>.

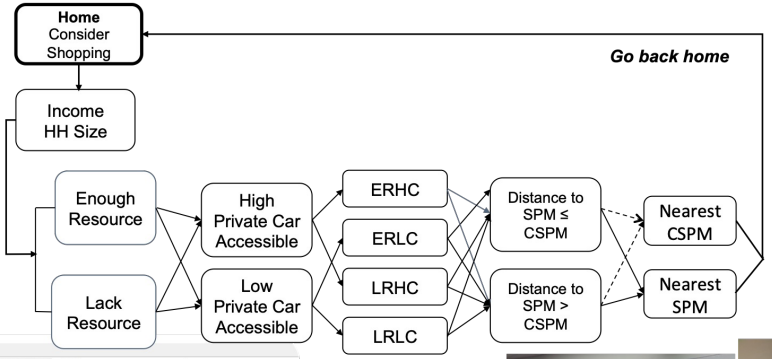


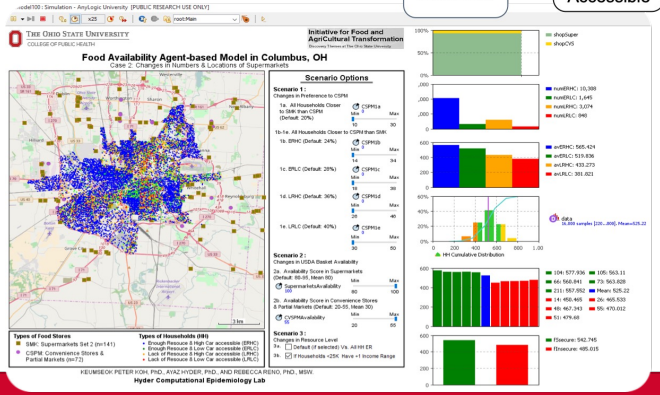
VIDEO

A very quick **<u>overview</u>** of how the CBSD approach was applied to develop a simulation model

Transforming the Food Environment for Better Health: A Systems Approach

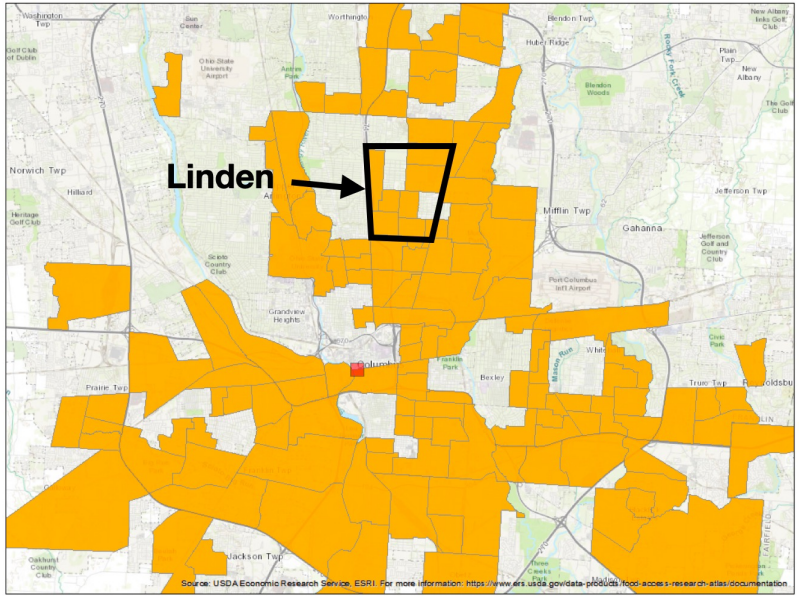
Joint work with Keumseok Peter Koh, PhD (postdoc)





Food insecurity

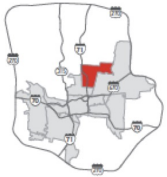
<u>Definition:</u> "is a household-level economic and social condition of limited or uncertain access to adequate food."—USDA definition



16% (range: 8.7%-20.4%) of Ohio households are food insecure.

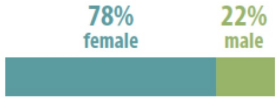
Addressing food insecurity means addressing access, availability and affordability of healthy food.

FOOD ACCESS, FOOD SECURITY, AND HEALTH



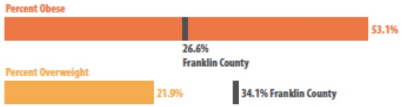
ZIP CODE 43211

33 People Surveyed



Percent of People Categorized Overweight and Obese

Figures based on self-reported height and weight measurements



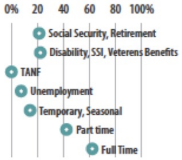
Household Employment

62.1%
Had someone in the household working a full-time job (more than 35 hours a week)

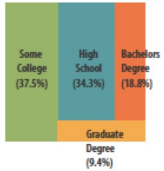
42.3%
Had someone in the household working a part-time job.

Household Income

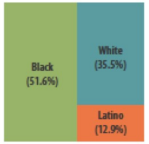
Household Income Sources



Participant's Education



Participant's Race/Ethnicity



Food Security

We used the six-item USDA food security module to assess food insecurity. Indications of food insecurity, such as having to skip meals, limit dietary intake, or adjust diets due to limited resources are categorized as Low Food Secure and Very Low Food Secure (formerly Food Insecure with Hunger).

*Feeding America estimates that Franklin County has a food insecurity rate of 17.9% (this number combines categories Low Food Secure and Extremely Low Food Secure)

Choosing Between Food and Other Expenditures

27.3%
Had to choose between buying food and paying for medicine or medical care

36.4%
Had to choose between buying food or paying for rent/mortage

Use of Community Food Resources

48.4%
Obtained food from a food pantry at least once in the last year

12.1%
Participated in a free community meal

45.5%
Visited a farmer's market or produce stand at least once in the last year

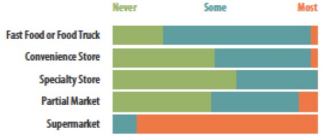
Barriers to Food Choice

79%
Said they were able to buy the food they wanted

18.2%
Said that food prices were a barrier

3%
Said that distance to a store was a barrier

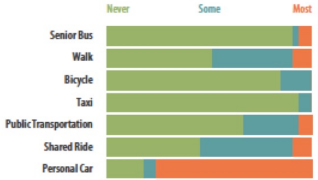
Where Food is Purchased



Food Purchasing Priorities

Factors that were important to participants when making food purchases. Categorized as Not Important, Slightly Important, Important, and Very Important.

Transportation to Food Stores



Average One-Way Distance Traveled to Purchase Food

4.25 Miles

Average One-Way Time Spent in Transit to Purchase Food

18.5 Minutes

Transforming the Food Environment for Better Health: A Systems Approach

Study Objectives

1. Develop an agent-based model that integrates multiple components of the local food system (e.g., food access, food purchase patterns, neighborhood environment, individual-level factors, and health status)

2. Perform scenario analysis (“What if?”) for alternative strategies to increase food access, improve healthy eating, and lower the risk of food insecurity.

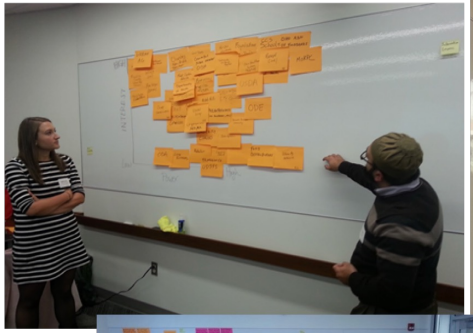
Transforming the Food Environment for Better Health: A Systems Approach

Levels of community engagement

Group model building: a participatory modeling approach that:

1. Brings together stakeholders that represent multiple facets of a complex problem.
2. Goes through a series of scripted exercises to build trust between participants, breakdown silos and take ownership of the problem.
3. Participants develop a computational model for the problem at hand and validate the model through cycles of feedback with modeling team.

Final product: A computational model that is owned by the participants for use in their decision-making and planning process in order to solve the complex problem.



CE: Community engagement perspective

CS/M: Computer scientist/Simulation Modeler's perspective

CE and CS/M: Diverse perspectives that lead to convergence by the end of the group model building workshops.

Profiles of GMB participants and attendance at GMB sessions 1 and 2

Role	Sector	Participant profile	Attendance
Participants	Public school	Senior-level expert	Both sessions
	Public health department	Senior-level expert	Both sessions
	Public health department	Entry-level professional	Session 1
	NGO	Nationwide NGO, senior-level expert	Both sessions
	NGO	Local NGO executive	Session 1
	NGO	Local NGO researcher	Session 2
	Academic	Urban planning professor	Session 1
	Academic	Social work professor	Both sessions
	Academic	Food science professor	Session 1
	Academic	University research center researcher	Both sessions
Facilitator/Modeler	Academic	Professor in public health/modeler, facilitator, and helper	Both sessions
	Academic	Researcher in public health/modeler, facilitator, and helper	Both sessions
	Academic	Researcher in public health/facilitator and helper	Both sessions
	Academic	Research administrative coordinator/helper	Session 1

The public agenda for the GMB workshops

1st GMB session	2nd GMB session
Check-in	Check-in
Introduction of participants	Welcome & Introduction
Introduction of project	Exercise 1: Review of 1st GMB
Exercise 1: Hopes & fears	Exercise 2: What we have done
Exercise 2: Identifying main barriers to food security	Exercise 3: Small group review/model exercise
Break	Break
Exercise 3: Mapping interactions	Exercise 4: Big group review/model exercise
Exercise 4: Key stakeholders	Exercise 5: Intervention Levers
Exercise 5: Intervention levers	Exercise 6: Wrap-up
Closing remarks & Adjourn	Closing remarks & Adjourn

Groups of main barriers identified by the GMB participants

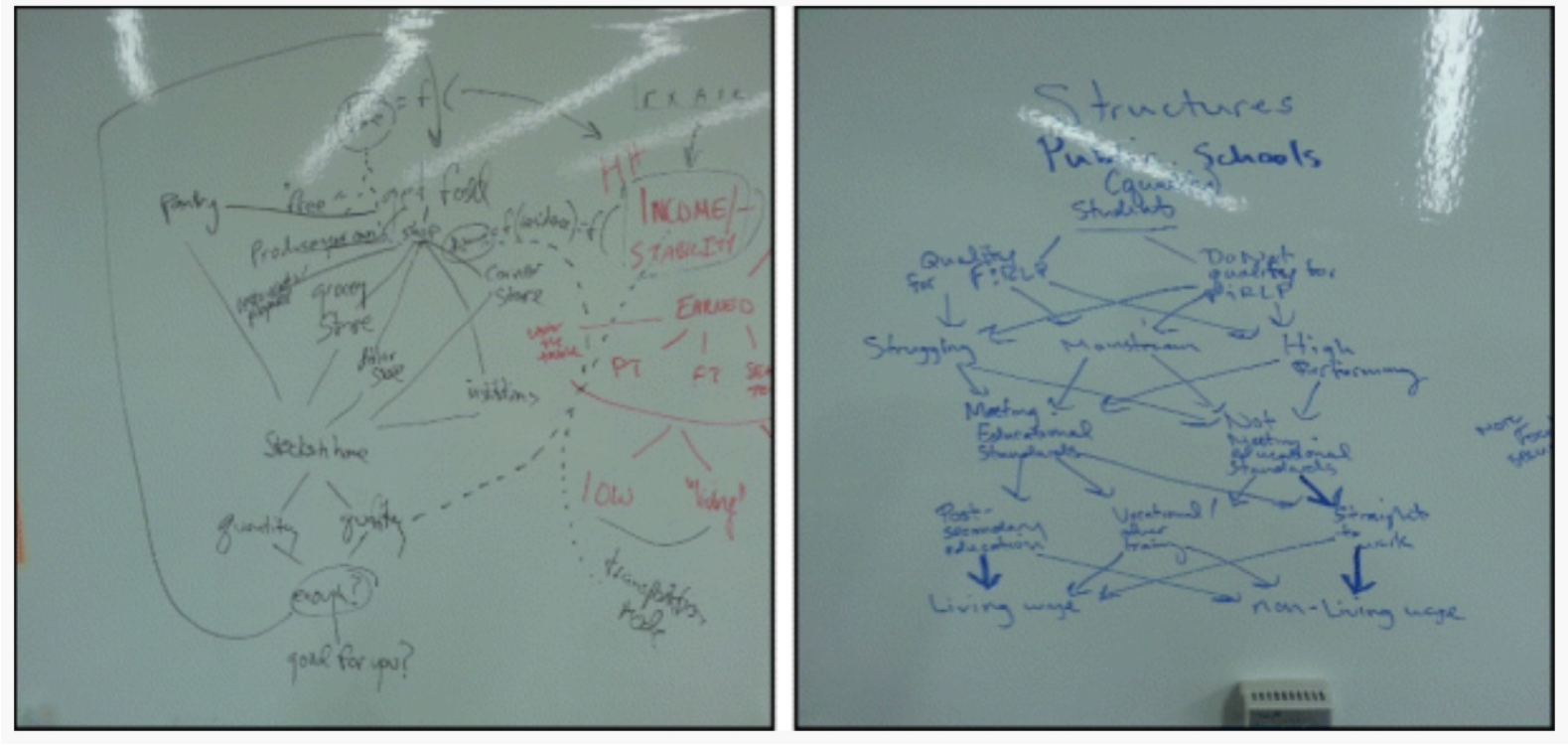
Rank by vote	Main barriers	Number of votes
1	Household income/income stability	7
2	Structural/institutional inequalities	6
2	Policy, advocacy & access in the food system	6
2	Inequities in food policies	6
5	System integration & sustainability	5
5	Barriers to physical access to foods	5
7	Individual behaviors, preferences, & knowledge	3
8	Shaming & blaming	2
9	Physical/mental health factors	0
9	Lack of support (formal/informal)	0
9	Diverse geographic makeup of the county	0

CE: Multiple policy options from multiple perspectives.

CS/M: Pain points of the model users

CE: A hand-drawn sketch of how household shopping behaviors and decision-making process.

CS: Algorithms for how agents make decisions in the simulation model.



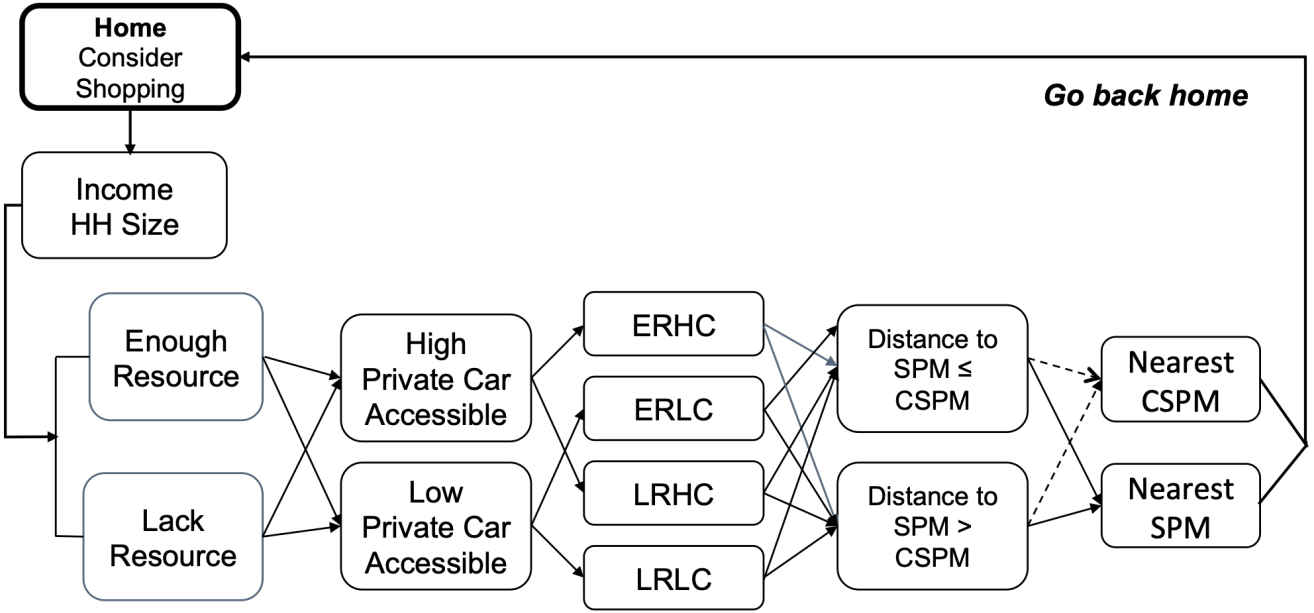
CE: Key stakeholders for addressing food insecurity.

CS: Who may access the simulation model and their level of interest is using the model?

Interest \ Power	Low		High
High	▪ Me ▪ Local food council	▪ Churches ▪ Food system activists ▪ State & City public health departments ▪ Foodbanks ▪ College extension	▪ Non-profit foundations ▪ Public schools
	▪ Family doctors ▪ Healthcare professionals ▪ County commissioners ▪ Local governments ▪ Independent grocers	▪ Suburban schools ▪ Philanthropists ▪ City council	▪ Federal government ▪ Federal health departments and agencies
Low		▪ Grocery chains ▪ Retailors ▪ Federal Congress ▪ Politicians	▪ Food distributors

Note: Specific organization names intentionally concealed by the authors for sensitivity.

Transforming the Food Environment for Better Health: A Systems Approach



Mental model —> Computer model

Transforming the Food Environment for Better Health: A Systems Approach

Sensitivity Analysis and Face Validation

Types		Categories	Behaviors/ Attributes				
			Monthly Shopping Frequencies	Movement Speed	Ability To Carry Items	SPM Preferences	
						If SPM* is farther than CSPM**	If SPM is closer than CSPM
	Household	Group 1: ERHC (Enough Resources, High Car Accessibility)	7	36km/h	100%	76% vs. 24%	80% vs. 20%
		Group 2: ERLC (Enough Resources, Low Car Accessibility)	8	3.6km/h	80%	72% vs. 28%	
		Group 3: LRHC (Low Resources, High Car Accessibility)	6	36km/h	100%	64% vs. 36%	
		Group 4: LRLC (Low Resources, Low Car Accessibility)	6	3.6km/h	80%	60% vs. 40%	
	Stores	Categories	Attributes				
			FSA score*** (% USDA TFP Items Available)				
		SPM	80-95 (mean: 80)				
		CSPM	20-55 (mean: 30)				

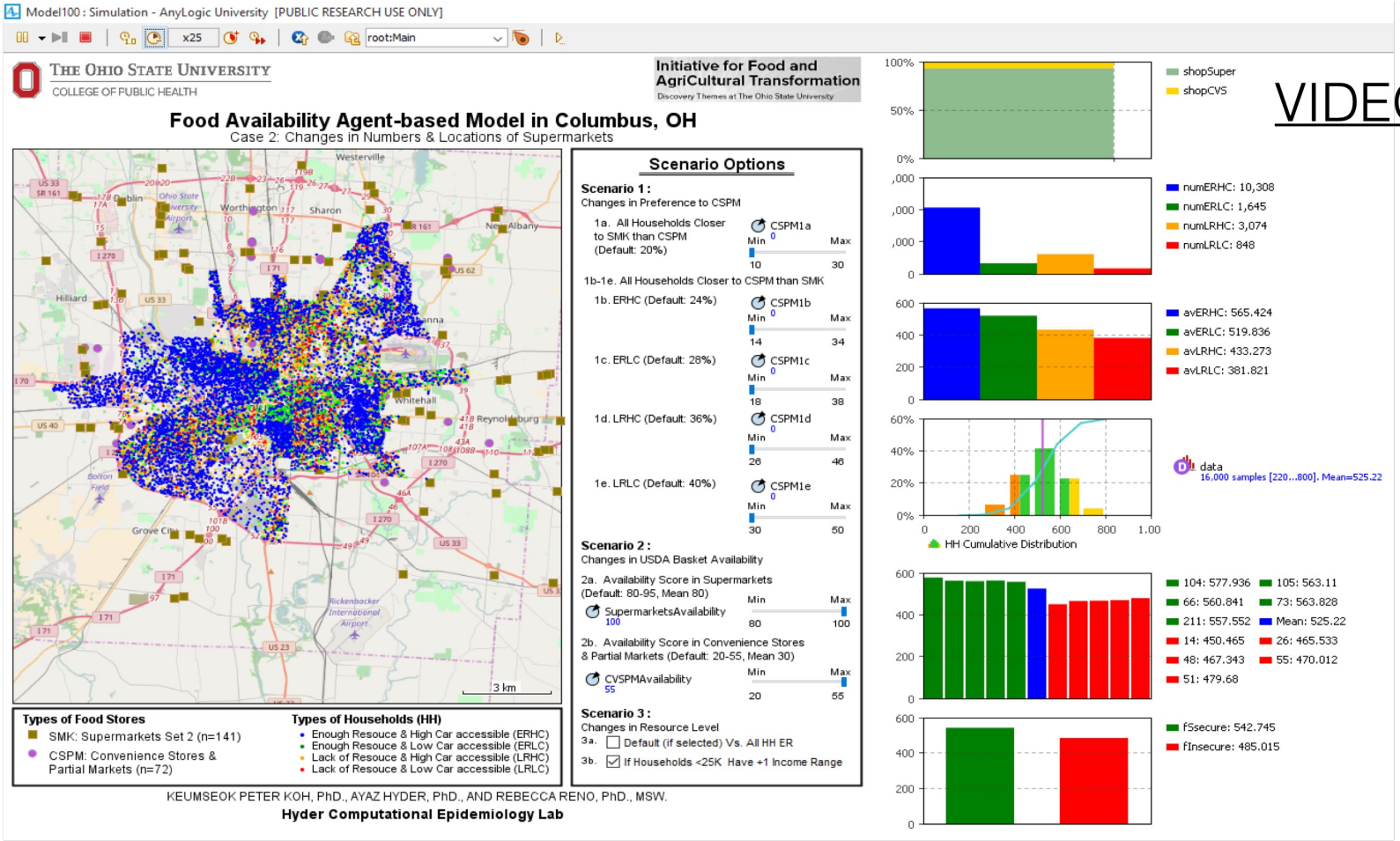
Notes. *SPM=Supermarket store; **CSPM = Convenience Stores & Partial Markets; ***FSA Score= Food Store Audits Score (% Availability of USDA Thrift Food Plan (TFP))

Transforming the Food Environment for Better Health: A Systems Approach

Sensitivity Analysis and Face Validation

Scenarios	Description
Base Model	Model with attributes/behavioral rules in Table 1
▪ Scenario 1	If low resources households visit less CSPMs and more SPMs (with the same SPM preferences as those of enough resources households (Group 1 & Group 2): • if Group 3 (LRHC) households' CSPM preference decrease from 36 to 24 (same as Group 1); and • if Group 4 (LRLC) households' CSPM preference decrease from 40 to 28 (same as Group 2)
▪ Scenario 2	If SPMs have more USDA TFP items available: all SPMs have a FSA score of 100
▪ Scenario 3	If CSPMs have more items available: all CSPMs have a FSA Score of 55 (maximum in the FSA)
▪ Scenario 4	If households less than household income of $25,000 have 1 increase in income level (i.e. this includes all three cases below): • in case households under household income of $10,000 earn $10,000-$14,999 • in case households with household income of $10,000-$14,999 earn $15,000-$24,999; and • in case households with household income of $15,000-$24,999 earn $25,000-$34,999

Notes. *SPM=Supermarket store; **CSPM = Convenience Stores & Partial Markets; ***FSA Score= Food Store Audits Score (% Availability of USDA Thrift Food Plan (TFP))


Model100 : Simulation - AnyLogic University [PUBLIC RESEARCH USE ONLY]
x25
root:Main
THE OHIO STATE UNIVERSITY
COLLEGE OF PUBLIC HEALTH
Initiative for Food and AgriCultural Transformation
Discovery Themes at The Ohio State University
Food Availability Agent-based Model in Columbus, OH
Case 2: Changes in Numbers & Locations of Supermarkets
Westerville
Dublin
Worthington
Sharon
New Albany
Hilliard
Whitehall
Reynoldsburg
Grove City
Rickenbacker International Airport
3 km
Types of Food Stores
SMK: Supermarkets Set 2 (n=141)
CSPM: Convenience Stores & Partial Markets (n=72)
Types of Households (HH)
Enough Resouce & High Car accessible (ERHC)
Enough Resouce & Low Car accessible (ERLC)
Lack of Resouce & High Car accessible (LRHC)
Lack of Resouce & Low Car accessible (LRLC)
Scenario Options
Scenario 1 :
Changes in Preference to CSPM
1a. All Households Closer to SMK than CSPM (Default: 20%)
CSPM1a
Min 10
Max 30
1b-1e. All Households Closer to CSPM than SMK
1b. ERHC (Default: 24%)
CSPM1b
Min 14
Max 34
1c. ERLC (Default: 28%)
CSPM1c
Min 18
Max 38
1d. LRHC (Default: 36%)
CSPM1d
Min 26
Max 46
1e. LRLC (Default: 40%)
CSPM1e
Min 30
Max 50
Scenario 2 :
Changes in USDA Basket Availability
2a. Availability Score in Supermarkets (Default: 80-95, Mean 80)
SupermarketsAvailability
Min 80
Max 100
2b. Availability Score in Convenience Stores & Partial Markets (Default: 20-55, Mean 30)
CVSPMAvailability
Min 20
Max 55
Scenario 3 :
Changes in Resource Level
3a. Default (if selected) Vs. All HH ER
3b. If Households <25K Have +1 Income Range
KEUMSEOK PETER KOH, PhD., AYAZ HYDER, PhD., AND REBECCA RENO, PhD., MSW.
Hyder Computational Epidemiology Lab
shopSuper
shopCVS
numERHC: 10,308
numERLC: 1,645
numLRHC: 3,074
numLRLC: 848
avERHC: 565.424
avERLC: 519.836
avLRHC: 433.273
avLRLC: 381.821
data
16,000 samples [220...800]. Mean=525.22
HH Cumulative Distribution
104: 577.936
105: 563.11
66: 560.841
73: 563.828
211: 557.552
Mean: 525.22
14: 450.465
26: 465.533
48: 467.343
55: 470.012
51: 479.68
fSsecure: 542.745
fInsecure: 485.015

VIDEO

Clunky dashboard to access the model in AnyLogic

Additional details in the following publications

- Koh K, Reno R, Hyder A. Designing an Agent-Based Model Using Group Model Building: Application to Food Insecurity Patterns in a U.S. Midwestern Metropolitan City. J Urban Health. 2018 Apr;95(2):278-289. doi: https://doi.org/10.1007/s11524-018-0230-1
- Koh, K., Reno, R. & Hyder, A. Examining disparities in food accessibility among households in Columbus, Ohio: an agent-based model. *Food Sec.* **11**, 317–331 (2019). https://doi.org/10.1007/s12571-019-00900-7

Pain points (post CBSD)

- Models are not used or have limited impact because decision-makers are not able to access them in a practical and timely manner.
- **Scenario:**

A food retail company announces plans to **close a grocery store** in a Columbus, Ohio neighborhood with **very high % of food-insecure households**.

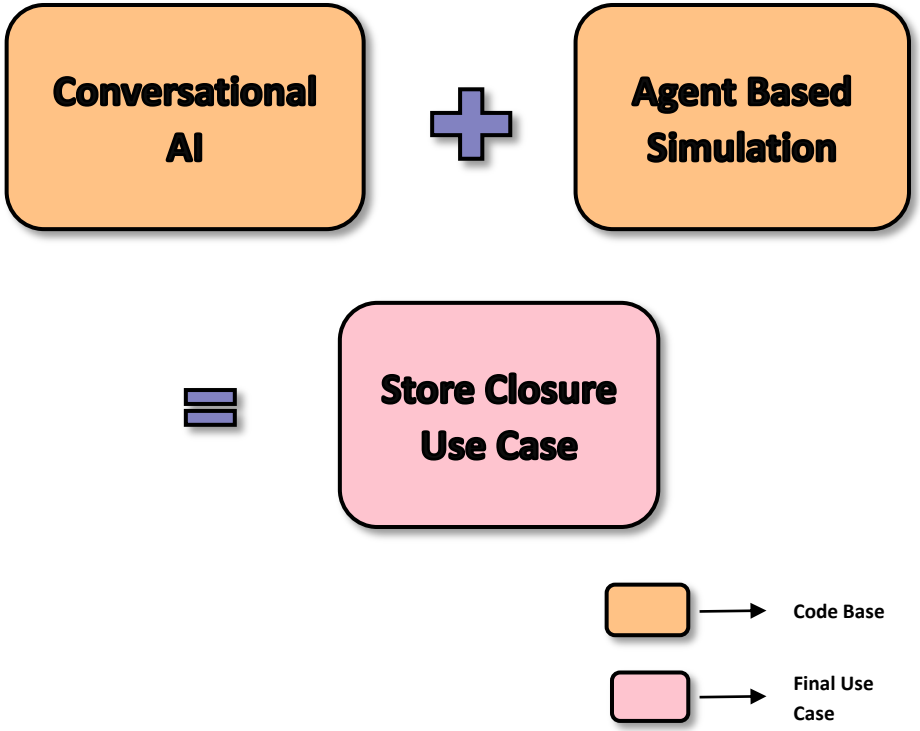
Now the health commissioner wants to know how the grocery store closure will affect community health so they can lobby the company to not close the grocery store or set up emergency food supply (e.g., pop-up food pantries) to reduce the impact on community health.

Our Solution (and use case via ICICLE)

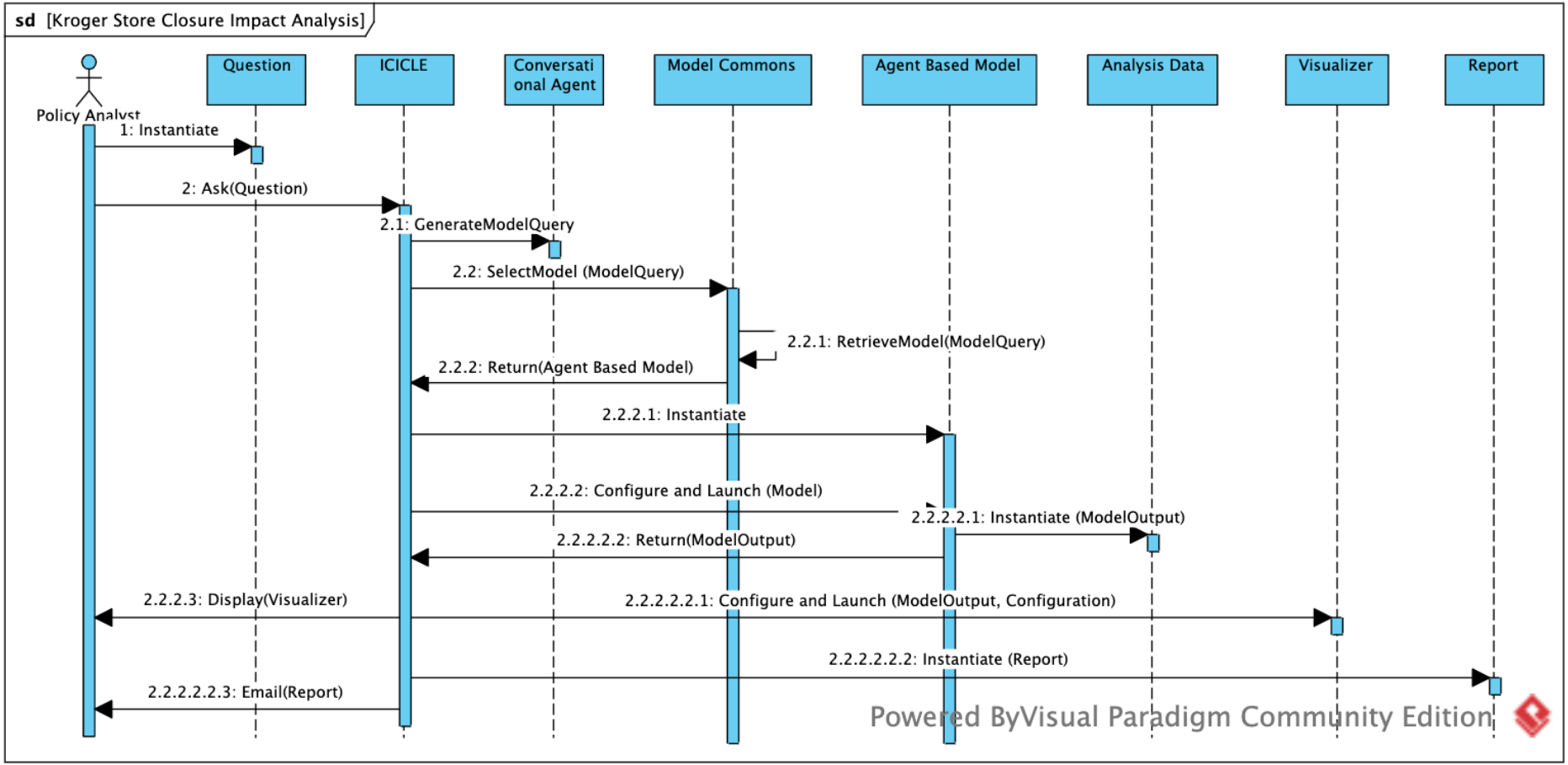
A conversational AI-enabled web interface that allows end users to run "What if?" scenarios based on an agent-based model for food insecurity.

Various Code base used for this use case

- *Conversational AI's code base*: This code base allows the agent-based model to understand the user's natural language query.
- *Agent Based Simulation* – This code base simulates the effect of complex interactions among households and environmental level factors on household level food availability.



Reference architecture diagram for this use case



Year 1 Goals and Accomplishments

Status	Type	Activities
In Progress	Milestone	Y1: Develop interfaces whereby a stakeholder submits a query to ICICLE;
In Progress	Milestone	Y1: Develop a set of workflows that identify a model from the model commons based on the query and run "What if?" analyses based on the selected model and query
In Progress	Milestone	Y1: Create visualization dashboards that summarizes the output from the "What if?" analysis.
Done	Milestone	Y1: Conversion of an existing computer simulation model for food insecurity, which was written in AnyLogic (a proprietary software) to an open-source programming language (Python).
In Progress	Milestone	Y1: Conduct performance tests on the simulation model and develop a knowledge graph for model parameters and outcomes.
Done	Milestone	Y1: Generate a conversational AI demo that can be used to query knowledge graphs for model parameters and outcomes, run "What if?" models using the agent-based model and conversational AI and provide reports (dashboards, documents, etc.) to be sent to stakeholders.

- **Existing agent-based model was developed using a participatory modeling approach with involvement of local food system stakeholders**
- **Allows for multiple scenarios (e.g., SNAP benefits)**
- **Scalable and extendable due to use of open-source MESA library**
- **Graduate student Harsh Panday (demo)**

- **Close partnership with Conversational AI group**
- **Extendable to other models in ICICLE**
- **Removes barriers to use of complex systems model**
- **Graduate student Amad Hussain (demo)**

- **Linkages with other ICICLE groups working on Model Commons, Conversational AI, CI, Reference Architecture, Visual Analytics**

Future plans

- **Integration** with existing regional food system model in Y2 or Y3
- **Leveraging ICICLE support** to add a conversational AI/food insecurity modeling project as part of a renewal application for a Clinical Translational Science Award at OSU (currently funded by NIH's National Center for Advancing Translational Science)
- **Community engagement plans** include integration of our conversational AI-enabled web interface within daily workflows of local health departments, regional food banks, and local food council (start in Ohio and expand to other Smart Foodshed partners in California and Wisconsin)

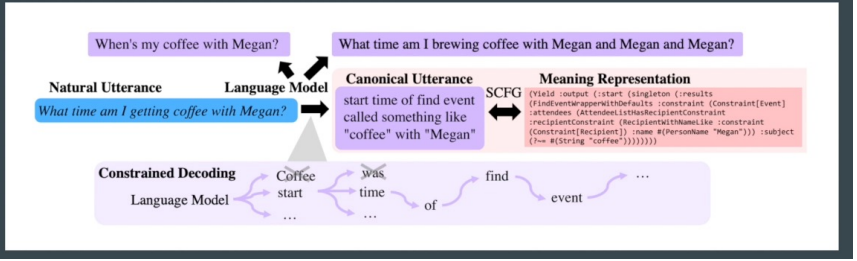
- **And now to the demos…**

Conversational AI demo

1. Overview of methods
2. Live demo of user input and system output

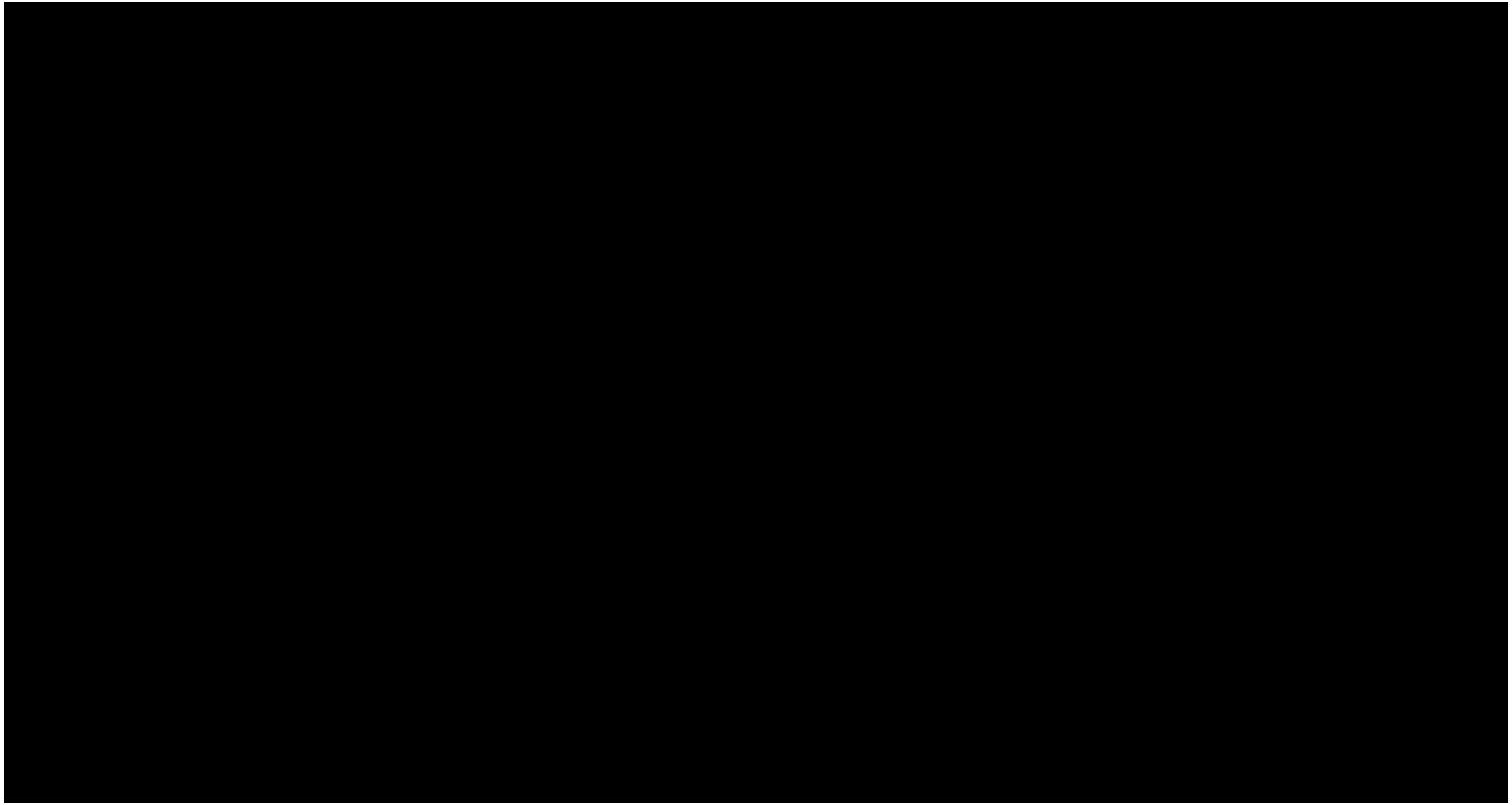
Foodshed Model Conversational System

- Focused on implementing few-shot semantic parsing
- Using Constrained Language Models Yield Few-Shot Semantic Parsers to guide initial efforts
  - Makes use of an intermediary step (canonical utterance) between natural language and meaning representation which allows for a few-shot pipeline
  - The canonical utterance is presided over by a rule-based grammar which allows it to be converted to the meaning representations
  - The intermediary step is generated through constrained decoding from the natural utterance step.
    - Constrained by what is valid according to the rule-based grammar
- Meaning representations will directly feed into python commands



Modeling demo

1. Live version of agent-base model
2. Example of report generated for end user



Result of the simulation

ICICLE – Food Policy Simulation Model Report

To: Joe Mazzola, Health Commissioner, Franklin County Public Health

From: ICICLE Food Policy Simulation Use Case Team (Contact: Ayaz Hyder at hyder.22@osu.edu)

Original Query → What will be the impact on food availability if a new grocery store opens on Moon Road?

Interpreted Query → Add a grocery store on Moon Road.

Key Insights:

Opening a new grocery store on ***Moon Road*** will increase the availability of food for all low resource households with or without a car. The impact on high resource households with a car is negligible and on high resource households without a car is a slightly lower availability of food.

Follow-up: Please contact our team (icicle@osu.edu) for additional insights.

Additional analyses for your consideration

Figure. Map of the neighorhood around Moon Road, Columbus, Ohio. Hash marks show simulated households with low resources that have access to a car (green) or do not have access to a car (yellow). Panel A shows map without new grocery store. Panel B shows map with new grocery store shown in black hash mark and circled for easier identificaiton.

The first page of the report will highlight the query inserted by the user and the changes made in the agent as the result of that query.

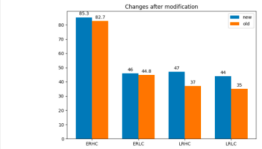
It will also give a summary of insights that the model collected after completing the simulation

The next page will show detailed insights with various graph comparison so that the user can understand the results better.

Description: This graph gives an overall comparison of the food availability before and after the modifications based on your query.

Pattern: The numbers over the bars are the total food available at the last step in each type of household.

Our recommendation: Overall positive impact of the new policy (opening a new grocery store on Moon Road, Columbus, Ohio) on food availability among vulnerable households.

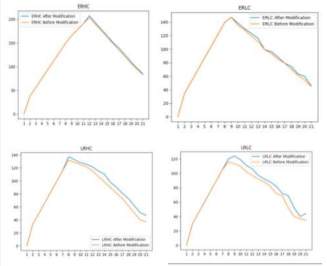


ERHC = Households with **high** level of resources and **have** access to a car.
ERLC = Households with **high** level of resources and **do NOT** have access to a car.
LRHC = Households with **low** level of resources and **have** access to a car.
LRLC = Households with **low** level of resources and **do NOT** have access to a car.

Description: These graphs show how long it will take for the positive impacts on food availability to take place as a results of the new policy on each type of households.

Pattern: In most household types, the positive impacts of the policy appear 6 weeks AFTER opening of the new store.

Our recommendation: Positive impact of the new policy on food availability among vulnerable households may not be apparent immediately but will eventually accumulate over several weeks.



Thank you for being here.

Questions?

Hyder.22@osu.edu