



Institute for Health
Metrics and Evaluation

Generative AI in Public Health: opportunities and risks in NLP, data analysis, and simulation

Abraham D. Flaxman

8/8/2023

Acknowledgement and Disclosures

Land acknowledgement: This research was conducted on the traditional lands of the Coast Salish people.

My grant and research support: US Census Bureau, Washington State Department of Health, Bill and Melinda Gates Foundation, The Alfred P. Sloan Foundation, NORC, NIH.

My other consulting: Janssen, Sanofi, SwissRe.

What is Generative Artificial Intelligence?

I'm not the right person to give a definition, but I think I have to give it a try.

Generative AI refers to the use of artificial intelligence techniques to create or generate new content, such as images, videos, text, music, or any other form of media.

Examples: Llama 2, ChatGPT, Stable Diffusion, StyleGAN

Capabilities of AI

“If a typical person can do a mental task with less than one second of thought, we can probably automate it using AI either now or in the near future”

---Andrew Ng, 2018

AI Opportunities in Public Health

AI can automate public health work that people:

- do not like doing
- cannot do fast enough
- cannot afford to do

Outline

- Natural Language Processing
- AI-Assisted Data Analysis
- Simulation with Generative AI Agents

Verbal Autopsy as a Case Study



- The Global Burden of Disease is a study that compares how much health is lost to each of over 200 different conditions, geographically and temporally
- To measure disease burden, we must know who is dying of what where.
- Death certificates exist for less than half of all deaths globally

Verbal Autopsy as a Case Study

Solution: ask



SECTION 5
 MODULE 1. GENERAL ILLNESS LEADING TO DEATH
 SPECIFIC QUESTIONS TO ELICIT SYMPTOMS AND SIGNS OF THE LAST ILLNESS

NO.	QUESTIONS AND FILTERS	CODING CATEGORIES	SKIP
501	Did _____ (NAME) have fever during her last illness?	YES.....1 NO2 DON'T KNOW.....8	502 502
501A	How many days/months before her death did the fever start and end?	START __ __ __ __ mons days END __ __ __ __ mons days DIED WITH FEVER.....9995 DON'T KNOW/UNSURE.....9998	
501B	How was the fever like?	HIGH.....1 MILD2 DON'T KNOW/UNSURE.....8	
501C	Was the fever continuous or on and off?	CONTINUOUS.....1 AFTER EVERY 1 - 2 DAYS.....2 AT NIGHT ONLY3 OTHER7	

Brief History of Verbal Autopsy

- Projects in Asia and Africa in the 1950s and 1960s used systematic interviews by physicians to assess causes of death
- Field workers at the Narangwal project in India labeled this technique “verbal autopsy” (VA)
- Subsequently spread and developed, particularly during the 1970s, when WHO suggested day reporting of health information by people with no medical background
- Remains the best available approach for assessing causes of death in communities in which most deaths occur at home

Outline

- Natural Language Processing
- AI-Assisted Data Analysis
- Simulation with Generative AI Agents

Natural Language Processing

- Free-text is somewhat common in global public health data
- Large language models make it easy to try to do something with it

Verbal Autopsy (VA)

- The challenge: find out what people are dying of
- The approach: ask someone who was present during the final illness
- The bottleneck: not many docs who have experience necessary to certify deaths based on verbal autopsy interview and they have better things to do with their time

Computer Certification of Verbal Autopsy

- Gates Foundation funded collection of a large database, suitable for training and validation: Murray et al, [Verbal Autopsy Data](#), 2011.
- Including open-text with personal information redacted: Flaxman, et al. [A de-identified database of 11,979 verbal autopsy open-ended responses](#), 2018.

Example of a verbal autopsy interview

Deceased was 53 Year Old

Male, reported to have:

- Asthma
- Puffiness of the Face All Over His
- More than Usual Protruding Belly

Underlying Cause (Gold Standard Level 1): COPD

• Difficulty Breathing (On-and-Off, Worse in Walking Position)

• Free Text: Asthma, Breath, Heart, Lung, Swell, Water

Computer Certification of Verbal Autopsy

****Input:****

the respondent explained that the deceased died on the 3 day after being injured himself with rod which caused his neck to become very stiff and died because of tetanus.

****Output:****



Cause of Death: Tetanus : Confidence 0.9

You are a virtual assistant helping researchers assign a cause of death based on given information. Your task is to analyze the provided context and determine the cause of death. You can only choose from the predefined list of causes:

[Cirrhosis, Epilepsy, Pneumonia, COPD, Acute Myocardial Infarction, Fires, Renal Failure, AIDS, Lung Cancer, Maternal, Drowning, Other Cardiovascular Diseases, Other Non-communicable Diseases, Falls, Road Traffic, Bite of Venomous Animal, Diabetes, Other Infectious Diseases, TB, Suicide, Other Injuries, Cervical Cancer, Stroke, Malaria, Asthma, Colorectal Cancer, Homicide, Diarrhea/Dysentery, Breast Cancer, Leukemia/Lymphomas, Poisonings, Prostate Cancer, Esophageal Cancer, Stomach Cancer, Measles, Other Defined Causes of Child Deaths, Violent Death, Other Digestive Diseases, Encephalitis, Sepsis, Other Cancers, Hemorrhagic fever, Meningitis, Birth asphyxia, Stillbirth, Preterm Delivery, Meningitis/Sepsis, Congenital malformation].

If there is a specific cause of death mentioned, return "Cause of Death: {Cause}: Confidence {X}" where X is a confidence score between 0 and 1. For example, "Cause of Death: Drowning : Confidence 0.9".

If there are symptoms mentioned without a clear cause of death, return "Symptoms noted: {Symptoms}". For example, "Symptoms noted: Chest Pain, Stomach Pain, Vomiting".

If there is no relevant information provided, return "No relevant info : Confidence {X}".

Given the provided context, determine the cause of death.

****Context:****

This was straightforward to prototype. But I bet we can do better!

See Kocoń et al, [ChatGPT: Jack of all trades, master of none](#)

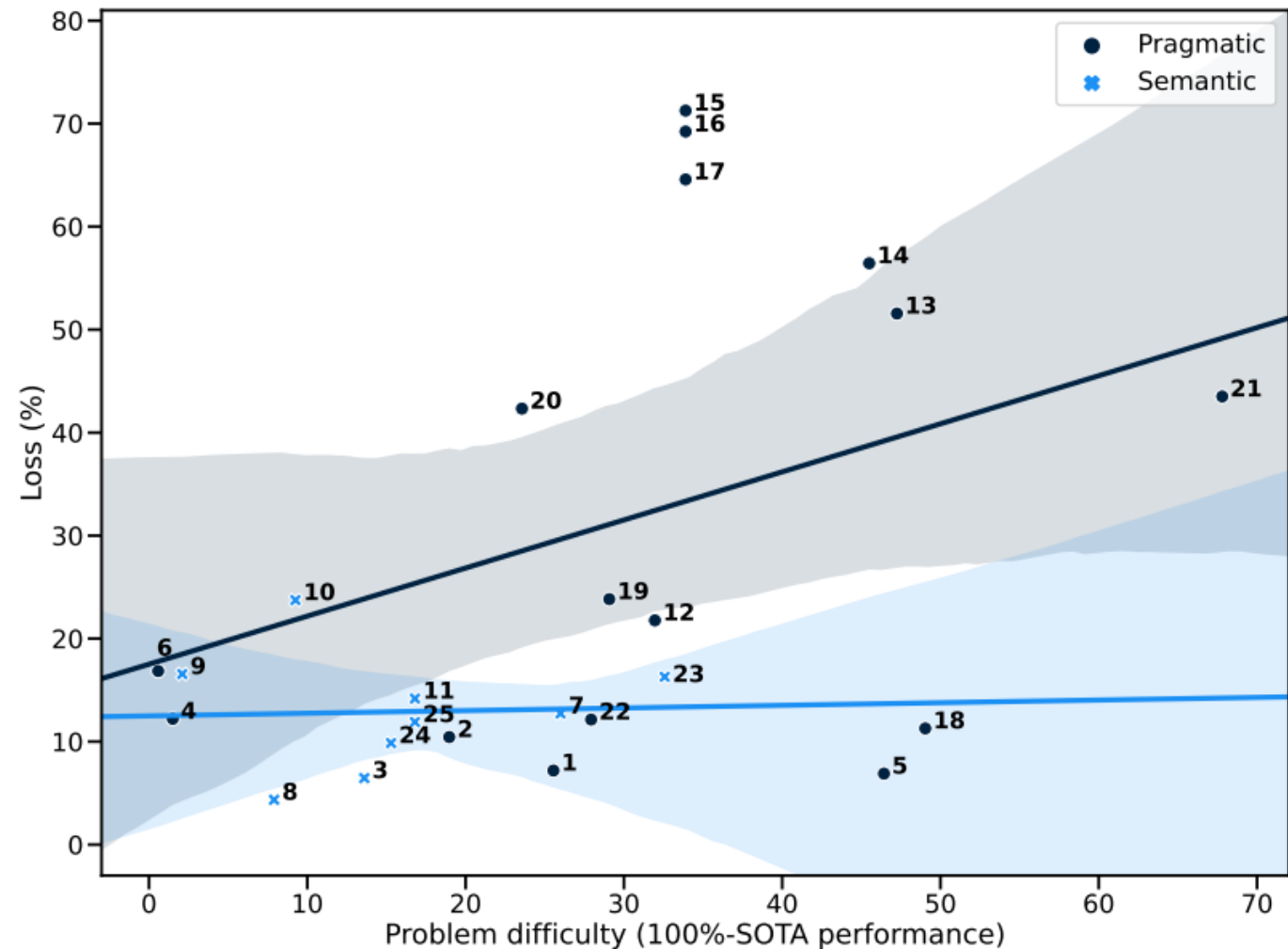


Figure 8: Correlations between the loss of ChatGPT performance compared to the SOTA method and difficulty of the task. Regression lines are drawn separately for pragmatic and semantic tasks. Each data point represents a single task with the index from Tab. 1.

Computer Certification of Verbal Autopsy

Lifting the curtain a little --- chatbots are not exactly what I need for this, but they seem close...are they?

$$p(w_{n+1} | w_1, w_2, \dots, w_n)$$

where w_i is a token, not a word. And causes of death are not always tokens...

With some additional affordances (benefits of openness in e.g. Llama 2 vs GPT4)

Additional GenAI opportunities in VA

- Adaptive interviewing, to get better data, and get it faster
- AI systems in collaboration with physicians --- better? opportunity to enhance training?
- Moral Distress of VA work, and emotional support for interviewee and interviewer

Other NLP challenges, and less natural languages

- Nurses Notes
- Triage for SMS-based advice during pregnancy
- Billing data and diagnosis codes

Outline

- Natural Language Processing
- AI-Assisted Data Analysis
- Simulation with Generative AI agents

AI-Assisted Data Analysis

Potential for AI-assisted data analysis to broaden participation in technical analyses central to evidence-based public health

Continuing VA example: what do you do with your VA data once you have collected it and mapped it to underlying cause of death?

AI-Assisted Data Analysis

Potential for AI-assisted data analysis to broaden participation in technical analyses central to evidence-based public health

Example 1: counting rows with missing data when considering using complete case analysis or multiple imputation

Example 2: exploring Llama 2 predictions with huggingface transformers library

AI-Assisted Data Analysis

How should we be writing libraries (and documentation) so that AI Assistants will be able to help users to use them?

This reminds me of an old fantasy novel

In “Earthsea” the study of wizardry requires learning the true names of things, in a language that is no longer spoken by most of the inhabitants of the Earthsea world.

Like what it might take to construct useful prompts for GenAI systems a generation from now...





provide flute music to accompany this poem

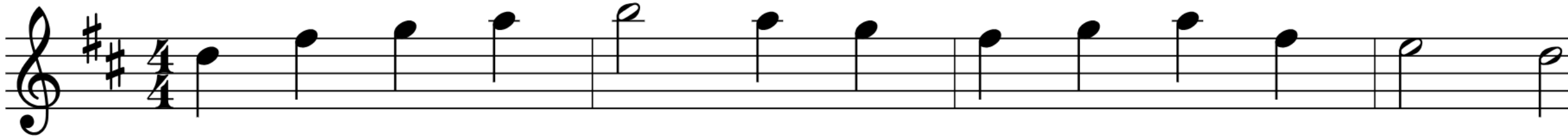


Creating flute music to match specific lyrics can be a complex process that involves understanding the rhythm, melody, and mood of the piece. While I can't create a precise musical score in text form without knowing more specific details about your musical preferences, I can provide a general guideline that might help a musician craft something to go along with the poem.



Can you compose a short tune (say four to eight bars) using ABC notation

Rubik's Cube Master



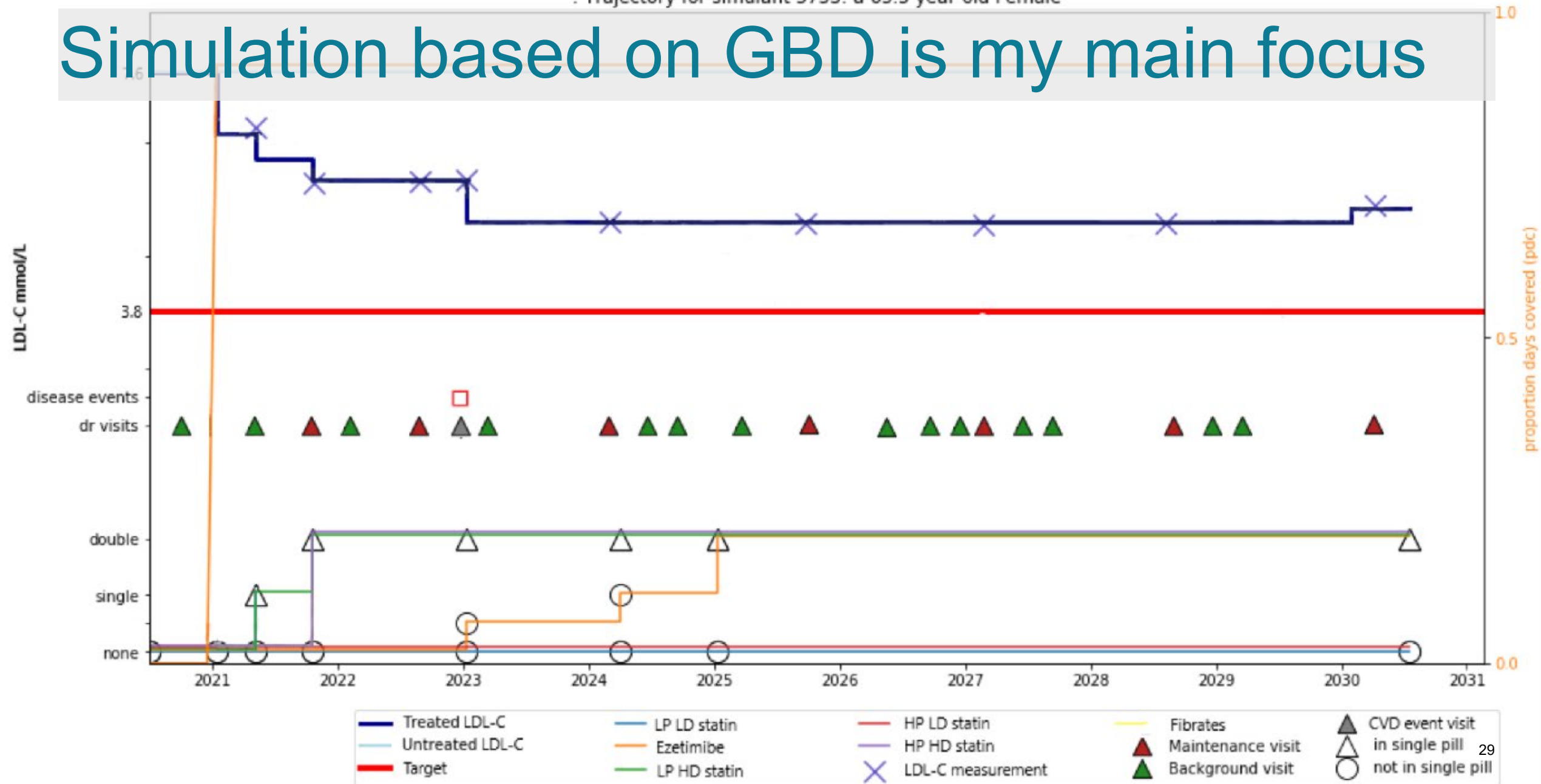
a cheerful, uplifting tune that could work as a motif for your poem. It's in 4/4 time and the key of C major.

Outline

- Natural Language Processing
- AI-Assisted Data Analysis
- Simulation with Generative AI agents

: Trajectory for simulant 5733: a 65.5 year-old Female

Simulation based on GBD is my main focus

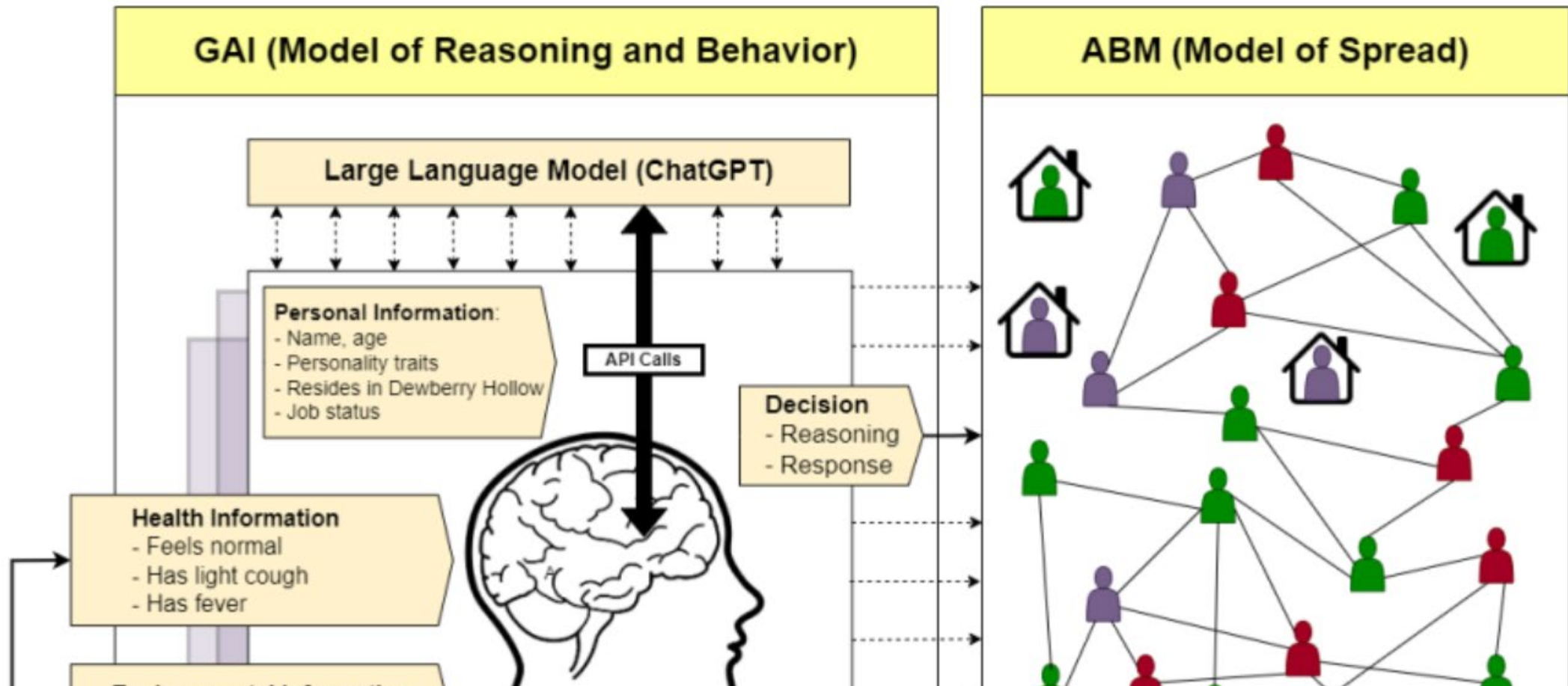


In work on COVID modeling, I have become convinced that modeling both “mass behavior” and “elite behavior” are an important area for future innovation



Simulation with Generative AI Agents

From Williams et al, [Epidemic Modeling with Generative Agents](#)



Simulation with Generative AI Agents

Verbal Autopsy example continues --- agents can experience symptoms based on disease, talk about them with others, remember learning about experiences of confidantes, etc

Results of a simulation like this could include some quantification of how much evidence a VA surveillance system is likely to provide. Could even include “elite decision makers” in sim to respond to VA-based evidence (?)

Simulation with Generative AI Agents

What to teach epi student this fall in Introduction to Epidemic Modeling??

Which also goes back to my interest in AI-Assisted programming...

AI Opportunities *and Threats*

- Natural Language Processing
- AI-Assisted Data Analysis
- Simulation with Generative AI Agents

Bake in biases: e.g. LLM approach to VA might rely on inaccurate stereotypes of disease presentation

Lack of understanding: e.g. analyst being assisted does not know what they have done; or AI is too opaque and it is not clear to where simulation results are generalizable.

Thank you!

Abraham D. Flaxman

abie@uw.edu

