

# Grounding large-language models with real-world public health knowledge using Chain of Cause Reasoning

Geoffrey H. Siwo

Research Assistant Professor

Department of Learning Health Sciences

Center for Global Health Equity

University of Michigan Medical School

# Disclosures

- Co-Founder, Equity holder: Helix Nanotechnologies & Anza Biotechnologies
- Stock options: Immusoft Corporation
- Advisory Board Member & Equity holder: Saguaro Capital Management
- Consulting fees: Google Research

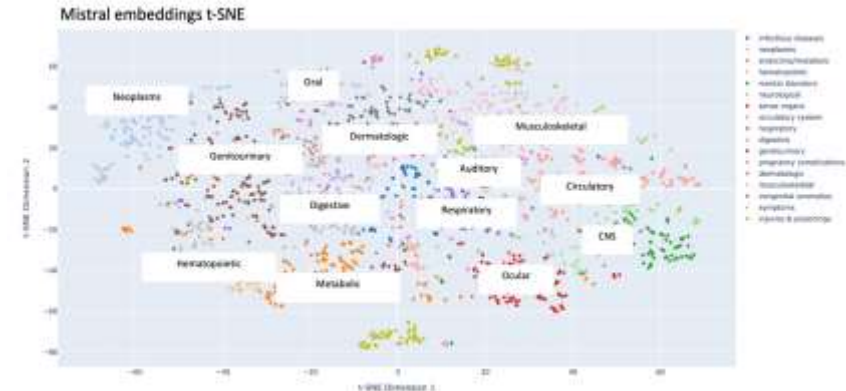
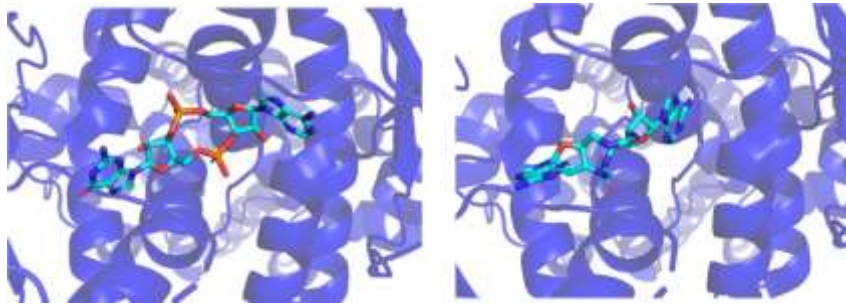
# LLM Research Focus: LLMs as tools for knowledge representation, generation and reasoning

## LLMs as tools for encoding and decoding data, information & knowledge

DNA, RNA & Protein Sequences

Small Molecules/ Drugs

Phenotypes, Diseases & Genes



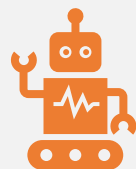
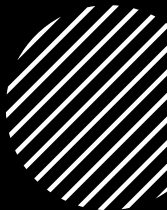
## LLMs as tools for reasoning over knowledge

Causal Reasoning with LLMs

Chain of Cause Reasoning



# Opportunities for LLMs in Global Health



LLM-based chatbots could provide general medical advice in future



LLMs could assist in disease diagnosis and treatment management



LLMs could help in educating the public about health

# Challenges

Healthcare organizations have no full visibility on how popular LLMs are developed including the algorithms, datasets for training, testing & validation

LLMs can be costly to operate and have a big impact on energy

No standardized approach to benchmarking LLMs

LLMs are largely trained on data from the internet, biased

LLMs can generate answers that sound factual yet confabulated – “Hallucination” and can be “tricked” to reveal sensitive information

# Fundamental constraints of LLMs

- Fundamental constraints: cannot be solved by more data
- Not factual, all responses are probabilistic based on 'next' word prediction
- Detached from realities of the natural world- Not everything that is plausible in theory makes practical sense

# Limitations of current evaluation

---



Top applications of LLMs in healthcare today perform shallow assessments e.g. how well LLMs answer questions (84% of studies)



No consideration of real patient care data



Good grades on answering a standard set of questions does not show LLMs use any reasoning to give correct answers, and so cannot be relied upon

**LLMs should be enabled to use  
reasoning before providing answers**



# A new approach: Chain of Cause Reasoning



Inspired by Daniel Kahneman's  
Thinking Fast, Thinking Slow

-Thinking Fast- Instinct

-Thinking Slow- Intentional, 'conscious'  
thought, planning



Current LLM systems are System 1 thinkers (Thinking Fast); We need  
System 2 LLMs for critical tasks



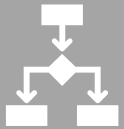
LLMs need to incorporate 'common sense', baseline groundtruth

# Comparison to Chain of Thought (CoT)

---



CoT- LLM provides step by step explanation of how it came to an answer



LLMs do not always use the step-by-step explanations they give- they generate explanations by predicting next series of words



LLMs CoT is frequently inconsistent with answers- Answers can be right but CoT wrong, and vice versa

# Chain of Cause Reasoning



LLM is first provided with an external causal graph, knowledge graph or asked to construct this between entities of interest

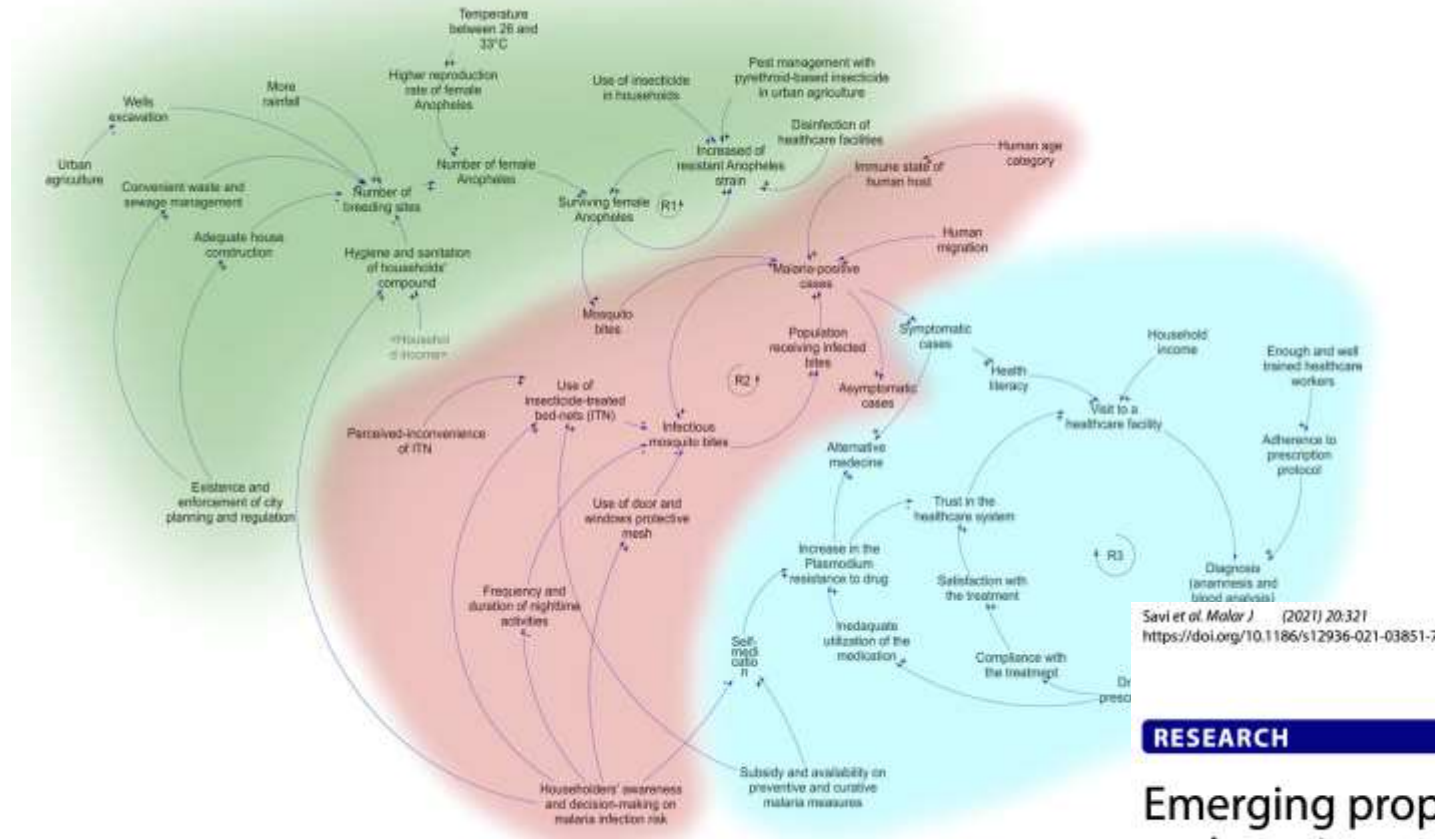


LLM is then given a question and asked to use the causal graph to come up with an answer



Approach 'forces' LLM to use causal reasoning and discourages LLM from cheating

# Expert curated causal graphs



Malaria Journal

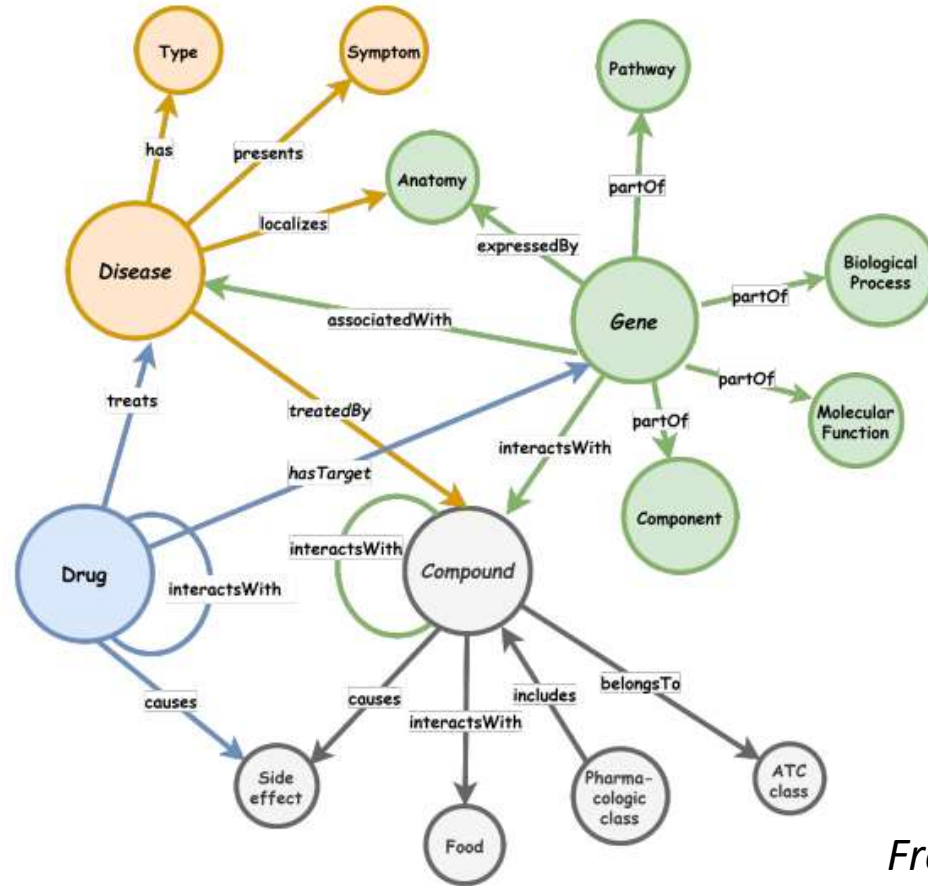
RESEARCH

Open Access

## Emerging properties of malaria transmission and persistence in urban Accra, Ghana: evidence from a participatory system approach

Merveille Koissi Savi<sup>1\*</sup>, Daniel Callo-Concha<sup>1,2</sup>, Henri E. Z. Tonnang<sup>3</sup> and Christian Borgemeister<sup>1</sup>

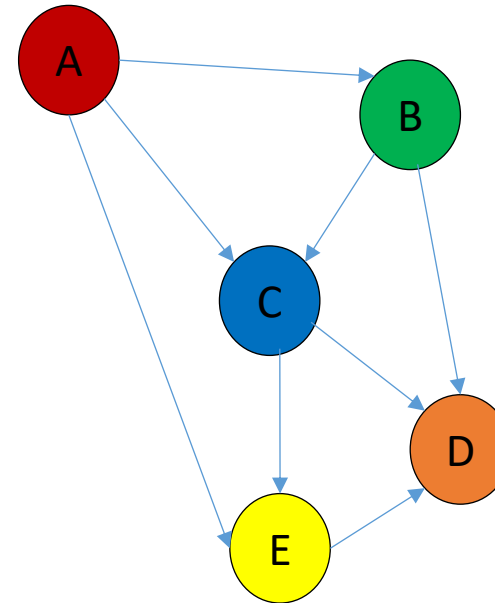
# Expert curated knowledge graphs



*From WiseCube*

# Causal Discovery from observational data using AI

Observational data  
e.g. Health &  
Demographic  
Surveys (HDSS)



Causal graph

# Causal graph generation using LLMs

- LLM is prompted to generate causal graph
- Causal graph can then be manually interrogated to determine whether inferred relationships are causal
- Augmentation of causal graphs by combining expert curated causal graphs with LLM-derived graphs

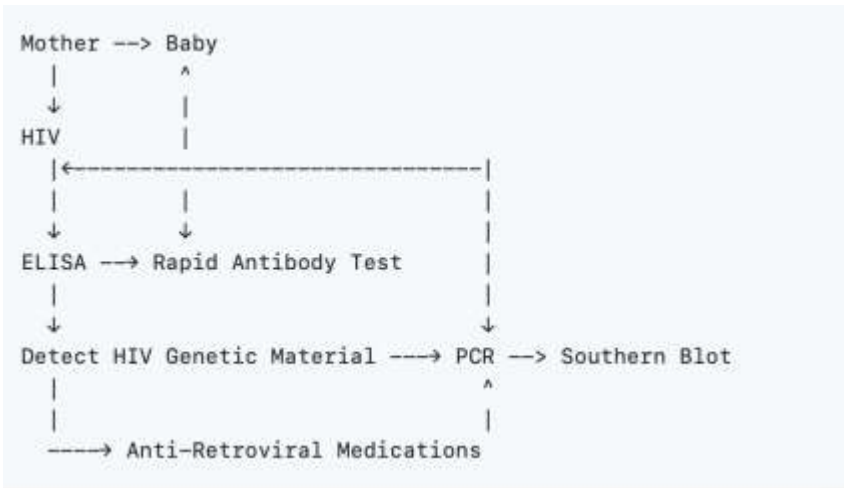
# Example- USMLE Question – ChatGPT

*A 27 year old HIV positive female gave birth to a 3.kg baby girl. The obstetrician is worried that the child may have been infected due to the mother's haphazard use of her anti-retroviral medications. Which of the following assays would a hospital use to detect the genetic material of HIV if the child had been infected? A. Enzyme-linked immunosorbent assay (ELISA). B. Rapid HIV antibody test. C. Polymerase chain reaction. D. Southern blot.*



# Example with UM GPT: Step 1

Draw a causal graph using ASCII relating the following: HIV, Mother, baby, anti-retroviral medications, detect HIV genetic material, ELISA, rapid antibody test, polymerase chain reaction, southern blot



### 1. Mother to Baby Transmission Node:

1. The "Mother" node would have an arrow pointing to the "Baby" node, indicating that HIV can be transmitted from an infected mother to her baby during pregnancy, childbirth, or breastfeeding.

### 2. HIV Detection and Testing Nodes:

1. The "HIV" node would have arrows pointing to both "Detect HIV genetic material" and the "ELISA" node, since the presence of HIV can be detected through these methods.
2. The "Detect HIV genetic material" node would have an arrow leading to the "Polymerase Chain Reaction (PCR)" node and the "Southern Blot" node, as both are methods used to detect the genetic material of HIV specifically.
3. The "ELISA" node would have an arrow pointing to the "Rapid antibody test" node because if the ELISA test is positive, a rapid antibody test might be used to confirm the result.

### 3. Anti-Retroviral Medications Node:

1. The "Anti-retroviral medications" node would have an arrow pointing back to the "HIV" node, indicating that these medications can suppress the viral load in an HIV-positive individual.
2. Additionally, there would be an arrow from the "Anti-retroviral medications" node to the arrow connecting the "Mother" and "Baby" nodes, highlighting that these medications can reduce the risk of mother-to-baby transmission of HIV.

### 4. Impact of Anti-Retroviral Medication on Testing Nodes:

1. There might also be arrows from the "Anti-retroviral medications" node to the testing nodes ("Detect HIV genetic material," "ELISA," "PCR," "Southern Blot"), representing the fact that effective anti-retroviral treatment can lower the viral load below the detectable levels.

# Example with UM GPT: Step 2

Using this causal graph, answer the following question. A 27 year old HIV positive female gave birth to a 3kg baby girl. The obstetrician is worried that the child may have been infected due to the mother's haphazard use of her anti-retroviral medications. Which of the following assays would a hospital use to detect the genetic material of HIV if the child had been infected?

- A. Enzyme-linked immunosorbent assay (ELISA).
- B. Rapid HIV antibody test.
- C. Polymerase chain reaction.
- D. Southern blot.

***In a newborn baby born to an HIV-positive mother, the standard test to directly detect the genetic material of HIV is the Polymerase Chain Reaction (PCR). This is because the standard antibody-based tests, such as ELISA (Enzyme-linked immunosorbent assay) and rapid HIV antibody tests, are not reliable indicators of infection in a newborn. The maternal antibodies against HIV can cross the placenta to the child and may give a positive result in these assays, despite the baby not being infected.***

***Therefore, the correct answer is C. Polymerase chain reaction (PCR). ....***

# Using GPT4o to generate knowledge graph (RDF\*)

@prefix : <http://example.org/hiv/>.

:Mother :hasCondition :HIV .

:Mother :canTransmitTo :Baby .

:AntiRetroviralMedications :reducesRiskOf :HIVTransmission .

:AntiRetroviralMedications :givenTo :Mother .

:Baby :testedFor :HIV .

DetectHIVGeneticMaterial :involves :PCR .

DetectHIVGeneticMaterial :involves :ELISA .

DetectHIVGeneticMaterial :involves :RapidAntibodyTest . DetectHIVGeneticMaterial :involves :SouthernBlot .

PCR :typeOf :Test .

ELISA :typeOf :Test .

RapidAntibodyTest :typeOf :Test . . :

SouthernBlot :typeOf :Test .

*\*RDF- Resource Description Format*  
*Subject : Predicate : Object*

# Using generated knowledge graph to answer question

Use the resulting RDF triples to answer the following question: "A 27 year old HIV positive female gave birth to a 3kg baby girl. The obstetrician is worried that the child may have been infected due to the mother's haphazard use of her anti-retroviral medications. Which of the following assays would a hospital use to detect the genetic material of HIV if the child had been infected?"

Based on the causal graph and RDF triples, the appropriate assays for detecting the genetic material of HIV in the baby would primarily be **PCR**. Therefore, the correct answer is:

## **C) Polymerase Chain Reaction (PCR)**

This conclusion aligns with the RDF triples indicating that PCR is involved in the detection of HIV genetic material (:DetectHIVGeneticMaterial :involves :PCR .) and the fact that it is a widely used technique in clinical diagnostics for such purposes.

# GPT4o prompting for malaria causal graph

Draw a causal graph linking the following factors for malaria transmission using ASCII: “urban agriculture”, “wells excavation”, “number of breeding sites”, “more rainfall”, “hygiene and sanitation of household compounds”, “adequate house constriction”, “convenient waste management”, “blood stage malaria vaccine”, “transmission blocking vaccine”, “pre-erythrocytic stage vaccine”, “treated bednets”, “adults”, “children under 5 years”, “children over 5 years”

Represent the causal graph using RDF triples

# LLM-generated RDF for malaria

@prefix : <http://example.org/malaria#> .

:More\_Rainfall :causes :Number\_of\_Breeding\_Sites .  
:Urban\_Agriculture :causes :Number\_of\_Breeding\_Sites .  
:Wells\_Excavation :causes :Number\_of\_Breeding\_Sites .  
:Hygiene\_and\_Sanitation\_of\_Household\_Compounds :reduces :Number\_of\_Breeding\_Sites .  
:Convenient\_Waste\_Management :reduces :Number\_of\_Breeding\_Sites .  
:Blood\_Stage\_Malaria\_Vaccine :reduces :Malaria\_Transmission .  
:Transmission\_Blocking\_Vaccine :reduces :Malaria\_Transmission .  
:Pre\_erythrocytic\_Stage\_Vaccine :reduces :Malaria\_Transmission .  
:Treated\_Bednets :reduces :Risk\_of\_Mosquito\_Bites .  
:Adequate\_House\_Construction :reduces :Risk\_of\_Mosquito\_Bites .  
:Risk\_of\_Mosquito\_Bites :causes :Malaria\_Transmission .  
:Adults :risk\_factor :Malaria\_Transmission .  
:Children\_Under\_5\_Years :risk\_factor :Malaria\_Transmission .  
:Children\_Over\_5\_Years :risk\_factor :Malaria\_Transmission .  
:Convenient\_Waste\_Management :causes :Adequate\_House\_Construction .  
:Hygiene\_and\_Sanitation\_of\_Household\_Compounds :causes :Adequate\_House\_Construction .

# Chain of Cause with LLM RDF triples

Using the generated RDF triples, what is the impact of liver stage drug resistance of malaria parasites on transmission blocking vaccines?

*Given the causal relationships shown in the RDF triples, the inferred relationships can be represented as:*

*:Liver\_Stage\_Drug\_Resistance :increases :Parasite\_Burden .*

*:Parasite\_Burden :increases :Gametocyte\_Level .*

*:Gametocyte\_Level :increases :Dependency\_on\_TBV .*

# Advantages of Chain of Cause Reasoning

Causal grounding

Tests knowledge of LLM beyond its ability to give the correct answer

Transparency

Explainability, logical basis for answers

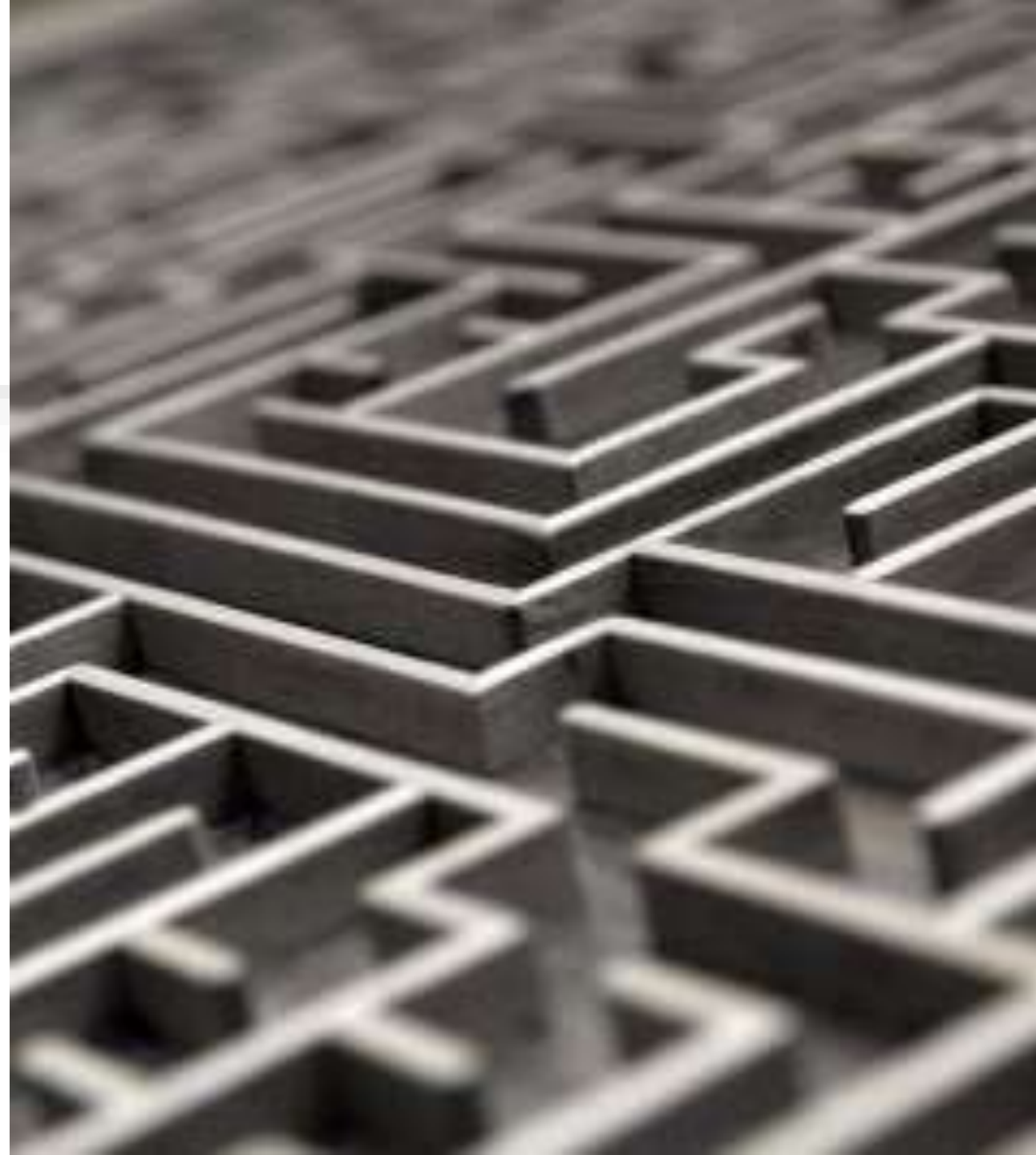
Minimize confabulation (hallucination)

Use of RDF could enable integration with other reasoning engines & knowledge graphs



# Limitations

- Requires a causal framework
- Dependent on existing LLMs- some variations between LLMs e.g. GPT vs Falcon vs Claude
- Requires specific entities of interest to be prespecified (this can be automated)



# Acknowledgements

## Center for Global Health Equity

- Dr. Akbar Waljee (PI: UZIMA DS-I Africa)

## Funding

- Bill & Melinda Gates Foundation
- NIH MICHR K12 Award
- Center for Global Health Equity, University of Michigan

