Causal AI for Web and Healthcare

Utkarshani Jaimini*, Usha Lokala*, Kaushik Roy, Amit Sheth
Artificial Intelligence Institute, University of South Carolina, USA
{ujaimini, nlokala, kaushikr}@email.sc.edu, {amit}@sc.edu. *denotes equal contribution.

#causality, #causalAI, #CausalKG, #explainability, #web, #healthcare
Website: http://aiisc.ai , http://wiki.aiisc.ai
Tutorial website: https://aiisc.ai/causalai/

UNIVERSITY OF South Carolina

AI INSTITUTE #AIISC UNIVERSITY OF SOUTH CAROLINA

Team

Utkarshani Jaimini
Ph.D. Student
Artificial Intelligence Institute
University of South Carolina

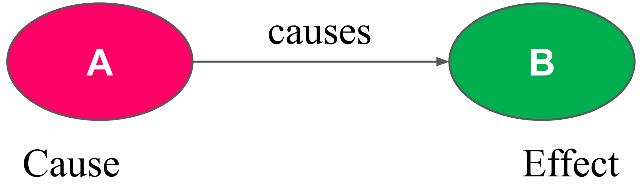
Usha Lokala
Ph.D. Student
Artificial Intelligence Institute
University of South Carolina

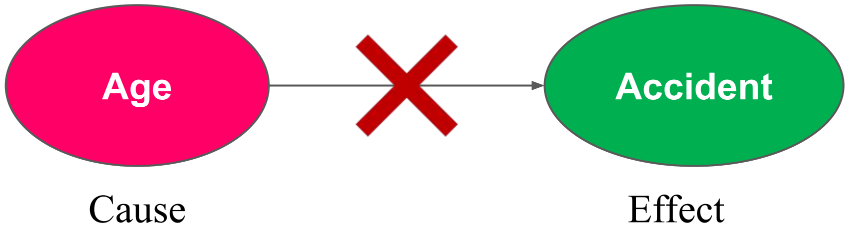
Kaushik Roy
Ph.D. Student
Artificial Intelligence Institute
University of South Carolina

Amit Sheth
Director
Artificial Intelligence Institute
University of South Carolina

Causality

- "**Causality** is an influence by which one event, process, state, or object (*a cause*) contributes to the production of another event, process, state, or object (an *effect*) where the cause is partly responsible for the effect, and the effect is partly dependent on the cause"
- Causality is at the core of everything we see, do, and imagine.
- Causality is a relationship "A" causes "B"
- Correlation is not Causation
- Younger drivers have high probability of being in an accident
  - Does not imply younger drivers cause accident





Why causality

Daily Mail .com

Home | Showbiz | Femail | Royals | Health | Science | Sports | Politics | U.K. | Money | V

Breaking News | Russia-Ukraine | Economy | Joe Biden | Donald Trump | Bryan Kohberger | Prince Harry | V

MORE evidence smokers are at less risk of Covid-19: Study of 90,000 infected patients in Mexico reveals adults addicted to cigarettes are 23% LESS likely to catch the virus

- **The research also showed smokers are less likely to suffer adverse outcomes**
- **It adds to the theory that smokers are somehow protected from the coronavirus**
- **International researchers have reported a low prevalence of smokers in patients**
- **Scientists say nicotine may be able to block the coronavirus from entering cells**
- **Doctors are keen to start trials of nicotine patches, but advise kicking the habit**
- **Here's how to help people impacted by Covid-19**

Observational Study > Nicotine Tob Res. 2021 Aug 4;23(8):1398-1404.
doi: 10.1093/ntr/ntab004.

Impact of Tobacco Smoking on the Risk of COVID-19: A Large Scale Retrospective Cohort Study

Nicolas Paleiron [1], Aurélie Mayet [2], Vanessa Marbac [3], Anne Perisse [3], Hélène Barazzutti [1], François-Xavier Brocq [4], Frédéric Janvier [5], Bertrand Dautzenberg [6], Olivier Bylicki [1]

Conclusions: Current smoking status was associated with a lower risk of developing Covid-19 but cannot be considered as efficient protection against infection. The mechanism of the lower susceptibility of smokers to SARS-CoV-2 requires further research.

Should we start smoking?

https://www.dailymail.co.uk/news/article-8398191/MORE-evidence-smokers-risk-Covid-19-23-likely-catch-virus.html , Paleiron N, Mayet A, Marbac V, et al. Impact of Tobacco Smoking on the Risk of COVID-19: A Large Scale Retrospective Cohort Study. *Nicotine Tob Res*. 2021;23(8):1398-1404. doi:10.1093/ntr/ntab004

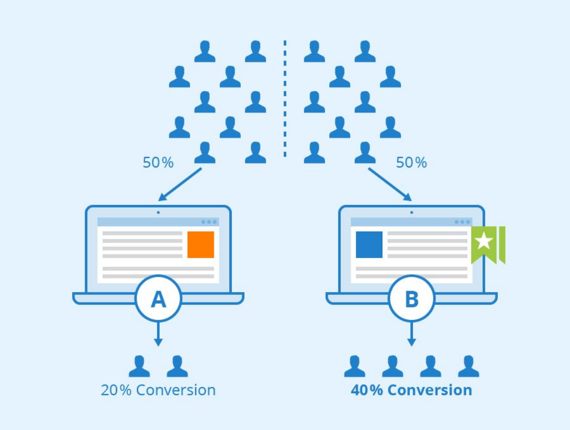
Why causality

- At the start of the pandemic, only healthcare workers (who did not smoke) and people with severe COVID-19 symptoms were tested
- Smokers with no COVID-19 symptoms were under-represented in this observed dataset
- Out of the ones who are tested, non-smokers are more likely to have COVID-19 than smokers
- The data imbalance leads to the correlation that smoking has reduced risk of COVID-19
- Relying on correlation from observation data can lead to embarrassing, costly, and dangerous mistake
- To overcome similar situations and to make actionable decisions
    - We need to understand cause and effect

https://www.rawpixel.com/image/3393875/free-photo-image-australia-australian-defence-force-cc0, https://www.flickr.com/photos/26344495@N05/50996232256

Use of causality in modern days

- User behaviour understanding in industries like Spotify, Amazon, Netflix, Intuit etc
- Randomized control trial to understand the effect of a drug
- Policy intervention
- Social media intervention to study the population dynamics



Observation data is not sufficient for the above use cases

We cannot go back in time and observer the changes

https://www.seobility.net/en/wiki/AB_Testing , https://pix4free.org/photo/15086/intervention.html

Tutorial Highlights

- Why statistical AI alone is not enough?
- Causal AI and causal knowledge graph as a step towards neuro-symbolic AI
- Can ontologies be used as inference for causal explanations?
- Can causal AI enable intervention planning and policy decisions making?
- Can causal inference serve as a bridge between prediction and decision making?

Table of Content

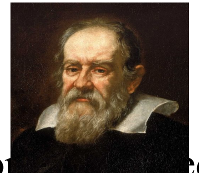
- Causal AI Primer
    - CausalKG: Causal knowledge graph
- Ontology and knowledge-based inferences for causal explanations
- Application of causal AI in web and healthcare use cases

Causal AI Primer

Causation and Statistics

Sir Francis Bacon

Aristotle



Galileo Galilei

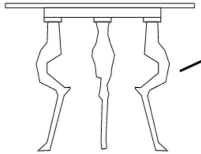
We have proper knowledge of the things only when we have understood its cause

What is it made out of?

Material Cause:
Wood

What is it good for?

Final Cause:
Dining



Formal Cause:
Design

What is it?
Design or form

Efficient Cause:
Carpentry

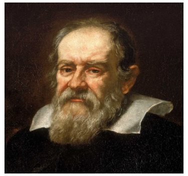
Where does change (motion) comes from?
Process to create

300 BC …. 1500 1600 …. 1900 1930 1960 1990

Causation and Statistics

Aristotle

Sir Francis Bacon



Galileo Galilei

George Udny Yule

Charles Spearman

Sewall Wright

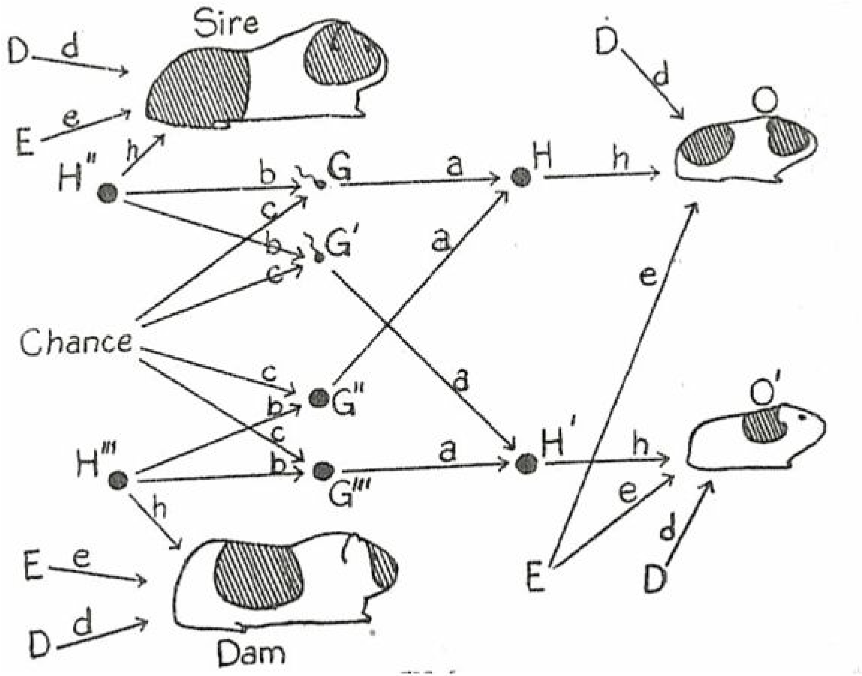


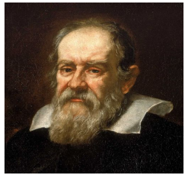
Figure I. Wright's First Path Diagram showing the influence of heredity and environment on the inheritance of color in the guinea pig (reproduced from Wright, 1920, p.328).

300 BC …. 1500 1600 …. 1900 1930 1960 1990

Causation and Statistics

Aristotle

Sir Francis
Bacon



Galileo Galilei

George Udny Yule

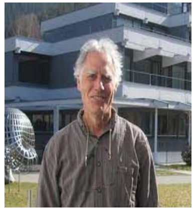
Charles Spearman

Sewall Wright

Sir Ronald A. Fisher

Jerzy Neyman

Graphical Causal
Models



Jamie Robins

Don Rubin

Potential
Outcomes

Judea Pearl

300 BC …. 1500 1600 …. 1900 1930 1960 1990

$\color{red}{\text{Causal}}$ Inference requires more than Probability

Prediction from Observation ≠ Prediction from Intervention

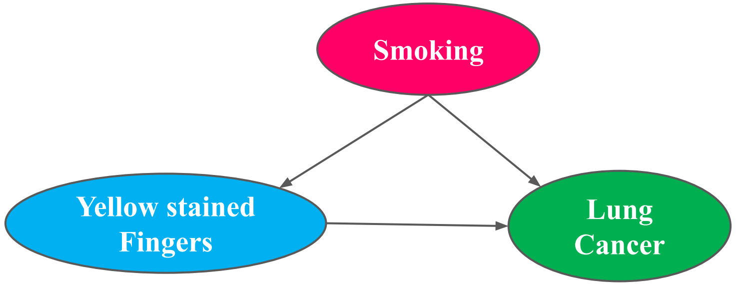
- $P(Y=y \mid X=x, Z=z) \neq P(Y=y \mid X_{set}=x, Z=z)$

P(Lung cancel 1960 = Yes | Tar stained fingers 1950 = no) (observation)

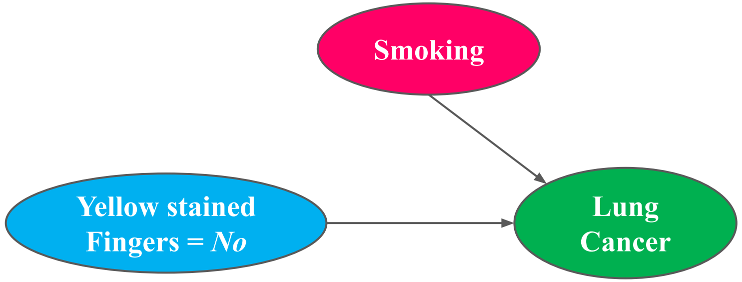
≠

P(Lung cancel 1960 = Yes | Tar stained fingers 1950_{set} = no) (intervention)

- Conditional probability is not equal to interventional probability



Observation



Intervention

We set their fingers to be clean, used soap
Lung cancer is not caused by fingers but by smoking etc.

Causal Prediction vs. Statistical Prediction

$$
\begin{tikzcd}
\text{Non-experimental data (Observational study)} \arrow[r] \arrow[dr] & P(Y,X,Z) \arrow[r] \arrow[dr] & P(Y=y \mid X=x, Z=z) \\
\text{Background Knowledge} \arrow[r] \arrow[ur] & \text{Causal Structure} \arrow[r] & P(Y=y \mid X_{set}=x, Z=z)
\end{tikzcd}
$$

Causal Estimation vs. Causal Search

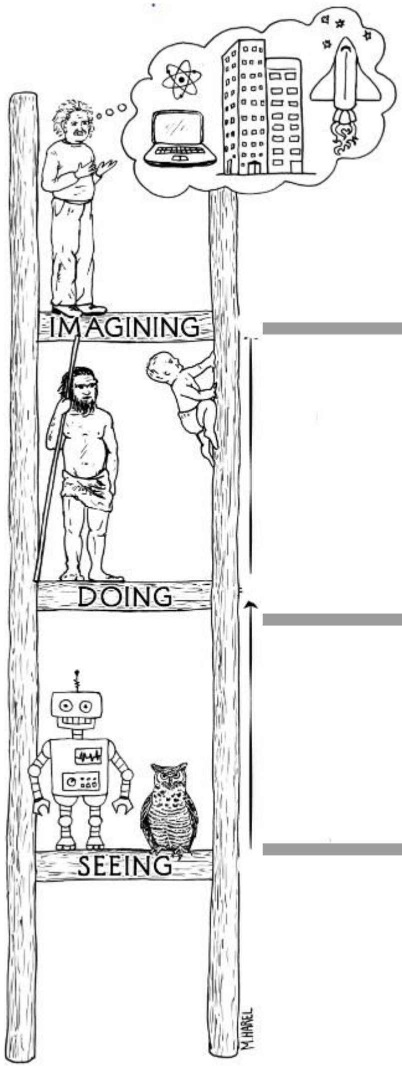
Estimation	Search
Causal Question: Effect of zivoduvine on survival among HIV-positive men (Hernan, et.al., 2000)	**Causal Question:** Which genes regulate flowering Arbidopsis?
Problem: Confounders (CD4 lymphocyte count) vary over time and are dependent on previous treatment of zivoduvine	**Problem:** over 25000 potential genes
Estimation method: marginal structural model, path specific policies, dynamic treatment regimes	**Estimation method:** graphical structural learning
Assumptions: Treatment measured reliably, measure covariates sufficient to capture major source of confounding, model of treatment given the past is accurate	**Assumptions:** RNA microarray measurement are reasonable proxy for gene expression Causal Markov assumption, etc
Output: Effect estimate with confidence interval	**Output:** Possible model which can be used for follow-up experiment 25 possible model out of which 13 were possible to tested because of seeds availability, 9 produced viable plants out of 4 had successful flowering time

Judea Pearl: Ladder of Causation

Neuro-symbolic AI

Using rich symbolic knowledge which can be used for reasoning task such as causality

Current AI



3. COUNTERFACTUALS, $P(y_x \mid x', y')$

ACTIVITY: Imagining, Retrospection, Understanding

QUESTIONS: What if I had done ….? What if I had acted differently? Was it X that caused Y? What if X had not occurred?

EXAMPLES: Was it the aspirin that stopped my headache? What if I had not smoked last 2 years?

2. INTERVENTION, $P(y \mid do(x), z)$

ACTIVITY: Doing, Intervening

QUESTIONS: What if I do ….? What if I do X? What would Y be if I do X?

EXAMPLES: If I take aspirin will my headache be cured? What if we ban cigarettes?

1. ASSOCIATION, $P(y \mid x)$

ACTIVITY: Seeing, Observing

QUESTIONS: What if I see….? How would seeing X change my belief in Y?

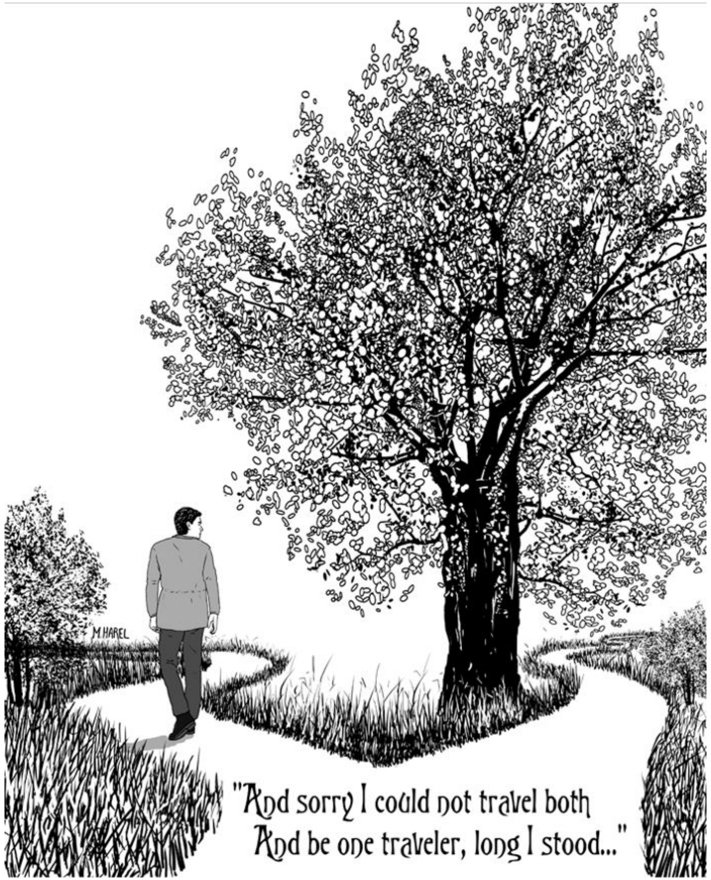
EXAMPLES: What does the symptom tell me about the disease? What does a survey tell us about the election?

Pearl, Judea, and Dana Mackenzie. *The book of why: the new science of cause and effect*. Basic books, 2018.

Counterfactuals

WHAT If? scenarios

- Human mind has an ability to conceive alternative, nonexistent worlds
- We can see what might have happened
  - Imagine, be prepared and act in counterfactuals scenarios



Counterfactuals
(The path not taken by Robert Frost)

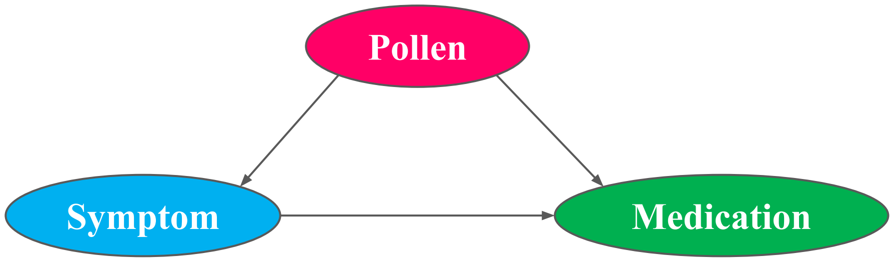
http://danamackenzie.com/amazing-science-art/

“Counterfactual reasoning requires Causal structure”

Causal Bayesian Network

A Causal Bayesian Network is a graphical representation to express causal knowledge

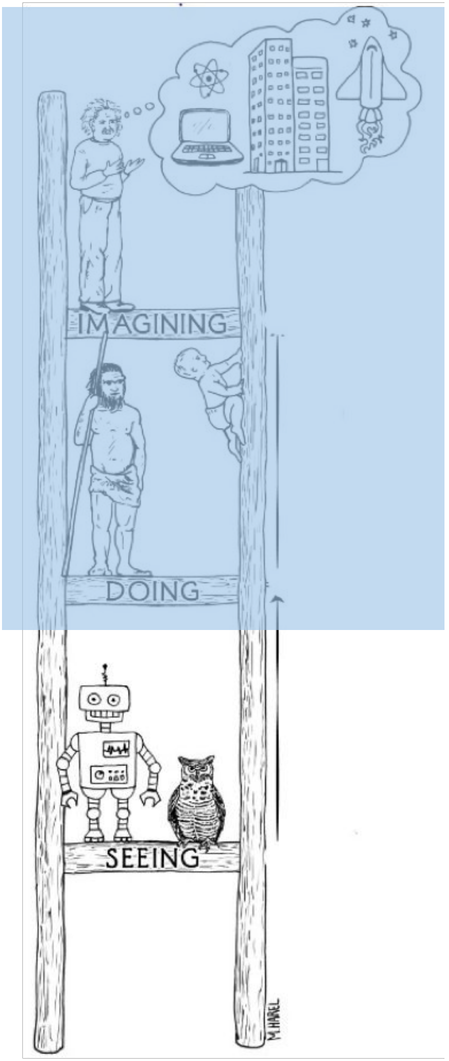
- Causal Bayesian Network (CBN) is a Bayesian network in which,
  - Each node is independent of all its non-descendants given its parents
  - The directed edges represents causal relationship between the corresponding nodes.



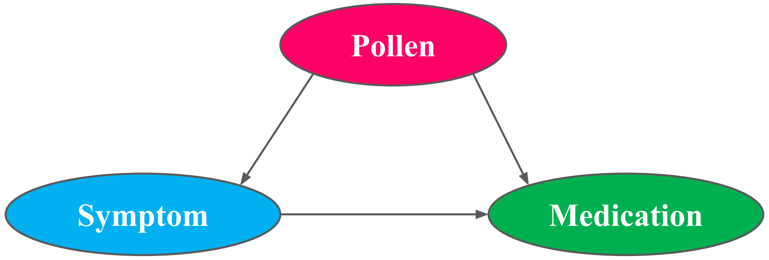
Asthma causal bayesian network

Pearl, Judea. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan kaufmann, 1988.

Ladder of Causation: Association



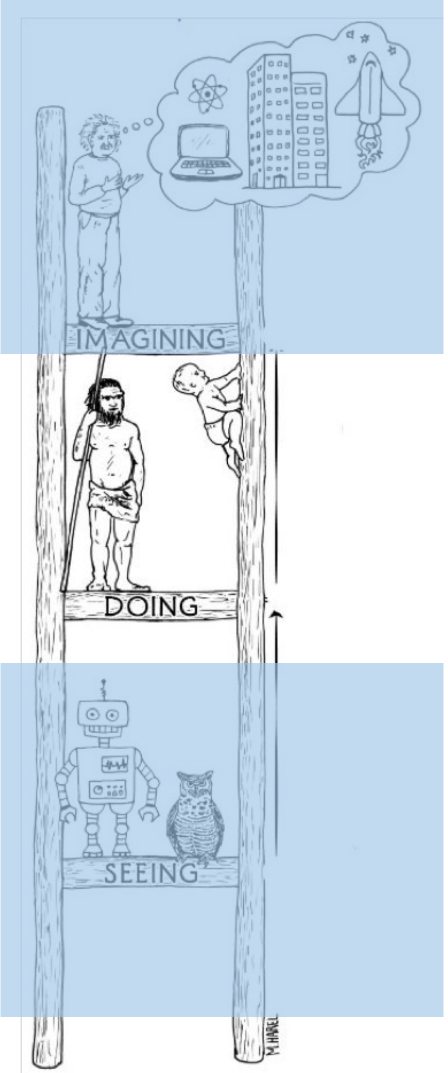
What does the presence of symptom tells us about the intake of medication by a patient?



P(Medication = Y | Symptom = Y)

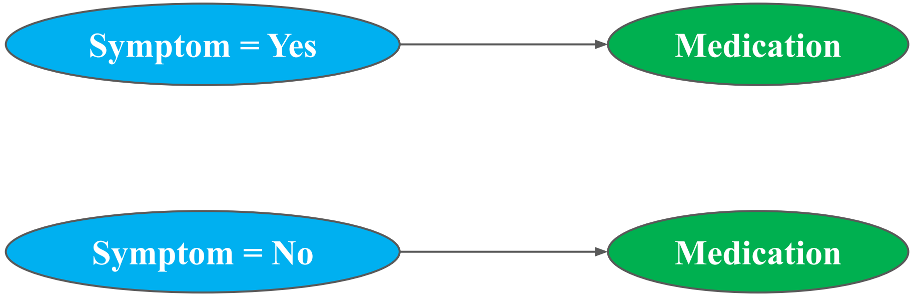
Probability of intake of medication given symptom is True (observed)

Ladder of Causation: Intervention



What would be the effect of symptom on medication intake ?

$$TCE_{No \rightarrow Yes} = E[Medication \mid do(Symptom = Yes)] - E[Medication \mid do(Symptom = No)]$$



Total Causal Effect

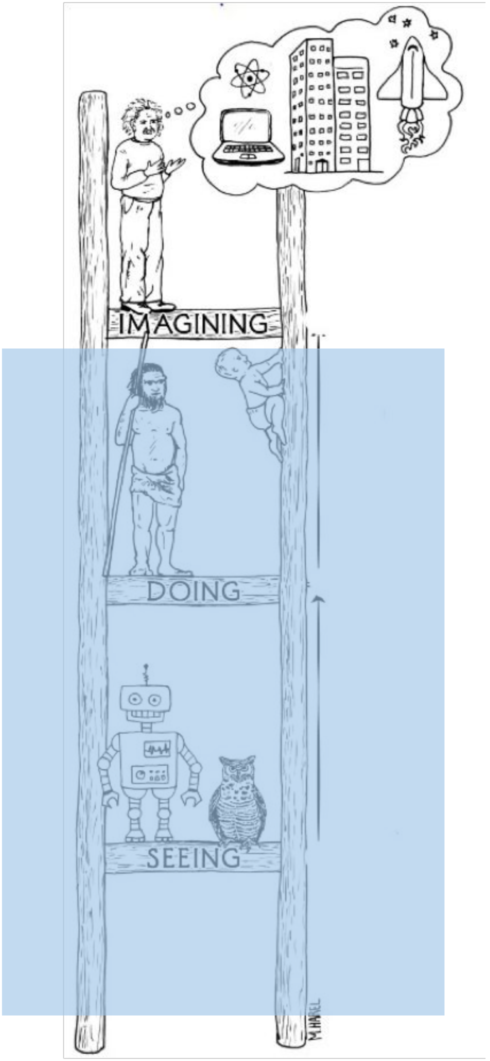
The Total Causal Effect (TCE) of a binary treatment T on an outcome Y is defined as the interventional contrast.

$$TCE_{0 \rightarrow 1} = E[Y \mid do(T = 1)] - E[Y \mid do(T = 0)]$$

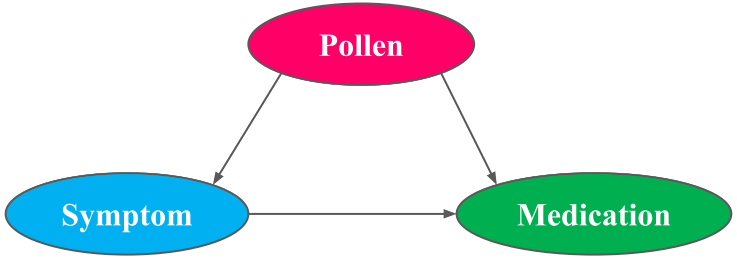
where the expectations are taken over the interventional distributions

$P(Y \mid do(T = 1))$ and $P(Y \mid do(T = 0))$

Ladder of Causation: Counterfactual



Given a patient has taken the medication, what is the chance they would not taken it if they didn't have a symptom?



P(Medication = N | Medication =Y, do(Symptom = N))

Ladder of Causation: Counterfactual

Natural Direct Effect (NDE)

Given the pollen in the outdoor environment, what is the chance the patient took the medication but the symptom was not due to pollen?

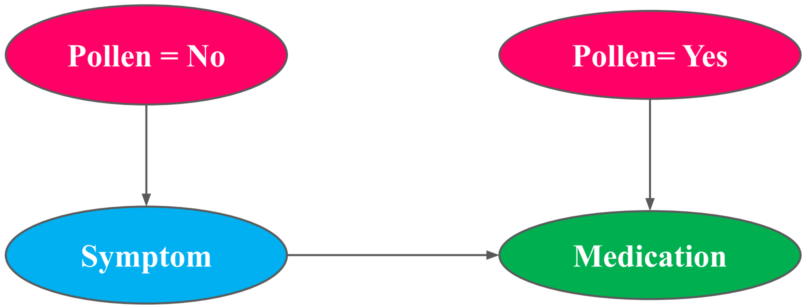
Natural Direct Effect (NDE) of a binary treatment T on an outcome Y with mediator X is given by the counterfactual contrast

$$NDE_{0\to 1} = E[Y_{X(0)} \mid do(T=1)] - E[Y \mid do(T=0)]$$

where the subscript X(0) refers to the counterfactual distribution of X but T been 0, and where the expectations are over both Y and X w.r.t. the corresponding interventional and counterfactual distributions.

- The effect of *treatment* on the *outcome* in the presence of the *mediator* (counterfactual = 0)

$$NDE_{No \to Yes} = E\,[Medication(Yes)_{Symptom\,(Yes,\,No)} \mid do(Pollen=True)] - E\,[Medication(Yes) \mid do(Pollen=False)]$$



Ladder of Causation: Counterfactual

Natural Indirect Effect (NIE)

Given there is no allergen (pollen) in the outdoor environment, what is the chance the patient took the medication but the symptom was due to an allergen (pollen)?

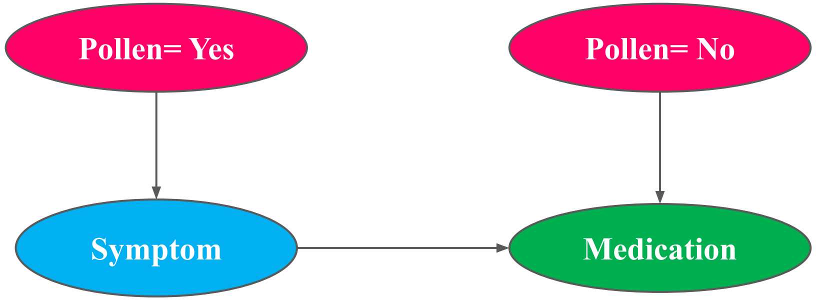
The Natural Indirect Effect (NIE) of a binary treatment T on an outcome Y with mediator X is given by the counterfactual contrast

$$NIE_{0\to 1} = E[Y_{X(1)} \mid do(T=0)] - E[Y \mid do(T=0)]$$

where the subscript X(1) refers to the counterfactual distribution of X had T been 1, and where the expectations are over both Y and X w.r.t. the corresponding interventional and counterfactual distributions

- The effect of *treatment* on the *outcome* in the presence of the *mediator* (counterfactual = 1)

$$NIE_{No\to Yes} = E[Medication(Yes)_{Symptom(Yes,\ Yes)} \mid do(Symptom = No)] - E[Medication(Yes) \mid do(Symptom = No)]$$



Causality and Knowledge Graph

CauseNet: causality graph extracted from the Web

CauseNet is large-scale open domain causal knowledge graph of causal relations between causal concepts, extracted from different web sources.

Linguistic Pattern						Relations
Cause dependency		Token/POS	Effect dependency			|E|
cause/N	-nsubj	cause/VB	+dobj	effect/N		904,385
cause/N	-nmod:with	associated/VBN	-acl	effect/N		892,908
cause/N	-nsubj	lead/VB	+nmod:to	effect/N		783,860
cause/N	-nsubj	led/VBD	+nmod:to	effect/N		724,978
cause/N	-nsubjpass	associated/VBN	+nmod:with	effect/N		692,666
cause/N	-nmod:by	caused/VBN	-acl	effect/N		598,639
cause/N	-nsubj	result/VB	+nmod:in	effect/N		552,352
cause/N	-nsubj	causes/VBZ	+dobj	effect/N		496,426
cause/N	-nsubj	leads/VBZ	+nmod:to	effect/N		491,340
cause/N	-nsubj	resulted/VBD	+nmod:in	effect/N		473,298

Top 10 patterns of causal relations: cause/N and effect/N refer to nouns within causal concepts, joined in a sentence fulfilling the respective pattern's dependencies.

(a)

Hostname	Category	$\|E\|$
sdbonline.org	Science	26,517
bionewsonline.com	Science	25,212
jci.org	Science	16,081
sec.gov	Regional	13,907
plosone.org	Science	12,722
molvis.org	Science	9,544
neurotransmitter.net	Reference	8,842
diseaseinformation.info	Reference	8,829
leninist.biz	Reference	8,033
lansbury.bwh.harvard.edu	Science	7,828

(b)

Domain	Category	Subdomains	$\|E\|$
researchtoday.net	Science	302	125,728
wordpress.com	Society	6,835	91,230
typepad.com	Society	5,687	72,357
hubpages.com	Society	4,370	40,473
nih.gov	Regional	368	40,280
deviantart.com	Arts	20,365	40,064
about.com	Reference	828	36,363
tripod.com	Society	1,877	31,131
sdbonline.org	Science	1	26,517
bionewsonline.com	Science	1	25,212

(c)

TLD	$\|E\|$
.com	5,597,297
.org	2,590,683
.net	793,937
.edu	766,731
.gov	320,263
.co.uk	229,834
.ca	185,661
.info	138,519
.org.uk	111,697
.ac.uk	85,304

(d)

Category	Domains	$\|E\|$
Science	121	296,330
Reference	118	240,033
Health	84	147,851
Society	80	129,058
Regional	34	76,754
Business	21	43,900
News	11	33,906
Computers	18	27,319
Shopping	9	14,078
Arts	3	8,017

(e)

Infobox Template	Articles	$\|E\|$
medical condition (new)	820	4,923
civil conflict	579	1,339
rail accident	452	530
event	380	495
wildfire	257	306
news event	146	170
oil spill	35	36
military conflict	23	32
birth control	13	26
bus accident	20	23

(f)

Wikipedia article title (ibxs.)	$\|E\|$
2013 Romanian protests (...)	23
Shock (circulatory)	19
Breast cancer	18
Constipation	17
Intracerebral hemorrhage	17
Protests against Donald Trump	17
Heat stroke	16
Scombroid food poisoning	16
Acute lymphoblastic leukemia	15
Bowel obstruction	15

(g)

Wikipedia article title (lists)	$\|E\|$
Flushing (physiology)	58
Mast cell activation syndrome	56
Coarse facial features	50
Hypotonia	47
Autistic catatonia	46
Livedo reticularis	46
Pallor	43
Delayed puberty	42
Eosinophilic myocarditis	42
Intraparenchymal hemorrhage	42

(h)

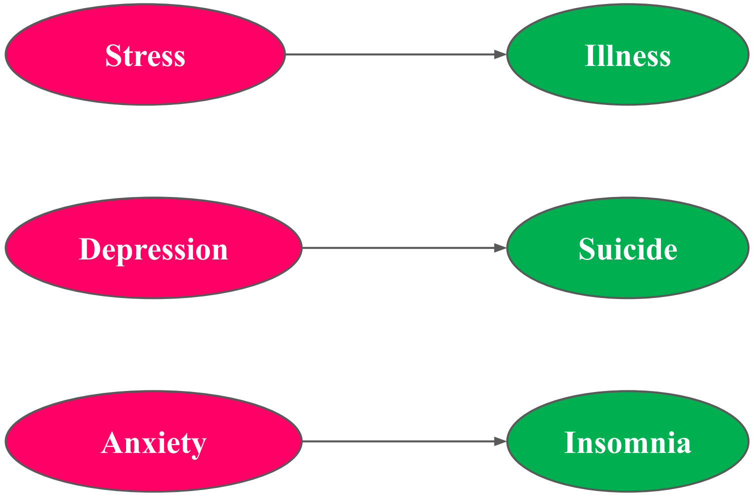
Wikipedia article title (texts)	$\|E\|$
Effects of global warming on human health	98
Hepatitis	79
Horse colic	77
Safety of electronic cigarettes	72
Nutritional neuroscience	71
Causes of cancer pain	70
Dog health	69
Long-term effects of alcohol consumption	69
Famine	67
Progeroid syndromes	60

Source analysis: The upper row shows the top sources of causal relations from ClueWeb12 in terms of (a) hostname, (b) domain, (c) top-level domain, and (d) DMOZ category, ordered by number of causal relations. The lower row shows sources of causal relations in Wikipedia: (e) infobox templates, and articles with causal relations in (f) infoboxes, (g) lists, and (h) texts.

Heindorf, Stefan, et al. "Causenet: Towards a causality graph extracted from the web." *Proceedings of the 29th ACM international conference on information & knowledge management*. 2020.

CauseNet: causality graph extracted from the Web

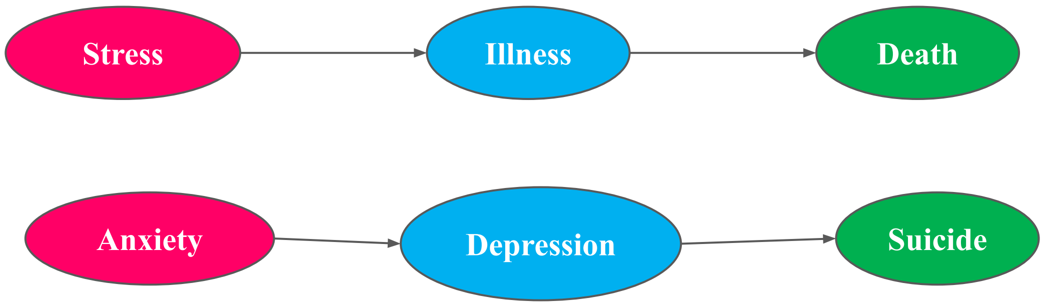
Cause	Effect
accident	death
drought	famine
injury	pain
disease	deaths
smoking	lung cancer
stress	illness
depression	suicide
anxiety	insomnia
bacteria	infection
diarrhea	dehydration



Shows 10 paths of length 1

CauseNet: Mediator path length 2

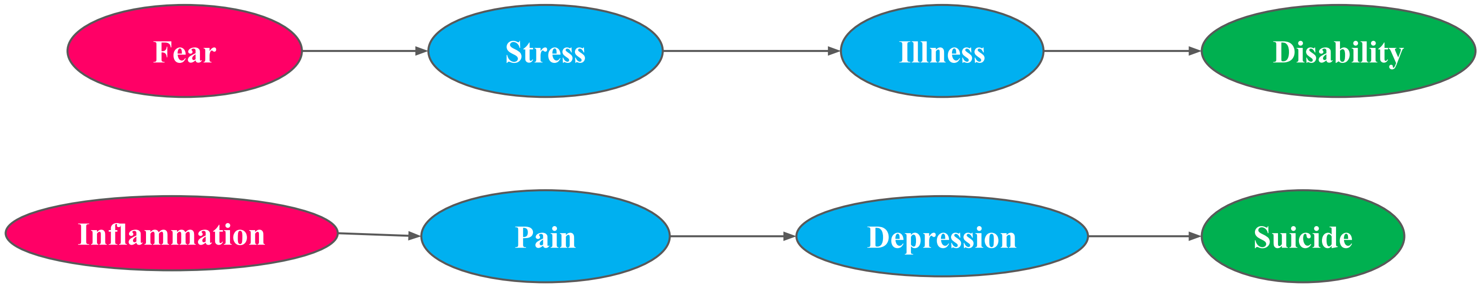
Cause	Mediator	Effect
stress	illness	death
accident	injury	pain
exposure	disease	deaths
bacteria	infection	inflammation
obesity	diabetes	blindness
anxiety	depression	suicide
global warming	drought	famine
diarrhea	dehydration	headaches
lightning	fire	damage
negligence	injuries	disability



Shows 10 paths of length 2

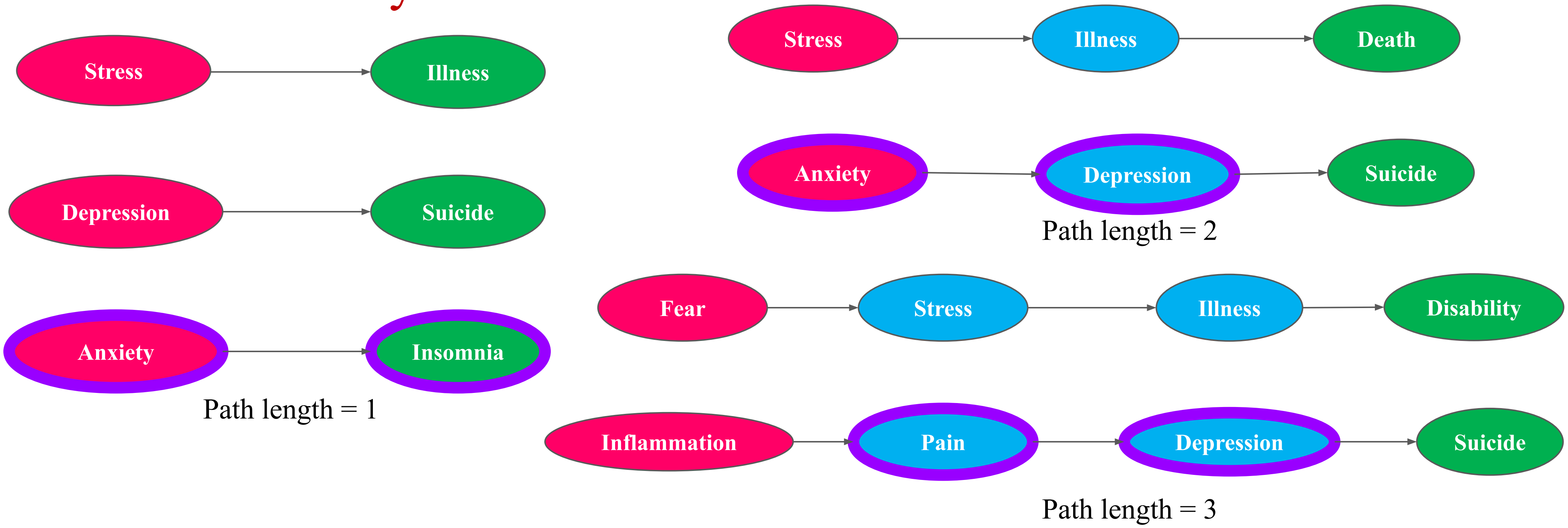
CauseNet: Mediator path length 3

Cause	Mediator 1	Mediator 2	Effect
negligence	accident	injury	death
bacteria	infection	disease	deaths
inflammation	pain	depression	suicide
fear	stress	illness	disability
greenhouse gases	global warming	drought	famine
lack of exercise	obesity	diabetes	blindness
lightning	fire	damage	cancer
virus	diarrhea	dehydration	headaches
anemia	fatigue	accidents	injuries
alcohol	problems	anxiety	insomnia



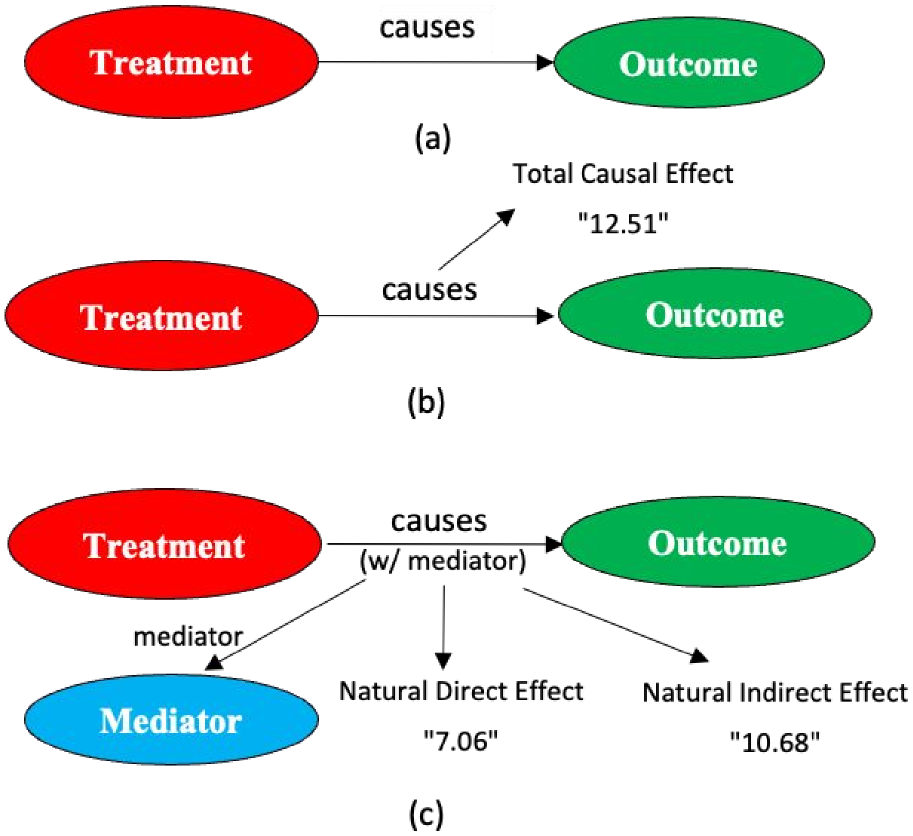
Shows 10 paths of length 3

CauseNet: Pathways



- Different causal chain sequence when increasing the length of the causal pathways
- Need for a Causal Bayesian Network built using the domain knowledge

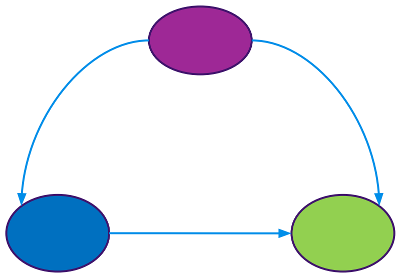
Current Causal representation



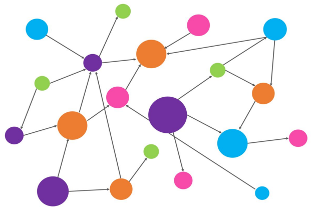
(a) Causal Representation as a single cause-effect relation.
(b), (c) Causality as a complex representation of causal effect associated with the different pathways.

Jaimini, Utkarshani, and Amit Sheth. "CausalKG: Causal Knowledge Graph Explainability using interventional and counterfactual reasoning." *IEEE Internet Computing* 26.1 (2022): 43-50.

Causal Knowledge Graph



Causal Bayesian Network



Knowledge Graph

CausalKG

Pros

Models the causal relations between the corresponding nodes

Cons

Different experts can suggest different causal model

Pros

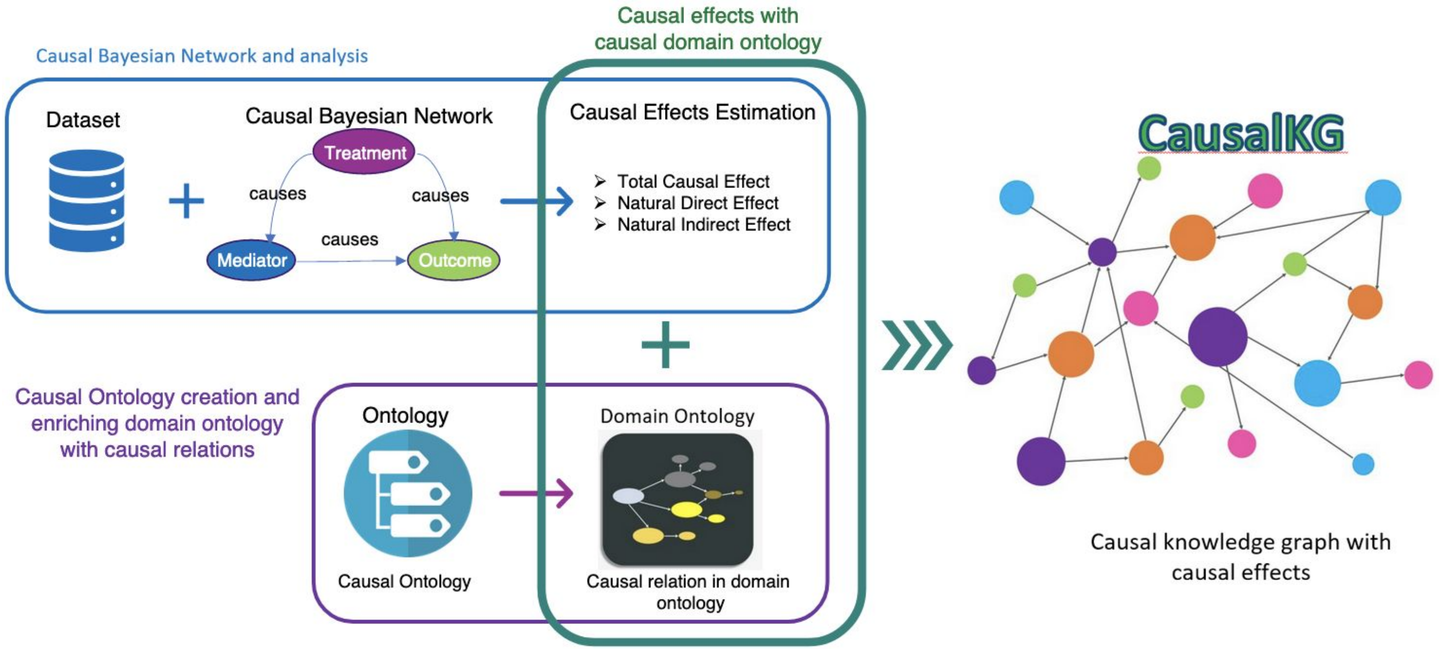
Extensive domain knowledge representation

Cons

Lacking causal knowledge representation techniques

Representation of causal facts from the dataset using causal Bayesian networks, causal ontology and knowledge representation technique for knowledge graph downstream tasks

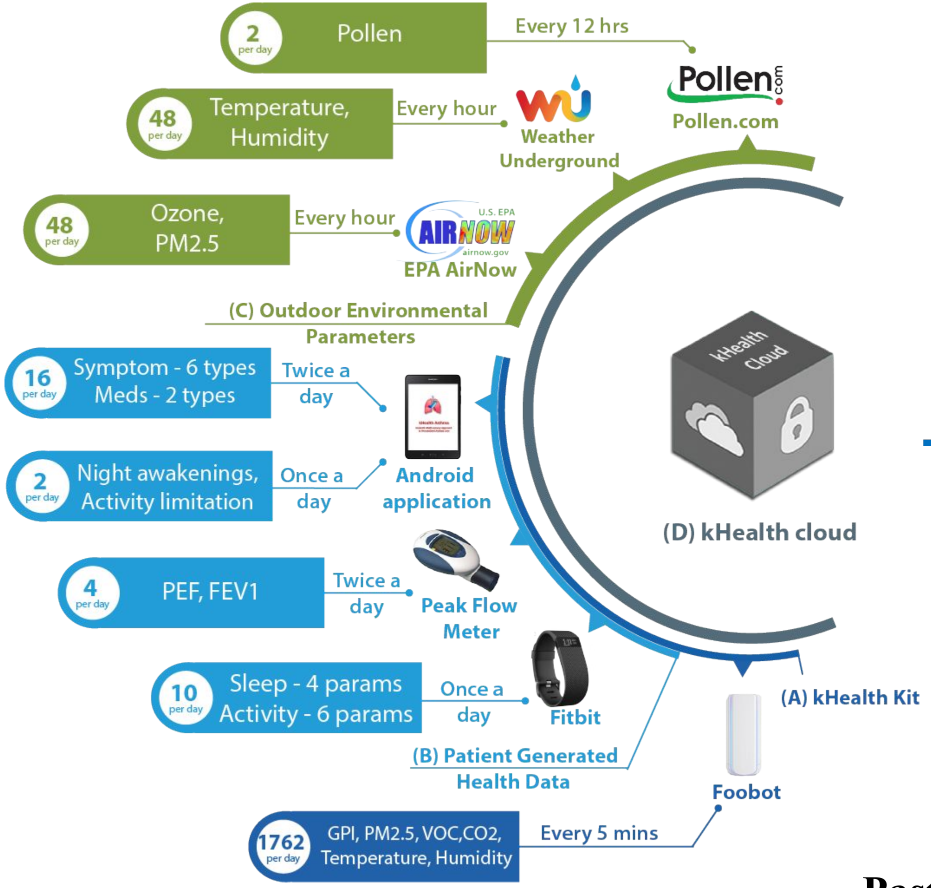
Causal Knowledge Graph Architecture



CausalKG framework consists of three main steps, 1) a **Causal Bayesian Network** and a domain-specific observational dataset, 2) **Causal Ontology** creation and enriching the domain ontology with causal relationships, and 3) Estimating the **causal effects** in the domain for a given context.

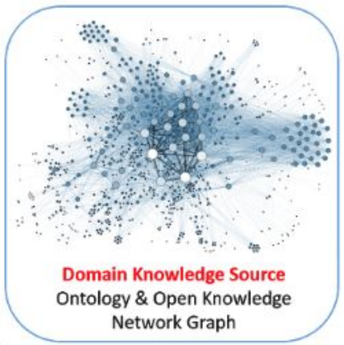
Jaimini, Utkarshani, and Amit Sheth. "CausalKG: Causal Knowledge Graph Explainability using interventional and counterfactual reasoning." *IEEE Internet Computing* 26.1 (2022): 43-50.

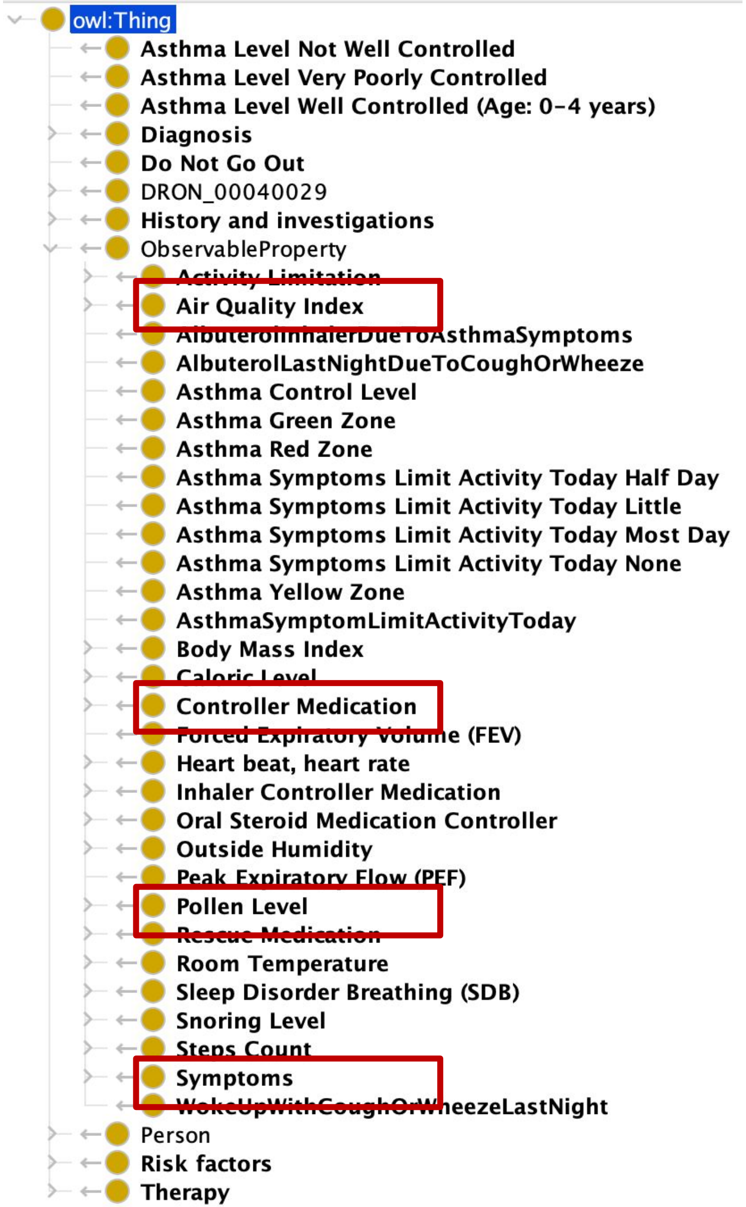
kHealth



+

Asthma Ontology



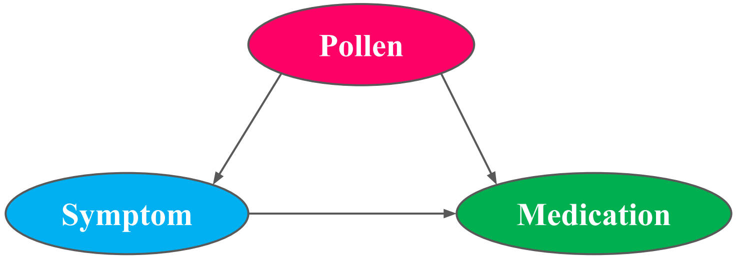


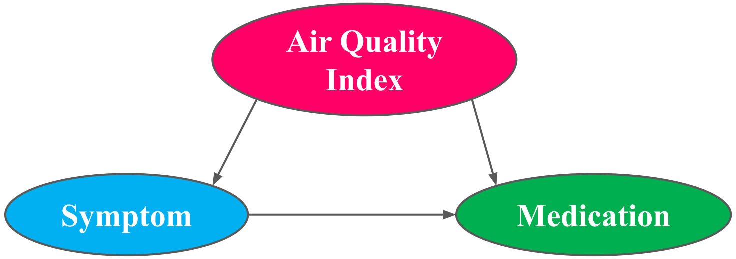
Many Sources of Highly Diverse Data
(& collection methods: Active + Passive):
Up to 1852 data points/ patient /day

Past observation based on correlation

Venkataramanan, Revathy, Krishnaprasad Thirunarayan, Utkarshani Jaimini, Dipesh Kadariya, Hong Yung Yip, Maninder Kalra, and Amit Sheth. "Determination of personalized asthma triggers from multimodal sensing and a mobile app: observational study." *JMIR pediatrics and parenting* 2, no. 1 (2019): e14300.

Personalized Causal Bayesian Network for Asthma

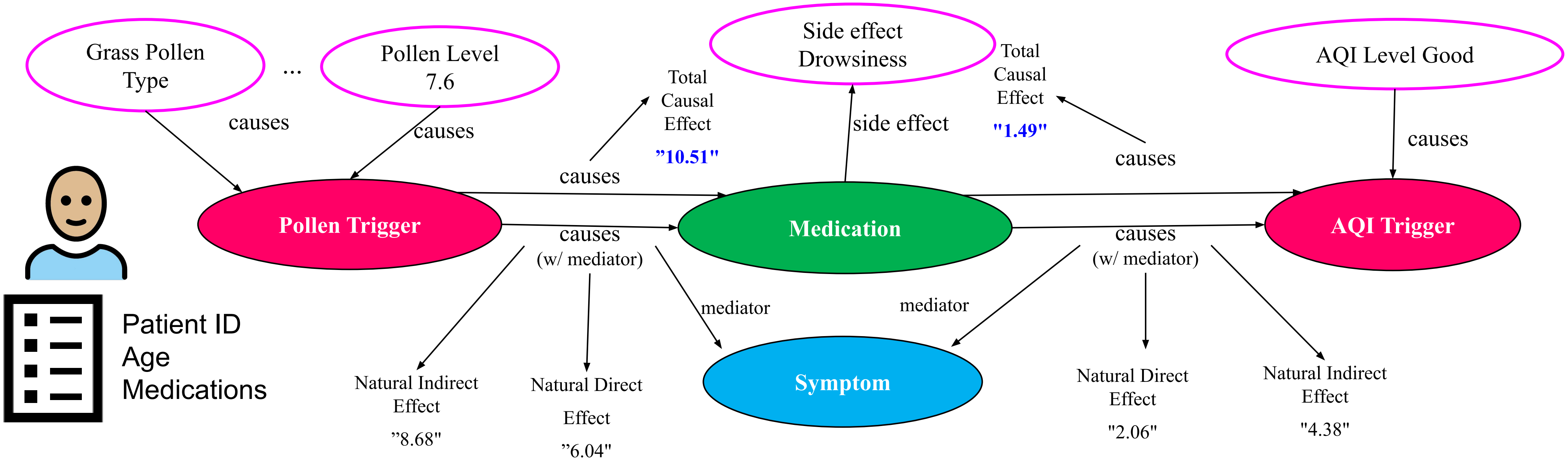




Personalized Causal Bayesian Network, a graphical representation of the causal relations and interaction between a patient's asthma **triggers** (such as pollen, AQI), asthma **symptom** (such as cough, wheeze, chest tightness, hard and fast breathing, and nose opens wide), and **medication** intake (such as controller, rescue, and allergy medication)

CausalKG for Asthma

Causal AI and causal knowledge graph as a step towards neuro-symbolic AI and explanation



CausalKG for a pediatric asthma scenario represented as a **hyper-relational graph**. Each node in the CausalKG is a concept in knowledge graph and is associated with a conditional probability estimated using the causal Bayesian network. The edges between the nodes represent the causal relationships between the concepts.

NDE: Given the pollen (AQI) in the outdoor environment, what is the chance the patient took the medication but the symptom was not due to pollen (AQI)?

- ➔ NDE for pollen (6.04) > NDE for AQI (2.06), pollen has been tested as a allergen for the patient
- ➔ Preventive measure which was common observation in the cohort

https://creazilla.com/nodes/43412-person-emoji-clipart ,

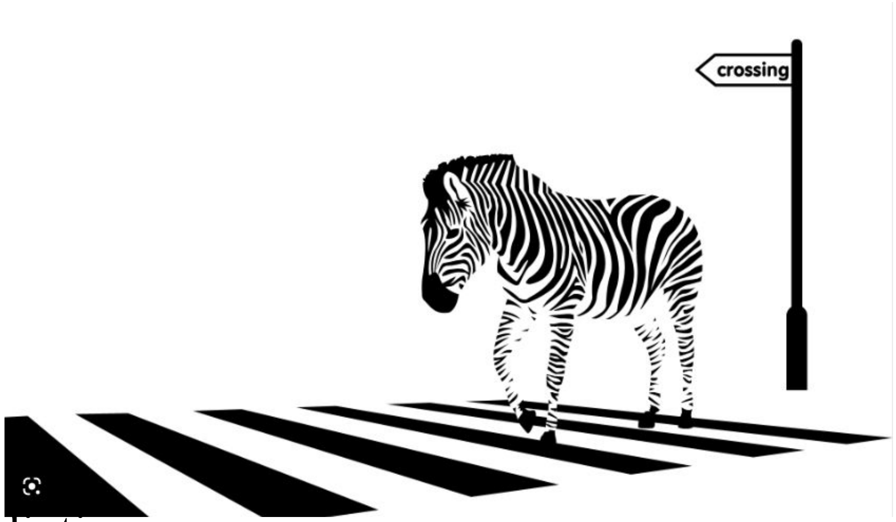
Ontology and Knowledge based inference

Causal Explanation

Why Causal Inference?

The correlational machine learning seeks to identify patterns, but it often identifies false patterns that do not have any actual causal relationship.

- ➔ Predictions from ML are insufficient to address causal inference.
- ➔ Correlation does not imply causation.
- ➔ Data alone is not enough for causal inference.
- ➔ Need domain knowledge and assumptions.
- ➔ Ontology based inference brings explanation to predictions.



“ If two events are related in a certain way, then one event may be the cause of the other. For example, if event A is a necessary condition for event B to occur, then A is likely to be the cause of B”

Traditional ML Explanations Compared to Knowledge Infused Approach

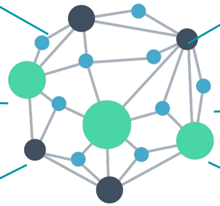
	Medical Diagnosis	Credit Risk Assessment	Customer Churn Prediction
ML	Symptoms Medical history Test results	Income Credit History Debt to Income Ratio	Usage patterns Customer demographics Satisfaction surveys
	Causal Relationships		
KG	Risk factors Comorbidities Treatments	Economic conditions Market trends Government regulations	Product features Customer support Marketing campaigns

Ontology Based Inference

Causal Relevance

Abduction

Causal Consistency



Causal Reasoning

Causal Coherence

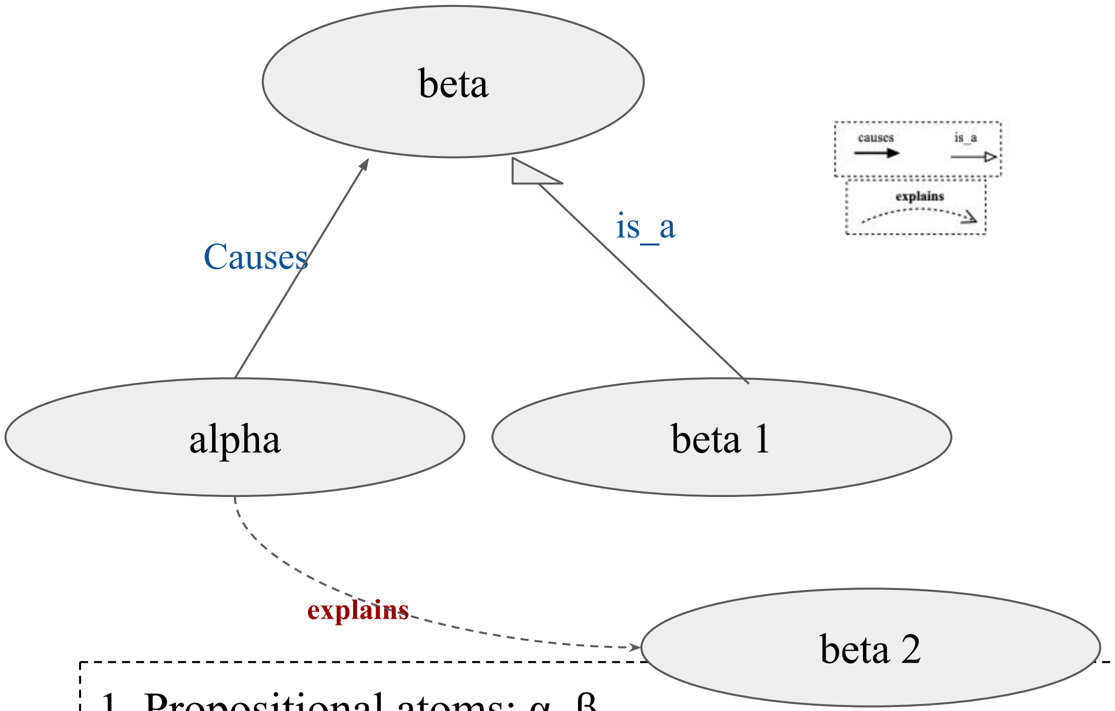
Causal Specificity

Ontology based Inference

Classes	**Properties**	**Relationships**	**Domain Specific**	**Causal Explanation**

Constructing Causal Inference System

We create an inference system that utilizes an IS-A hierarchy ontology to provide explanations based on causal statements. To achieve this, we introduce a basic logical language that enables the expression of causation between two facts and the explanation of one fact by another.



1. Propositional atoms: α, β,
2. Causal atoms: α causes β.
3. Ontological atoms: α →IS−A β.
4. Explanation atoms: α explains β

Causal Inference System Creation

Automatic or semi-automatic methods will depend on the complexity of the problem,the available data.

Test and Refine
Test the causal inference system using additional data and refine it as necessary to improve its accuracy and effectiveness.

Causal Explanation
Requires careful attention to the domain of interest, the available data, and the methods used to identify and evaluate causal relationships.

Steps to Causal Inference

Bayesian networks, decision trees, or rule-based systems

Identify Causal Relationships

Evaluate Causal Relationships

Define Domain

Gather Data

Ontology based Causal Inference - Methods

- One of the primary approaches to constructing causal models based on ontology is the **use of Bayesian networks**. In this approach, causal relationships are represented by nodes in a directed acyclic graph. Methods to learn the structure of Bayesian networks, include constraint-based, score-based, and hybrid approaches.

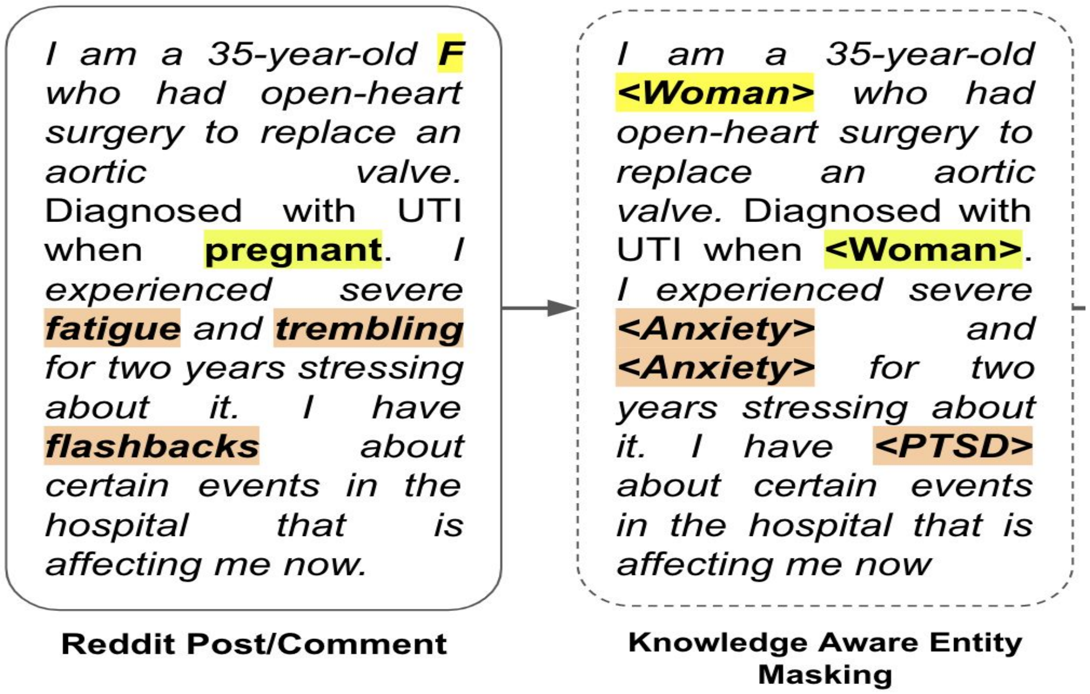
- **Ontology-based causal discovery algorithms** use domain knowledge encoded in ontologies to guide the search for causal relationships in data. The ontology can help identify relevant variables and relationships that may not be obvious from the data alone.

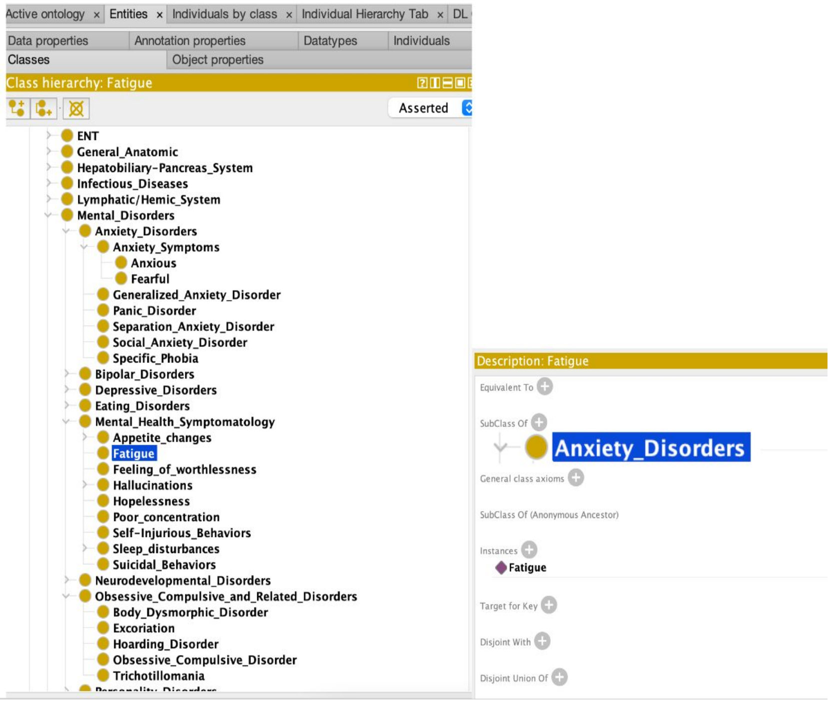
- **Abductive reasoning** - To find the most plausible explanation for a given set of observations.

- **Hybrid approaches** - for example, combining Bayesian networks and ontology-based causal discovery to model causal relationships.

ontology-based inference for causal explanation has the potential to provide a more structured and comprehensive way of understanding and modeling causal relationships.

Ontology based inference for Causal Explanation - Mental Health Domain





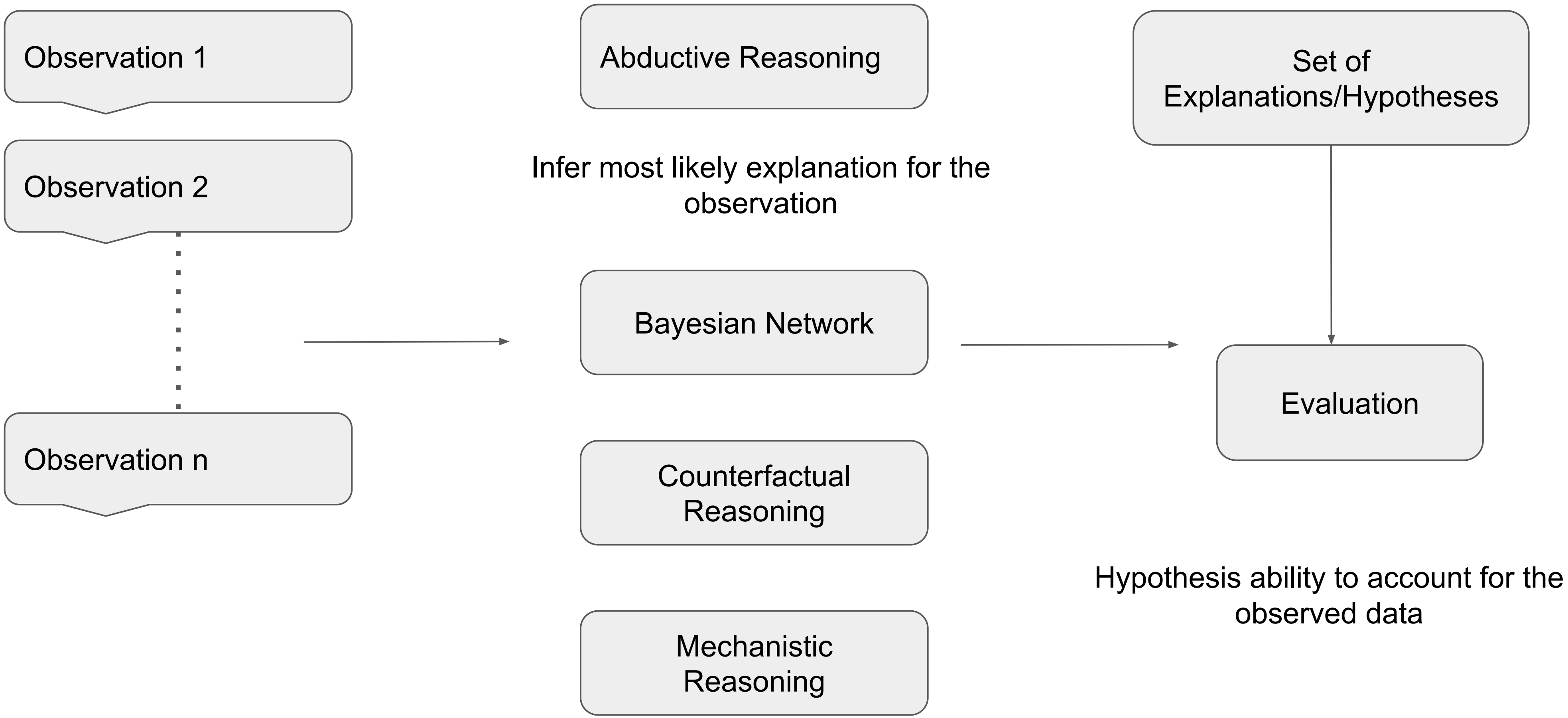
TASK: Explain *Gender and Mental Health Prediction?*

(**Q**): How is the Post expressing depression , anxiety or PTSD, if so Can we find the gender language?

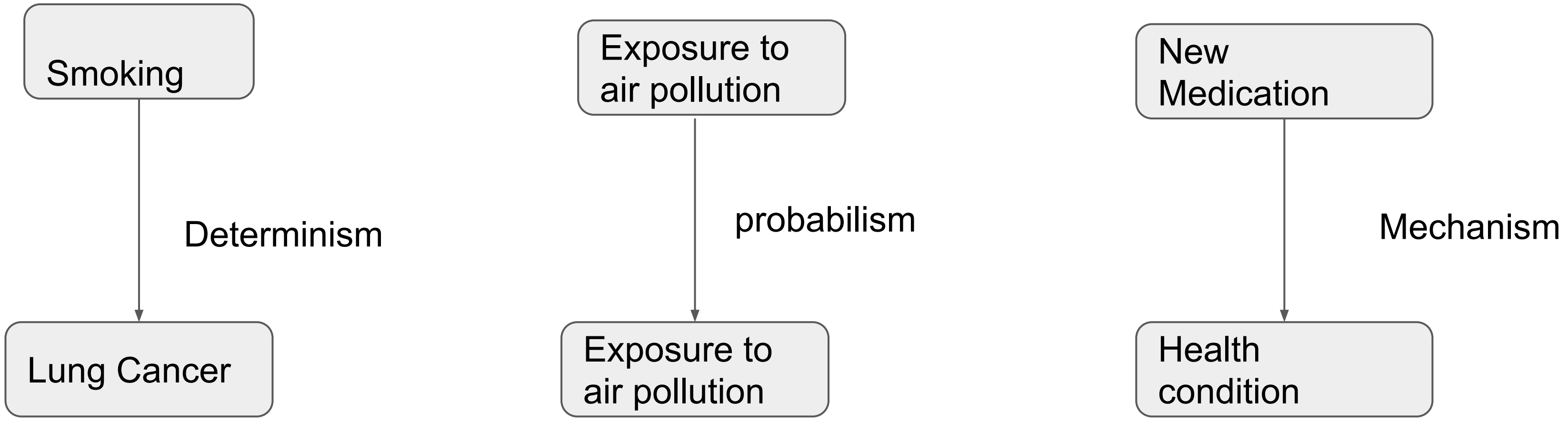
(**A**): Concepts fatigue, trembling relate to Anxiety. Concepts 'F' and Pregnant relates to Woman.

Cite: Lokala, Usha, et al. "A computational approach to understand mental health from reddit: knowledge-aware multitask learning framework." *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 16. 2022.

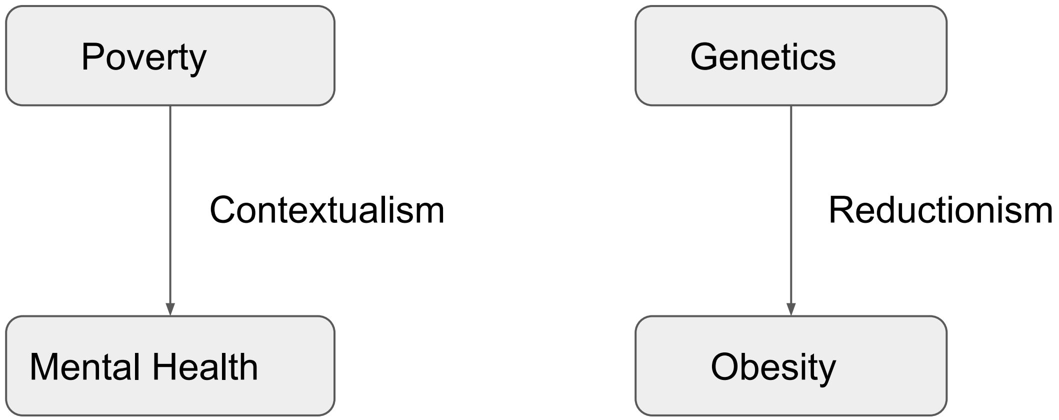
Formal inference Patterns - Causal Claims to Explanations



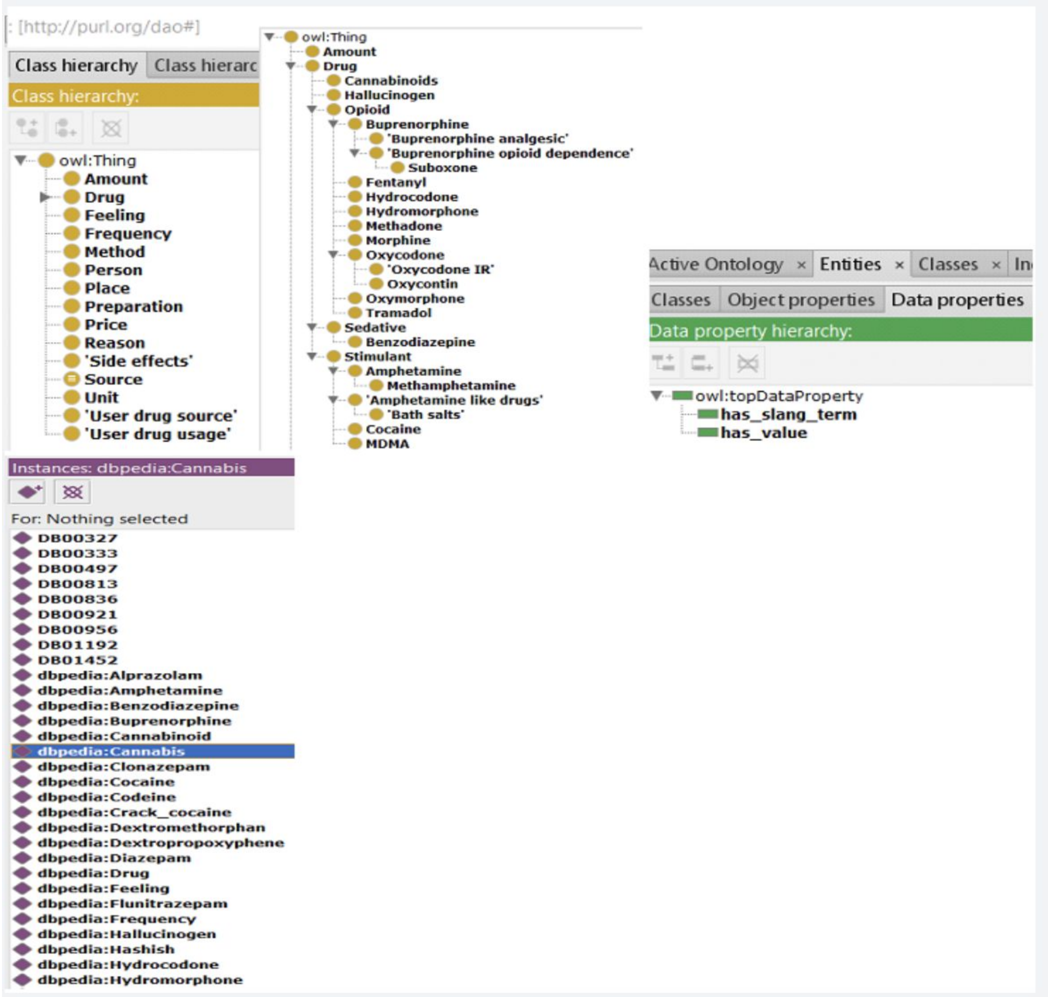
Ontological Principles - justifying Explanations/Outcome



"How different principles may be more appropriate depending on the specific context and nature of the phenomenon being studied?"



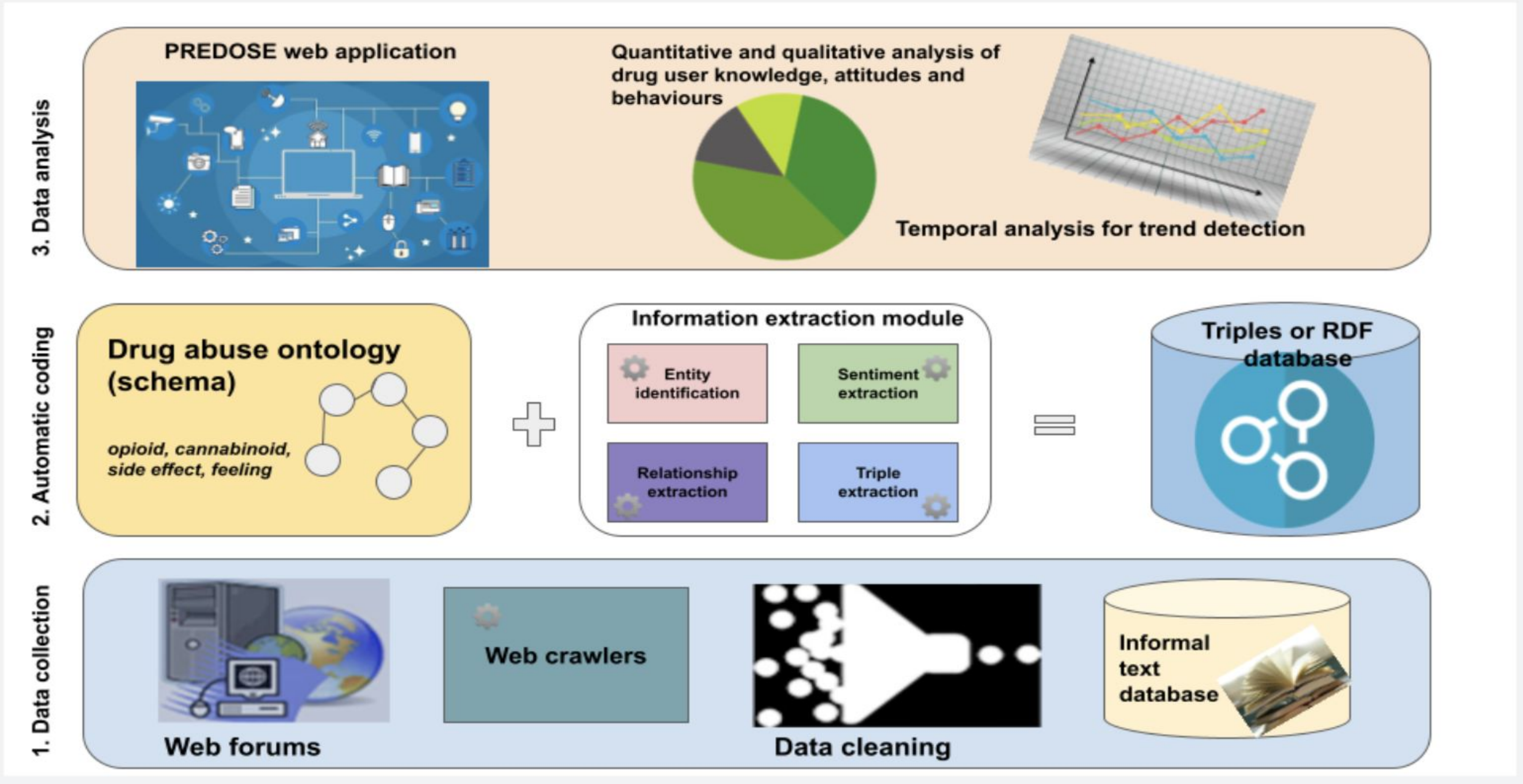
Example: Drug Abuse Ontology



Metric	Count, n	Description
Ontology metrics		
Axioms	4876	Combined logical and nonlogical axiom count
Logical axiom count	3478	The number of logical axioms
Declaration axioms count	1185	The number of declaration axioms
Class count	316	The number of distinct classes, object properties, data properties and individuals that are mentioned in the ontology
Object property count	12	The number of distinct classes, object properties, data properties and individuals that are mentioned in the ontology
Data property count	13	The number of distinct classes, object properties, data properties and individuals that are mentioned in the ontology
Individual count	845	The number of distinct classes, object properties, data properties and individuals that are mentioned in the ontology
Class axiom		
SubClassOf	313	The number of SubClassOf axioms in the ontology. A subclass axiom states that a class is a subclass of another class
Individual axioms		
Data property assertion	2317	A data property assertion states that the individual is connected by the data property expression to the literal.
ClassAssertion	830	A class assertion states that the individual is an instance of the class expression.
AnnotationAssertion	213	An annotation assertion states that the annotation subject is an anonymous individual with the annotation property and value.

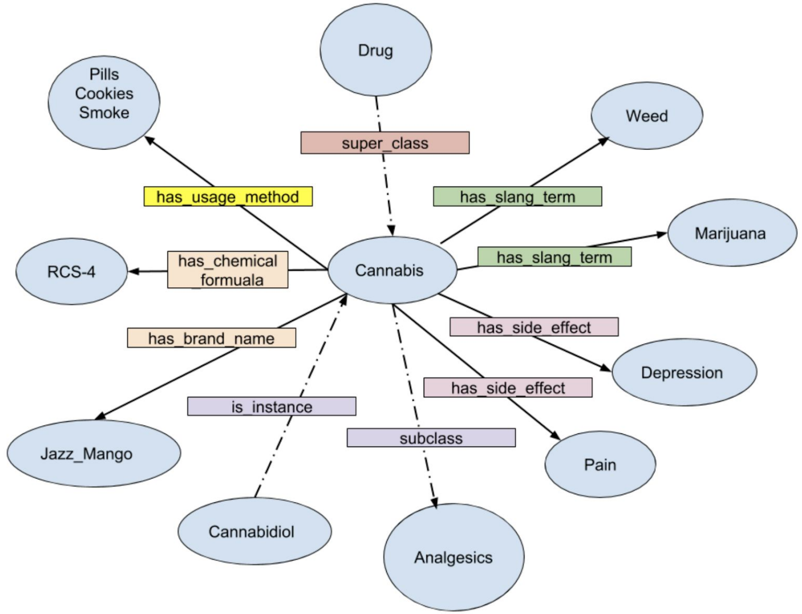
Cite: Lokala U, Lamy F, Daniulaityte R, Gaur M, Gyrard A, Thirunarayan K, Kursuncu U, Sheth A Drug Abuse Ontology to Harness Web-Based Data for Substance Use Epidemiology Research: Ontology Development Study JMIR Public Health Surveill 2022;8(12):e24938

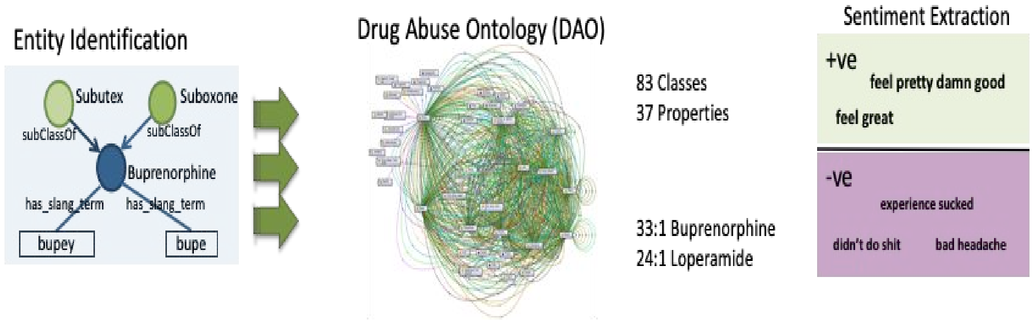
Ontology based Causal Inference - How one fact leads to another?



Cite: Lokala, Usha, et al. "Drug abuse ontology to harness web-based data for substance use epidemiology research: ontology development study." *JMIR public health and surveillance* 8.12 (2022): e24938.

Ontology based annotations for Causal Extraction - Drug Use Domain





I was sent home with 5 x 2 mg Suboxones. I also got a bunch of phenobarbital (I took all 180 mg and it didn't do shit except make me a walking zombie for 2 days). I waited 24 hours after my last 2 mg dose of Suboxone and tried injecting 4 mg of the bupe. It gave me a bad headache, for hours, and I almost vomited. I could feel the bupe working but overall the experience sucked.

Of course, junkie that I am, I decided to repeat the experiment. Today, after waiting 48 hours after my last bunk 4 mg injection, I injected 2 mg. There wasn't really any rush to speak of, but after 5 minutes I started to feel pretty damn good. So I injected another 1 mg. That was about half an hour ago. I feel great now.

Ontology as inference for analyzing web-based data and use the extracted wisdom to inform public health surveillance as insights and actionable items.

DIVERSE DATA TYPES

ENTITIES	
DOSAGE	PRONOUN
INTERVAL	Route of Admin.
RELATIONSHIPS	SENTIMENTS

Triples

Codes	Triples (subject-predicate-object)
Suboxone used by injection, negative experience	Suboxone Injection-causes-Cephalalgia
Suboxone used by injection, amount	Suboxone Injection-dosage amount-2mg
Suboxone used by injection, positive experience	Suboxone Injection-has_side_effect-Euphoria

»» Can ontologies be used as inference for causal explanations?

Cite: . Cameron D, Smith GA, Daniulaityte R, Sheth AP, Dave D, Chen L, et al. PREDOSE: a semantic web platform for drug abuse epidemiology using social media. J Biomed Inform 2013 Dec;46(6):985-997 [FREE Full text] [doi:10.1016/j.jbi.2013.07.007]

Cite: Yadav, Shweta, et al. "“When they say weed causes depression, but it’s your fav antidepressant”: knowledge-aware attention framework for relationship extraction." PloS one 16.3 (2021): e0248299.

Ontology based Causal Inference - How one fact discovers new fact?

PREDOSE: Loperamide-Withdrawal Discovery

Loperamide is used to self-medicate to from Opioid Withdrawal symptoms

"But I just wanted to tell you that loperamide WILL WORK. I take 105 mg of methadone/day, and recently have been running out early due to a renewed interest in IVing that shit. 200mg of lope 100 pills will make me almost 100 again. It brings the sickness down to the level of, say, a minor flu. Sleep returns, restlessness dissipates. Sometimes a mild opiation is felt."

"Back in the day when I would run out of pills early I would take 8-10 Lopermide tabs and get some pretty good relief from w/d."

"If you take a shitload of loperamide like 10-20 pills at once in withdrawal, you'll get relief from some of the physical symptoms. Im not sure exactly how it works, but it's definitely MORE than just relieving the GI symptoms. Im guessing if you just bombard your blood with it, SOME of it has to make it through? Not sure."

"Normally around 100 milligrams of loperamide will get me out of withdrawals."

"Loperamide alone is enough to keep me well without being miserable, IF I megadose."

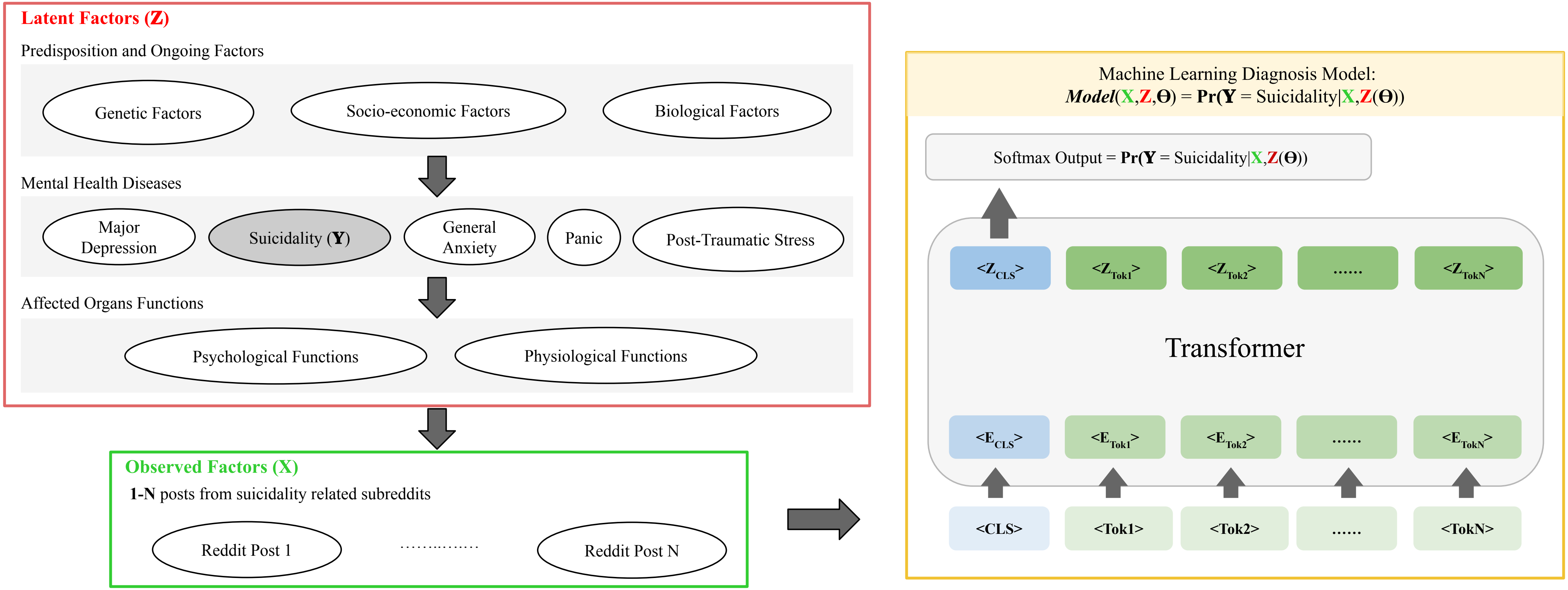
Imodium

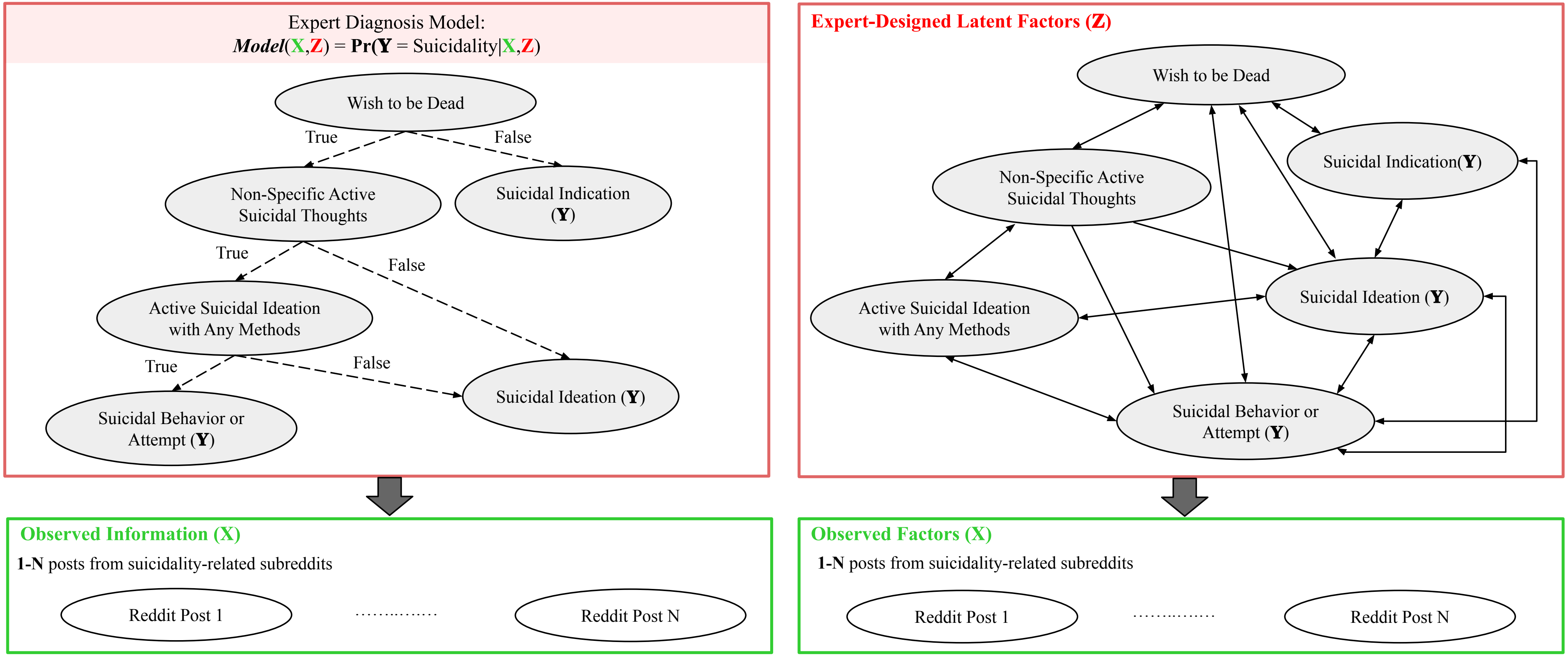
"This loperamide has saved my life during w/ds.... and made me even more careless with my monthly meds."

Cite: Daniulaityte R, Carlson RG, Falck RS, Cameron DH, Udayanaga S, Chen L, et al. A web-based study of self-treatment of opioid withdrawal symptoms with loperamide. Wright State University. 2012. URL: https://corescholar.libraries.wright.edu/ knoesis/624/

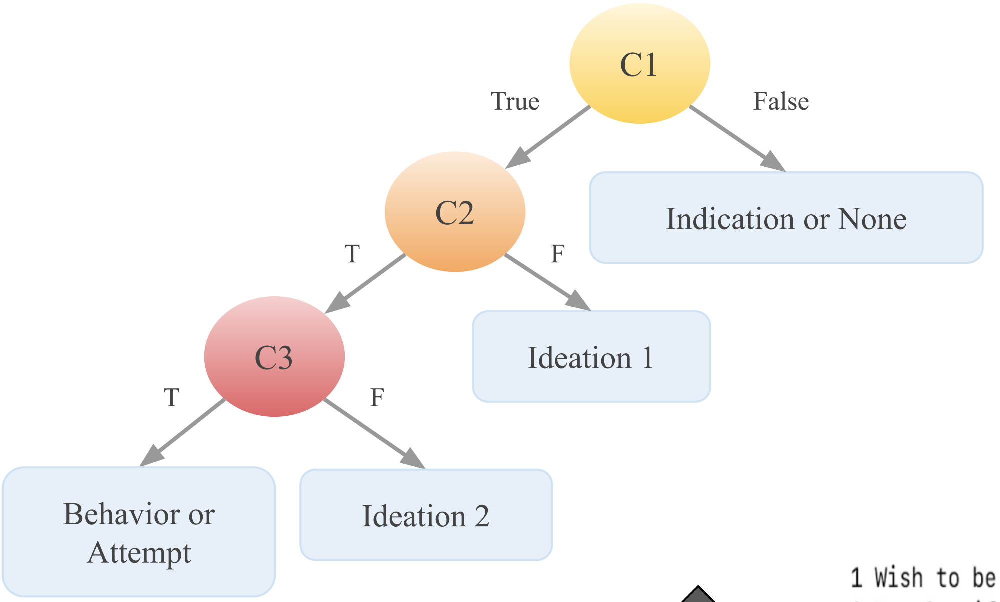
Applications in web and healthcare

Causal Process Knowledge Infused Reasoning For Mental Health Triaging


Latent Factors (Z)
Predisposition and Ongoing Factors
Genetic Factors
Socio-economic Factors
Biological Factors
Mental Health Diseases
Major Depression
Suicidality (Y)
General Anxiety
Panic
Post-Traumatic Stress
Affected Organs Functions
Psychological Functions
Physiological Functions
Observed Factors (X)
1-N posts from suicidality related subreddits
Reddit Post 1
Reddit Post N
Machine Learning Diagnosis Model:
Model(X,Z,Θ) = Pr(Y = Suicidality|X,Z(Θ))
Softmax Output = Pr(Y = Suicidality|X,Z(Θ))
<Z_CLS>
<Z_Tok1>
<Z_Tok2>
<Z_TokN>
Transformer
<E_CLS>
<E_Tok1>
<E_Tok2>
<E_TokN>
<CLS>
<Tok1>
<Tok2>
<TokN>


Expert Diagnosis Model:
Model(X,Z) = Pr(Y = Suicidality|X,Z)
Wish to be Dead
True
False
Non-Specific Active Suicidal Thoughts
Suicidal Indication (Y)
True
False
Active Suicidal Ideation with Any Methods
True
False
Suicidal Behavior or Attempt (Y)
Suicidal Ideation (Y)
Expert-Designed Latent Factors (Z)
Wish to be Dead
Suicidal Indication(Y)
Non-Specific Active Suicidal Thoughts
Active Suicidal Ideation with Any Methods
Suicidal Ideation (Y)
Suicidal Behavior or Attempt (Y)
Observed Information (X)
1-N posts from suicidality-related subreddits
Reddit Post 1
Reddit Post N
Observed Factors (X)
1-N posts from suicidality-related subreddits
Reddit Post 1
Reddit Post N

Tree Paths Representing Different Suicidality Contexts



Columbia Suicide Severity Rating Scale (CSSRS) Concepts

C1: Wish to be Dead

C2: Non-Specific Active Suicidal Thoughts

C3: Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act

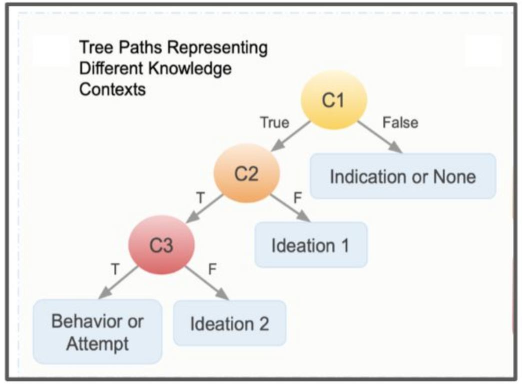
```
1 Wish to be Dead
2 Non-Specific Active Suicidal Thoughts
3 Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
(pytorch) kau[redacted]ok-Pro-2 PKiL % python3 infer.py
100%|████████████████████████████████████████| 3/3 [00:02<00:00,  1.48it/s]
========================================
You entered:  i have been there and got nothing same as my life.i have a gun.im not
on a ledge or something but i have a gun on my lap.
========================================
color coding scheme
123
========================================
 i have been there and got nothing same as my life
 i have a gun
 im not on a ledge or something but i have a gun on my lap

========================================
Inference: Behavior or Attempt
```

Explanation for Inference

```
========================================
You entered:  i have been there and got nothing same as my life.i have a gun.im not
on a ledge or something but i have a gun on my lap.
========================================
```

Context: Concept 1 (T), Concept 2 (T), Concept 3 (T) →
Behavior or Attempt

Application in Web & Health- Suicidality Context Identification



1. Learn the parameters that maximize the likelihood
2. Much fewer parameters to learn and Strongly convex objective
3. Easily Explained to the End-User

$$P(y \mid x, \{\theta_i\}) = \sum_{y \in Outcomes} p_y \prod_{i=1}^{3} \sum_{x_{sub} \in x} \left(cos\_sim\left(x^R_{sub}, q^R_i\right) \geq \theta_i \right) \geq 0.5$$

$$\mathcal{L}(x, y, \{\theta_i\}) = P(y \mid x, \{\theta_i\}) log(P(y \mid x, \{\theta_i\})) + (1 - P(y \mid x, \{\theta_i\})) log(1 - P(y \mid x, \{\theta_i\}))$$

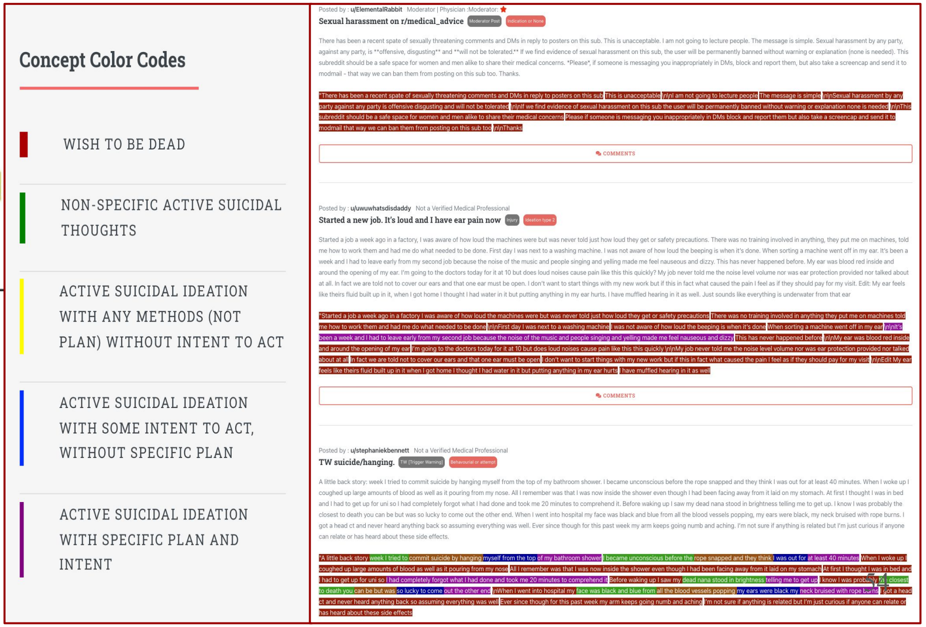
```
1 Wish to be Dead
2 Non-Specific Active Suicidal Thoughts
3 Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
(pytorch) kaushikroy@MacBook-Pro-2 PKiL % python3 infer.py
100%|████████████████████████████████████████| 3/3 [00:02<00:00,  1.48it/s]
=========================================
You entered:  i have been there and got nothing same as my life.i have a gun.im not
on a ledge or something but i have a gun on my lap.
=========================================
color coding scheme
123
=========================================
 i have been there and got nothing same as my life
 i have a gun
 im not on a ledge or something but i have a gun on my lap
=========================================
Inference: Behavior or Attempt
```

Context: Concept 1 (T), Concept 2 (T), Concept 3 (T) → ***Behavior or Attempt***

Applications in Other Domains

- Cooking/Nutrition Management
- Autonomous Driving
- Game Playing

CPR - Reddit Web Plug-in



Future Work:

1. Efficient Re-discovery of Causal Models
2. Efficient learning methods to modify parameters
3. Enhancing explanation visuals for the end-user

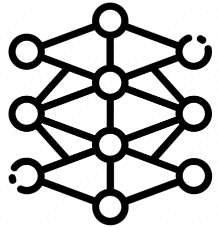
https://mentalhealthcpr.herokuapp.com scan QR code

Explainability

Causal Explainability

- Current AI approaches rely on statistical correlations that are often spurious and can’t be explained.
- Causality in AI systems using knowledge based approach can assist in better explainability, and providing support for intervention and counterfactuals, leading to improved understanding of AI systems by humans.

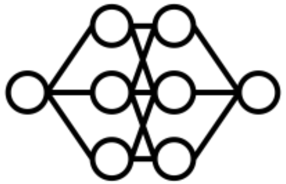
Really struggling with my **bisexuality** which is causing **chaos in my relationship** with a girl. Being a fan of LGBTQ community, I am equal to **worthless** for her. I'm now starting to get drunk because I can't cope with the obsessive, intrusive thoughts**,** and need to get it out of my head.



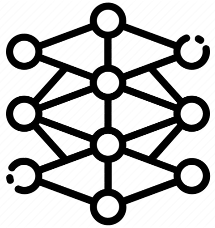
Reasoning over Model:

Why model predicted SI?
Unknown

Suicidal Ideation ? Yes: 0.71 , No: 0.29



Really struggling with my **bisexuality** which is causing **chaos in my relationship** with a girl. Being a fan of LGBTQ community, I am equal to **worthless** for her. I'm now starting to get drunk because I can't cope with the **obsessive, intrusive thoughts,** and need to get it out of my head.

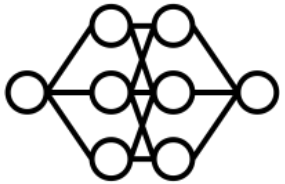


Reasoning over Model:

Why model predicted Depression?
Unknown

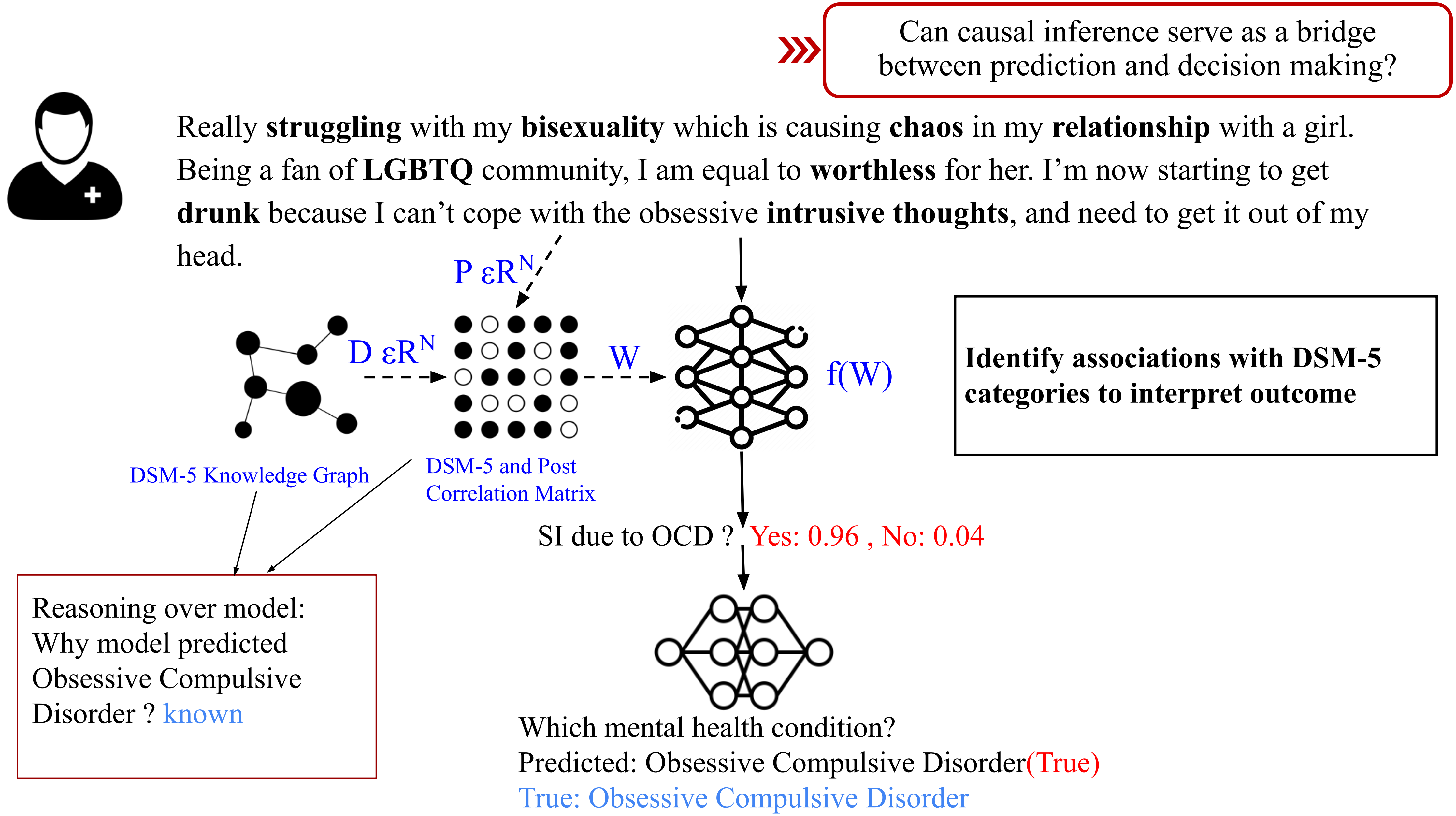
Is mental health related ? Yes: 0.71 , No: 0.29

Highlighted terms based on attention matrix

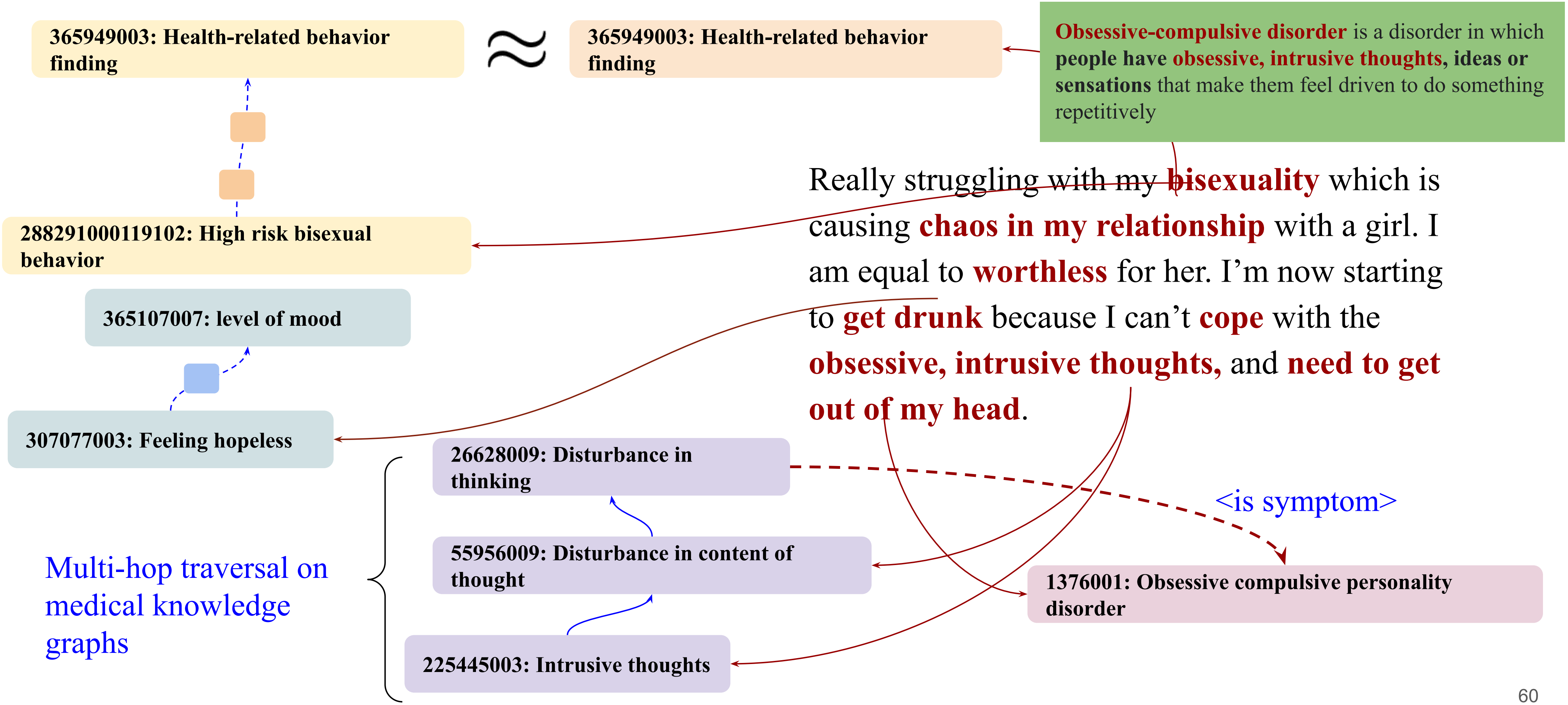


Which mental health condition?
Predicted: Depression (False)
True: Obsessive Compulsive Disorder

Manas Gaur, Ugur Kursuncu, Amanuel Alambo, Amit Sheth, Raminta Daniulaityte, Krishnaprasad Thirunarayan, and Jyotishman Pathak. "" Let Me Tell You About Your Mental Health!" Contextualized Classification of Reddit Posts to DSM-5 for Web-based Intervention." CIKM 2018. [CIKM 2018]

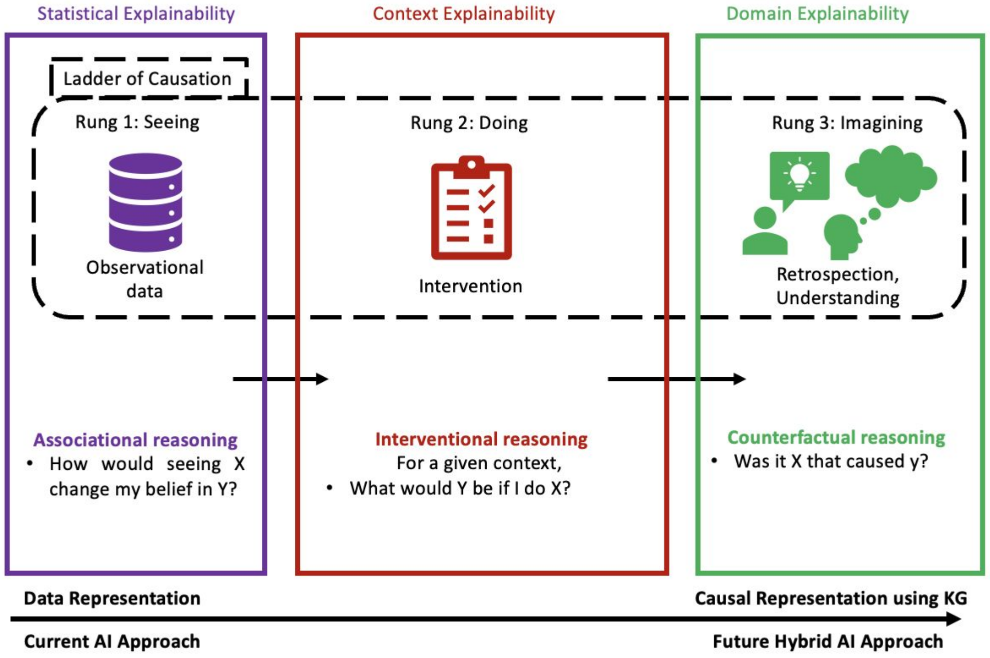

Can causal inference serve as a bridge between prediction and decision making?
Really **struggling** with my **bisexuality** which is causing **chaos** in my **relationship** with a girl. Being a fan of **LGBTQ** community, I am equal to **worthless** for her. I'm now starting to get **drunk** because I can't cope with the obsessive **intrusive thoughts**, and need to get it out of my head.
P εR^N
D εR^N
W
f(W)
Identify associations with DSM-5 categories to interpret outcome
DSM-5 Knowledge Graph
DSM-5 and Post Correlation Matrix
SI due to OCD ? Yes: 0.96 , No: 0.04
Reasoning over model: Why model predicted Obsessive Compulsive Disorder ? known
Which mental health condition?
Predicted: Obsessive Compulsive Disorder(True)
True: Obsessive Compulsive Disorder

Explain interpretation through linking to KG and definitions



Causal Explainability

- ***Statistical explainability*** generates an explanation for the co-occurrence of a given phenomenon based on the statistical (or associational) methods such as correlations.
- ***Context explainability*** is a means to generate a human-understandable explanation taking the context information of a given observation into account.
- ***Domain explainability*** explains the underlying causal relations using observational data, domain knowledge, and counterfactual reasoning.



Jaimini, Utkarshani, and Amit Sheth. "CausalKG: Causal Knowledge Graph Explainability using interventional and counterfactual reasoning." *IEEE Internet Computing* 26.1 (2022): 43-50.

Takeaways

- Statistical AI alone is not enough
- Causal AI can be used for generating explanations
- Causal AI can be used for complex causal pattern extraction
- Causal AI can be used for causal entity associations
- Causal AI can be used for intervention planning using web data in the application area of epidemiology, mental health, as examples

Thank You
Open to questions

ChatGPT and Causality

why do people not believe that the cock crowing caused the sun to rise, even though the cock crows before the sun rises?

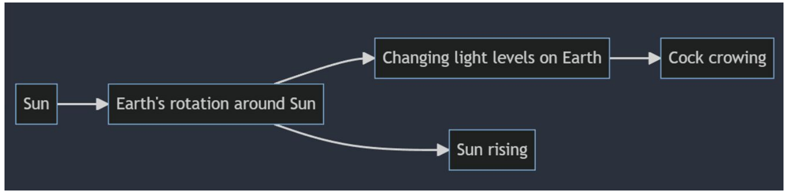
People do not believe that the cock crowing caused the sun to rise because the relationship between the two events is not causative, but rather temporal. The sun rising and the cock crowing are simply correlated, but the rising of the sun is a result of the Earth's rotation and its position relative to the sun, not the crowing of the cock. The sun would still rise even if there were no crowing roosters.

Sure! Here is a causal diagram to explain the relationship between the Earth's rotation around the sun, the cock crowing, and the sun rising using Mermaid syntax:

lua Copy code

```lua
graph LR
  Sun[Sun] --> Earth_rotation[Earth's rotation around Sun]
  Earth_rotation --> Light_levels[Changing light levels on Earth]
  Light_levels --> Cock_crowing[Cock crowing]
  Earth_rotation --> Sun_rising[Sun rising]
```

This diagram shows that the sun rising is a result of the Earth's rotation around the sun, which in turn causes the changing light levels on the Earth. The cock crowing is a response to the changing light levels, but it does not cause the sun to rise. The sun rises as a result of the Earth's rotation around the sun, regardless of whether or not the cock is crowing.



- Sun does not *cause* the earth's rotation
- Earth's rotation is a result of the residual momentum of the material from which our solar system is formed
- Earth revolution around the sun is due to the gravitational pull between them

https://www.causalmap.app/post/chatgpt-causal-of-course

Potential Outcome Framework (Rubin-Neyman Causal Model)

- Each unit (individual) x_i have two potential outcomes:
  - $Y_0(x_i)$ is the potential outcome had the unit not been treated: "*Control outcome*"
  - $Y_1(x_i)$ is the potential outcome had the unit been treated: "*Treated outcome*"

- Conditional average treatment effect (also known as individual treatment effect) for unit *i*:

$$CATE(x_i) = E_{Y_1 \sim p(Y_1|x_i)}[Y_1|x_i] - E_{Y_0 \sim p(Y_0|xi)}[Y_0|x_i]$$

- Average treatment effect:

$$ATE = E[Y_1 - Y_0] = E_{x \sim p(x)}[CATE(x)]$$

$$= E_{x \sim p(x)}[E[Y_1|x, T = 1] - E[Y_0|x, T = 0]]$$

Potential Outcome Framework (Rubin-Neyman Causal Model)

- Each unit (individual) x_i have two potential outcomes:
  - $Y_0(x_i)$ is the potential outcome had the unit not been treated: "*Control outcome*"
  - $Y_1(x_i)$ is the potential outcome had the unit been treated: "*Treated outcome*"

- Observed factual outcome:

$$y_i = t_i Y_1(X_i) + (1 - t_i) Y_0(X_i)$$

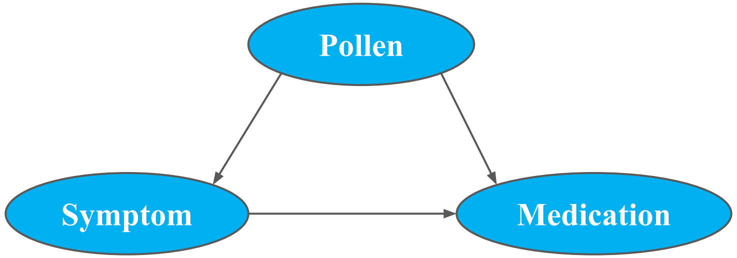
- Unobserved counterfactual outcome:

$$y^{CF}{}_i = (1 - t_i) Y_1(X_i) + t_i Y_0(Xi)$$

Fundamental problem of Causal Inference- *We only observe one of the two outcomes*

Ladder of Causation: Counterfactual

Given a patient has taken the medication, what is the chance they would not take it if the pollen was not present in the outdoor environment?



P(Medication = N | Medication =Y, do(Symptom = N))

Natural Direct Effect