



The Knowledge Graph Conference

# Causal Knowledge Graph (KG)- Infused Prescriptive Analytics

Prepared for a Presentation at 2022 Knowledge Graph Conference, May 6<sup>th</sup>

**Dr. Victor Z. Chen, CFA**

Director of Data Analytics and Insights, Experimental Design and Causal Inference

**Fidelity Investments**

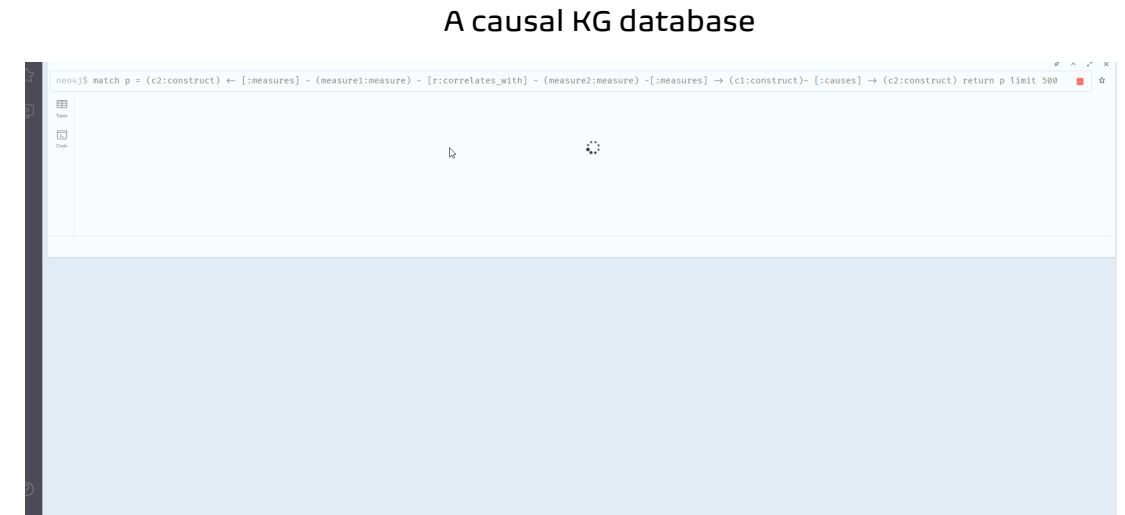
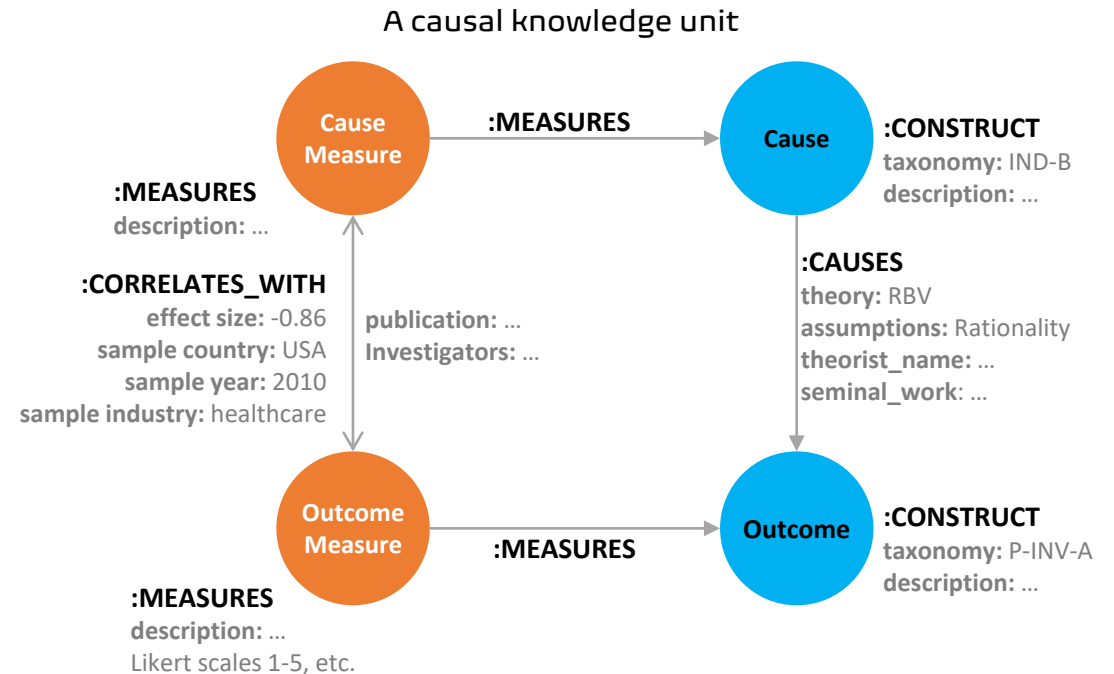
# Agenda

- 1. What is causal knowledge graph (KG)?
- 2. How to construct a causal KG?
- 3. How to infuse a causal KG into data analytics?

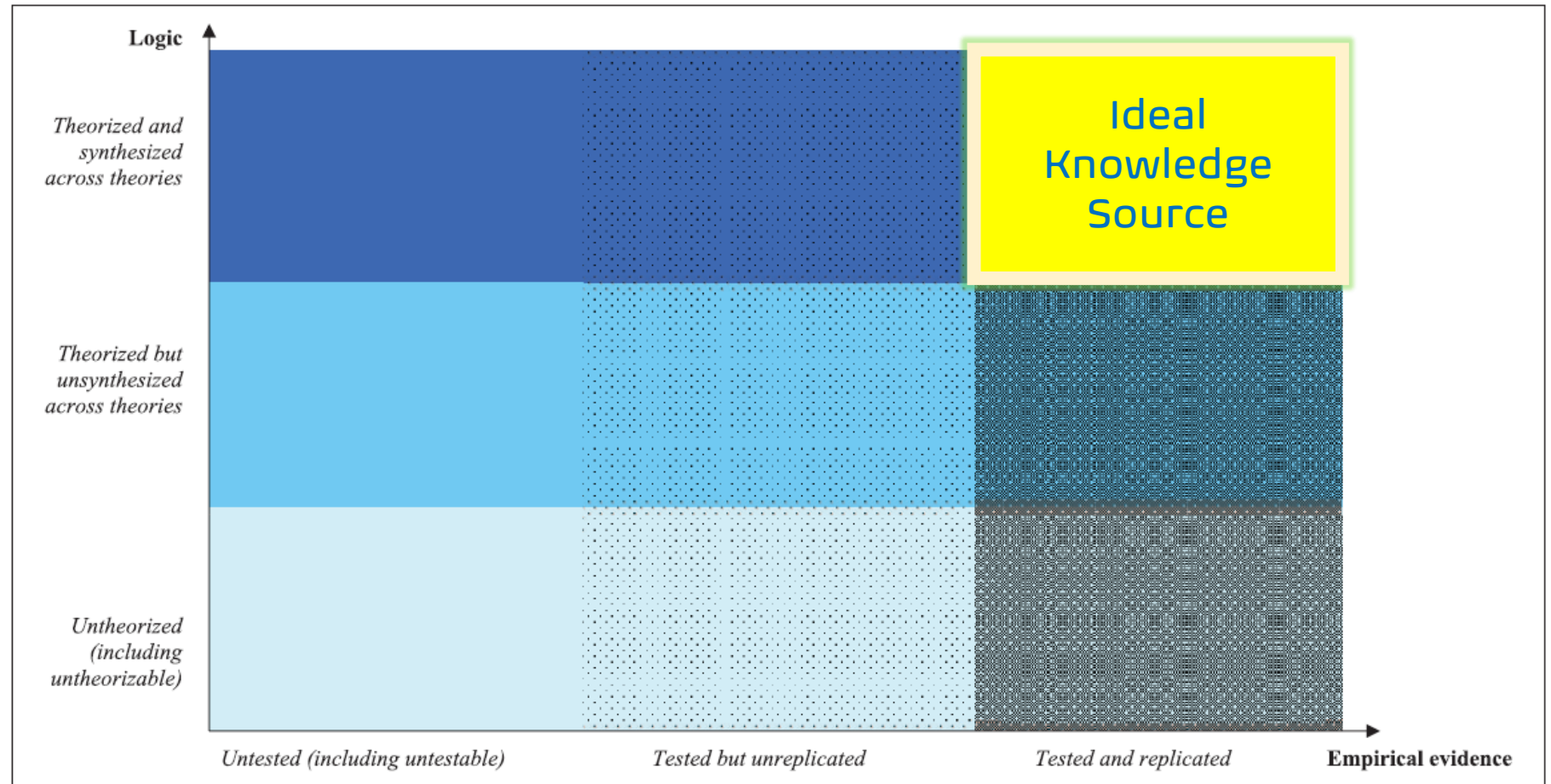
# 1. What is a Causal KG?

A unified and cumulative representation of knowledge

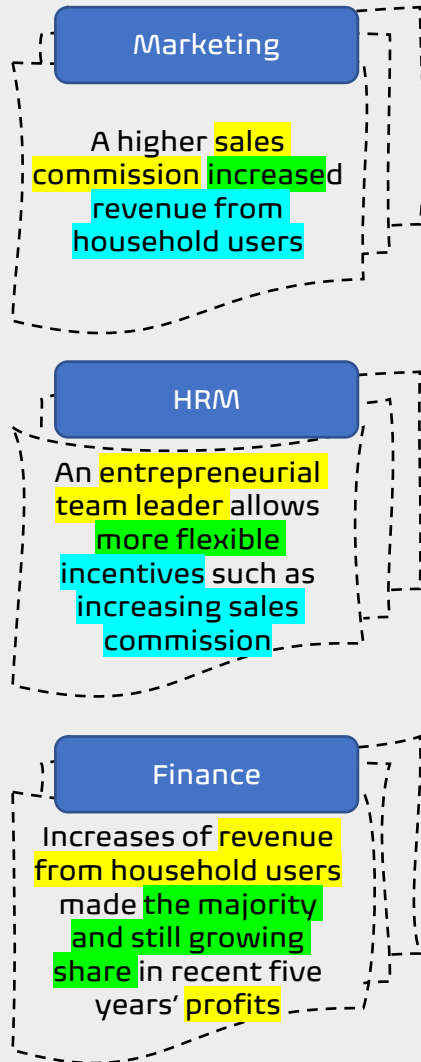
- Entities (e.g., constructs and variables)
- Directed Links:
  - Ontological relationships (e.g., taxonomy, semantics)
  - Causal logic (e.g., common sense, theoretical hypothesis, expert assumptions)
  - Prior empirical evidence (e.g., experiments and Bayesian meta-analysis)



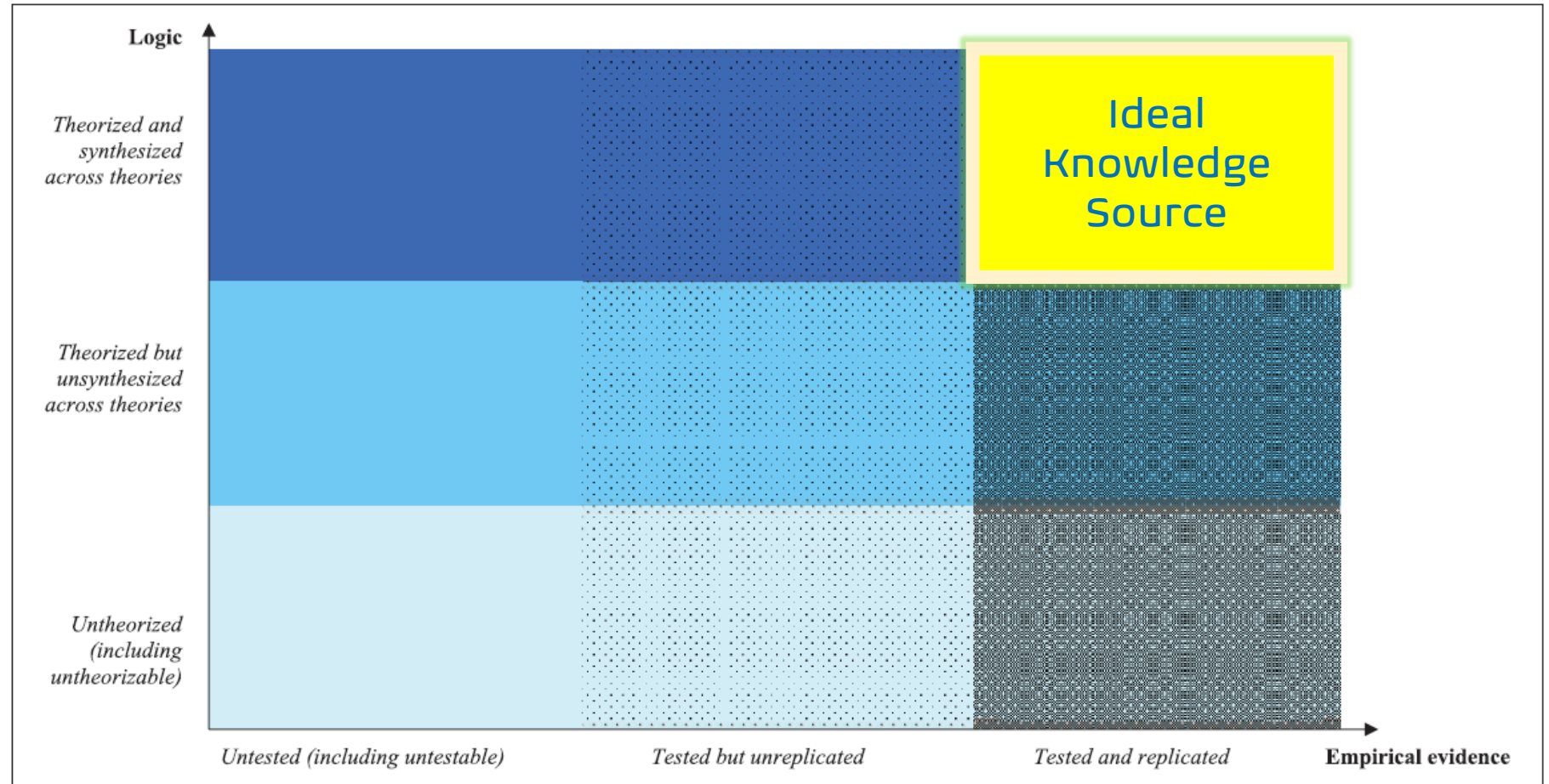
## 2. How to construct a Causal KG?



Source: Chen and Hitt (2021)



## 2. How to construct a Causal KG?



Source: Chen and Hitt (2021)

## Sourcing

### Marketing

A higher sales commission increased revenue from household users

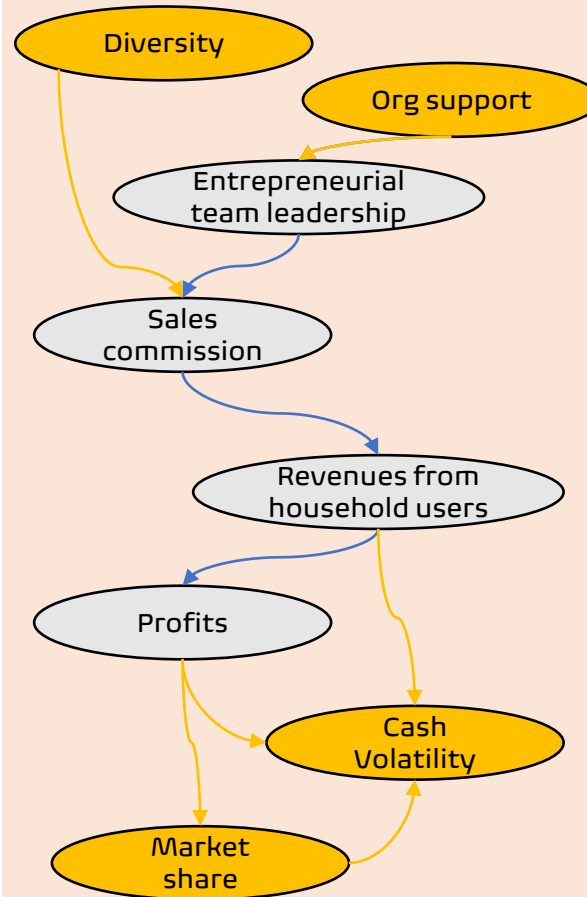
### HRM

An entrepreneurial team leader allows more flexible incentives such as increasing sales commission

### Finance

Increases of revenue from household users made the majority and still growing share in recent five years' profits

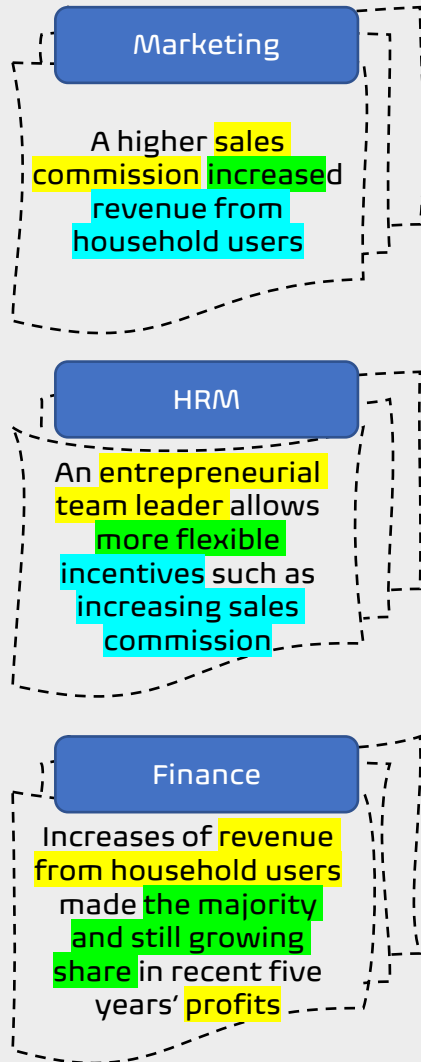
## Graphing



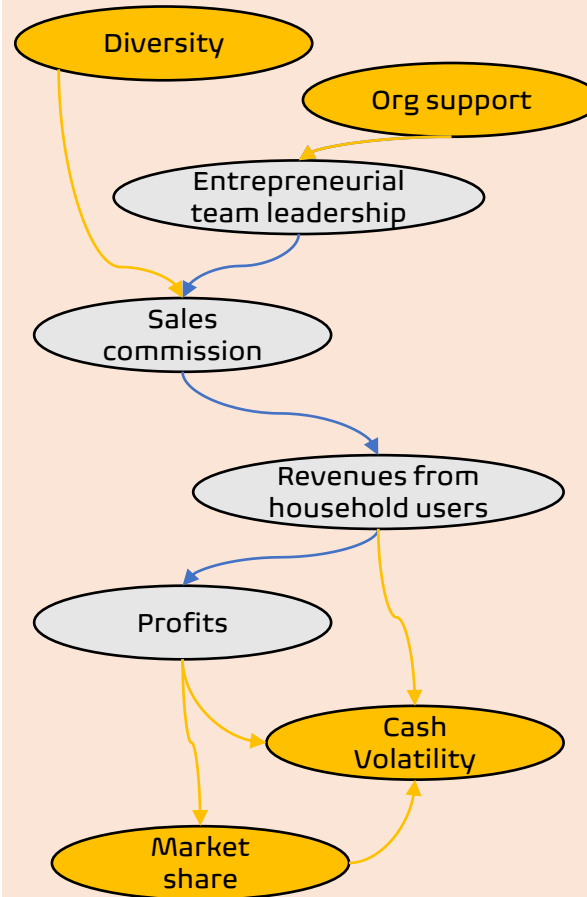
## Synthesis

## Analysis

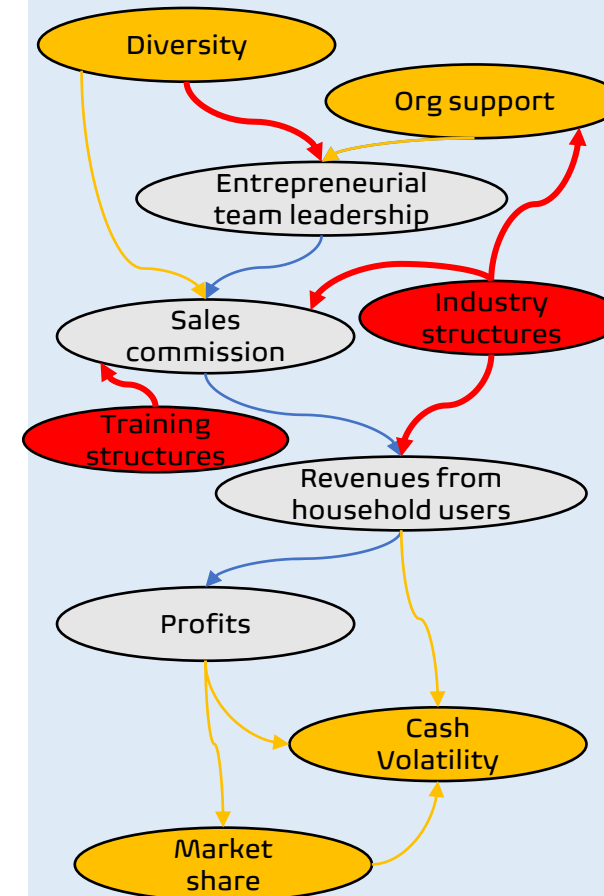
## Sourcing



## Graphing



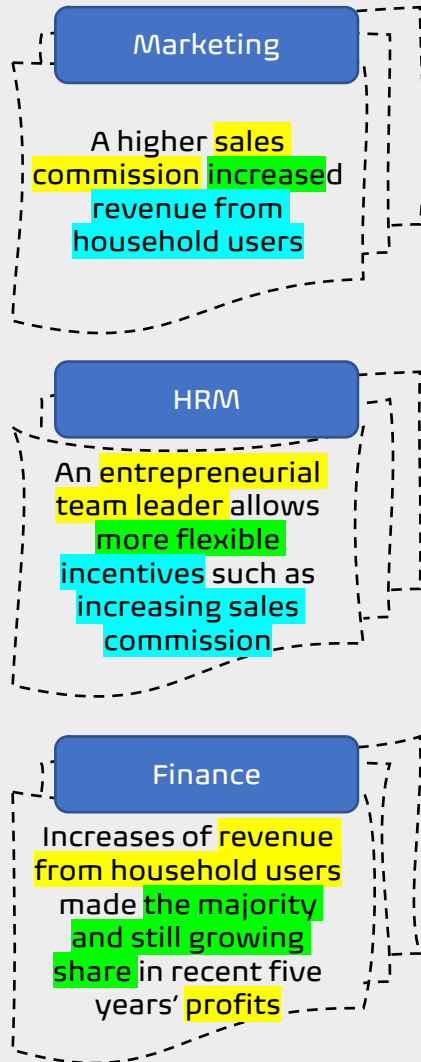
## Synthesis



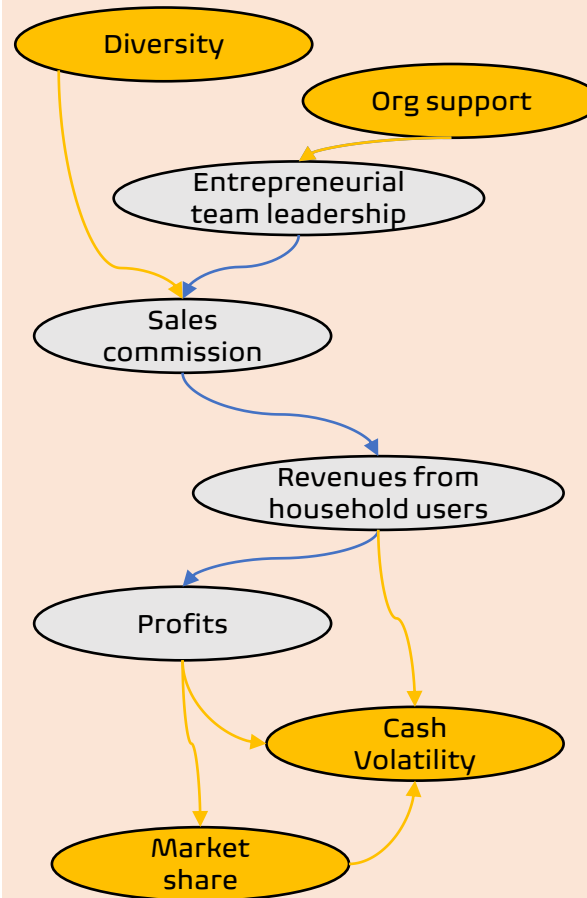
## Analysis



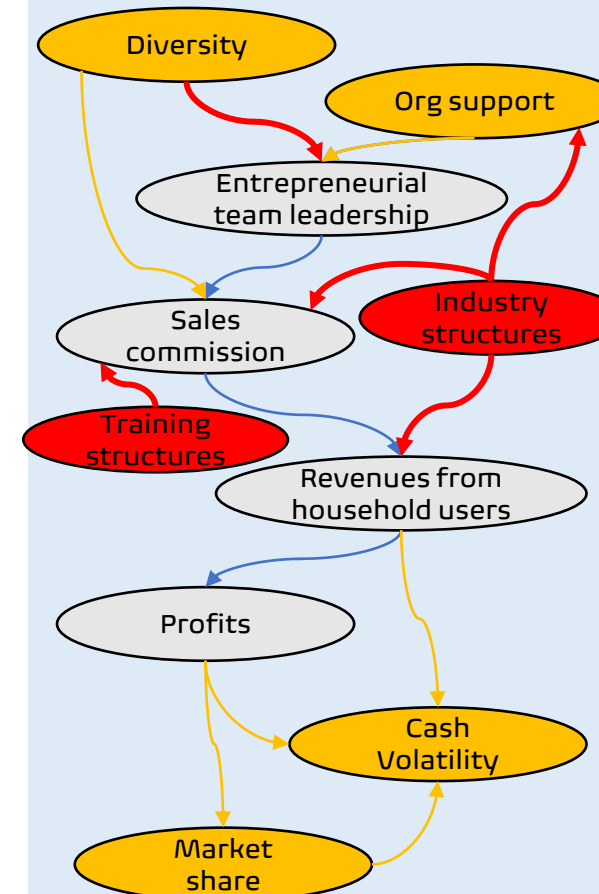
## Sourcing



## Graphing

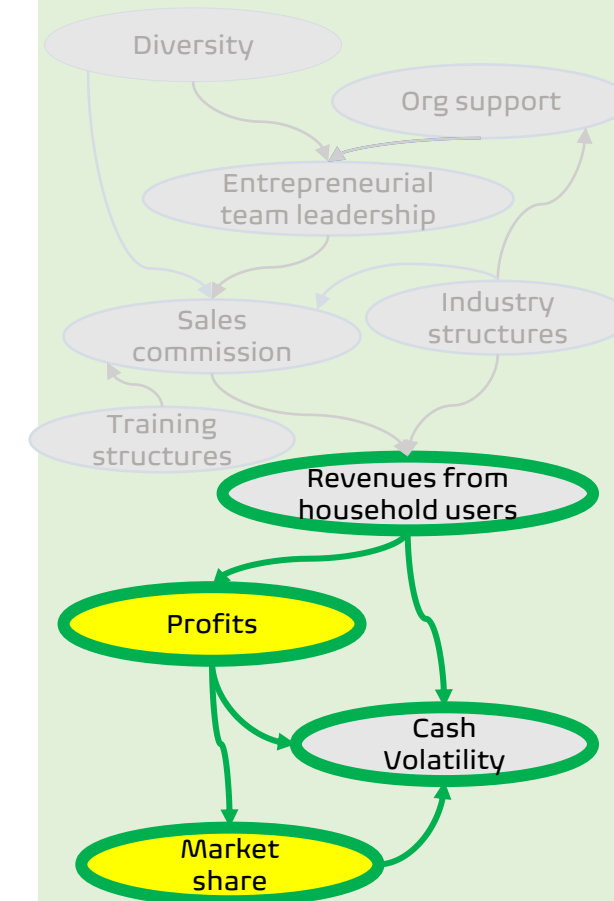


## Synthesis



## Analysis

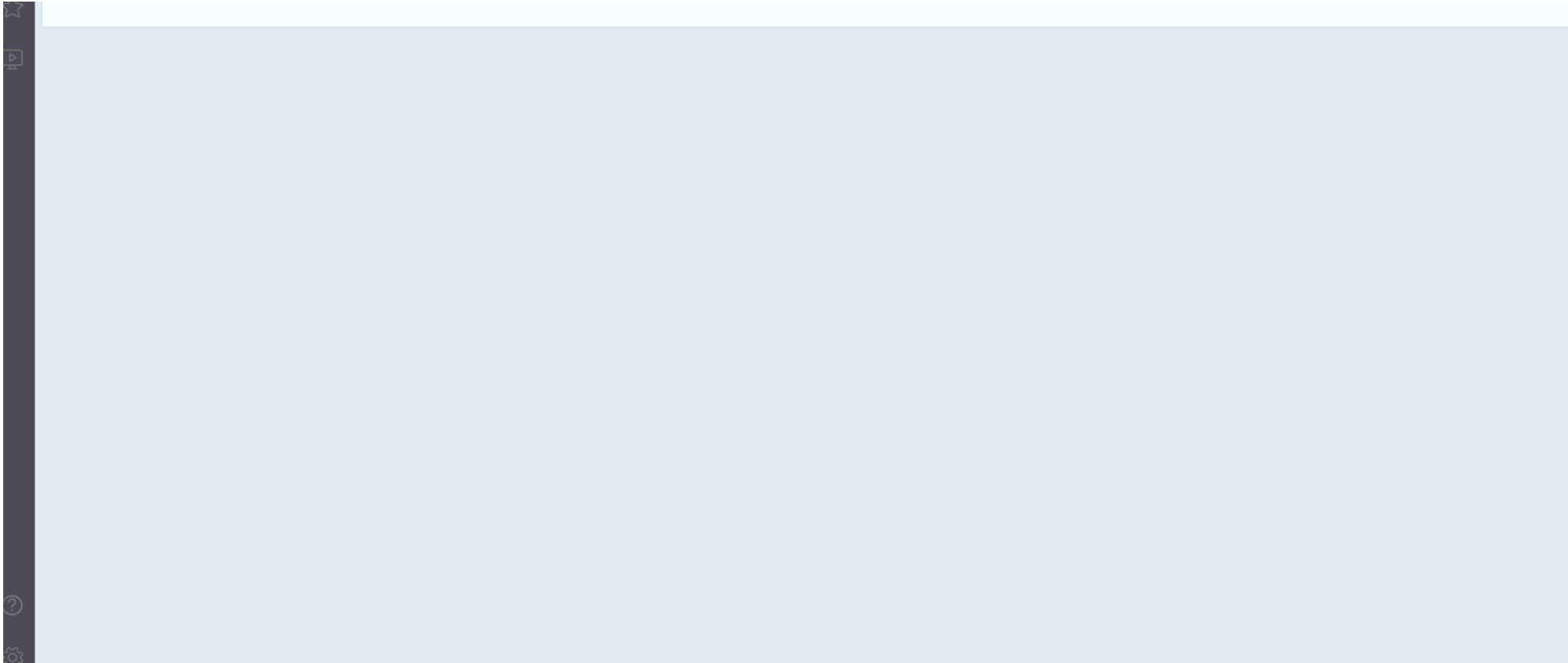
### Causal Graph-based Data Science





## 2. How to construct a Causal KG?

Step 1. Constructing and loading a unified taxonomy



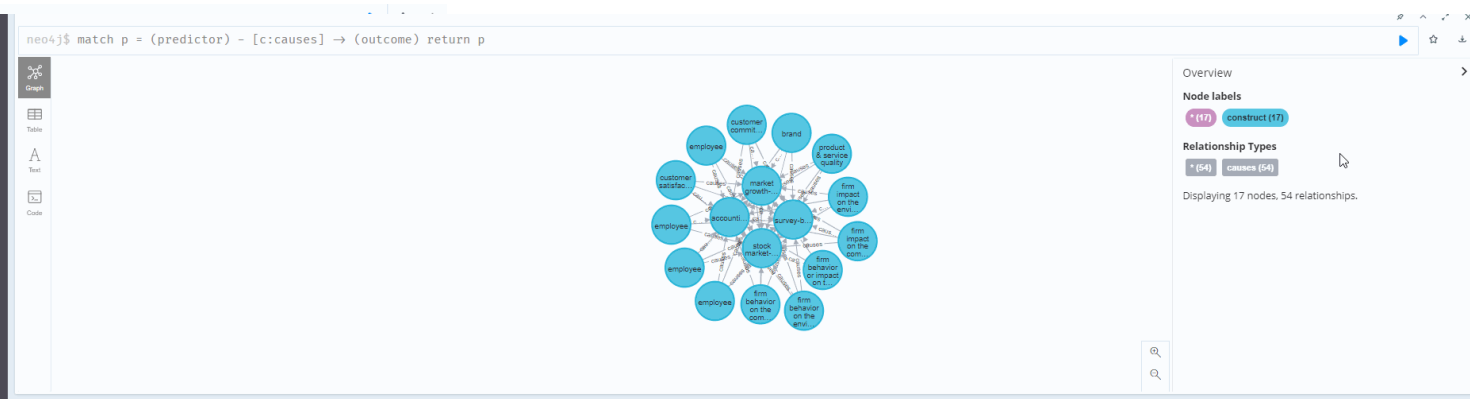
## 2. How to construct a Causal KG?

### Step 2. Adding theoretical/causal assertions

```
neo4j$ match p = (predictor) - [c:causes] -> (outcome) return predictor.name, outcome.name, c.theory
```

	predictor.name	outcome.name
1	"firm behavior on the community"	"stock market-based performance"
2	"firm behavior on the community"	"survey-based financial performance"
3	"firm behavior on the community"	"market growth-based performance"
4	"firm behavior on the community"	"accounting-based performance"
5	"firm behavior on the environment"	"accounting-based performance"
6	"firm behavior on the environment"	"market growth-based performance"

Started streaming 54 records in less than 1 ms and completed after 3 ms.



## 2. How to construct a Causal KG?

### Step 3. Linking constructs with measures/variables

```
neo4j$ match p = (measure:measure) -[:measures] - (construct:construct) return construct.name, measure.name, measure.description
```

	construct.name	measure.name	measure.description
1	"firm behavior on the community"	"social performance"	"We conduct one of only very few studies of the antecedents of corporate reputation drawing on a multivariate data set that describes the financial performance, social performance, size, R&D and advertising intensity, and industry of sample companies."
2	"firm behavior on the community"	"soccdis"	"Level of corporate social responsibility disclosure for firm i in period t: 0 = poor, 1 = good, 2 = excellent"
3	"firm behavior on the community"	"csr"	"We used the 13-item scale developed by Singhapakdi, Vitell, Rallapalli, and Kraft (1996) to measure the perceptions of top management toward CSR (ranging from 1 = totally disagree, to 5 = totally agree; sample items include 'Social responsibility and profitability can be compatible,' 'Being ethical and socially responsible is the most important thing a firm can do,' and 'Business has a social responsibility to the community'")"
4	"firm behavior on the community"	"csr"	"We define CSR as those business activities that take into account the interests of society"
5	"firm behavior on the community"	"government-aimed csr"	"On operational terms, global retailers donate cash or provide services to the community"
6	"firm behavior on the community"	"community responsibilities"	"Give money to charities in the communities where we operate"

Started streaming 1116 records after 1 ms and completed after 2 ms, displaying first 1000 rows.

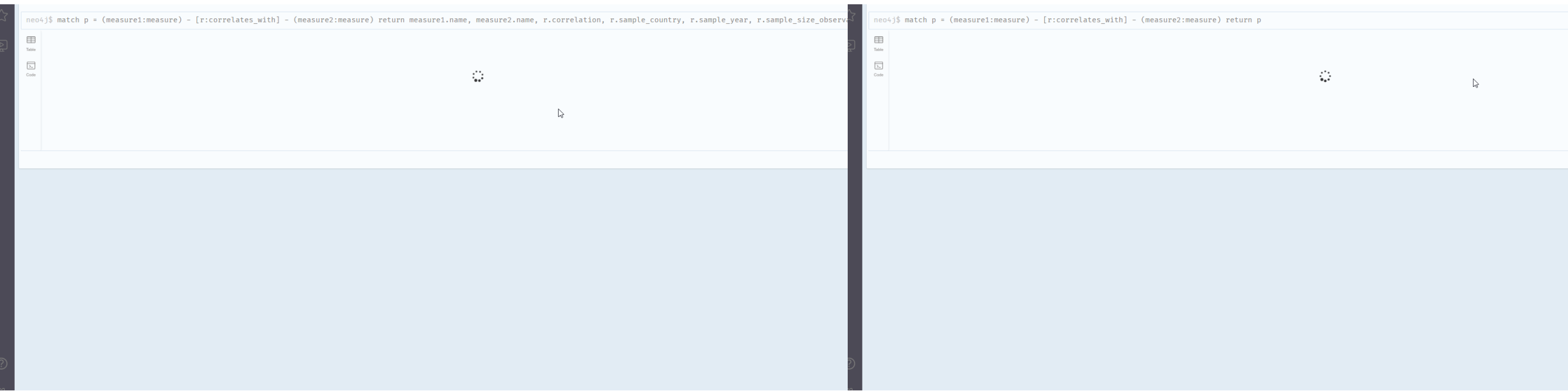
```
neo4j$ match p = (measure:measure) -[:measures] - (construct:construct) return p
```

"p"
[{"name": "social performance", "description": "We conduct one of only very few studies of the antecedents of corporate reputation drawing on a multivariate data set that describes the financial performance, social performance, size, R&D and advertising intensity, and industry of sample companies.", "name_raw": "Social performance", "level1": "firm behavior on the community", "description": "the actions a firm has taken or reported on the community (with impact unknown)", "level10": "community benefits", "level0": "organizational performance"}]
[{"name": "soccdis", "description": "Level of corporate social responsibility disclosure for firm i in period t: 0 = poor, 1 = good, 2 = excellent", "name_raw": "SOCODIS", "level1": "firm behavior on the community", "description": "the actions a firm has taken or reported on the community (with impact unknown)", "level10": "community benefits", "level0": "organizational performance"}]
[{"name": "csr", "description": "We used the 13-item scale developed by Singhapakdi, Vitell, Rallapalli, and Kraft (1996) to measure the perceptions of top management toward CSR (ranging from 1 = totally disagree, to 5 = totally agree; sample items include 'Social responsibility and profitability can be compatible,' 'Being ethical and socially responsible is the most important thing a firm can do,' and 'Business has a social responsibility to the community'")", "name_raw": "CSR", "level1": "firm behavior on the community", "description": "the actions a firm has taken or reported on the community (with impact unknown)", "level10": "community benefits", "level0": "organizational performance"}]

MAX COLUMN WIDTH: 100%

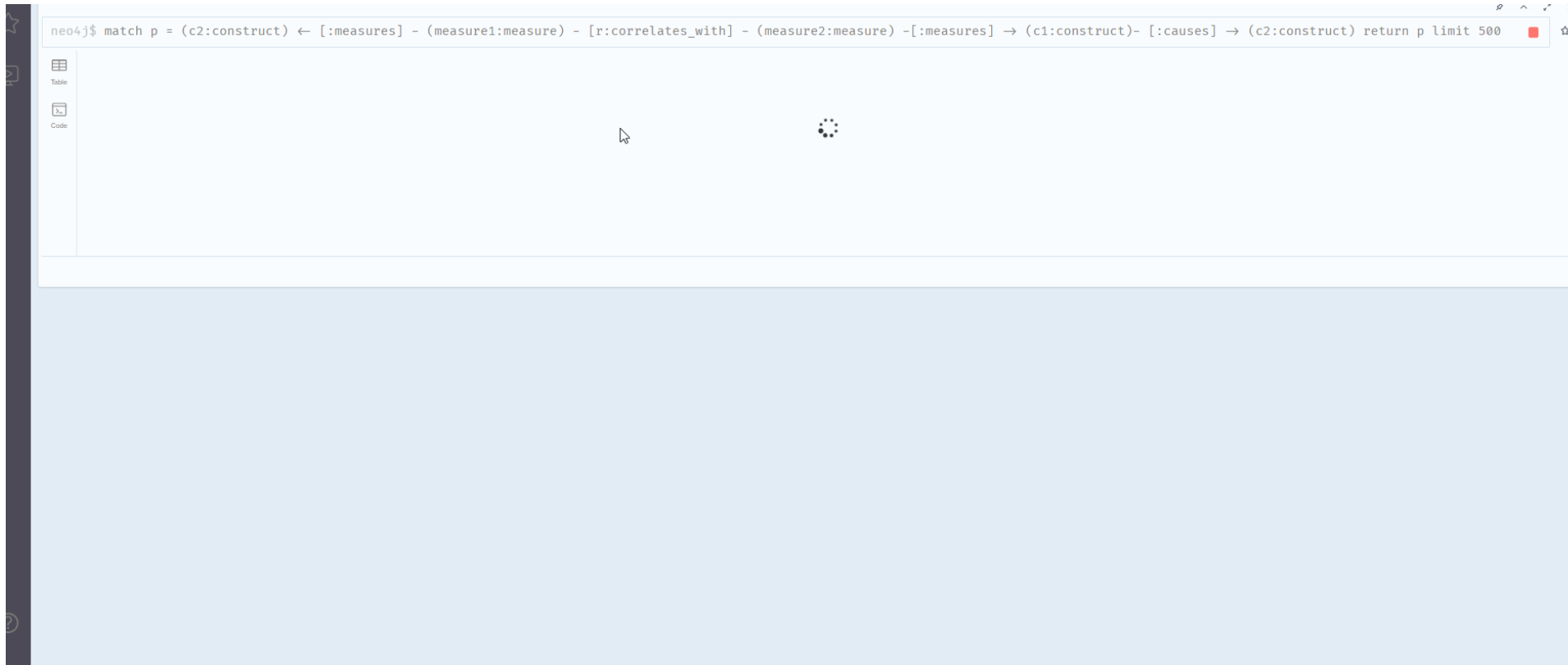
## 2. How to construct a Causal KG?

### Step 4. Adding statistical evidence and sample meta data

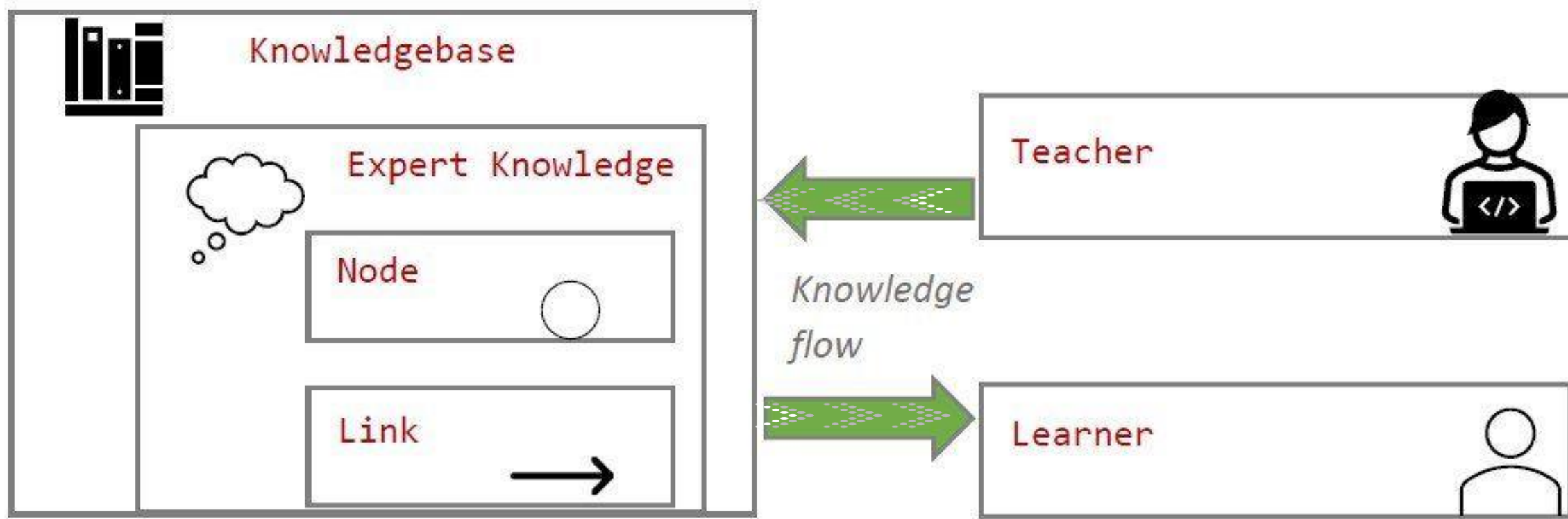


## 2. How to construct a Causal KG?

Here you go! A complete causal knowledge graph

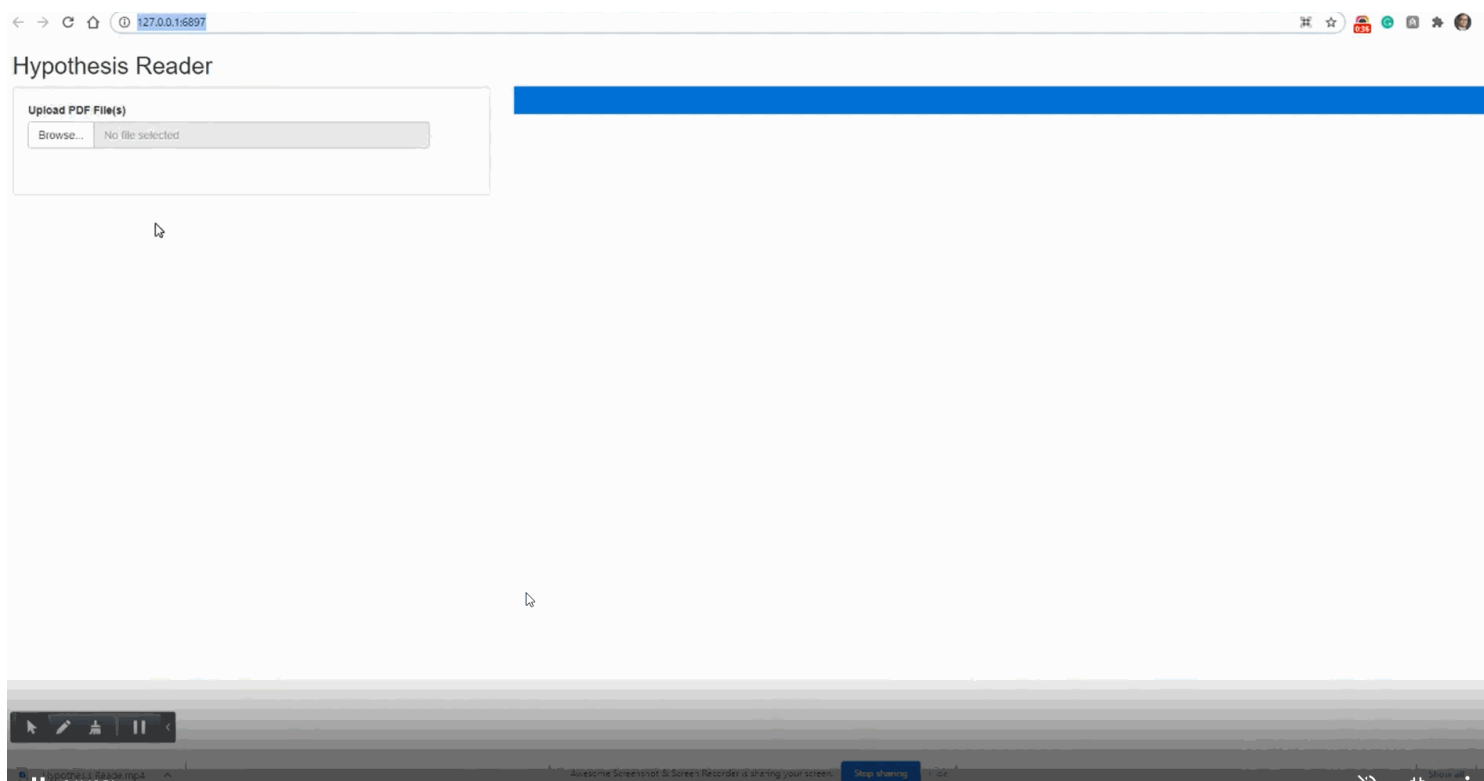


## 2. How to construct a Causal KG? An Expert Community Approach



## 2. How to construct a Causal KG? A Machine Learning Approach

### 1. Machine Reading of Hypotheses



arXiv.org > cs > arXiv:2106.16102 Search...

[Help](#) | [Advanced](#)

**Computer Science > Information Retrieval**

*[Submitted on 30 Jun 2021 (v1), last revised 11 Jul 2021 (this version, v2)]*

**Machine Reading of Hypotheses for Organizational Research Reviews and Pre-trained Models via R Shiny App for Non-Programmers**

Victor Zitian Chen, Felipe Montano-Campos, Wlodek Zadrozny, Evan Canfield



## 2. How to construct a Causal KG? A Machine Learning Approach

### 2. Entity Classification

	Model	P	R	F1
Engineered features	XGBoost	0.74	0.71	0.72
	Multinomial naïve Bayes	0.80	0.81	0.80
	SVM	0.81	0.81	0.81
	Logistic regression	0.81	0.81	0.81
Latent features	FastText	0.84	0.79	0.81
	Fine-tuned BERT	0.59	0.51	0.40
	SpaCy embeddings with LR	0.74	0.75	0.75
	Elmo embeddings with LR	0.78	0.78	0.78

arXiv.org > cs > arXiv:2107.05133

Search...  
Help | Advanced Search

Computer Science > Computation and Language

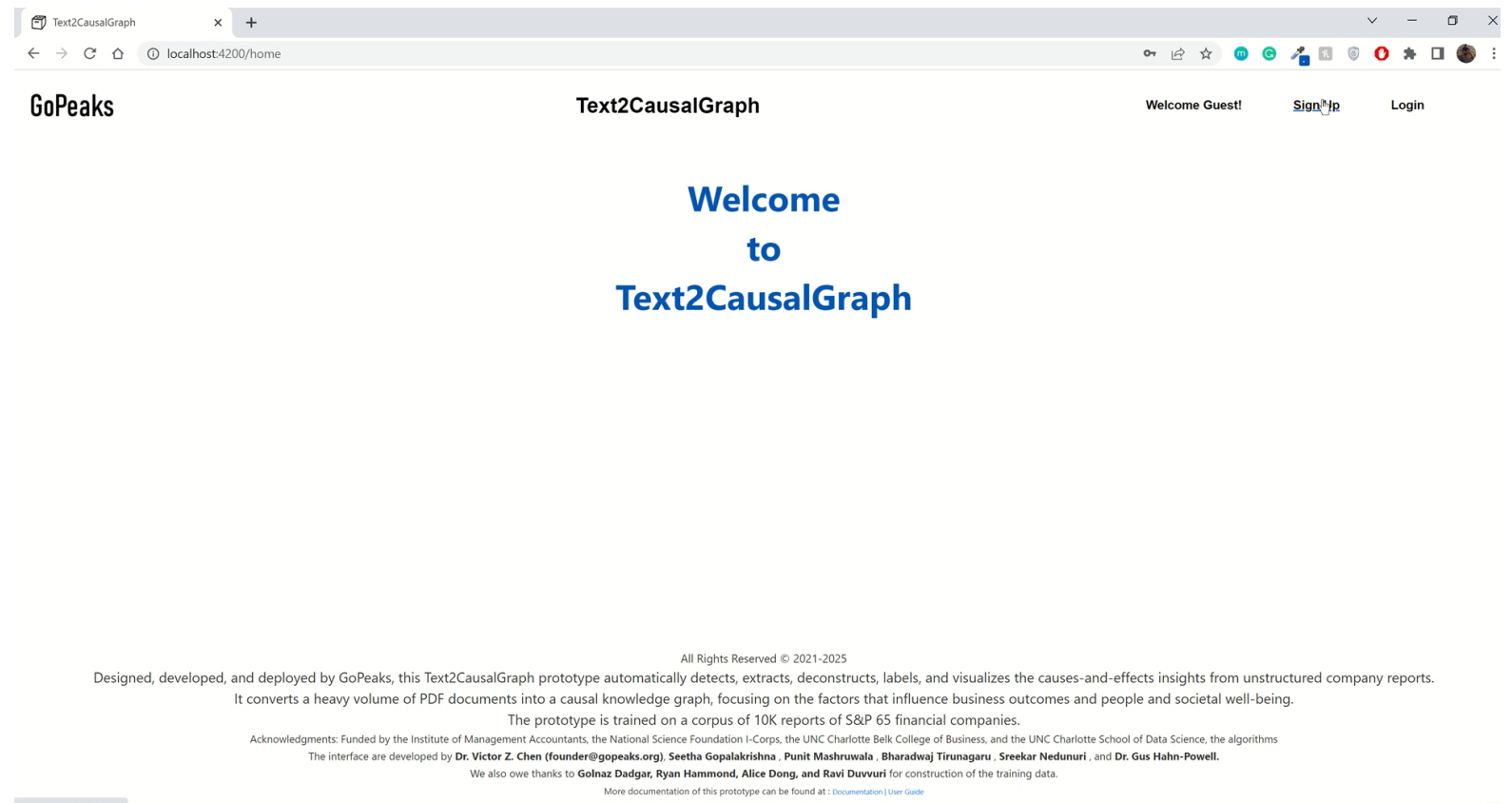
[Submitted on 11 Jul 2021 (v1), last revised 24 Aug 2021 (this version, v2)]

**Computer-assisted construct classification of organizational performance concerning different stakeholder groups**

Seethalakshmi Gopalakrishnan, Victor Chen, Gus Hahn-Powell, Bharadwaj Tirunagar

## 2. How to construct a Causal KG? A Machine Learning Approach

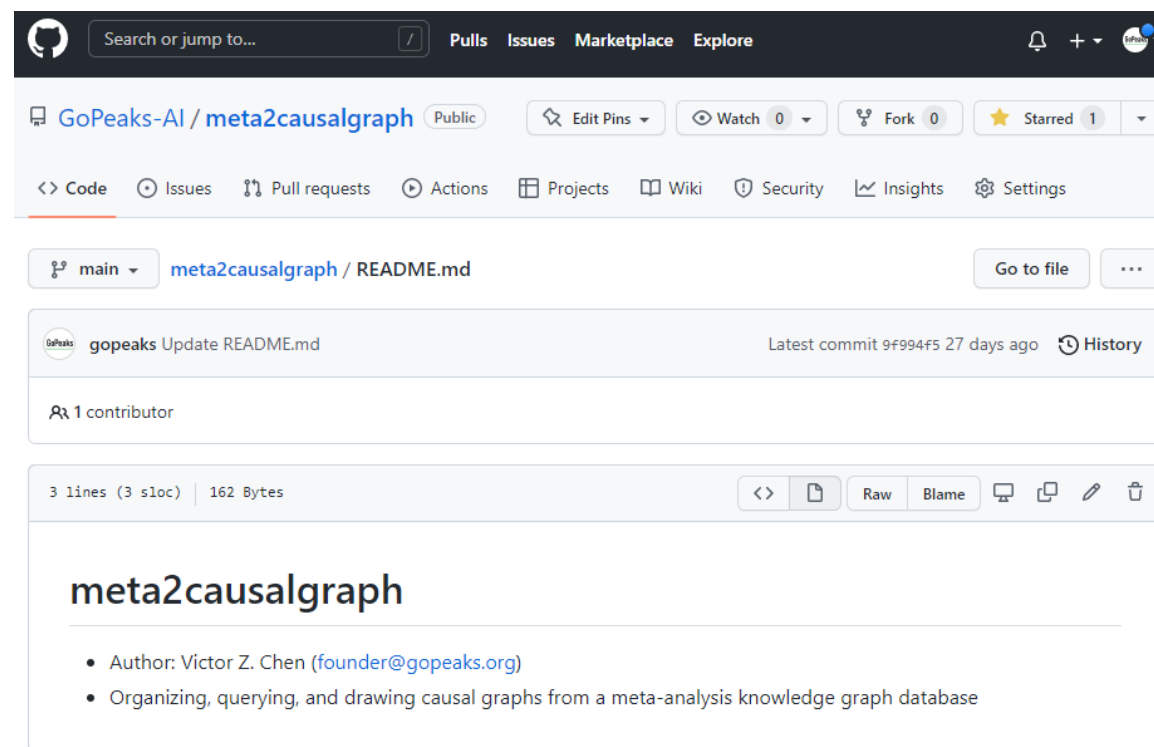
### 3. Converting Text into KG Database



## 2. How to construct a Causal KG?

A demonstration using Python and Neo4j

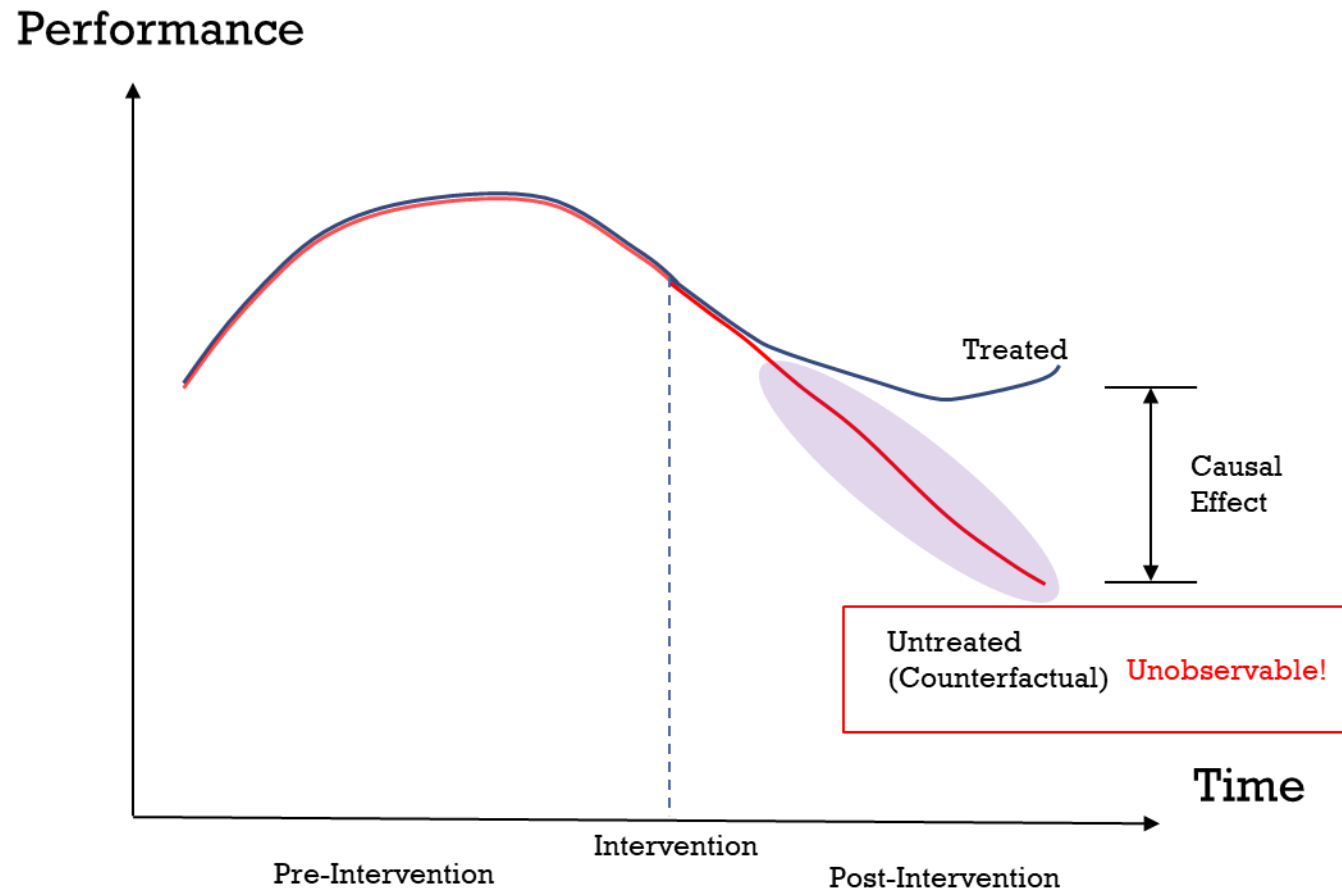
<https://github.com/GoPeaks-AI/meta2causalgraph>



### 3. How to infuse Causal KG into prescriptive analytics?

The essence of causal analysis:

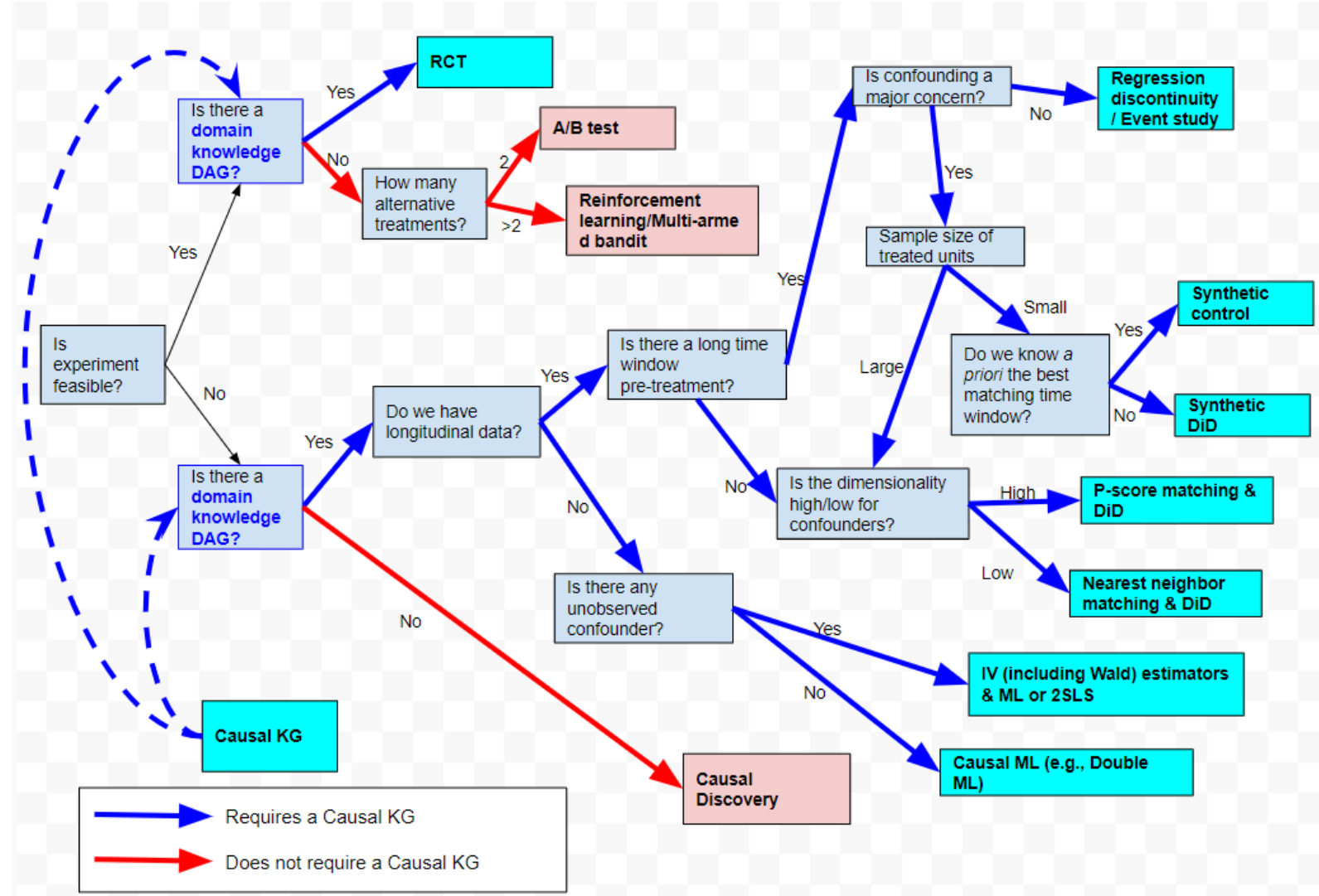
- Self-counterfactual comparison
  - Balancing: Similar distribution in covariates
  - Debiasing: Removing the confounding



### 3. How to infuse Causal KG into prescriptive analytics?

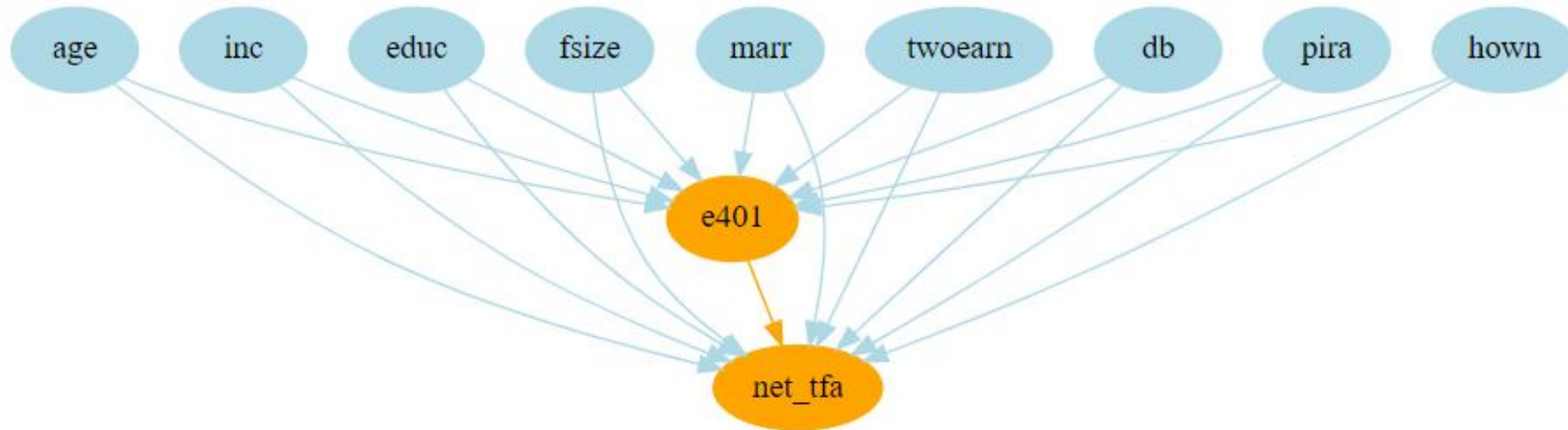
Causal KG may enable and improve many causal analytics methods by:

- Identifying what covariates need to be balanced
- Identifying what confounding biases there are in observational data



### 3. How to infuse Causal KG into prescriptive analytics?

A demonstration using Jupyter Notebook and Neo4j



Want to learn more?



GoPeaks.org



[github.com/GoPeaks-AI](https://github.com/GoPeaks-AI)

**towards**  
data science

@victorzitianchen