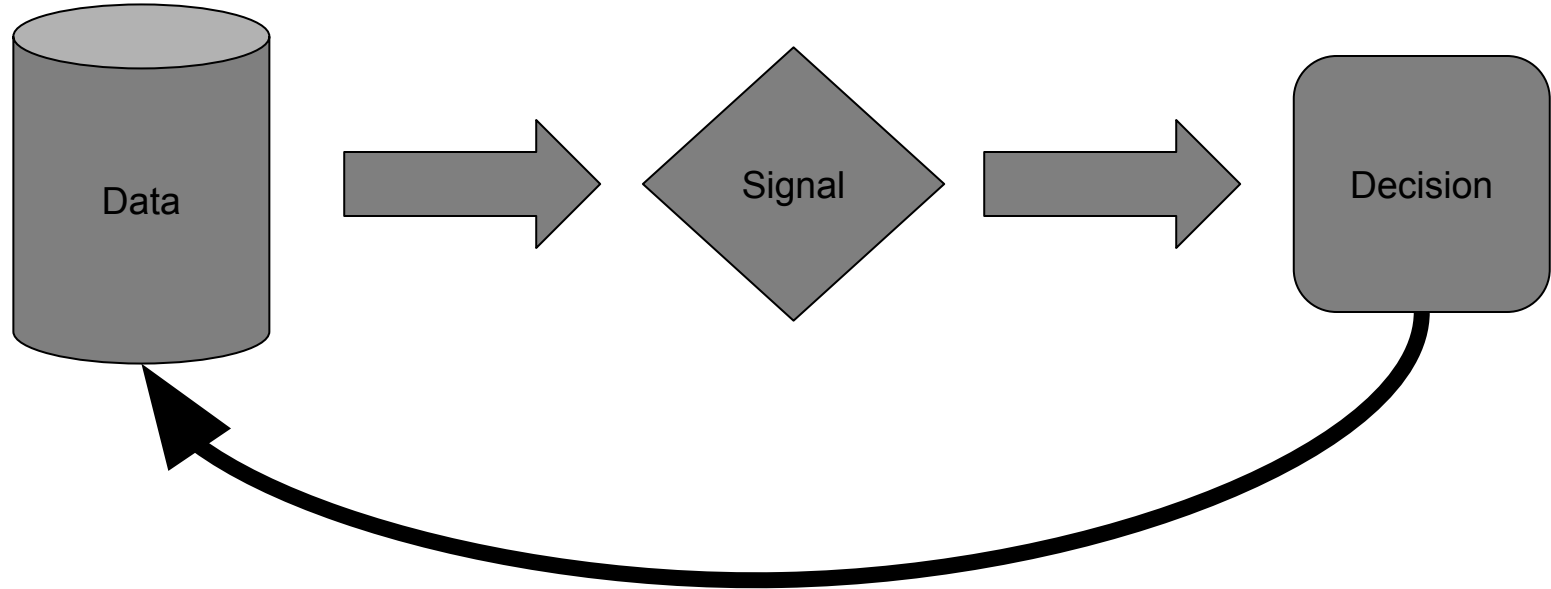
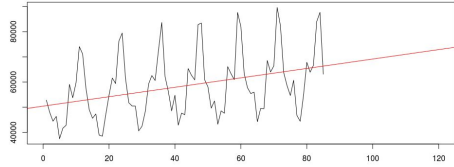


# Knowledge Graph driven Discovery

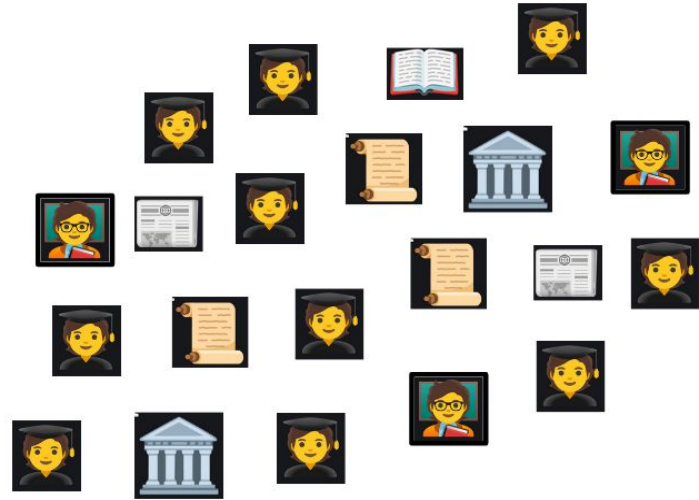
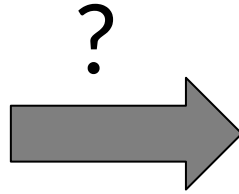
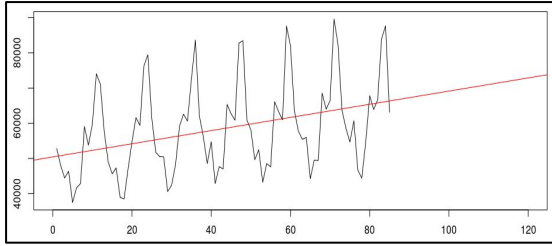
Alexander Belikov, PhD



# Decision making process



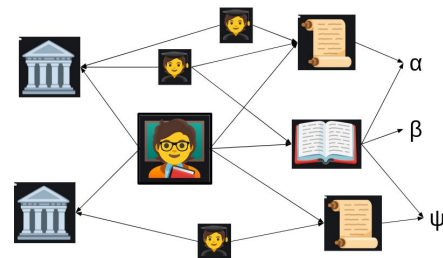
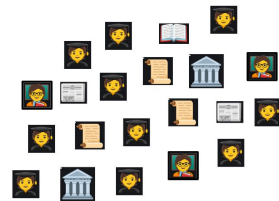
# What if data is complex?



**It should be organized!**

# What do problems do KG solve?

- Organize (normalize) unstructured data
- Facilitate access, visualize
- Generate insights



## Applications at Different Scales

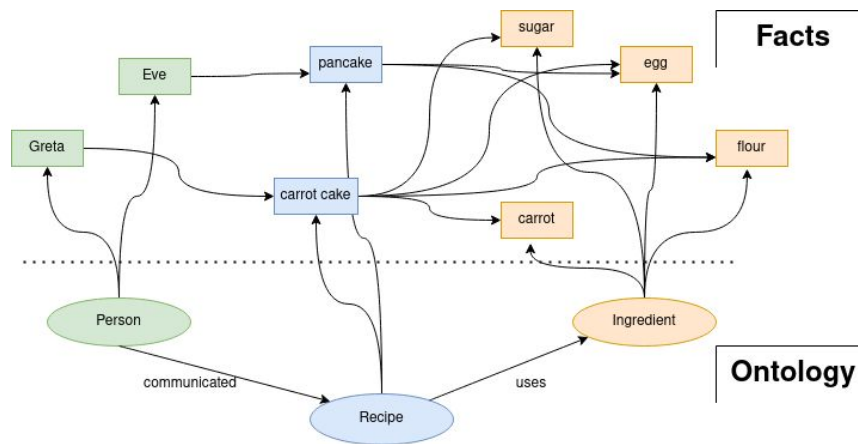
- Personal Knowledge Graphs
  - Enhance Learning Process
- Enterprise Knowledge Graphs
  - Project Management
  - Opportunity Selection / Optimization
- Domain Knowledge Graphs
  - Advance Domain Knowledge
  - Validation / Verification / Fact checking



# What is a Knowledge Graph?

A knowledge graph, also known as a semantic network, represents **a network of real-world entities**—i.e. objects, events, situations, or concepts—and illustrates the relationship between them. This information is usually stored in a graph database and visualized as a graph structure, prompting the term knowledge “graph.”

Ontology is the backbone of KG



# How do we create a KG?

Inputs:

1. structured: rdfs, grounded KGs, ontologies
2. semi-structured input: json, xml, csv
3. unstructured: text, digitally created pdf, scanned text, audio, audio records, emails

## Ideal case

structured data that uses the same ontology - can be directly added to KG

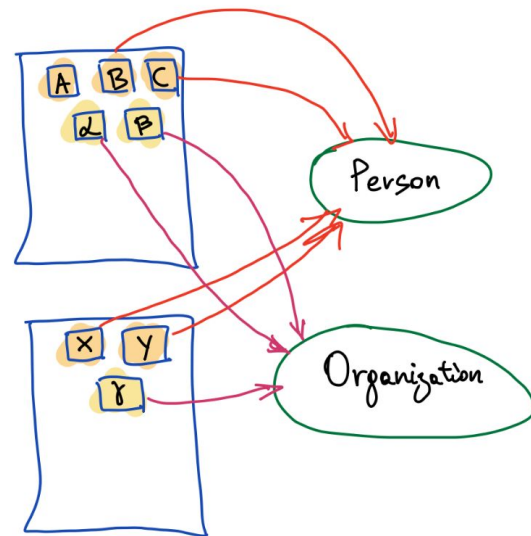
## Less ideal case

“To make pancakes use one cup of flour, two eggs and two tablespoons of sugar”

# Organizing incoming data as KG (grounding)

Not trivial when input data is not structured:

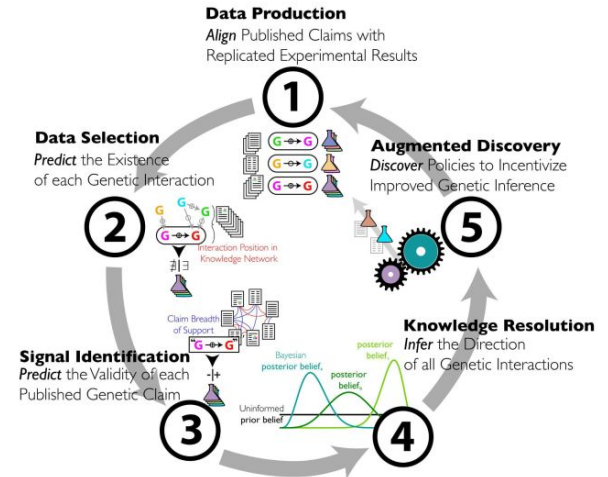
- identify named entities, relations
- map named entities to ontology
- map relation to ontology
- possibly enhance ontology



# Generating signals from KGs : Examples

## Discovery and Selection

1. Science
  - a. predict technologies that will be important tomorrow
  - b. identify correct statements in literature
2. Finance
  - a. identify perspective startups
  - b. financial instruments
3. Management
  - a. facilitate HR policies, who should be promoted
  - b. what is the best skill to learn, given my background?
4. Operations
  - a. logistics optimization

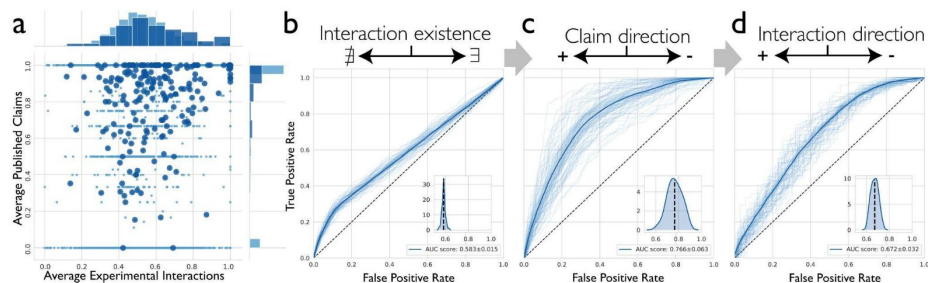




# How to generate signals from KG?

Supervised or unsupervised? Transfer learning?

1. Graph derived features used by Vanilla ML
2. Graph neural networks
  - a. embeddings
  - b. evolution on graphs, MPNN methods



# References

- [KG tutorial from WSDM](#)
- [Dynamic networks in Science](#)
- [Prediction of robust scientific facts from literature](#)