



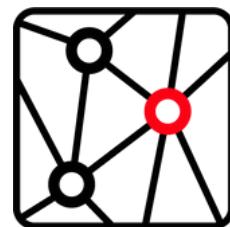
# From Theory to Practice

## A Practitioner's Journey with Knowledge Graphs



**Patryk Wielopolski**

*DataWalk*



ML in PL Conference 2024

7 - 10 October / Warsaw, Poland

# Motivation

Different cities, ...



# Motivation

Different cities, ...



The same question:

What are Knowledge Graphs?

# Today's Speaker

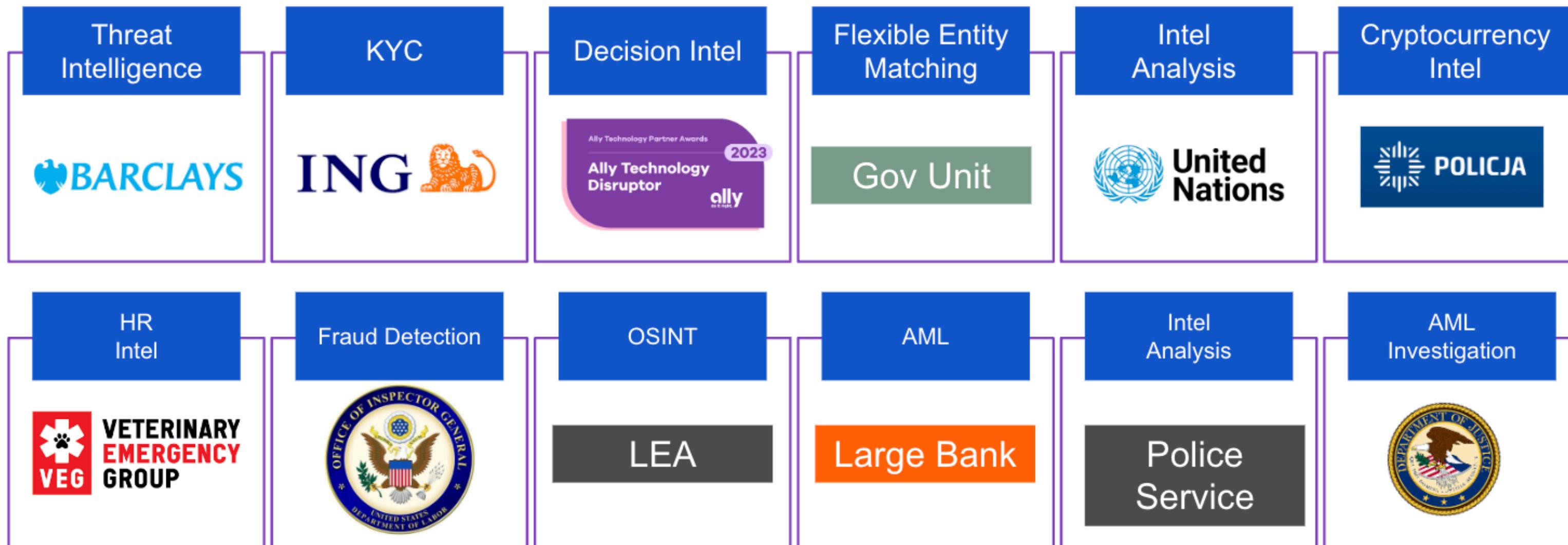


6 years at DataWalk as Data Scientist,  
Solution Architect, R&D Team Leader



Wrocław-based Start-Up  
Knowledge Graph Platform

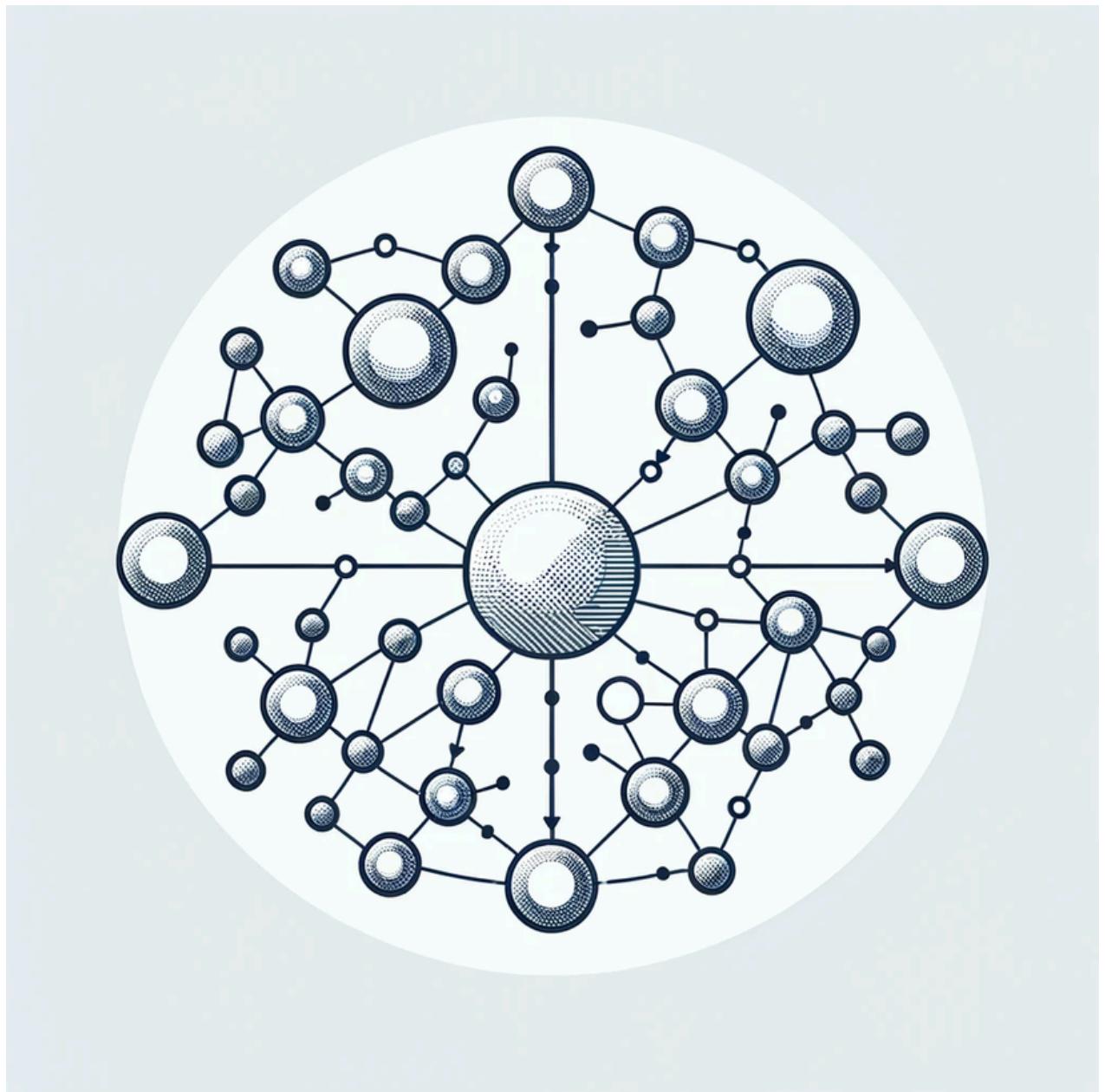
# Examples of Implementations



# Knowledge Graphs

# Knowledge Graph

Semantically enriched information organized as interconnected entities, properties, and relationships in a graph-structured data model in order to mirror human intuition.



# Knowledge Graph

Semantic Layer

Technology Layer

# Theory of Knowledge Graphs

## Note on Technology Layer

# Graph Databases

## Graph Data Models



# Graph Databases

## Graph Data Models

### Property Graph

*Entities, Relationships,  
Properties*

#### Examples

Linked In

VISA Transaction Network  
Transportation Networks



# Graph Databases

## Graph Data Models

### Property Graph

*Entities, Relationships,  
Properties*

#### Examples

Linked In  
VISA Transaction Network  
Transportation Networks



### Semantic Web KG\*

*Triples: Subject, Predicate,  
Object*

#### Examples

WikiData  
DBpedia  
Drug Discovery



# Theory of Knowledge Graphs

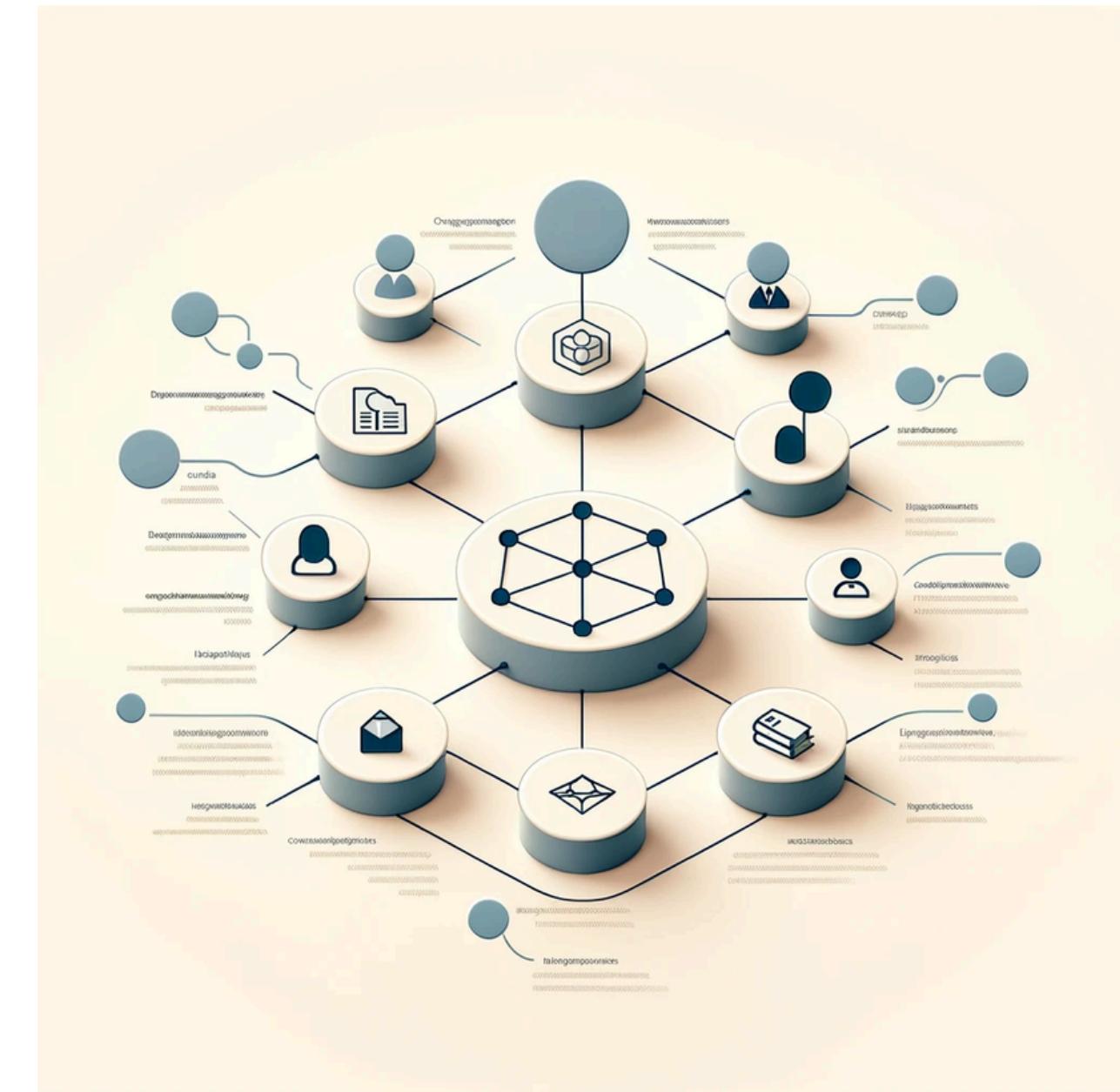
## Semantic Layer

# Ontologies

**Ontologies** are formal representations of knowledge with a rich set of **relationships** among various **entities** and **concepts**, to enable machines to understand and interpret the **semantics** of data.

By using standardized ontologies businesses can create **shared understandings** and achieve better data **integration, interoperability**, and **reasoning**.

Semantic Web and **OWL (Ontology Web Language)** are widely known standards for describing ontologies.



# Domain Ontology & Instance Data

## Domain Ontology

A formal representation of knowledge as a set of concepts within a domain, defining the relationships among those concepts to enable shared understanding and data interoperability across applications.

# Domain Ontology & Instance Data

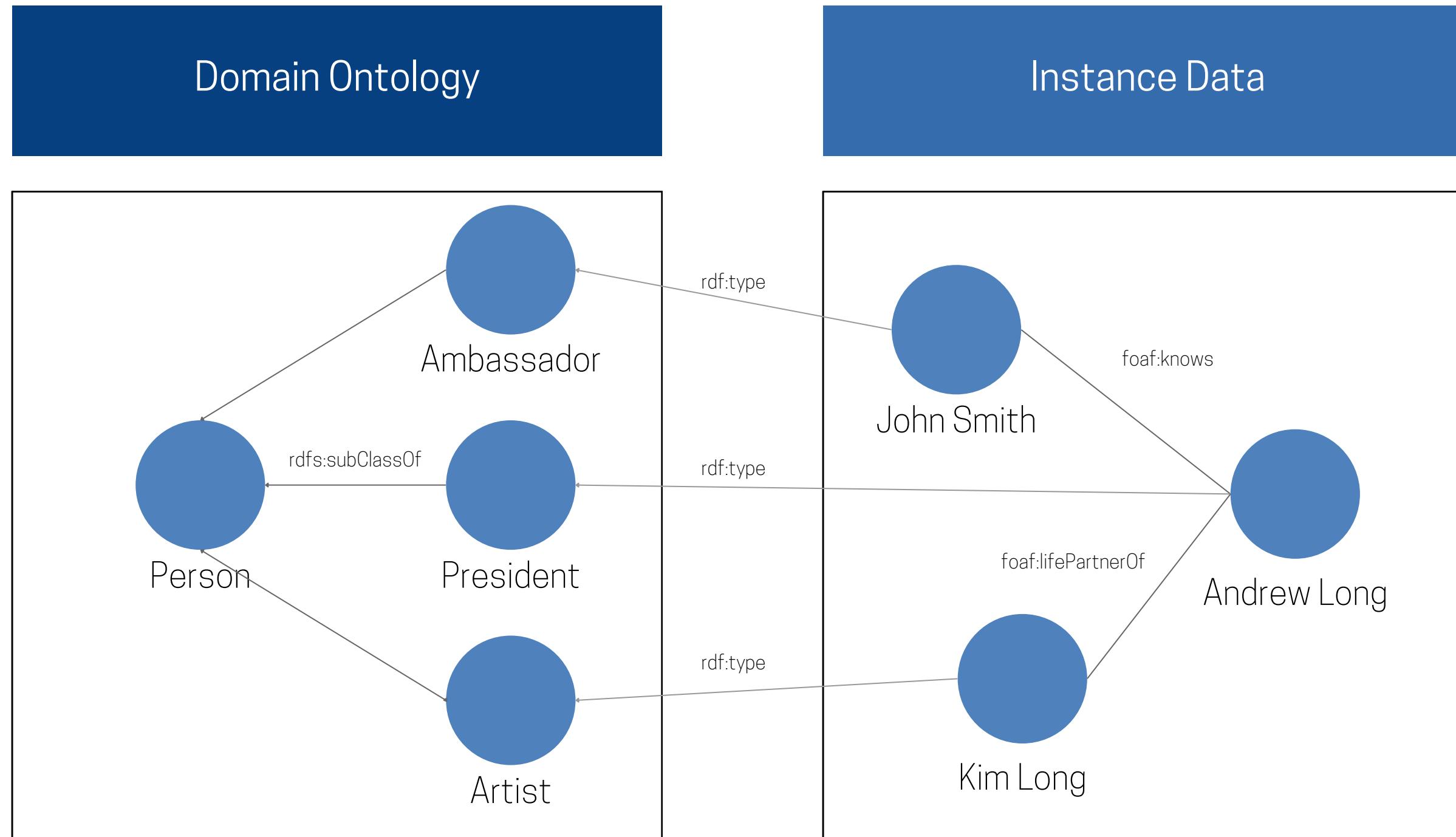
## Domain Ontology

A formal representation of knowledge as a set of concepts within a domain, defining the relationships among those concepts to enable shared understanding and data interoperability across applications.

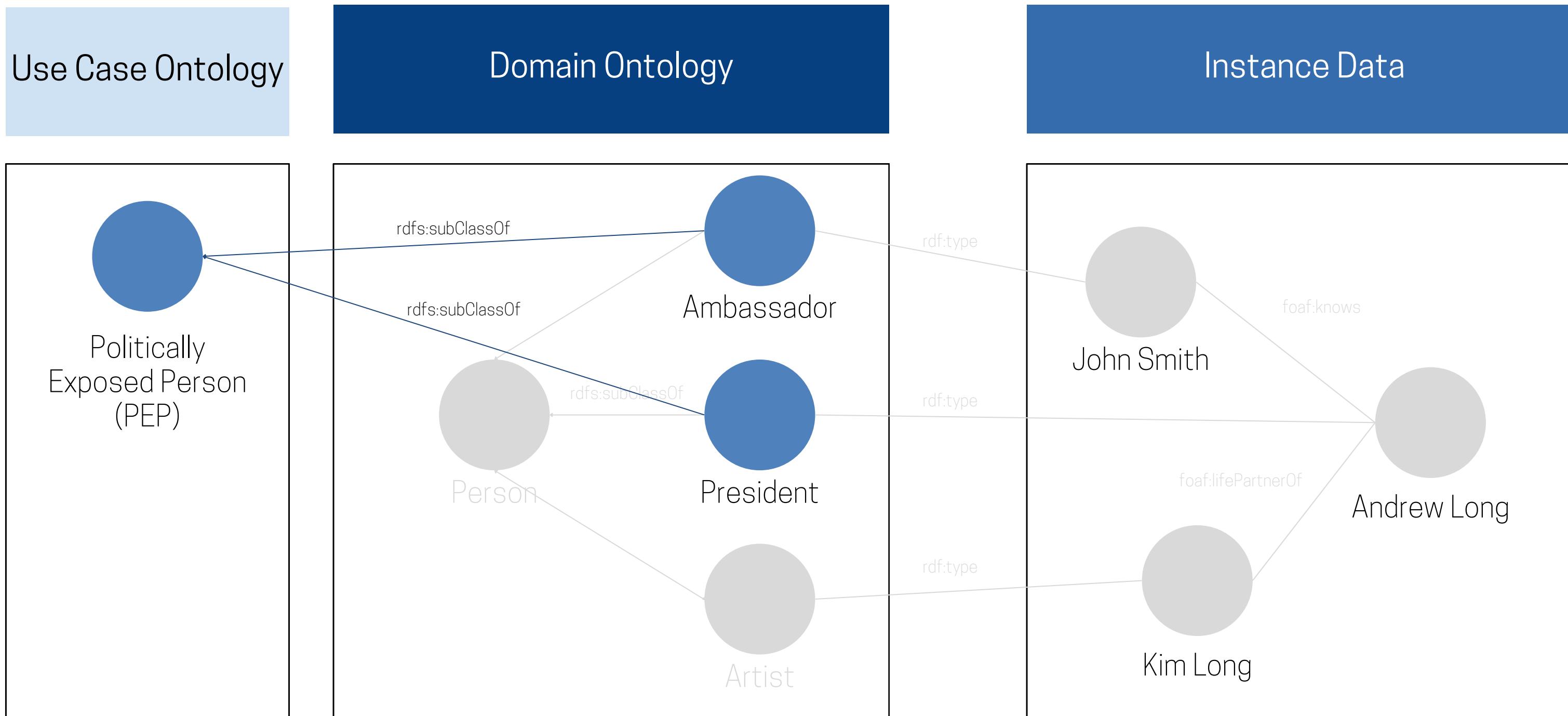
## Instance Data

Specific examples of data that instantiate the concepts and relationships defined in a domain ontology, representing real-world entities and their interconnections.

# Domain Ontology & Instance Data



# Use Case Ontology

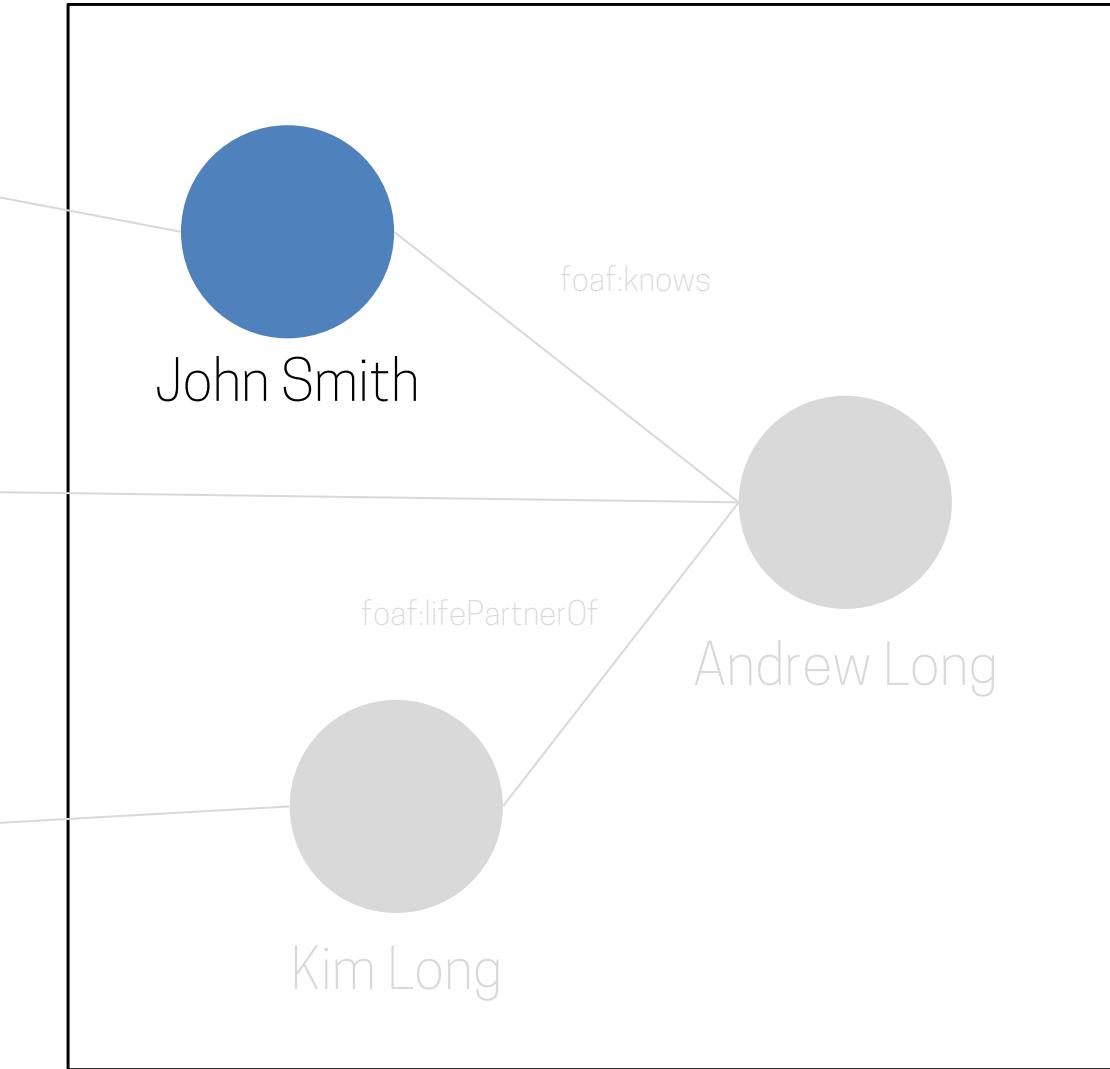
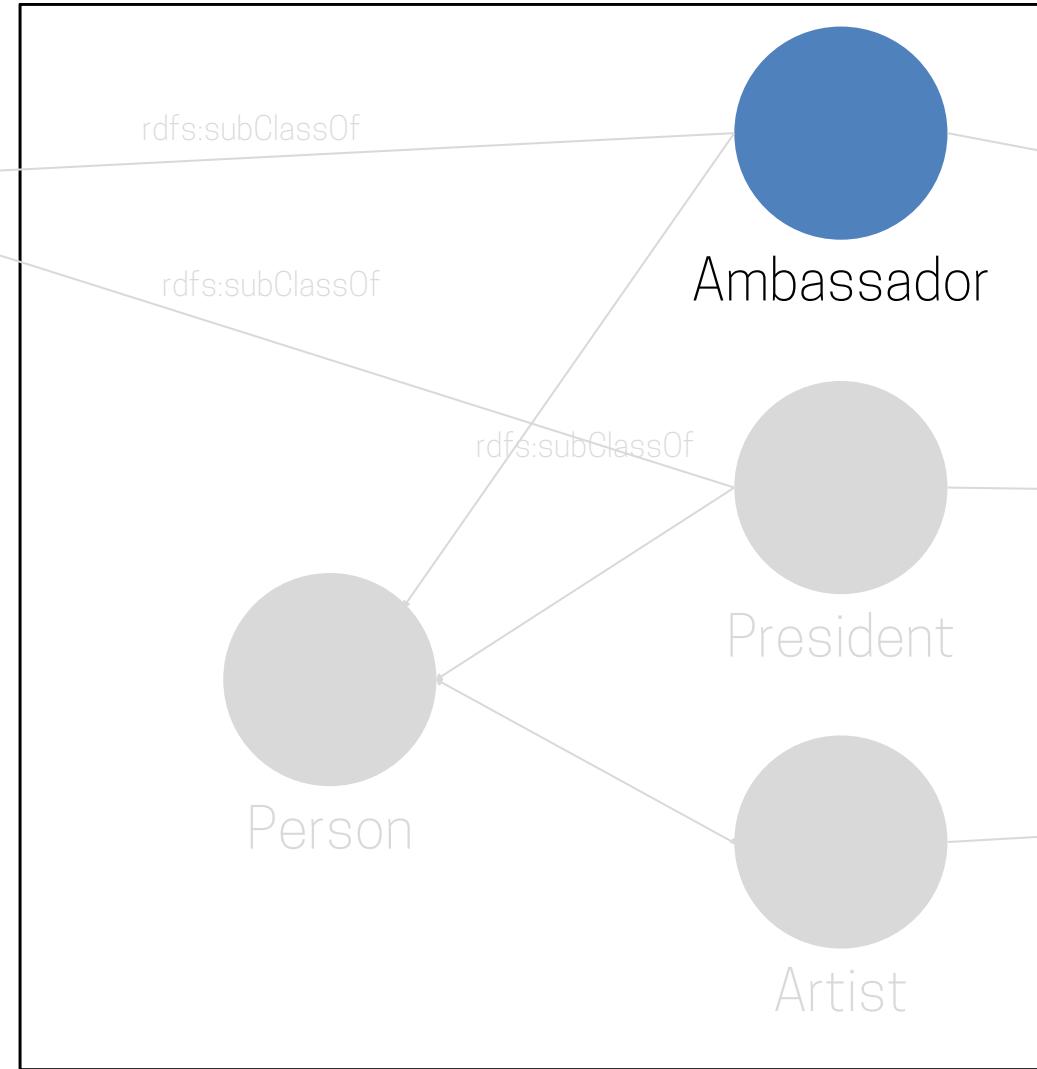
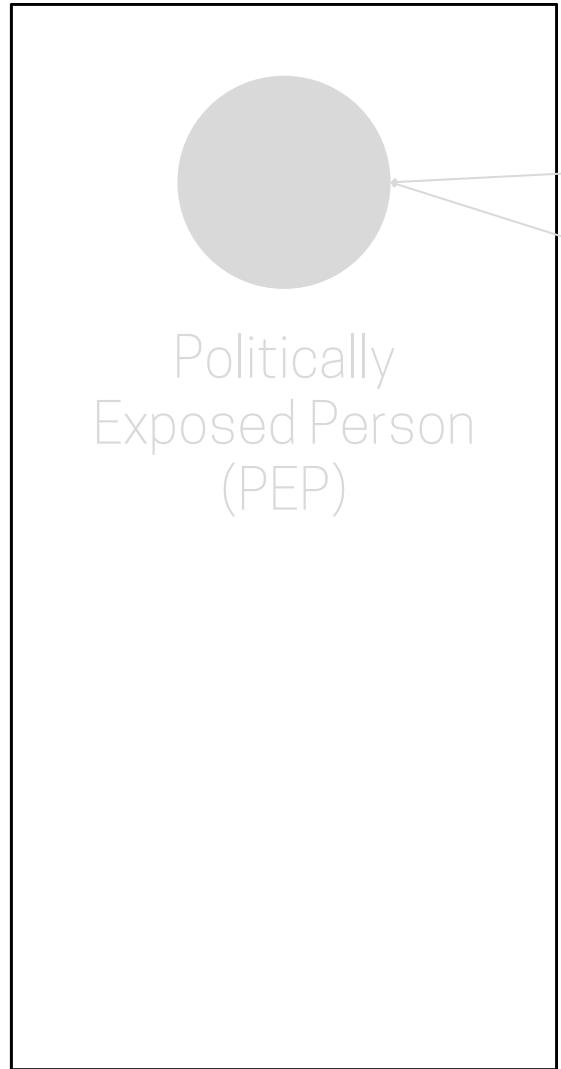


# Inference

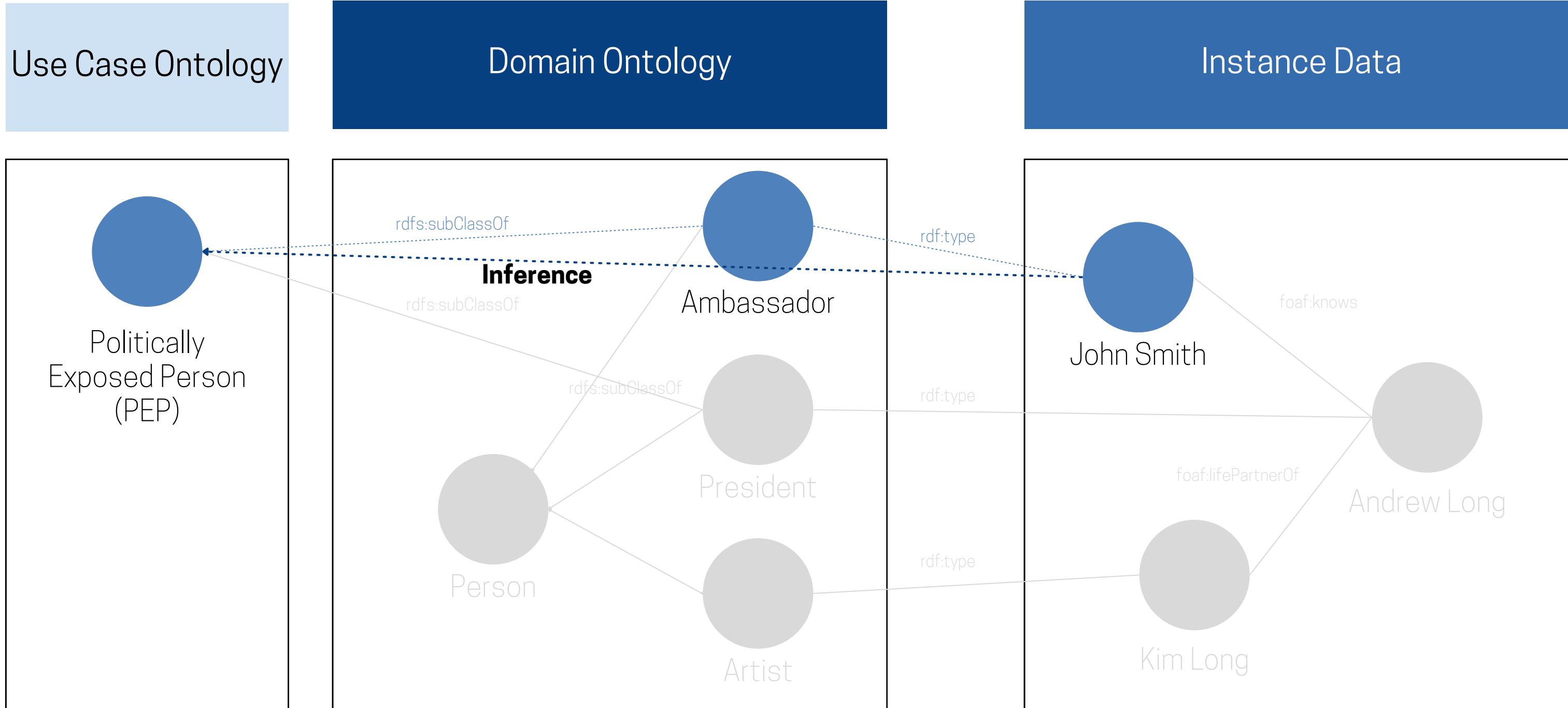
Use Case Ontology

Domain Ontology

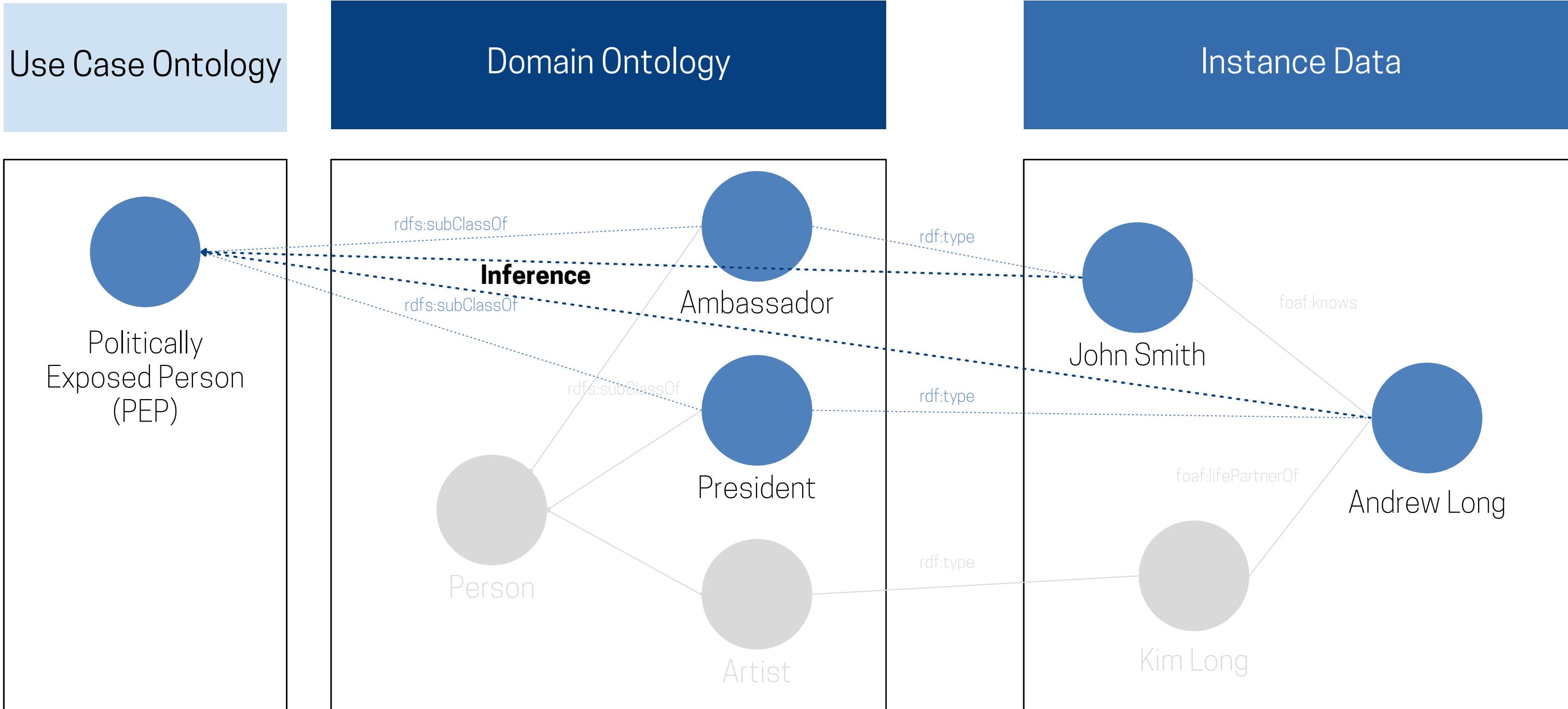
Instance Data



# Inference



# Inference



Inference:  
The same is for **Andrew Long**.

# Practice of Knowledge Graphs

## Implementation Process

# Implementation Process

Disclaimer

- Enterprise Data, i.e., Mess and Billions of Records
- Sometimes DataWalk-specific
- Results of observing implementations for a few years

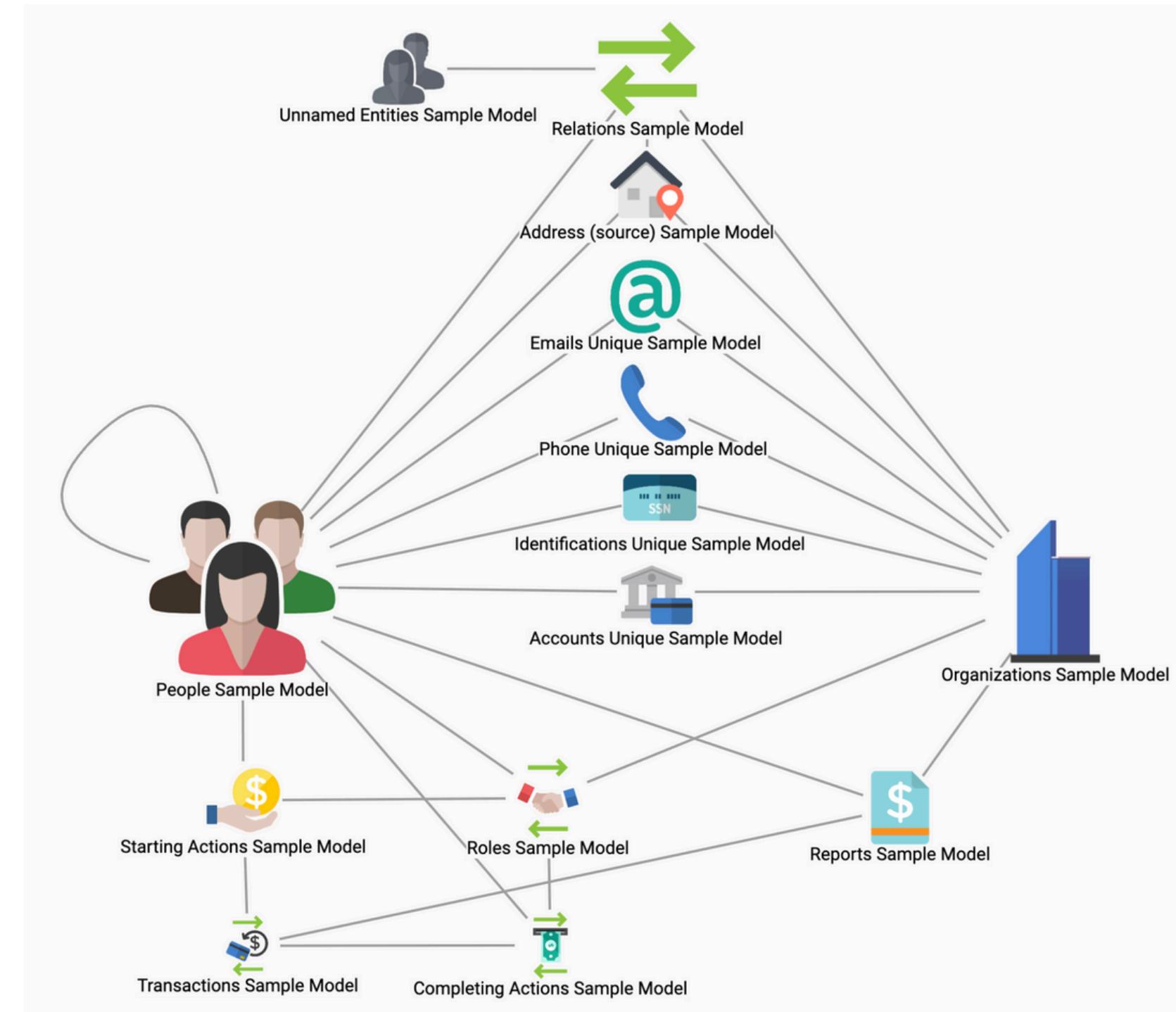
# Implementation Process

## Overview



# Implementation Process

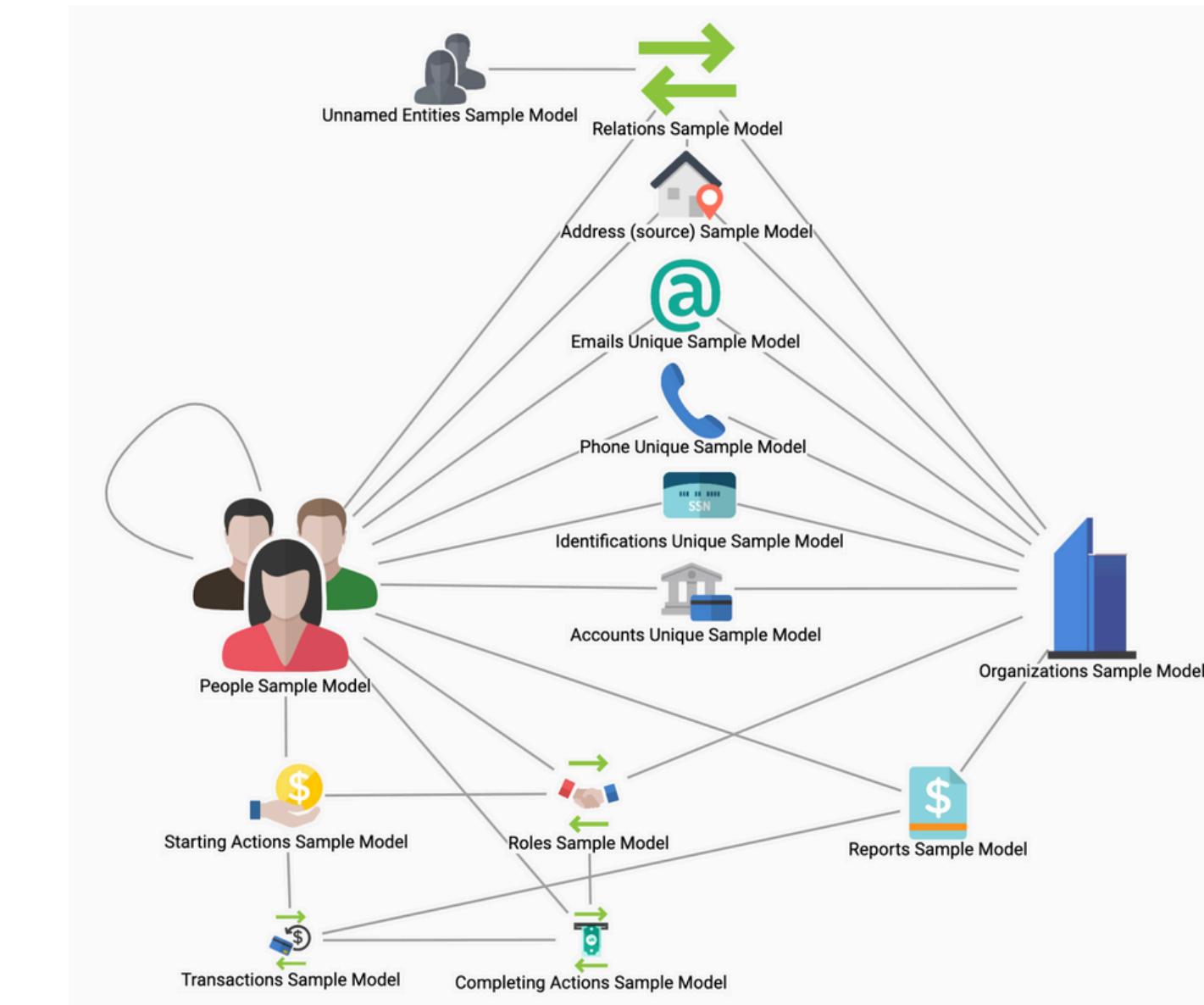
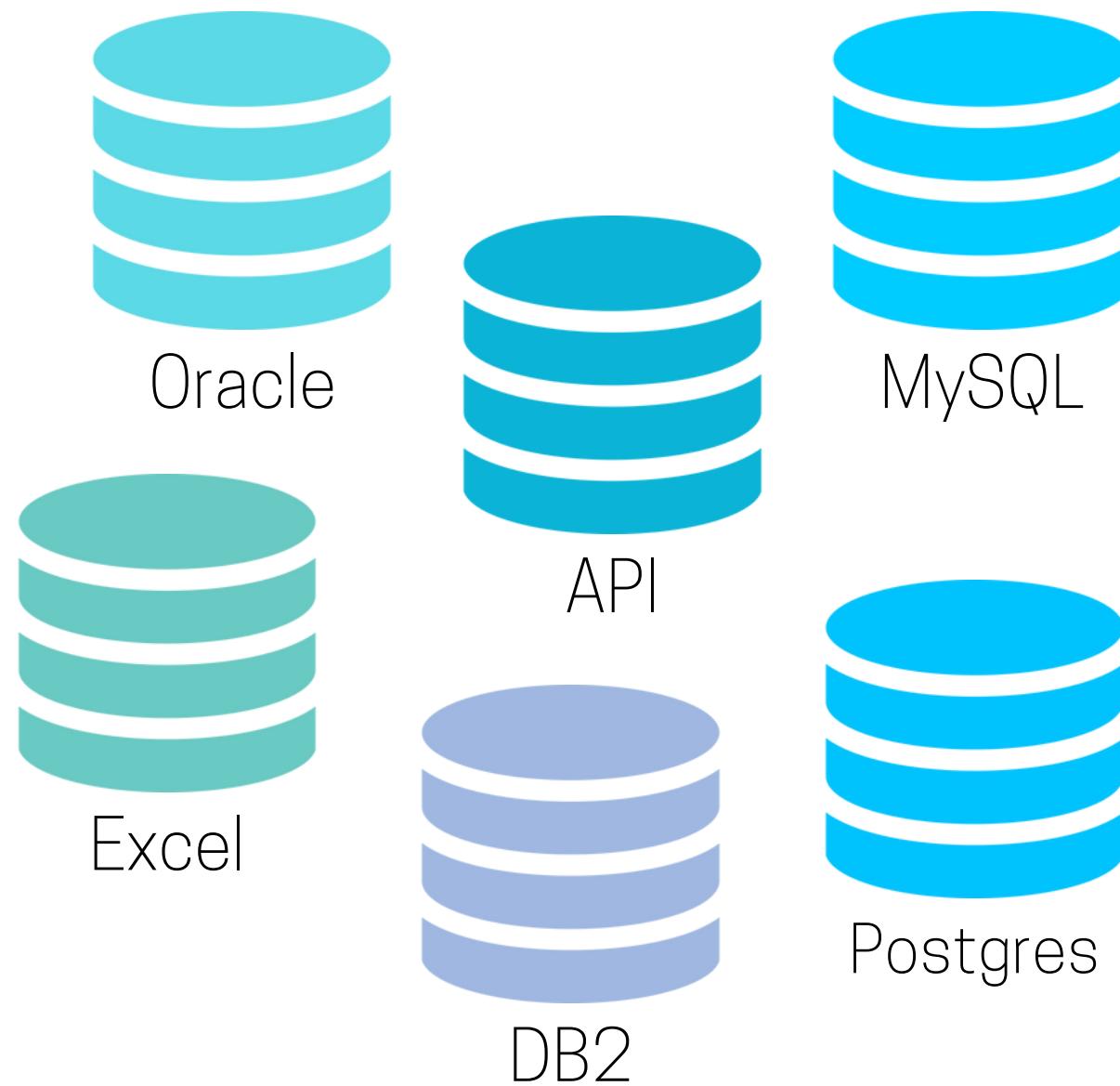
Codifying Expert Knowledge  
🔑 Competency Queries



iTelos  
Good Reference

# Implementation Process

## Mapping Sources & Data Loading



*Important Part but takes time...*

# Implementation Process

## Data Cleaning

### openvenues/ **libpostal**

A C library for parsing/normalizing street addresses around the world. Powered by statistical NLP and open geo data.

52  
Contributors

277  
Issues

4k  
Stars

418  
Forks



Address Parser

### google/ **libphonenumber**

Google's common Java, C++ and JavaScript library for parsing, formatting, and validating international phone numbers.

140  
Contributors

441  
Used by

17k  
Stars

2k  
Forks



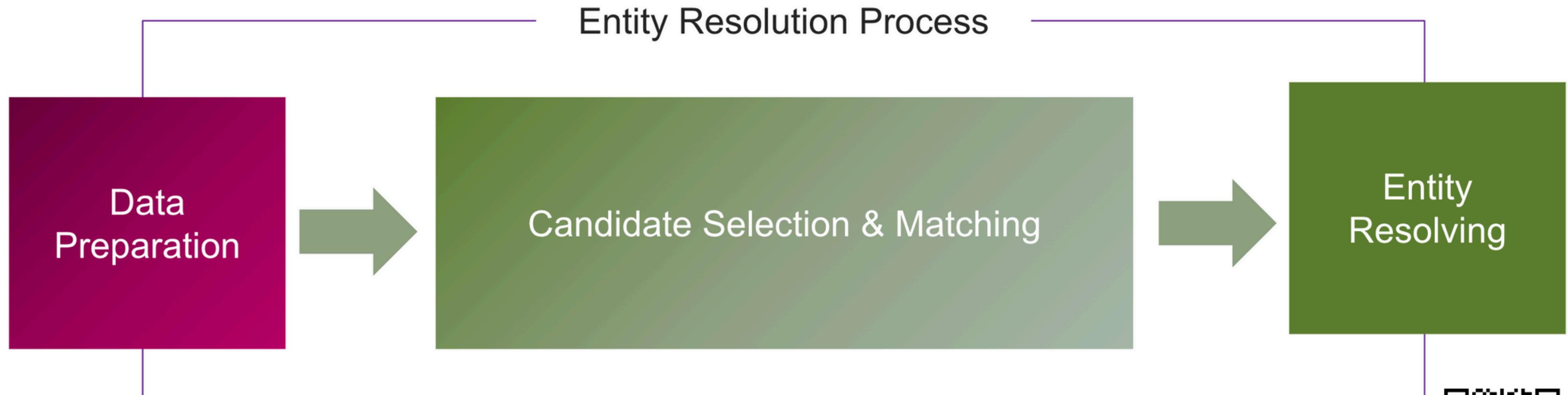
Phone Number Parser

*Note: C/C++ Implementation  
#DataVolumes #LawOfPhysics*

# Implementation Process

## Entity Resolution

Entity Resolution is the process of **identifying** and **linking** records that refer to **the same real-world entity**.



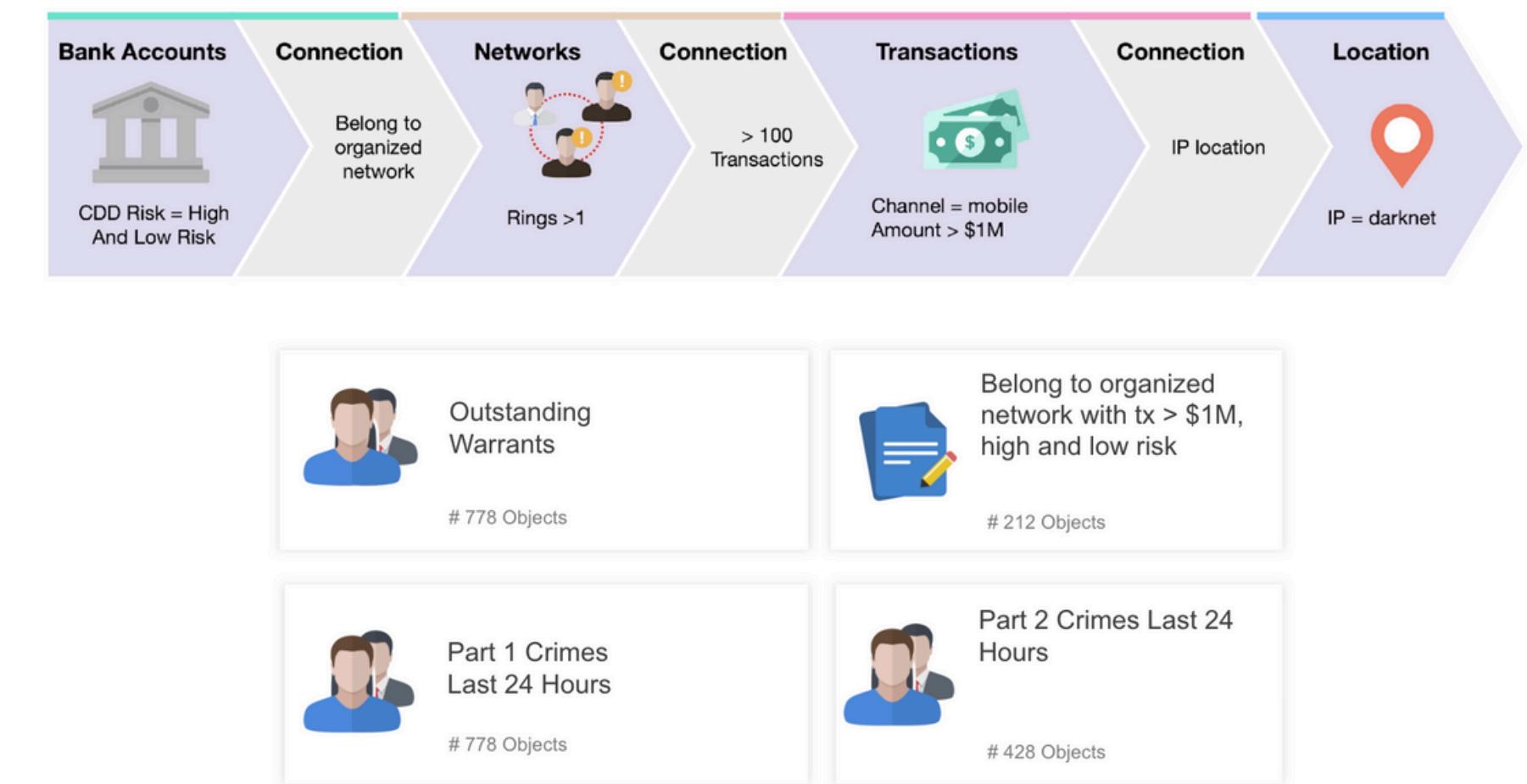
*ER Guide*

# Implementation Process

## Analytics

The screenshot shows a detailed summary dashboard for John Nelson Smith. At the top, there are three status indicators: 'REGISTERED WEAPONS?' (NO), 'OUTSTANDING WARRANT?' (YES), and 'ARMED & DANGEROUS?' (YES). Below these are sections for 'IDENTIFYING INFORMATION' (including a photo and basic details like First: John, Last: Nelson, etc.), 'ADDRESS HISTORY' (listing multiple addresses from 2009 to 2015), 'ARREST HISTORY' (listing multiple arrests from 2007 to 2015), 'KNOWN ASSOCIATES' (listing Bill Smith and Jan Byron with their locations), and 'LINKED PLACES' (a map showing various locations). A sidebar on the left provides a comprehensive list of personal details such as race, gender, height, weight, hair, eyes, DOB, age, marks, gang, cellphone, and email.

360 Summary Dashboard



Querying the Knowledge Graph  
Creating Rule-based Systems

# Practice of Knowledge Graphs

## More Analytics

# **Practice of Knowledge Graphs**

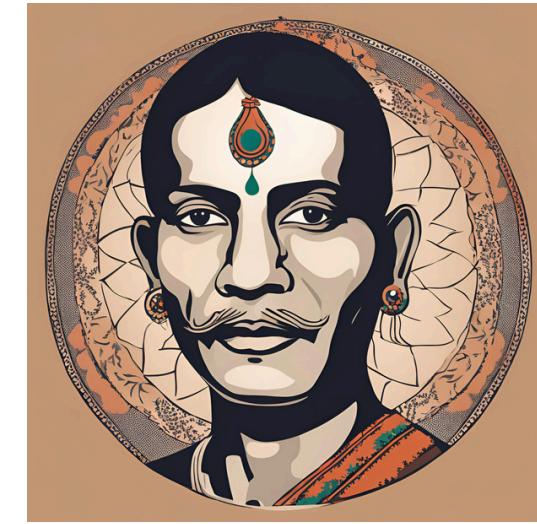
More Analytics

- Graph Analytics for Law Enforcement
- AI for Law Enforcement

# Graph Analytics for Law Enforcement



John  
Smith

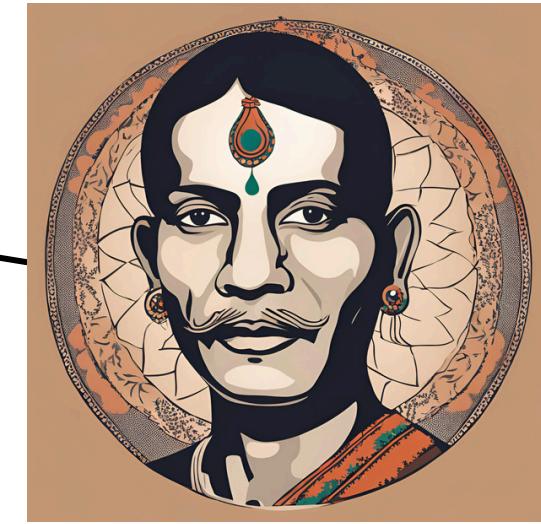


Andrew  
Long

# Graph Analytics for Law Enforcement

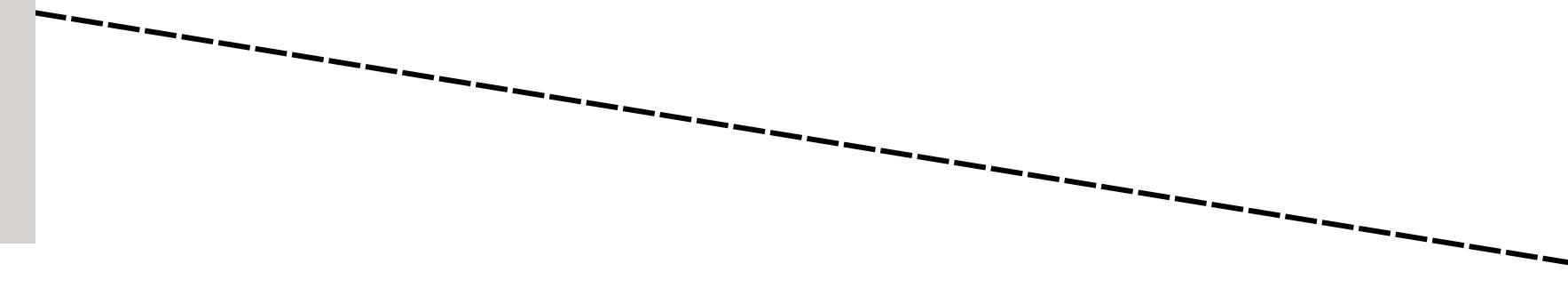


John  
Smith



Andrew  
Long

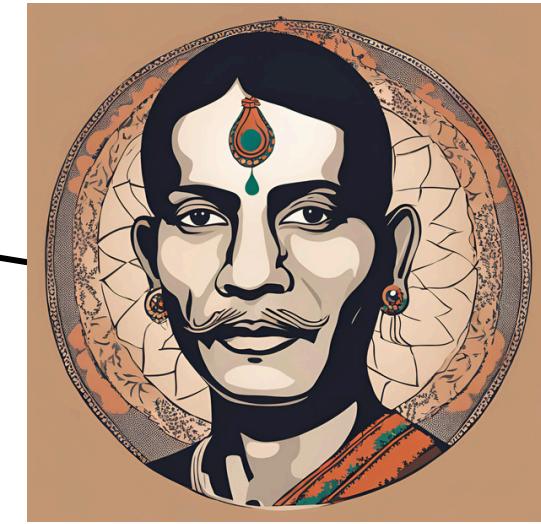
*Is there a path between those two people?*



# Graph Analytics for Law Enforcement



John  
Smith



Andrew  
Long

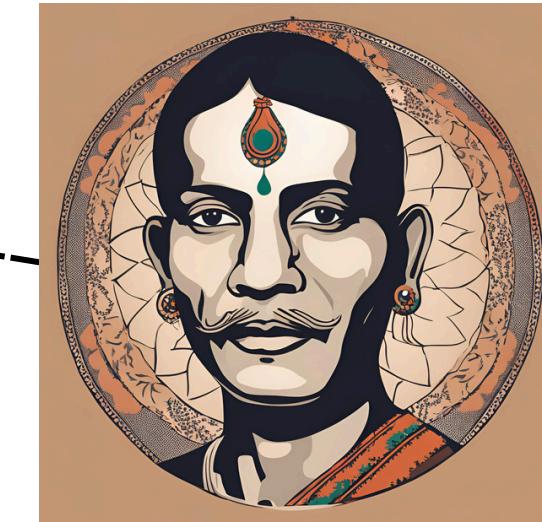
*Is there a path between those two people?*

*Consider all paths in the system,  
i.e., 100M+ entities and relationships*

# Graph Analytics for Law Enforcement



John  
Smith



Andrew  
Long

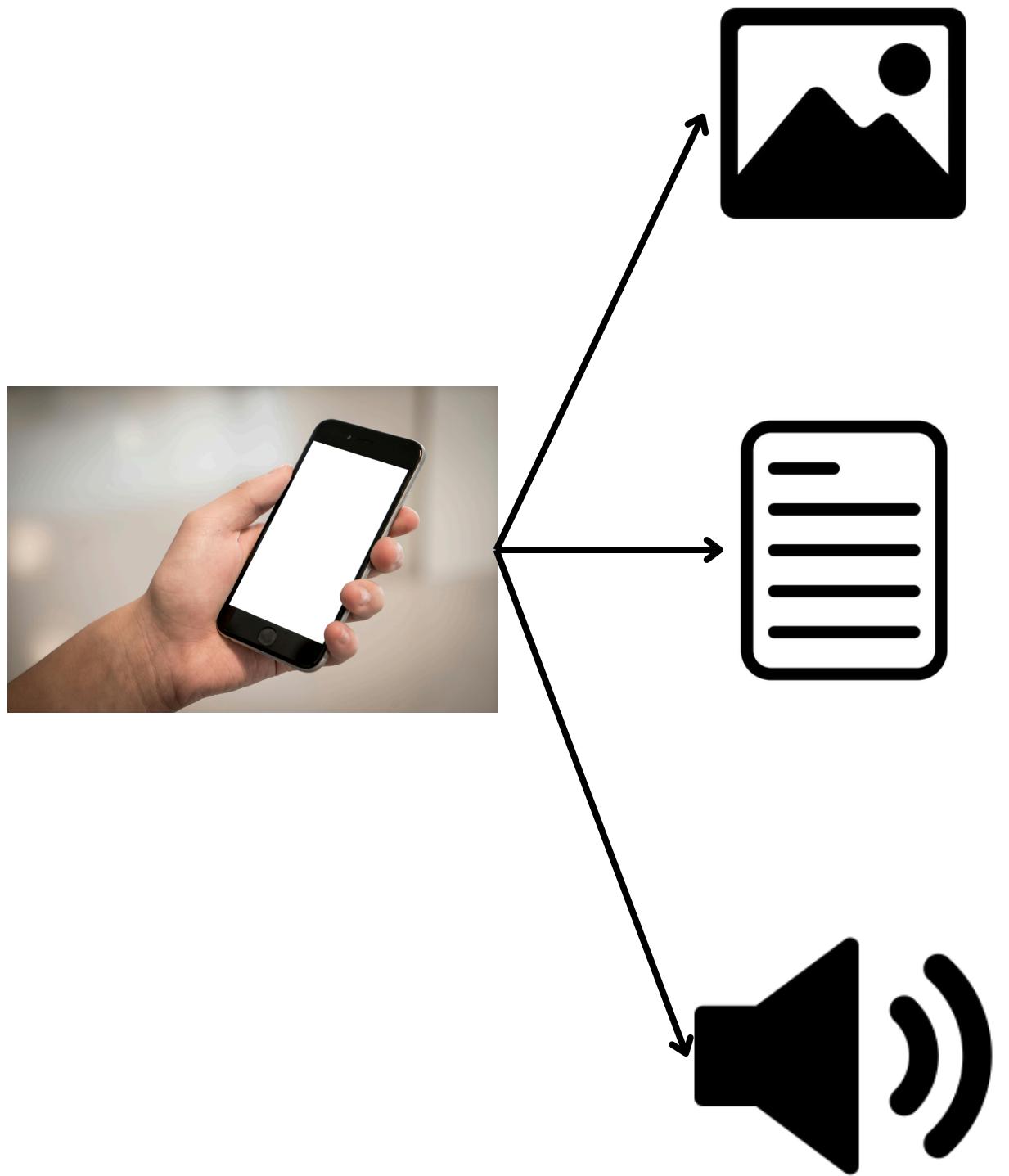
*Is there a path between those two people?*

*Consider all paths in the system,  
i.e., 100M+ entities and relationships*

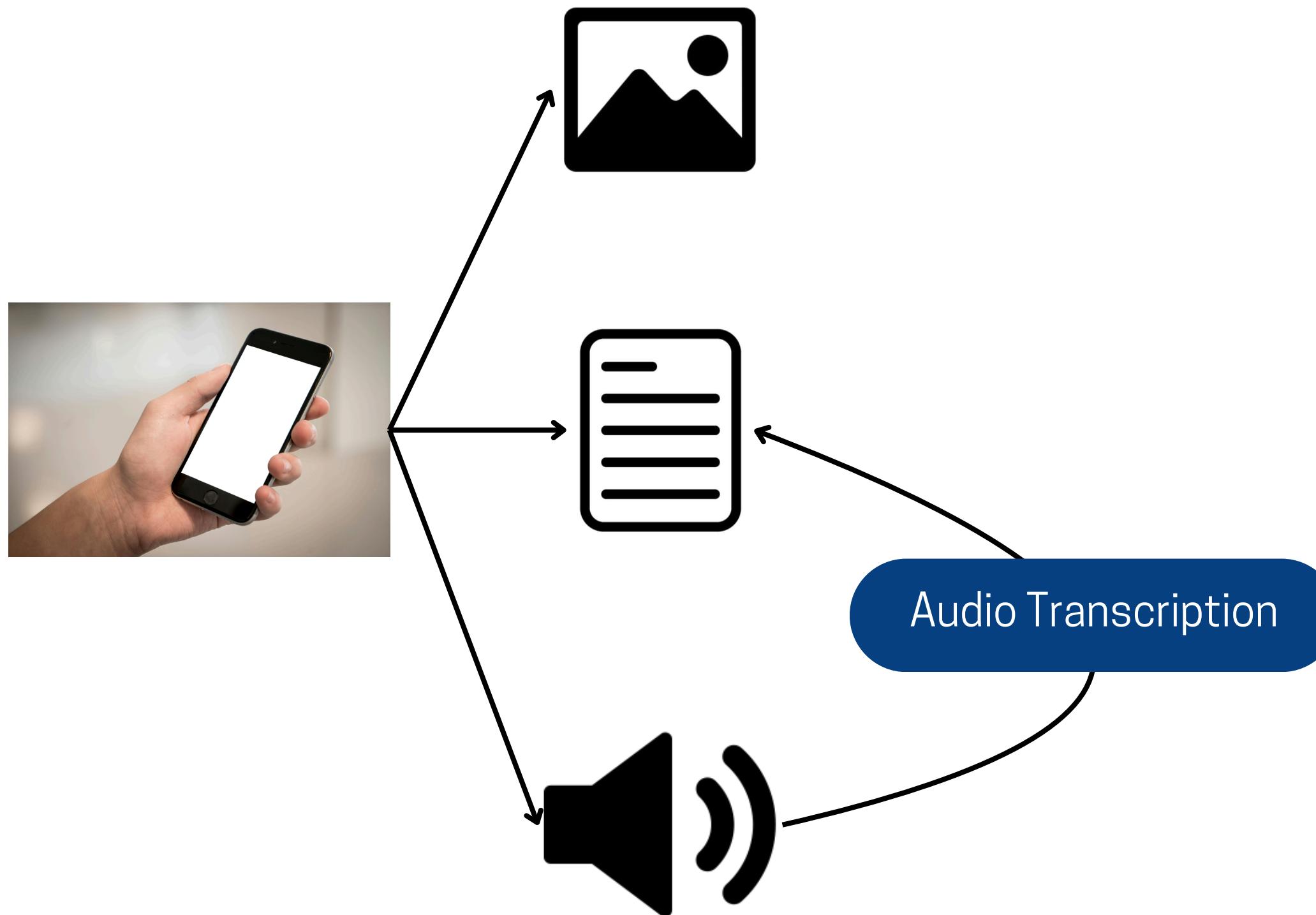
## **Effect?**

*Investigations solved in one click vs. weeks of manual work.*

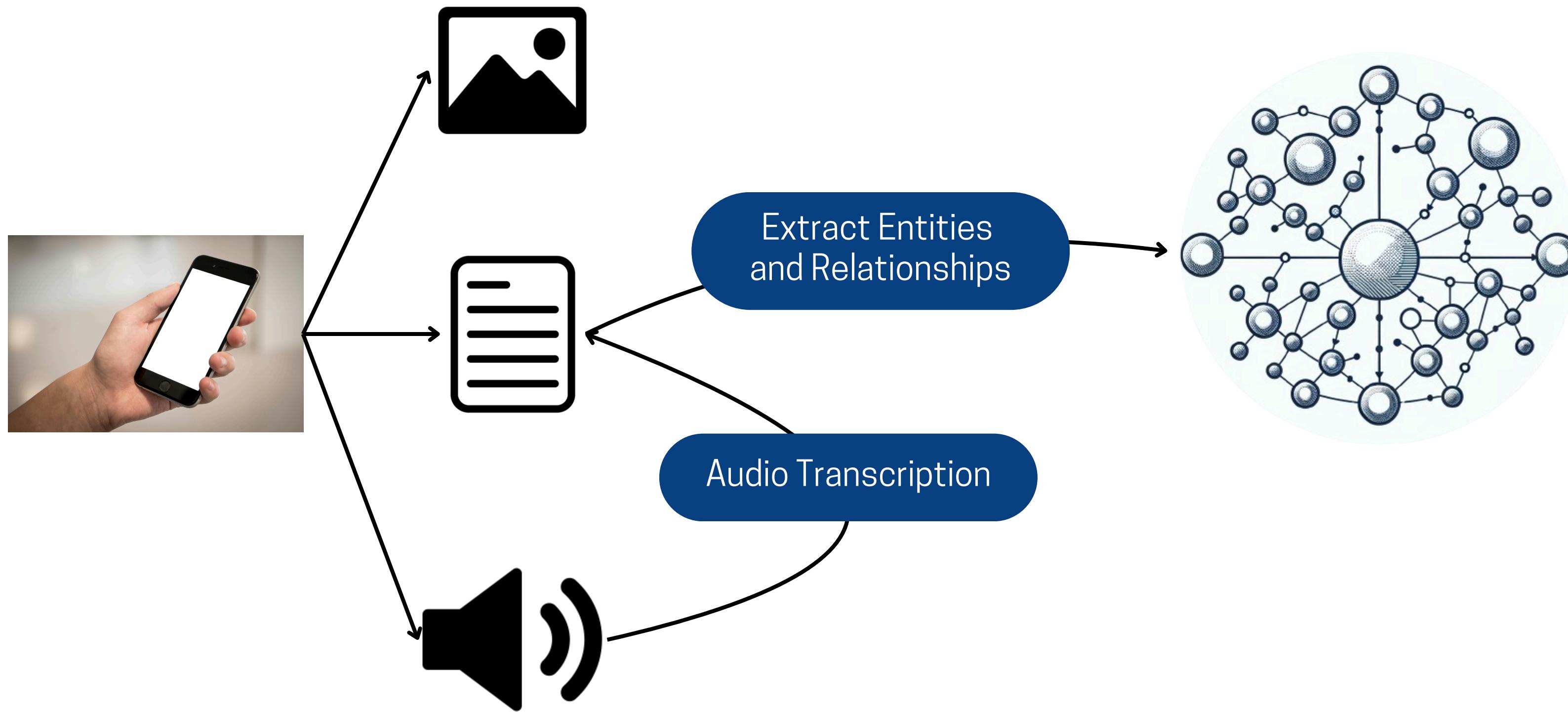
# AI for Law Enforcement



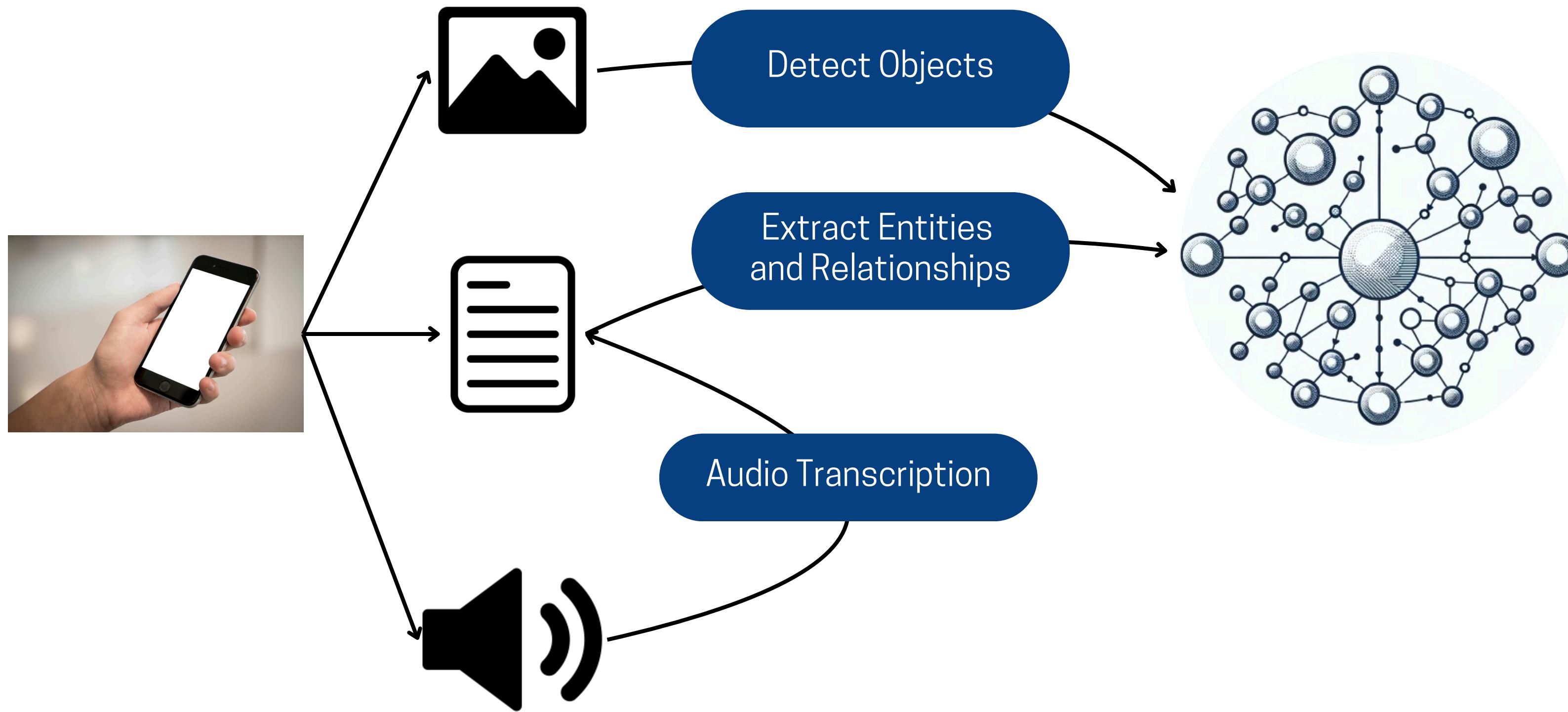
# AI for Law Enforcement



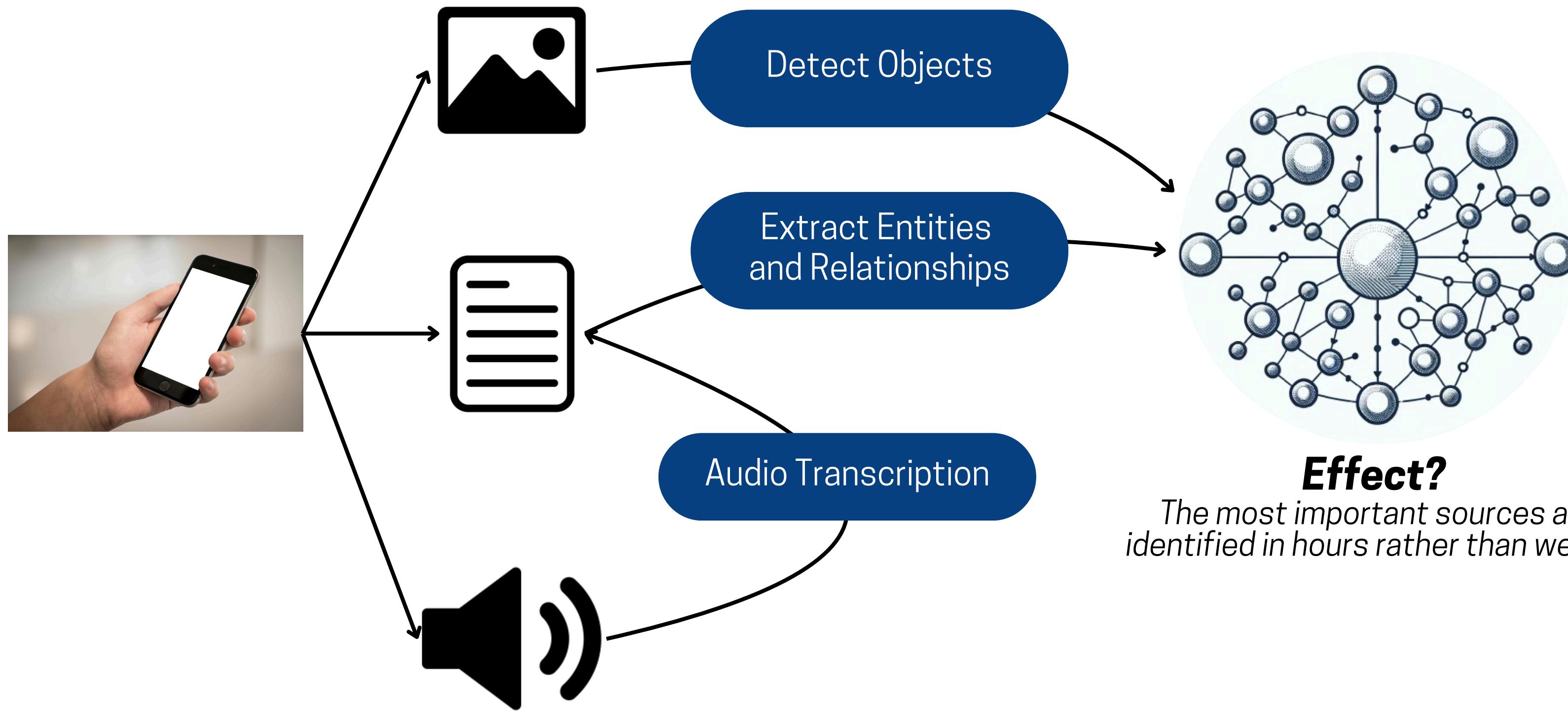
# AI for Law Enforcement



# AI for Law Enforcement



# AI for Law Enforcement



## **Effect?**

*The most important sources are identified in hours rather than weeks.*



# Summary

- You understand that Knowledge Graphs can serve as a powerful foundation for analytics.
- You learned about the Knowledge Graph creation process.
- You know important keywords: Knowledge Graph, Ontology, Inference, Entity Resolution
- You found useful parsing tools to check: libpostal, libphonenumber
- You got inspired! 

Thank you  
for your attention!

Questions?

Let's Discuss!

