



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

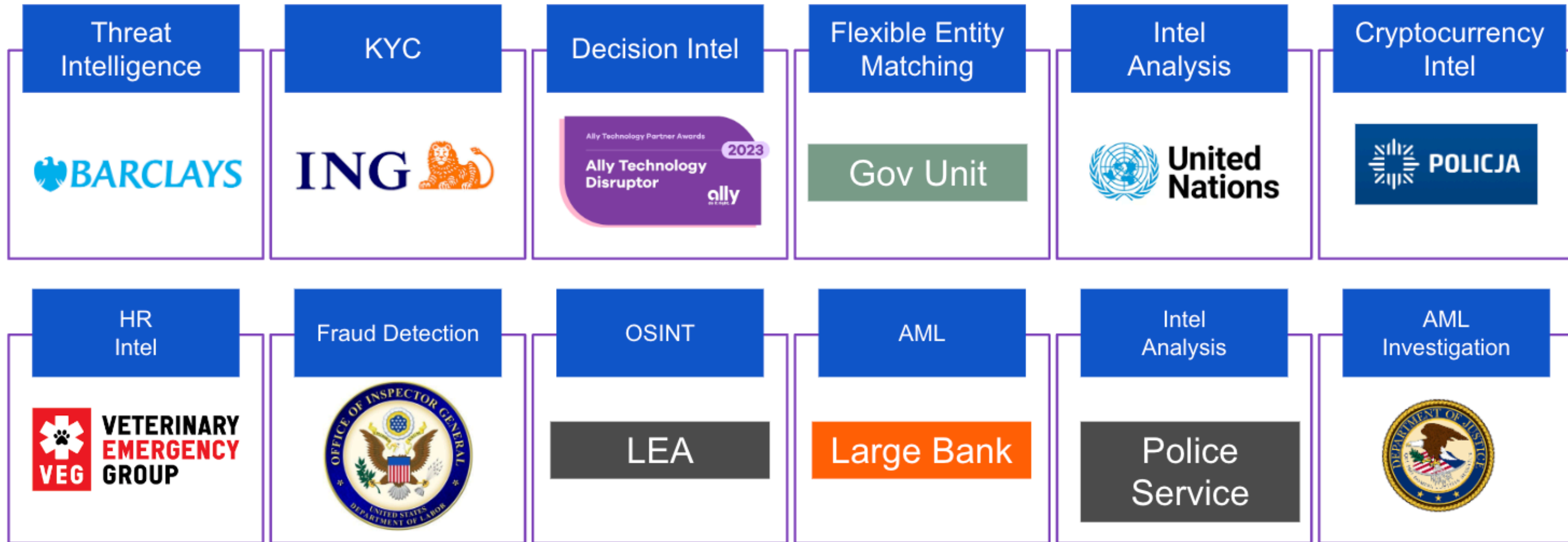


6 years at DataWalk as Data Scientist,
Solution Architect, R&D Team Leader



Wrocław-based Start-Up
Knowledge Graph Platform

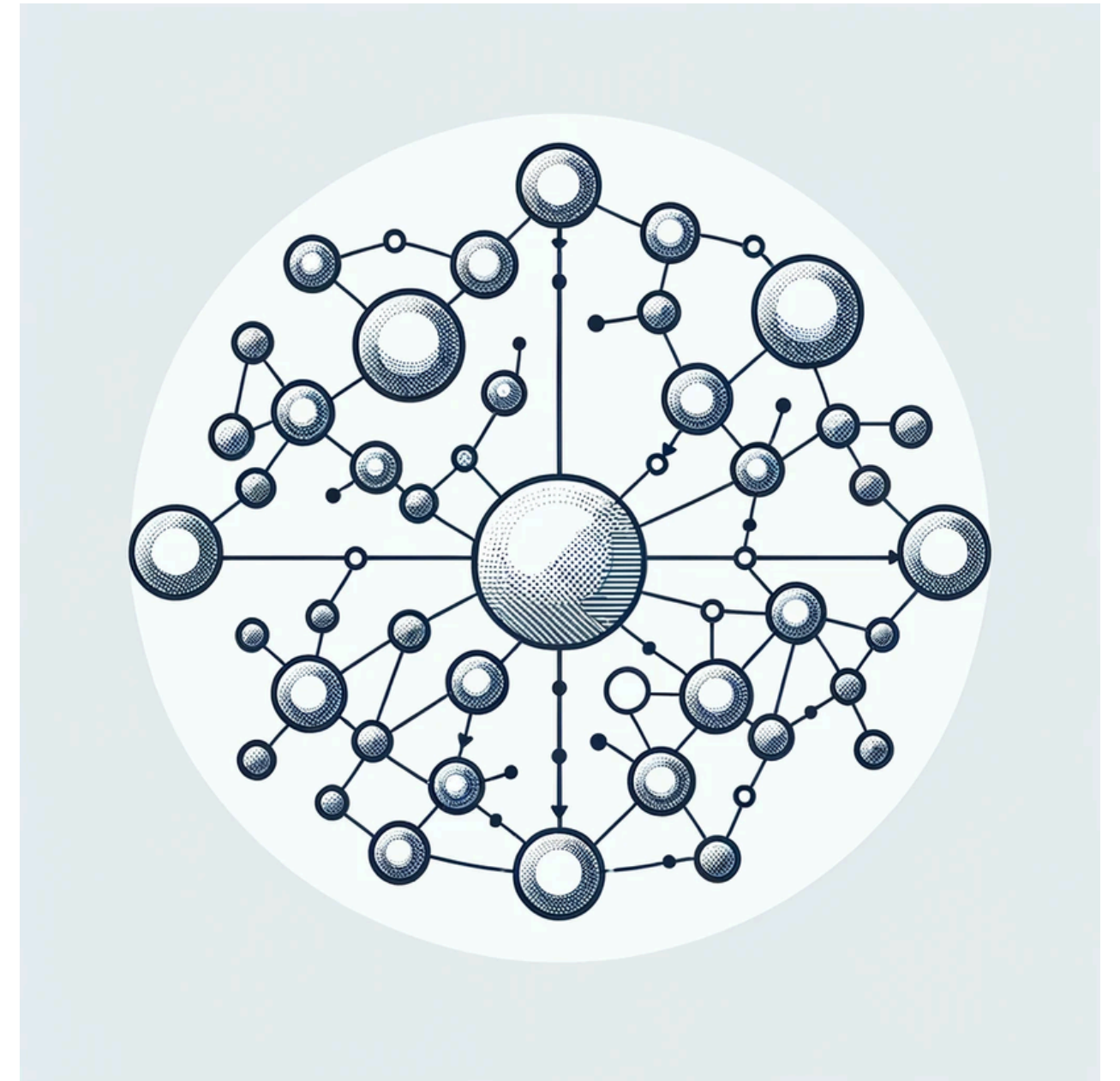
Examples of Implementations



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

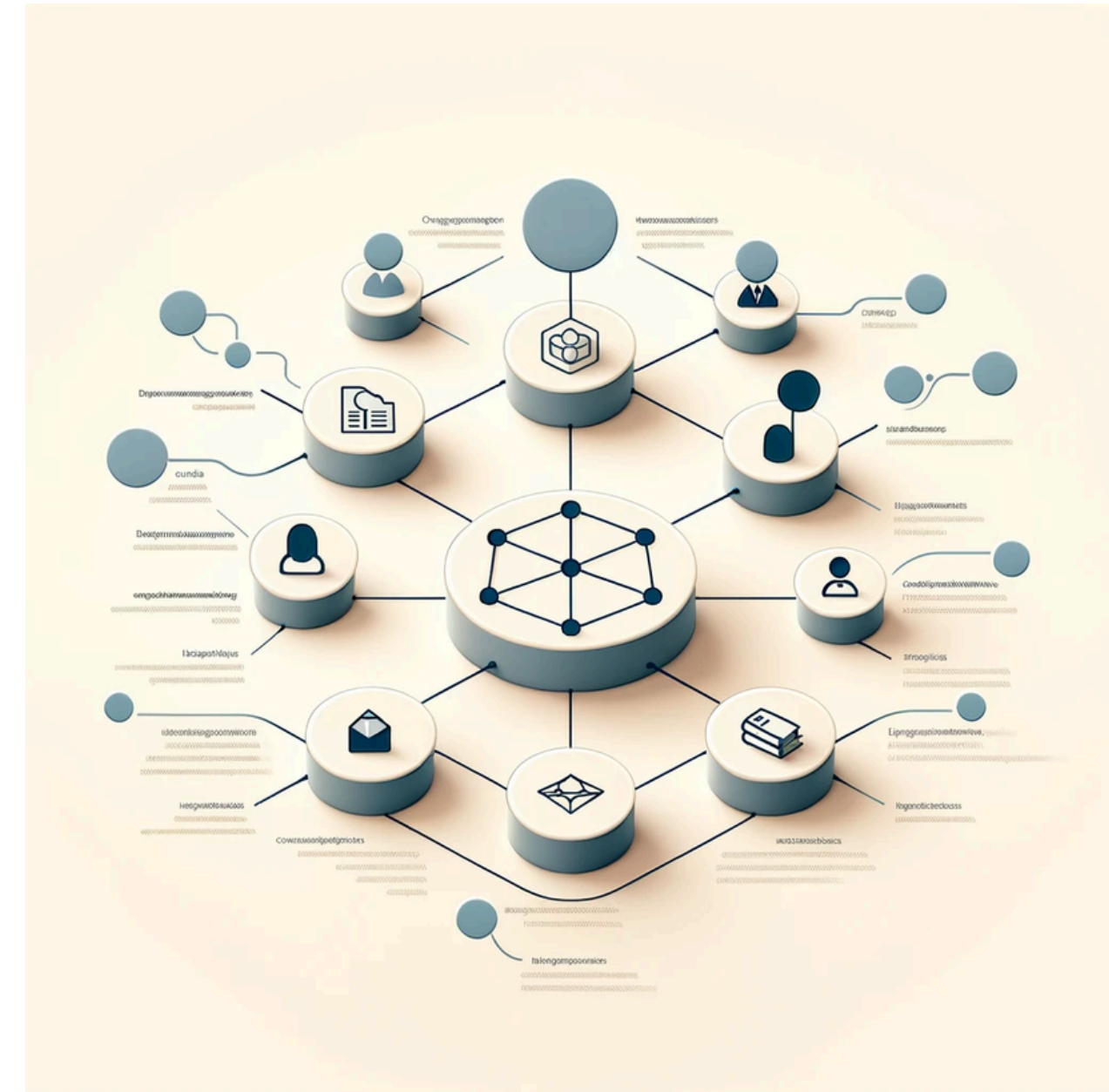
Semantic Layer

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

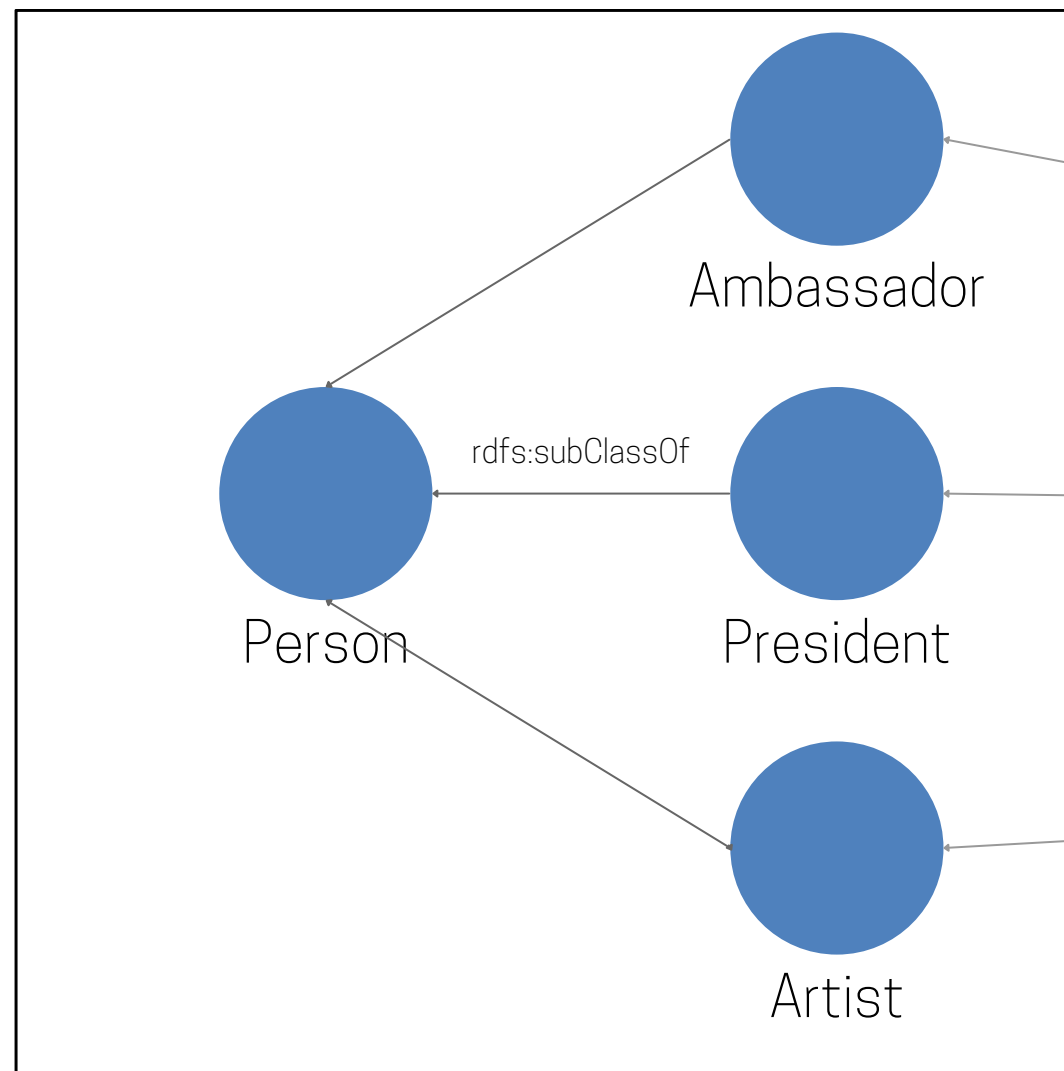
A formal representation of knowledge as a set of concepts within a domain, defining the relationships among those concepts to enable shared understanding and data interoperability across applications.

Instance Data

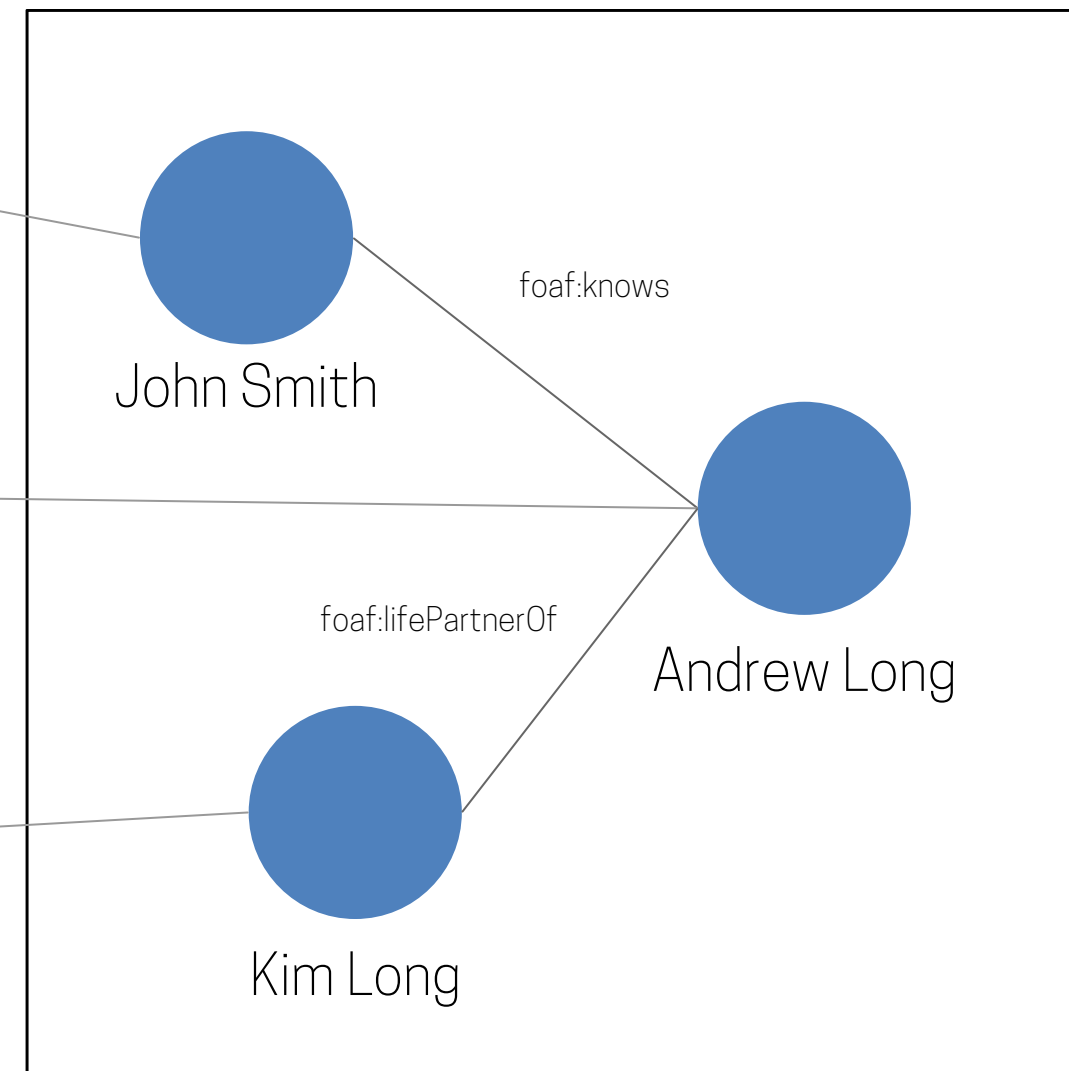
Specific examples of data that instantiate the concepts and relationships defined in a domain ontology, representing real-world entities and their interconnections.

Domain Ontology & Instance Data

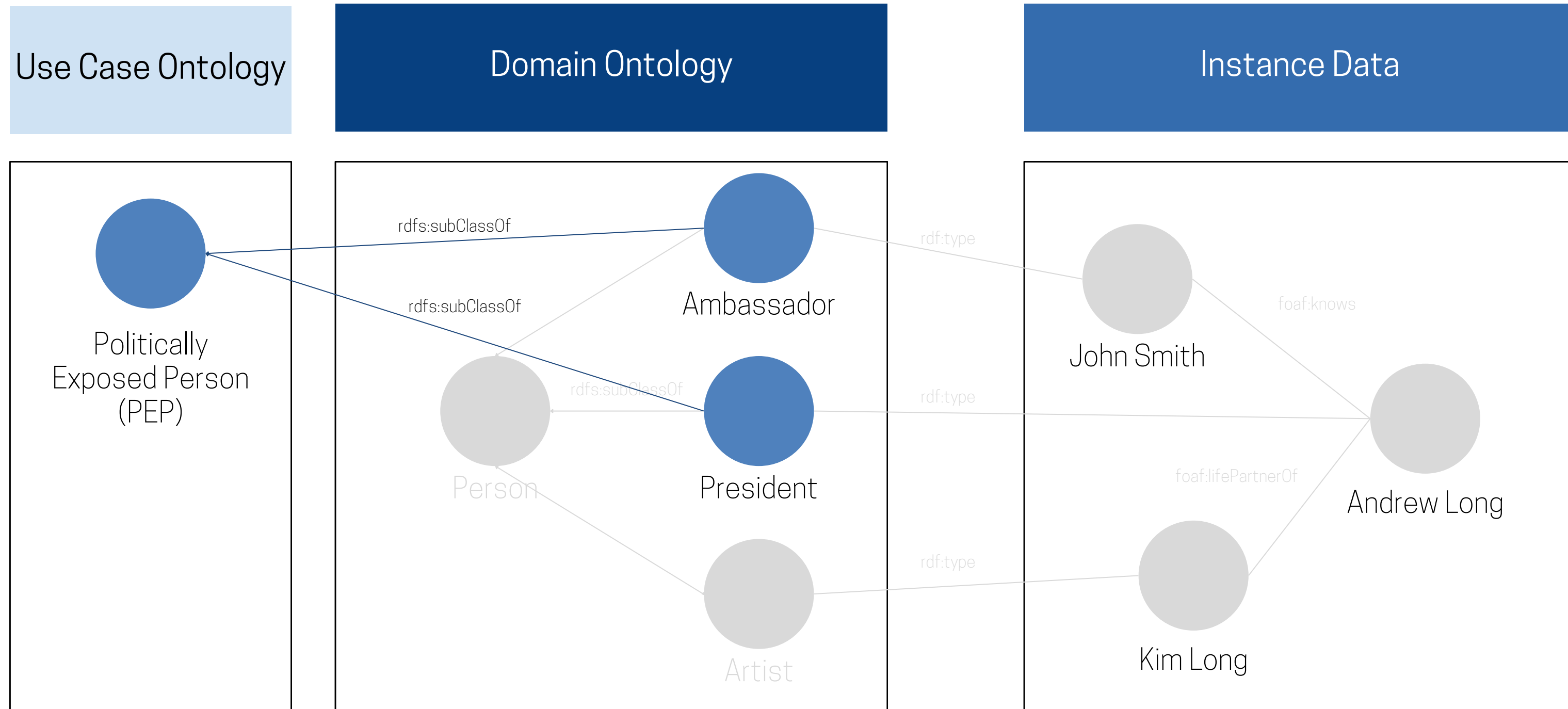
Domain Ontology



Instance Data



Use Case Ontology

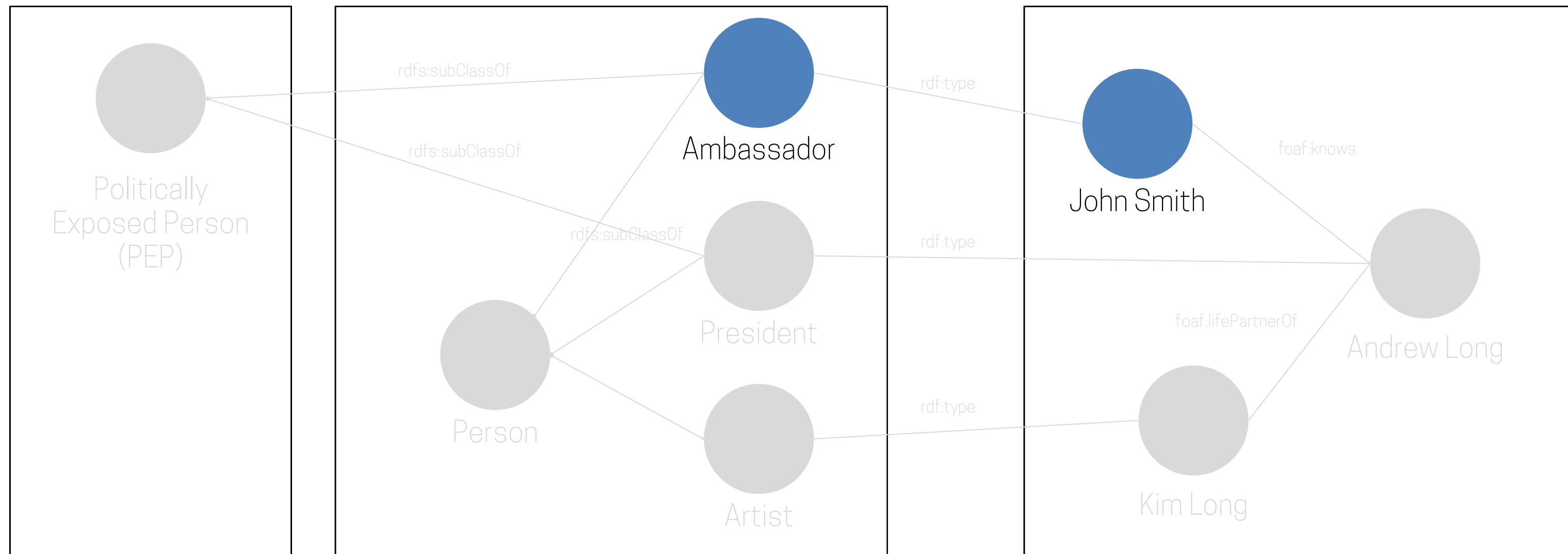


Inference

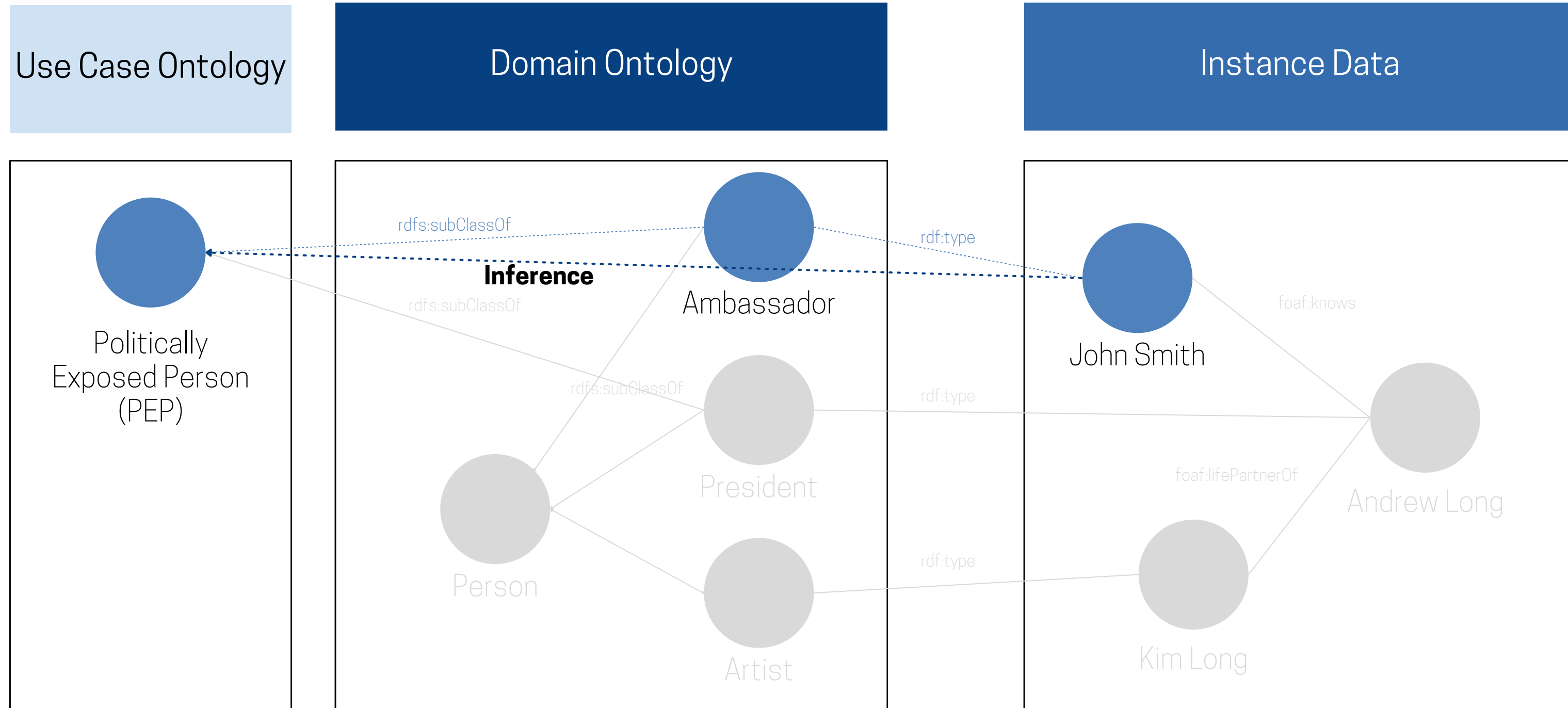
Use Case Ontology

Domain Ontology

Instance Data

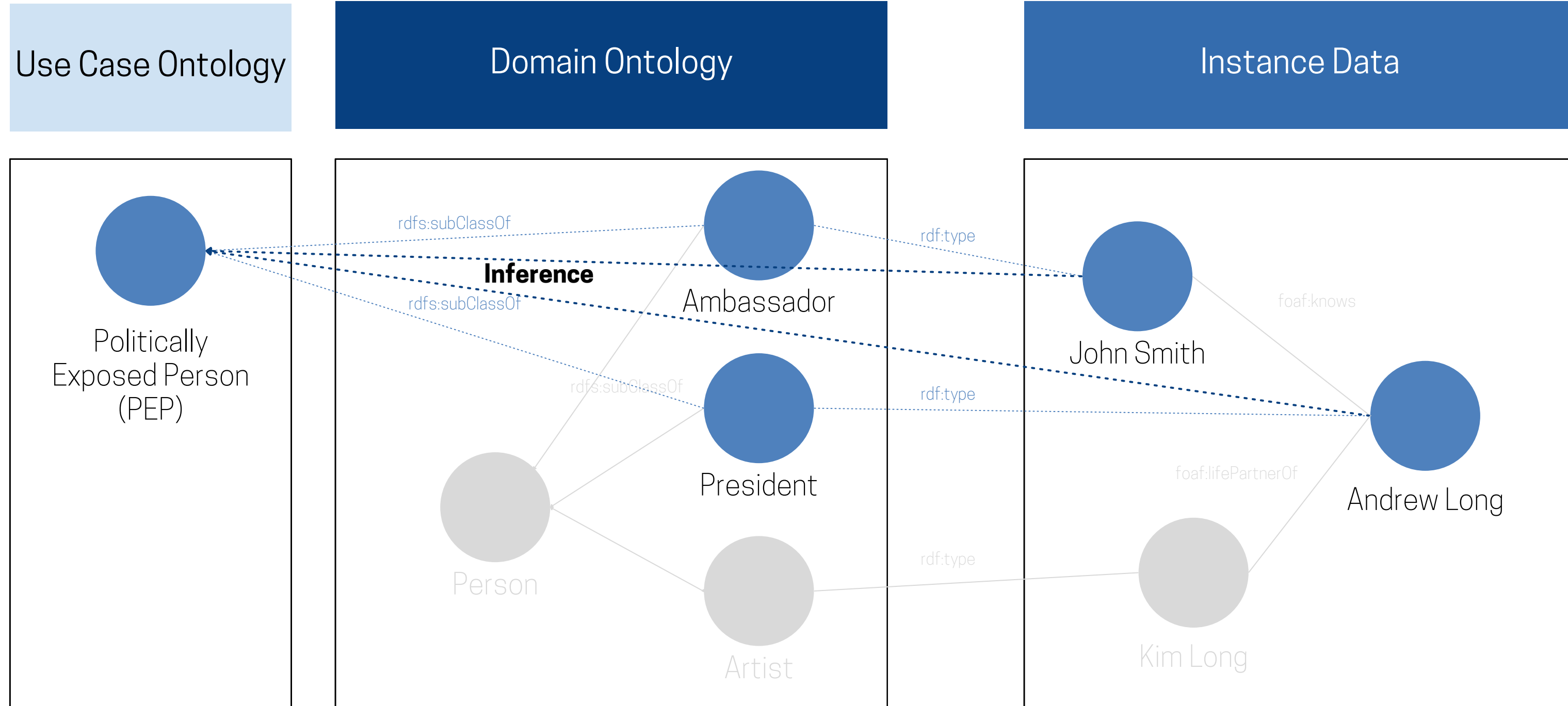


Inference



Inference:
John Smith is a **Politically Exposed Persons (PEP)**
because he is a **Ambassador** and custom ontology defines all **Ambassadors** as **PEPs**.

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

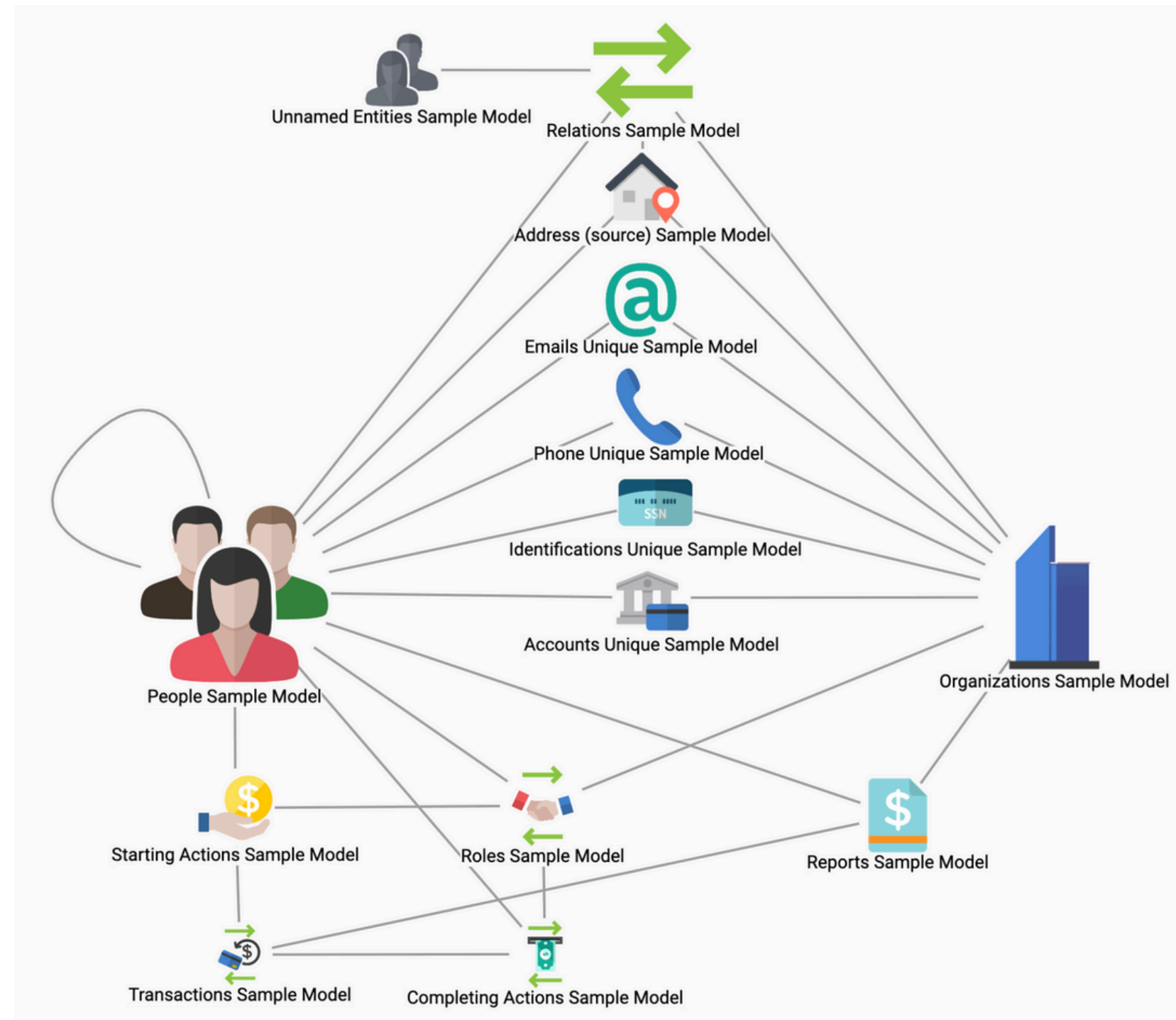


Implementation Process

Ontology Creation

Codifying Expert Knowledge

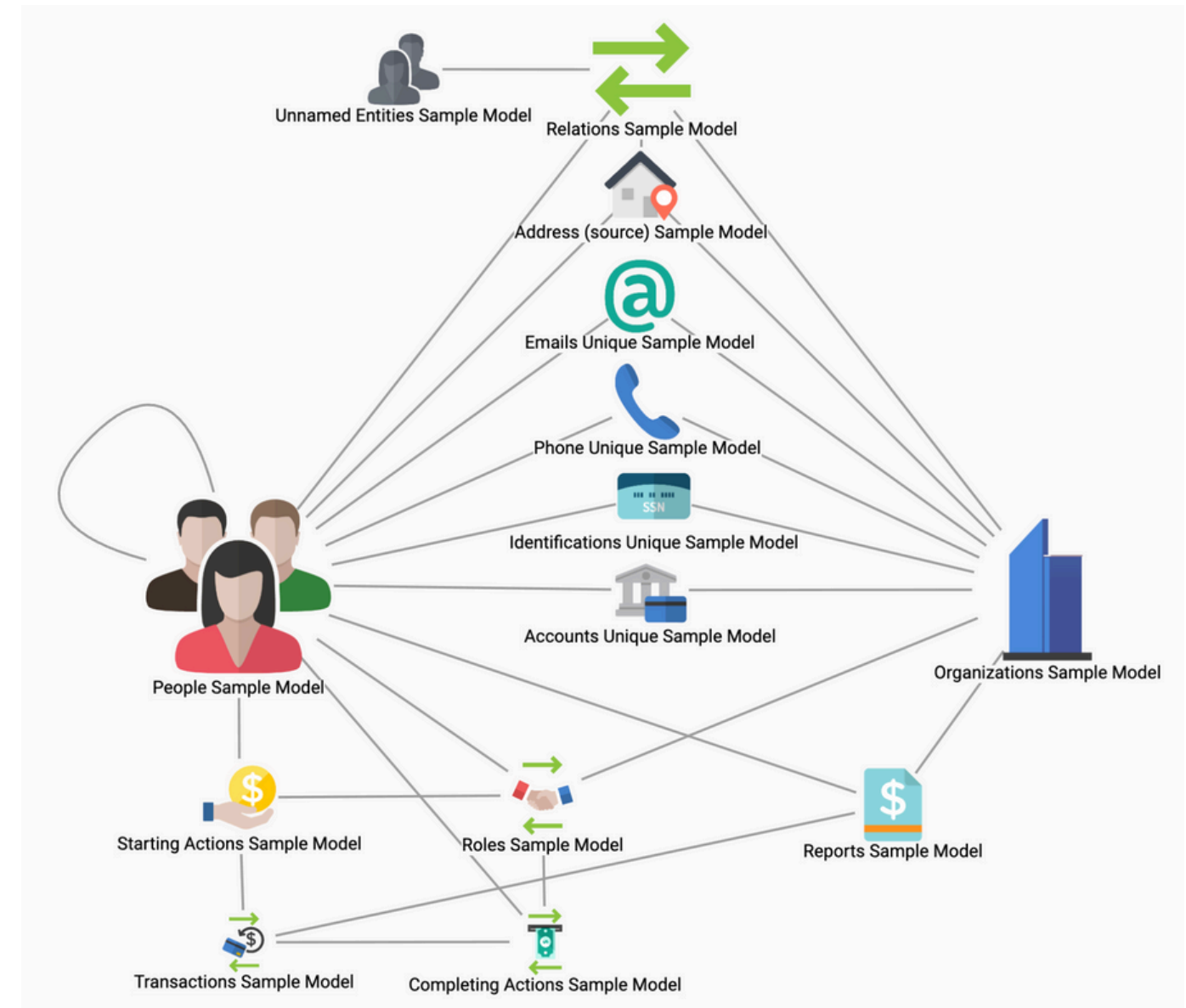
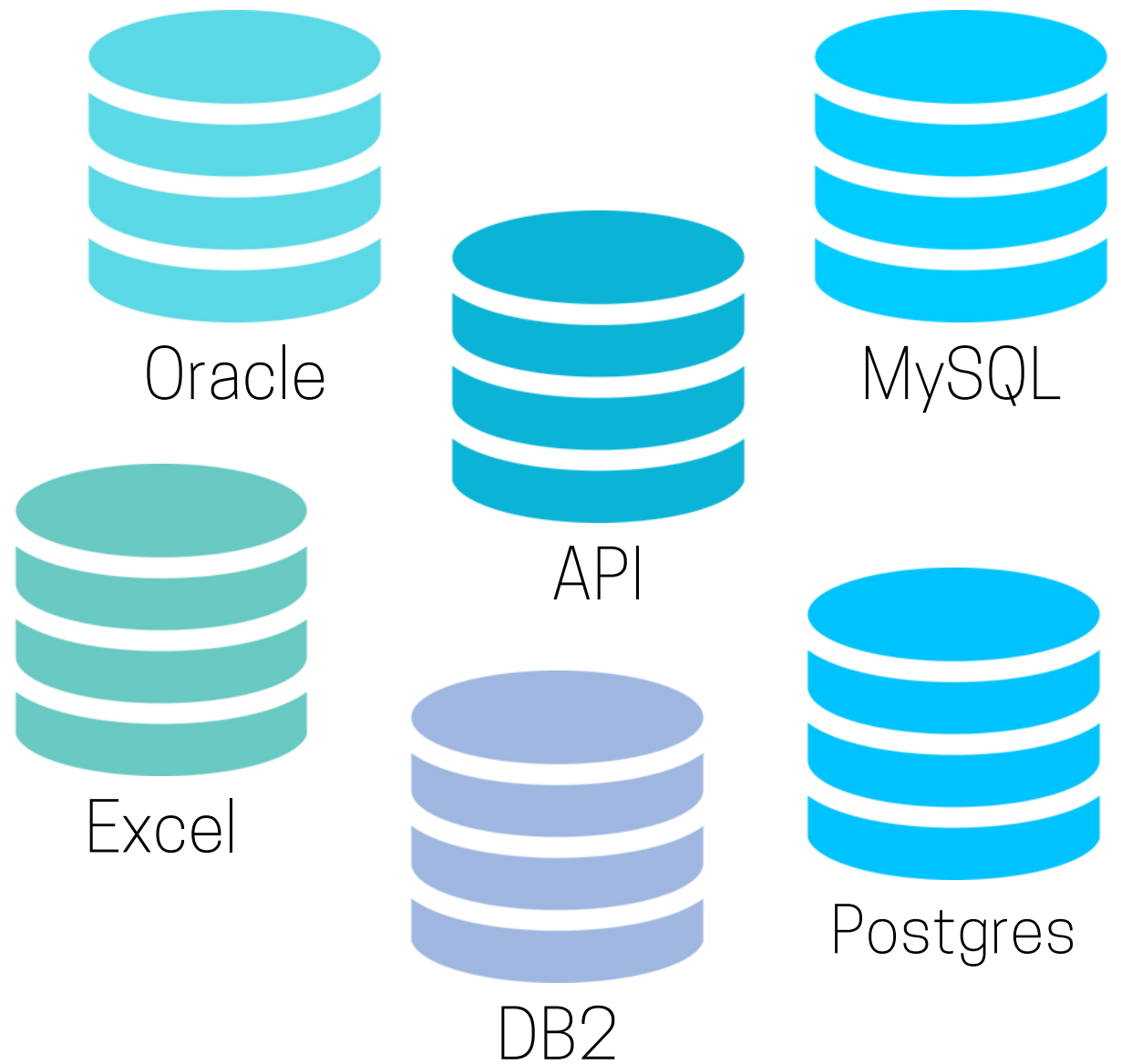
 *Competency Queries*



*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



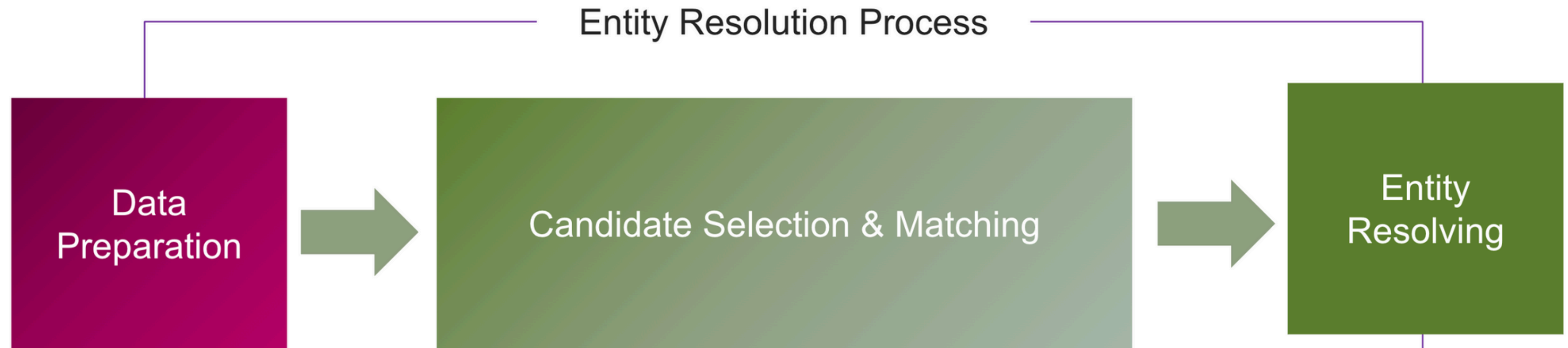
Phone Number Parser

*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

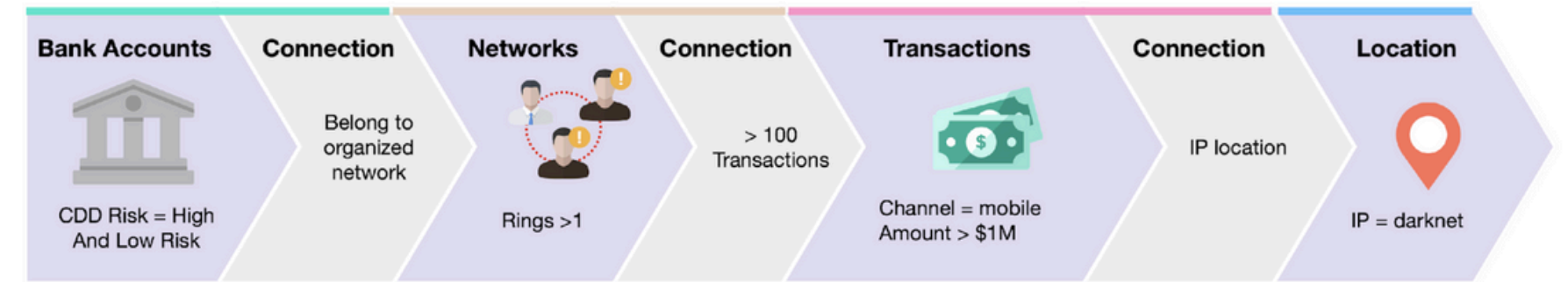
Implementation Process

Analytics

The screenshot shows a comprehensive profile for John Nelson Smith. It includes identifying information such as a mugshot, physical characteristics (DOB: 5/6/92, Height: 6'1", Weight: 220, Hair: BLK, Eyes: BLU), and contact details. Key sections include:

- REGISTERED WEAPONS?** NO
- OUTSTANDING WARRANT?** YES
- ARMED & DANGEROUS?** YES
- ADDRESS HISTORY:** A table with columns for From, To, Street, Apt, City, and State, listing addresses in Atlanta, GA and Miami, FL.
- ARREST HISTORY:** Includes entries for Battery (3/1/12) and ADW (4/2/15) in Atlanta, GA.
- KNOWN ASSOCIATES:** Lists Bill Smith and Jan Bryan.
- VEHICLES:** Lists a Ford F150 (2008) with GA license 1234567.
- LINKED PLACES:** A map showing various locations.

360 Summary Dashboard



Four summary cards displaying query results:

- Outstanding Warrants:** # 778 Objects
- Part 1 Crimes Last 24 Hours:** # 778 Objects
- Part 2 Crimes Last 24 Hours:** # 428 Objects
- Belong to organized network with tx > \$1M, high and low risk:** # 212 Objects

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

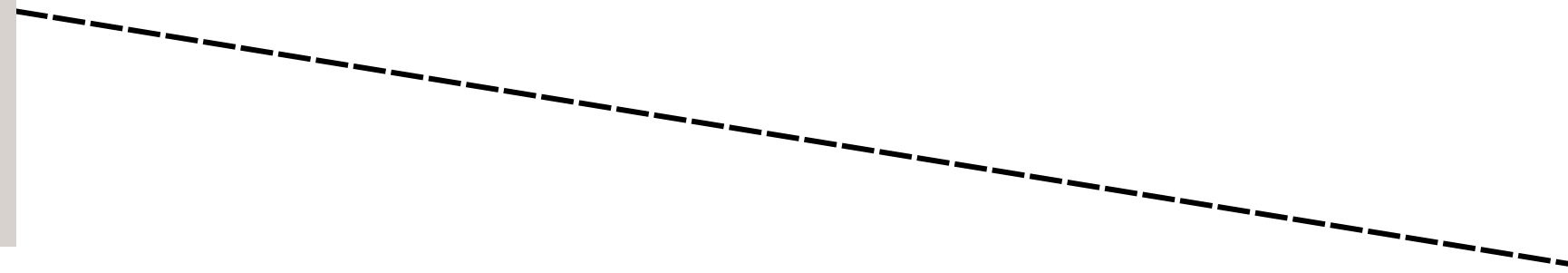


Andrew
Long

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

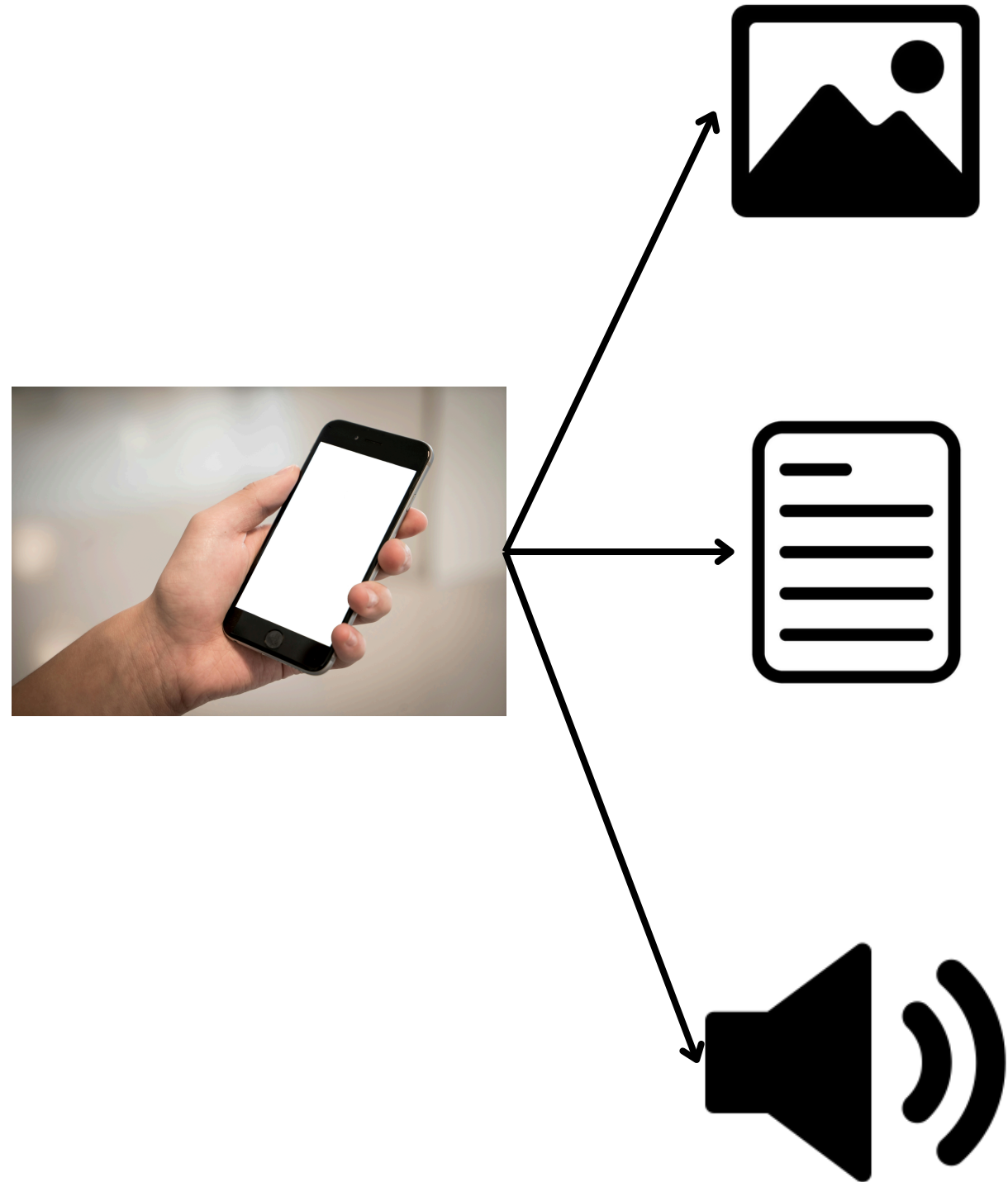
Is there a path between those two people?

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

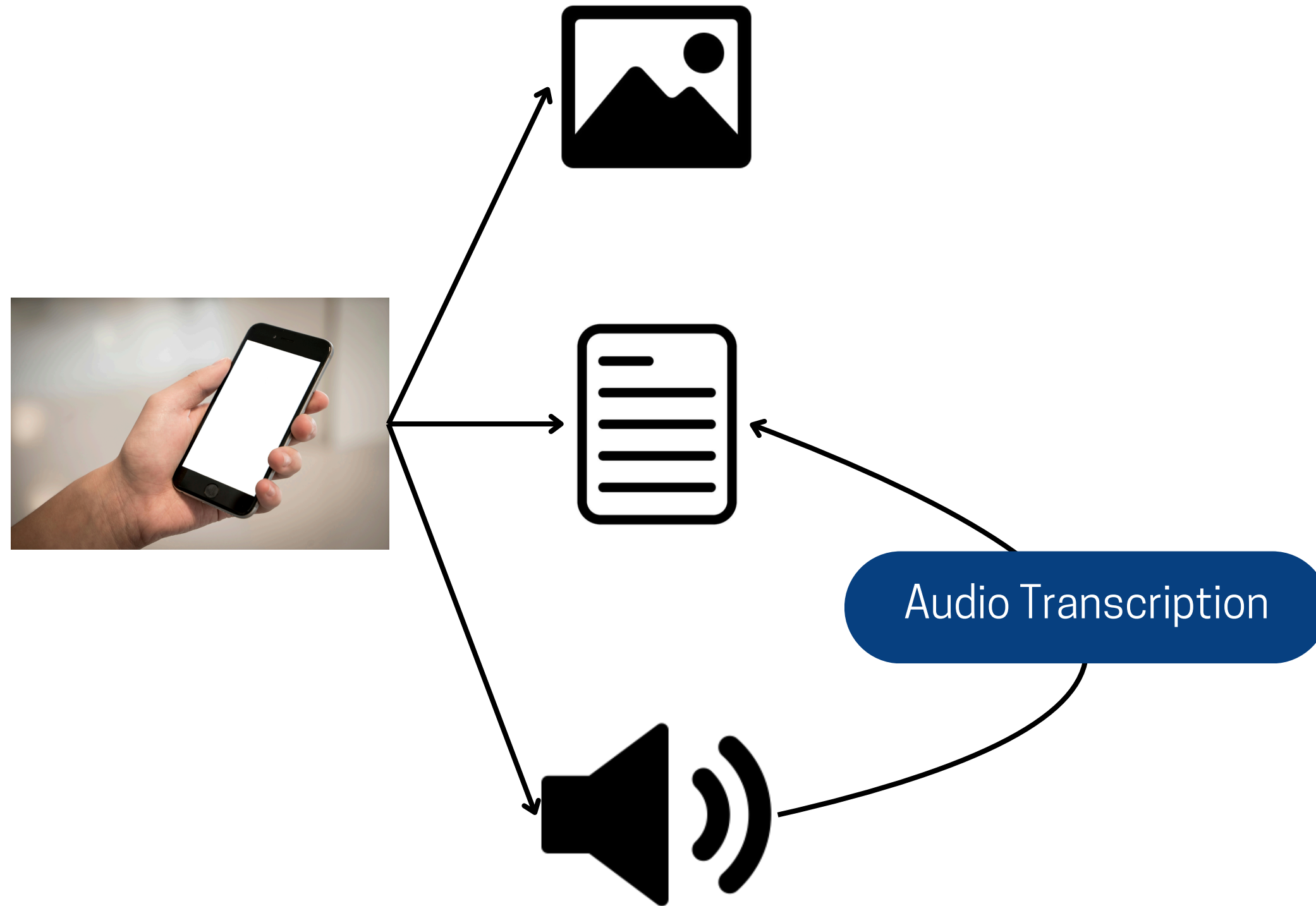
Effect?

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

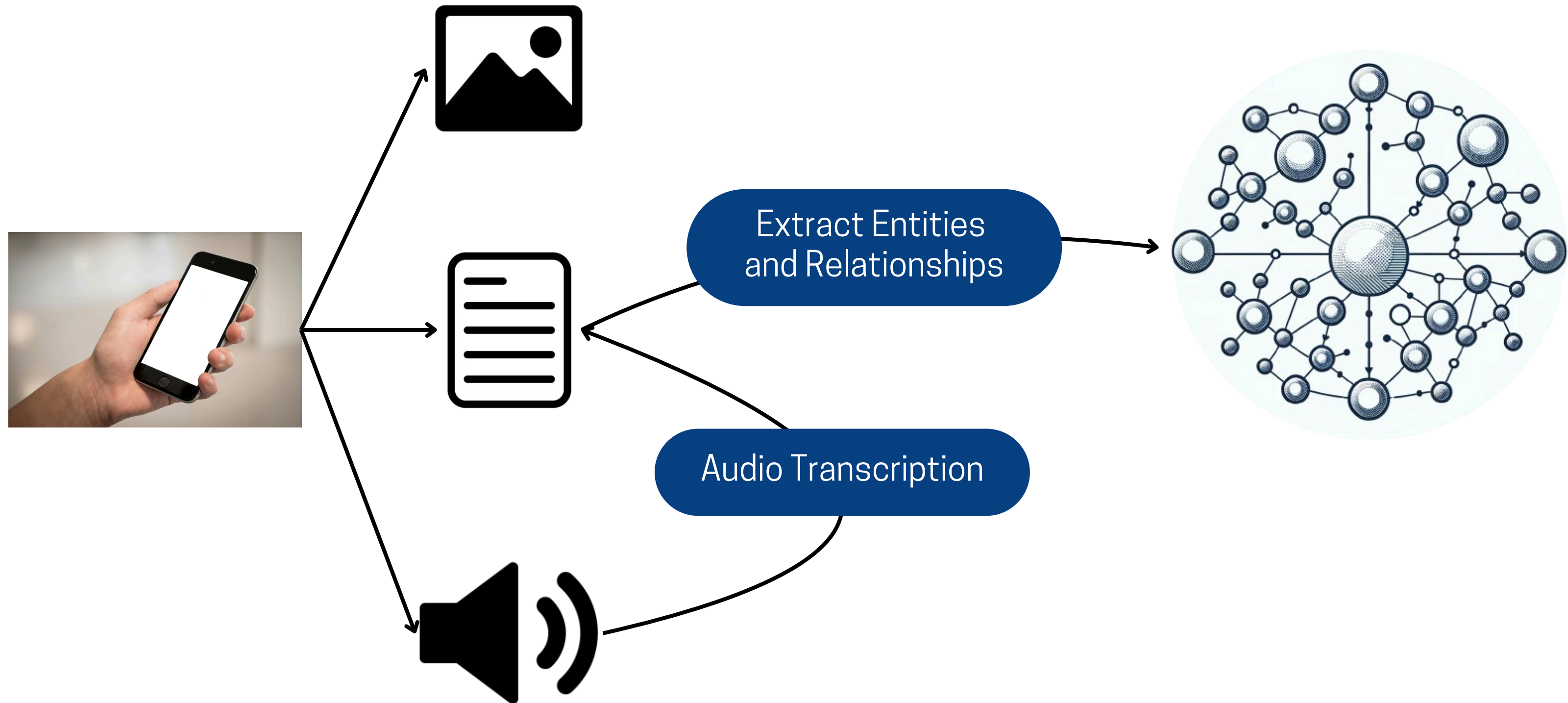
AI for Law Enforcement



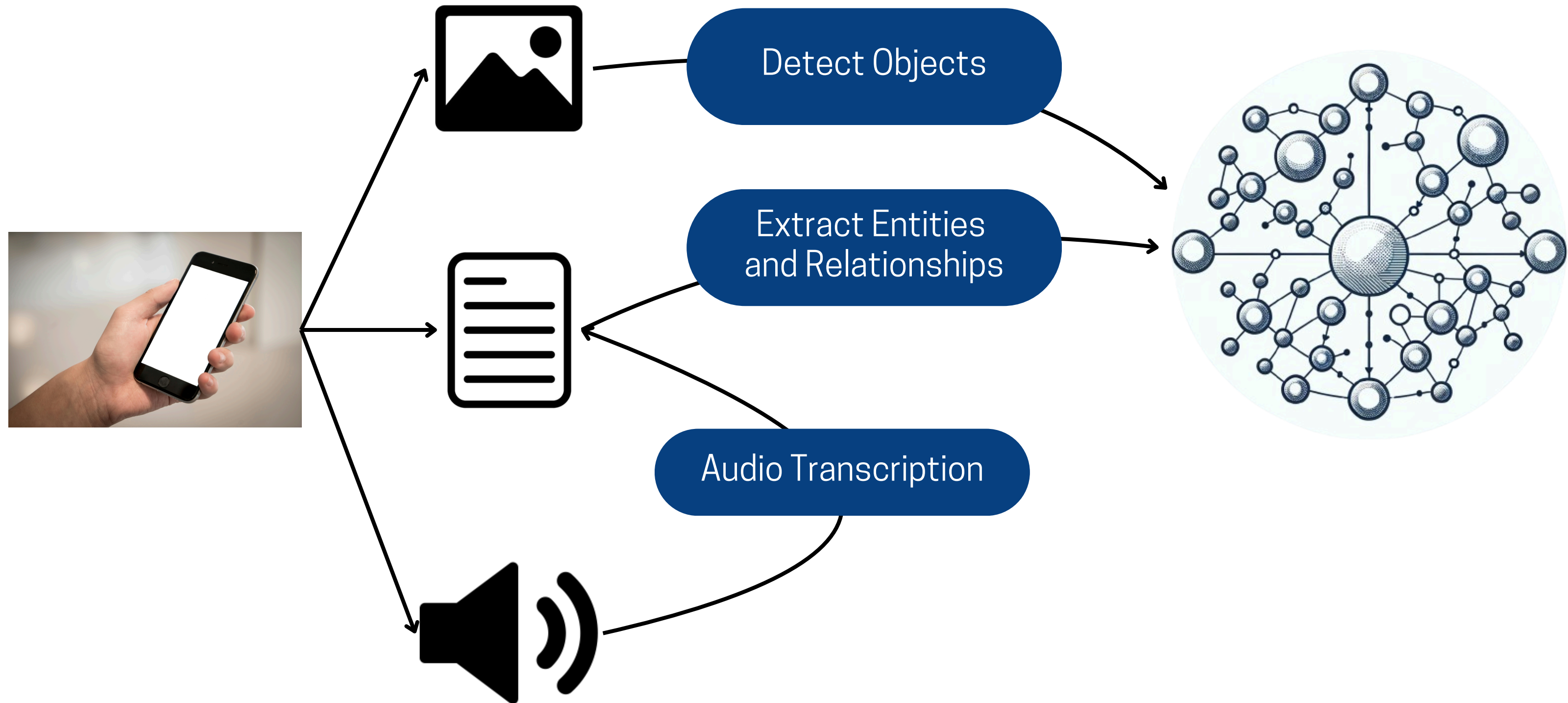
AI for Law Enforcement



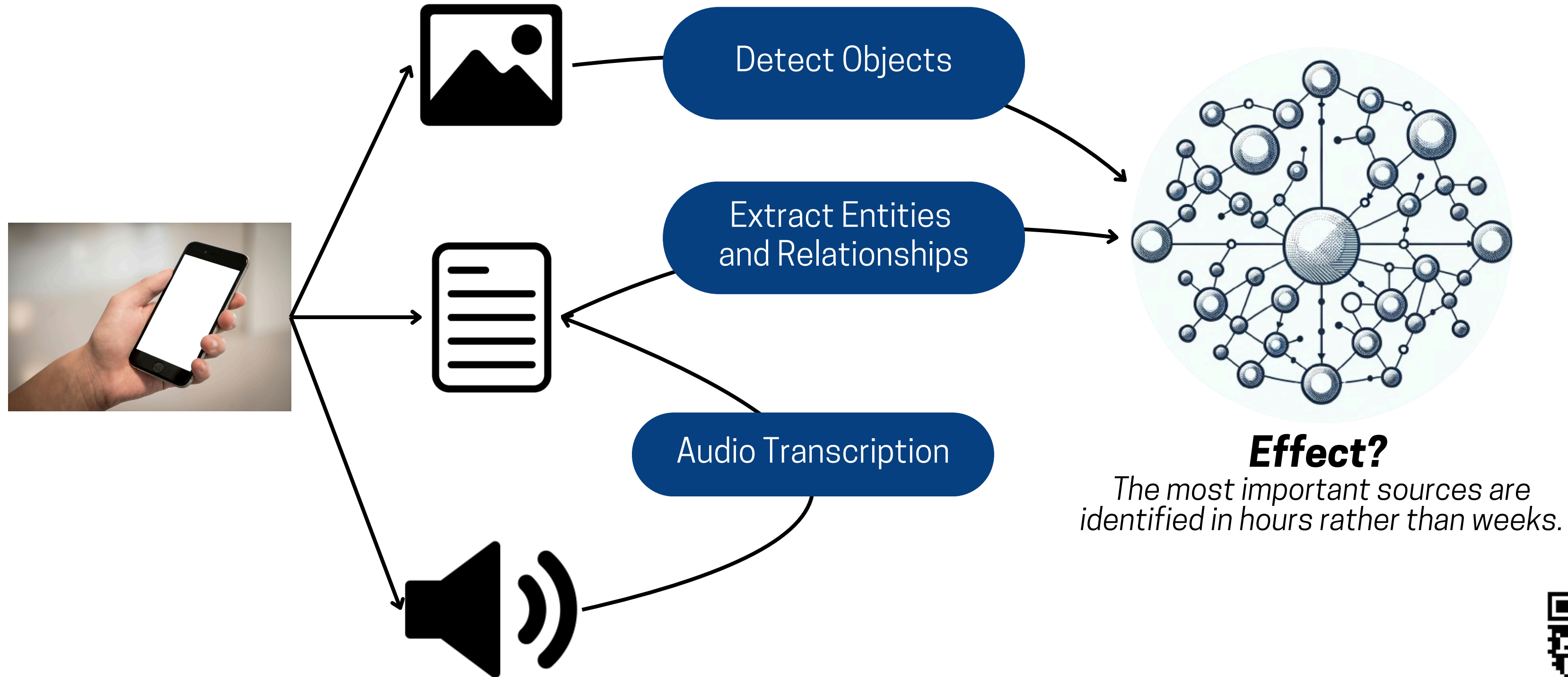
AI for Law Enforcement



AI for Law Enforcement



AI for Law Enforcement



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!

