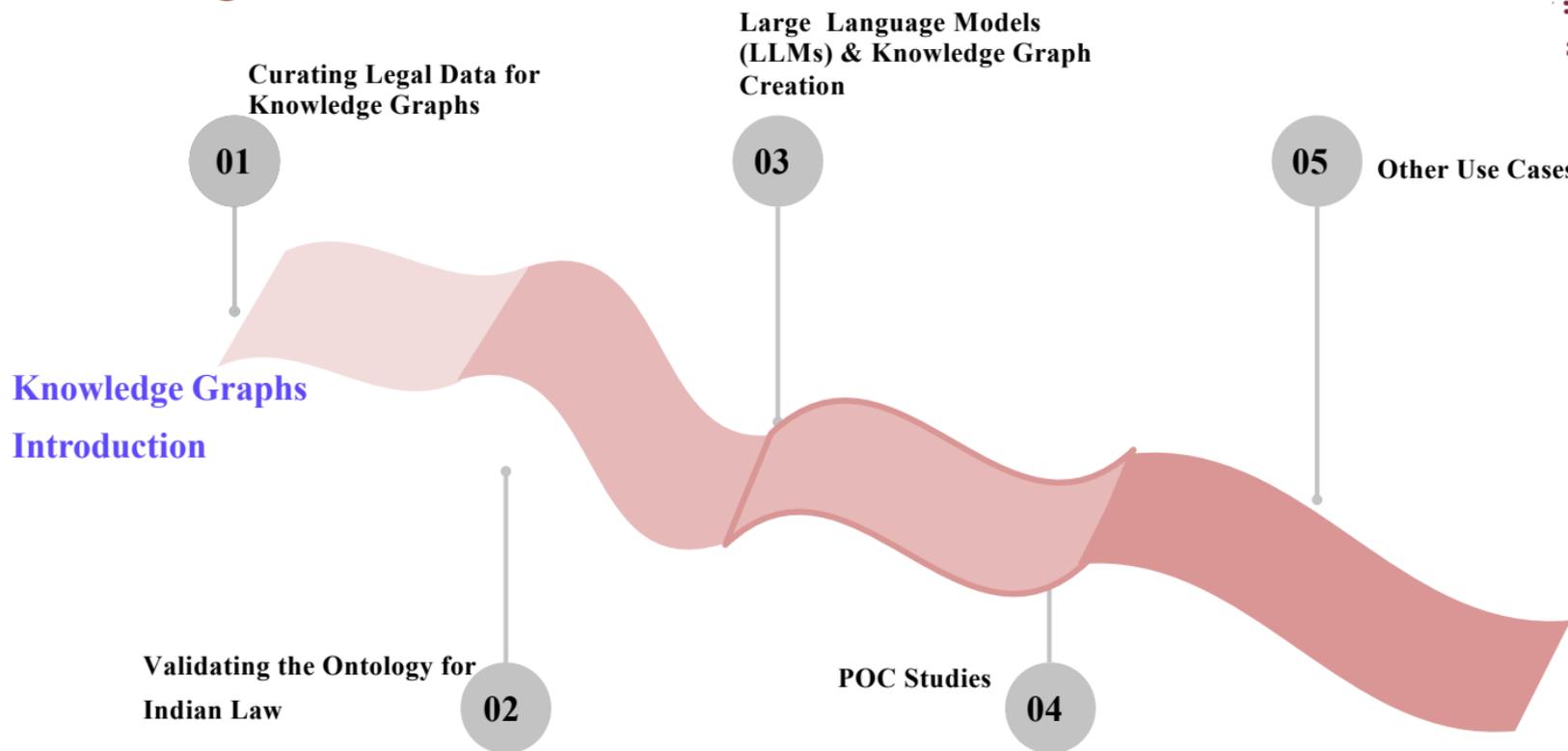




Building  
Legal Knowledge Graphs  
for India

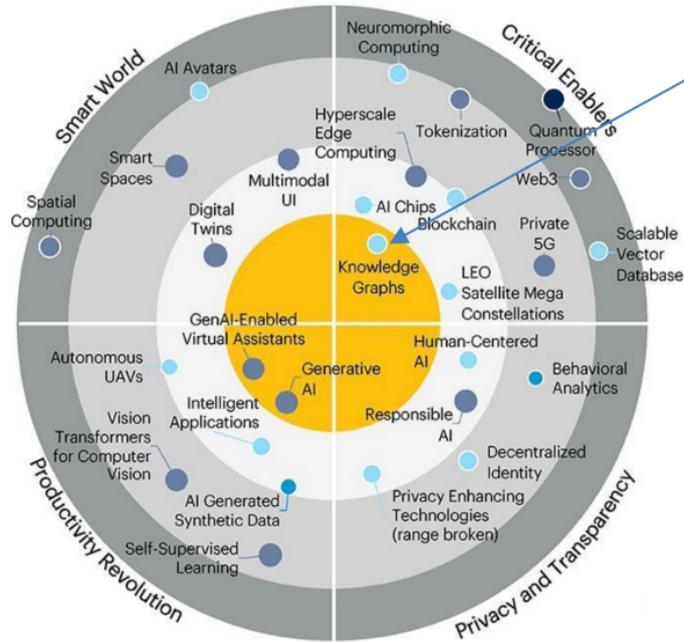
# Agenda



# Knowledge Graphs

- Knowledge graphs are collections of facts (called "triples"), **made for computers**
- **Triples** are composed of a predicate that links one concept with another or with a value
  - **Relation predicates** link two concepts and give the collection its graph structure
  - **Attribute predicates** link a concept and a value
- Concepts and predicates are (self-)defined by facts: i.e., their semantics is **grounded** in the other triples that contain the concept (its "graph neighborhood")
- There are four common **types** of knowledge graphs:
  - **Taxonomies** (many individual entities, only one entity category is focused, one kind of predicate)
  - **Ontologies** (no individual entities, many category entities, many kinds of predicates)
  - (simple) **Knowledge Graphs** (many individuals, many categories, many kinds of predicates)
  - **Knowledge Hypergraphs** (many individuals and categories, predicates that link two or more entities)

# Emerging Tech Impact Radar for 2024



Gartner's 2024 Emerging Tech Impact Radar report identified **knowledge graphs** as the key software enabler and important investment to enable generative AI adoption.

**Knowledge graphs** are *critical enablers* since it –

- 1) reduces cost of LLM usage,
- 2) improves accuracy of LLMs,
- 3) provides domain-specific terminology for LLMs,
- 4) generates up-to-date responses; and
- 5) ensures transparency and interpretability.

# Agenda



**Curating Legal Data for Knowledge Graphs**

**01**

**Large Language Models (LLMs) & Knowledge Graph Creation**

**03**

**05**

**Other Use Cases**

**Knowledge Graphs Introduction**

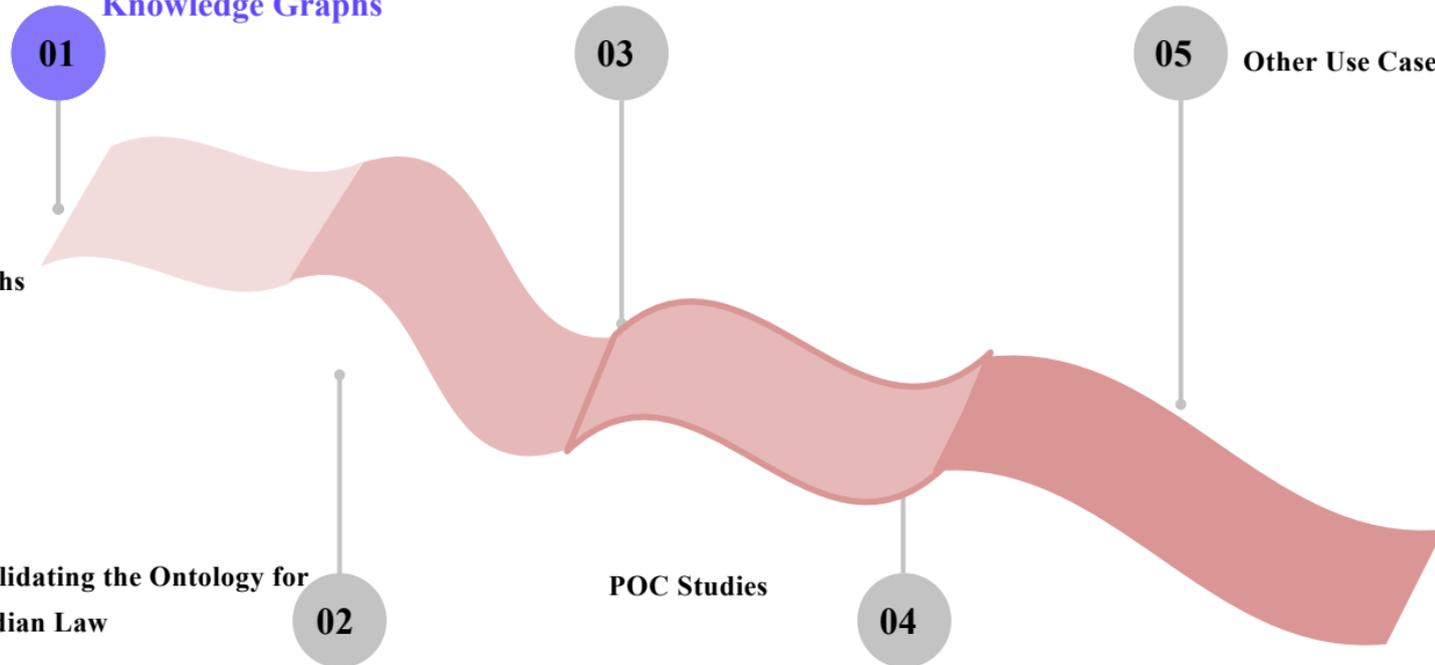
**Validating the Ontology for Indian Law**

**02**

**POC Studies**

**04**

**Governance**

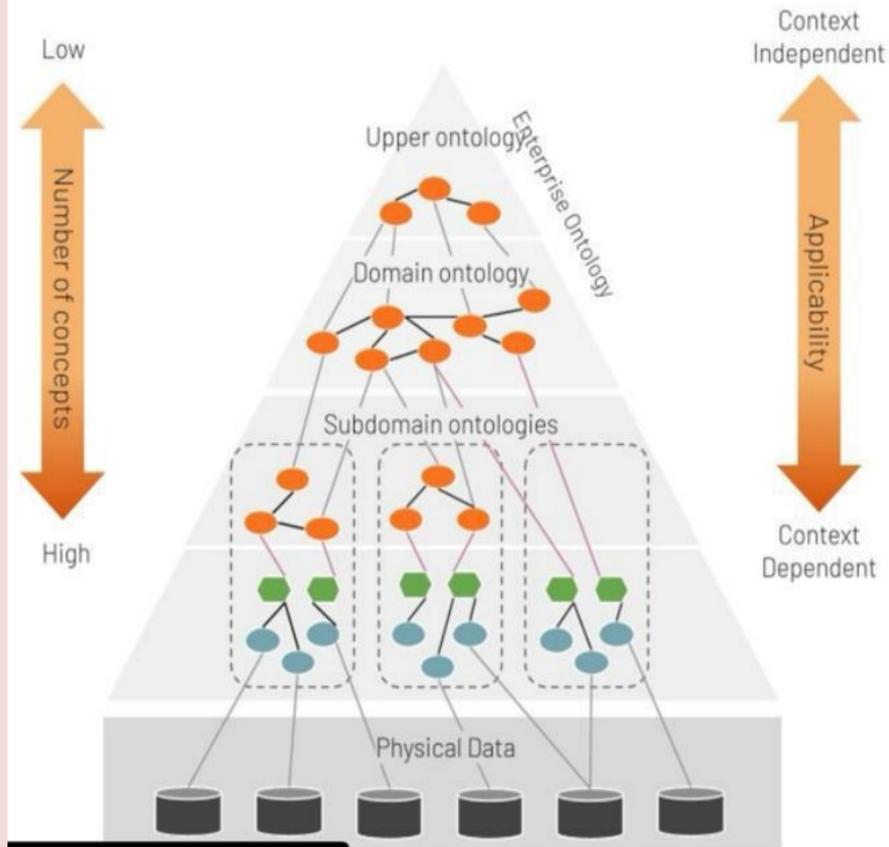


# Stage 1

## Curating Data for Legal Knowledge Graphs

Physical Data products like judgments, Acts, Rules, Bills, circulars, notifications, etc. are pivotal elements in the incremental and distributed construction of a knowledge graph.

1. Aggregate technical metadata
2. Create the semantic link between physical data asset and business concepts modeled in the SALI ontology



ONTOLOGY + DATA = KNOWLEDGE GRAPH

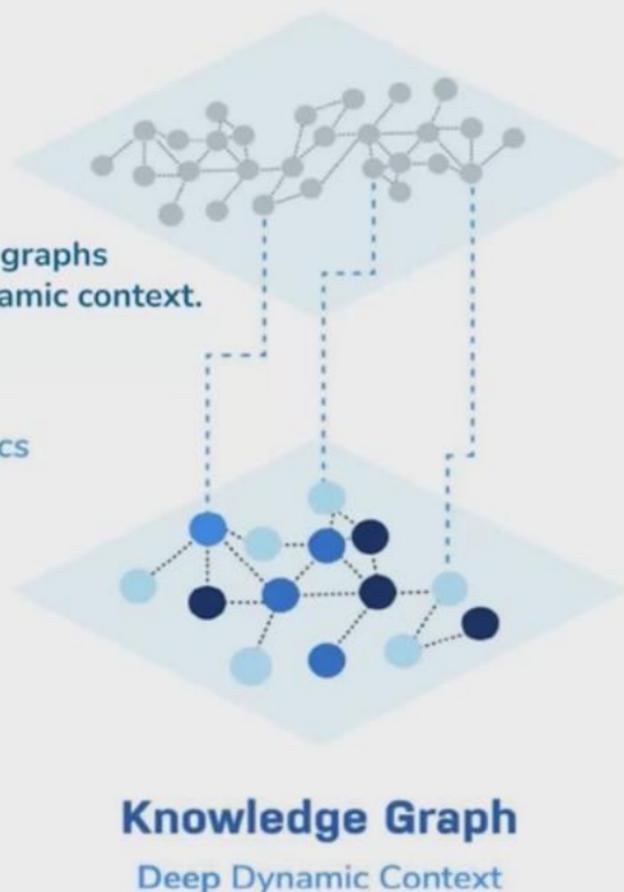
# Knowledge Graphs Give you Facts

Ontology	Data	Knowledge Graph
Defines features, attributes, and relationships from the real world	Key Value pairs of events that happened or relationships that exist	Real world data mapped to an ontology that computers can use for reasoning



Knowledge graphs provide deep, dynamic context.

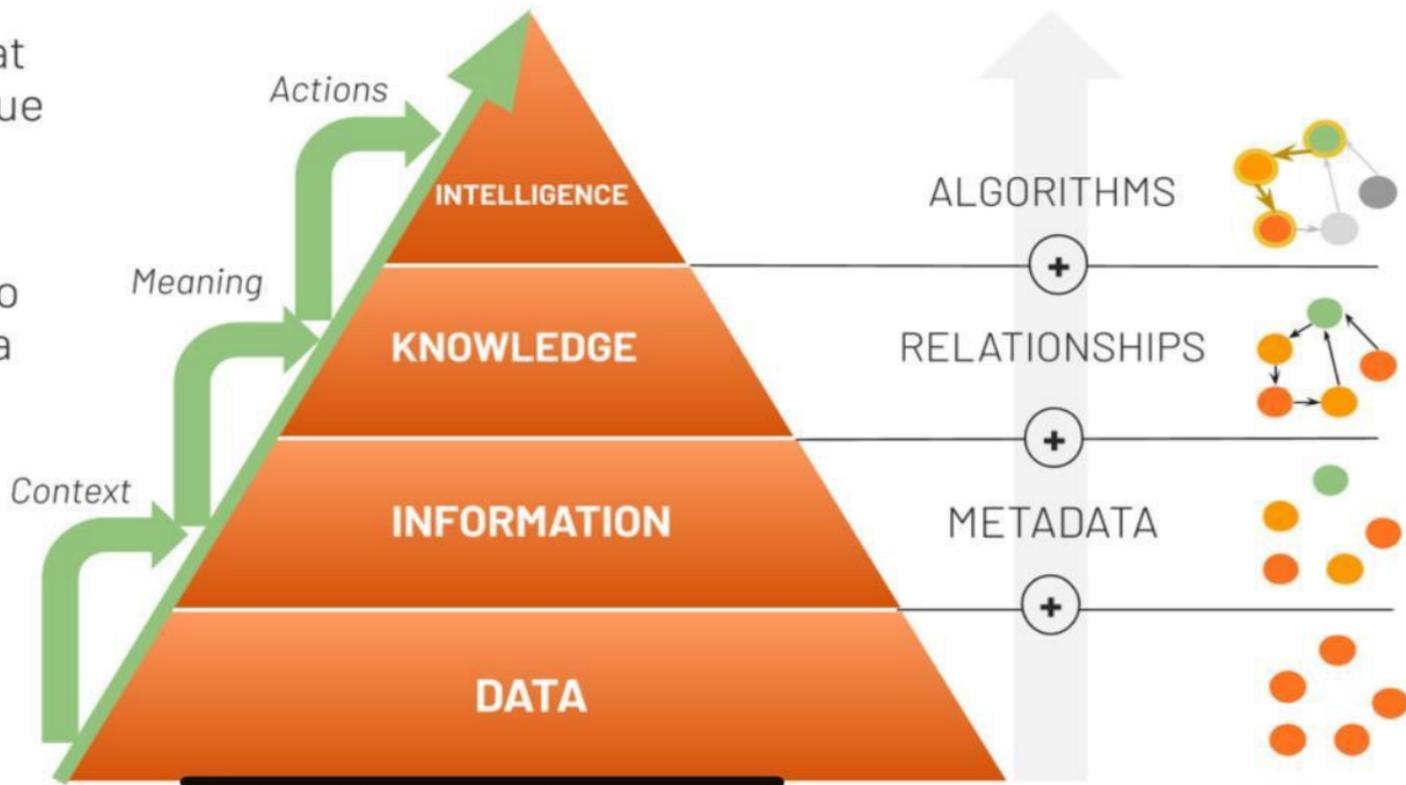
**+**  
Semantics



# Information architecture

Data is an asset that only unlocks its value when put to use.

Data management cannot be limited to just managing data



# Agenda



Curating Legal Data for Knowledge Graphs

01

Large Language Models (LLMs) & Knowledge Graph Creation

03

05

Other Use Cases

Knowledge Graphs Introduction

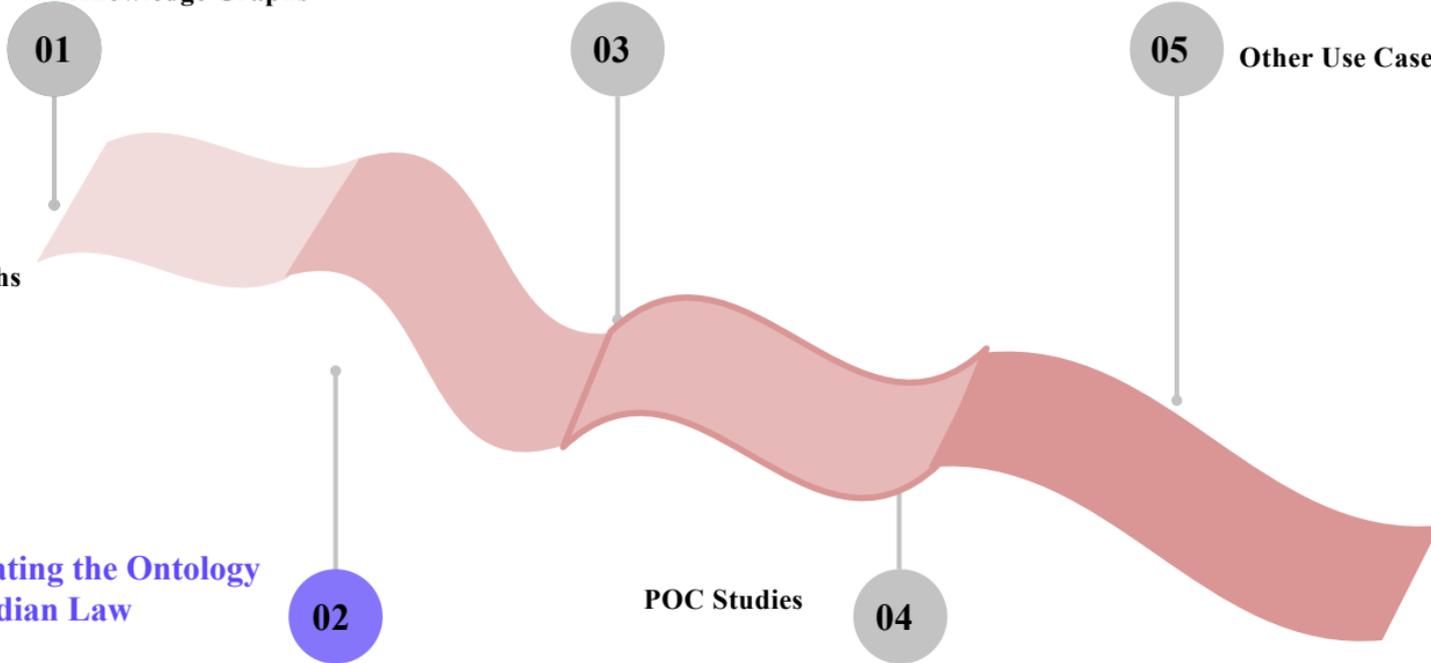
Validating the Ontology for Indian Law

02

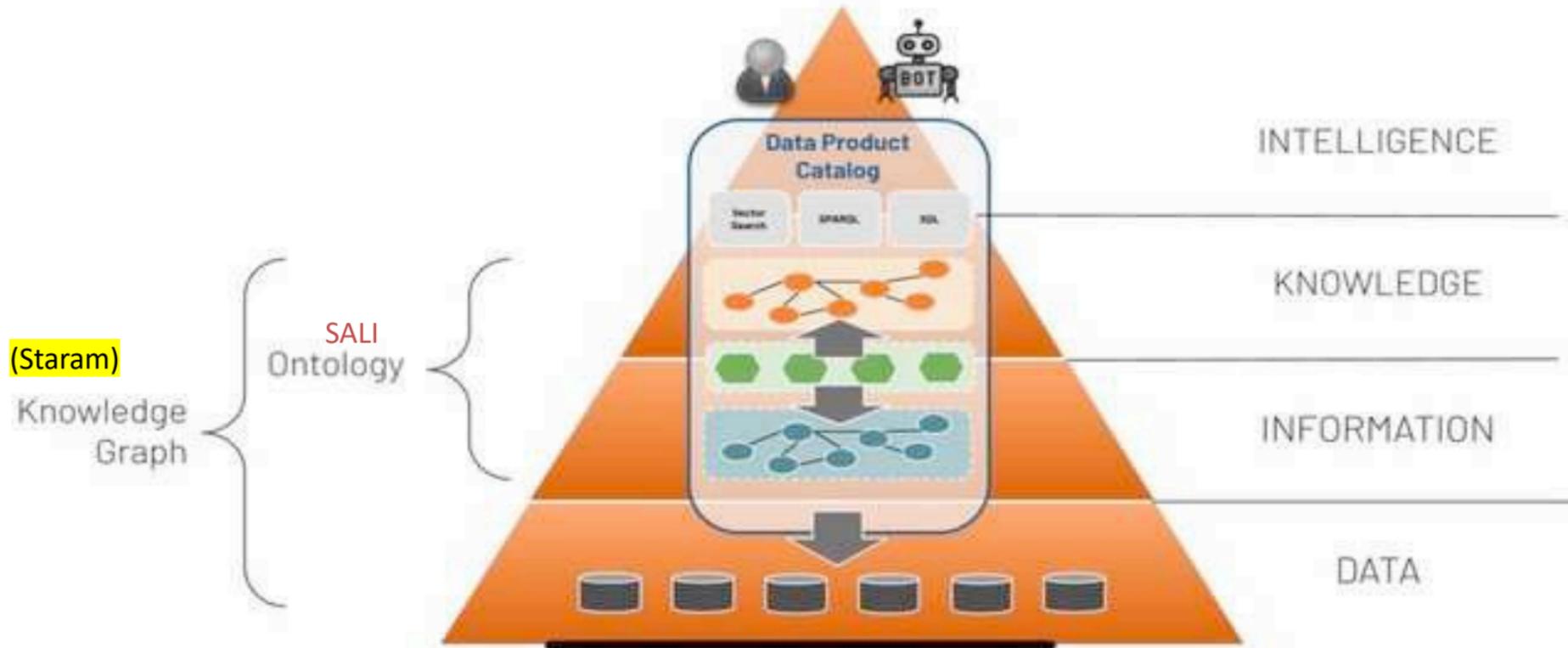
POC Studies

04

Governance



## Stage 2- Validating the Ontology



## Data Sources



Unstructured

## LLM Graph Builder



LLMs

## Knowledge Graph



Neo4j Aura



Bloom



Graph DB



Graph Data Science

## Chat w/ GraphRAG



LLMs

## End User App / UseCase

Enterprise Knowledge Search

Customer Service

Ticket Triaging

Recommendations

News Content &  
Discovery

FAQ Bots

Judgment Search

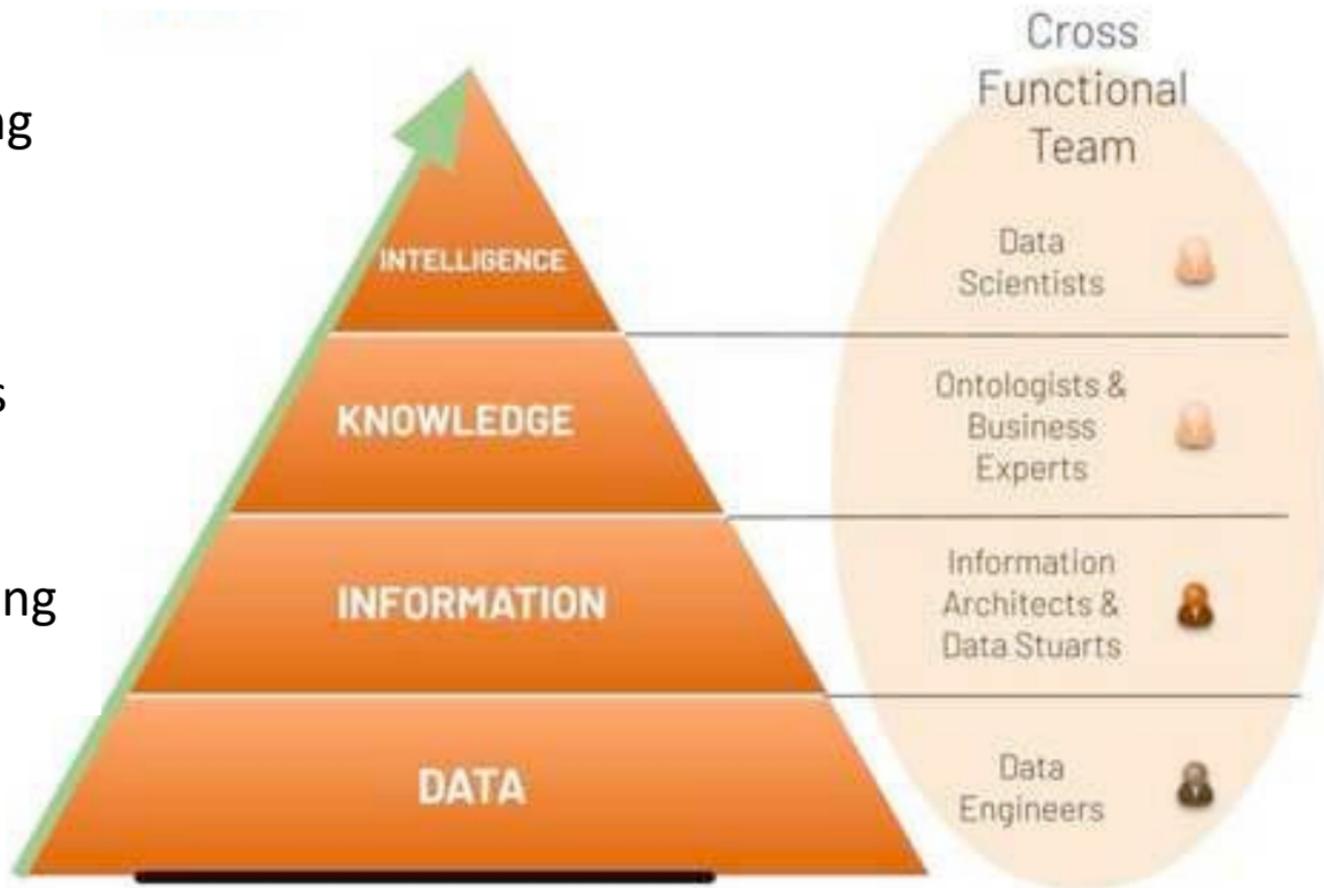
Predictive  
Analytics

Contracting

Compliance

## Validating the Knowledge Graphs

This involves cross functional teams having exposure to law and work in legal research/technology would validate the KGs created by LLMs. This requires a holistic approach understanding nuances of legal data

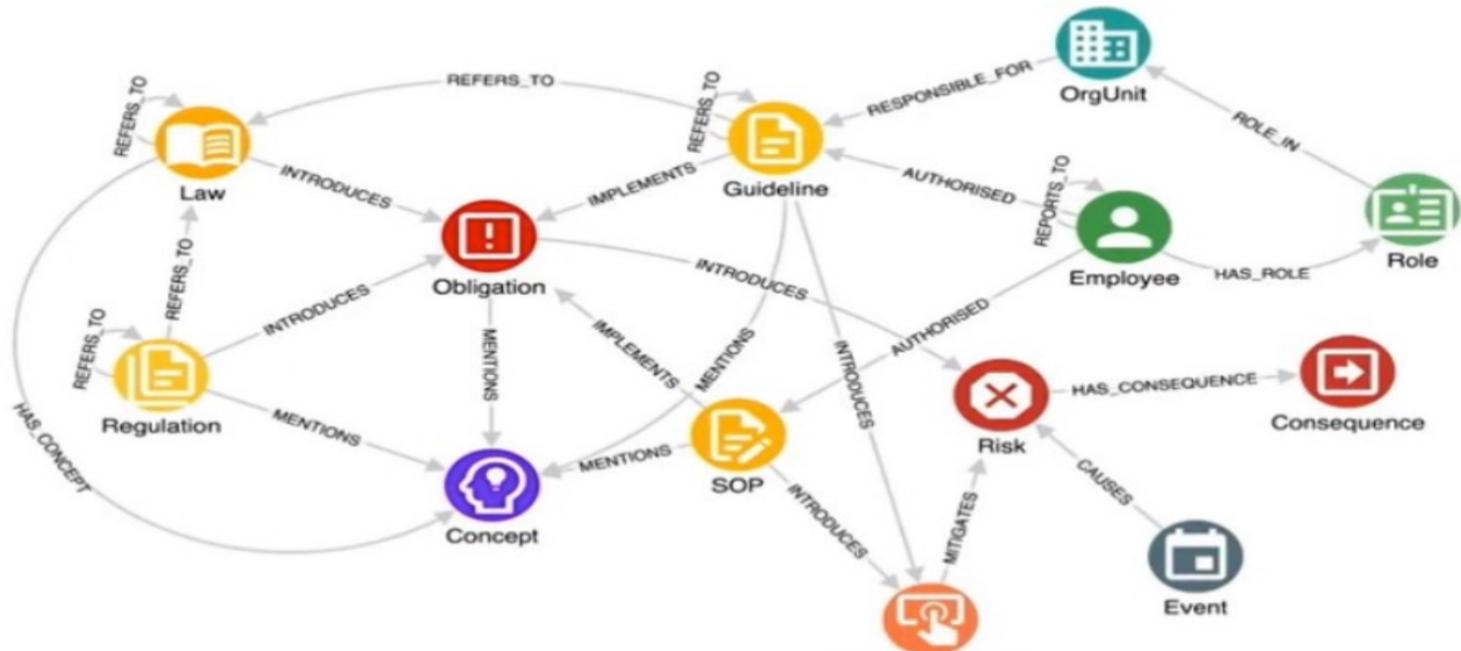


## Sample: Knowledge Graph for Compliance

-**Entities:** Laws, Regulations, Standards, Organizations, Processes

-**Relationships:** IsPartOf, CompliesWith, Governs, Enforces

-**Attributes:** Dates, Jurisdictions, Compliance Status



# Agenda



Curating Legal Data for Knowledge Graphs

01

Large Language Models (LLMs) & Knowledge Graph Creation

03

05

Other Use Cases

Knowledge Graphs Introduction

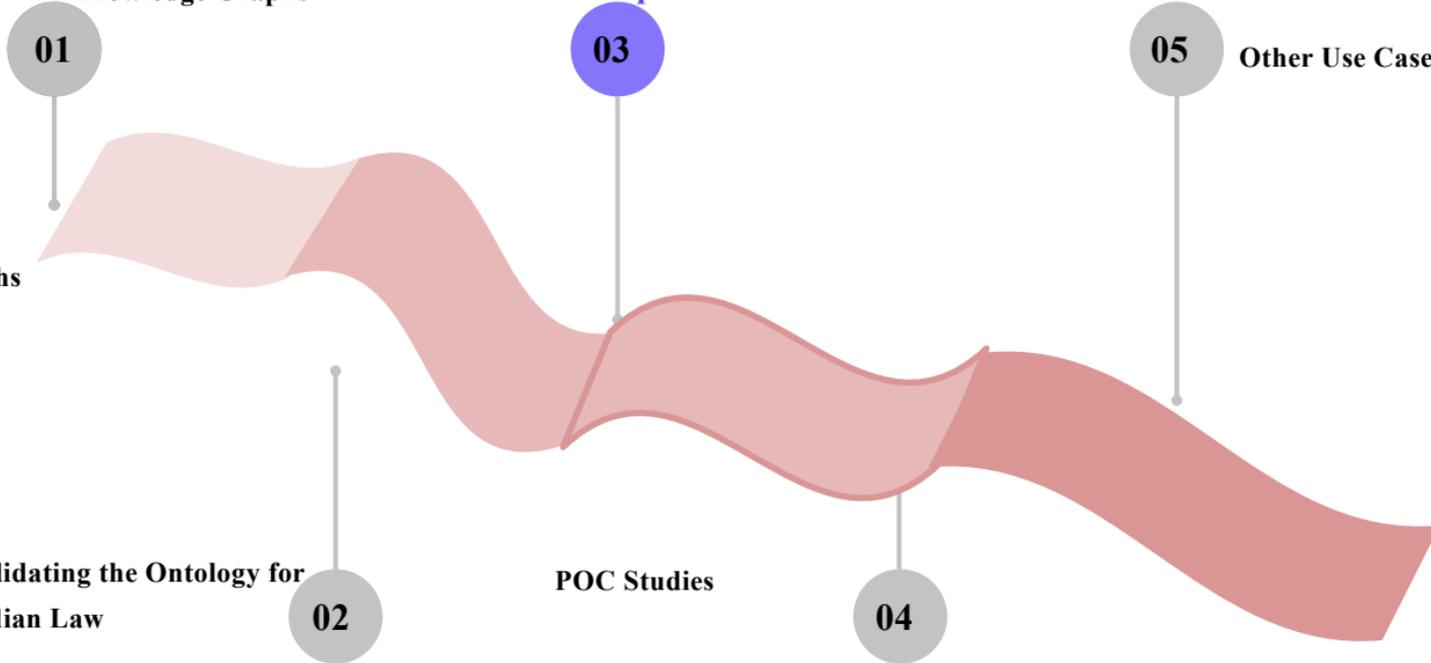
Validating the Ontology for Indian Law

02

POC Studies

04

Governance



## LLMs & Knowledge Graph Creation

We would use LLMs to speed the process of **generating knowledge graph triples** - i.e., a draft of a knowledge graph.

- LLMs can produce valid RDF, SKOS, AMR, SPARQL, etc. output
- LLMs can't tell us which concepts to focus on
- Accuracy of generated triples varies by predicate, so we still need human curation
- Coherence of related groups of triples varies, so we still need curation
- LLMs are like MT for translators: **they help a lot but aren't publication-ready**

# Unifying KGs and LLMs – 3 Frameworks

Unifying large language models (LLMs) and knowledge graphs (KGs) fit in **three main frameworks**:

## 1) KG-enhanced LLMs:

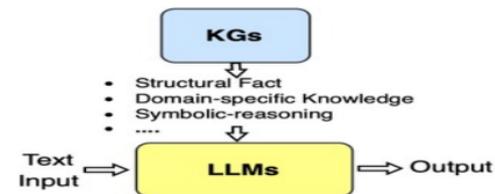
- Incorporating KGs into LLMs during **pre-training** or inference stages
- Using **KGs to analyze and interpret LLMs**
- Aims to improve LLMs' knowledge awareness, performance and interpretability

## 2) LLM-augmented KGs:

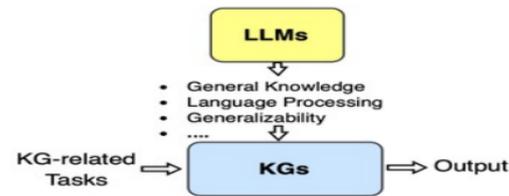
- Applying LLMs to **enhance various KG-related tasks** like **embedding, completion, construction, question answering**
- Aims to handle limitations of conventional KG methods in processing text, unseen entities etc.

## 3) Synergized LLMs + KGs:

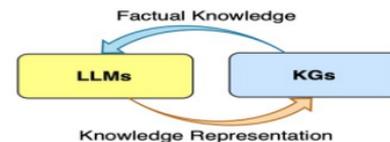
- Integrating LLMs and KGs into a unified model/framework
- *Mutual enhancement* of capabilities e.g. linguistic knowledge of LLMs + factual knowledge of KGs
- Aims for **bidirectional** reasoning combining strengths of both approaches



a. KG-enhanced LLMs



b. LLM-augmented KGs



c. Synergized LLMs + KGs

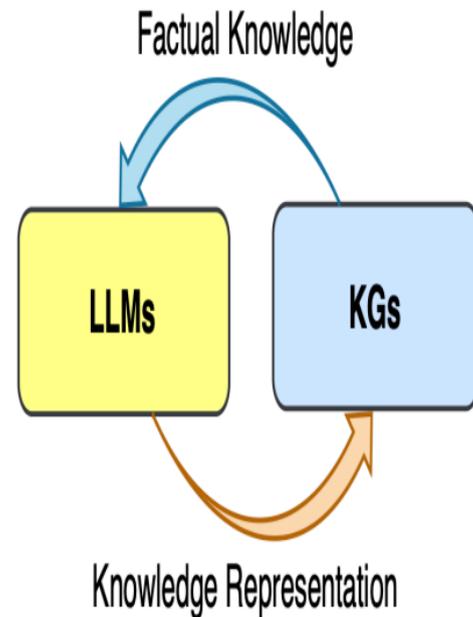
# Preferred Approach for POC - Synergizing LLMs and KGs

## For knowledge representation:

- **Jointly pre-train LLMs and KGs** to align their representations
- Methods like KEPLER and JointGT propose joint objectives to embed text and KGs in a shared space

## For reasoning:

- Apply LLMs and KGs together to **combine their reasoning strengths**
- Use LLMs to process text and retrieve relevant KG facts, then reason over facts using KG structure
- In logical reasoning, LLMs can generate logic queries executed on KGs, with results fused back into LLMs



**c. Synergized LLMs + KGs**

# Rapidly Build a Knowledge Graph for AI Use Cases



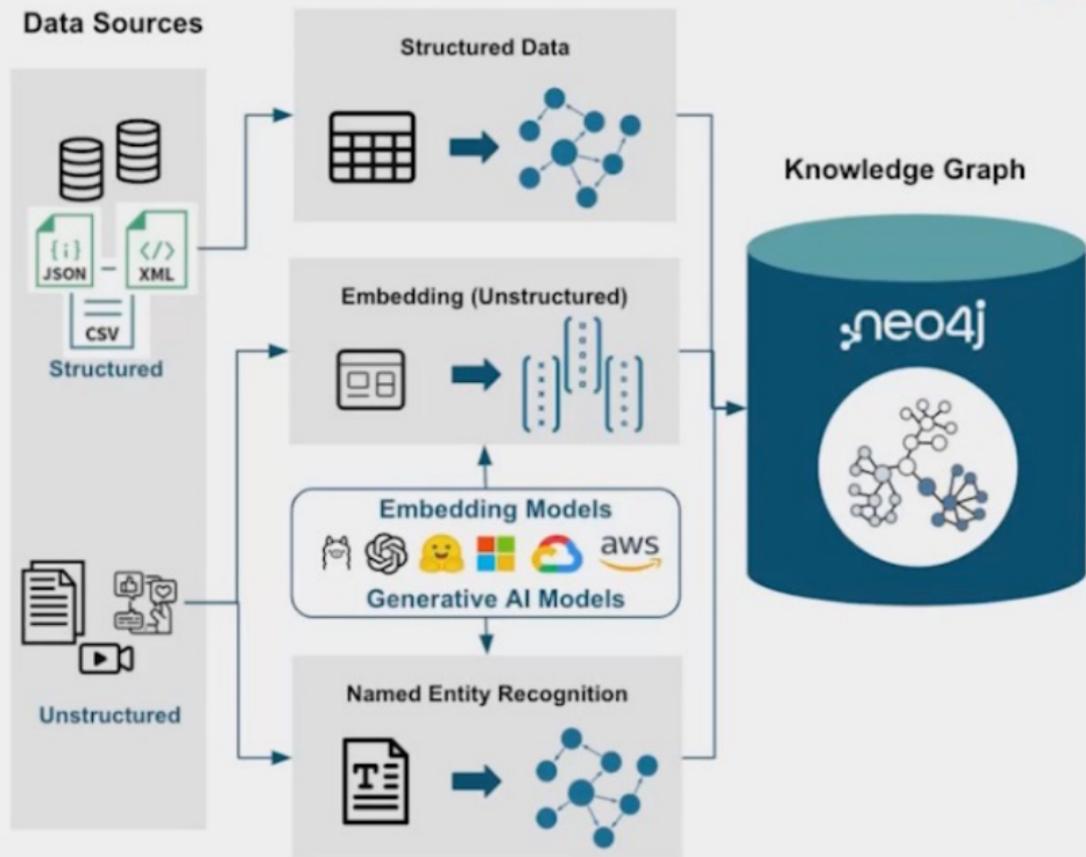
Jumpstart knowledge graph creation from unstructured data with named entity recognition



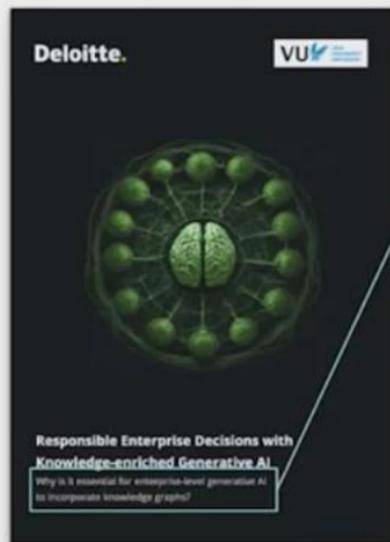
Rapidly model structured data as nodes and relationships in a knowledge graph



Ingest text, audio, and other content embeddings as node vector properties and index them for use in RAG



# GenAI + KG Association



Deloitte. VU

Responsible Enterprise Decisions with Knowledge-enriched Generative AI

Why is it essential for enterprise-level generative AI to incorporate knowledge graphs?

Why is it essential for enterprises to incorporate knowledge graphs



Forbes

PROFESSIONAL SERVICES

### The Role Of Knowledge Graphs In Overcoming LLM Limitations

Charles Breque Forbes Councils Member  
Forbes Business Council  
COUNCIL POST | Membership (For-Board)

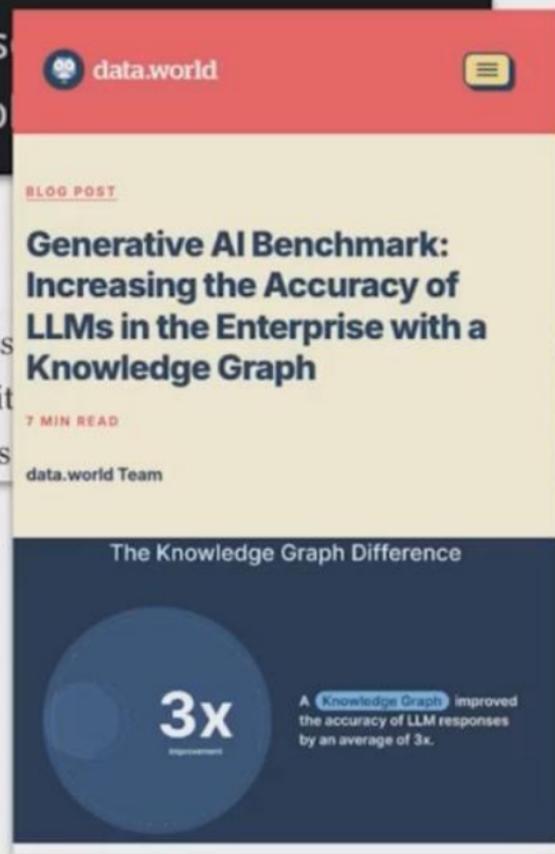
Dec 1, 2023, 10:00am EST

By Charles Breque, CEO and Founder of Legislate.



While Large Language Models (LLMs) like LLaMA 2 have shown remarkable prowess in understanding and generating text, they have a critical limitation. These models require

Regardless of the size of your business, incorporating knowledge graphs with your LLMs can significantly improve their performance and accuracy. This is because knowledge graphs provide a structured and interconnected representation of information, which can help LLMs better understand the context and relationships between different pieces of data. By leveraging knowledge graphs, businesses can unlock the full potential of their LLMs and achieve more accurate and reliable results.



data.world

BLOG POST

### Generative AI Benchmark: Increasing the Accuracy of LLMs in the Enterprise with a Knowledge Graph

7 MIN READ

data.world Team

#### The Knowledge Graph Difference

**3x** improvement

A Knowledge Graph improved the accuracy of LLM responses by an average of 3x.

Regardless of the size of your business, incorporating knowledge graphs with your LLMs can significantly improve their performance and accuracy. This is because knowledge graphs provide a structured and interconnected representation of information, which can help LLMs better understand the context and relationships between different pieces of data. By leveraging knowledge graphs, businesses can unlock the full potential of their LLMs and achieve more accurate and reliable results.

# Agenda



Curating Legal Data for Knowledge Graphs

01

Large Language Models (LLMs) & Knowledge Graph Creation

03

05

Other Use Cases

Knowledge Graphs Introduction

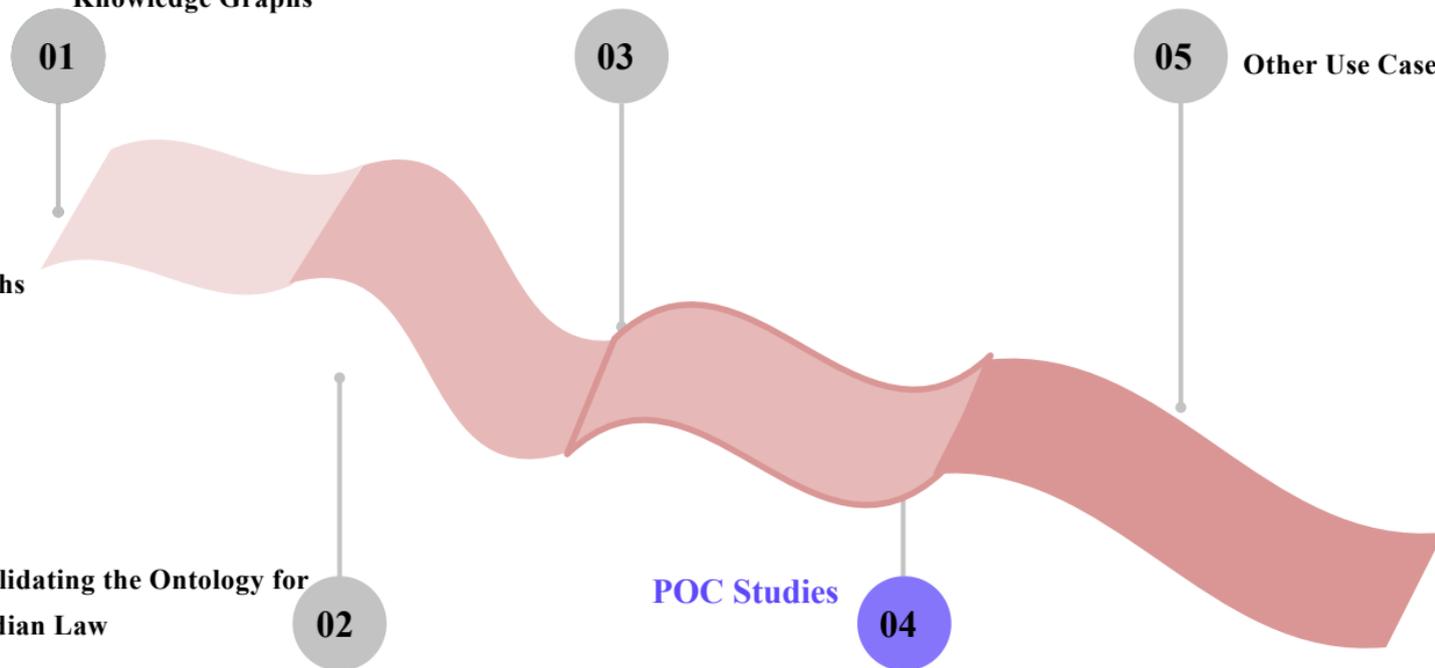
Validating the Ontology for Indian Law

02

POC Studies

04

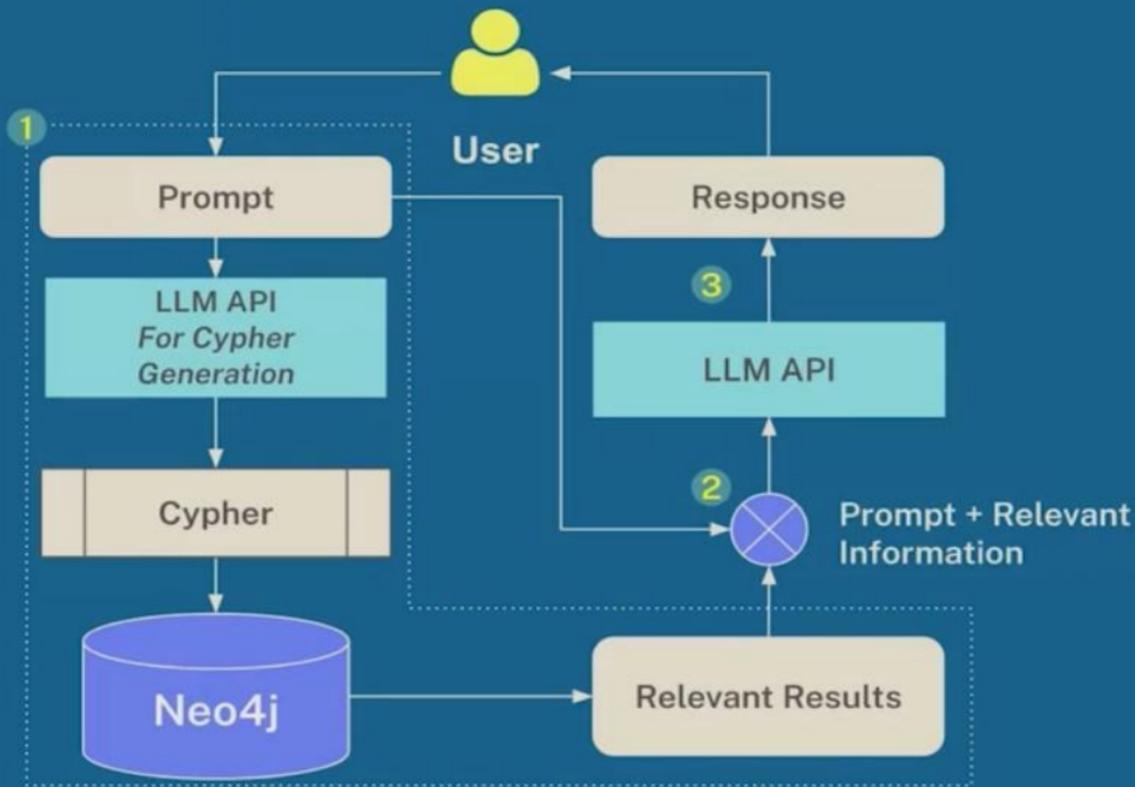
Governance



## Stage 4 - POC Study Objectives

1. Evaluate whether Text 2Cypher we could could create a scalable process to build a knowledge graph (KG) on the SALI framework so that it can capture and represent relationships and dependencies. Our first study is limited to cheque bouncing cases (NIA judgments).
2. Check if the NIA KGs on SALI can-
  - capture relationships and dependencies in cheque bouncing cases,
  - provide semantic analysis and reasoning capabilities for LLMs,
  - answer complex queries based on logical operators; and
  - perform advanced reasoning and knowledge discovery.
  - enhance diverse applications like search, recommendations, and AI assistants;
3. Whether backpropagation of missing nodes in KGs could be achieved simultaneously .

# GraphRAG with Text2Cypher



- 1 Retrieve relevant results from Neo4j using LLM to generate Cypher (Neo4j's query language)
- 2 Combine relevant results with prompt
- 3 Instruct LLM to only use the relevant results to generate response

# Agenda



**Curating Legal Data for Knowledge Graphs**

**01**

**Large Language Models (LLMs) & Knowledge Graph Creation**

**03**

**05**

**Other Use Cases**

**Knowledge Graphs Introduction**

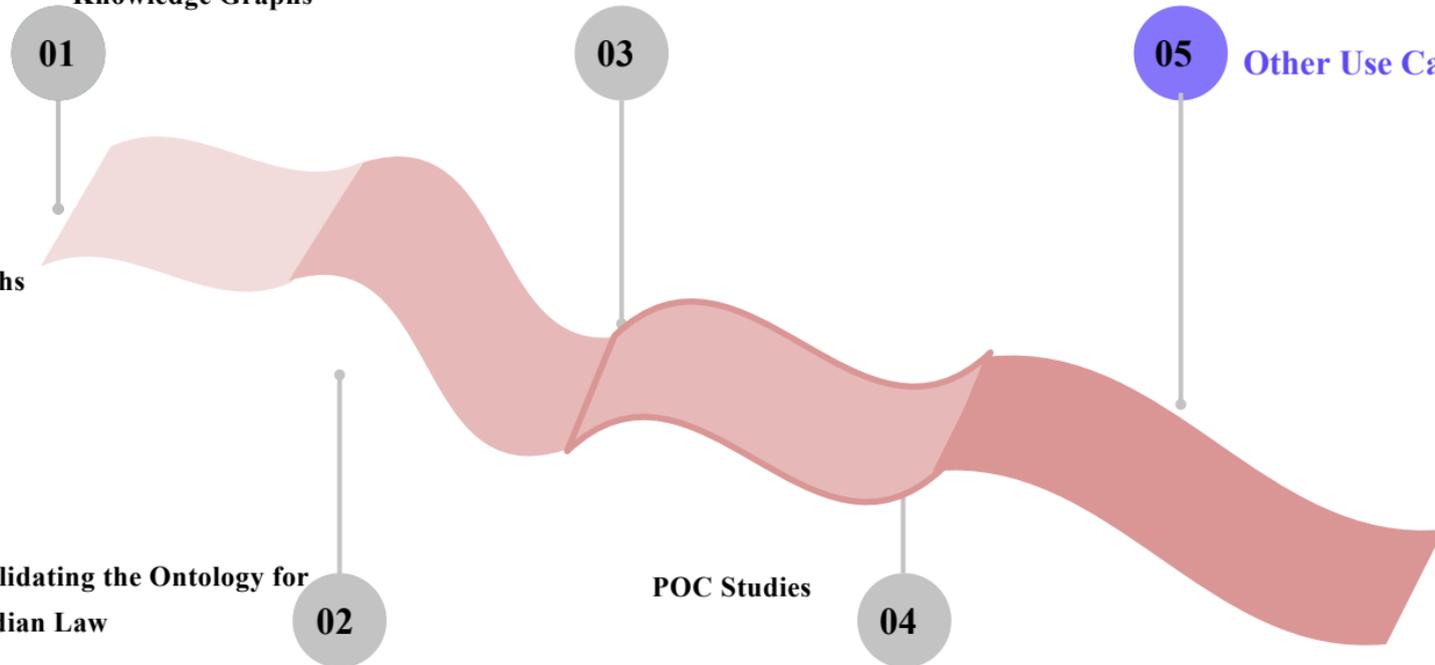
**Validating the Ontology for Indian Law**

**02**

**POC Studies**

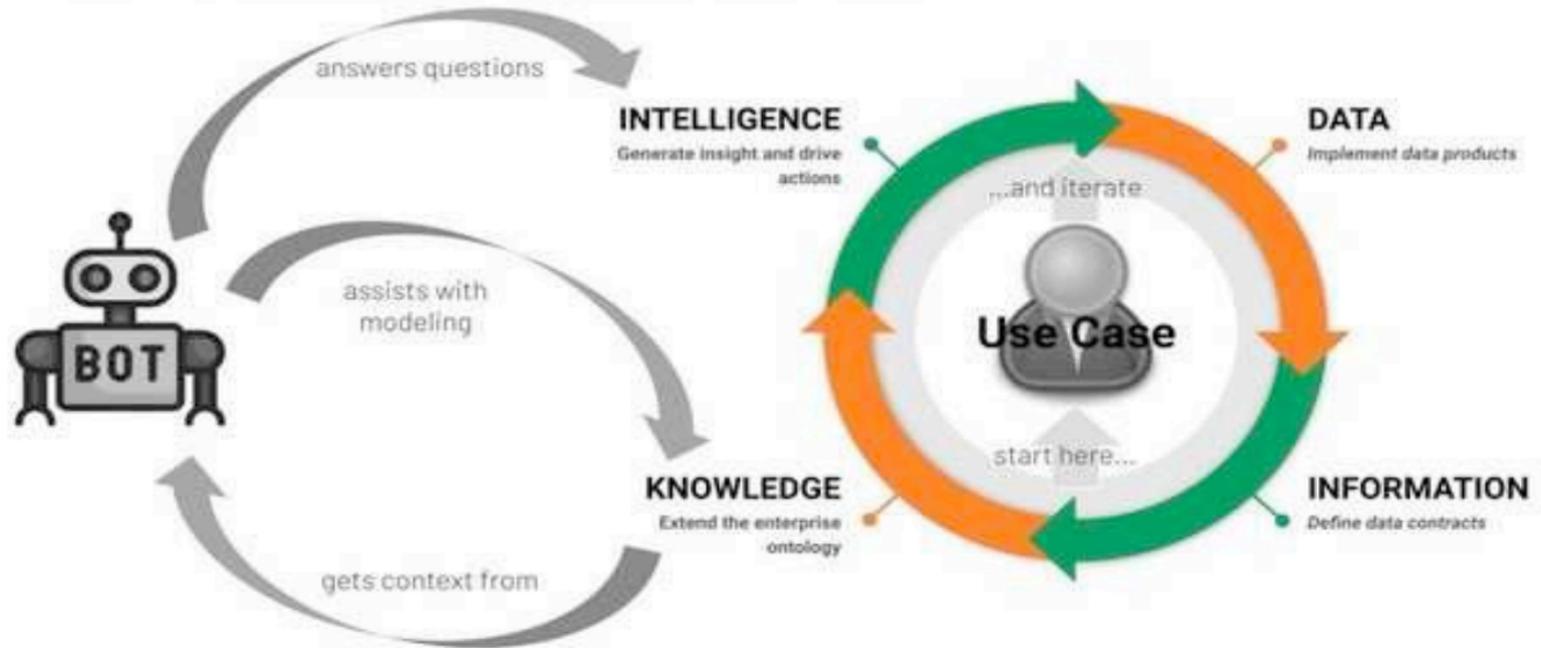
**04**

**Governance**



# Stage 5 - Other Use Cases

## Building the Knowledge Graph Flywheel



# Evaluating Alternative GraphRag Approaches

## Text2Cypher

Convert natural language prompts to explicit Cypher queries for retrieval



## Graph Vectors

Use graph embeddings for retrieval - incorporating both structured and unstructured data in vector similarity search



## Vector Search with Graph Context

Use graph patterns to retrieve items related to vector search results



## Graph Filtering

Use graph patterns and properties to pre/post filter vector search results. Also includes hybrid search



## Agents and Multi-Step

Combine multiple tools and GraphRAG patterns together for intelligent GenAI applications. Leverage corrective GraphRAG, semantic layers, and more

# Demonstrating Other Use Cases of LKG

**Assist partners/ community to use LKGs for solutions for additional use cases like**

- **Clustering**- Enable docket management and clustering of similar cases
- **Recommend Claims/Defences** – Based on the facts of the document, can LLMs recommend and draft relevant claims/defences most applicable in the scenario.
- **Recommend necessary rules/statute** – Based on the suggested claim/defence, can LLMs suggest to Judge or Lawyer the relevant rules applicable to given case scenario.
- **Authorities** - With NIA case laws annotated with Rhetorical Role (RR) should suggest relevant precedents for each potential claim/defence.
- **Predictive Analytics** - Determine outcomes of cases associated with certain claim/defenses. System should be able to predict the likelihood of success for the potential claim/defenses.
- **Review** – Identify potential issues with the cited authority and quotations/arguments.
- **Analysis** – Analyse files and guide judiciary/ lawyers on which claim/defences to allow/emphasize or potentially which ones to drop.