neo4j

Using Ontologies to Guide Knowledge Graph Creation

Dr. Jesús Barrasa / AI Field CTO

An **ontology** is a shared, explicit, formal model of a domain

What's in an ontology?

categories, classes, entity types, node types

Customer, Location, Product

properties, attributes, characteristics of categories

customer_id, product_description, latitude

relationships, connections between categories

purchases (Customer - Product)

What’s in an ontology? (optionally)

Instance (reference) data

US, ZA, ES, GB, Product_250346

Additional axioms (taxonomy)

Gold Customer ⊂ Customer, SARS-CoV-2 ⊂ Coronavirus

Annotations for relationship and properties

Depends_on is Transitive, “LI Connected” is symmetric

What are ontologies for?

Standard way of representing Information / Knowledge

- Interoperability
- Inferencing

Example ontologies

Finance

General

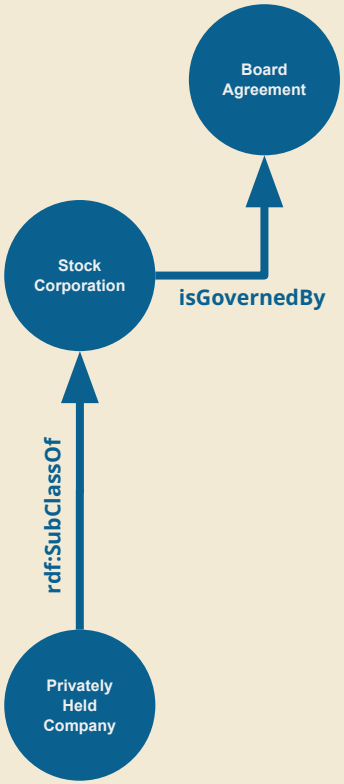
Life Sci

What does an ontology Look Like?

```
fibo:PrivatelyHeldCompany a owl:Class ;
  rdfs:subClassOf fibo:StockCorporation ;
  rdfs:label "privately held company" ;
  skos:definition "corporation whose issued shares are all held
       by a family or a small group of investors and, therefore,
       cannot be bought by the public" ;
  ns1:explanatoryNote "For British or Commonwealth companies… " ;
  ns1:synonym "closed corporation", "privately held corporation" .
```

```
fibo:StockCorporation a owl:Class ;
  rdfs:subClassOf [
    a owl:Restriction ;
    owl:onProperty fibo:isGovernedBy ;
    owl:someValuesFrom fibo:BoardAgreement ] .
```

```
fibo:BoardAgreement a owl:Class ;
  rdfs:label "board agreement" ;
  skos:definition "formal, legally binding agreement between
        members of the Board of Directors of the organization" .
```



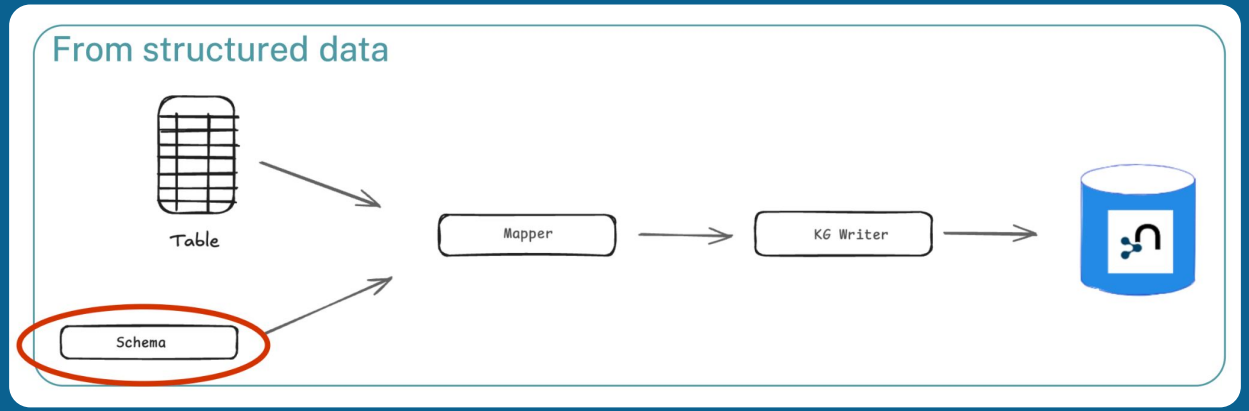
Ontologies for KG construction

The data ingestion pipeline maps the source data to the target model defined by the ontology

Structured data (databases, APIs…)

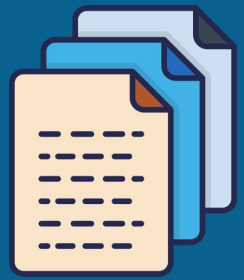
Done through structural mapping rules from source structures to the structures defined in the ontology

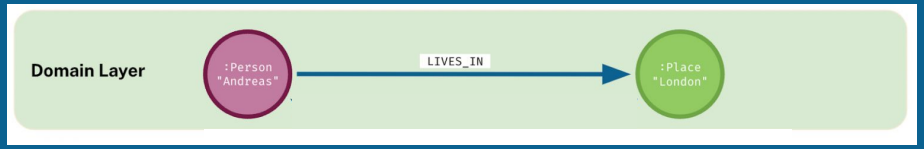
- Table (or projection, or document) to category or node type
- FK (or nested document) to Relationship type
- Column or key to attribute
- …

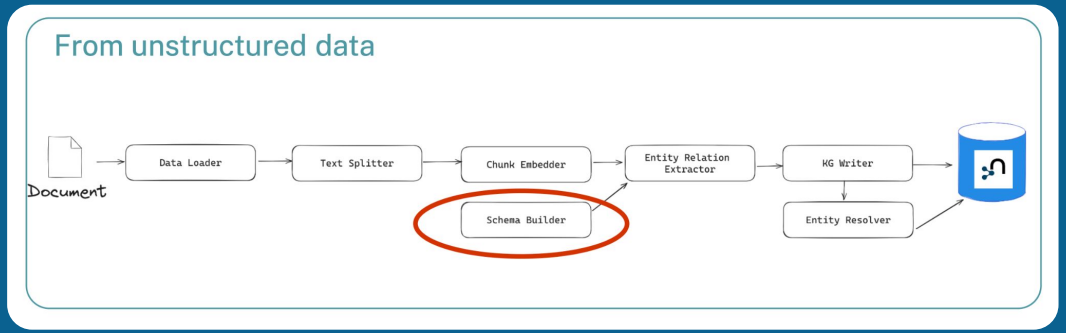


Unstructured data (NL text)

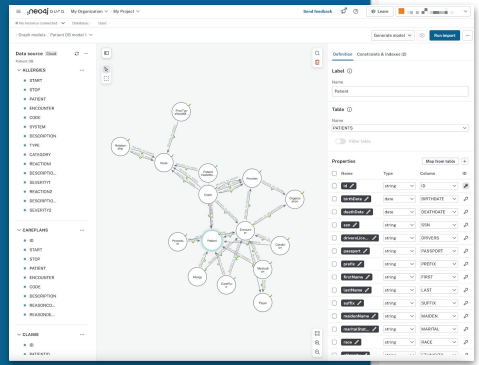
Done through entity extraction and entity disambiguation.



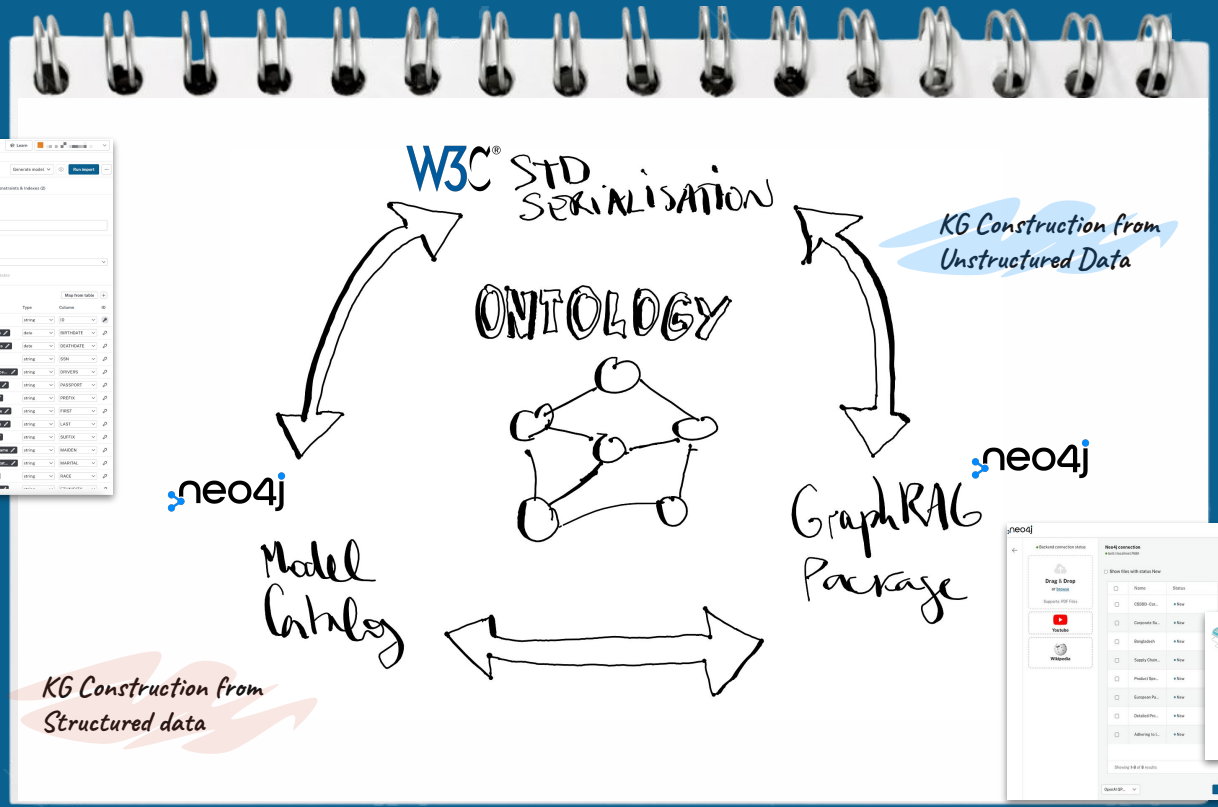


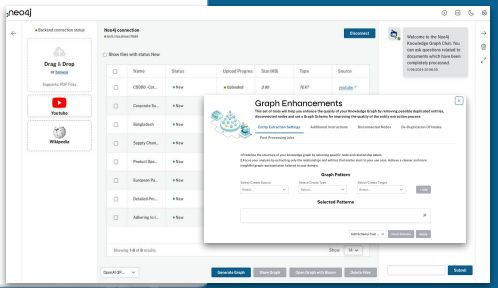


An ontology can take different forms



https://console.neo4j.io/



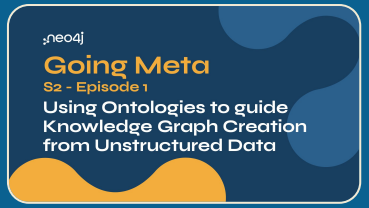


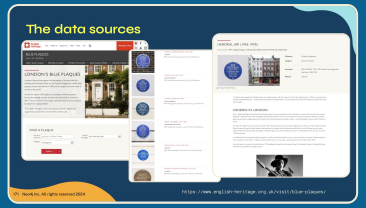
https://llm-graph-builder.neo4jlabs.com

Guided Examples

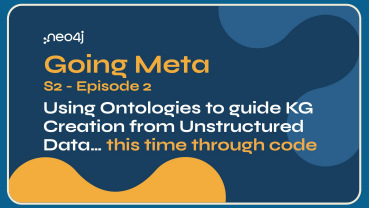
https://github.com/jbarrasa/goingmeta

From unstructured data (no code)
S2-Ep1

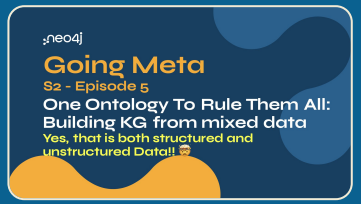


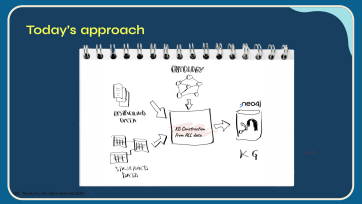


From unstructured data (programmatic)
S2-Ep2

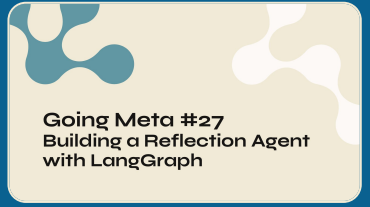


From structured data (no code)
S2-Ep5





From structured data (programmatic)
S1-Ep5 & S1-Ep27



Demo time!

Key Takeaways

Ontologies can take many different forms. Use them!

Avoid blind KG construction from unstructured data. Always ontology driven.

KG construction from unstructured data will benefit from some scaffolding (that typically comes from reference/master data)

- in an ontology
- in a structured data source

Ontologies can also drive data validation

Q&A