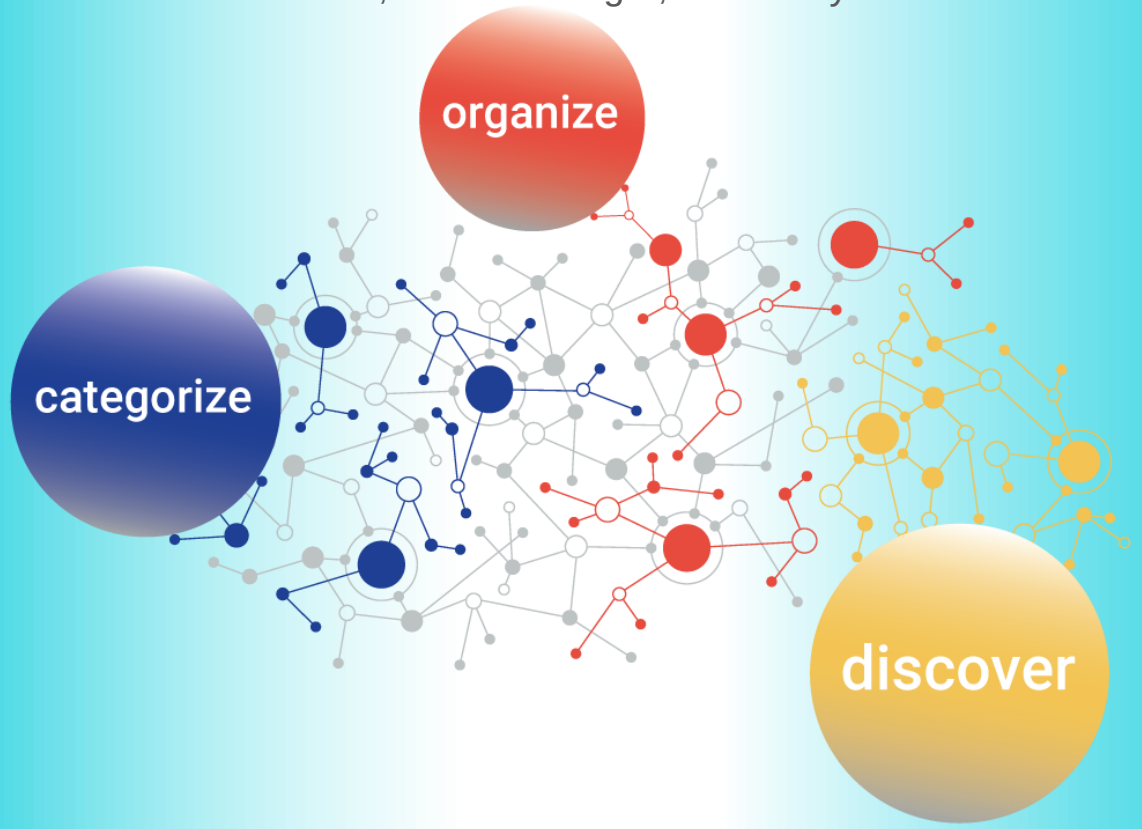
synaptica®

ISKO UK Meetup

Knowledge Graphs, Part I: Theories and Definitions

Tuesday, 14 July 2020

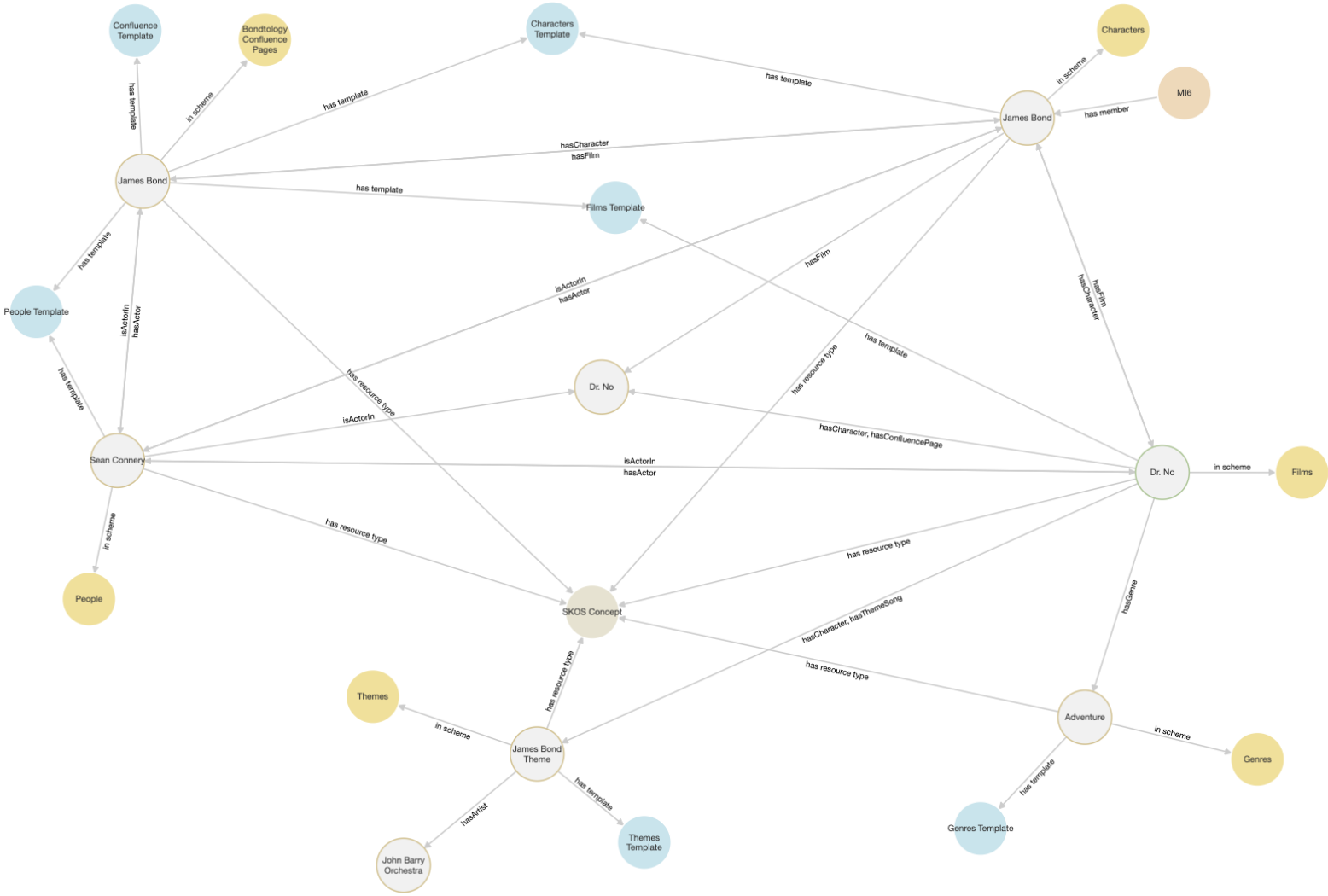
Bob Kasenchak, Senior Manager, Client Solutions
Ahren E. Lehnert, Senior Manager, Text Analytics Solutions



Webinar Objectives

In this webinar, we will cover

- knowledge graph definitions
- ontology definitions
- knowledge graph technologies



The Presenters

Ahren E. Lehnert
Senior Manager, Text Analytics Solutions

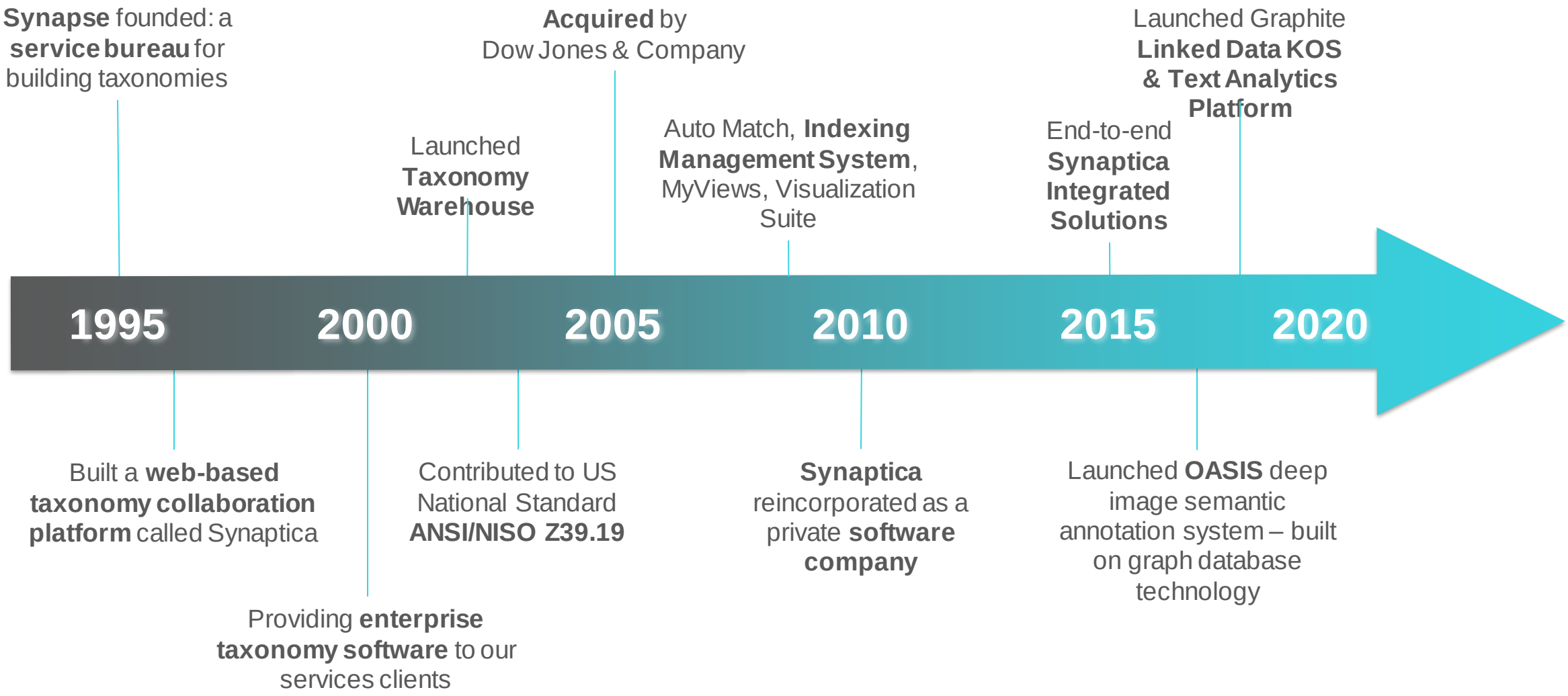
Information and knowledge management professional with over fifteen years' experience in taxonomy, text analytics, search, and content and records management. Ahren is Senior Manager, Text Analytics Solutions at Synaptica, LLC, a provider of taxonomy, ontology, and text analytics products and services.

@ahrenlehnert
ahren.lehnert@synaptica.com
www.synaptica.com

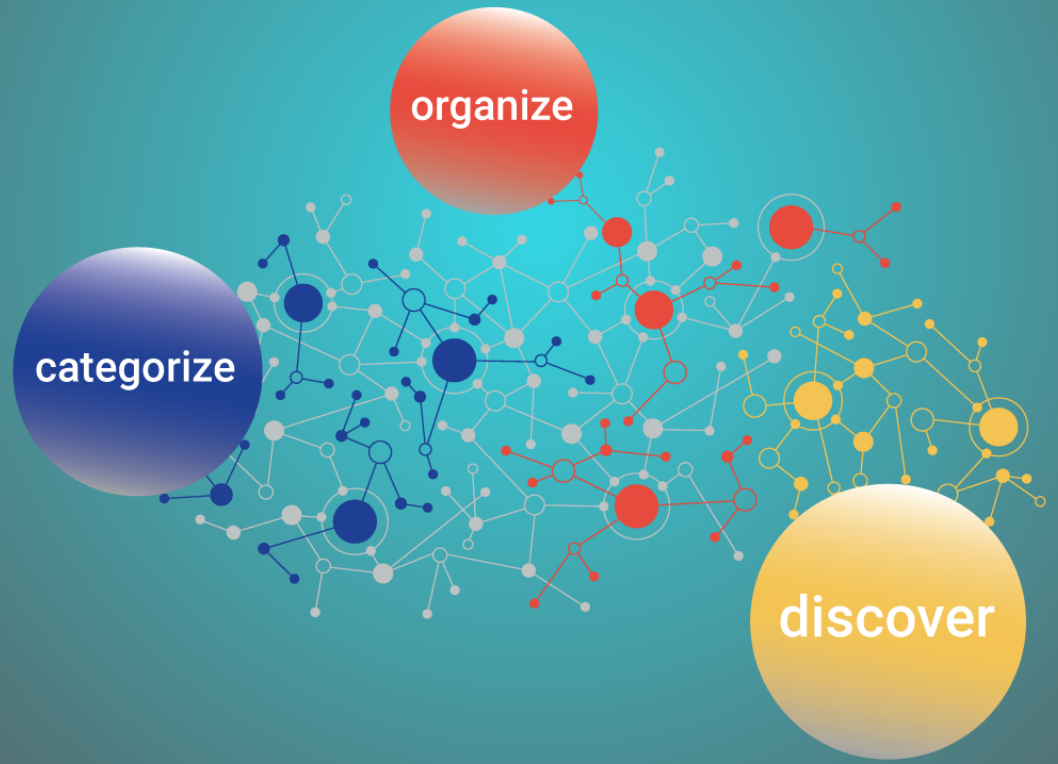
Bob Kasenchak
Senior Manager, Client Solutions

Taxonomist and ontologist with a focus on vocabulary development, indexing, software, and text analytics. Developed vocabulary projects for publishers and other enterprise clients.

@taxobob
bob.kasenchak@synaptica.com
www.synaptica.com


Synapse founded: a **service bureau** for building taxonomies
Launched **Taxonomy Warehouse**
Acquired by Dow Jones & Company
Auto Match, **Indexing Management System**, MyViews, Visualization Suite
End-to-end **Synaptica Integrated Solutions**
Launched Graphite **Linked Data KOS & Text Analytics Platform**
1995
2000
2005
2010
2015
2020
Built a **web-based taxonomy collaboration platform** called Synaptica
Providing **enterprise taxonomy software** to our services clients
Contributed to US National Standard **ANSI/NISO Z39.19**
Synaptica reincorporated as a private **software company**
Launched **OASIS** deep image semantic annotation system – built on graph database technology

We help people to organize, categorize and discover knowledge.



Taxonomies are Knowledge Organization Systems (KOS)

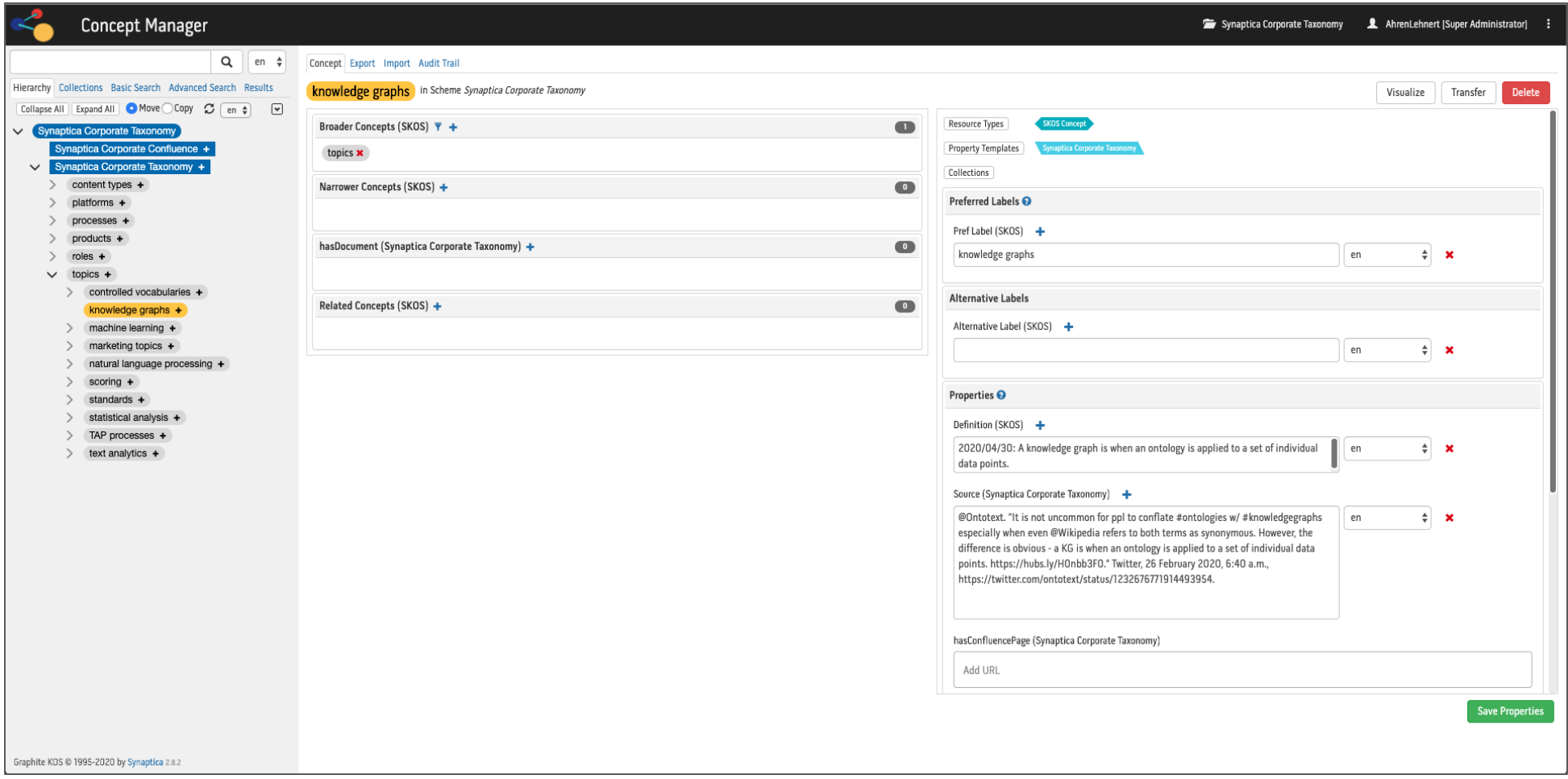
When we do taxonomy, we use industry standard data models to centralize and standardize the terminology used in our enterprise

We define and unambiguously label enterprise terminology, and then we organize it into hierarchical and associative concept schemes, which we call taxonomies or ontologies

These schemes help us to understand how concepts, people, places, products, processes, and organizations all relate to one another

Definitions

We're taxonomists, so you know we'd have definitions, right?



What Is a Knowledge Graph? - Theories

Information Science	Computer Science
Knowledge graphs are information provided directly to the user to enrich or enhance search or knowledge discovery.	Knowledge graphs are systems used for inferencing, reasoning, and machine learning.
A knowledge graph is when an ontology is applied to a set of individual data points. (1)	*A knowledge graph acquires and integrates information into an ontology and applies a reasoner to derive new knowledge. (2)*
Ontologies and taxonomies are always there at the start. (3)	*Many would argue that the divide between ontology and knowledge graph has nothing to do with size or semantics, but rather the very nature of the data. For example, dividing all class structures and relationship definitions into one group and all instance-level data into another might fulfill their idea of an ontology and knowledge graph, respectively — one to be used for inference, and the other to be queried for examples. (7)*
Ontologies actually form the specifications of the labels and the graph patterns that you use in your knowledge graph...	*...and this formal specification allows you to do the inferencing on them. (4)*
Knowledge graphs are a specific subclass of graphs, also known as semantic graphs. They come with metadata, schema, global identifier...	*...and reasoning capabilities. (5)*
Smart Data: (IRI, ontology, standard data type, open standards, designed for interoperability) (6)	*Automation: (supports machine learning and reasoning, powers event-driven data processing) (6)*

What Is a Knowledge Graph? - Synaptica's Position

Information Science
Ontology-first approach to knowledge graph engineering
By itself, an ontology is the structural underpinning of a knowledge graph...
Therefore, the initial ontology structure must be modelled and planned first
The concepts in an ontology are the interlinked entities described in a formal structure
The ontology structure connects metadata, data, and content; your ontology is a content model and mapping of your domain(s)
Human-in-the-loop model in which an ontology can be expanded through connections to external data or content and put into action through representation, visualization, querying, and reasoning, it is a knowledge graph

What Is an Ontology?

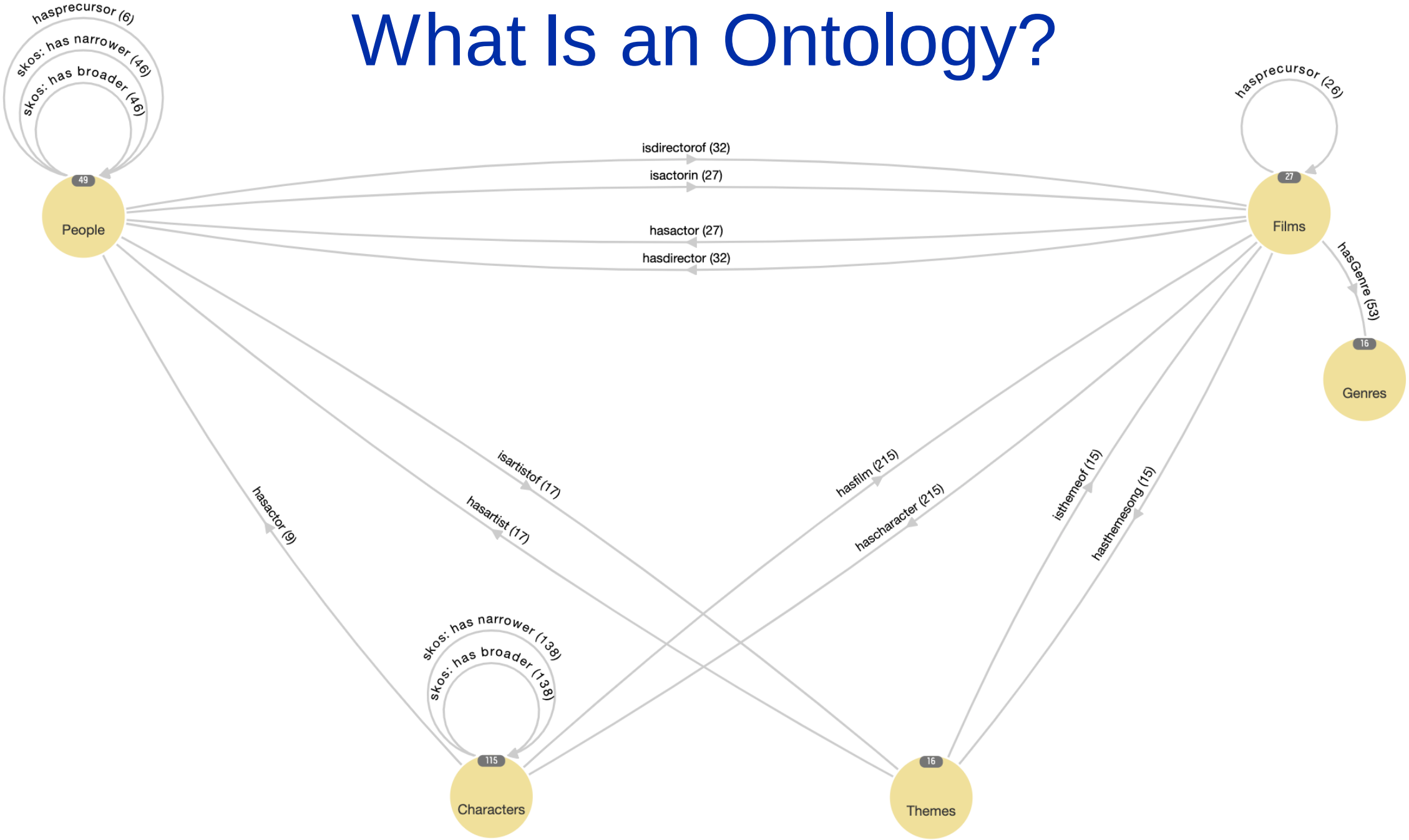
*In computer science and information science, an ontology encompasses a representation, formal naming and definition of the categories, properties and relations between the concepts, data and entities that substantiate one, many or all domains of discourse.**

*"Ontology (information science)." *Wikipedia,* Wikimedia Foundation, Inc., 3 August 2019, https://en.wikipedia.org/wiki/Ontology_(information_science).

What Is an Ontology? - Synaptica's Position

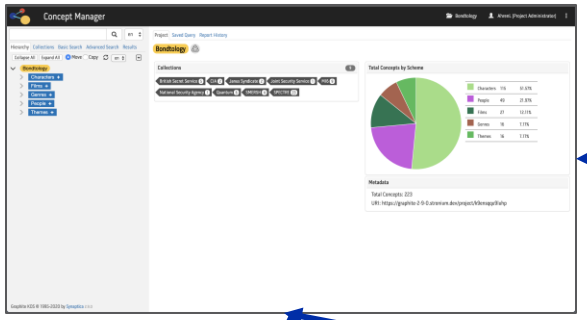
Ontology Management
In most use cases, an organization uses ontologies to describe what it knows and what it wants to know
Typically, organizations require domain ontologies, built specifically to the domain or domains in which they operate
An ontology scheme, the definitions of which properties and relationships can be used, are repeatable and reusable
When an ontology scheme is populated by concept instances, and those instances connect to Linked Data or external data or content, it is a knowledge graph

What Is an Ontology?

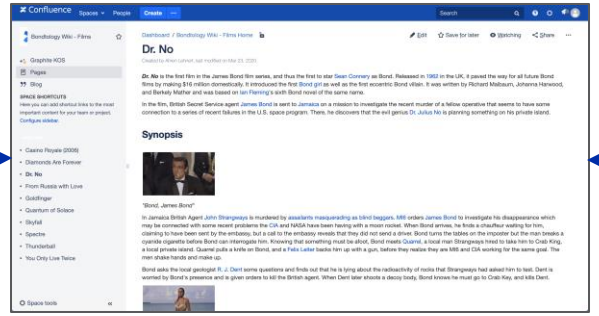


Knowledge Graph

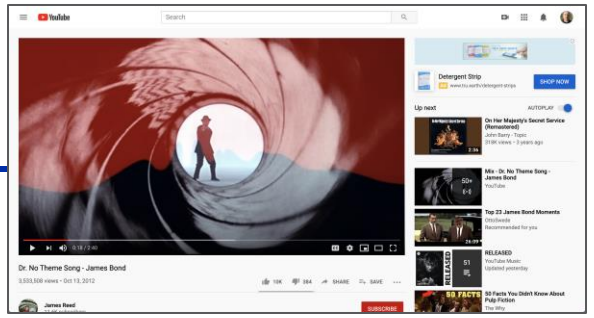
Ontology Management System

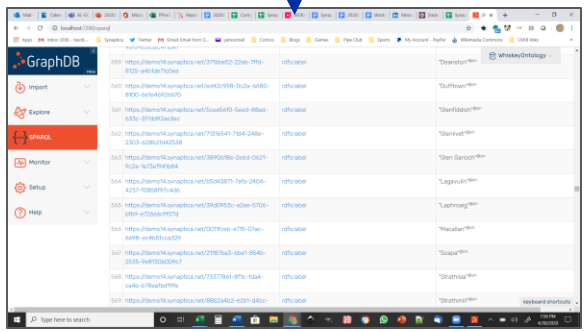


Confluence Wiki

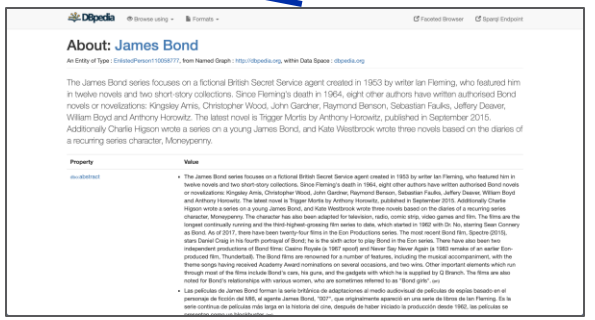


YouTube





Graph Database
(Ontotext GraphDB)



DBPedia (Linked Data)

What Is a Graph?

A graph data structure consists of a finite (and possibly mutable) set of vertices (also called nodes or points), together with a set of unordered pairs of these vertices for an undirected graph or a set of ordered pairs for a directed graph.

*These pairs are known as edges (also called links or lines), and for a directed graph are also known as arrows. The vertices may be part of the graph structure, or may be external entities represented by integer indices or references.**

*"Graph (abstract data type)." *Wikipedia,* Wikimedia Foundation, Inc., 3 August 2019, https://en.wikipedia.org/wiki/Graph_%28abstract_data_type%29.

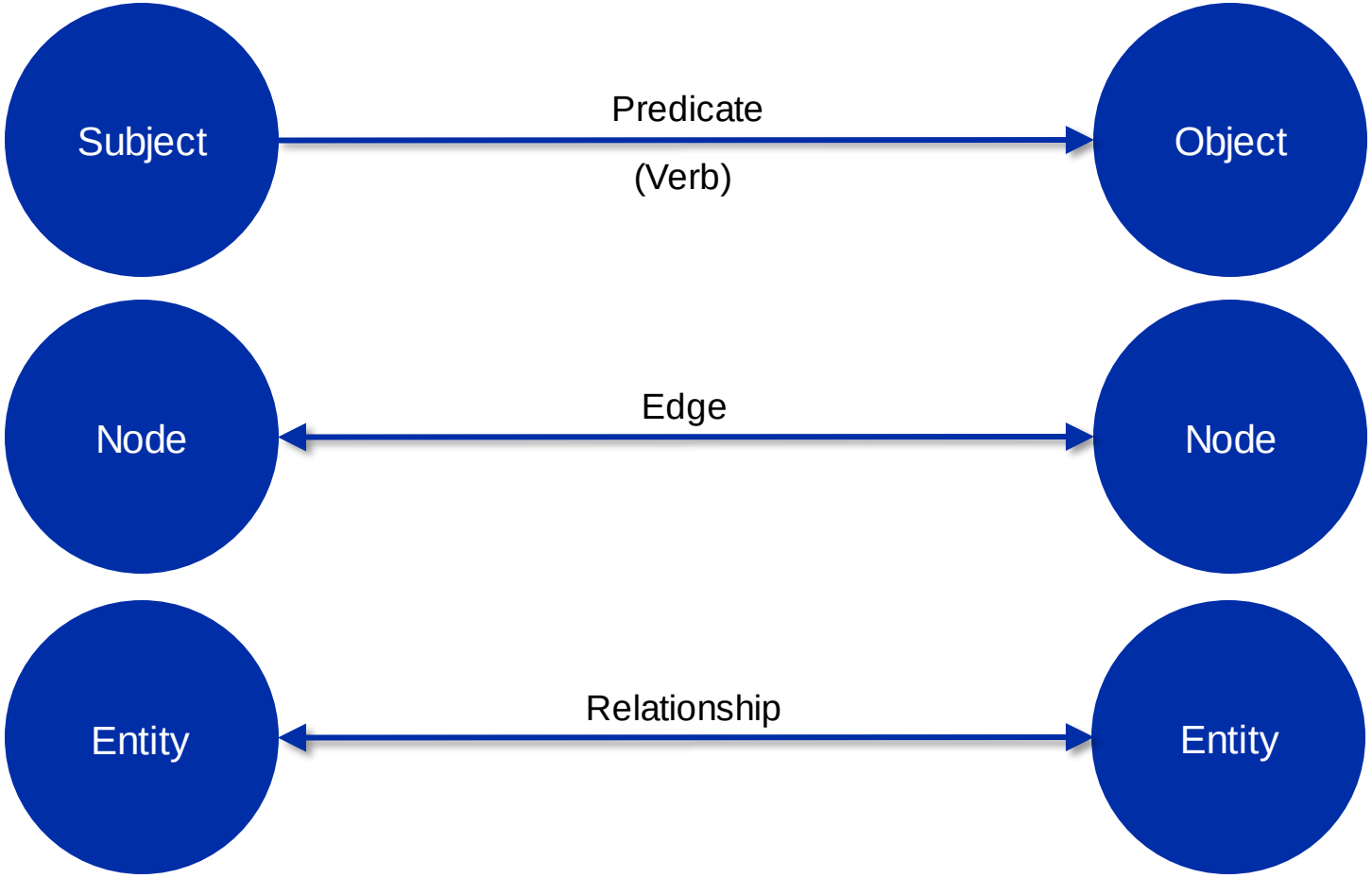
What Is a Graph?

A graph is a set of vertices (nodes, entities) and edges (relationships).

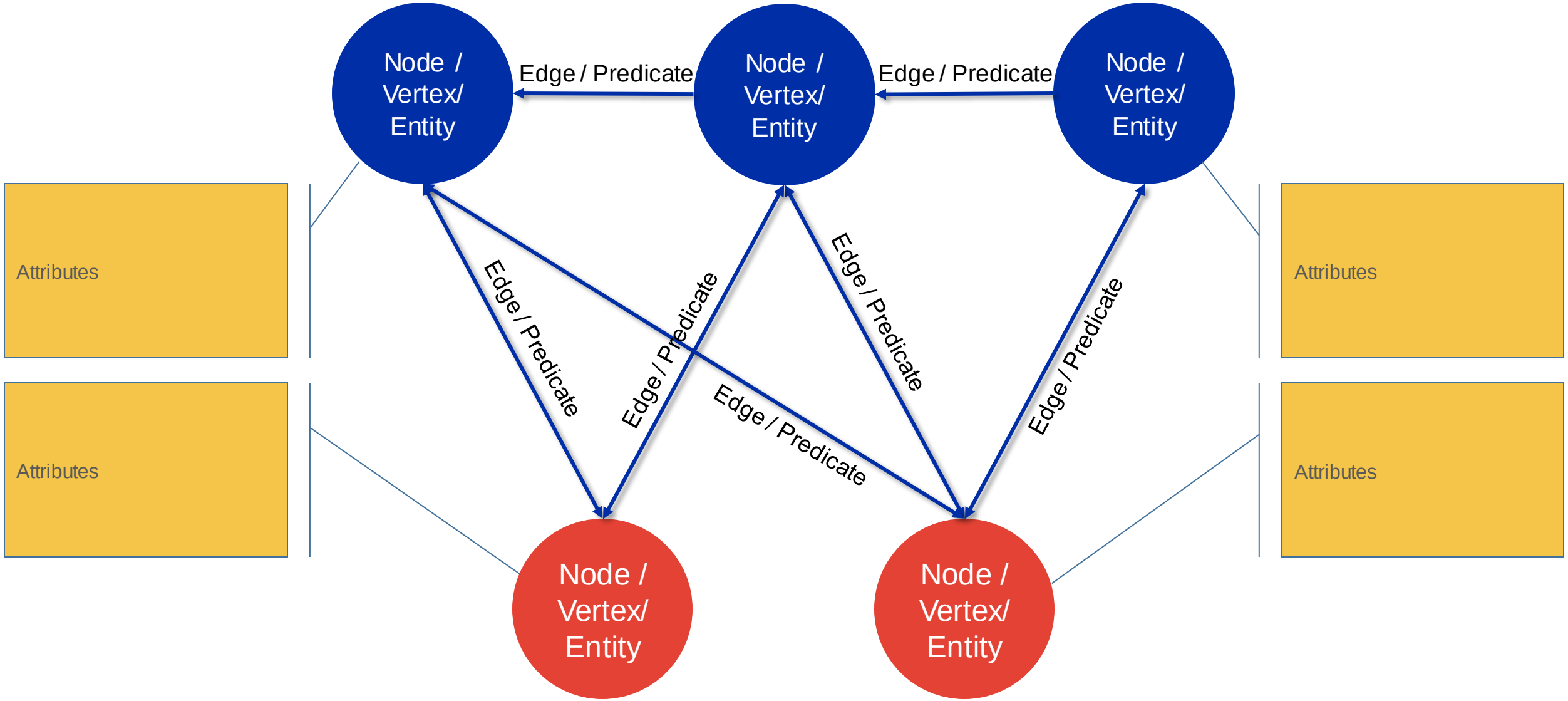
*Graphs represent data in a way that's closer to our mental models.**

**"Traversing the Land of Graph Computing and Databases." *Humans of Data,* 31 July 2019, https://humansofdata.atlan.com/2019/07/what-is-graph-database-and-graph-tech/.

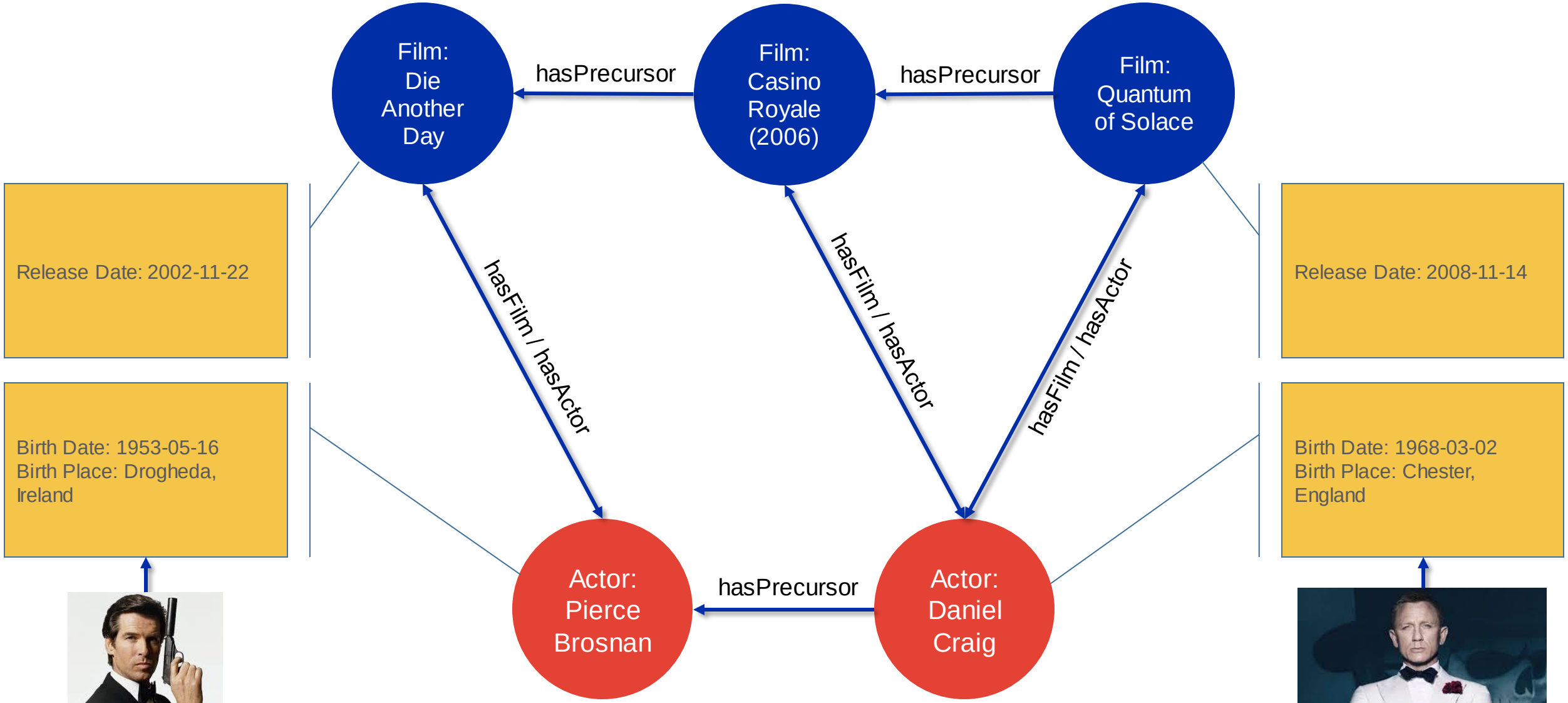
Triples


Subject
Predicate
(Verb)
Object
Node
Edge
Node
Entity
Relationship
Entity

Knowledge Graph: Structure



Knowledge Graph: Bond. James Bond



Google Knowledge Graph

*The Knowledge Graph is a knowledge base used by Google and its services to enhance its search engine's results with information gathered from a variety of sources.**

- Started in May 2012
- Sources
  - Google Image Search
  - Wikipedia & Wikidata
  - Movies (IMDB?)
  - Social Media Profiles
  - Google Search
  - Preferred Sources: Sponsored Ads, LyricFind, Google Books, etc.

James Bond

Film series

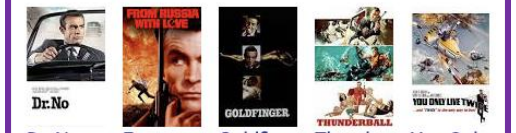
The James Bond film series is a series of spy films based on the fictional character of MI6 agent James Bond, "007", who originally appeared in a series of books by Ian Fleming. It is one of the longest continually-running film series in history, having been in on-going production from 1962 to the present. Wikipedia

Cast: Léa Seydoux, Lashana Lynch, Naomie Harris, MORE

Directors: Sam Mendes, Cary Joji Fukunaga, Martin Campbell, MORE

Production companies: Eon Productions, Metro-Goldwyn-Mayer, MORE

Movies View 20+ more



Dr. No 1962 | From Russia with Love 1963 | Goldfin... 1964 | Thund... 1965 | You Only Live Twice 1967

Profiles

YouTube Facebo...

People also search for View 10+ more

Mission: Imposs... | Bourne | Knives Out | Top Gun | Harry Potter

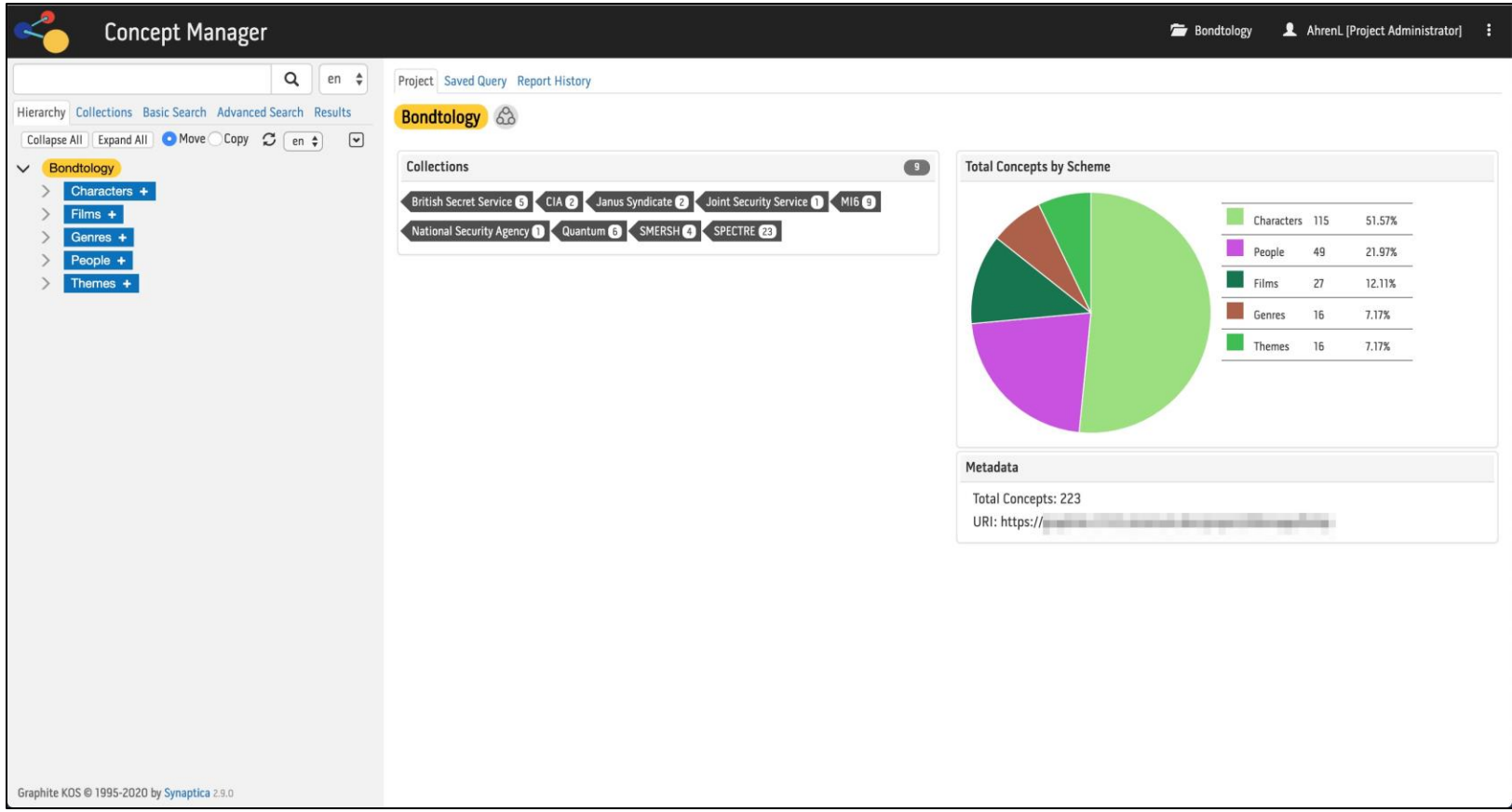
*"Knowledge Graph." *Wikipedia,* Wikimedia Foundation, Inc., 4 August 2019, https://en.wikipedia.org/wiki/Knowledge_Graph.

Ontology Management

- An ontology management system is a software application for creating and curating Knowledge Organization Systems (KOS) based on Linked Data and Semantic Web standards that stores the KOS in an RDF graph database, also known as a triple-store
- Ontology management systems lets users adopt publicly available ontologies from the Linked Open Data cloud. This can drastically reduce the time and effort required to design and configure Knowledge Organization Systems

Ontology Management

- Easily accessible and intuitive UI
- RDF is permissive, so it is necessary to apply constraints and rules



Graph Databases

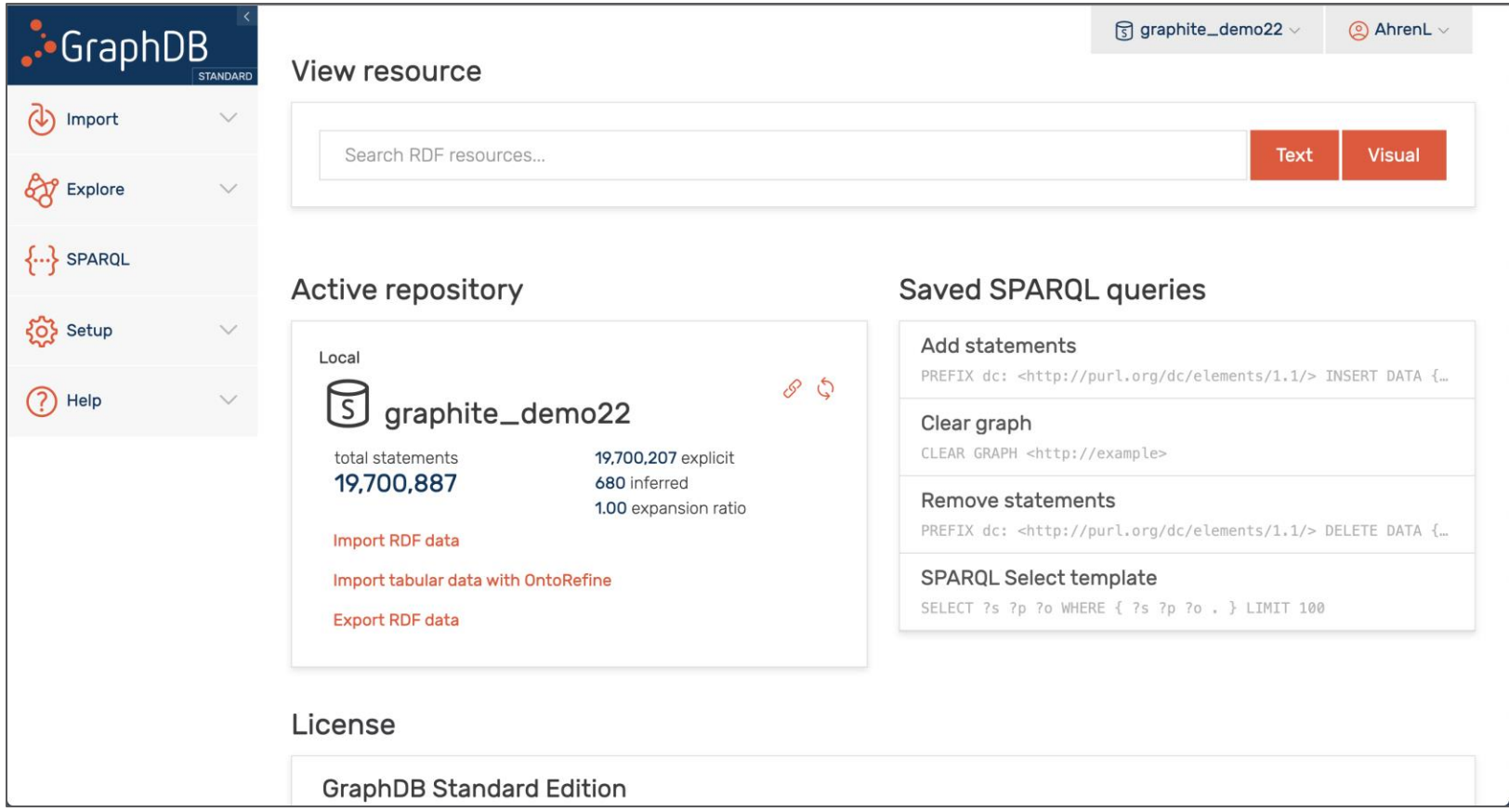
*In computing, a graph database (GDB) is a database that uses graph structures for semantic queries with nodes, edges, and properties to represent and store data.**

*"Graph database." *Wikipedia,* Wikimedia Foundation, Inc., 23 March 2020, https://en.wikipedia.org/wiki/Graph_database.

What Is GraphDB?

- *GraphDB is a family of highly efficient, robust and scalable RDF databases. It streamlines the load and use of linked data cloud datasets, as well as your own resources.**

- GraphDB can be used to query and visualize ontologies created in Graphite



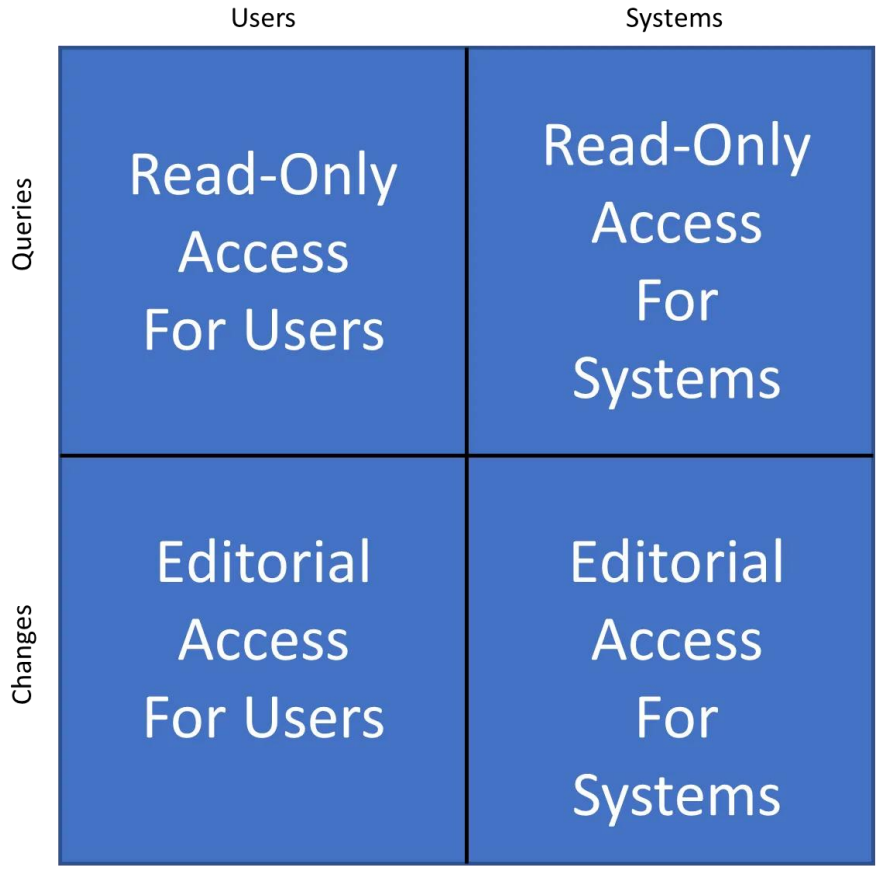
**About GraphDB*. OntoText. 24 September 2019. http://graphdb.ontotext.com/documentation/free/about-graphdb.html.

Internal and External Systems

- Databases
  - External sources, such as DBpedia
  - Internal sources housing structured information used by your organization

- Content Management Systems
  - Internal systems with employee-produced documents or web pages
  - Integration for the tagging of content, providing consistent labeling and relationships between objects

Access to the Graph

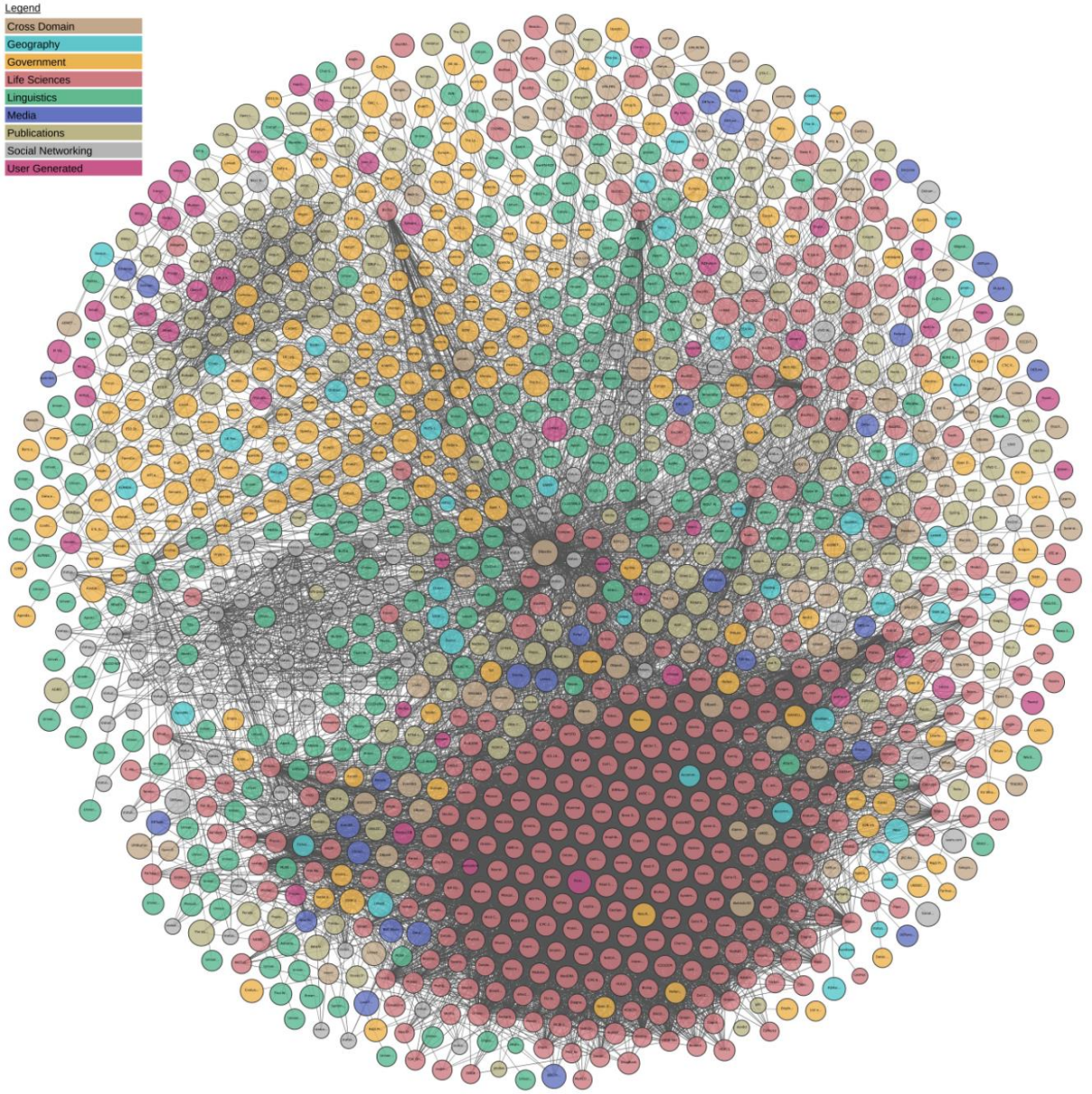


Linking to Internal & External Data

- APIs
  - Connect to external systems but also provide constraints and access rules
- Search
  - Allow user queries via SPARQL
  - Allow user queries through search or other interfaces
  - Foundational technology for bringing knowledge graph content together in the UI

What Is Linked Data?

*Linked Data is about using the Web to connect related data that wasn't previously linked, or using the Web to lower the barriers to linking data currently linked using other methods.**



**Linked Data*. Linked Data. 24 September 2019. http://linkeddata.org/home.

What Is Linked Data?

The term Linked Data refers to a set of best practices for publishing structured data on the Web. These principles have been coined by Tim Berners-Lee in the design issue note Linked Data. The principles are:

- Use URIs as names for things
- Use HTTP URIs so that people can look up those names
- When someone looks up a URI, provide useful information
- Include links to other URIs so that they can discover more things

Use standards for the representation and the access to data on the Web.

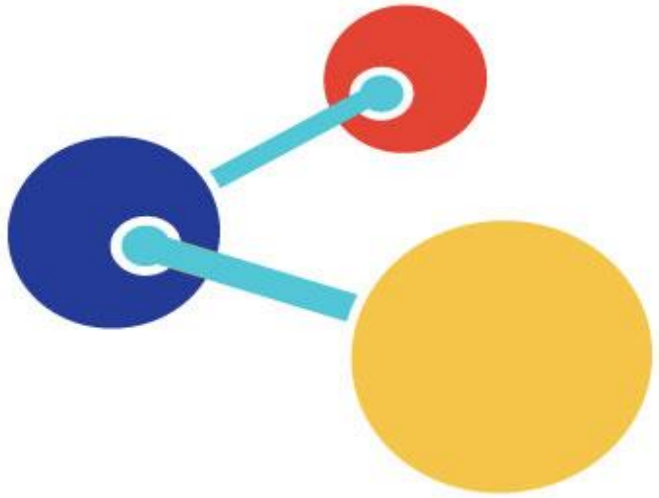
Use hyperlinks between data from different sources. These hyperlinks connect all Linked Data into a single global data graph.*

**Linked Data*. W3C. 24 September 2019. https://www.w3.org/wiki/LinkedData.

In Summary...

- Knowledge graphs are ontology-first
- A knowledge graph is when an ontology is connected to other data and content
- Knowledge graph technologies include ontology management software, graph databases, and sources of information in the form of content, Linked Data, etc.

Questions?



@ahrenlehnert
ahren.lehnert@synaptica.com
www.synaptica.com

@taxobob
bob.kasenchak@synaptica.com
www.synaptica.com

Resources – Knowledge Graphs

- (5) Anadiotis, George. Linked Data Orchestration Blog/Year of the Graph newsletter. https://linkeddataorchestration.com/the-year-of-the-graph/.
- (3) Aasman, Jans. "Knowledge Graphs". Stanford CS 520 Knowledge Graphs. 31 March 2020. Accessed 6 July 2020. https://www.youtube.com/watch?v=bvwjG-3qAmY&feature=youtu.be.
- CS 520 Knowledge Graphs. 31 March 2020 - 2 June 2020. https://web.stanford.edu/class/cs520/.
- Deus, Helena. *Knowledge Graphs and Machine Learning - ISWC 2018 trip report*. 18 October 2018. 30 September 2019. https://www.linkedin.com/pulse/knowledge-graphs-machine-learning-iswc-2018-trip-report-helena-deus/.
- (2) Ehrlinger, Lisa, and Wolfram Wöß. "Towards a Definition of Knowledge Graphs." *SEMANTICS 2016: Posters and Demos Track*, Leipzig, Germany, September 13-14, 2016. Johannes Kepler University Linz, http://ceur-ws.org/Vol-1695/paper4.pdf.
- (6) Enterprise Knowledge Graph Foundation. *Characteristics of an Enterprise Knowledge Graph*. 23 June 2020. Accessed 7 July 2020. https://www.ekgf.org/20200623-ekgf-launch-webinar.
- Guilin Qi et al. Knowledge Graph. https://books.google.com/books/about/Knowledge_Graph.html?id=6CsetAEACAAJ&source=kp_book_description.
- J.Z. Pan et al., eds. Exploiting Linked Data and Knowledge Graphs in Large Organisations. https://www.springer.com/gp/book/9783319456522.
- (1) @Ontotext. "It is not uncommon for ppl to conflate #ontologies w/ #knowledgegraphs especially when even @Wikipedia refers to both terms as synonymous. However, the difference is obvious – a KG is when an ontology is applied to a set of individual data points. https://hubs.ly/H0nbb3F0." *Twitter, 26 February 2020, 6:40 a.m.,* https://twitter.com/ontotext/status/1232676771914493954.
- (7) Open Data Science. "Where Ontologies End and Knowledge Graphs Begin." 16 October 2018. Accessed 9 July 2020. https://medium.com/predict/where-ontologies-end-and-knowledge-graphs-begin-6fe0cdede1ed.
- (4) Vrandecic, Denny. "Knowledge Graphs". Stanford CS 520 Knowledge Graphs. 31 March 2020. Accessed 6 July 2020. https://www.youtube.com/watch?v=bvwjG-3qAmY&feature=youtu.be.

Resources – Ontology, Linked Data, Semantic Web

- Allemang & Hendler. Semantic Web for the Working Ontologist. http://www.kevenlw.name/downloads/Ontologist.pdf.
- Keet, C. Maria. An Introduction to Ontology Engineering. https://people.cs.uct.ac.za/~mkeet/files/OEbook.pdf.
- "Ontology (information science)." *Wikipedia,* Wikimedia Foundation, Inc., 3 August 2019, https://en.wikipedia.org/wiki/Ontology_(information_science).
- Knowledge Organization Systems. https://link.springer.com/article/10.1007/s00799-018-0241-2.
- *Linked Data*. Linked Data. 24 September 2019. http://linkeddata.org/home.
- Linked Open Data. https://lod-cloud.net/.
- "Ontology (information science)." *Wikipedia,* Wikimedia Foundation, Inc., 3 August 2019, https://en.wikipedia.org/wiki/Ontology_(information_science).
- Wood, Zaidman & Ruth. Linked Data. https://www.manning.com/books/linked-data.

Resources – SKOS, OWL & RDF

- SKOS
  - https://www.w3.org/TR/skos-reference/
  - http://www.w3.org/2004/02/skos/core#

- OWL
  - https://www.w3.org/TR/2012/REC-owl2-primer-20121211/
  - *OWL Web Ontology Language Reference*. W3C. 11 May 2020. https://www.w3.org/TR/owl-ref/#Class.

- RDF
  - "RDF." *W3C Semantic Web, 7 May 2020*, https://www.w3.org/RDF/.
  - "XML RDF." *W3Schools.com, 7 May 2020*, https://www.w3schools.com/xml/xml_rdf.asp.
  - "Enter the Realm of Semantic Web Languages." *Medium,* 11 May 2020, https://medium.com/swlh/enter-the-realm-of-semantic-web-languages-ee94ee68f544.

- Schema.org: https://schema.org/

Resources – Graphs & Graph Databases

- "Graph (abstract data type)." *Wikipedia,* Wikimedia Foundation, Inc., 3 August 2019, https://en.wikipedia.org/wiki/Graph_%28abstract_data_type%29.
- "Traversing the Land of Graph Computing and Databases." *Humans of Data,* 31 July 2019, https://humansofdata.atlan.com/2019/07/what-is-graph-database-and-graph-tech/.
- O'Reilly. Graph Databases. https://graphdatabases.com/.
- "Graph database." *Wikipedia,* Wikimedia Foundation, Inc., 23 March 2020, https://en.wikipedia.org/wiki/Graph_database.
- Ontotext: https://www.ontotext.com/knowledgehub/fundamentals/what-is-rdf-triplestore/.

Resources - SPARQL

- DuCharme, Bob. Learning SPARQL. http://www.learningsparql.com/.
- Wikibooks: SPARQL. https://en.wikibooks.org/wiki/SPARQL.