

Building & Mining Knowledge Graphs

KEN4256

Lecture 1: Introduction to Knowledge Graphs



Maastricht University

Institute of Data Science

2024

© 2024 by Michel Dumontier and the Institute of Data Science at Maastricht University is licensed under Attribution 4.0 International
To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

This license requires that reusers give credit to the creator. It allows reusers to distribute, remix, adapt, and build upon the material in any medium or format, even for commercial purposes.

id: KEN4256_L1

version: 1.2024.0

created: January 8, 2019

last modified: March 26, 2024

published on: March 26, 2024

Course Instructors



Dr. Michel Dumontier

Distinguished Professor

Institute of Data Science

Expertise: knowledge graphs, ontologies, linked data, machine learning, deep learning, personalised medicine

michel.dumontier@maastrichtuniversity.nl



Vincent Emonet

Data Science Developer

Institute of Data Science

Expertise: semantic web, ontology + repositories, data integration, devops

vincent.emonet@maastrichtuniversity.nl



Dr. Rena Yang

Postdoctoral Fellow

Institute of Data Science

Expertise: machine learning, statistical modeling

rena.yang@maastrichtuniversity.nl



Shervin Mehryar

PhD Student

Institute of Data Science

Expertise: knowledge graphs, semantic web, deep learning, optimization, neuro-symbolic agents

shervin.mehryar@maastrichtuniversity.nl



Rishabh Jakhar

PhD Student

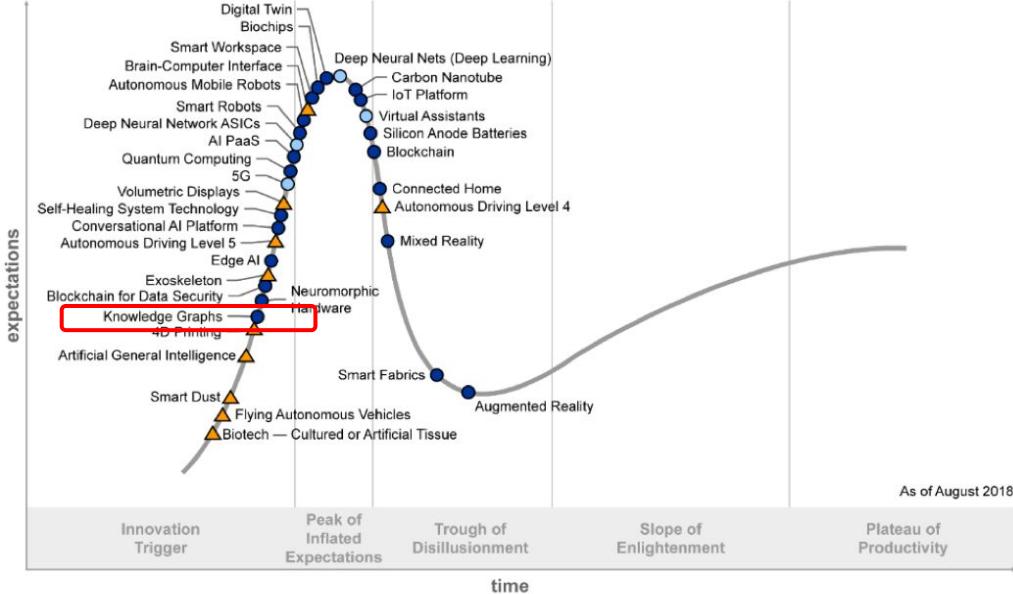
Institute of Data Science

Expertise: deep learning, image processing, numerical optimization

rishabh.jakhar@maastrichtuniversity.nl

Lecture outline

1. The ‘What’ and ‘Why’ of knowledge graphs
2. Knowledge graph representations
3. Course logistics
4. Use of AI



Hype Cycle for Artificial Intelligence, 2023



gartner.com

Source: Gartner
© 2023 Gartner, Inc. and/or its affiliates. All rights reserved. 2079794

Gartner

History around the term knowledge graph

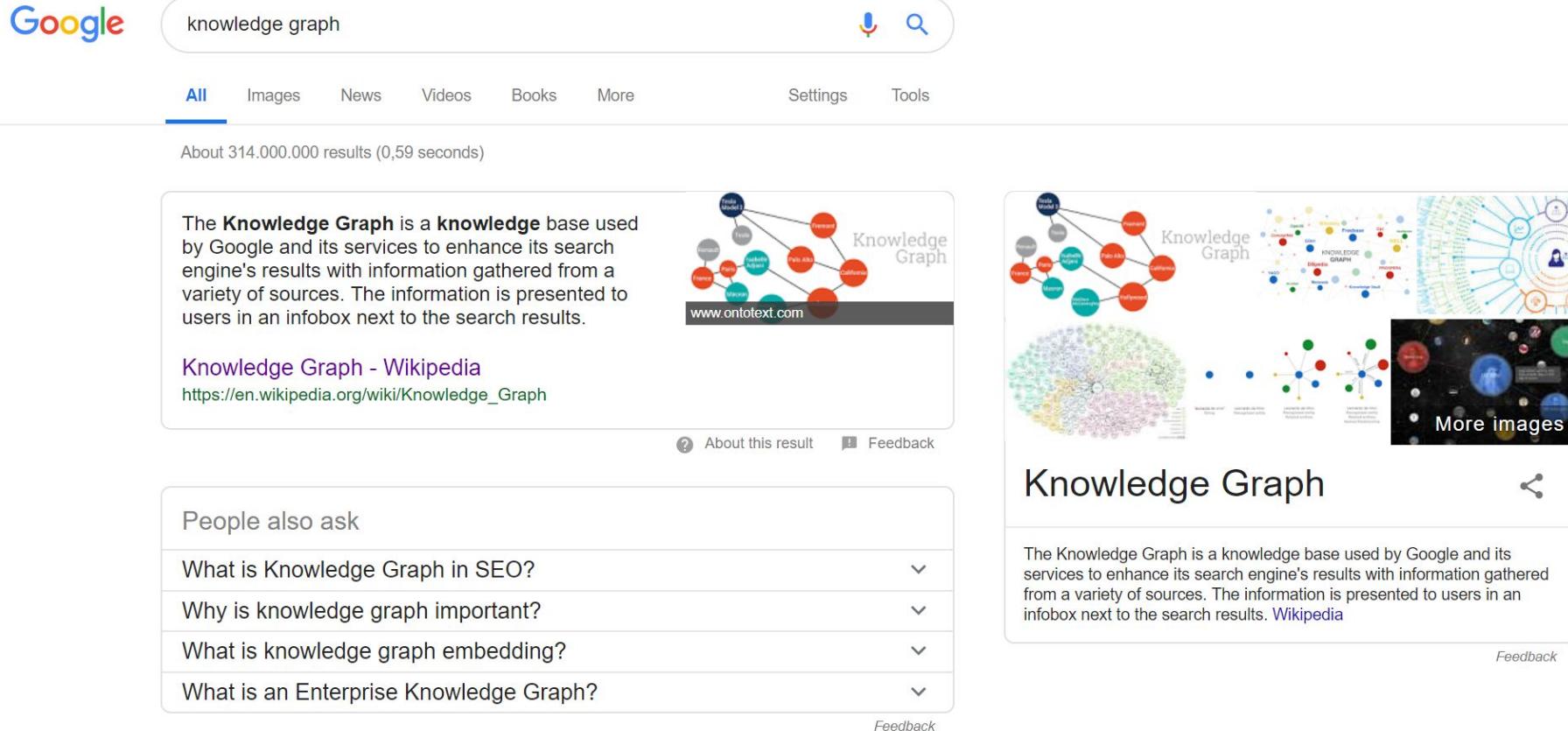
1974, Marchi and Miguel defined a knowledge graph as a mathematical structure with **vertices** as *knowledge units* connected by **edges** that represent the *prerequisite relation*.

In 2007, DBpedia and Freebase were created as graph based knowledge repositories (but did not use the term).

In 2012, Google introduces *the “Knowledge graph”* ...



<https://www.youtube.com/watch?v=mmQI6VGvX-c>



Netherlands - Wikipedia

<https://en.wikipedia.org/wiki/Netherlands> ▾

Location of the European **Netherlands** (dark green). – in Europe (green & dark grey) – in the European Union (green). Location of the Dutch special ...

Official languages: Dutch

Sovereign state: Kingdom of the Netherlands

Capital and largest city: Amsterdam; 52°22'N ...

Currency: Euro (€) (EUR); US dollar (\$) (USD) ...

Kingdom of the Netherlands · History of the Netherlands · Netherlands · Amsterdam

Visit the Netherlands: Destinations, tips and inspiration - Holland.com

<https://www.holland.com/global/tourism.htm> ▾

Welcome to Holland.com, the official website of the **Netherlands** Board of Tourism and Conventions, where you receive all the information for your visit to ...

20 Oct - 28 Oct **Dutch Design Week** Eindhoven

27 Oct - 28 Oct **Störrig Festival**

31 Oct - 4 Nov **Affordable Art Fair**

Government.nl: Information from the Government of The Netherlands

<https://www.government.nl/> ▾

Are you an international (non-EU) student and almost graduating or ending your research in the **Netherlands**? Watch back this webinar on The Orientation Year ...

Netherlands | World | The Guardian

<https://www.theguardian.com/world/netherlands> ▾

They've explored pollution in the Urals, freight trains in China, and daily life in the world capital of fake news. Now the Dutch art crusaders have imagined ...

Netherlands - Wikitravel



Netherlands

Country in Europe

The Netherlands, a country in northwestern Europe, is known for a flat landscape of canals, tulip fields, windmills and cycling routes.

Amsterdam, the capital, is home to the Rijksmuseum, Van Gogh Museum and the house where Jewish diarist Anne Frank hid during WWII. Canalside mansions and a trove of works from artists including Rembrandt and Vermeer remain from the city's 17th-century "Golden Age."

Capital: Amsterdam

Dialing code: +31

Official regional languages: West Frisian; Papiamento; English;

Points of interest: Keukenhof, Anne Frank House, Rijksmuseum, MORE

Did you know: Netherlands has the fifth-largest natural gas exports (53,650,000,000 cu m) in the world. [wikipedia.org](#)

People also search for



Belgium



Europe



Germany



Holland



Curaçao

Feedback

[powered by schema.org](#)

Google Knowledge Graph Applications



netherlands

Google Search

I'm Feeling Lucky

Netherlands - Wikipedia

<https://en.wikipedia.org/wiki/Netherlands> ▾

Location of the European Netherlands (dark green). – in Europe (green & dark grey) – in the European Union (green). Location of the Dutch special ...

Official languages: Dutch Sovereign state: Kingdom of the Netherlands
Capital and largest city: Amsterdam; 52°22'N ... Currency: Euro (€) (EUR); US dollar (\$) (USD) ...

[Kingdom of the Netherlands](#) · [History of the Netherlands](#) · [Netherlands](#) · [Amsterdam](#)



Netherlands

Country in Europe

The Netherlands, a country in northwestern Europe, is known for a flat landscape of canals, tulip fields, windmills and cycling routes. Amsterdam, the capital, is home to the Rijksmuseum, Van Gogh Museum and Rembrandt House. Other major cities include Rotterdam, Hague and Utrecht.

Caravaggio mansions and a trove of works from artists including Rembrandt and Vermeer remain from the city's 17th-century "Golden Age."

Capital: Amsterdam

Dialing code: +31

Official regional languages: West Frisian; Papiamento; English;

Points of interest (selected): Anne Frank House, Rijksmuseum, MORE

Data: Natural resources have the fifth largest natural gas exports

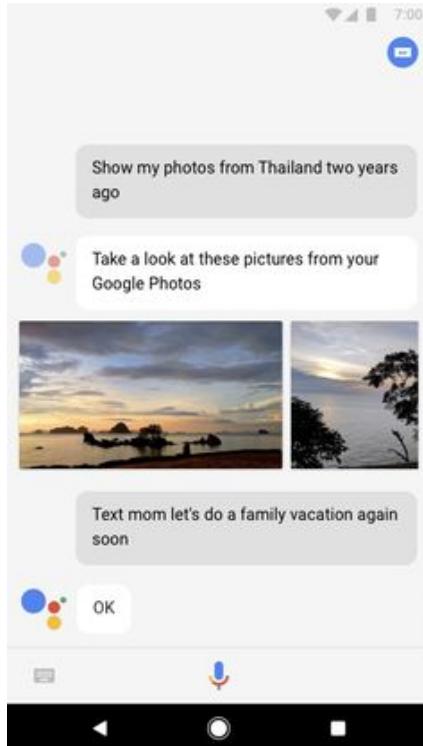
(53,600,000,000 m³) in the world; [wikipedia.org](#)

People also search for



Feedback

Google Search



Google Voice Assistant



Google Home Services

Personal Assistants



who is the best dutch football player?



All

Images

News

Videos

Shopping

More

Settings

Tools

About 51.100.000 results (0,74 seconds)

According to thetoptens.com

[View 10+ more](#)



Johan
Cruyff



Arjen
Robben



Dennis
Bergkamp



Frank de
Boer



Marco van
Basten



Ruud van
Nistelrooy



Robin van
Persie



Source: kgtutorial.github.io

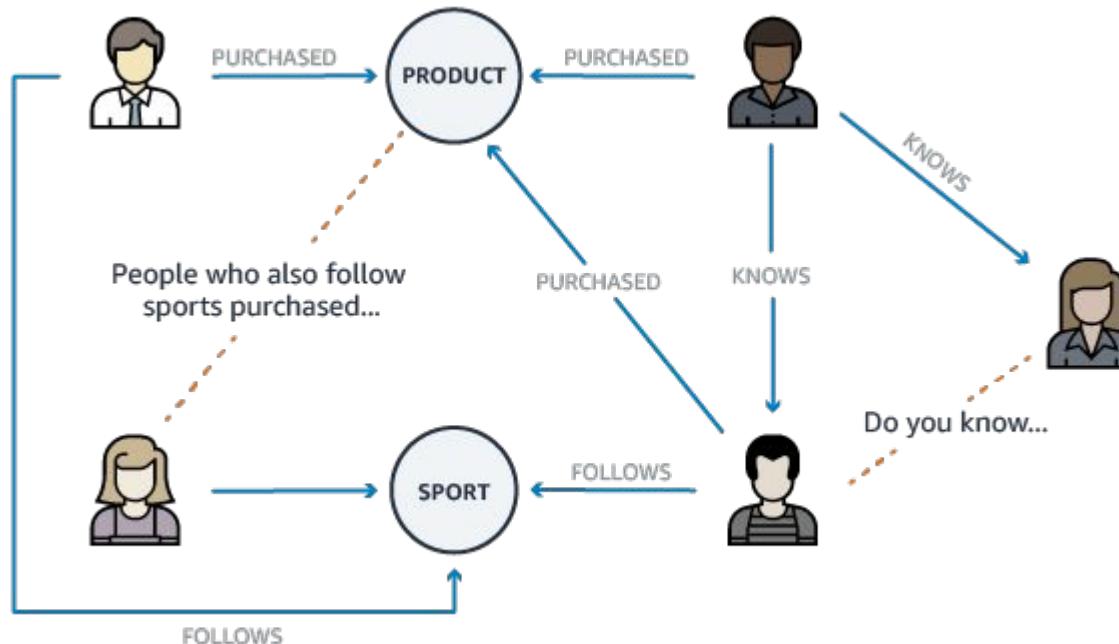
Semantic Question answering



- IBM Research undertook a challenge to build a computer system that could compete at the human champion level in real time on the American TV quiz show, Jeopardy
- Led to building of the DeepQA architecture and the implementation of Watson
- DeepQA leverages semistructured and structured content, which include databases, taxonomies, and ontologies, such as DBpedia, WordNet and the Yago ontology

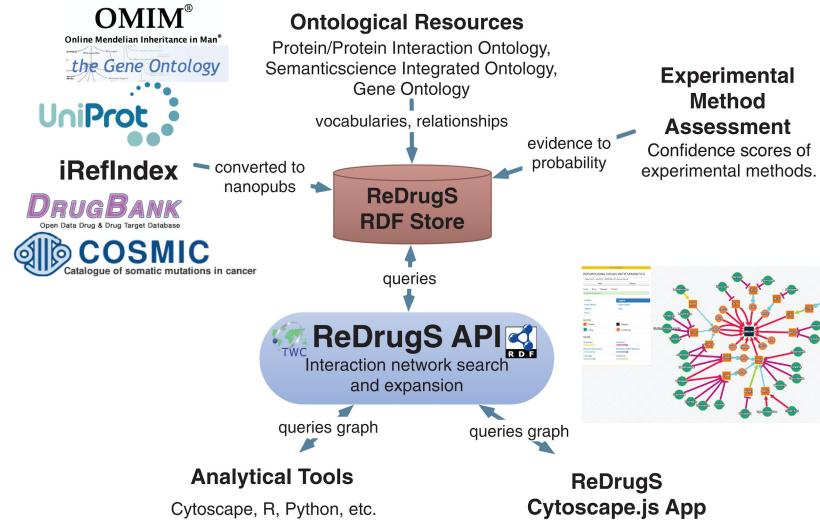
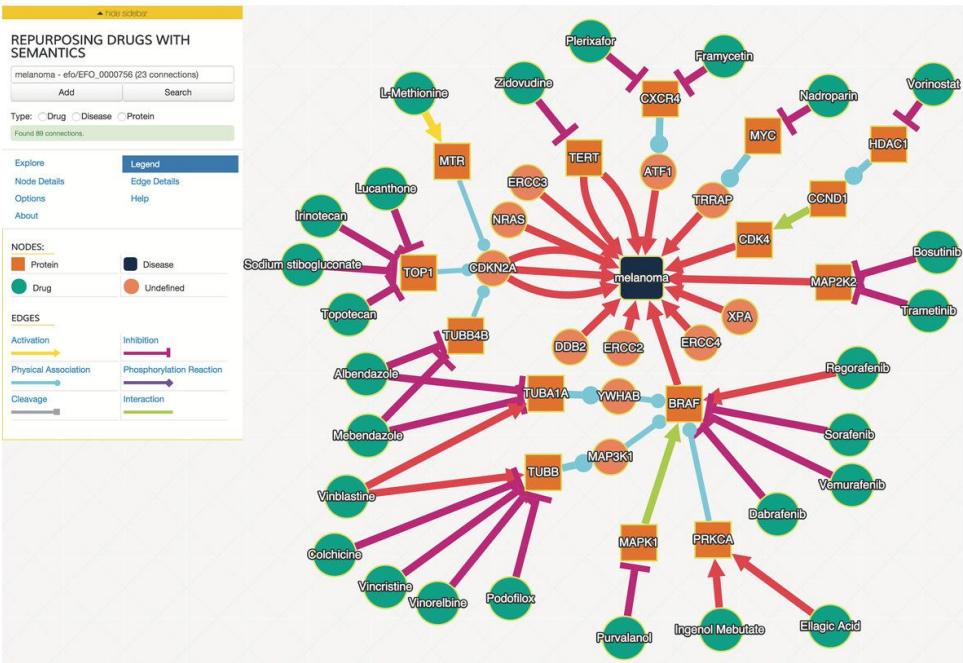
<http://www.aaai.org/Magazine/Watson/watson.php>

Recommendation Systems



Source: <https://geomarketing.com>

Discovery science



slido



What interests you about KGs?

- ⓘ Click **Present with Slido** or install our [Chrome extension](#) to activate this poll while presenting.

knowledge graph terminology

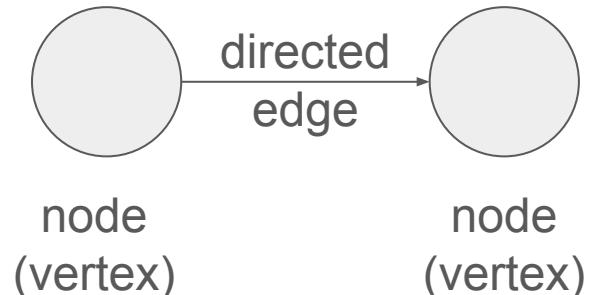
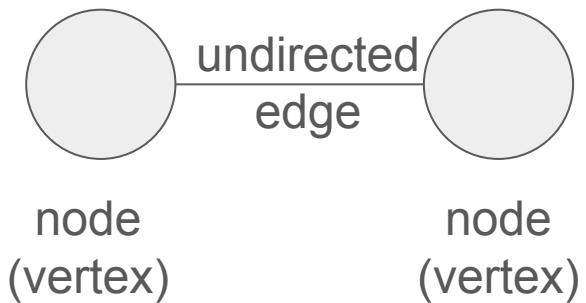
Graph (or Network)

A graph **G** is composed of two parts (**V,E**) where **V** is a set of nodes (vertices) and **E** is a set of edges (arcs, lines).

An **edge** is a pair (x,y) , where x and y belong to V .

An **undirected graph** is a graph in which edges do not have direction. The nodes are unordered pairs in the definition of every edge.

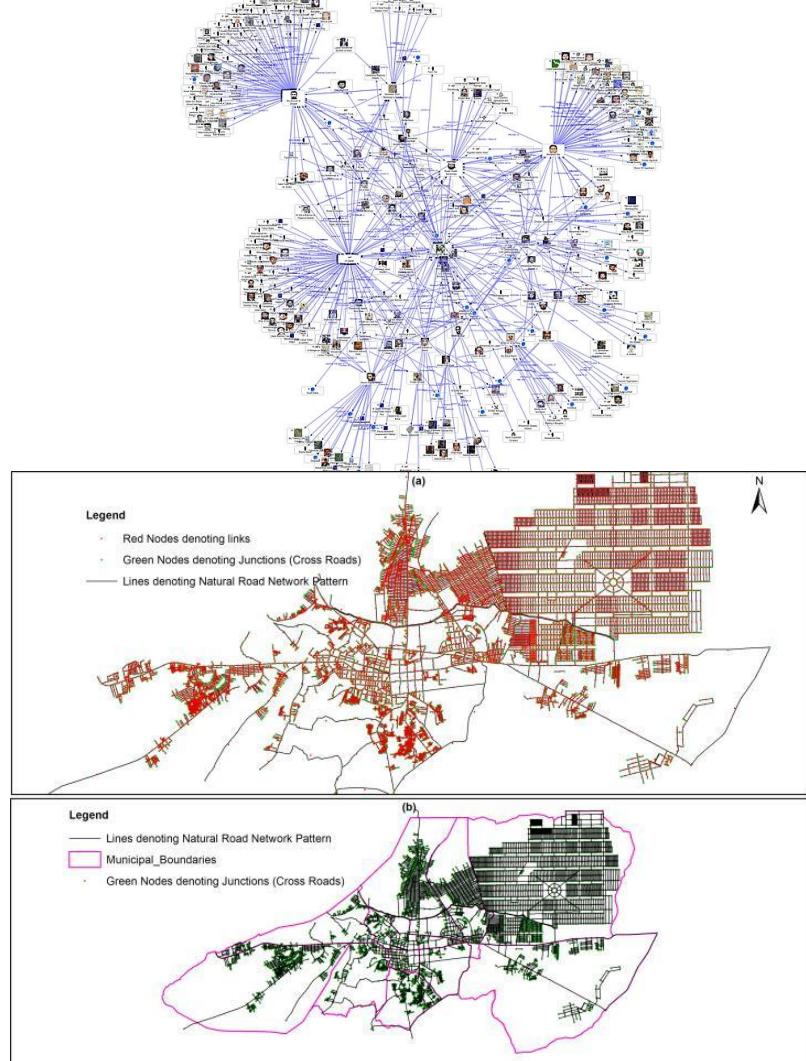
A **directed graph** is a graph in which an edge has a direction. The nodes are ordered pairs in the definition of every edge.



Graphs

Graphs can be used in various applications, such as computer networks, road maps, social networks, etc., to represent the connections between different entities.

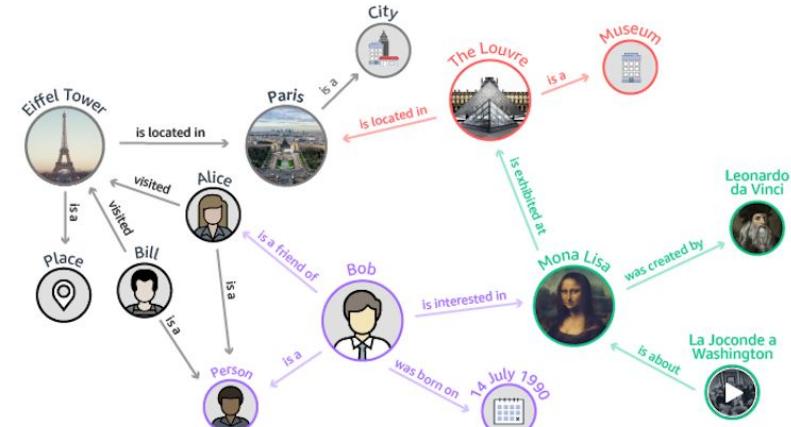
The graph data structure is amenable to specific types of analysis such as clustering, community detection, shortest path, page rank.



Knowledge Graphs

A **knowledge graph** is a type of graph designed specifically designed to capture the meaning (or semantics) of entities, their relationships and their attributes.

Unlike a basic graph, which could theoretically represent any kind of data in a network form, a knowledge graph is focused on **semantic knowledge representation**, enabling complex queries and reasoning about entities and their relationships.





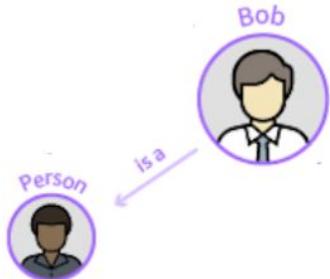
Bob



Bob is an **Entity**
(aka Thing aka Individual)

Bob is a **Person**

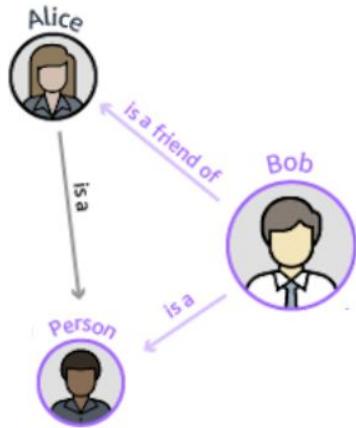
Person is a **Type** (aka **class**, **category**)



“is a” is a **Relation** (aka **property**, **predicate**)

Bob is **an instance of** the **Type Person**

Bob is **an Individual**



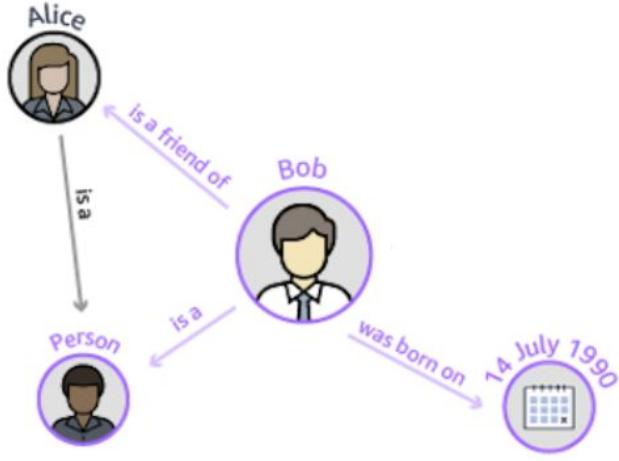
Alice is a Person

Alice is an Instance of Person

Bob is a friend of Alice

"is a friend of" is a Relation

the relation "is a friend of" **holds** between Bob and Alice

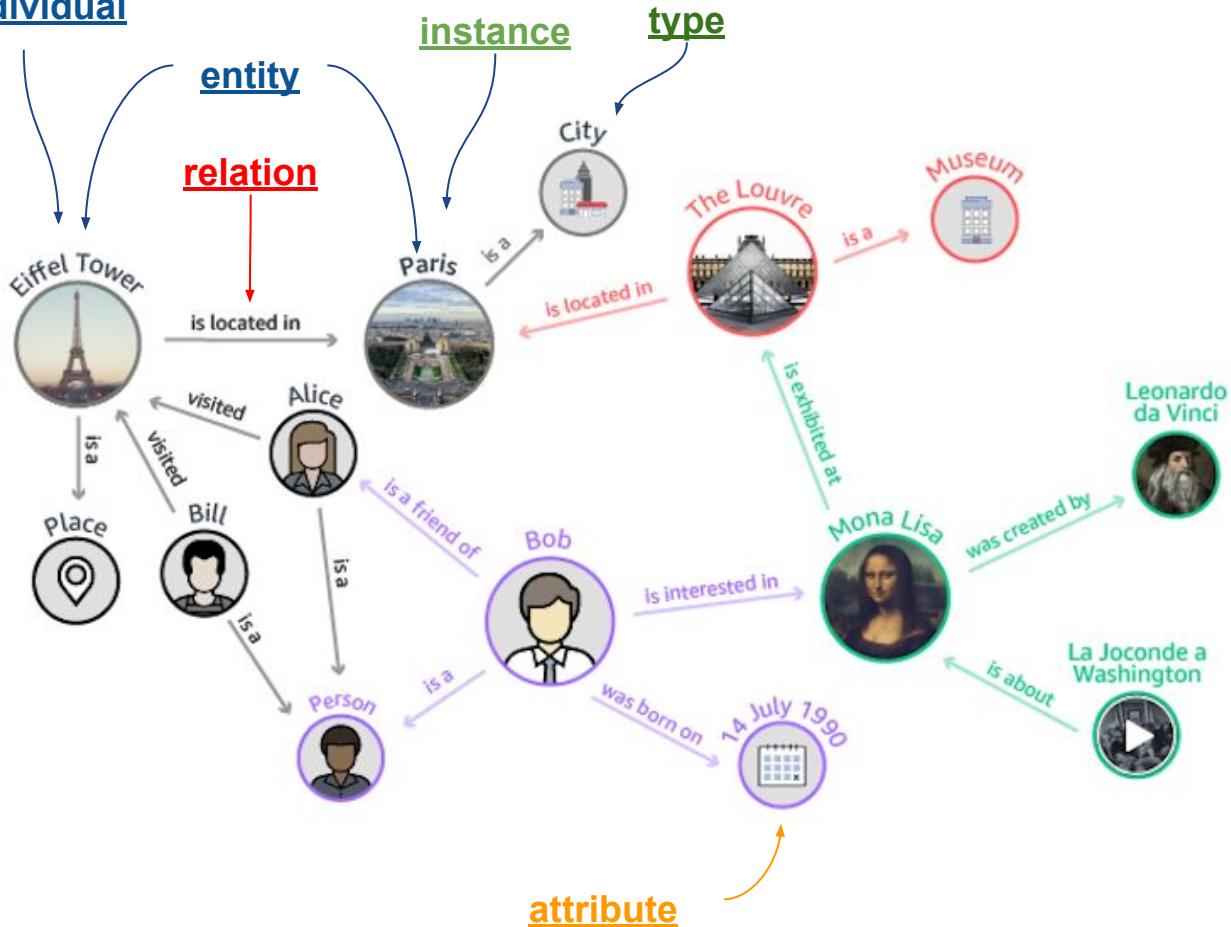


Bob was born on 14 July 1990

“was born on” is an **Attribute (aka Property)**

14 July 1990 is a **Date**
A **Date** is a **Datatype**

Individual

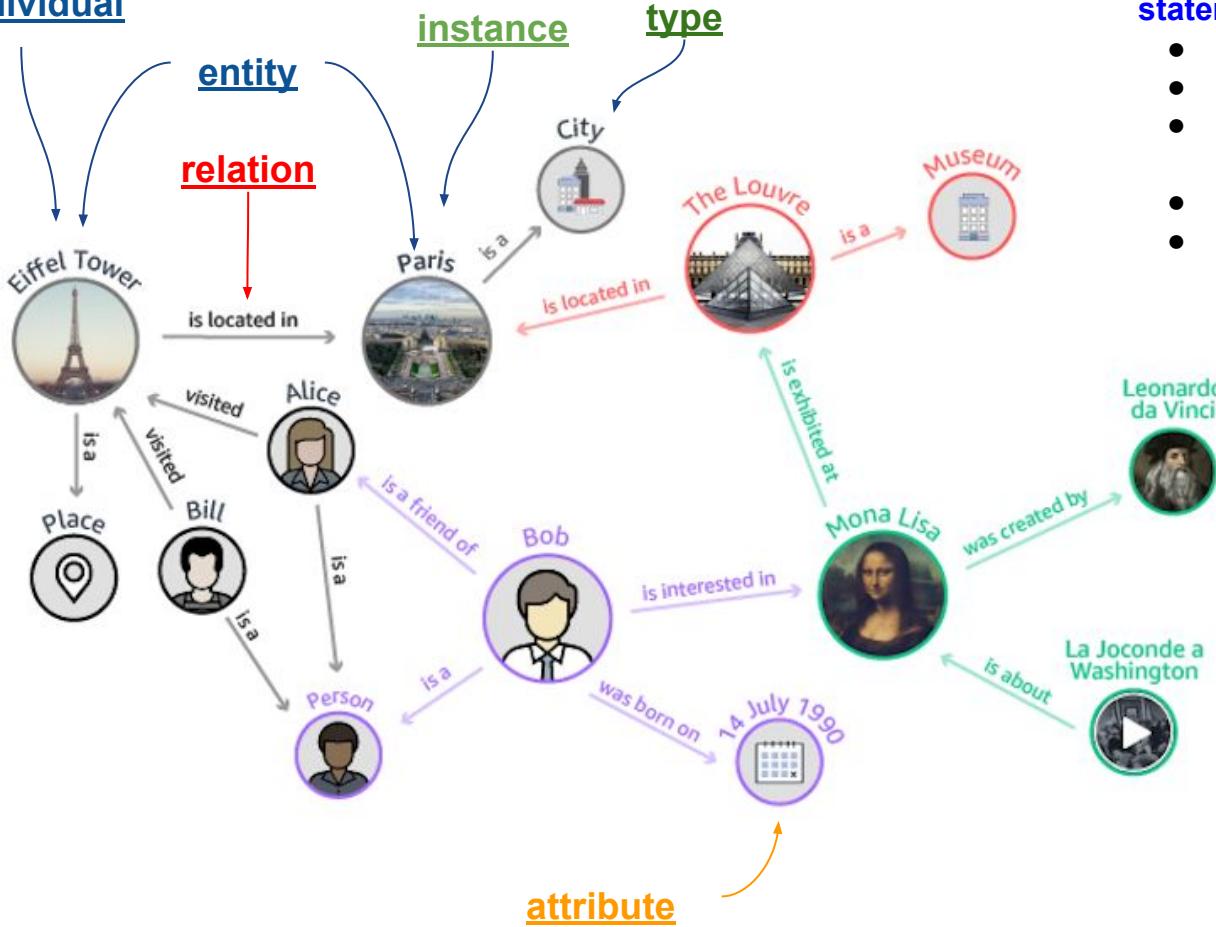


A (widely agreed) definition for the term knowledge graph is lacking.

“We define a knowledge graph as: a graph, composed of **a set of assertions (edges labeled with relations) that are expressed between entities (vertices)**, where the meaning of the graph is encoded in its structure, the relations and entities are unambiguously identified, a limited set of relations are used to label the edges, and the graph encodes the provenance, especially justification and attribution, of the assertions.”

McCusker et al. (2018), What is a Knowledge Graph?

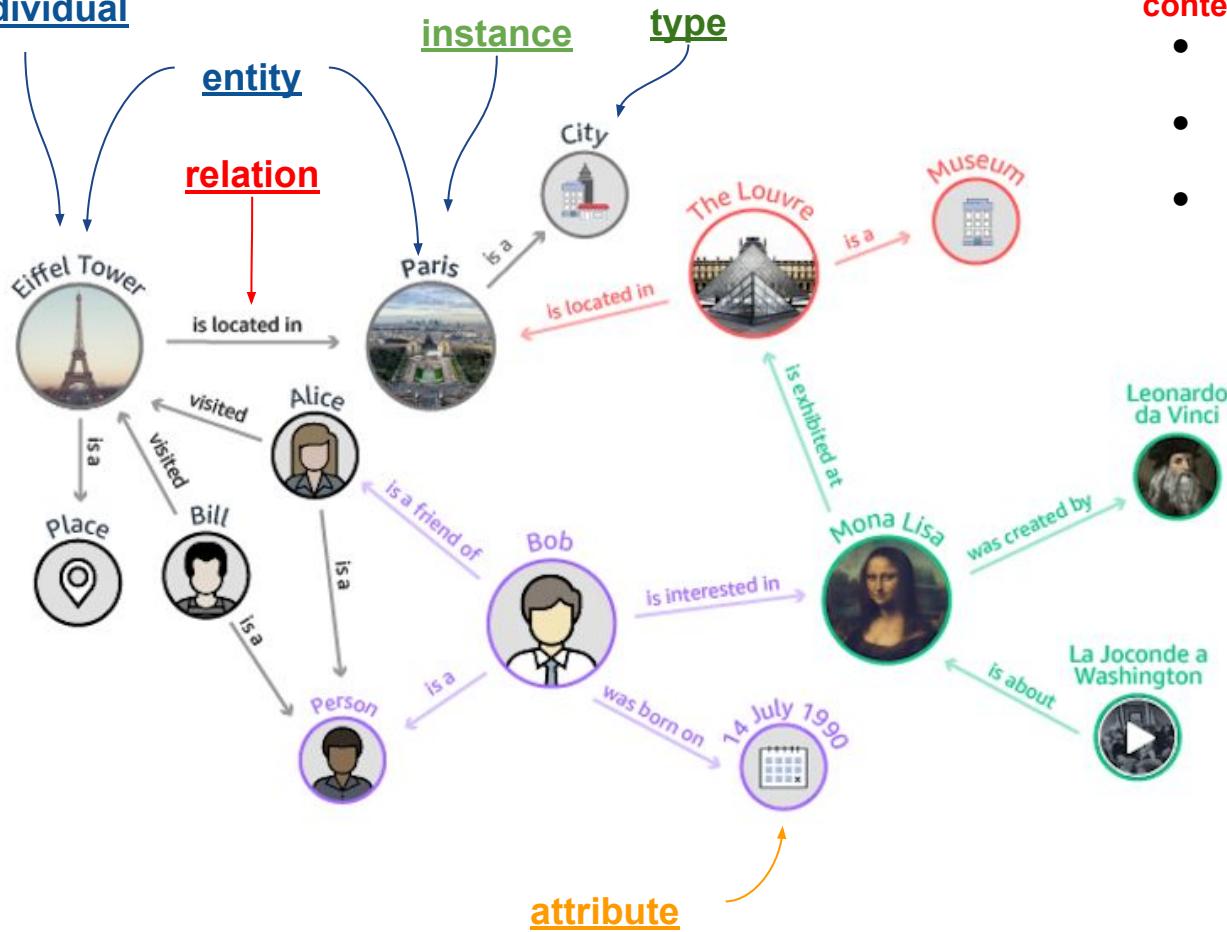
Individual



statement provenance:

- when was the statement made?
- who made the statement?
- where were they when the statement was made?
- how was it reported?
- what are the terms of (re)use?

Individual



context:

- when (in which time interval) does the relation hold true?
- is there a particular place in which this relation holds true?
- are there other agents that must be present for it to hold true?



You

what is a knowledge graph?



ChatGPT

A knowledge graph is a structured way to represent knowledge in a graph format, where entities (such as objects, events, situations, or concepts) are represented as nodes, and the relationships between these entities are represented as edges connecting the nodes. This format allows for a flexible and intuitive representation of complex networks of information.

Knowledge graphs are widely used in various applications, including search engines, recommendation systems, social network analysis, natural language processing, and more. They enable machines to understand and infer information about the world in a more human-like way, by leveraging the relationships and properties defined within the graph.

Here's a breakdown of the key components of a knowledge graph:

- **Nodes (or vertices):** Represent the entities or concepts within the knowledge domain. Entities can be tangible objects like "Paris" or "Apple Inc." or abstract concepts like "happiness" or "economics."
- **Edges (or links):** Represent the relationships or associations between the entities. For example, an edge might connect the "Leonardo DiCaprio" node to the "Titanic" node with the relationship "acted in."
- **Properties (or attributes):** Information that describes the nodes. For example, the node representing "Albert Einstein" might have properties such as "birthdate: March 14, 1879" and "field: Physics."

Knowledge graphs are powerful because they enable complex queries and analyses by understanding the relationships and attributes of the entities within them. They are foundational to many semantic web technologies and have become integral to enhancing the capabilities of artificial intelligence systems to process and understand vast amounts of information.

A Knowledge Graph uses Ontologies to infer new facts

An **ontology** is a formal representation of knowledge amenable to automated reasoning. Ontology defines the kinds of things that exist and what relations may or may not hold between them.

The Ontology adds background knowledge to the KG: Very useful to infer new facts, and to check the quality (semantic coherence) of the KG.

Consider
KG Facts:

- The Eiffel tower is-a **Place**
- The Louvre is-a **Museum**
- Paris is-a **City**

An Ontology could define:

- Every Museum is a **Place**
- Every City is a **Place**

In a logic that defines:

- **is-a as a transitive property**, such that x is-a y, y is-a z, then x is-a z

Therefore, we can infer that

- The Louvre is a **Place**
- Paris is a **Place**

slido



What primarily distinguishes a knowledge graph from a basic graph?

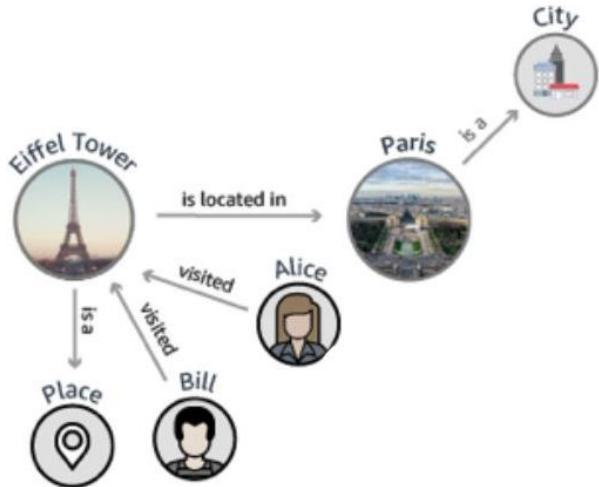
- ⓘ Click Present with Slido or install our [Chrome extension](#) to activate this poll while presenting.

Why represent knowledge as graphs?

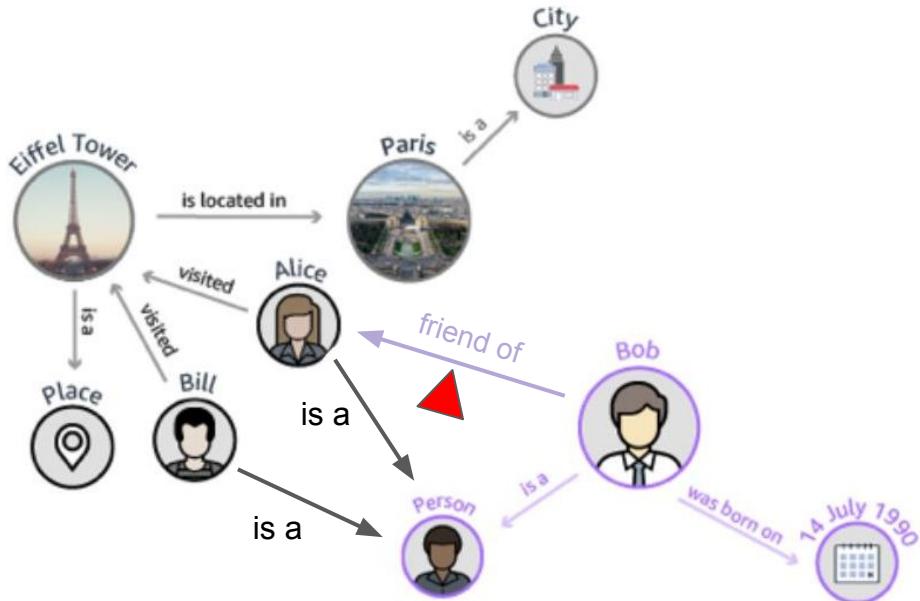
Graphs as a KR paradigm

- Accessible to humans *and* machines
 - Graph paradigm is intuitive for humans to understand
 - Graphs can be encoded with a formal semantics for machine inference
- Facilitates integration of heterogeneous knowledge
 - Use of shared identifiers
 - Use of shared vocabularies
 - **Creating links** between entities from different datasets
- Enables analysis
 - Decades of research on graph analysis & querying algorithms can be reused to explore, mine and analyse KGs

travel network



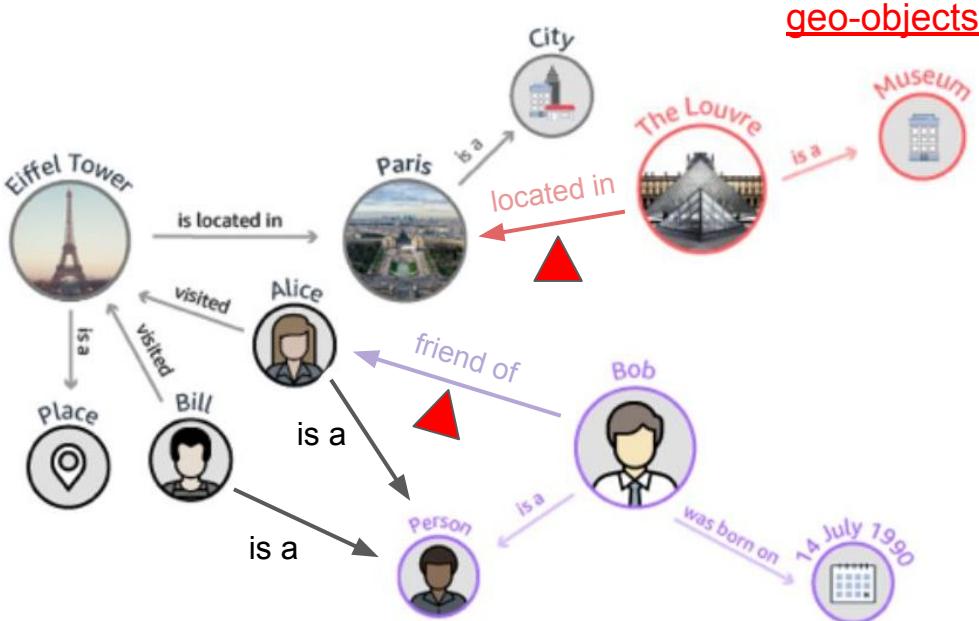
travel network



social network

▲ linking relations

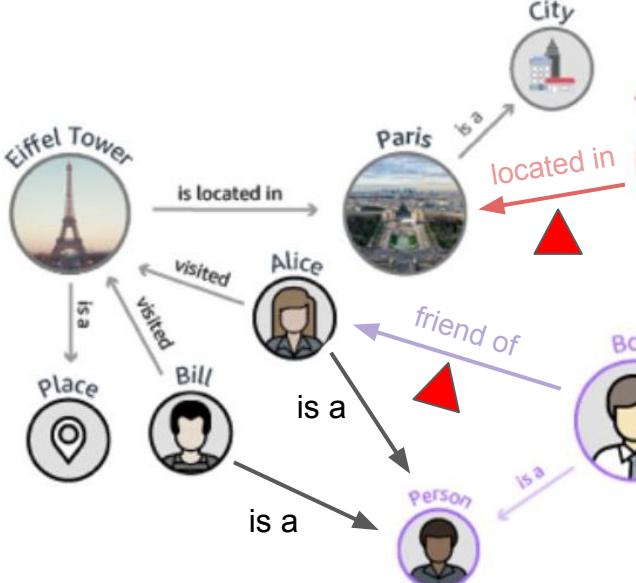
travel network



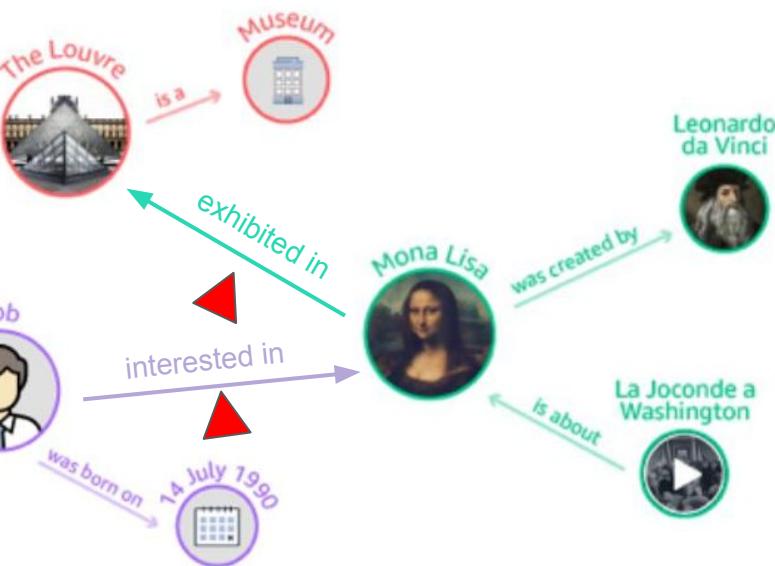
geo-objects

▲ linking relations

travel network



geo-objects



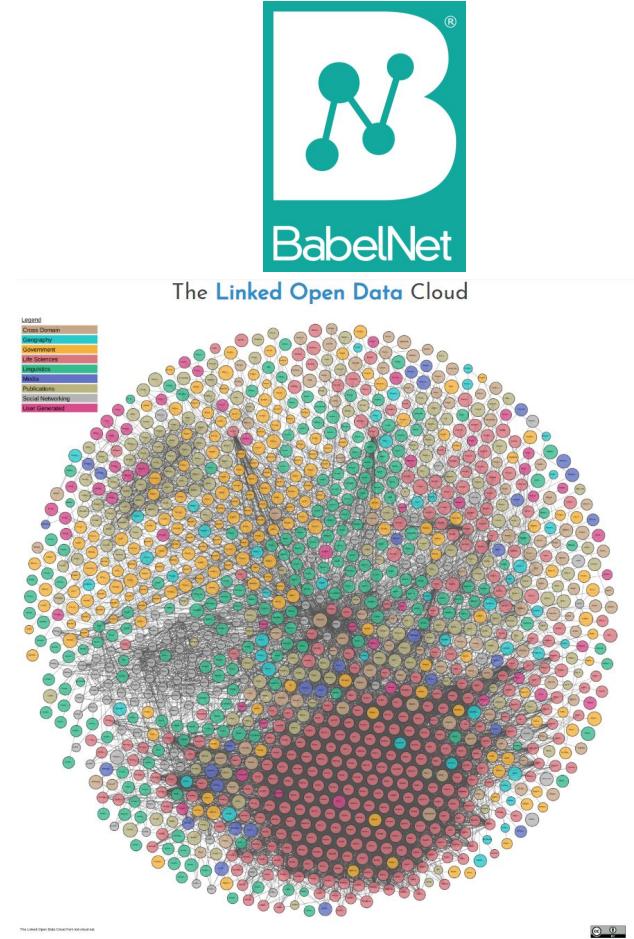
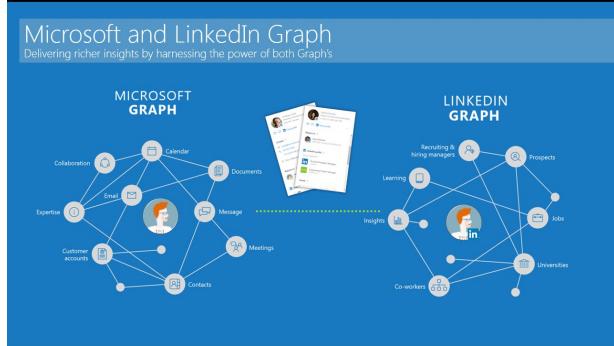
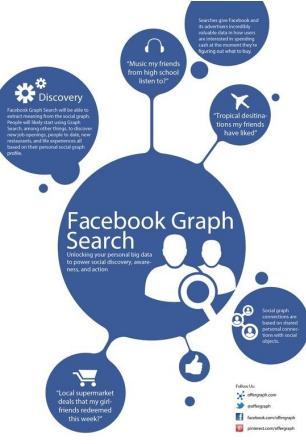
art history

social network

▲ linking relations

A photograph of five large, cylindrical metal grain silos standing in a field of tall, dry, golden-brown grass. The silos have dark, conical roofs and are arranged in a slightly staggered line from left to right. The sky above is a clear, pale blue.

Data silos – not made for sharing





WIKIPEDIA

The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item
Cite this page

Print/export
Create a book
Download as PDF
Printable version

In other projects

Wikimedia Commons
Wikinews
Wikiquote
Wikivoyage

Languages

Deutsch

Article Talk

Read

Edit

View history

Search Wikipedia



Wiki Loves Love: Documenting festivals and celebrations of love on Commons.
Help Wikimedia and win prizes by sending photos.

Netherlands

From Wikipedia, the free encyclopedia

Coordinates: 52°19'N 5°33'E

This article is about the constituent country. For the sovereign state, see Kingdom of the Netherlands. For other uses, see Netherlands (disambiguation).
Not to be confused with the region of Holland, or the Low Countries.

The **Netherlands** (Dutch: *Nederland*, [neːdərlãnt] (listen)) is a country located mainly in Northwestern Europe. Together with three island territories in the Caribbean (Bonaire, Sint Eustatius and Saba), it forms a constituent country of the Kingdom of the Netherlands. The European portion of the Netherlands consists of twelve provinces and borders Germany to the east, Belgium to the south and the North Sea to the northwest, with maritime borders in the North Sea with Belgium, Germany and the United Kingdom.^[12] The five largest cities in the Netherlands are Amsterdam, Rotterdam, The Hague, Utrecht and Eindhoven. Amsterdam is the country's capital,^[13] while The Hague holds the seat of the States General, Cabinet and Supreme Court.^[14] The Port of Rotterdam is the largest port in Europe and the largest outside Asia.^[15]

'Netherlands' literally means 'lower countries', referring to its low and flat topography, with only about 50% of its land exceeding 1 metre (3 ft 3 in) above sea level and nearly 17% being below sea level.^[16] Most of the areas below sea level, known as *polders*, are the result of land reclamation beginning in the 16th century. With a population of 17.25 million living within a total area of roughly 41,500 square kilometres (16,000 sq mi)—of which the land area is 33,700 square kilometres (13,000 sq mi)—the Netherlands is one of the most densely populated countries in the world. Nevertheless, it is the world's second-largest exporter of food and agricultural products after the United States, owing to its fertile soil, mild climate, and intensive agriculture.^{[17][18]}

The Netherlands was the third country in the world to have representative government and has been a parliamentary constitutional monarchy with a unitary structure since 1848. Historically, the country has a tradition of pillarisation and a long record of social tolerance, having legalised abortion, prostitution and euthanasia, while maintaining a progressive drug policy. The Netherlands abolished the death penalty in 1870, allowed women's suffrage in 1917 and became the world's first country to legalise same-sex marriage in 2001. It is a founding member of the EU, Eurozone, G10, NATO, OECD and WTO, as well as a part of the Schengen Area and the trilateral Benelux Union. It hosts several intergovernmental organisations and international courts, many of which are centered in The Hague, which is consequently dubbed 'the world's legal capital'.^[19] Its mixed-market advanced economy had the thirteenth-highest per capita income globally. One of the world's most prosperous countries, the Netherlands ranks among the highest in international indexes of press freedom,^[20] economic freedom,^[21] human development, quality of life, as well as happiness.^[22] Its strong performance is owed in large part to a generous welfare state that provides universal healthcare, public education and infrastructure, in addition to social benefits. It is also known for its *polder model*, the country's leading socioeconomic model based on consensus decision-making.

Netherlands

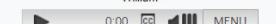
Nederland (Dutch)



Flag Coat of arms

Motto: *'Je maintiendrai'* (French)^[a]
'Ik zal handhaven' (Dutch)
'I will uphold'
Other historic mottoes: [\[show\]](#)

Anthem: *'Wilhelmus'* (Dutch)
'William'



0:00 ▶ MENU



Location of the European Netherlands (dark green)
– in Europe (green & dark grey)
– in the European Union (green)

Capital and largest city	Amsterdam ^[c] 52°22'N 4°53'E
Government seat	The Hague ^[c]
Official languages	Dutch ^[2]
Official regional languages	West Frisian Papiamento English ^[d]
Recognised regional languages	Limburgish Dutch Low Saxon ^[d]
Ethnic groups (2017 ^[5])	77.39% Dutch 9.88% Other Europeans 2.34% Turks 2.29% Moroccans 2.13% Indonesians 2.05% Surinamese 0.90% Dutch Caribbeans 0.23% Other Americans 2.80% Other
Religion (2015 ^[6])	50.1% Irreligion 43.8% Christianity 4.9% Islam 1.1% Sikhism, Hinduism, Buddhism, Judaism
Demonym(s)	Dutch
Sovereign state	Kingdom of the Netherlands
Government	Unitary parliamentary constitutional monarchy
• Monarch	Willem-Alexander
• Prime Minister	Mark Rutte (VVD)
• Deputy Prime Ministers	Hugo de Jonge (CDA) Kajsa Ollongren (D66) Carola Schouten (CU)
• Vice President of the Council of State	Thom de Graaf (D66)
Legislature	States General Senate House of Representatives
Independence from Spain	
• Proclaimed	26 July 1581
• Recognised	30 January 1648
• Kingdom established	16 March 1815
• Liberation Day	5 May 1945
• Admitted to the United Nations	10 December 1945
• Constituent country	15 December 1954
• Incorporation of the Caribbean Netherlands	10 October 2010



- Crowdsourced community effort to extract structured content from information created in various Wikimedia projects
- Structured information resembles an open knowledge graph which is available for everyone on the Web
- English version of the DBpedia knowledge base describes
 - 4.58 million things,
 - out of which 4.22 million are classified in a consistent ontology,
 - including 1,445,000 persons, 735,000 places , 241,000 organizations etc.
- Localized versions of DBpedia in 125 languages
 - 8.8 billion fact, large-scale-cross-domain knowledge graph

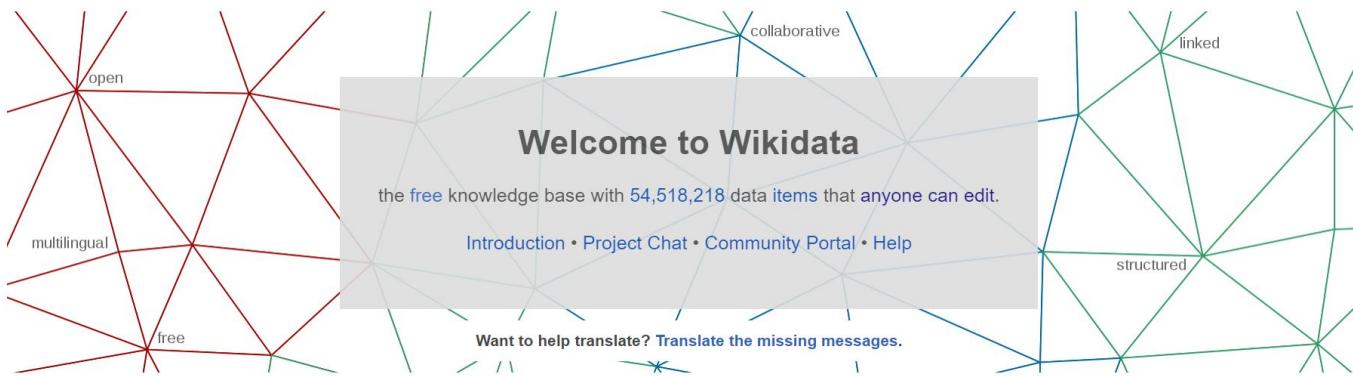
About: Maastricht

An Entity of Type : List of municipalities of the Netherlands, from Named Graph : <http://dbpedia.org>, within Data Space : dbpedia.org

Maastricht (Dutch: [maːst̪rɪxt̪] ; Limburgish (incl. Maastrichtian): Mestreech [mə'st̪reːχ]; French: Maestricht (archaic); Spanish: Mastrique (archaic)) is a town and a municipality in the southeast of the Netherlands. It is the capital city of the province of Limburg. Maastricht is located on both sides of the Meuse river (Dutch: Maas), at the point where the Jeker River (French: Geer) joins it.

Property	Value
dbo:abstract	<ul style="list-style-type: none">▪ Maastricht (Dutch: [maːst̪rɪxt̪] ; Limburgish (incl. Maastrichtian): Mestreech [mə'st̪reːχ]; French: Maestricht (archaic); Spanish: Mastrique (archaic)) is a town and a municipality in the southeast of the Netherlands. It is the capital city of the province of Limburg. Maastricht is located on both sides of the Meuse river (Dutch: Maas), at the point where the Jeker River (French: Geer) joins it. Maastricht developed from a Roman settlement to a Medieval religious centre, a garrison town and an early industrial city. Today, Maastricht is well-regarded as an affluent cultural center. Maastricht has 1677 national heritage sites (Rijksmonumenten), which is the second highest number in a Dutch town, after Amsterdam. It has become known, by way of the Maastricht Treaty, as the birthplace of the European Union, European citizenship, and the single European currency, the euro. The town is popular with tourists for shopping and recreation, and has a large growing international student population. Maastricht is a member of the Most Ancient European Towns Network and is part of the Meuse-Rhine Euroregion, which includes nearby cities Aachen, Eupen, Hasselt, Liège and Tongeren. (en)
dbo:anthem	<ul style="list-style-type: none">▪ dbr:Maastricht
dbo:areaCode	<ul style="list-style-type: none">▪ 043
dbo:country	<ul style="list-style-type: none">▪ dbr:Netherlands
dbo:elevation	<ul style="list-style-type: none">▪ 49.000000 (xsd:double)
dbo:governingBody	<ul style="list-style-type: none">▪ dbr:Municipal_council_(Netherlands)
dbo:isPartOf	<ul style="list-style-type: none">▪ dbr:Limburg_(Netherlands)
dbo:leaderName	<ul style="list-style-type: none">▪ dbr:Annemarie_Penn-te_Strake
dbo:leaderTitle	<ul style="list-style-type: none">▪ Mayor (en)

<http://dbpedia.org/page/Maastricht>



■■■ Welcome!

Wikidata is a free and open knowledge base that can be read and edited by both humans and machines.

Wikidata acts as central storage for the **structured data** of its Wikimedia sister projects including Wikipedia, Wikivoyage, Wikisource, and others.

Wikidata also provides support to many other sites and services beyond just Wikimedia projects! The content of Wikidata is [available under a free license](#), exported using standard formats, and can be interlinked to other [open data sets](#) on the linked data web.

Wikidata Query Service
[Examples](#)
[Help](#)
[More tools](#)
⇄ English

```

1 #Cats
2 SELECT ?item ?itemLabel
3 WHERE
4 {
5   ?item wdt:P31 wd:Q146.
6   SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en". }
7 }
```

▶

⟳

🕒
?
144 results in 1002 ms
</> Code
Download
Link

item	itemLabel
wd:Q28792126	Gli
wd:Q30600575	Orlando

E.g. <https://www.wikidata.org/wiki/Q1309>

Graph Knowledge Representation

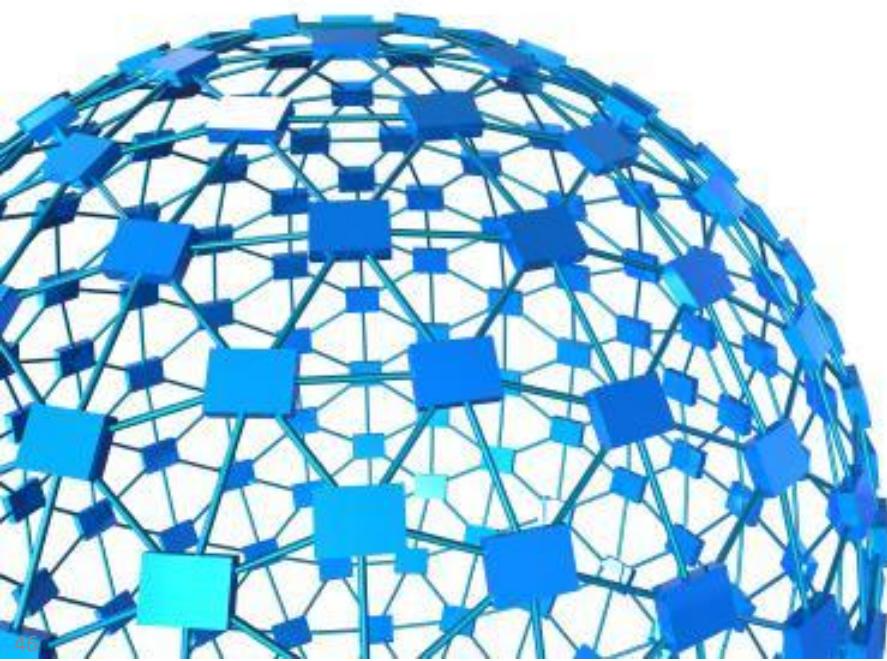
Long history of KR & graph representation

- Existential Graphs (1896; Peirce)
- Semantic Networks
- Conceptual Graphs
- Cognitive Semantic Networks
- Structured Inheritance Networks
- Multilayered Extended Semantic Networks (MultiNets)
- Basic Conceptual Graphs
- Full Conceptual Graphs
- Hierarchical Semantic Form
- **Semantic Web Technologies**
- Property Graphs

The Semantic Web is a portal to the web of knowledge

standards for publishing, sharing and querying
facts, expert knowledge and services

scalable approach for the discovery
of *independently constructed,
collaboratively described,
distributed knowledge
(in principle)*



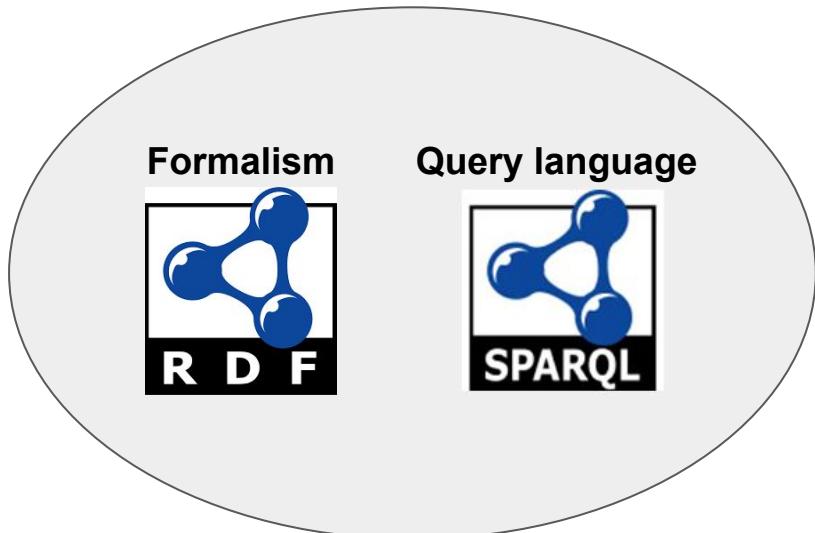
Resource Description Framework

Graph-based method for knowledge representation on the world wide web.

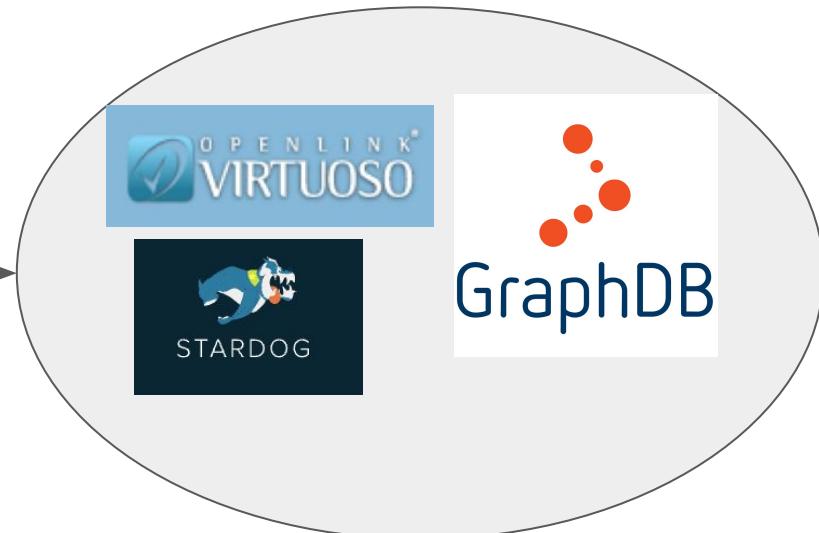
- All things are named with **Internationalized Resource Identifiers** (IRIs)
- Binary relations are *directed*, and expressed using a *named* property (annotations on the named property are shared in all graphs containing it).
- Values of relations can be a IRI, blank node, or a Literal



Semantic knowledge graphs

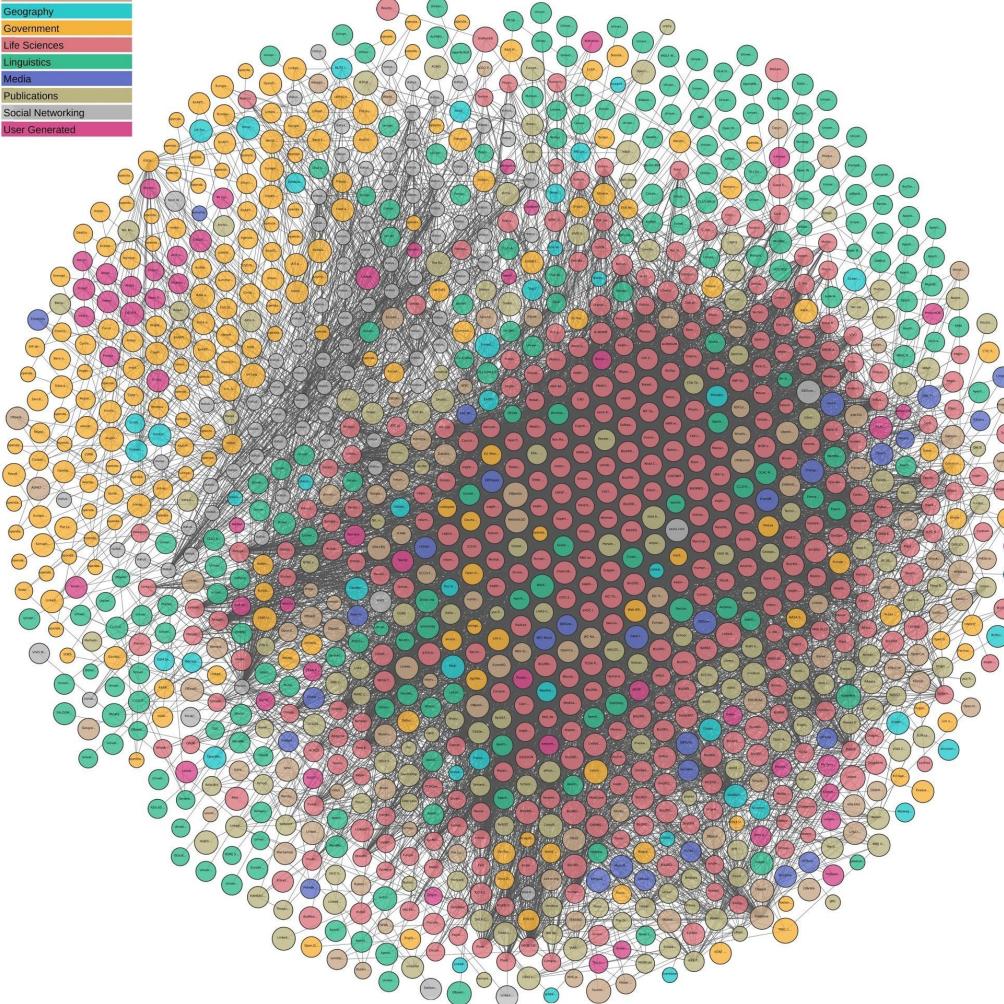


RDF store implementations



See <https://www.w3.org/wiki/LargeTripleStores>

Geography
Government
Life Sciences
Linguistics
Media
Publications
Social Networking
User Generated

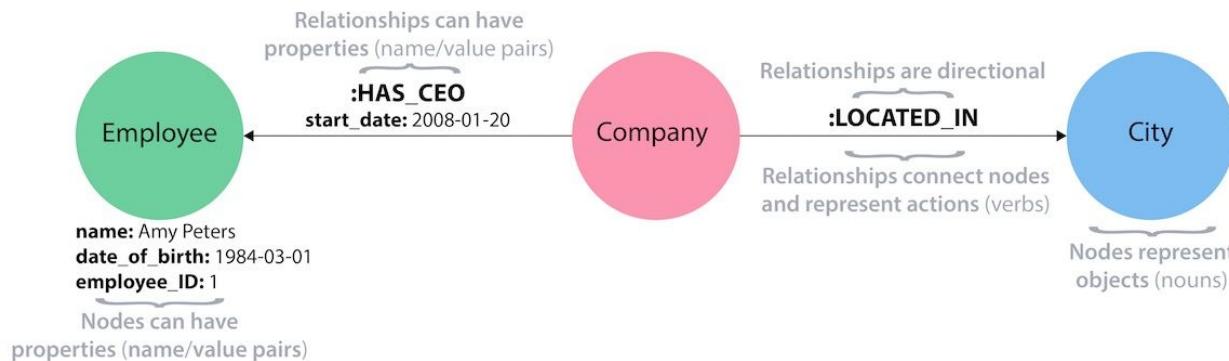


<https://lod-cloud.net/>

1,255 datasets
with 16,113 links
(as of end 2022)

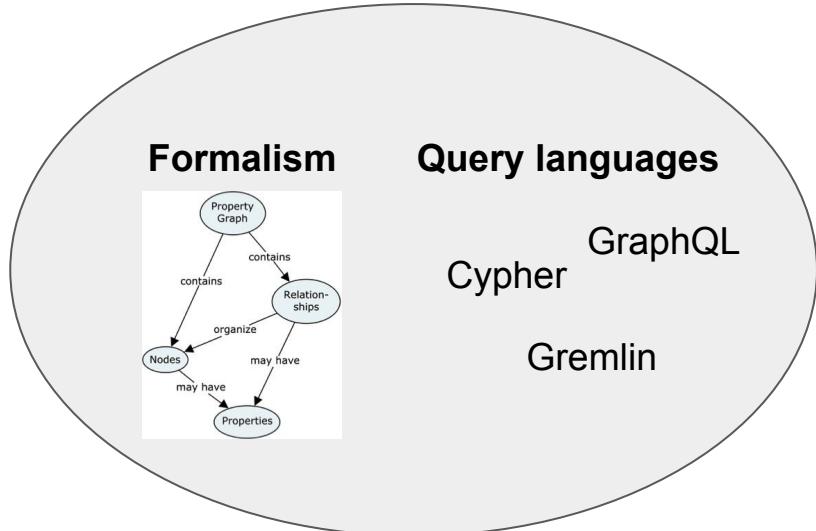
Property Graph

a property graph is an **attributed, multi-relational graph**. edges are labeled and both vertices and edges can have any number of key/value properties associated with them

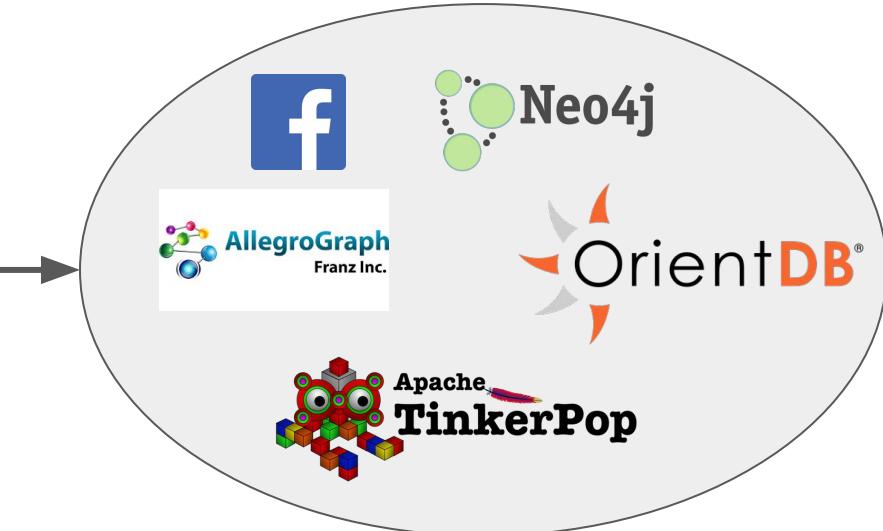


Limitations: No explicit semantics for the graph, identifiers are not necessarily unique, no semantics for types and relations

Property graphs



Graph database implementations



Summary

- Knowledge Graphs are very old and very new at the same time
- Many definitions of knowledge graphs
- Many representations for knowledge graphs
- Many instances of knowledge graphs
- Many ways to access knowledge graphs

You will learn how to build and mine specific representations and instances of knowledge graphs in a very short period of time!

slido



What do you hope to learn in
this course?

- ⓘ Click **Present with Slido** or install our [Chrome extension](#) to activate this poll while presenting.

Questions?