

# Knowledge Graphs

Lecture 3 - Querying RDFS with SPARQL

Excursion 2: DBpedia Knowledge Graph

Prof. Dr. Harald Sack & Dr. Mehwish Alam

FIZ Karlsruhe - Leibniz Institute for Information Infrastructure

AIFB - Karlsruhe Institute of Technology

Autumn 2020



KIT  
Leibniz-Institut für Technologie



Leibniz-Institut für Informationsinfrastruktur

# Knowledge Graphs

## Lecture 3: Querying RDF(S) with SPARQL

### 3.1 How to Query RDF(S)

#### Excursion 2: DBpedia Knowledge Graph

#### Excursion 3: Wikidata Knowledge Graph

### 3.2 Complex Queries with SPARQL

### 3.3 More Complex SPARQL Queries

### 3.4 SPARQL Subqueries and Property Paths

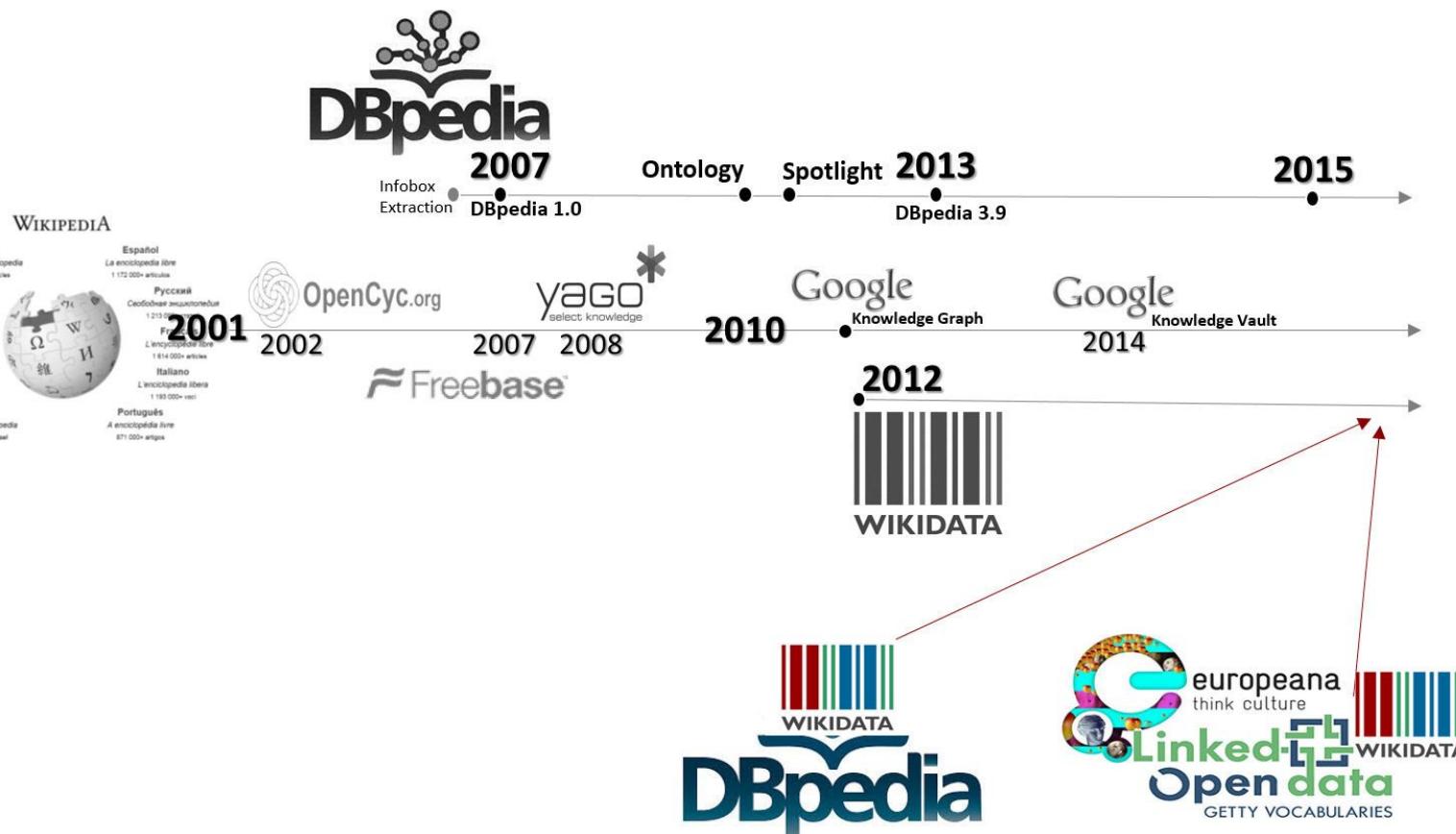
### 3.5 RDF Databases

### 3.6 SPARQL is more than a Query Language



- a central hub in the Web of Data

# Knowledge Bases Timeline



# DBpedia Knowledge Graph

**English version of the DBpedia Knowledge Graph** (as by January 2020)

- describes 6.6 million things,
- of which 5.5 million are classified in a consistent ontology
- including 1.5 million persons,
- 840,000 places (including 513,000 populated places),
- 496,000 creative works
  - including 139,000 music albums,
  - 111,000 films and
  - 21,000 video games,
- 286,000 organizations
  - including 70,000 companies and 55,000 educational institutions,
- 306,000 species and
- 6,000 diseases.



<https://wiki.dbpedia.org/develop/datasets>

# From Wikipedia to DBpedia

[http://en.wikipedia.org/wiki/Carbon\\_dioxide](http://en.wikipedia.org/wiki/Carbon_dioxide)



**WIKIPEDIA**  
The Free Encyclopedia

Screenshot of the Wikipedia page for Carbon dioxide. The page includes a sidebar with links like Main page, Contents, and Interaction, and a main content area with text about the chemical properties and sources of CO<sub>2</sub>. A red box highlights the right-hand sidebar, which contains chemical structures, names, identifiers, and other data.

**Chemical Structure:**

$\text{O}=\text{C}=\text{O}$   
116.3 pm

**Names:**

- Carbonic acid gas
- Carbonic anhydride
- Carbonic oxide
- Carbon oxide
- Carbon(V) oxide
- Dry ice (solid phase)

**Identifiers:**

CAS Number	124-38-9 ✓
3D model (JSmol)	Interactive image ↗ Interactive image ↗
3DMet	B01131 ↗
Beilstein	1900390
Reference	
ChEBI	CHEBI:165266 ✓
ChEMBL	ChEMBL1231871 ↗ ✗
ChemSpider	274 ↗ ✓
ECHA InfoCard	100.004.271 ↗
EC Number	204-696-9
E number	E290 (preservatives)
Gmelin Reference	989
EGG	D00004 ↗ ✓



[http://dbpedia.org/resource/Carbon\\_dioxide](http://dbpedia.org/resource/Carbon_dioxide)

# From Wikipedia to DBpedia

[http://dbpedia.org/resource/Carbon\\_dioxide](http://dbpedia.org/resource/Carbon_dioxide)

DBpedia Browse using ▾ Formats ▾ Faceted Browser Sparql Endpoint

## About: Carbon dioxide

An Entity of Type : [chemical compound](#), from Named Graph : [http://dbpedia.org](#), within Data Space : [dbpedia.org](#)

Carbon dioxide (chemical formula CO<sub>2</sub>) is a colorless and odorless gas vital to life on Earth. This naturally occurring chemical compound is composed of a carbon atom covalently double bonded to two oxygen atoms. Carbon dioxide exists in Earth's atmosphere as a trace gas at a concentration of about 0.04 percent (400 ppm) by volume. Natural sources include volcanoes, hot springs and geysers, and it is freed from carbonate rocks by dissolution in water and acids. Because carbon dioxide is soluble in water, it occurs naturally in groundwater, rivers and lakes, in ice caps and glaciers and also in seawater. It is present in deposits of petroleum and natural gas.

Property	Value
<a href="#">dbo:abstract</a>	<ul style="list-style-type: none"> <li>▪ Carbon dioxide (chemical formula CO<sub>2</sub>) is a colorless and odorless gas vital to life on Earth. This naturally occurring chemical compound is composed of a carbon atom covalently double bonded to two oxygen atoms. Carbon dioxide exists in Earth's atmosphere as a trace gas at a concentration of about 0.04 percent (400 ppm) by volume. Natural sources include volcanoes, hot springs and geysers, and it is freed from carbonate rocks by dissolution in water and acids. Because carbon dioxide is soluble in water, it occurs naturally in groundwater, rivers and lakes, in ice caps and glaciers and also in seawater. It is present in deposits of petroleum and natural gas.</li> <li>▪ Kohlenstoffdioxid oder Kohlendioxid ist eine chemische Verbindung aus Kohlenstoff und Sauerstoff mit der Summenformel CO<sub>2</sub>. In Wasser gelöst wird es umgangssprachlich als „Kohlensäure“ bezeichnet. Kohlenstoffdioxid ist ein unbrennbares, saures, farb- und geruchloses Gas.</li> </ul>

# DBpedia Naming Conventions

[https://en.wikipedia.org/wiki/Carbon\\_dioxide](https://en.wikipedia.org/wiki/Carbon_dioxide)



**WIKIPEDIA**  
The Free Encyclopedia

---

[http://dbpedia.org/resource/Carbon\\_dioxide](http://dbpedia.org/resource/Carbon_dioxide)

Entity Identifier



[http://dbpedia.org/page/Carbon\\_dioxide](http://dbpedia.org/page/Carbon_dioxide)

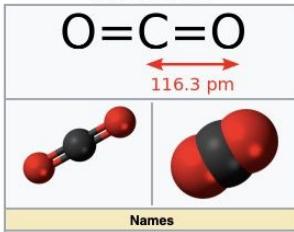
HTML version

[http://dbpedia.org/data/Carbon\\_dioxide](http://dbpedia.org/data/Carbon_dioxide)

RDF/XML version

# Wikipedia Infoboxes

## Carbon dioxide



Other names  
 Carbonic acid gas  
 Carbonic anhydride  
 Carbonic oxide  
 Carbon oxide  
 Carbon(IV) oxide  
 Dry ice (solid phase)

## Identifiers

CAS Number	124-38-9
3D model (JSmol)	<a href="#">Interactive image</a>
3DMol	<a href="#">Interactive image</a>
Belisted	B01131
Reference	1900390
ChEBI	CHEBI:16526
ChEMBL	ChEMBL1231871
ChemSpider	274
ECHA InfoCard	100-004-271
EC Number	204-696-9
E number	E290 (preservatives)
Gmelin Reference	989
KEGG	D00004
MeSH	Carbon-dioxide
PubChem CID	280
RTECS number	FF6400000
UNII	142M471B3J
UN number	1013 (gas), 1845 (solid)
CompTox Dashboard (EPA)	DTXSID4027028
InChI	[show]
SMILES	[show]
Properties	[show]

## Greta Thunberg



Greta Thunberg in April 2019

<b>Born</b>	Greta Ernman Thunberg 3 January 2003 (age 16) Stockholm, Sweden
<b>Occupation</b>	Student and climate activist
<b>Movement</b>	School strike for climate
<b>Parent(s)</b>	Svante Thunberg Malena Ernman
<b>Relatives</b>	Olof Thunberg (grandfather)

## Karlsruhe

Carlsruhe



Karlsruhe Palace, view over Karlsruhe,  
Schlossplatz, Konzerthaus, Crown of Baden



Flag



Coat of arms

## Location of Karlsruhe



## An Inconvenient Truth



Theatrical release poster

<b>Directed by</b>	Davis Guggenheim
<b>Produced by</b>	Laurie David Lawrence Bender Scott Z. Burns
<b>Written by</b>	Al Gore
<b>Starring</b>	Al Gore
<b>Music by</b>	Michael Brook
<b>Cinematography</b>	Bob Richman Davis Guggenheim
<b>Edited by</b>	Jay Cassidy Dan Swietlik
<b>Production company</b>	Lawrence Bender Productions Participant Productions
<b>Distributed by</b>	Paramount Classics
<b>Release date</b>	May 24, 2006
<b>Running time</b>	97 minutes <sup>[1]</sup>

## School strike for climate

FridaysForFuture  
Part of the climate movement



Maximum number of school strikers per country:  
 100+  
 1000+  
 10 000+  
 100 000+

**Date** Since August 2018, mostly on Fridays, sometimes on Thursdays, Saturdays or Sundays

**Location** International

**Caused by** Political inaction against global warming

**Goals** Climate change mitigation

**Methods** Student strike

**Status** Active

## Parties to the civil conflict

Youth

## Lead figures

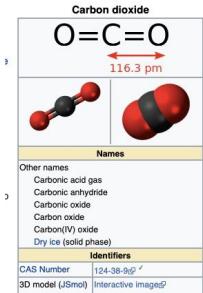
Greta Thunberg

## Number

estimated 1 400 000 (for 15 March 2019)<sup>[1]</sup>

# DBpedia Category System - DBpedia Ontology Classes

dbr:Carbon\_dioxide



rdf:type

dbo:Chemical\_compound

rdfs:subClassOf

dbo:Chemical\_substance

rdfs:subClassOf

owl:Thing

Prefixes:

rdf: <<http://www.w3.org/1999/02/22-rdf-syntax-ns#>>  
 dbr: <<http://dbpedia.org/resource/>>  
 dbo: <<http://dbpedia.org/ontology/>>  
 owl: <<http://www.w3.org/2002/07/owl#>>  
 rdfs <<http://www.w3.org/2000/01/rdf-schema#>>

# DBpedia Category System - Wikipedia Categories

[https://en.wikipedia.org/wiki/Carbon\\_dioxide](https://en.wikipedia.org/wiki/Carbon_dioxide)

Categories: Carbon dioxide | Acid anhydrides | Acidic oxides | Coolants | Fire suppression agents | Greenhouse gases | Household chemicals | Inorganic solvents | Laser gain media | Nuclear reactor coolants | Oxocarbons | Propellants | Refrigerants | Gaseous signaling molecules | Heterocumulenes | E-number additives

dbr:Carbon\_dioxide → dbc:Greenhouse\_gases  
 dct:subject

dbc:Greenhouse\_gases

↓ skos:broader

dbc:Global\_warming

↓ skos:broader

dbc:Climate\_change

↓ skos:broader

dbc:Global\_environmental\_issues

Prefixes:

dct: <http://purl.org/dc/terms/subject>

dbr: <http://dbpedia.org/resource/>

dbc: <http://dbpedia.org/resource/Category:>

skos: <http://www.w3.org/2004/02/skos/core#>

# DBpedia Infobox Extraction - Infobox Properties

An Inconvenient Truth



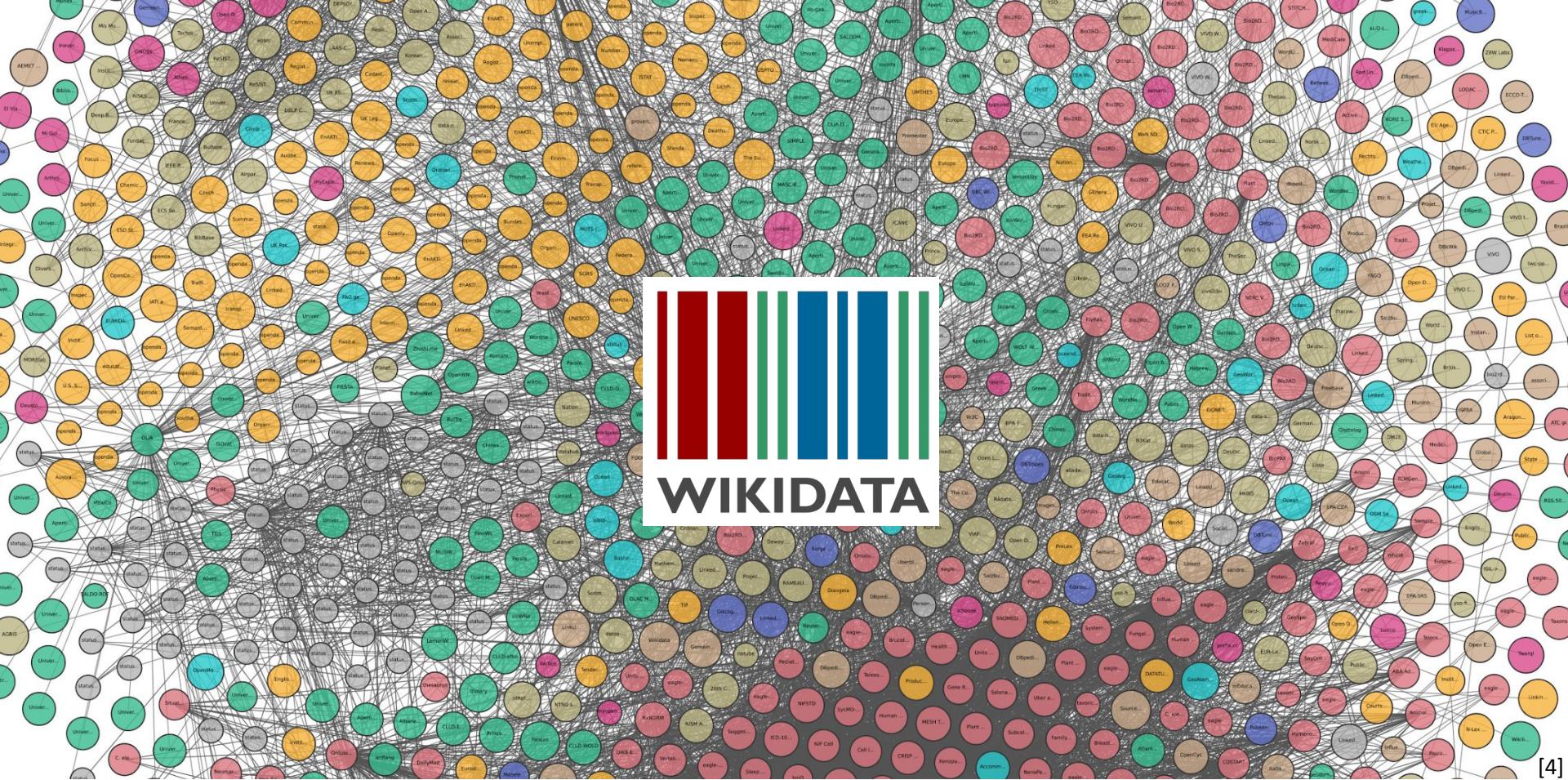
Theatrical release poster

Directed by	Davis Guggenheim	<a href="http://dbpedia.org/resource/Davis_Guggenheim">http://dbpedia.org/resource/Davis_Guggenheim</a>
Produced by	Laurie David Lawrence Bender Scott Z. Burns	
Written by	Al Gore	
Starring	Al Gore	
Music by	Michael Brook	
Cinematography	Bob Richman Davis Guggenheim	

[http://dbpedia.org/resource/An\\_Inconvenient\\_Truth](http://dbpedia.org/resource/An_Inconvenient_Truth)

<http://dbpedia.org/property/director>

[http://dbpedia.org/resource/Davis\\_Guggenheim](http://dbpedia.org/resource/Davis_Guggenheim)



[4]

## Next Lecture:

## Excursion 3: Wikidata Knowledge Graph

### Picture References:

- [1] DBpedia logo, wiki.dbpedia.org, DBpedia Team [Public Domain]  
<https://commons.wikimedia.org/wiki/File:DBpediaLogo.svg>
- [2] The Linked Open Data Cloud, [lod-cloud.net](https://lod-cloud.net/), [CC-BY]  
<https://lod-cloud.net/clouds/lod-cloud.svg>
- [3] Andrea Huang, A preliminary study on Wikipedia Dbpedia and Wikidata, SlideShare [CC\_BY\_SA 4.0]  
<https://www.slideshare.net/andreasinica/a-preliminary-study-on-wikipedia-dbpdeia-and-wikidata>
- [4] Wikidata Logo, Planemad, [Public Domain]  
<https://commons.wikimedia.org/wiki/File:Wikidata-logo-en.svg>