Knowledge Graph

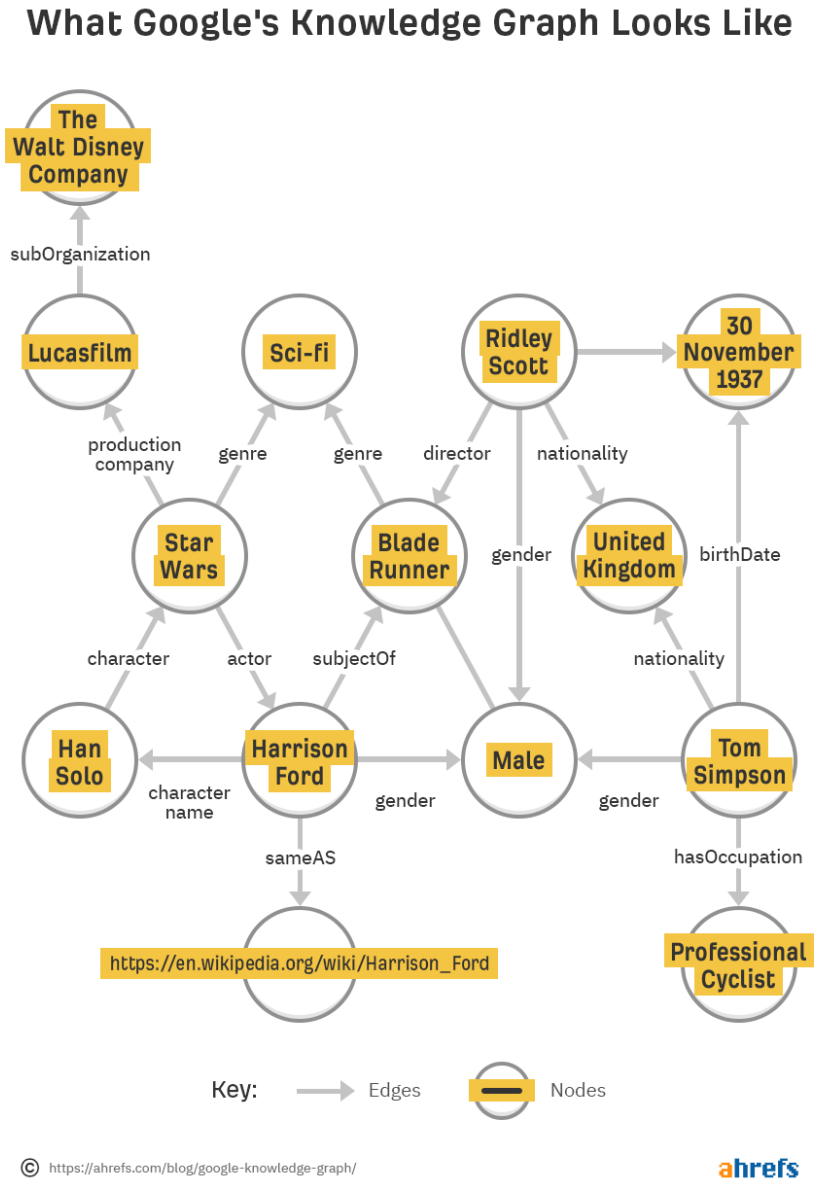
- What Knowledge Graph (KG)?
- Data sources
- Linked Data
- Google Knowledge Graph and the Web of Data

Knowledge Graph

- **knowledge graph represents a collection of interlinked descriptions of entities – objects, events or concepts.**
- Knowledge graphs combine characteristics of several data management paradigms:
  - **Database**, because the data can be explored via structured queries;
  - **Graph**, because they can be analyzed as any other network data structure;
  - **Knowledge base**, because they bear formal semantics, which can be used to interpret the data and infer new facts.

What is Knowledge Graph:

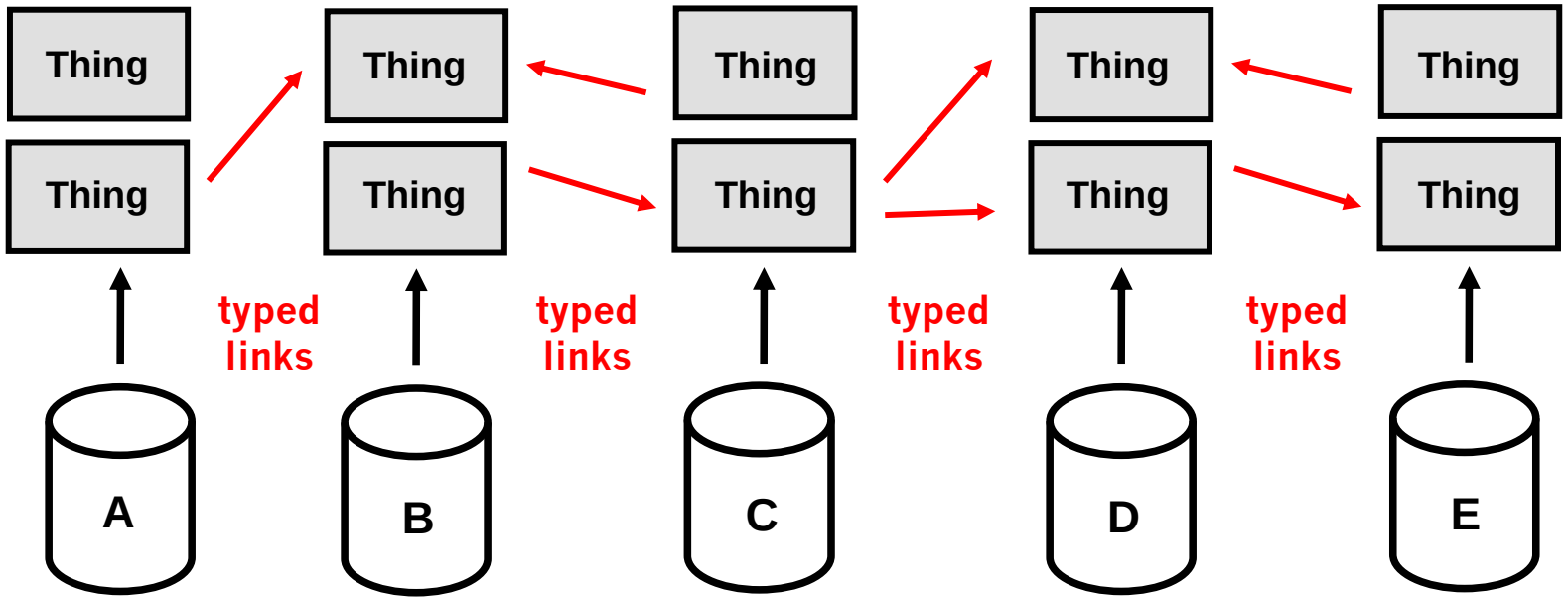
Knowledge Graph is a knowledge base of entities and relationships between them.



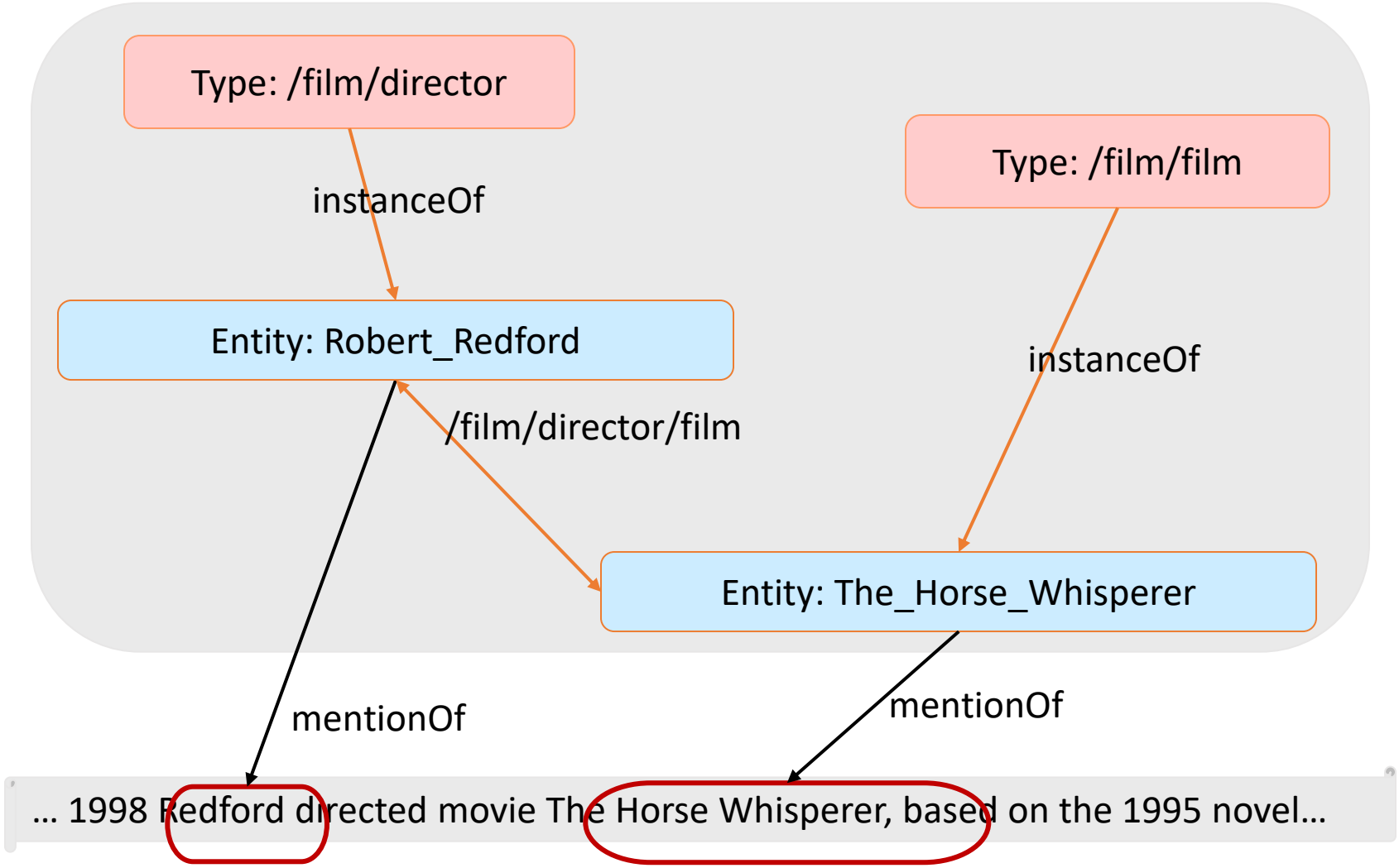
Linked Data

Use Semantic Web technologies to

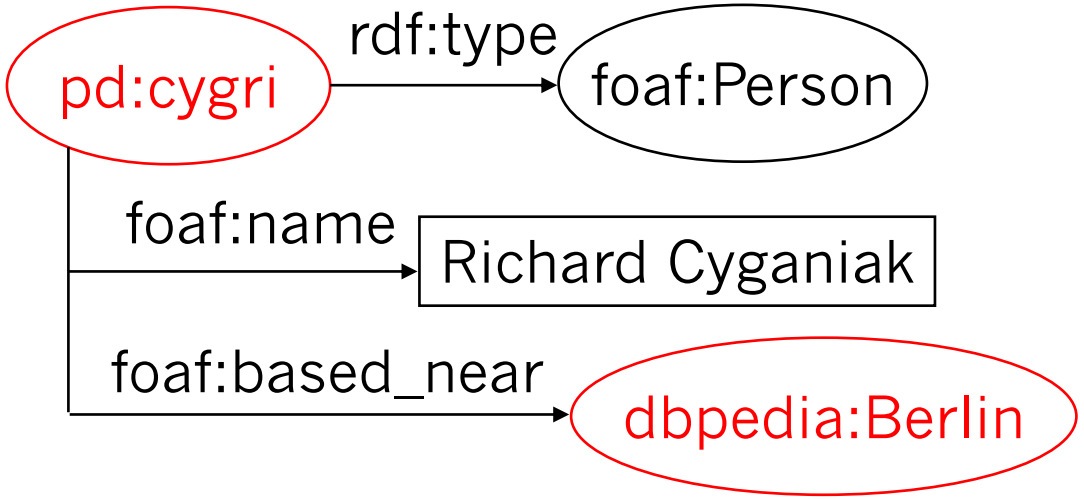
1. **publish structured data on the Web,**
2. **set links between data from one data source to data within other data sources.**



Annotated Web Resources



Data objects are identified with HTTP URIs

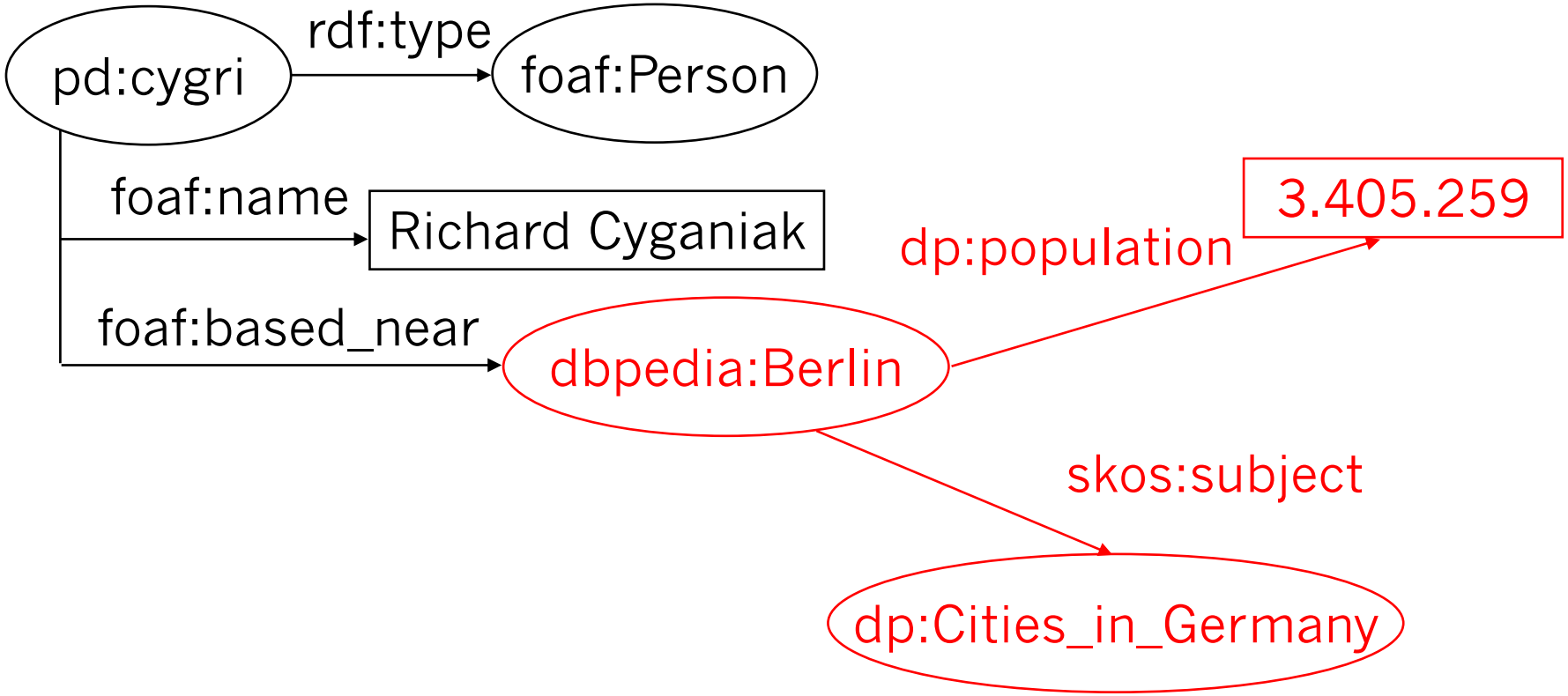


pd:cygri = http://richard.cyganiak.de/foaf.rdf#cygri

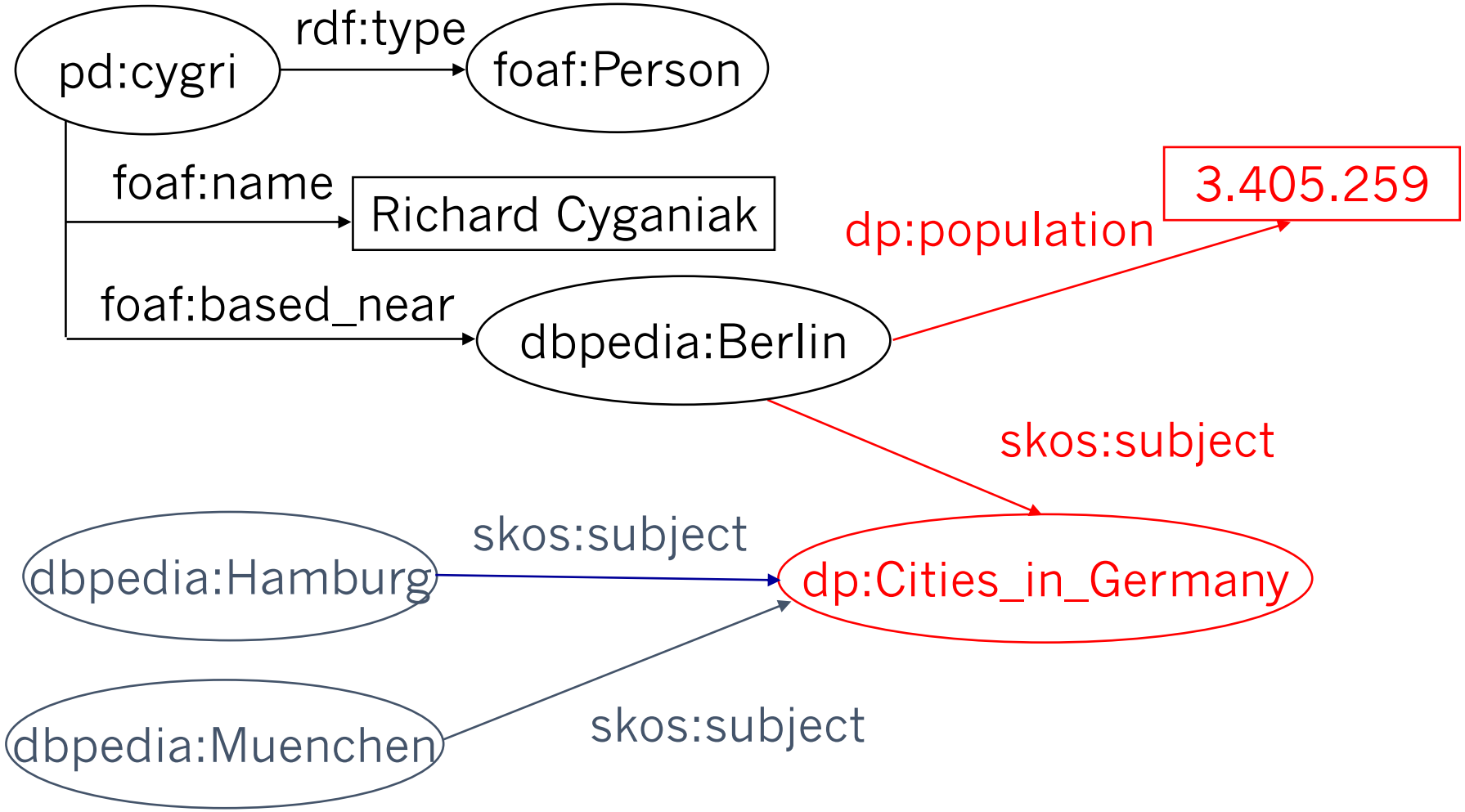
dbpedia:Berlin = http://dbpedia.org/resource/Berlin

Forms an RDF link between two data sources.

Expanding data extraction over the Web

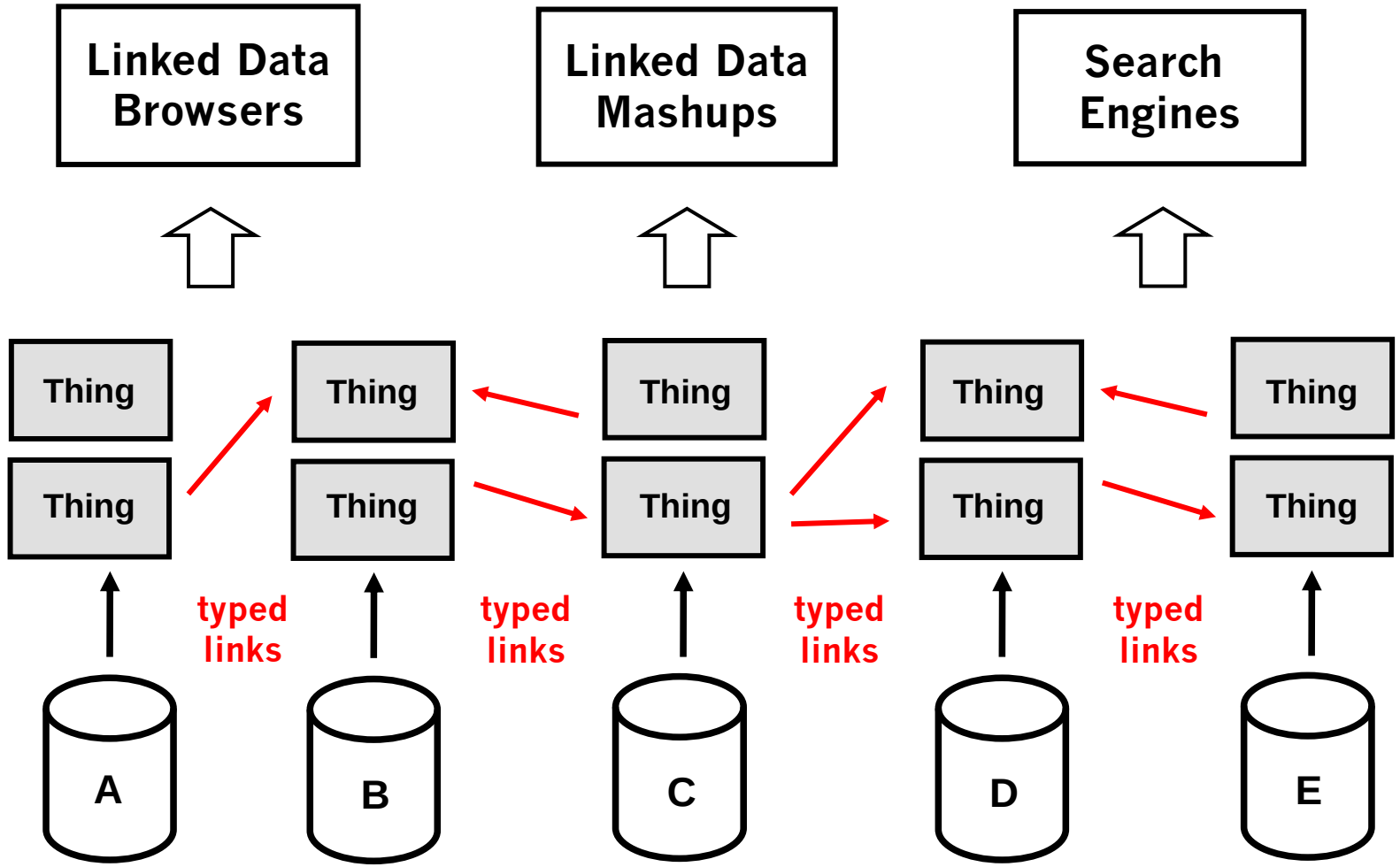


Expanding data extraction over the Web

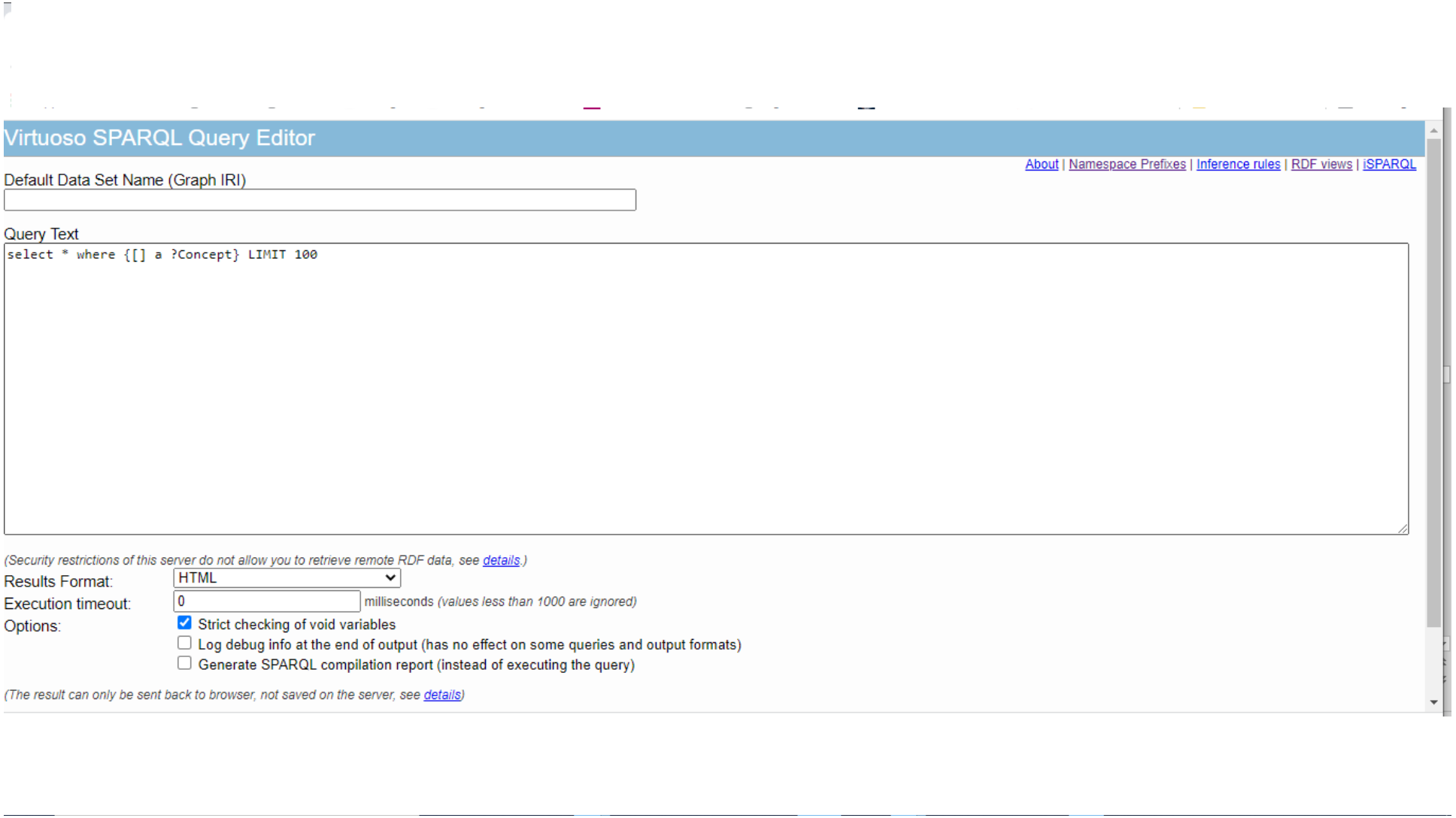


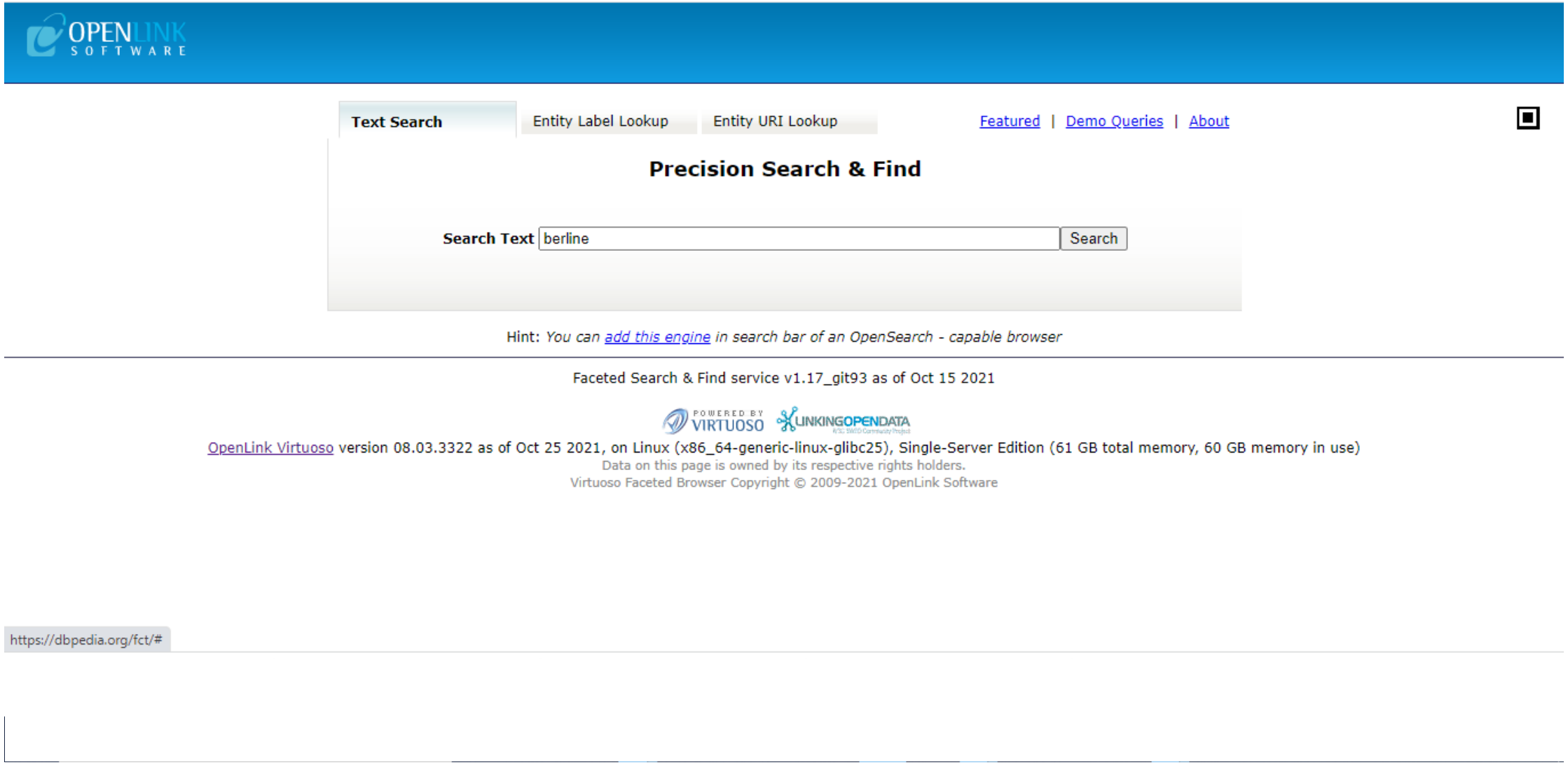
Applications

- **What can I do with this?**



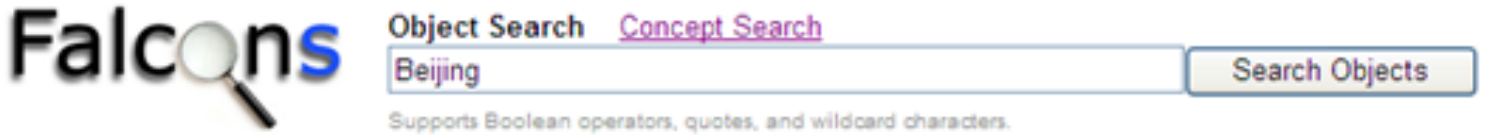
SPARQL Endpoints




OPENLINK SOFTWARE
Text Search
Entity Label Lookup
Entity URI Lookup
Featured | Demo Queries | About
Precision Search & Find
Search Text
berline
Search
Hint: You can add this engine in search bar of an OpenSearch - capable browser
Faceted Search & Find service v1.17_git93 as of Oct 15 2021
POWERED BY VIRTUOSO
LINKINGOPENDATA
OpenLink Virtuoso version 08.03.3322 as of Oct 25 2021, on Linux (x86_64-generic-linux-glibc25), Single-Server Edition (61 GB total memory, 60 GB memory in use)
Data on this page is owned by its respective rights holders.
Virtuoso Faceted Browser Copyright © 2009-2021 OpenLink Software
https://dbpedia.org/fct/#

Property	Value
dbo:wikiPageID	•761572 (xsd:integer)
dbo:wikiPageRevisionID	•918581648 (xsd:integer)
dbp:wikiPageUsesTemplate	•dbt:About •dbt:Commons_category •dbt:Portal •dbt:Cat_main
rdf:type	•skos:Concept
rdfs:label	•Berlin (en)
skos:broader	•dbc:Cities_in_Germany •dbc:Wikipedia_categories_named_after_populated_places_in_Germany •dbc:Capitals_in_Europe •dbc:German_city-states •dbc:German_state_capitals •dbc:States_of_Germany •dbc:Wikipedia_categories_named_after_capitals
skos:prefLabel	•Berlin (en)
prov:wasDerivedFrom	•wikipedia-en:Category:Berlin?oldid=918581648&ns=14
is dbo:wikiPageWikiLink of	•dbr:Berlin •dbr:Permanent_Forest_Contract
is dct:subject of	•dbr:Berlin •dbr:Permanent_Forest_Contract

Falc



All

Artifact Capital City City Document Group
Institution Landmark Location Noun Synset Ontology
Organization Person Publication Subject System

Objects **1** - **10** of **8634** for your search **Beijing** (**1.223** seconds)

Beijing
Types: Capital, City
Labels: 北京" || Pekin || Пекин" || 北京市" || Pequim || Pechino || **Beijing** || Pékin" || Peking || Pekín"
http://dbpedia.org/resource/**Beijing** - Described in 184 documents

Beijing
Types: Subject,
Labels: **Beijing**
http://ontoworld.org/wiki/Special:URIResolver/**Beijing** - Described in 11 documents

Beijing Guoan
Types: Club
Labels: **Beijing** Hyundai || 北京国安" || 北京国安足球俱乐部" || **Beijing** Guoan
http://dbpedia.org/resource/**Beijing**_Guoan - Described in 30 documents

Beijing
The declaration of this URI may be unauthorized.
Types: Capital City
Labels: **Beijing**
http://lonely.org/russia#**Beijing** - Described in 5 documents

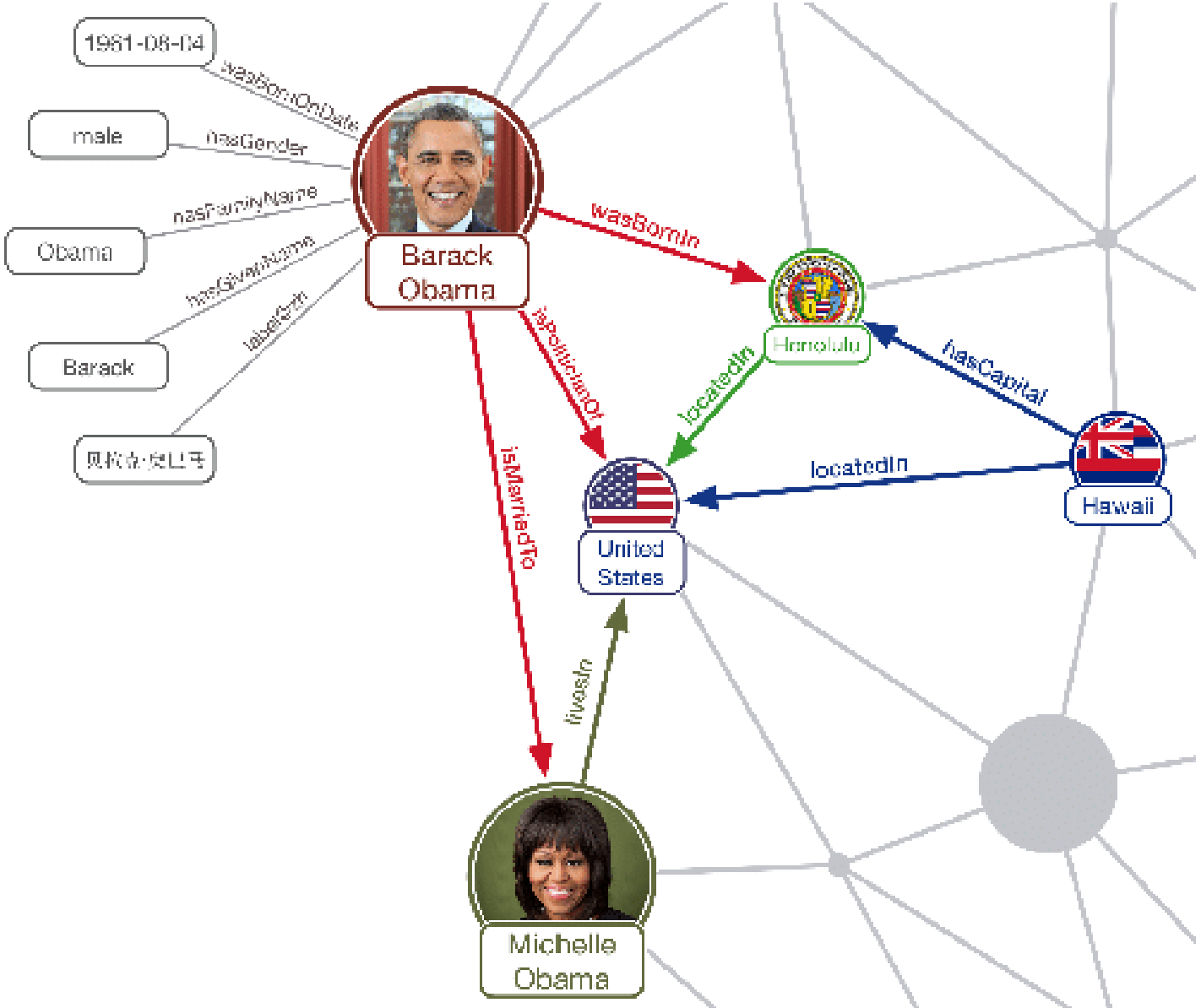
What do we need?

- Heavily used vocabularies
    - Schema.org, Wikidata, DBpedia, YAGO, …

- **Schema.org** provides a collection of shared vocabularies.
- Launched in June 2011 by Bing, Google and Yahoo
- To Create a common set of schemas for webmasters to mark-up with structured data their websites.

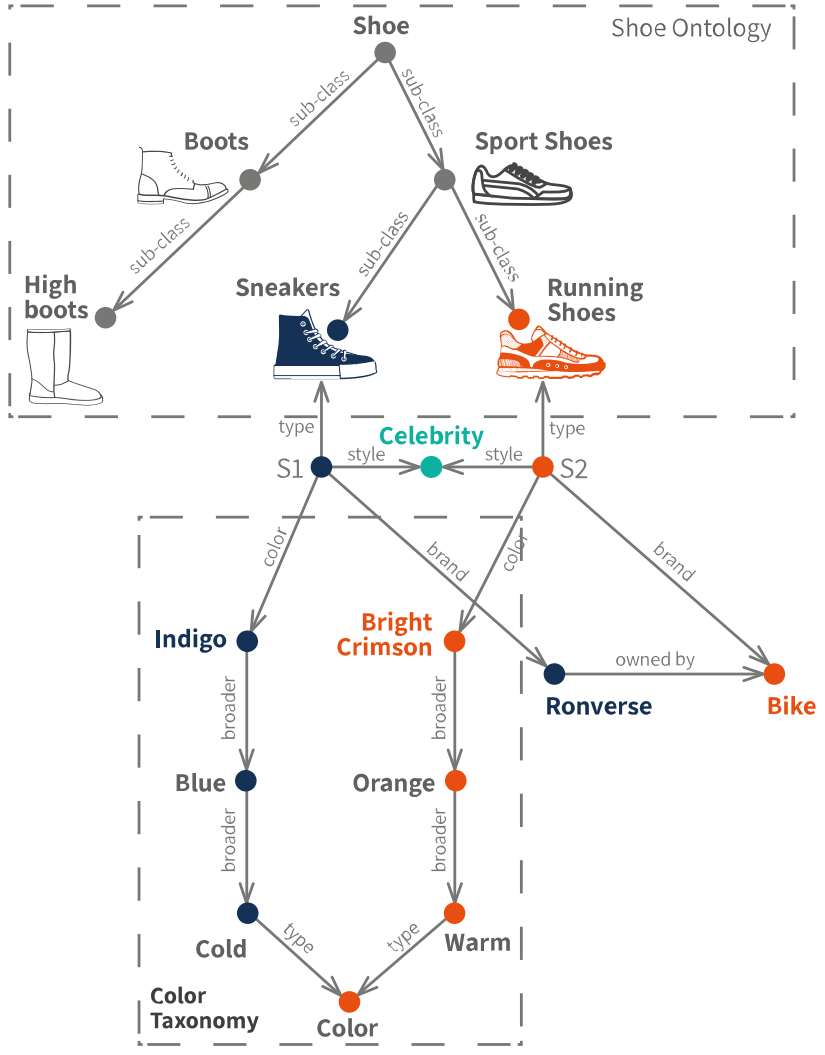
Equality statements (owl:sameAs, skos:exactMatch)

Example of Knowledge Graph

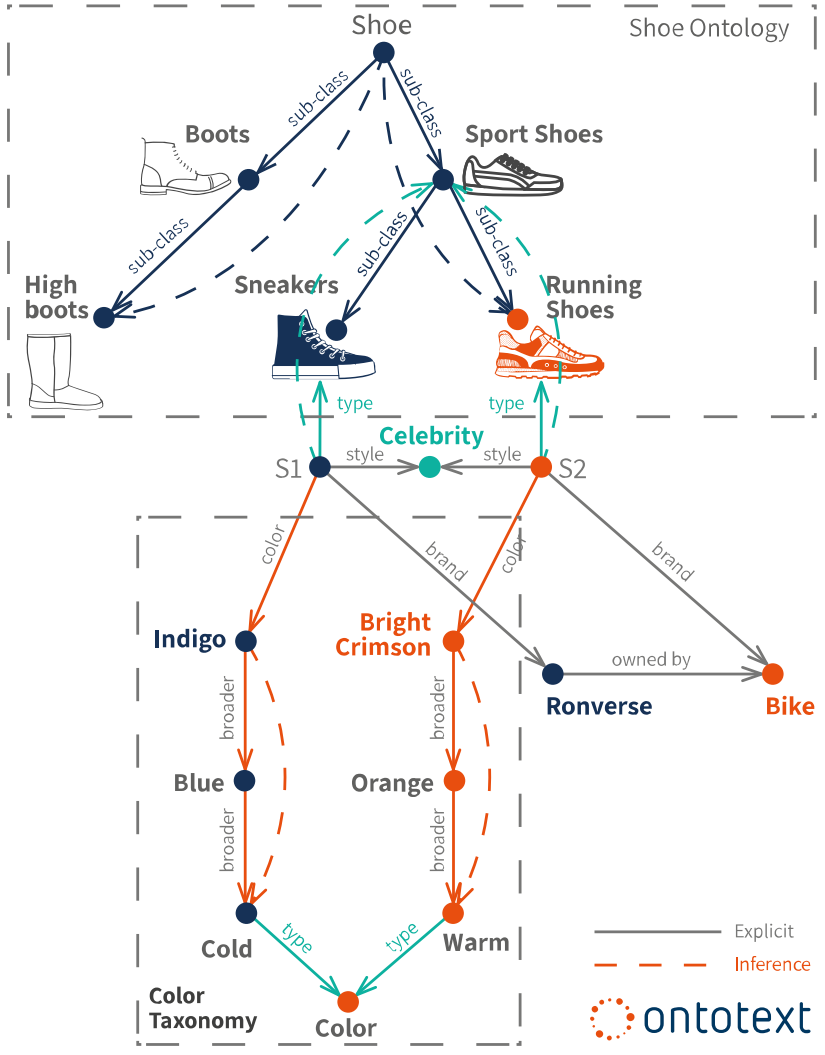


Plain Graph

Knowledge Graph



Knowledge Graph with Inference



Examples of Big Knowledge Graphs

Google Knowledge Graph. Google made this term popular with the announcement of its knowledge graph in 2012. However, there are very few technical details about its organization, coverage and size. There are also very limited means for using this knowledge graph outside Google's own projects.

DBPedia. This project leverages the structure inherent in the infoboxes of Wikipedia to create an enormous dataset of 4.58 things (link https://wiki.dbpedia.org/about) .
This dataset is at the heart of the Open Linked Data movement.

Geonames. Under a creative commons, users of Geonames dataset have access to 25 million geographical entities and features.

Wordnet. One of the most well-known lexical databases for the English language, often used to enhance the performance of NLP and search applications.

DBpedia

- Each thing in the DBpedia data set is identified by a URI reference of the form

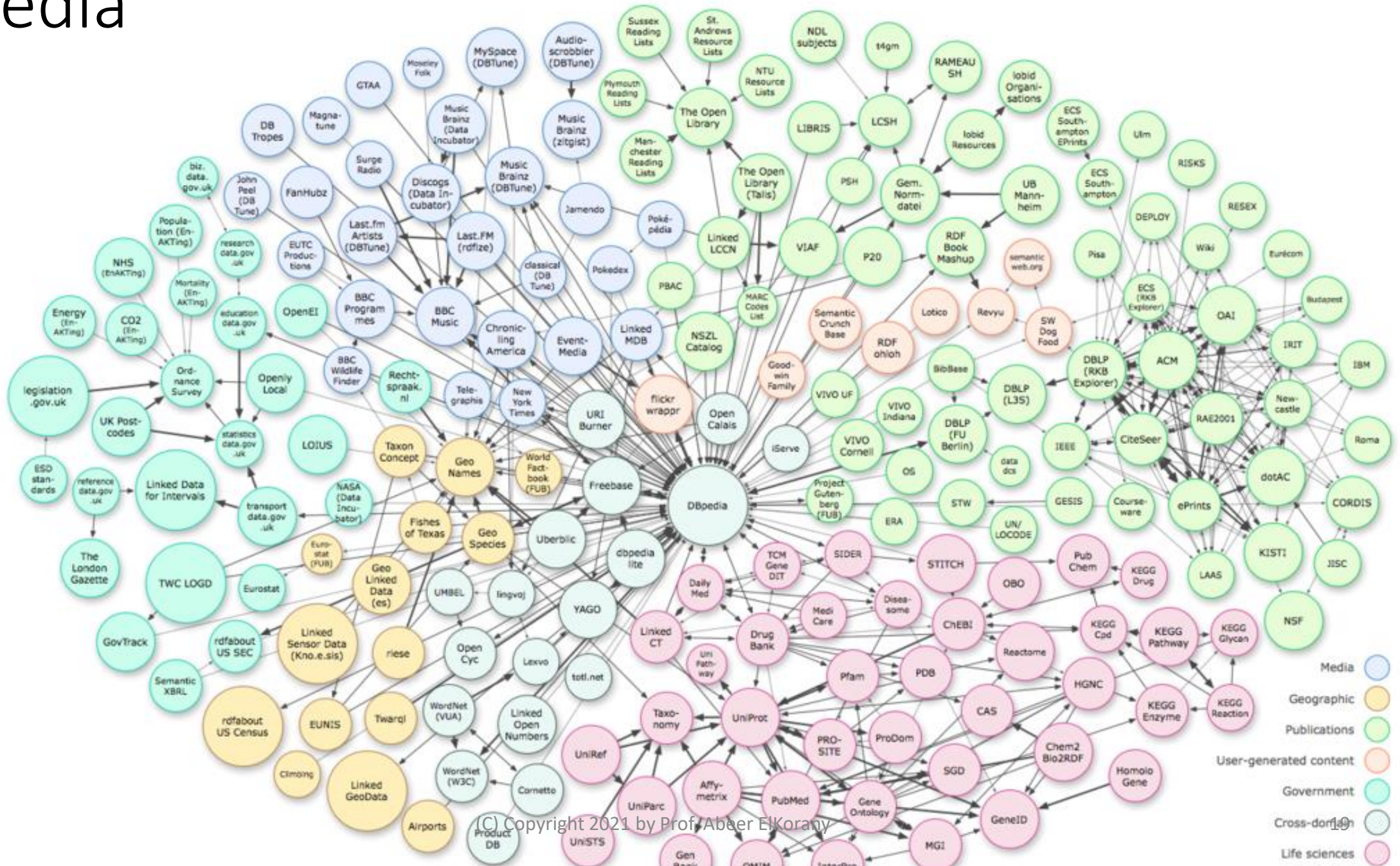
  http://dbpedia.org/resource/Name

- where Name is taken from the URL of the source Wikipedia article, which has the form

  http://en.wikipedia.org/wiki/Name

- Thus, each resource is tied directly to an English-language Wikipedia article.

DBpedia



DBpedia

Class	Examples
City	Cambridge, Berlin, Manchester
Country	Spain, Iceland, South Korea
Politician	George W. Bush, Nicolas Sarkozy, Angela Merkel
Musician	AC/DC, Diana Ross, Röyksopp
Music album	Led Zeppelin III, Like a Virgin, Thriller
Director	Woody Allen, Oliver Stone, Takashi Miike
Film	Pulp Fiction, Blindness, Breakfast at Tiffany's
Book	The Lord of the Rings, The Adventures of Tom Sawyer, The Holy Bible
Computer Game	Tetris, World of Warcraft, Sam & Max hit the Road
Technical Standard	HTML, RDF, URI

Case Study : Google Knowledge Graph (GKG)

- *"A huge knowledge graph of interconnected entities and their attributes"*.

  Amit Singhal, Senior Vice President at Google

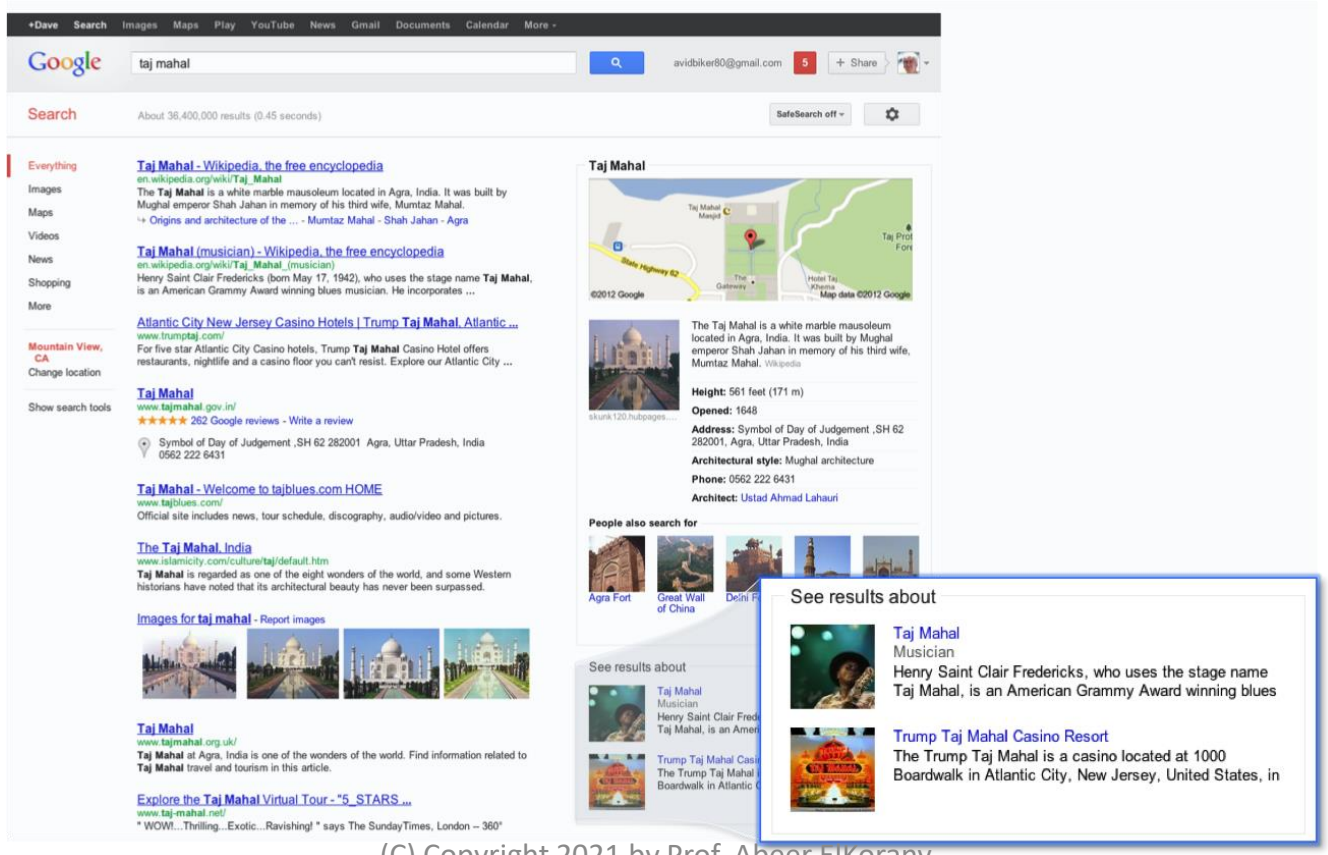
- *"A knowledge based used by Google to enhance its search engine's results with semantic-search information gathered from a wide variety of sources"*

  http://en.wikipedia.org/wiki/Knowledge_Graph

What is it?

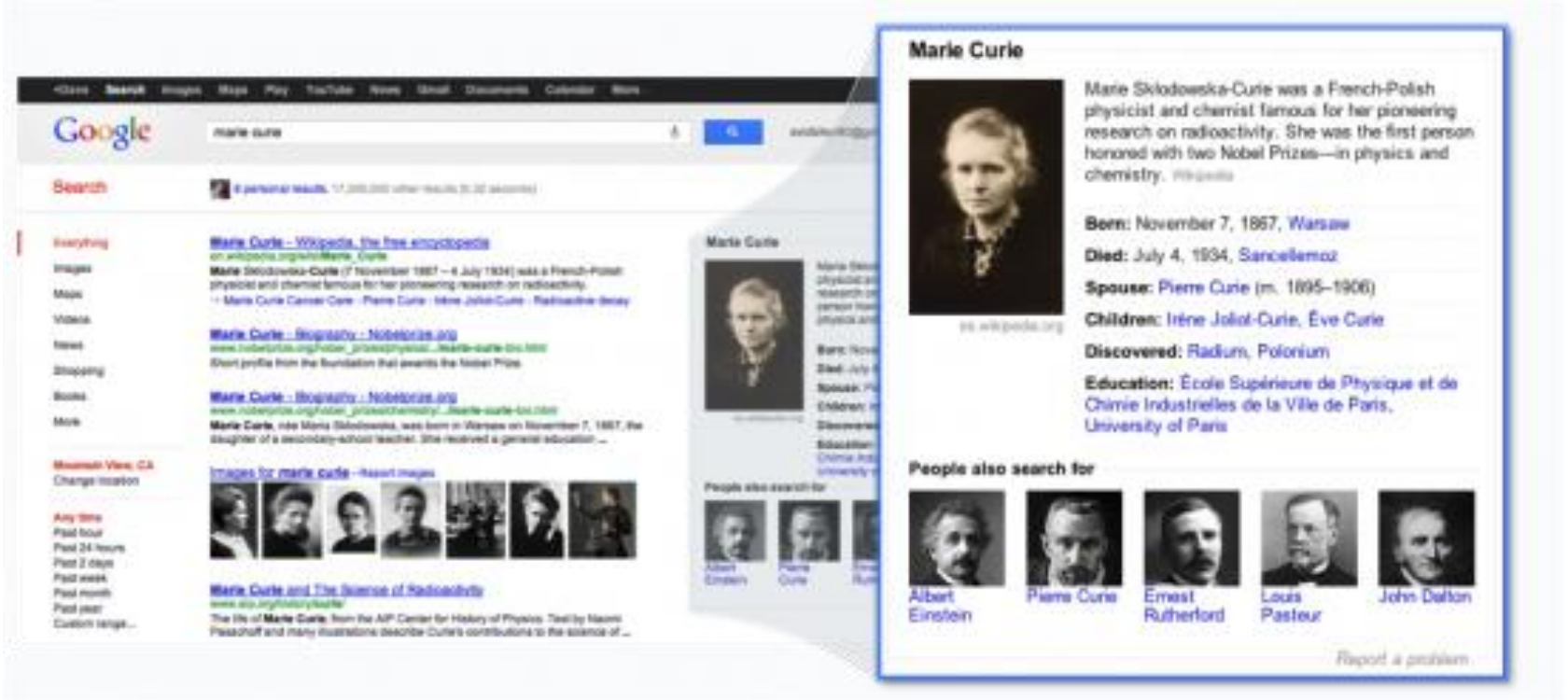
GKG enhances Google Search in **three main ways**:

- ***Find the right thing***
  - deals with the ambiguity of the language



What is it?

- ***Summaries***
  - summarize relevant content around that topic, including key facts about the entity



What is it?

- ***Deeper and broader information***
  - reveal new facts
  - anticipate what the next questions and provide the information beforehand (based on what other users asked before)