

Semantic Search (Part I)

[DAT640] Information Retrieval and Text Mining

Krisztian Balog
University of Stavanger

October 14, 2019

Discussion

Question

What do you think *semantic search* means?

Examples of semantic search

Query: buy books

Google search results for "buy books" in 2005:

- Web** Results 1 - 10 of about 259,000,000 for **buy books** (0.15 seconds)
- Compare book prices** Sponsored Link
www.lookbookstores.com Compare prices at dozens of stores Save up to 90% on books & textbooks
- Buy books** Sponsored Links
College Planning, SAT Preparation Financial Aid - Books & More www.CollegeBoard.com
- Amazon.com: Online Shopping for Electronics, Apparel, Computers ...** Online shopping from the earth's biggest selection of books, magazines, music, DVDs, ...
Redeem or buy a gift certificate. • Visit our Help department. ...
www.amazon.com/exec/obidos/submit/home/html - 56k - 30 Sep 2005 - Cached - Similar pages
- Amazon.com Books** New & used textbooks, biographies, children's ...
Online shopping for millions of new & used books on thousands of topics at everyday low ... Redeem or buy a gift certificate. • Visit our Help department. ...
www.amazon.com/exec/obidos/tg/browse/-/283155 - 56k - 30 Sep 2005 - Cached - Similar pages
- Barnes & Noble.com - Home** Browse All Books By Subject ... Special Offer. Buy 2, Get the 3rd Free. Buy two paperbacks and get a third for free! Browse our special collection of Modern ...
www.barnesandnoble.com/ - 54k - 30 Sep 2005 - Cached - Similar pages
- Powells Books - Used, New, and Out of Print** New, used, and out of print books; many categories with detailed descriptions and illustration of selected books.

Google result page in 2005

Google search results for "buy books" today:

- All** Images Maps Shopping Books More Settings Tools
- About 5,160,000,000 results (0.64 seconds)
- Adlibris online bookstore | Norges største bokutvalg | adlibris.com**
www.adlibris.com *
More than 10 million titles - Quick deliveries and low prices. Over 10 million titles. Trygg e-handel.
Nyheter
De seneste bøkene tilverkene - Mata røster og rank levering
- Aktuelt i media**
Balanse det prates om nå
Se hvilke bøker som synes i media
- Book Depository | Free Delivery Worldwide | bookdepository.com**
www.bookdepository.com *
Discover 18 Million Books with Free Delivery on All Orders. Over 18 Million Titles. Leading Online Bookstore. Everyday Low Prices. Free Delivery To Norway. Types: Romance, Crime & Thriller, Children's Books, Food & Drink, Travel & Holiday, Fantasy, Fantasy Books - Children's Books - Travel & Holiday Guides - The Bargain Shop - Coming Soon
- Hours** Sort by +
- A Outland**
Book Store - Tredissemestrene
4.5 km - Stavanger 51 86 17 11
Open - Closes 2PM
WEBSITE **DIRECTIONS**
- B Notabene Tredissemestrand**
Book Store - Tredissemestrene
4.7 km - Stavanger 52 90 05 80
Open - Closes 8PM
WEBSITE **DIRECTIONS**
- C Notabene Akadem Torglamassen**
Book Store - Akadem Torglamassen
4.4 km - Stavanger 51 89 09 88
Open - Closes 8PM
WEBSITE **DIRECTIONS**

Google result page today

Examples of semantic search

Query: *flight berlin chicago*

Google search results for 'flight berlin chicago':

- Best Fares to Berlin, Germany**
Best Fares to Berlin, Germany. Sign Up For Best Fare Alerts Track Best Fares to ...
[Chicago, IL - O'Hare \(ORD\), \\$356.00*](#) **Chicago, IL - all (CHI), \$356.00*** ...
[search.travel.yahoo.com/bin/search?&search?int=us&dc=BER&xt_main=Berlin&source=37k](#) - Cached - Similar pages
- Cheap Flights from Chicago to Berlin, Chicago Flights | Travelzoo ...**
Find Chicago Flights through Travelzoo SuperSearch. SuperSearch recommends the best airline for your itinerary.
[flights.travelzoo.com/Air/Germany/Berlin_BER/USA_IL/Chicago_CHI/ - 10k](#) - Cached - Similar pages
- Cheap airline tickets to Berlin, Germany departing from Chicago ...**
Cheap airline tickets to Berlin, Germany from Chicago, Illinois. Compare cheap airline tickets for flights to Berlin, Germany departing from Chicago, ...
[www.cheapflights.com/flights/Berlin/Chicago/ - 57k](#) - Cached - Similar pages
- Cheap airline tickets to Berlin, Germany**
Cheap airline tickets to Berlin, Germany at Cheapflights.com. Compare flights to Berlin, Germany from the USA.
[www.cheapflights.com/flights/Berlin/ - 35k](#) - 30 Sep 2005 - Cached - Similar pages
- Gridskipper, the Urban Travel Guide**
Scouring the world for discount flights, chic hotels - and pretty people. Gridskipper, the detailed travel guide.

Google result page in 2005

Google search results for 'flight berlin chicago':

- Berlin-Chicago from €189 - eDreams.com**
[www.edreams.com/Berlin_Chicago](#) - Rating for edreams.com 3.6 - 2,649 reviews
Compare & Save on Your Next Flight. Limited Offers. Book Your Flight Best Price Guarantee. Customer Service 10/7. Cheap Flights from €189. More than 750 Airlines. Flight + Hotel Offers.
- Book Now & Get Best Deals**
Top Airline Offers from eDreams. We Find the Best Deals for You!
- Best Offers to/from Paris**
Compare & Save on Your Next Flight. Book Your Flight to Paris and Save
- Flights from Berlin, Germany (all airports) to Chicago, USA (all airports)**
[www.google.no/flight](#)

From	To	Date	Airline	Flight Time	Price
Berlin, Germany (all airports)	Chicago, USA (all airports)	Mon, October 8	WOW	17h 15min	from NOK 3,531
			Aer Lingus	14h 35min	from NOK 5,030
			LOT	14h 30min	from NOK 5,407
			Turkish Airlines	17h 30min	from NOK 5,969
			Other airlines	11h 55min	from NOK 6,617

→ Search flights

Cheap Flights from Berlin to Chicago from \$357 | (BER - CHI) - KAYAK
Compare Prices & Book Flights to Chicago from Germany
Fly from Berlin to Chicago on WOW air from \$357, Icelandair from \$401, KLM from \$517, Delta from \$720. Search and find deals on flights to Chicago.
Flight Price: \$345 Airlines: WOW air, Icelandair, KLM, Delta, Aer ...

Google result page today

Semantic search

A broad view on semantic search:

Definition

Semantic search encompasses a variety of methods and approaches aimed at aiding users in their information access and consumption activities, by understanding their context and intent.

- “Search with meaning”
 - Beyond literal matches
 - Understanding what the query actually means
 - Searching for *things* instead of *strings*
- Our notion of semantics: references to meaningful, i.e., machine understandable structures

Entity-oriented search

Entity-oriented search to refer to a broad range of information access tasks where entities are used as information objects, instead of or in addition to documents

Definition

Entity-oriented search is the search paradigm of organizing and accessing information centered around entities, and their attributes and relationships.

- Note: entity-oriented search is a subset of semantic search

Examples of entity-oriented search

Google Search

Web Images Maps News Videos More Search tools

About 78,200,000 results (0.59 seconds)

Lisbon - Official Website - visitlisboa.com
www.visitlisboa.com/TouristInfo ▾
Discover Golf, Music, Surf And Food Discover Everything About Lisbon!

Lisbon - Wikipedia, the free encyclopedia
<https://en.wikipedia.org/wiki/Lisbon> ▾ Wikipedia ▾
Lisbon (/t̪ɪzben/; Portuguese: Lisboa, IPA: [l̪iʒˈboɐ]) is the capital and the largest city of Portugal, with a population of 552,700 within its administrative limits ...
Belém Tower - Tagus - Belém - Alcântara

Images for lisbon Report images



[More images for lisbon](#)

Lisbon, Portugal - Lonely Planet
www.lonelyplanet.com/portugal/lisbon ▾ Lonely Planet ▾
Spread across steep hillsides that overlook the Rio Tejo, Lisbon offers all the delights you'd expect of Portugal's star attraction, yet with half the fuss of other ...
Top things to do in Lisbon - Best places to stay in Lisbon - Portugal Image gallery

Lisbon Tourism: Best of Lisbon, Portugal - TripAdvisor
www.tripadvisor.com ▾ ... ▾ Central Portugal ▾ Lisbon District ▾ TripAdvisor ▾
Lisbon Tourism: TripAdvisor has 486539 reviews of Lisbon Hotels, Attractions, and Restaurants making it your best Lisbon resource.



Lisbon

Capital of Portugal

Lisbon, Portugal's hilly capital, is a coastal city known for its cafe culture and soulful Fado music. From imposing São Jorge Castle, the view encompasses the old city's pastel-colored buildings, Tagus Estuary and the Ponte 25 de Abril suspension bridge. Nearby, the National Azulejo Museum displays 5 centuries of decorative ceramic tiles. And just outside Lisbon is a string of Atlantic beaches, from Cascais to Estoril.

Area: 32.74 mi²
Weather: 61°F (16°C), Wind NE at 7 mph (11 km/h), 82% Humidity
Local time: Tuesday 10:41 AM
Hotels: 3-star averaging \$50, 5-star averaging \$140. [View hotels](#)
Population: 530,847 (2012) [UNdata](#)

Points of interest View 15+ more



- Belém Tower
- Jerónimos Monastery
- São Jorge Castle
- Lisbon Oceanarium
- Praça do Comércio

Examples of entity-oriented search

Google

Web Maps Images News Shopping More Search tools

Lisbon / Points of interest

Belém Tower Jerónimos Monastery São Jorge Castle Lisbon Oceanarium Praça do Comércio Calouste Gulbenkian Foundation Museu Calouste Gulbenkian Lisbon Zoo Santa Justa L.

Things To Do In Lisbon - visitlisboa.com
www.visitlisboa.com/ThingsToDo ▾
Find Out What To Visit And Where To Sleep. Everything About Lisbon In Lisbon: Associação de Turismo de Lisboa

Lisbon Things To See - City Tours, Day Trips, and more
www.viator.com/lisbon ▾
4.4 ★★★★★ rating for viator.com
Book Lisbon Things to Do on Viator.
24/7 Live Support · Verified Reviews & Photos · Low Price Guarantee
Viator.com has 1,456,772 followers on Google+
Sintra Tours - Lisbon Tours - Lisbon Fado Show - Porto Wine Tours

Best Things to Do in Lisbon - holidaylettings.co.uk
www.holidaylettings.co.uk/rentals ▾
Find, Book & Enjoy The Best Rentals Book Securely or Enquire Online Now
Late Deals - Holiday Destinations - Payment Protection - Holiday Ideas

The Top 10 Things to Do in Lisbon - TripAdvisor - Lisbon ...
www.tripadvisor.com/Attractions-g189158-Activities-Lisbon... ▾ TripAdvisor ▾
Hotels near Oceanário de Lisboa. Hotels near Torre de Belém. Hotels near Castelo de São Jorge. Hotels near Mosteiro dos Jerónimos. Hotels near Alfama. Hotels near Tram 28. Hotels near Bairro Alto. Hotels near Museu da Fundação Calouste Gulbenkian. Oceanário de Lisboa - Miradouro da Senhora do Monte - Mosteiro dos Jerónimos

Best Things to Do in Lisbon | U.S.News Travel
www.usnews.com/travel-best-things/lisbon-best-things

Lisbon
Capital of Portugal

Lisbon, Portugal's hilly capital, is a coastal city known for its cafe culture and soulful Fado music. From imposing São Jorge Castle, the view encompasses the old city's pastel-colored buildings, Tagus Estuary and the Ponte 25 de Abril suspension bridge. Nearby, the National Azulejo Museum displays 5 centuries of decorative ceramic tiles. And just outside Lisbon is a string of Atlantic beaches, from Cascais to Estoril.

Area: 32.74 mi²
Weather: 61°F (16°C), Wind NE at 7 mph (11 km/h), 82% Humidity
Hotels: 3-star averaging \$50, 5-star averaging \$140. [View hotels](#)
Local time: Tuesday 10:42 AM
Population: 530,847 (2012) [Unidata](#)
Colleges and Universities: [University of Lisbon](#), More

Examples of entity-oriented search

The screenshot shows a Google search results page for the query "kl1799". The search bar at the top contains "kl1799". Below the search bar, the "All" tab is selected, along with other options like Maps, Videos, Shopping, News, More, and Search tools. A search result for "KLM Flight 1799" is displayed, showing flight details from AMS to MUC. The flight is listed as "On-time - departs in 10 hours 50 mins". The departure information is: Departs Amsterdam, Tuesday, May 10, Time: 6:40 PM, Terminal: 1, Gate: C7. The arrival information is: Arrives Munich, Tuesday, May 10, Time: 8:05 PM, Terminal: 1, Gate: -. Below this, there is a link to "KLM flight KL1799 - Flightradar24" with the URL <https://www.flightradar24.com/data/flights/kl1799/>. The result also includes a snippet from FlightAware and FlightStats, and a link to "KL1799 / KLM1799 — KLM Royal Dutch Airlines — Plane Fin..." with the URL <https://planefinder.net/data/flight/KL1799>.

KLM Flight 1799
On-time - departs in 10 hours 50 mins

AMS ✈ MUC

Departs Amsterdam, Tuesday, May 10	Arrives Munich, Tuesday, May 10
Time 6:40 PM	Time 8:05 PM
Terminal 1	Terminal 1
Gate C7	Gate -

Showing airport times Feedback

KLM flight KL1799 - Flightradar24
<https://www.flightradar24.com/data/flights/kl1799/> ▾ Flightradar24 ▾
KL1799 (KLM) - Live flight status, scheduled flights, flight arrival and departure times, flight tracks and playback, flight route and airport.

KLM (KL) #1799 FlightAware
<https://flightaware.com/live/flight/KLM1799> ▾ FlightAware ▾
KLM (KL) #1799 Flight Tracker (KLM1799) Flight Tracker (en route flights, arrivals, departures, history) with live maps and aircraft photos.

KL 1799 Flight Status - FlightStats
www.flighthstats.com/FlightStatus/FlightStatusByFlight.do?... ▾ FlightStats ▾
Check the current status of flight (KL) KLM 1799 complete with live maps, weather and more.

KL1799 / KLM1799 — KLM Royal Dutch Airlines — Plane Fin...
<https://planefinder.net/data/flight/KL1799>
Plane Finder Data has the latest real-time information on flight KL1799 / KLM1799 (KLM Royal Dutch Airlines)

KL1799 schedule. (KLM flight: Amsterdam -> Munich)
info.flightradar24.net/flight/KLM_KL_1799 ▾
KL 1799 Non-stop Embraer EMB 170 / EMB 190 (EMJ) 1:25 Effective from 2016-10-30 ...
KL 1799 Non-stop Fokker 70 (F70) 1:25 Effective 2016-10-31 through ...

Why?

- From a **user perspective**
 - Entities are natural units for organizing information
 - We care about and mostly think in terms of real-world things and their connections
- From a **machine perspective**
 - Entities allow for a better understanding of search queries, of document content, and even of users (e.g., their context and preferences)
 - Entities enable search engines to be more intelligent

What is an entity?



people



locations



organizations



products

What is an entity?

Commonly accepted definition:

Definition

An *entity* is an object or concept in the real world that can be distinctly identified.

- Issues

- What does the “real world” mean? (Is “Superman” an entity or not?)
- Answering this will likely lead to a long philosophical discussion about “existence”

What is an entity?

Pragmatic, data-oriented definition:

Definition

An *entity* is a uniquely identifiable object or thing, characterized by its name(s), type(s), attributes, and relationships to other entities.

Our universe is restricted to some particular registry of entities:

Definition

An *entity catalog* is a collection of entries, where each entry is identified by a unique ID and contains the name(s) of the corresponding entity.

Named entities vs. concepts

- Two main classes of entities may be distinguished
 - *Named entities* are real-world objects that can be denoted by a proper noun
 - For example, specific persons, locations, organizations, products, events, etc.
 - *Concepts* are abstract objects, including, but not limited to
 - Mathematical and philosophical concepts (e.g., "distance," "axiom," "quantity")
 - Physical concepts and natural phenomena (e.g., "gravity," "force," "wind")
 - Psychological concepts (e.g., "emotion," "thought," "identity"), and social concepts (e.g., "authority," "human rights," "peace")
- This distinction is mostly of a philosophical nature. From a technical perspective, the exact same methods may be used for names entities and concepts.

Properties of entities

- **Unique identifier**
 - There must be a one-to-one correspondence between each entity identifier (ID) and the (real-world or fictional) object it represents
 - For example, social security number, product EAN, MAC address, etc.
- **Name(s)**
 - Names do not uniquely identify entities; multiple entities may share the same name
 - The same entity may be known by more than a single name (e.g., “Barack Obama,” “President Obama”)
 - Alternative names are called *surface forms* or *aliases*
- **Type(s)**
 - Entities may be categorized into multiple *entity types*
 - Types can also be thought of as containers (semantic categories) that group together entities with similar properties
 - Analogy to object-oriented programming: an entity of a type is like an instance of a class
 - Entity types are often organized in a hierarchical structure (*type taxonomy*)

Properties of entities (2)

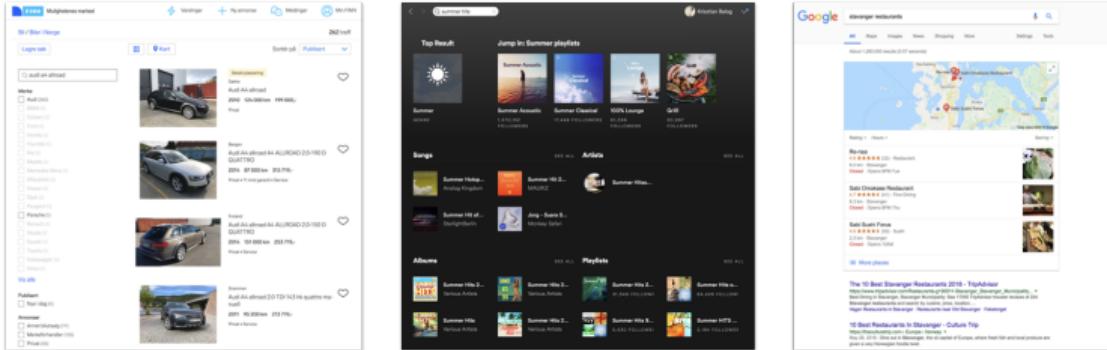
- **Attributes**
 - Entities are characterized by attributes
 - Different types of entities typically have different sets of attributes
 - People: date and place of birth, weight, height, parents, spouses, etc.
 - Places: latitude, longitude, population, postal code(s), country, continent, etc.
 - Attributes always have *literal values*
- **Relationships**
 - May be seen as “typed links” between entities (or attributes where the value is another entity)
 - For example, parents of a person, capital of a country, manufacturer of a product, etc.

Entity-oriented search tasks

- Entities as the unit of retrieval
 - Entity retrieval
- Entities for knowledge representation
 - Entity linking
 - Knowledge base population
- Entities for an enhanced user experience
 - Query assistance
 - Recommendations

Entity retrieval

- Task: given a search query, return a ranked list of entities (instead of documents)



Entity linking

- Task: recognize mentions of entities in text and assign to these unique identifiers from a knowledge repository

The screenshot shows the DBpedia Spotlight web application. At the top, there's a logo featuring a stylized book with orange dots above it, followed by the text "DBpedia Spotlight". Below the logo is a search bar containing the text: "First documented in the 13th century, Berlin was the capital of the Kingdom of Prussia (1701-1918), the German Empire (1871-1918), the Weimar Republic (1919-33) and the Third Reich (1933-45). Berlin in the 1920s was the third largest municipality in the world. After World War II, the city became divided into East Berlin -- the capital of East Germany -- and West Berlin, a West German exclave surrounded by the Berlin Wall from 1961-89. Following German reunification in 1990, the city regained its status as the capital of Germany, hosting 147 foreign embassies." The interface includes several controls: a "Confidence" slider set to 0.5, a "Language" dropdown set to English, a checkbox for "n-best candidates" which is unchecked, and two buttons labeled "SELECT TYPES..." and "ANNOTATE". In the bottom right corner of the main text area, there's a "BACK TO TEXT" button.

Architecture of an entity-oriented search system

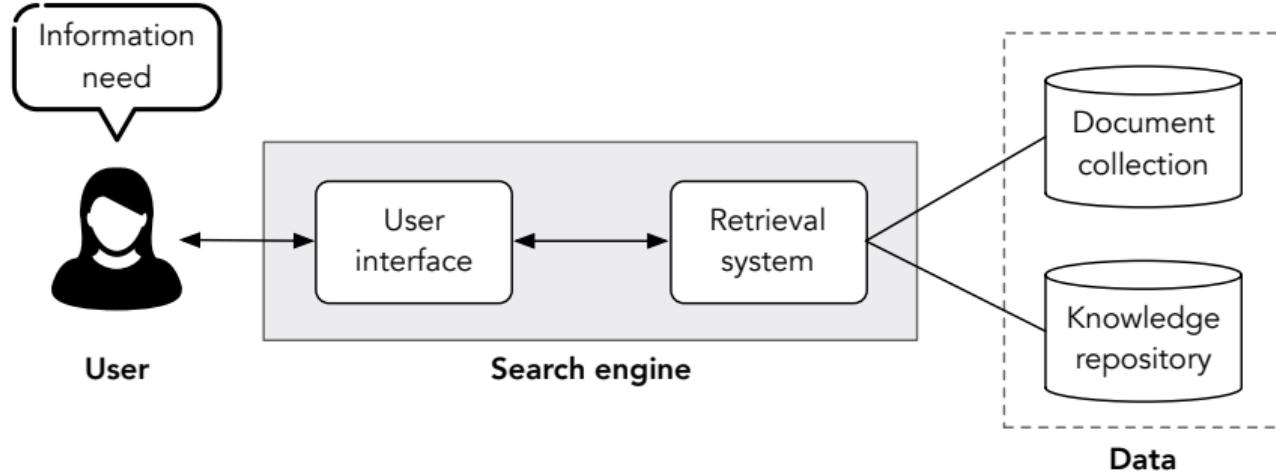


Figure: Illustration is taken from (Balog, 2018) [Fig. 1.3]

Representing properties of entities

Information about entities can be represented and stored in semi-structured or in structured form

Definition

A *knowledge repository* (KR) is a catalog of entities that contains entity type information, and (optionally) descriptions or properties of entities, in a semi-structured or structured format.

- Classic example: Wikipedia
 - Each article in Wikipedia is an entry that describes a particular entity
 - Articles are also assigned to categories (which can be seen as entity types)
 - Wikipedia articles also contain information about attributes and relationships of entities, but not in a structured form

Representing properties of entities

To organize and store information about entities in a structured form, entities may be represented as a set of statements (facts or assertions)

Definition

A *knowledge base* (KB) is a structured knowledge repository that contains a set of facts (assertions) about entities.

- Note: all knowledge bases are also knowledge repositories, but the reverse is not true
- Conceptually, entities in a knowledge base may be seen as nodes of a graph, with the relationships between them as (labeled) edges
 - When this graph nature is emphasized, a knowledge base may also be referred to as a *knowledge graph* (KG)

Representing properties of entities

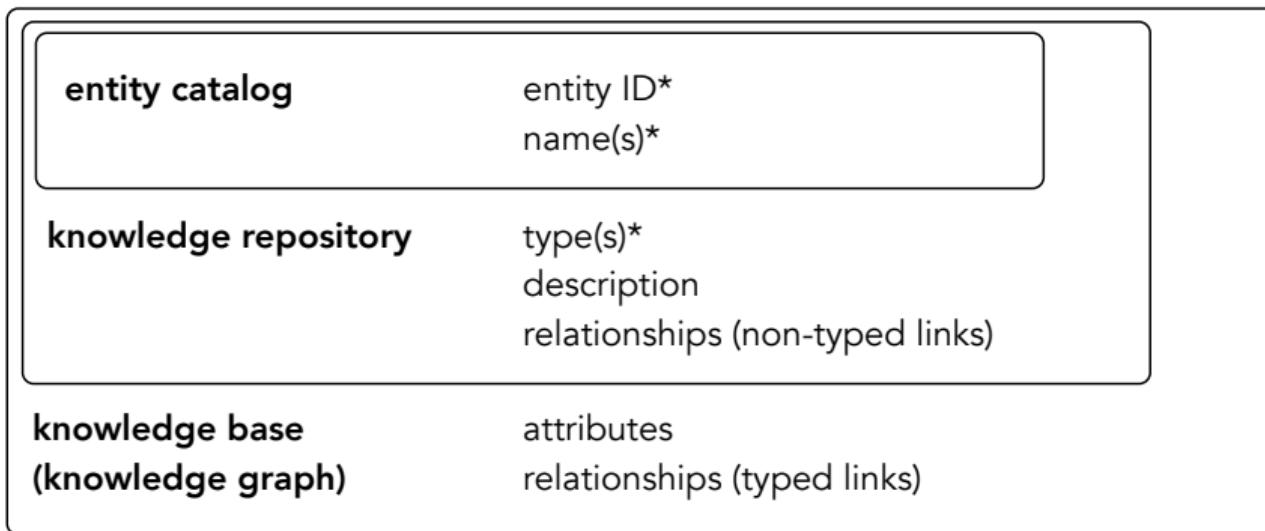


Figure: Illustration is taken from (Balog, 2018) [Fig. 1.2]

Wikipedia

- One of the most popular web sites in the world and a trusted source of information for many people
- Content is created through the collaborative effort of a community of users, facilitated by a *wiki* platform
- Available in nearly 300 languages, although English is by far the most popular, with over five million articles

Discussion

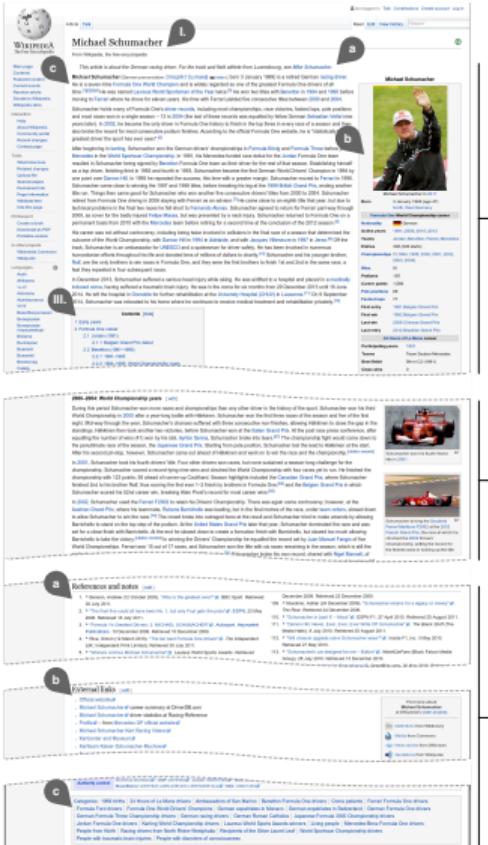
Question

Why is Wikipedia relevant for entity-oriented search?

- One of the most popular web sites in the world and a trusted source of information for many people
- Content is created through the collaborative effort of a community of users, facilitated by a *wiki* platform
- Available in nearly 300 languages, although English is by far the most popular, with over five million articles

Wikipedia as a knowledge repository

- Most of Wikipedia's entries can be considered as (semi-structured) representations of entities



The anatomy of a Wikipedia article

- Title
- Lead section
 - Disambiguation links
 - Infobox
 - Introductory text
- Table of contents
- Body content
- Appendices and bottom matter
 - References and notes
 - External links
 - Categories

Exercise #1

- Make entities in Wikipedia searchable using Elasticsearch
- Code skeleton on GitHub: [exercises/lecture_13/exercise_1.ipynb](https://github.com/.../exercises/lecture_13/exercise_1.ipynb)
(make a local copy)

Knowledge bases

Knowledge base

- A data repository for storing entities and their properties in structured format
- A set of assertions about the world, describing specific entities and their relationships
- Conceptually, it forms a graph (“knowledge graph”)

Resource description Framework (RDF)

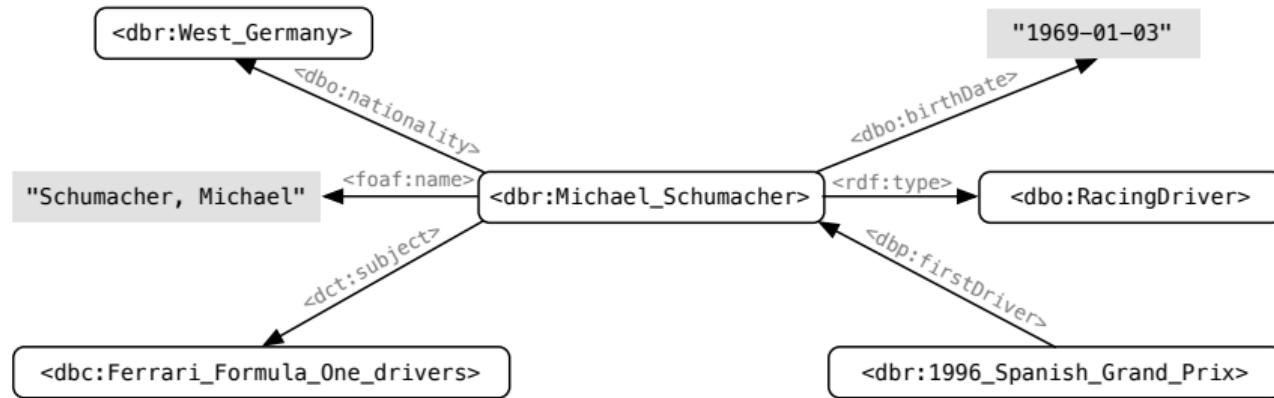
- A language designed to describe “things” (which are referred to as *resources*)
- Each resource is assigned a *Uniform Resource Identifier* (URI), making it uniquely and globally identifiable
- Each RDF statement is a triple, consisting of subject, predicate, and object components
 - **Subject:** always a URI, denoting a resource
 - **Predicate:** always a URI, corresponding to a relationship or property of the subject resource
 - **Object:** either a URI (referring to another resource) or a literal

Example

Michael Schumacher (born 3 January, 1969) is a retired German racing driver, who raced in Formula One for Ferrari.

subject	predicate	object
<dbr:Michael_Schumacher>	<foaf:name>	"Schumacher, Michael"
<dbr:Michael_Schumacher>	<dbo:birthPlace>	<dbr:West_Germany>
<dbr:Michael_Schumacher>	<dbo:birthDate>	"1969-01-03"
<dbr:Michael_Schumacher>	<rdf:type>	<dbo:RacingDriver>
<dbr:Michael_Schumacher>	<dct:subject>	<dbc:Ferrari_Formula_One_drivers>

Example



Related technologies

- RDF describes the instance level in the knowledge base
- RDFS and OWL are vocabularies for ontological modeling
 - An ontology is a means to formalizing knowledge. Building blocks of an ontology include classes, instances, relations, attributes, restrictions, and rules and axioms.
- Serializations for RDF data: Notation-3, Turtle, N-Triples, RDFa, and RDF/JSON
- SPARQL is a structured query language for retrieving and manipulating RDF data
- Triplestores are special-purpose databases designed for storing and querying RDF data

Public knowledge bases

- Cyc
 - Started in 1984 with the goal to manually build a knowledge base of everyday common knowledge
 - ... still building and far from complete
 - “one of the most controversial endeavors of the artificial intelligence history”

Public knowledge bases

- DBpedia
 - Extracted from Wikipedia (mostly from infoboxes) using a set of manually constructed mapping rules
 - Community effort, users collaboratively create and edit the mapping rules
 - Available in multiple languages
 - Contains over 5 million entities (English)

Example

About: Stavanger

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

Stavanger /stə'væŋər/ (Norwegian pronunciation: [sta'vɑŋər]) is a city and municipality in Norway. The city is the third-largest urban zone and metropolitan area in Norway (through conurbation with neighbouring Sandnes) and the administrative centre of Rogaland county. The municipality is the fourth most populous in Norway. Located on the Stavanger Peninsula in Southwest Norway, Stavanger counts its official founding year as 1125, the year the Stavanger Cathedral was completed. Stavanger's core is to a large degree 18th- and 19th-century wooden houses that are protected and considered part of the city's cultural heritage. This has caused the town centre and inner city to retain a small-town character with an unusually high ratio of detached houses, and has contributed signif

Property	Value
dbo:PopulatedPlace/areaMetro	▪ 2598.0
dbo:PopulatedPlace/areaTotal	▪ 71.0
dbo:PopulatedPlace/areaUrban	▪ 77.98
dbo:abstract	▪ Stavanger /stə'væŋər/ (Norwegian pronunciation: [sta'vɑŋər]) is a city and municipality in Norway. The city is the third-largest urban zone and metropolitan area in Norway (through conurbation with neighbouring Sandnes) and the administrative centre of Rogaland county. The municipality is the fourth most populous in Norway. Located on the Stavanger Peninsula in Southwest Norway, Stavanger counts its official founding year as 1125, the year the Stavanger Cathedral was completed. Stavanger's core is to a large degree 18th- and 19th-century wooden houses that are protected and considered part of the city's cultural heritage. This has caused the town centre and inner city to retain a small-town character with an unusually high ratio of detached houses, and has contributed significantly to spreading the city's population growth to outlying parts of Greater Stavanger. The city's rapid population growth in the late 20th century was primarily a result of Norway's booming offshore oil industry. Today the oil industry is a key industry in the Stavanger region and the city is widely referred to as the Oil Capital of Norway. The largest company in the Nordic region, Norwegian energy company Statoil is headquartered in Stavanger. Multiple educational institutions for higher education are located in Stavanger. The largest of these is the University of Stavanger. Domestic and international military installations are located in Stavanger, among these is the North Atlantic Treaty Organisation's Joint Warfare Center. Other international establishments, and especially local branches of foreign oil and gas companies, contribute further to a significant foreign population in the city. Immigrants make up 11.3% of Stavanger's population. Stavanger has since the early 2000s consistently had an unemployment rate significantly lower than the Norwegian and European average. In 2011, the unemployment rate was less than 2%. The city is also among those that frequent various lists of expensive cities in the world, and Stavanger has even been ranked as the world's most expensive city by certain indexes. Stavanger is served by international airport Stavanger Airport, Sola, which offers flights to cities in most major European countries, as well as a limited number of intercontinental charter flights. The airport was named most punctual European regional airport by flightstats.com in 2010. Every two years, Stavanger organizes the Offshore Northern Seas (ONS), which is the second largest exhibition and conference for the energy sector. Gladmat food festival is also held each year and is considered to be one of Scandinavia's leading food festivals. The city is also known for being one of the nation's premier culinary clusters. Stavanger 2008

Public knowledge bases

- Wikidata
 - Operated by the Wikimedia Foundation
 - Its goal is to provide the same information as Wikipedia, but in a structured format
 - Wikidata considers “claims” not “facts”
 - Each claim must be supported by a reference
 - Claims can contradict each other and coexist, thereby allowing opposing views to be expressed (e.g., different political positions)

Example

Wikidata

Main page
Community portal
Project chat
Create a new item
Create a new lexeme
Recent changes
Random item
Query Service
Nearby
Help
Donate
Print/export
Create a book
Download as PDF
Printable version

Tools
What links here
Related changes
Special pages
Permanent link
Page information
Concept URI
Cite this page

Item Discussion

Stavanger (Q25416)

municipality in Rogaland, Norway
Stavanger, Rogaland, Norway | Stavanger, Norway

In more languages

Configure

Language	Label	Description	Also known as
English	Stavanger	municipality in Rogaland, Norway	Stavanger, Rogaland, Norway Stavanger, Norway
Hungarian	Stavanger	No description defined	
Danish	Stavanger	kommune i Rogaland i Norge	
Norwegian Bokmål	Stavanger	kommune i Rogaland i Norge	St. Svitnus by Stavanger kommune Sankt Svitnus by

All entered languages

Statements

instance of municipality of Norway

+ 0 references + add reference + add value

image Vaagen-mod.jpg
800 × 509; 267 KB

+ 0 references + add reference

Wikipedia (66 entries) edit

- af Stavanger
- ar ستافن
- bat-smg Stavangeris
- be-x-old Ставангер
- be Ставангер
- bg Ставангер
- bpy ସ୍ଟାଵଙ୍ଗେର
- br Stavanger
- bs Stavanger
- ca Stavanger
- ceb Stavanger (munisipyo)
- ce ცრანერ
- cs Stavanger
- cy Stavanger
- da Stavanger
- de Stavanger
- el Σταύρος
- en Stavanger
- eo Stavanger
- es Stavanger
- et Stavanger
- eu Stavanger
- fa ستافن
- fi Stavanger
- fo Stavanger
- fro Stavanger
- fr Stavanger
- fy Stavanger
- gl Stavanger
- ha Stavanger
- he סטונגראן
- hr Stavanger
- hu Stavanger

Proprietary knowledge bases

- Google Knowledge Graph
 - "... the knowledge graph is one of Google's biggest search milestones of the last decade..."—Amit Singhal, Google's director of search
- Facebook Entity Graph
- Microsoft Satori
- ...

Connecting knowledge bases

- The same entity may be present in multiple knowledge bases
- A special predicate <owl:sameAs> can be used to connect URIs across different knowledge bases

subject	predicate	object
<dbr:Michael_Schumacher>	<owl:sameAs>	<fb:m.053w4>
<dbr:Michael_Schumacher>	<owl:sameAs>	<>wikidata:Q9671>

The Web of Data

- Increasingly more data is being exposed on the Web in the form of semantic annotations
 - Microdata, RDFa, JSON-LD
- Strong incentive for websites for marking up their content with semantic metadata: It allows search engines to better understand their content
- Standardization: development of schema.org
 - A common vocabulary used by major search providers (including Google, Microsoft, and Yandex) for describing commonly used entity types (including people, organizations, events, products, books, movies, recipes, etc.)

Example

Easy Chicken Satay Recipe - Allrecipes.com



allrecipes.com/recipe/132929/easy-chicken-satay/ ▾

★★★★★ Rating: 4.7 - 282 reviews - 2 hrs 45 mins - 418 cal

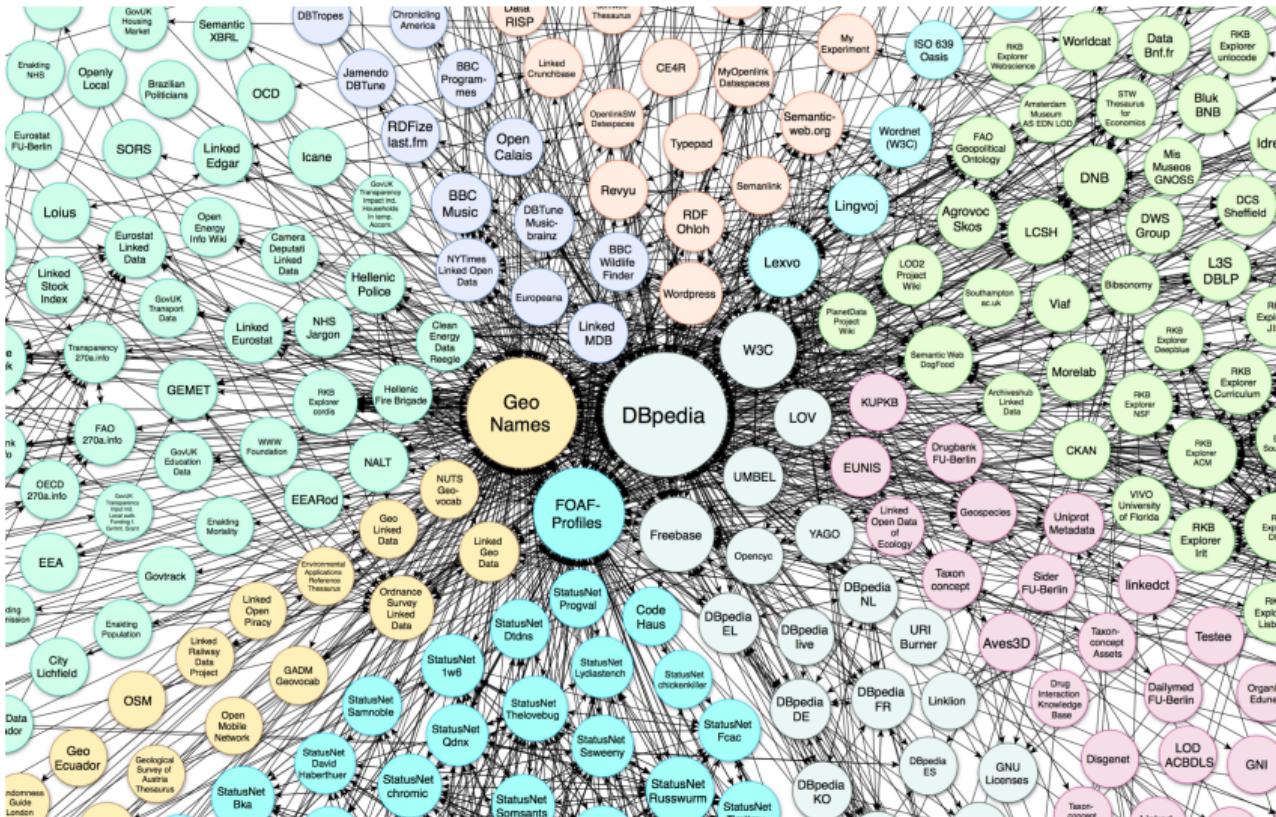
Bring 1 cup coconut milk, 1 tablespoon curry powder, peanut butter, **chicken** stock, and 1/4 cup brown sugar to a simmer in a saucepan over medium-high heat. Simmer for 5 minutes, stirring constantly, until smooth and thickened. Remove from heat and stir in lime juice and soy sauce; season to taste with salt.

```
<section class="ar_recipe_index full-page" itemscope
        itemtype="http://schema.org/Recipe">
    <link href="http://allrecipes.com/recipe/132929/easy-chicken-satay/"
          itemprop="url" />
    <meta itemprop="mainEntityOfPage" content="True" />
```

The Semantic Web

- Historically, data made available in RDF format was referred to as Semantic Web data
- One of the founding principles behind the Semantic Web is that data should be interlinked
- The term *Linked Data (LD)* refers to a set of best practices for publishing structured data on the Web
 - This is facilitated by the special “same-as” predicate
 - A knowledge base published using LD principles should be called *Linked Dataset*
- These “same-as” links connect all Linked Data into a single global data graph
- *Linked Open Data (LOD)* (a casual synonym for the Web of Data) emphasizes the fact that Linked Data is released under an open license

Linked Open Data



Reading

- Entity-Oriented Search (Balog)¹
 - Chapters 1 and 2

¹PDF: <https://rd.springer.com/content/pdf/10.1007%2F978-3-319-93935-3.pdf>