



metaphacts

Smart Data. Smart Apps. Smart Decisions.

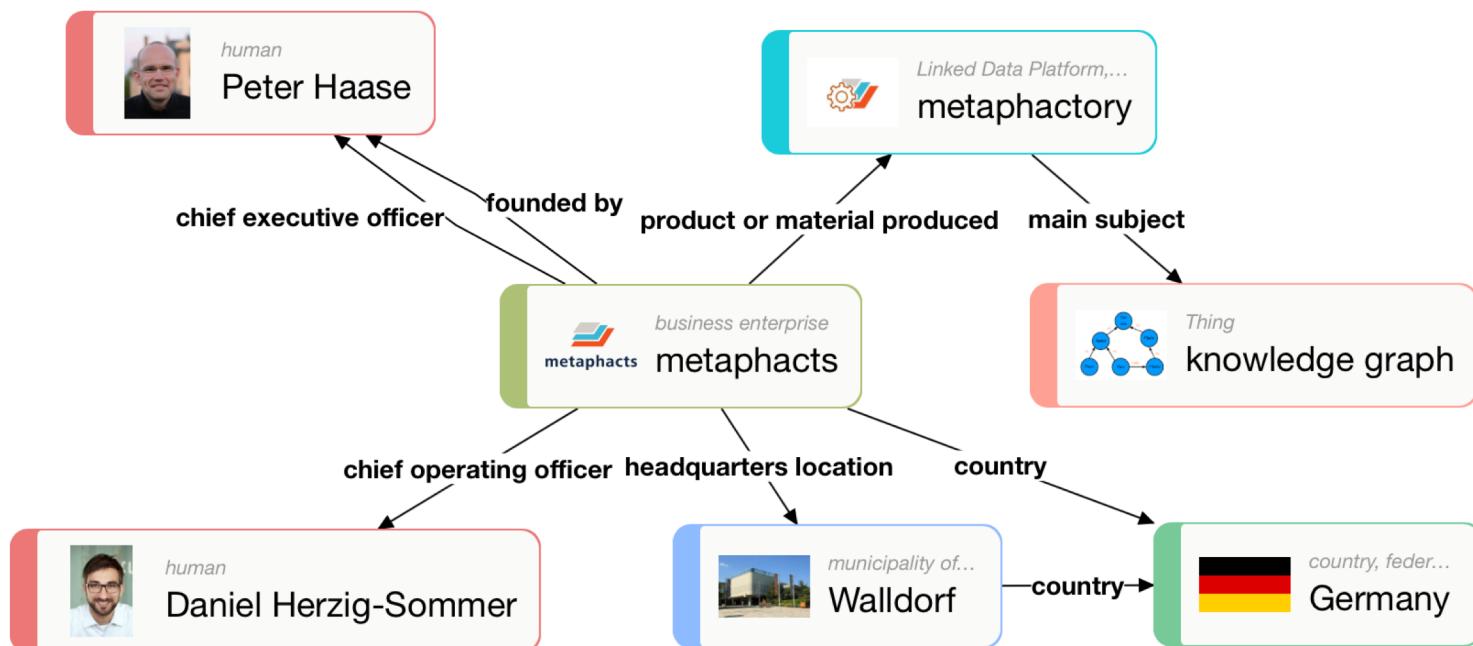
Training “Semantic Technologies”

- An Overview of Semantic Technologies & Knowledge Graphs
- Semantic Technologies for the Life Sciences
- Technologies and Standards: RDF, OWL and SPARQL
- Introduction to SPARQL
- Hands-on Exercises with SPARQL
- Outlook: Concepts and Tooling for end user oriented search, visualization, data analytics

- Write a SPARQL query that ...
 1. Returns isoforms of Hyaluronan synthase 3
 2. Finds proteins phosphorylated and located in the cell membrane
 3. Returns proteins with more than 2 intramembrane regions
 4. Returns proteins located in mitochondrion and are associated with a disease
 5. Returns proteins associated with heart diseases

COMPANY FACTS

- Founded in Q4 2014
- Headquartered in Walldorf, Germany
- Currently ~10 people
- Platform for knowledge graphs and knowledge graph applications





THE TEAM

The metaphacts team offers **unmatched experience and know-how around enterprise knowledge graphs** for our clients in areas such as business, engineering & manufacturing, life sciences, and cultural heritage.

THE PRODUCT

The metaphactory is our end-to-end platform to **ease onboarding into the world of enterprise knowledge graphs** – from knowledge graph management, through rapid application development, to end-user oriented interaction.

THE VALUE

Built entirely on **open standards** and technologies, our platform **supports data engineers along the entire lifecycle** of knowledge graphs and allows to rapidly build data-driven **end-user facing applications**.



Peter Haase

- Interest and experience in ontologies, semantic technologies and Linked Data
- PhD in KR and semantic technologies
- 15 years in academic research and software development
- Contributor to W3C OWL 2 ontology language

NewProt: A protein engineering portal

Andreas Schwarte¹, Hanka Venselaar², Peter Haase¹, Gert Vriend²

¹fluid Operations AG, Walldorf, Germany - ²CMBI, Nijmegen, Netherlands

Introduction

The NewProt project designs and implements a one-stop-shop portal for all *in silico* protein engineering work with a strong focus on point-mutations. The NewProt project is supported by several commercial and academic partners.

The NewProt portal is a platform that gives users access to a broad range of protein engineering tools and services. On an embedded website users can interactively work with these tools and perform their computations supported by integrated workflows. The portal provides a tight integration of standard tools for protein engineering and workflow development.

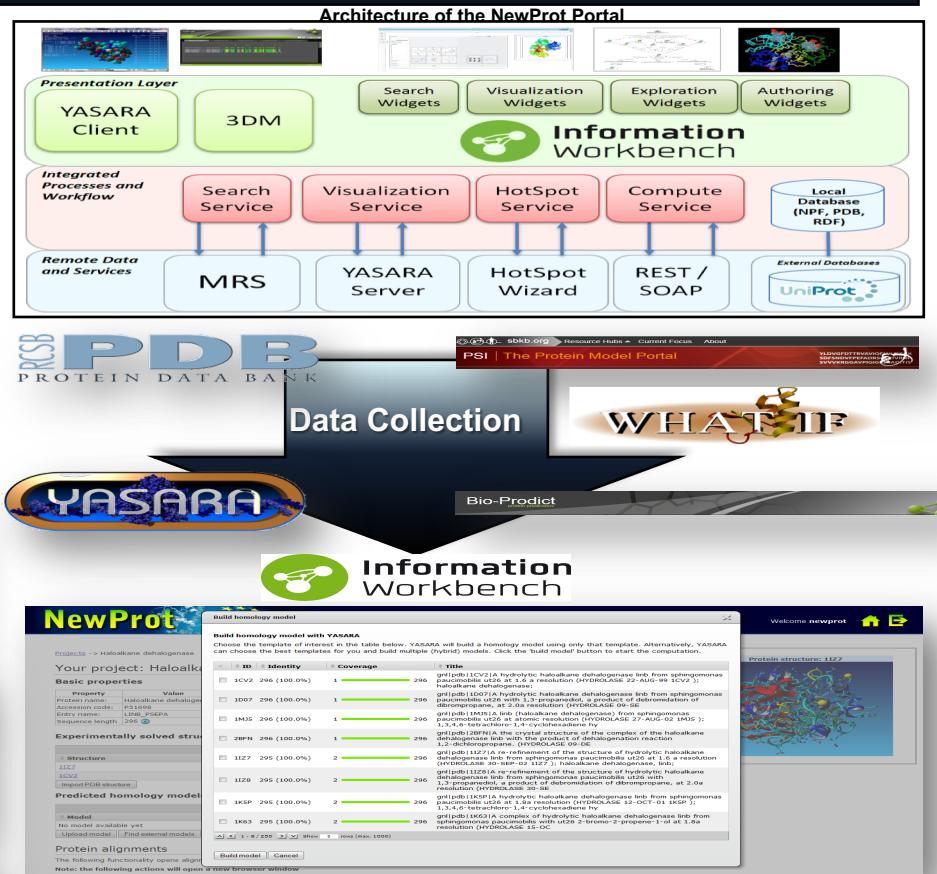
The concept of the NewProt project is to combine and integrate the best European software into a homogeneous portal for *in silico* protein engineering.

Use cases

- Collaborative work on projects
- Analysis of experimentally solved structures
- Analysis and building of homology models
- Visualization with YASARA
- Alignment of Proteins by 3DM
- Calculations using WHAT IF
- Mutation Analysis by HOPE
- HotSpot Computation by HotSpot Wizard
- Integration of compute services
- Project management, sharing and storage of files



The screenshot shows the NewProt portal interface. At the top, there are links for "Structure", "Alignments", "Predicted homology models", "Model", "Alignments", "Predicted homology models", and "Data visualization". Below these are sections for "Protein alignments", "Predicted homology models", and "Uploaded resources". A large central area is labeled "Data Visualization" and displays a 3D molecular model. To the right, there is a table titled "Experimentally solved structures" with columns for ID, Identity, Coverage, and Resolution. At the bottom, there is a "YASARA" interface showing a 3D molecular model and various toolbars.



NPF - The NewProt Format for Achieving Interoperability

The NewProt Format (NPF) is an XML-based file format for data exchange between the different software tools. The format allows residues to be represented as a mapping from Residue to Residue-Value.

```
<residueValue isHotspot="true">
<residue>
  <numbers>7</numbers>
  <chain>A</chain>
  <type>ILE</type>
  <parent_residue>ILE</parent_residue>
  <insertion_code></insertion_code>
  <model_number>0</model_number>
</residue>
<value>
<customValue name="prolineMutation">1.557</customValue>
</value>
</residueValue>
```

OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax (Second Edition)

W3C Recommendation 11 December 2012

This version:

<http://www.w3.org/TR/2012/REC-owl2-syntax-20121211/>

Latest version (series 2):

<http://www.w3.org/TR/owl2-syntax/>

Latest Recommendation:

<http://www.w3.org/TR/owl-syntax>

Previous version:

<http://www.w3.org/TR/2012/PER-owl2-syntax-20121018/>

Editors:

[Boris Motik](#), University of Oxford

Peter F. Patel-Schneider, Nuance Communications

[Bijan Parsia](#), University of Manchester

Contributors: (in alphabetical order)

Conrad Bock, National Institute of Standards and Technology (NIST)

[Achille Fokoue](#), IBM Corporation

[Peter Haase](#), FZI Research Center for Information Technology

[Rinke Hoekstra](#), University of Amsterdam

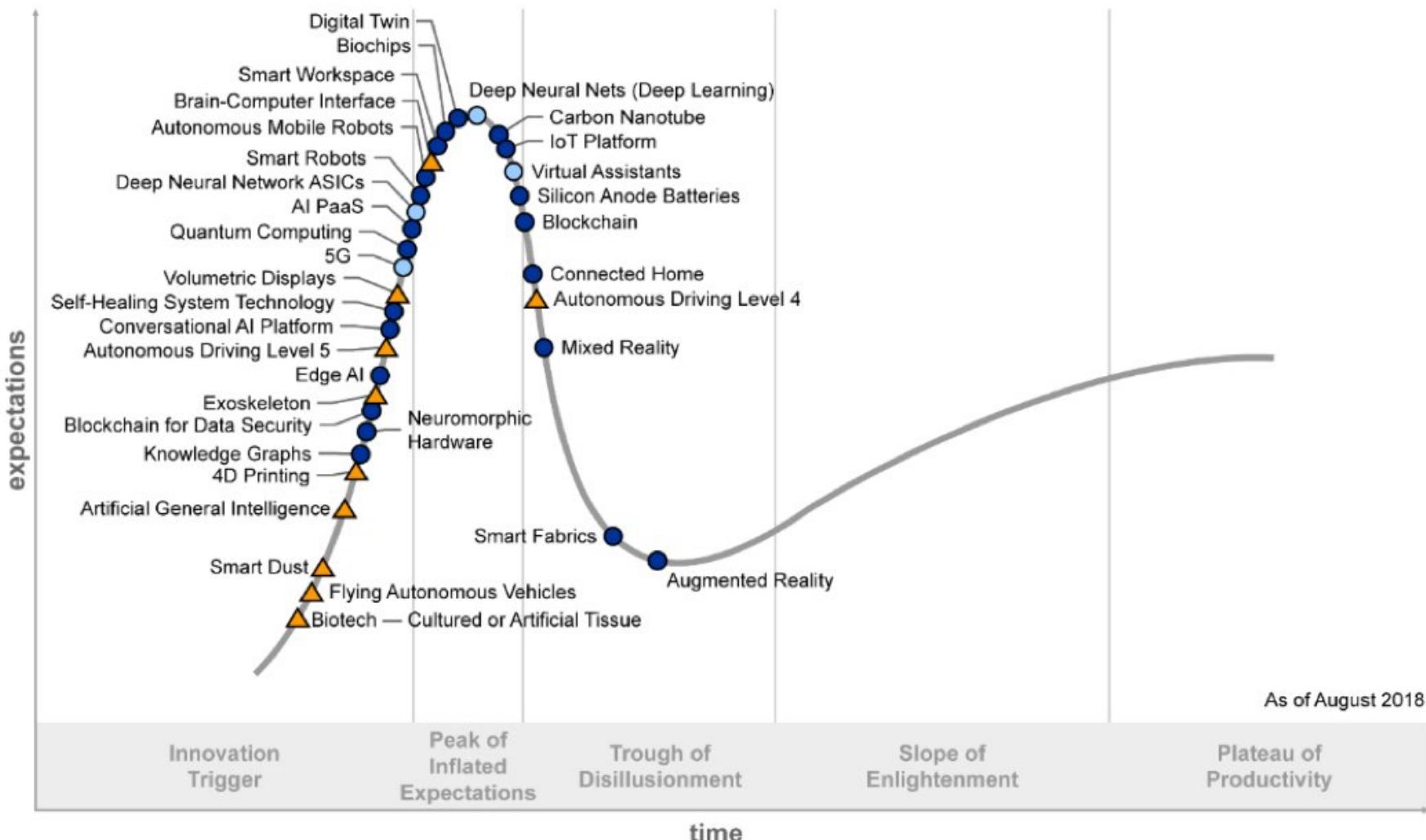
[Ian Horrocks](#), University of Oxford

[Alan Ruttenberg](#), Science Commons (Creative Commons)

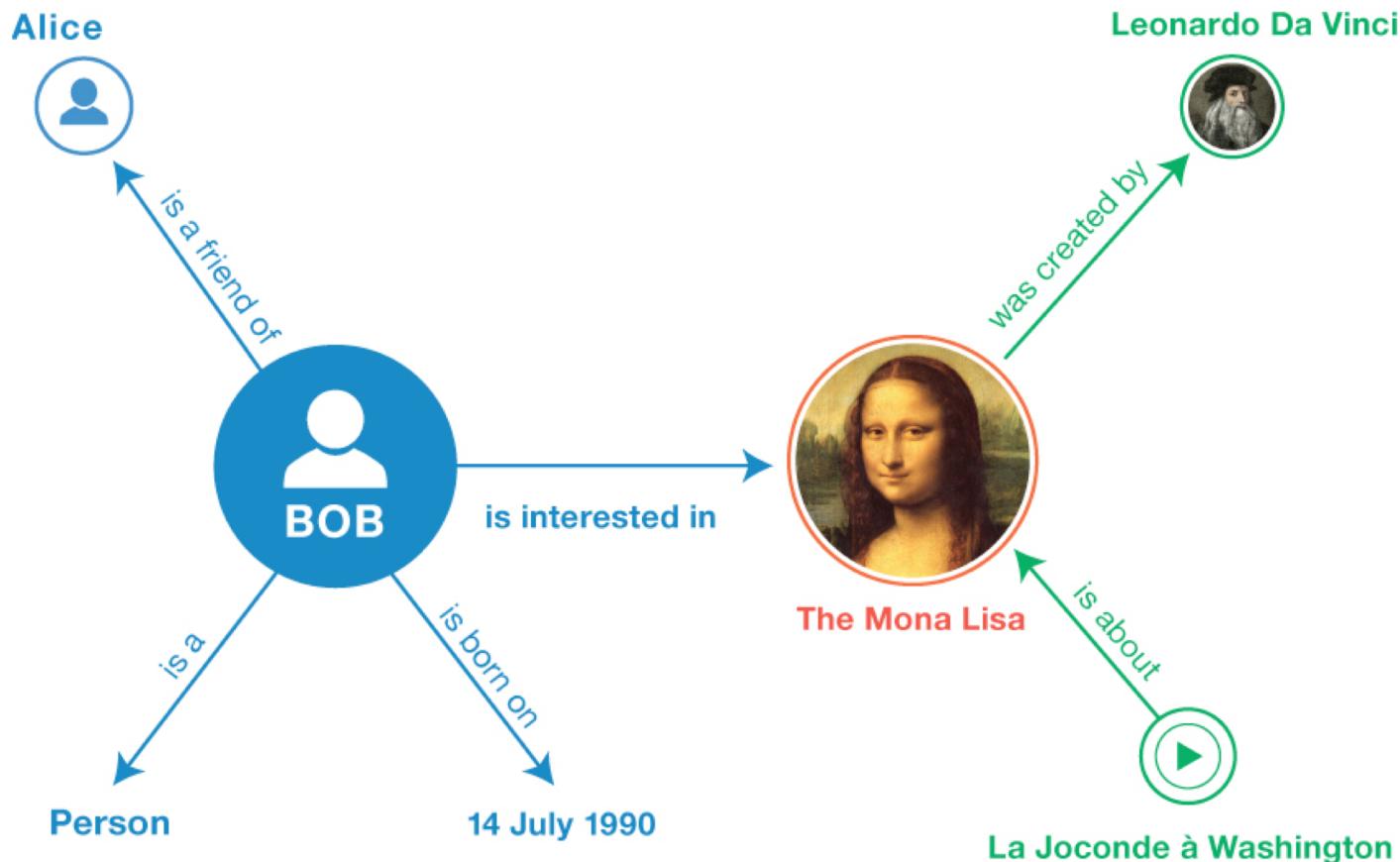
[Uli Sattler](#), University of Manchester

[Michael Smith](#), Clark & Parsia

Knowledge Graphs - Hype Technology?



A (very small) Knowledge Graph

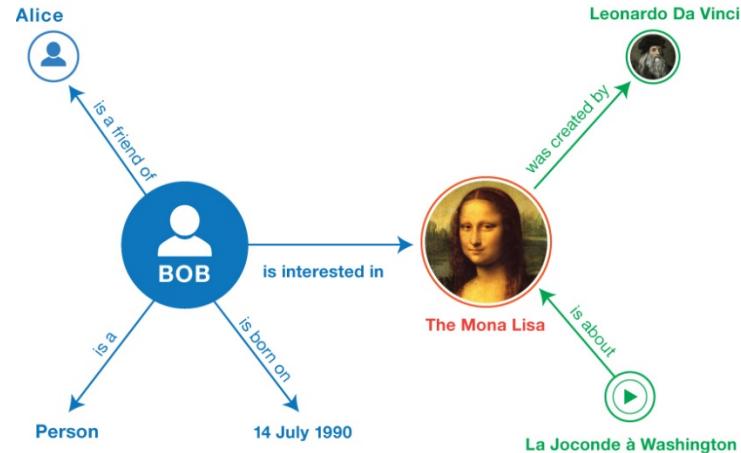


<http://www.w3.org/TR/2014/NOTE-rdf11-primer-20140225/example-graph.jpg>

- We need a **structured and formal representation of knowledge**
- We are surrounded by **entities**, which are connected by **relations**
- **Graphs are a natural way** to represent entities and their relationships
- Graphs can **capture a broad spectrum of data** from structured to unstructured
- Graphs can be managed **efficiently**

What are Knowledge Graphs?

- **Semantic descriptions of entities and their relationships**
- Uses a **knowledge representation formalism** (RDF, RDF-Schema, OWL)



- **Entities**: real world objects (things, places, people) and abstract concepts (genres, religions, professions)
- **Relationships**: graph-based data model where relationships are first-class
- **Semantic descriptions**: types and properties with a well-defined meaning (e.g. through an ontology)
- Possibly axiomatic knowledge (e.g. rules) to support automated reasoning

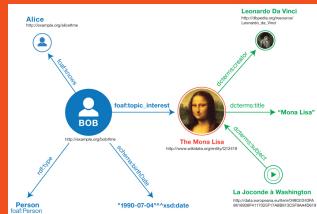
Applications

- Semantic Search
- Question Answering
- Analytics
- Dashboards
- Knowledge Sharing
- Knowledge Management

Algorithms

- Inferencing
- Machine Learning
- Entity Recognition
- Disambiguation
- Text Understanding
- Recommendations

Knowledge Graph



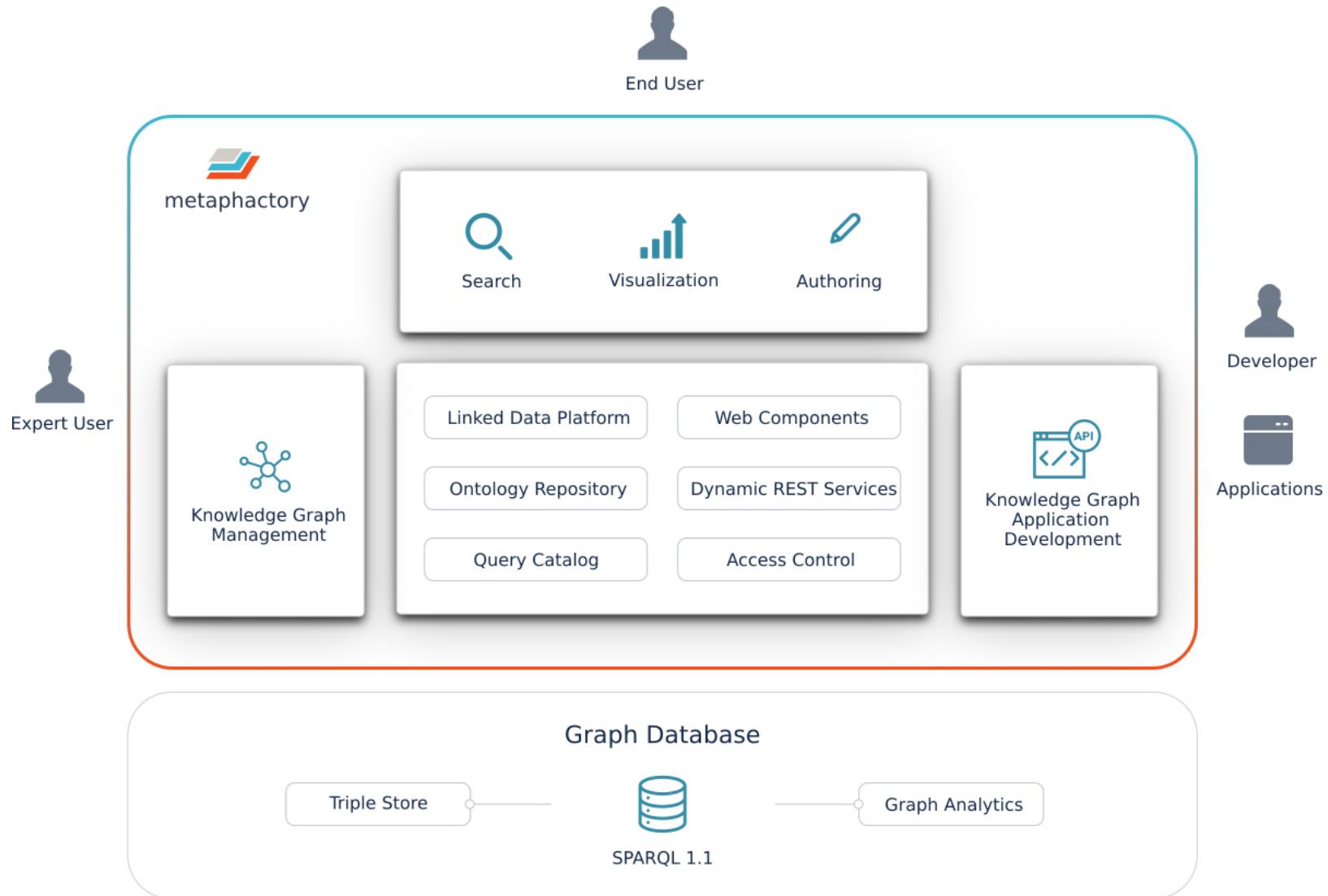
- Entities
- Relationships
- Semantic Descriptions

Data Sources

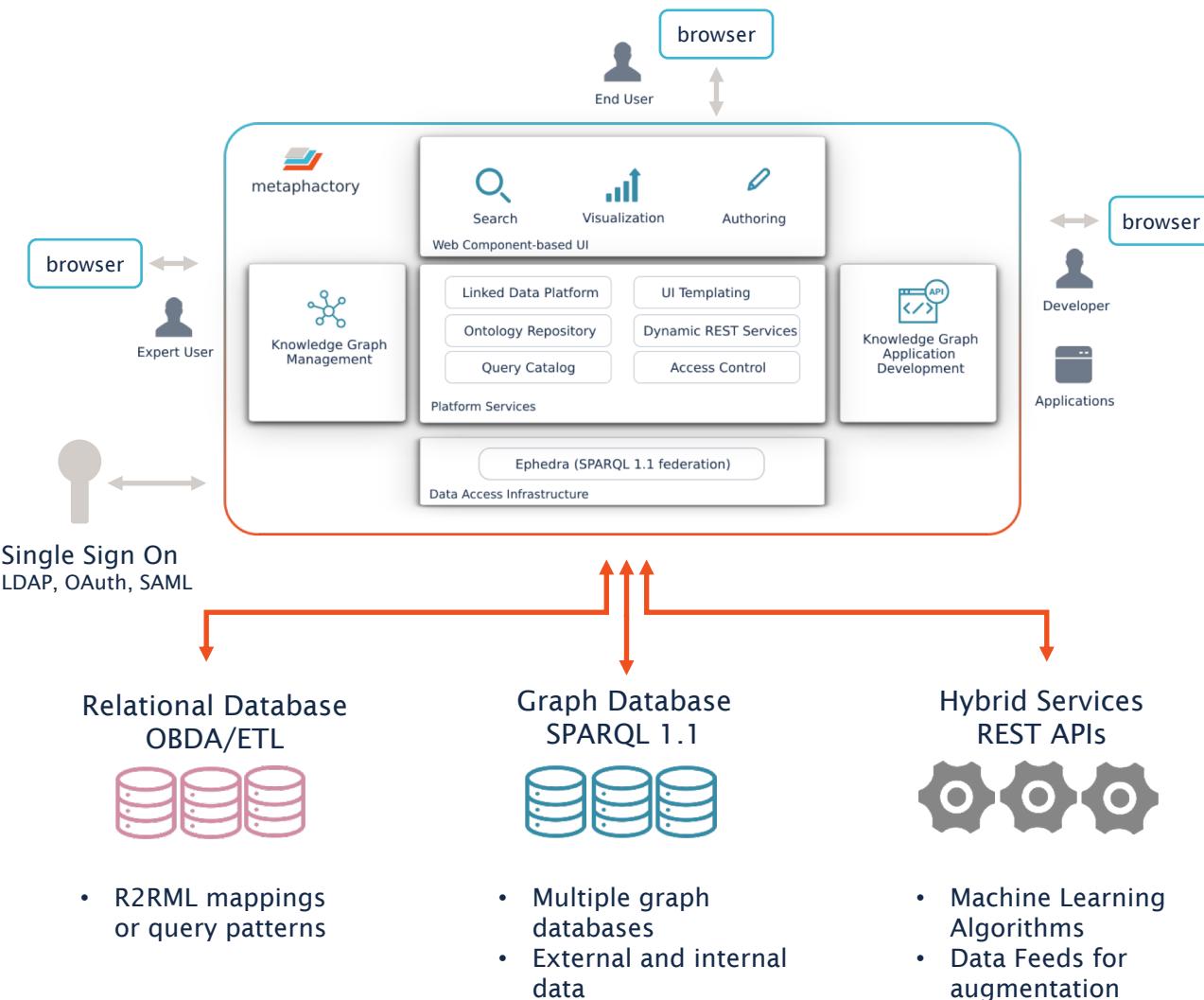
Data Transformation, Integration
Natural Language Processing



metaphactory Platform Architecture



Enterprise Knowledge Graphs Span Multiple Data Spaces



Advantages of Enterprise Knowledge Graphs

- Unlock isolated data silos
- Query across data sources
- 360° view on data

Ephedra – metaphactory's federation engine

Virtual and materialized integration of multiple data sources

- Graph databases
- Relational databases
- Compute services
- REST APIs

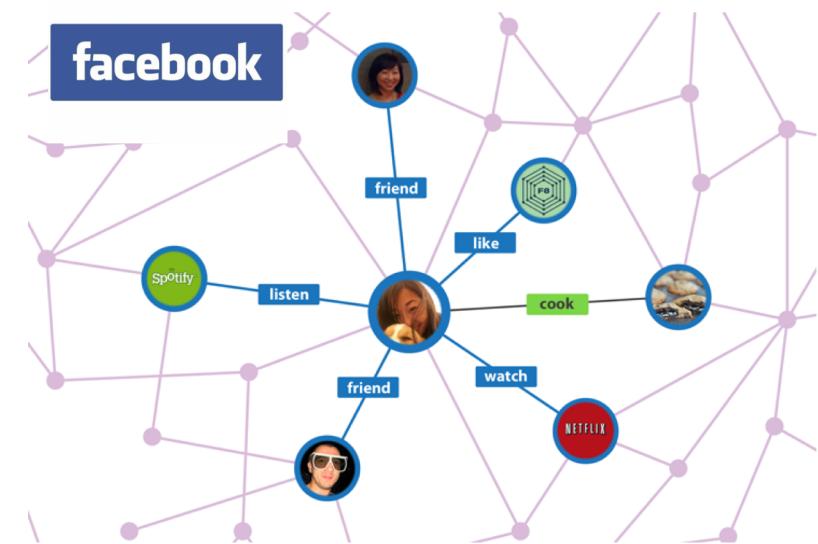
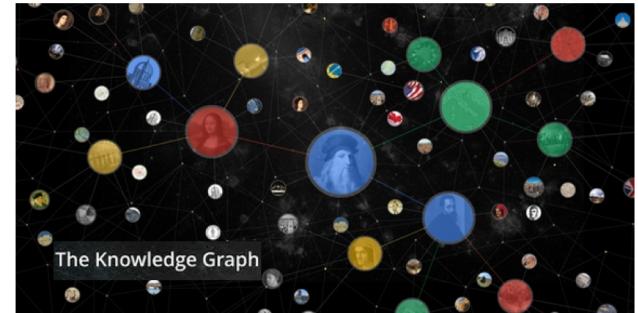
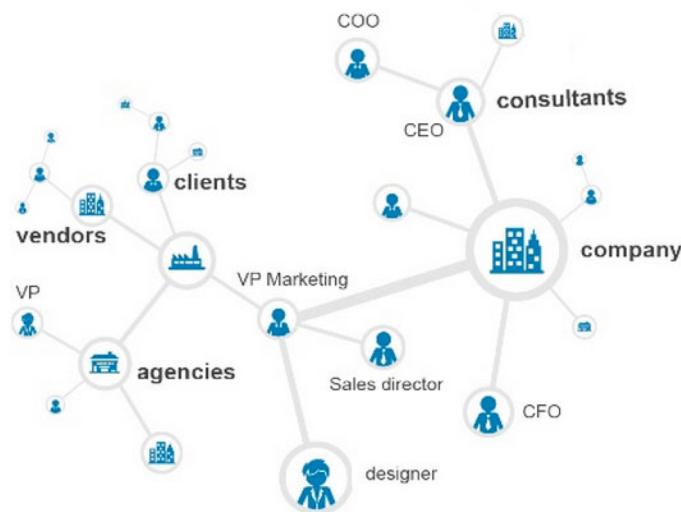


metaphacts

Smart Data. Smart Apps. Smart Decisions.

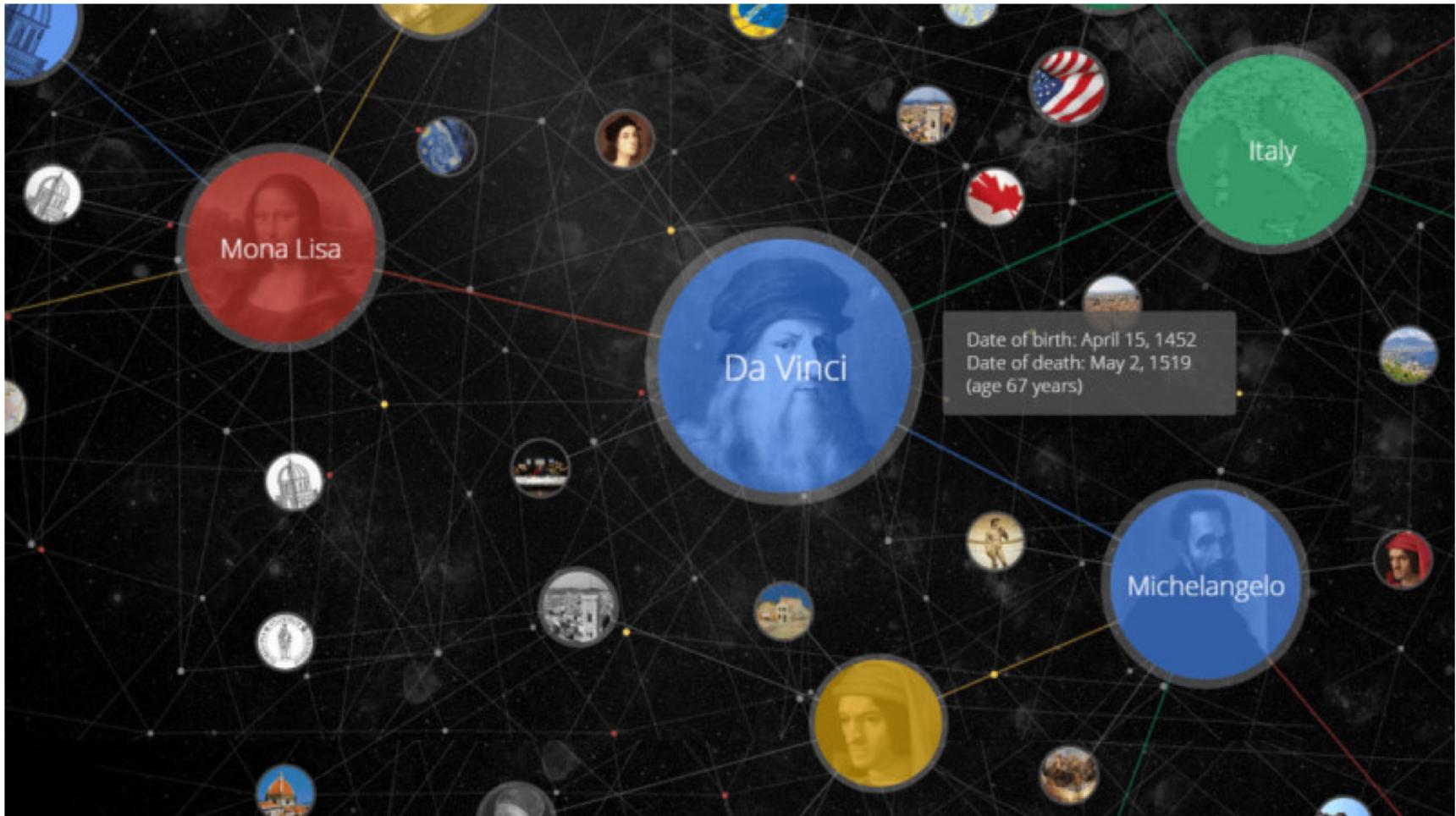
Examples and Applications

The Rise of Knowledge Graphs

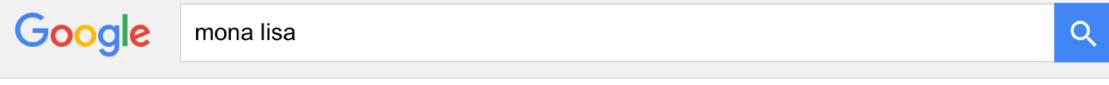


- Web search providers start to build **Knowledge Graphs**
 - Manual / community curation
 - Freebase, DBpedia, Wikidata
 - Annotations / markup
 - schema.org
 - Automated extraction
 - Machine Learning / Information Extraction
- Applications
 - Entity search
 - Improvements to text retrieval
 - Vertical search
 - Related entity recommendations

Google Knowledge Graph



Entity Search and Summarizations



All Images Videos News Maps More ▾ Search tools

About 21,500,000 results (1.11 seconds)

Mona Lisa - Wikipedia

https://en.wikipedia.org/wiki/Mona_Lisa ▾

The Mona Lisa is a half-length portrait of a woman by the Italian Renaissance artist Leonardo da Vinci, which has been acclaimed as "the best known, the most ...
[Louvre](#) · [Mona Lisa](#) · [Mona Lisa \(disambiguation\)](#) · [The Theft of the Mona Lisa](#)

Mona Lisa – Portrait of Lisa Gherardini, wife of Francesco del ... - Louvre

www.louvre.fr/en/.../mona-lisa-portrait-lisa-gherardini-wife-francesco-del-giocondo ▾

The history of the Mona Lisa is shrouded in mystery. Among the aspects which remain unclear are the exact identity of the sitter, who commissioned the portrait, ...

Mona Lisa | painting by Leonardo da Vinci | Britannica.com

<https://www.britannica.com/topic/Mona-Lisa-painting> ▾

Aug 17, 2016 - Mona Lisa, oil painting on a poplar wood panel by the Italian painter, draftsman, sculptor, architect, and engineer Leonardo da Vinci, probably the world's most-famous painting. ... Mona Lisa, oil painting on a poplar wood panel by Leonardo da Vinci, A third suggestion was that ...

Mona Lisa | Leonardo da Vinci | High Renaissance: Florence and ...

<https://www.khanacademy.org/.../celebrity-art-leonardo-s-mona-l...> ▾

Learn about the history and key elements of the Mona Lisa, one of the most famous paintings in the world.

Is the Mona Lisa Hiding Another Portrait? - artnet News

<https://news.artnet.com/art-world/leonardo-da-vinci-mona-lisa-portrait-hidden-388235> ▾

Dec 8, 2015 - According to French scientist Pascal Cotte, who has analyzed the painting for over 10 years, da Vinci's 'Mona Lisa' is hiding another portrait ...



Mona Lisa

Work of art

The Mona Lisa is a half-length portrait of a woman by the Italian Renaissance artist Leonardo da Vinci, which has been acclaimed as "the best known, the most visited, the most written about, the most ...
[Wikipedia](#)

Artist: [Leonardo da Vinci](#)

Dimensions: 77 cm x 53 cm

Location: [The Louvre](#) (since 1797)

Created: 1503

Media: Oil paint

Periods: High Renaissance, Renaissance

Discovering Related Entities

Google leonardo da vinci paintings

All Images Videos Books News More ▾ Search tools

Leonardo da Vinci / Artwork

Most popular first ▾

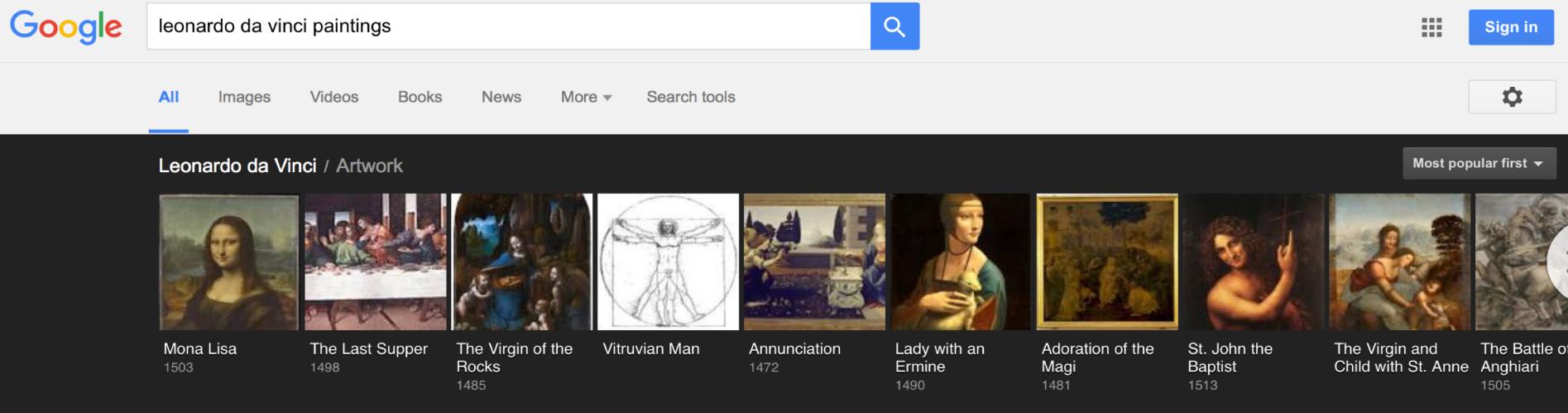


Image	Name	Year
Mona Lisa	Mona Lisa	1503
The Last Supper	The Last Supper	1498
The Virgin of the Rocks	The Virgin of the Rocks	1485
Vitruvian Man	Vitruvian Man	
Annunciation	Annunciation	1472
Lady with an Ermine	Lady with an Ermine	1490
Adoration of the Magi	Adoration of the Magi	1481
St. John the Baptist	St. John the Baptist	1513
The Virgin and Child with St. Anne	The Virgin and Child with St. Anne	
The Battle of Anghiari	The Battle of Anghiari	1505

List of works by Leonardo da Vinci - Wikipedia

https://en.wikipedia.org/wiki/List_of_works_by_Leonardo_da_Vinci ▾

The Last Supper is Leonardo da Vinci's most famous work of religious art and his only surviving mural.

Leonardo da Vinci (baptised Leonardo di ser Piero da Vinci) (pronunciation (help·info)), (April 15, 1452 – May 2, 1519) was one of the leading artists of the High Renaissance.

[Major extant works](#) · [Disputed attributions](#) · [Lost works](#) · [Some recent attributions](#)

Leonardo da Vinci - Wikipedia

https://en.wikipedia.org/wiki/Leonardo_da_Vinci ▾

Jump to [Painting](#) - Leonardo was, and is, renowned primarily as a painter. Among his works, the [Mona Lisa](#) is the most famous and most parodied portrait and [The Last Supper](#) the most reproduced religious painting of all time.

Leonardo Da Vinci - The complete works

www.leonardoda-vinci.org/ ▾

Leonardo Da Vinci - Homepage. The complete [works](#), large resolution images, ecard, rating, slideshow and more! One of the largest [Leonardo Da Vinci](#) ...

[Biography](#) · [Self Portrait](#) · [The Works](#) · [The Last Supper 1498](#)



Leonardo da Vinci

Mathematician

Leonardo di ser Piero da Vinci, more commonly Leonardo da Vinci or simply Leonardo, was an Italian polymath whose areas of interest included invention, painting, sculpting, architecture, science, music, ...
[Wikinedia](#)

Factual Answers

when was leonardo da vinci born

All Images News Videos Shopping More Search tools

About 2,960,000 results (0.97 seconds)

Leonardo da Vinci / Born

April 15, 1452, Anchiano

Michelangelo March 6, 1475 Leonardo DiCaprio November 11, 1974 William Shakespeare April 1564

Feedback

People also ask

How old was Leonardo da Vinci when he died?

What did Leonardo da Vinci do?

Why is Leonardo da Vinci well known?

What things did Leonardo da Vinci study?

Leonardo da Vinci

Mathematician

Leonardo di ser Piero da Vinci, more commonly Leonardo da Vinci or simply Leonardo, was an Italian polymath whose areas of interest included invention, painting, sculpting, architecture, science, music, ... [Wikipedia](#)

Born: April 15, 1452, [Anchiano](#)

Died: May 2, 1519, [Clos Lucé, Amboise, France](#)

Buried: [Chapel of Saint-Hubert](#)

Siblings: [Giovanni Ser Piero](#), [Guglielmo Ser Piero](#), more

Quotes

Simplicity is the ultimate sophistication.

As a well-spent day brings happy sleep, so a life well spent brings happy death.

Learning never exhausts the mind.

Artwork

[View 15+ more](#)





formel ascorbinsäure



Alle Bilder News Maps Videos Mehr Einstellungen Tools

Ungefähr 433.000 Ergebnisse (0,54 Sekunden)

Ascorbinsäure / Formel

C₆H₈O₆

Feedback

Ascorbinsäure – Wikipedia

<https://de.wikipedia.org/wiki/Ascorbinsäure> ▾

Ascorbinsäure ist ein farb- und geruchloser, kristalliner, gut wasserlöslicher Feststoff mit saurem Geschmack. Sie ist eine organische Säure, genauer eine ...

[Geschichte](#) · [Eigenschaften](#) · [Physiologische Bedeutung](#) · [Stoffwechsel im Detail](#)

Ascorbinsäure - Chemie.de

www.chemie.de/lexikon/Ascorbinsäure.html ▾

Ascorbinsäure (in der Biochemie auch Ascorbat) ist eine organische Festkörpersäure. Da sie leicht oxidierbar ist, wirkt sie antioxidativ. Sie ist sehr gut ...

[Geschichte](#) · [Vorkommen](#) · [Physiologische Bedeutung](#) · [Chemische Eigenschaften](#)

Ascorbinsäure in Chemie | Schülerlexikon | Lernhelper

<https://www.lernhelper.de/schuelerlexikon/chemie/artikel/ascorbinsaeure> ▾

Ascorbinsäure ist eine ringförmige Verbindung, die besser unter dem Namen Vitamin C bekannt ist. Obwohl der Stoff keine Carbonsäure ist, kann das Molekül ...

Ascorbinsäure (Vitamin C) - Seilnacht

www.seilnacht.com › [Chemikaliendatenbank](#) ▾

Strukturformel Ascorbinsäure Weißes, kristallines Pulver Vorkommen Höhere Pflanzen und Tiere, Molmasse 176,124 g/mol. AGW keine Angaben Dichte 1,65 ...

Ascorbinsäure



Ascorbinsäure ist ein farb- und geruchloser, kristalliner, gut wasserlöslicher Feststoff mit saurem Geschmack. Sie ist eine organische Säure, genauer eine vinyloge Carbonsäure; ihre Salze heißen Ascorbate. [Wikipedia](#)

Formel: C₆H₈O₆

Molmasse: 176,12 g/mol

PK_s-Wert: 4,25

Vorkommen: Obst, Gemüse, Grüner Tee

Folgen bei Mangel: Skorbut (Morbus Möller-Barlow), Schwächung des Bindegewebes

Trivialname: Vitamin C

Täglicher Bedarf: 100 mg

Feedback

THE ECONOMIC GRAPH





metaphacts

Smart Data. Smart Apps. Smart Decisions.

Examples and Applications

Wikidata

Query the Knowledge of Wikipedia like a Database



Wikipedia page

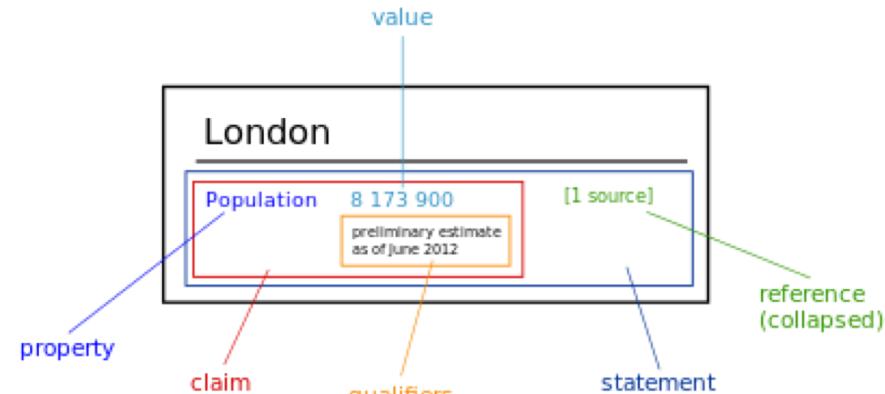
The screenshot shows the Wikipedia article for "London". The page title is "London". Below the title, it says "From Wikipedia, the free encyclopedia". The main content starts with a brief introduction about London being the capital city of England and the United Kingdom. It then discusses its history, population, and status as a global city. There are several sections of text with footnotes and references. On the right side of the page, there is a large image of the London skyline featuring the Shard, the River Thames, and Buckingham Palace. At the bottom of the page, there is a section titled "Central London skyline, Buckingham Palace, and Palace of Westminster" with a smaller image of the same scene.

A query against Wikipedia

The screenshot shows the metaphacts interface. The top navigation bar includes "metaphacts", "sparql", and "Login". The main query results are displayed under the heading "Largest cities in the world with a female mayor". A "Filter Results" input field is present. The results are presented in a table with three columns: "city", "mayor", and "population". The data shows the following cities:

city	mayor	population
Madrid	Ana Botella	3165235
Paris	Anne Hidalgo	2240621
Houston	Annise Parker	2195914
Warsaw	Hanna Gronkiewicz-Waltz	1726581
Prague	Adriana Krnáčová	1243201
Valencia	Rita Barberá Nolla	792303
Taguig	Lani Cayetano	644473
Dresden	Helma Orosz	531982
Zürich	Corine Mauch	400028
Resistencia	Aída Beatriz Máxima Ayala	300000

- **Collecting structured data.** Unlike the Wikipedias, which produce encyclopedic articles, Wikidata collects data, in a structured form.
- **Collaborative.** The data in Wikidata is entered and maintained by Wikidata editors, who decide on the rules of content creation and management in Wikidata supporting the notion of verifiability.
- **Free.** The data in Wikidata is published under the [Creative Commons](#)
- **Large.**
 - 26 million entities
 - 150 million statements
 - 130 million labels
 - 350 languages
 - >1500 million triples



London (Q84)

capital city of England and the United Kingdom

London, England | London, UK | London, United Kingdom

[edit]

Wikipedia (225 entries) [edit]

[Collapse]

► In more languages

Statements

topic's main Wikimedia portal

Portal:London	[edit]
▼ 0 references	[add reference]
	[add]

instance of

metropolitan area	[edit]
▼ 0 references	[add reference]
city with millions of inhabitants	[edit]
► 1 reference	
metropolis	[edit]
▼ 0 references	[add reference]
financial centre	[edit]
▼ 0 references	[add reference]
capital	[edit]
country	United Kingdom
	England
▼ 0 references	[add reference]
	[add]

ab Лондан

ace London

af Londen

als London

am Անջն

ang Lunden

an Londres

arc ლონდონი

ar لندن

arz لندن

ast Londres

ay London

az London

bar London

bat_smg Londons

ba Лондон

bcl Londres

be_x_old Лёндан

be Горад Лондан

bg Лондон

bh লংডন

bi London

bn লন্ডন

bo ښکوټنۍ

br Londrez

bs London

bxr Лондон

ca Londres

cbk_zam London

cdo Lüng-düng

ceb London

ce لوندون

chr Іđ Առահմէծ

ckb սանլ

co Londra

crh London

Leonardo da Vinci (Q762)

Italian Renaissance polymath

Leonardo I da Vinci I Leonardo de Vinza I Léonard I Lionardo I Leonardo d'Avinci I Леонардо I Leonardo da Vince I Liyūnārdū Dāvīnshī I Leonardo di Ser Piero da Vinci I Léonard de Vinci I Leonardo da Vinchi I Leonardo D'Vinci I Leonardo Da Vinci I Leonardo De Vinci I Leonardo de Vinci I da Vinci Leonardo I Leonard Davincy I Lieh-ao-na-to I Leonardo di ser Piero da Vinci

► In more languages

Wikipedia (202 entries)

- ab Leonardo da Vinchi
- af Leonardo da Vinci ☀
- als Leonardo da Vinci
- am ሌያኖርና የስኔንድ
- an Leonardo da Vinci
- arc በታሪክ በትኩረት
- ar ليوناردو دا فينشي
- arz ليوناردو دافينتشي
- ast Leonardo da Vinci
- as লিওনার্ডো দা বিন্চি
- ay Leonardo da Vinci
- azb لئوناردو داوینچی
- az Leonardo da Vinci
- bar Leonardo da Vinci
- bat_smg Leounards da Vénčés
- ba Леонардо да Винчи
- bcl Leonardo Da Vinci
- be_x_old Леанарда да Вінчы
- be Леанарда да Вінчы
- bg Леонардо да Винчи ☀
- bh লিয়েনার্ডো দা বিন্চি
- bi Leonardo da Vinci
- bn লিওনার্দো দা ভিঞ্চি
- bo ລາຍෝනාර්දො ດາ ວິນຈີ
- br Leonardo da Vinci
- bs Leonardo da Vinci
- bxr Leonardo da Vinchi
- ca Leonardo da Vinci ☀
- cbk_zam Leonardo da Vinci
- ceb Leonardo da Vinci
- ce Леонардо да Винчи
- ckb Leonardo da Vinci
- cm Leonardo da Vinci
- cs Leonardo da Vinci
- cv Леонардо да Винчи
- cy Leonardo da Vinci
- da Leonardo da Vinci
- de Leonardo da Vinci
- el Λεονάρδο ντα Βίντσι
- en Leonardo da Vinci
- es Leonardo da Vinci
- et Leonardo da Vinci
- fa لئوناردو دافینچی
- fi Leonardo da Vinci
- fr Leonardo da Vinci
- he ליאונרדו דה וינצ'י
- hi लियोनार्डो दा विन्चि
- hr Leonardo da Vinci
- it Leonardo da Vinci
- ja レオナルド・ダ・ヴィンチ
- ka ლეონარდო და ვინჩი
- km Leonardo da Vinci
- kn ಲೆಾನರ್ಡೋ ದಾ ವಿನ್ಚಿ
- ko 레오나르도 다 비ン치
- lt Leonardo da Vinci
- lv Leonardo da Vinci
- mg Leonardo da Vinci
- ml ലീനാർഡോ ദാ വിൻചി
- mr དേഹനാർട്ട് ດ ວິນຈີ
- ne ଲୋନାର୍ଡୋ ଦା ବିନ୍ଚି
- nl Leonardo da Vinci
- no Leonardo da Vinci
- pl Leonardo da Vinci
- pt Leonardo da Vinci
- ro Leonardo da Vinci
- ru Леонардо да Винчи
- sr Leonardo da Vinci
- sv Leonardo da Vinci
- th ແລ້ວນາຣດ້ ດາ ວິນຈີ
- tr Leonardo da Vinci
- uk Леонардо да Вінчі
- vi Leonardo da Vinci
- zh 雅各布·列奧納多·達芬奇

Statements

instance of

human

2 references

imported from	English Wikipedia
stated in	artist list of the National Museum of Sweden
publication date	12 February 2016
retrieved	27 February 2016
reference URL	http://kulturnav.org/ba0098eb-25b5-4c6a-8af7-2d7d884bf1fc

image

Leonardo self.jpg

0 references

sex or gender

male

3 references

imported from	Virtual International Authority File
stated in	Integrated Authority File
retrieved	9 April 2014
stated in	artist list of the National Museum of Sweden
publication date	12 February 2016



Vitamin C (Q199678)

nutrient found in citrus fruits and other foods

L-ascorbic acid | ascorbate | Cevelin | Cetane-Caps TC | Ascorbicap | Cetane | Cebione | Cenolate | Cecon | Acide ascorbique | L-Ascorbate | Acidum ascorbinicum | L-Ascorbic Acid | Ascoltin | Acidum ascorbicum | Acido ascorbico | Ascorbinsaeure | Ascorbate | Ascorbic Acid | L-(+)-ascorbic acid | Cevex | Vitamin C / ascorbic acid

► In more languages

Statements

instance of

chemical compound	edit
► 1 reference	

pharmaceutical drug	edit
► 1 reference	

+ add value

subclass of

vitamin	edit
▼ 0 references	
+ add reference	

+ add value

chemical formula

C ₆ H ₈ O ₆	edit
► 1 reference	

+ add value

chemical structure

L-Ascorbic acid.svg	edit
► 1 reference	

+ add value

topic's main category

Category:Vitamin C	edit
▼ 0 references	
+ add reference	

+ add value

Wikipedia (68 entries) [edit](#)

af Vitamin C

ar فيتامين C

az C vitaminı

bg Витамин C

bn বিটামিন সি

bs Vitamin C

ca Vitamina C

cs Vitamín C

de Vitamin C

dv วิตามินซี

el Βιταμίνη C

en Vitamin C

eo Vitamino C

es Vitamina C

et C-vitamiini

eu C vitamina

fa ویتامین C

fi C-vitamiini

fr Vitamine C

gl Vitamina C

he C ויטמין C

hi विटामिन सी

hr Vitamin C

hy Վիտամին C

id Vitamin C

ja ビタミンC

lv Vitamin C

kn ವಿಟಾಮಿನ್ ಸಿ

ko 비타민 C

ky Витамин C

lt Vitaminas C

lv C vitamīns

mk Витамин Ц

ml വിടാമിൻ സി

ms Vitamin C

ne ଭିଟାମିନ୍ ସି

nn C-vitamin

no Vitamin C

oc Vitamina C

Wikidata Query Service



<https://query.wikidata.org>

Wikidata Query Examples Prefixes Tools Help English

```
1 #Paintings by Leonardo da Vinci
2 #defaultView:ImageGrid
3 SELECT *
4 WHERE
5 {
6 ?item wdt:P31 wd:Q3305213 .
7 ?item wdt:P170 wd:Q762 .
8 ?item wdt:P18 ?pic .
9 }
```

Press [CTRL-SPACE] to activate auto completion.

Data updated **a few seconds** ago

Run Clear 35 Results in 138 ms Display Download Link

commons:Mona Lisa, by Leonardo da Vinci, from C2RMF retouched.jpg
Q wd:Q12418

commons:Lascapigliata.jpg
Q wd:Q797902

commons:Leonardo da Vinci attributed - Madonna Litta.jpg
Q wd:Q1327676

commons:La Battaglia di Anghiari di Leonardo.jpg
Q wd:Q2397593

commons:Leonardo Da Vinci - Vergine delle Rocce (Louvre).jpg
Q wd:Q11935346

commons:St. John the Baptist in the Wilderness (Leonardo da Vinci).jpg
Q wd:Q12419

commons:Madonna of the Yarnwinder (Leonardo da Vinci).jpg
Q wd:Q12420

commons:St. Anne with the Virgin and Child (Leonardo da Vinci).jpg
Q wd:Q12421

commons:The Virgin of the Rocks (Leonardo da Vinci).jpg
Q wd:Q12422



metaphacts

Smart Data. Smart Apps. Smart Decisions.

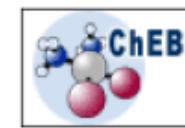
Semantic Technologies for the Life Sciences and Pharma



How do we integrate these resources?



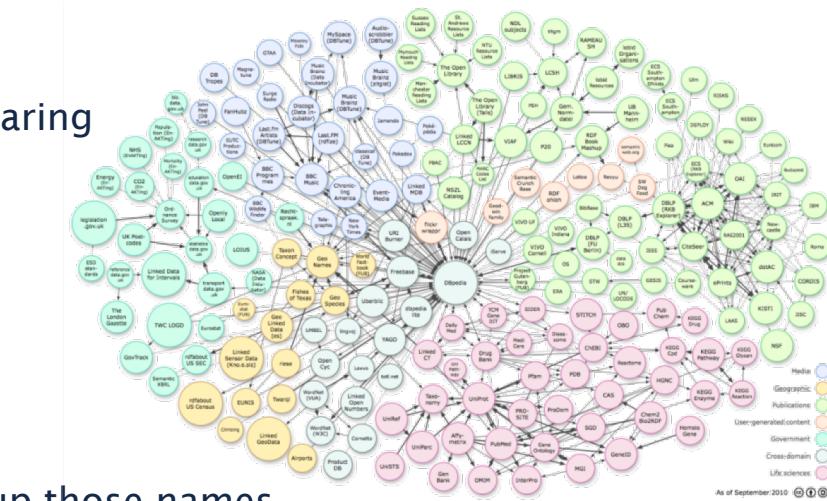
UniSTS
Integrating Markers and Maps



GGCTATTATATAGCTCGATCGAT
ACACCGCGCGGATAGCATGACTGATC
CTCGCTTCCTTGTATGATGATGATC
TTGGGAGAGAGAGGGGGGAAATGCAAG
Nucleotide



- **Linked Data**
- Set of standards, principles for publishing, sharing and interrelating structured knowledge
- From data silos to a Web of Data
- Linked Data principles
 1. Use **URIs** as names for things
 2. Use **HTTP URIs** so that people can look up those names.
 3. When someone looks up a URI, provide useful information, using the standards: **RDF, SPARQL**
 4. Include **links** to other URIs, so that they can discover more things.
- Linked Data as abstraction layer for virtualized data access across data spaces
- Platform independent data access
- Enables data portability and interoperability across current data silos

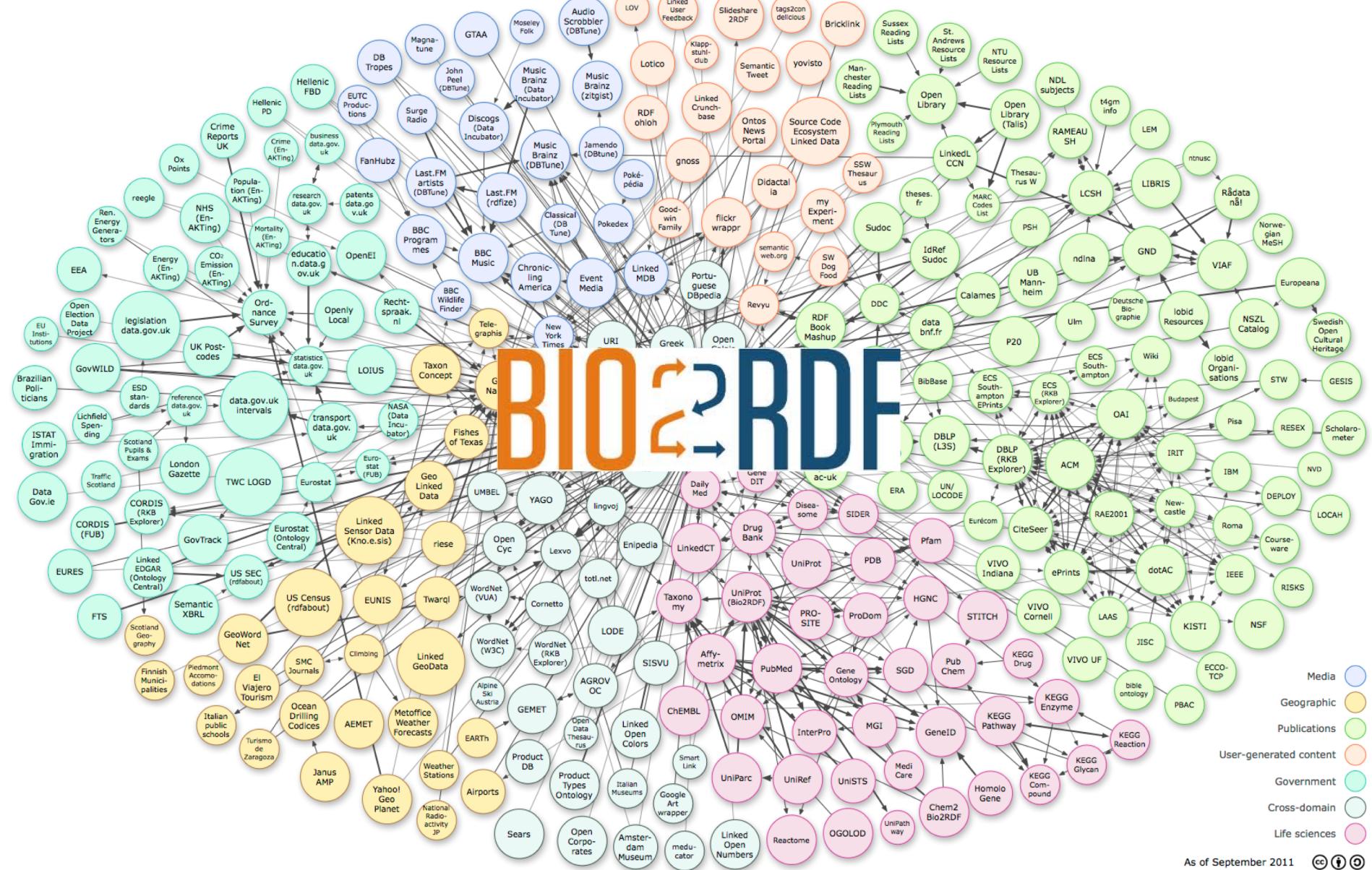




metaphacts

Linked Open Data Cloud

BIO2RDF



Allows one to talk about anything

Uniform Resource Identifier (URI) can be used as entity names

<http://bio2rdf.org/uniprot:P05067>

is a name for **Amyloid precursor protein**

uniprot:P05067

<http://bio2rdf.org/omim:104300>

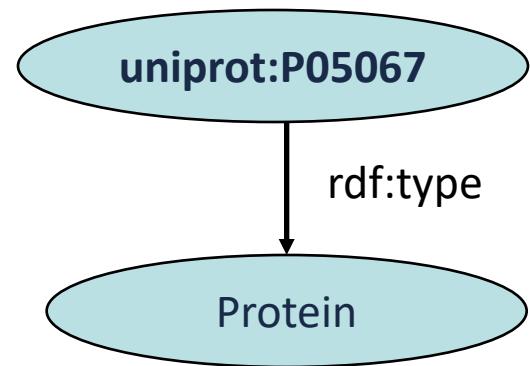
is a name for **Alzheimer disease**

omim:104300

Allows one to express statements

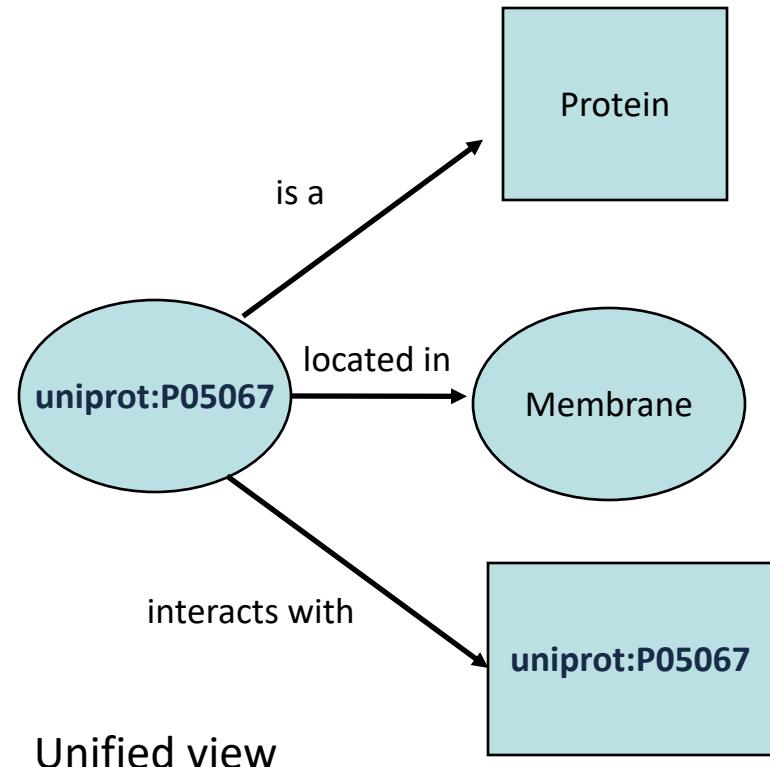
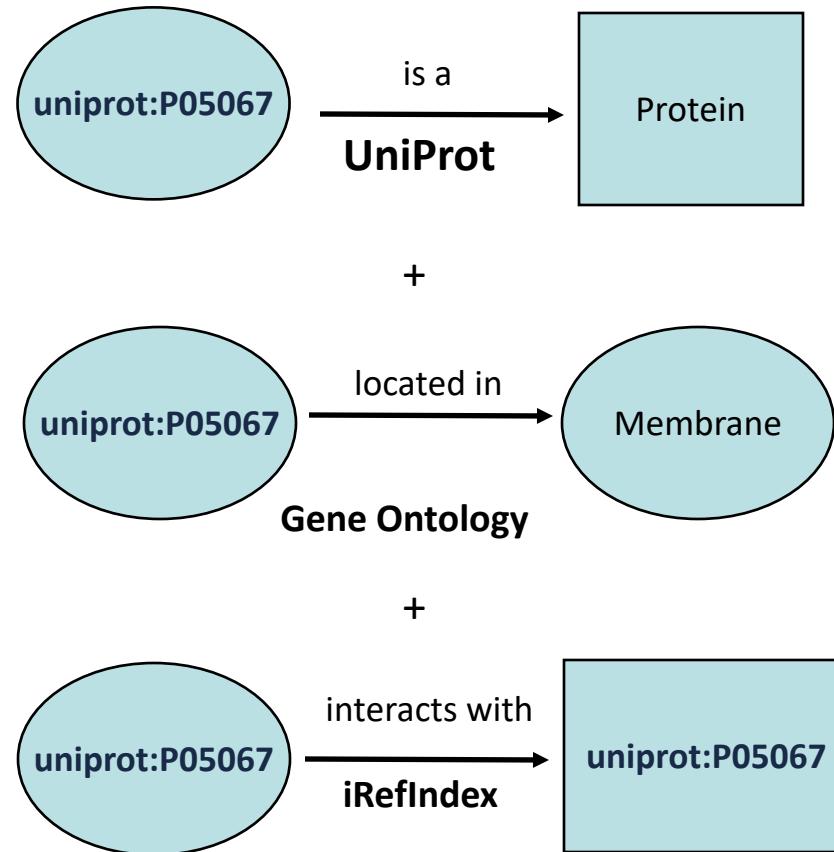
A RDF statement consists of:

- *Subject*: resource identified by a URI
- *Predicate*: resource identified by a URI
- *Object*: resource or literal



Multi-Source Data Integration

depends on consistent naming



Sample SPARQL Query



Select all proteins that are linked to a curated molecular interaction, to inflammatory response and to a target of an existing drug

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
PREFIX biopax2: <http://www.biopax.org/release/biopax-level2.owl#>
PREFIX uniprot: <http://purl.uniprot.org/core/>
PREFIX drugbank: <http://www4.wiwiss.fu-berlin.de/drugbank/resource/drugbank/>

SELECT distinct ?fullname
WHERE {
    ?interaction biopax2:PARTICIPANTS ?participant .
    ?participant biopax2:PHYSICAL-ENTITY ?physicalEntity .
    ?physicalEntity skos:exactMatch ?protein .
    ?protein uniprot:classifiedWith <http://purl.uniprot.org/go/0006954>.
    ?protein uniprot:recommendedName ?name.
    ?name uniprot:fullName ?fullname .
    ?protein uniprot:mnemonic ?mnemonic .
    ?target drugbank:swissprotName ?mnemonic .
}
```

SPARQL Query

Results for [PREFIX rdf...](#) (44)

fullname

- Adenosine receptor A1
- Adenosine receptor A2a
- Arachidonate 15-lipoxygenase
- Annexin A1
- Aldehyde oxidase
- B2 bradykinin receptor
- Complement C5
- Tumor necrosis factor receptor superfamily member 5
- CD40 ligand
- Cysteine dioxygenase type 1
- C-C chemokine receptor type 5
- Cannabinoid receptor 2
- Epoxide hydrolase 2
- Histamine H1 receptor
- Bifunctional heparan sulfate N-deacetylace/N-sulfotransferase 1
- Interferon alpha-2
- Interleukin-1 receptor antagonist protein
- Interleukin-5
- Interleukin-8
- C-X-C motif chemokine 10
- Integrin alpha-L
- Integrin beta-2
- Kininogen-1
- Leukotriene A-4 hydrolase
- Pyrin
- Macrophage migration inhibitory factor
- Nuclear factor NF-kappa-B p105 subunit
- Ras-related C3 botulinum toxin substrate 1



- **One-stop portal** for consolidated access to target-related information such as function, expression, interactions, localisation, sequence, etc.
- **Linked Data and ontology-based data access**
 - Integration of UniProt, neXtProt, Medline, Pubchem, DrugBank databases
 - Data is **linked as an RDF graph**, enabling queries across the sources
 - Ontologies to provide an **integrated view**
- Portal front-end for **end-user data access**
 - Specific **visualizations** including tables, graphs, charts, etc.
 - Google-like **search** with semantic disambiguation and auto-completion
 - Structured queries for complex information with **visual query builder**
 - **Shared queries** managed in a query catalog

Target Dashboard: Search Interface



SANOFI Target Dashboard Toolbox ▾ Account ▾

akt|

X

AKT1 | Homo sapiens
Synonyms: AKT; AKT1; CWS6; PKB; PKB-ALPHA; PRKBA; RAC; RAC-ALPHA
<http://identifiers.org/ncbigene/207>

Akt1 | Rattus norvegicus
Synonyms: Akt; Akt1
<http://identifiers.org/ncbigene/24185>

Akt1 | Mus musculus
Synonyms: Akt; Akt1; LTR-akt; PKB; PKB/Akt; PKBalphalpha; Rac
<http://identifiers.org/ncbigene/11651>

Akt1s1 | Rattus norvegicus
Synonyms: Akt1s1; PRAS40
<http://identifiers.org/ncbigene/292887>

Akt1s1 | Mus musculus
Synonyms: 110012j22Rik; AI227026; Akt1s1; Lobe; Lobel; PRAS40
<http://identifiers.org/ncbigene/67605>

AKT1S1 | Homo sapiens
Synonyms: AKT1S1; Lobe; PRAS40
<http://identifiers.org/ncbigene/84335>

- Google-like “Instant Search”
- Semantic disambiguation

Target Dashboard: Summary Page



SANOFI

Target Dashboard

Toolbox ▾ Account ▾

AKT1
AKT serine/threonine kinase 1

General Information

AKT1 -> AKT serine/threonine kinase 1

Gene symbol: AKT1 [HGNC](#)
Synonyms: AKT , CWS6 , PKB , PKB-ALPHA , PRKBA , RAC , RAC-ALPHA [NCBI](#)
Species: Homo sapiens [NCBI](#)
Chromosomal location: 14q32.33 [NCBI](#)
Family: Belongs to the protein kinase superfamily. AGC Ser/Thr protein kinase family. RAC subfamily. [neXtprot](#)
External Identifiers: P31749 (Uniprot) , 391 (HUGO) , ENSG00000142208 (Ensembl) , 207 (Entrez)

Summary

Program
Data for LC program code information from eSAR datasource is not yet available..

Function Summary
AKT1-specific substrates have been recently identified, including palladin (PALLD), which phosphorylation modulates cytoskeletal organization and cell motility; prohibitin (PHB), playing an important role in cell metabolism and proliferation; and CDKN1A, for which phosphorylation at 'Thr-145' induces its release from CDK2 and cytoplasmic relocalization. These recent findings indicate that the AKT1 isoform has a more specific role in cell motility and proliferation. Phosphorylates CLK2 thereby controlling cell survival to ionizing radiation.
AKT1 is one of 3 closely related serine/threonine-protein kinases (AKT1, AKT2 and AKT3) called the AKT kinase, and which regulate many processes including metabolism, proliferation, cell survival, growth and angiogenesis. This is mediated through serine and/or threonine phosphorylation of a range of downstream substrates. Over 100 substrate candidates have been reported so far, but for most of them, no isoform specificity has been reported. AKT is responsible of the regulation of glucose uptake by mediating insulin-induced translocation of the SLC2AA/GLUT4 glucose transporter to the cell surface. Phosphorylation of PTPN1 at 'Ser-50' negatively modulates its phosphatase activity preventing dephosphorylation of the ins...[read more](#)

Expression
184 proteome known by neXtProt.

Homology
This gene has **12** homologs according to HomoloGene. Ensembl knows about **64** orthologs and **9** paralogs.

Pathways
This gene is involved in **98** pathways from Reactome..

Drug and Disease
Linked to **3** targets from ChEMBL: [CHEMBL2111353](#) (59 activity types) , [CHEMBL4282](#) (6917 activity types) , [CHEMBL3038463](#) (2 activity types) . See all compound related properties for the targets.

- <https://wikidata-new.graprhs.com/resource/app:ChemicalDrawing>
- Structured Search



- **What:** a comprehensive resource that complements SIB/EBI Swiss-Prot human protein annotation efforts. We expect neXtProt to become a central resource for human protein-centric information;
- **How:**
 - by mining, in the most appropriate way and with stringent quality criteria, many high-throughput data resources.
We plan to add additional protein/protein and protein/small molecules interactions, proteomics data, pathways/networks information, variation data (such as SNP frequencies), siRNA screen data, phylogenetic profiling, etc.;
 - by integrating experimental results from an extensive network of collaborating laboratories.

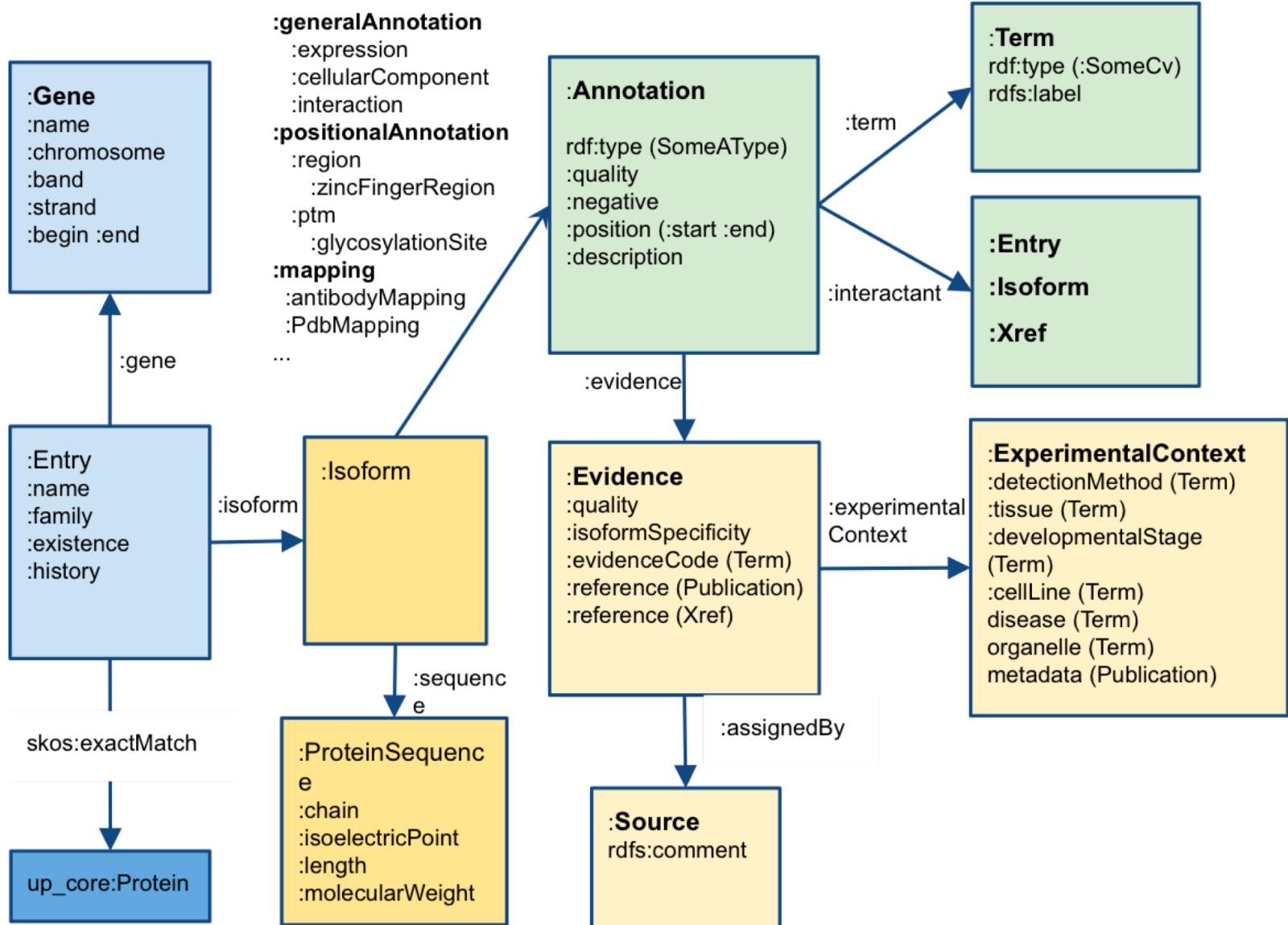
Nextprot Data Model

References

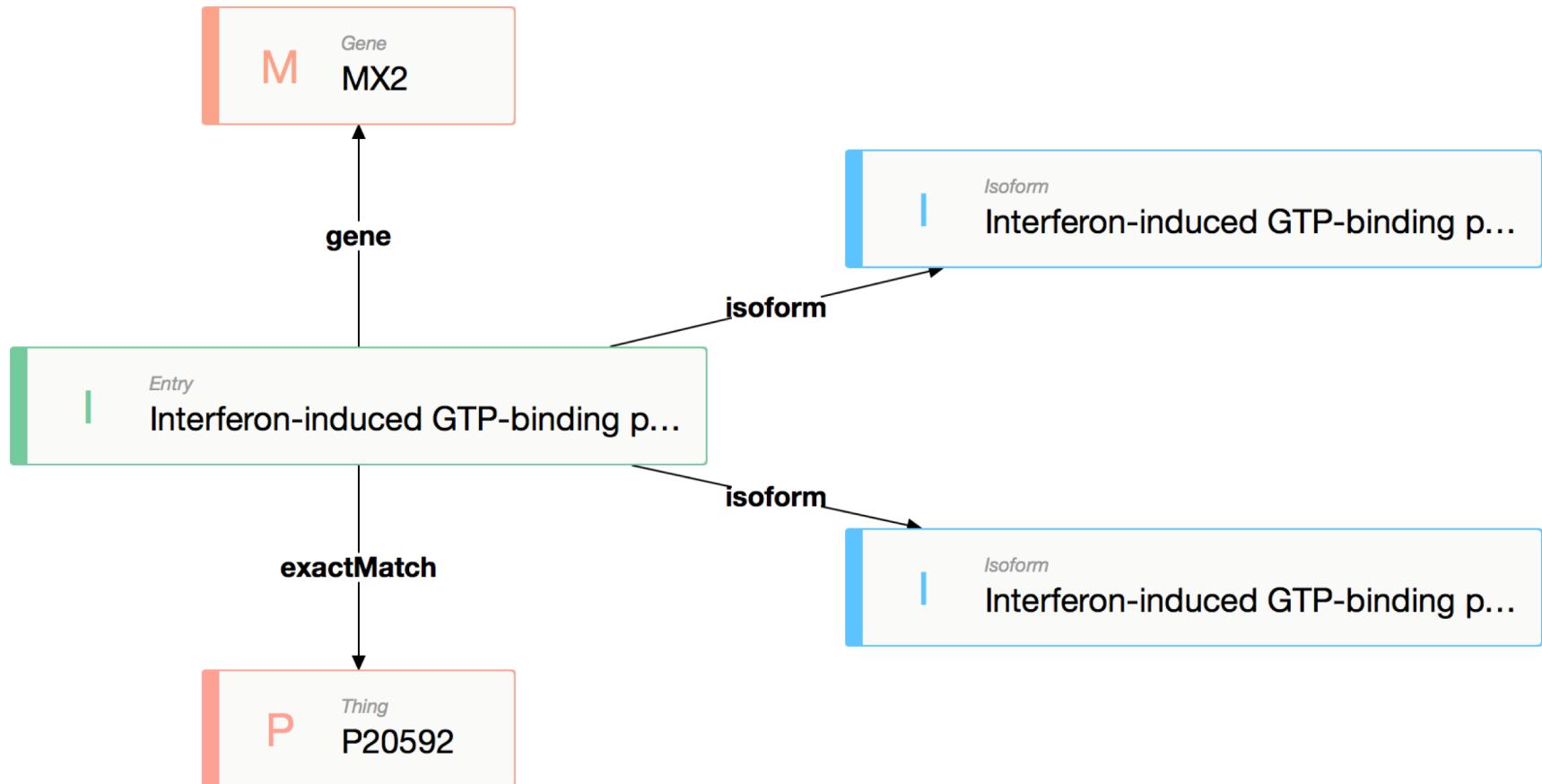
:Identifier
:provenance
:accession

:Publication
:author
:title
:journal
:volume
:first/lastPage
:year

:Xref
:provenance
:accession



Sample Graph



metaphactory

Home Edit Page

nextprot

Gene (Protein) Entry Diseases Uniprot keyword Subcellular locations Human anatomy All Terms

Nextprot schema overview

Ontology Overview

Nextprot data model documentation site

Namespace prefixes

Sample queries

Sample Resources

Type at least 3 characters to search. For example, "akt1"

Gene: MX2
Entry: Interferon-induced GTP-binding protein Mx2
Isoform: NX_P20592-1

Gene: HAS3
Entry: Hyaluronan synthase 3
Isoform: NX_O00219-1

You can use the following form to navigate to any resource in metaphactory by providing its full IRI.

E.g., <http://nextprot.org/rdf/gene/> Navigate

<http://nextprot.metaphacts.cloud/>

- <http://wikidata.metaphacts.com/>
- <http://nextprot.metaphacts.cloud/>

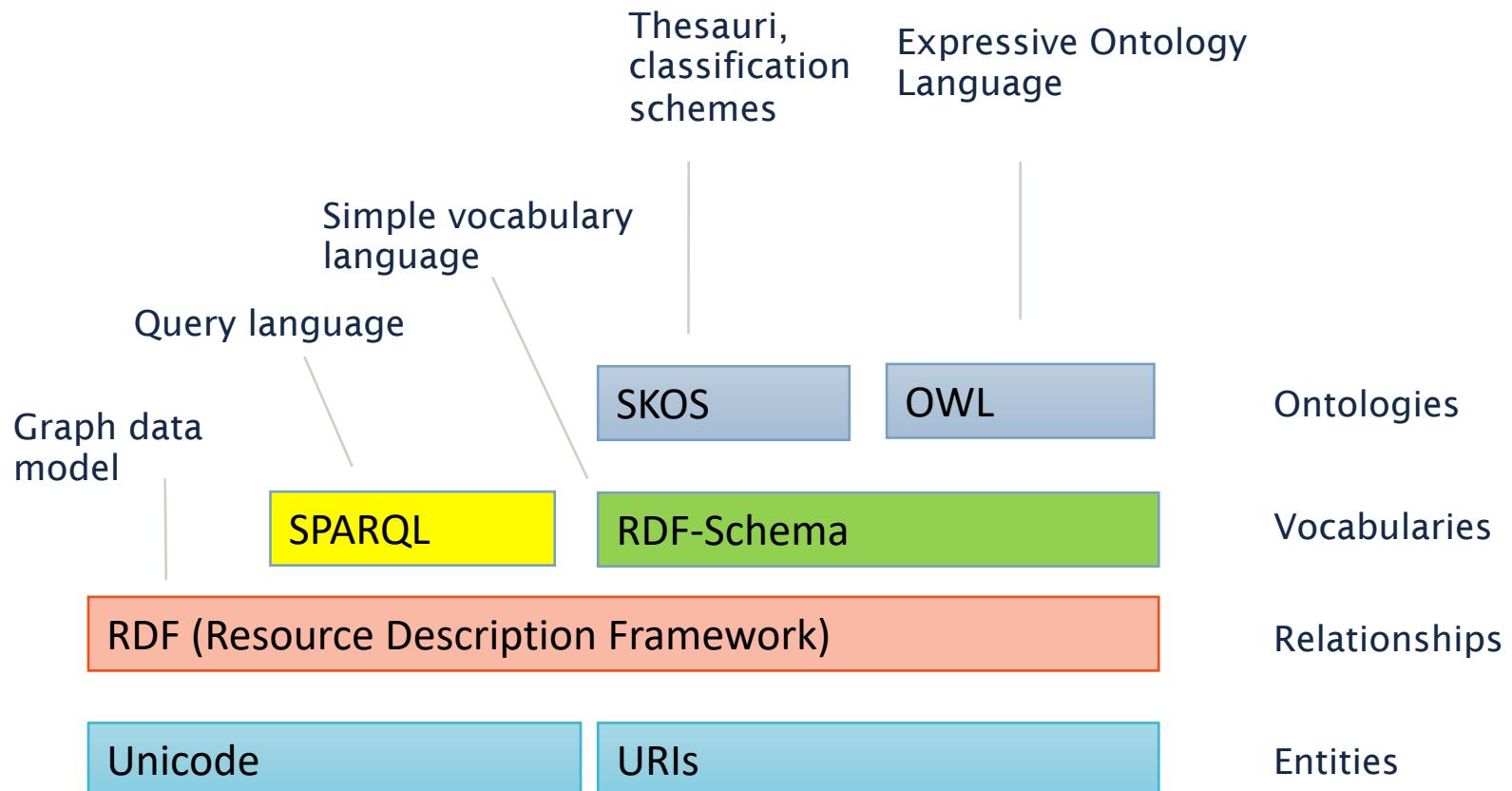


metaphacts

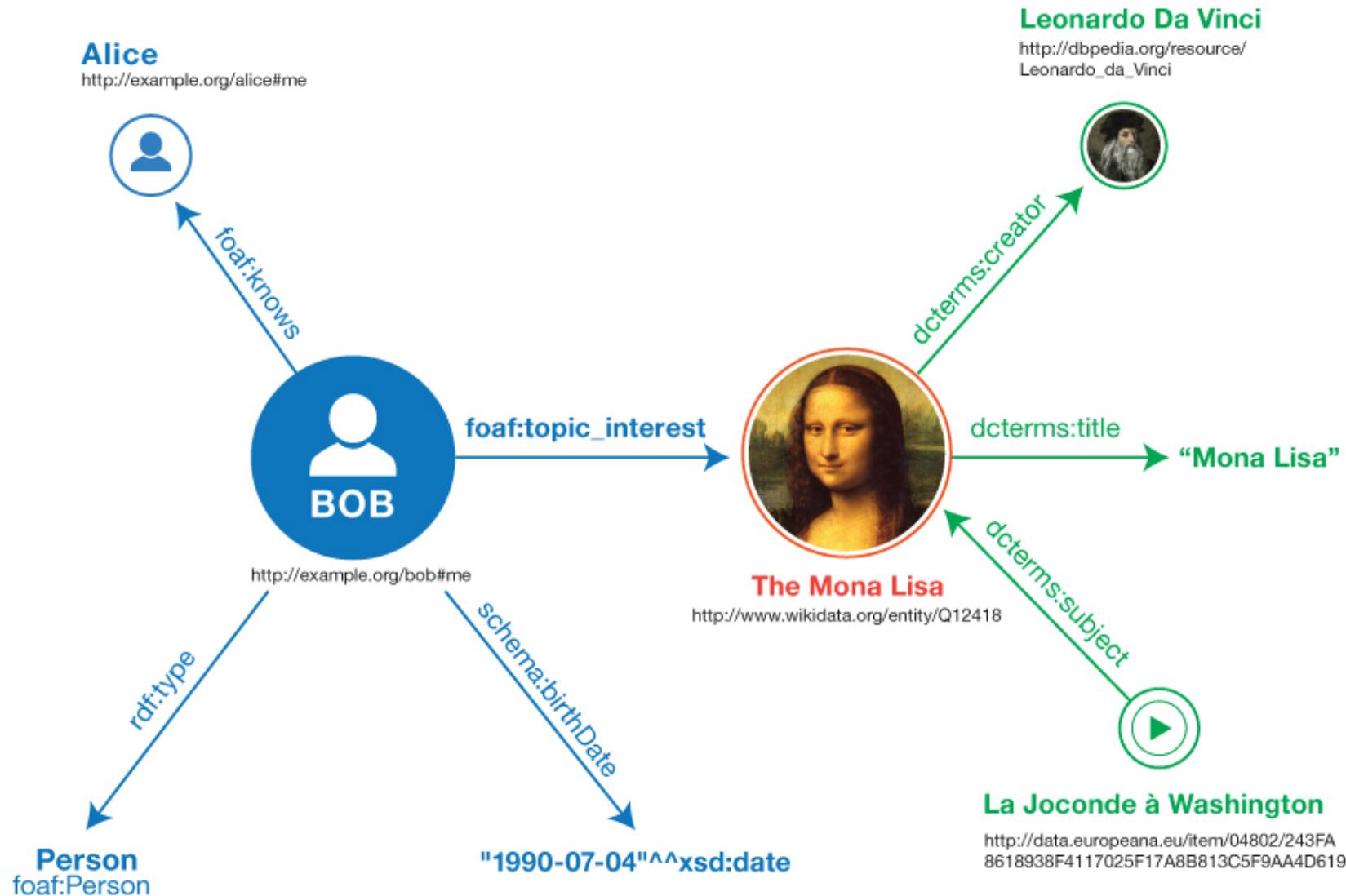
Smart Data. Smart Apps. Smart Decisions.

Standards and Principles

Knowledge Graphs Built on the Semantic Web Layer Cake



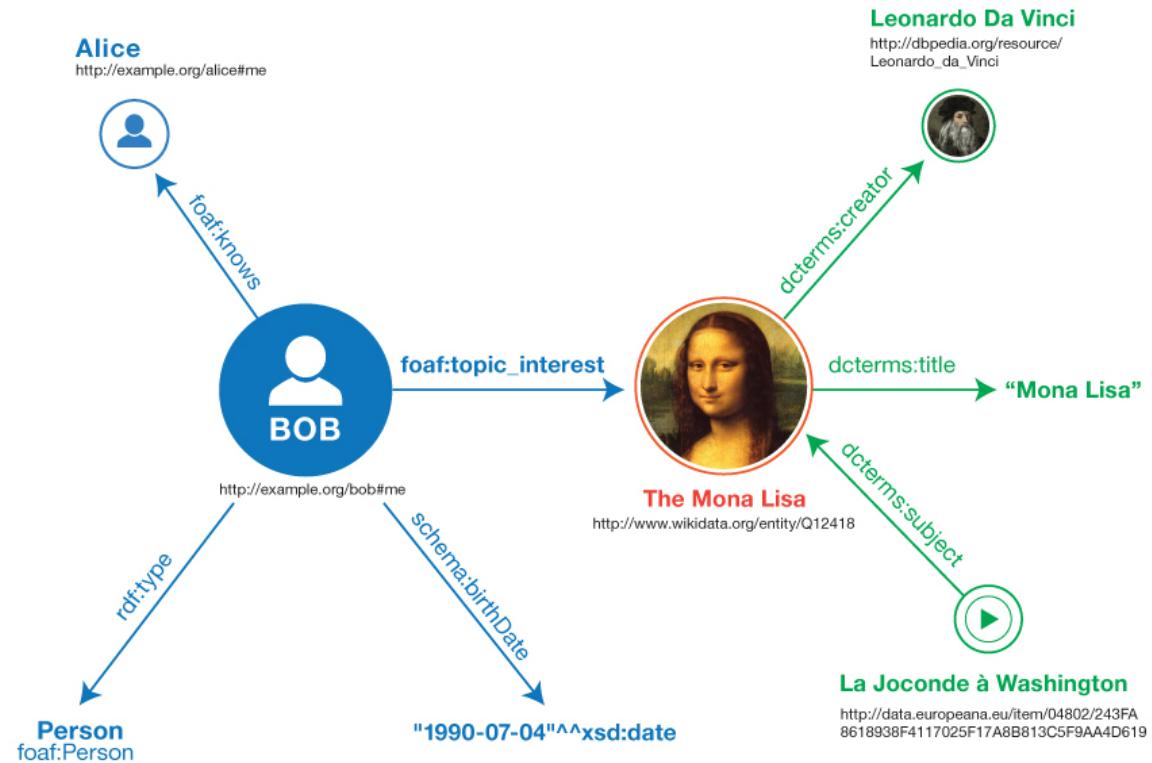
Our Knowledge Graph again (a bit more technical)



Our Knowledge Graph again (a bit more technical)

Graph consists of:

- Resources (identified via URIs)
- Literals: data values with data type (URI) or language (multilinguality integrated)
- Attributes of resources are also URI-identified (from *vocabularies*)
- Various data sources and vocabularies can be arbitrarily mixed and meshed
- URIs can be shortened with namespace prefixes; e.g. schema: → <http://schema.org/>



Our Graph – a bit more technical...



Allows one to talk about anything

Uniform Resource Identifier (URI) can be used to identify entities

<http://nextprot.org/rdf/gene/ENSG00000183486>
is a name for MX2

http://nextprot.org/rdf/entry/NX_P20592
is a name for the Interferon-induced GTP-binding protein Mx2

Allows one to express statements

An RDF statement consists of:

- **Subject**: resource identified by a URI
- **Predicate**: resource identified by a URI
- **Object**: resource or literal

N-Triples Syntax

```
<http://nextprot.org/rdf/entry/NX_P20592> —————  
<http://www.w3.org/2004/02/skos/core#exactMatch>  
<http://purl.uniprot.org/uniprot/P20592> .
```

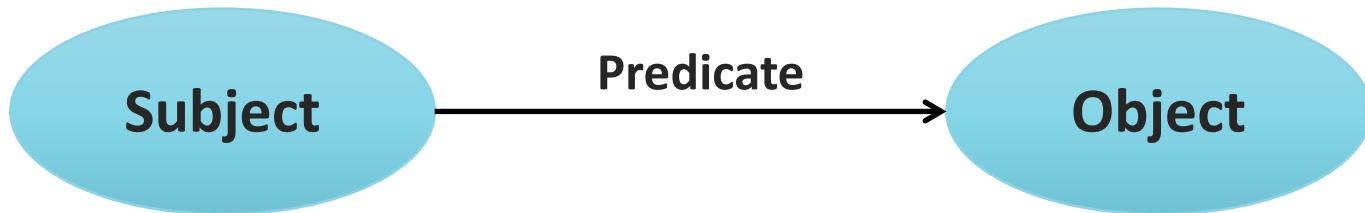
URIs are given in angle brackets in N-Triples.

```
<http://nextprot.org/rdf/entry/NX_P20592>  
<http://www.w3.org/2000/01/rdf-schema#label>  
"Interferon-induced GTP-binding protein Mx2" .
```

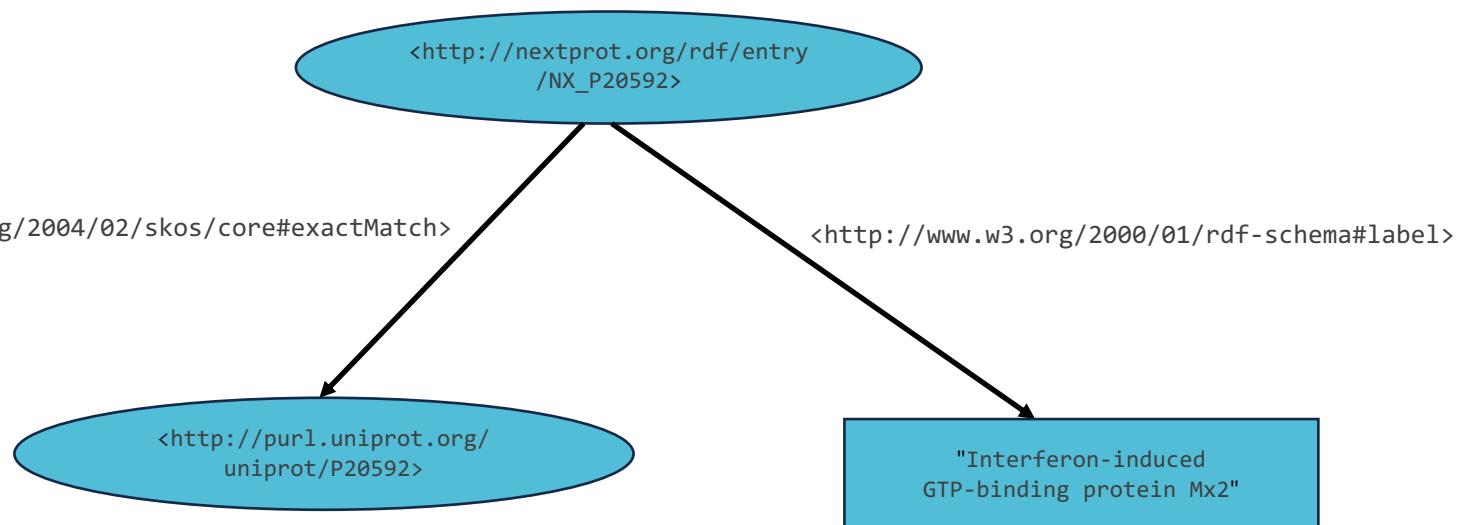
In N-Triples every statement is terminated with a full stop.

Literals are given in quotes in N-Triples.

- Every set of RDF assertions can then be drawn and manipulated as a (labelled directed) graph:
 - **Resources** – the subjects and objects are nodes of the graph.
 - **Predicates** – each predicate use becomes a label for an arc, connecting the subject to the object.



Resource Description Framework (RDF)



- Turtle is a syntax for RDF more readable.
- Since many URIs share same basis we use **prefixes**:

```
@prefix nextprot: <http://nextprot.org/rdf#> .  
@prefix owl: <http://www.w3.org/2002/07/owl#> .  
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .  
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .  
@prefix dc: <http://purl.org/dc/elements/1.1/> .  
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .  
@prefix entry: <http://nextprot.org/rdf/entry/> .  
@prefix isoform: <http://nextprot.org/rdf/isoform/> .  
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .  
@prefix uniprot: <http://purl.uniprot.org/uniprot/> .  
  
entry:NX_P20592 skos:exactMatch uniprot:P20592 .  
  
entry:NX_P20592 rdfs:label "Interferon-induced GTP-binding protein Mx2" .
```

<http://nextprot.metaphacts.cloud/resource/Prefixes>

Also has a simple *shorthand* for class membership:

```
@prefix : <http://nextprot.org/rdf#> .  
@prefix entry: <http://nextprot.org/rdf/entry/> .  
entry:NX_P20592 a :Entry.
```

Is **equivalent** to:

```
<http://nextprot.org/rdf/entry/NX_P20592>  
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>  
<http://nextprot.org/rdf#Entry> .
```

When multiple statements apply to **same subject** they can be abbreviated as follows:

```
entry:NX_P20592 a :Entry;
```

```
skos:exactMatch uniprot:P20592;
```

Same subject

```
:isoform isoform:NX_P20592-1, isoform:NX_P20592-2 .
```

Same subject & predicate

Turtle also provides a simple syntax for **datatypes** and **language tags** for literals, respectively:

```
isoform:NX_P20592-1 :canonicalIsoform "true"^^xsd:boolean .  
isoform:NX_P20592-2 :canonicalIsoform "false"^^xsd:boolean .  
  
entry:NX_P20592 rdfs:label  
"Interferon-induced GTP-binding protein Mx2"@en .  
entry:NX_P20592 rdfs:label  
"Interferon-induziertes GTP-bindendes Protein Mx2"@de .
```

From http://api.nextprot.org/entry/NX_P20592.ttl

```
entry:NX_P20592 a :Entry;
  skos:exactMatch uniprot:P20592;
  :swissprotPage unipage:P20592;
  :isoform isoform:NX_P20592-1 ;
  :isoform isoform:NX_P20592-2 ;
  .
```

```
isoform:NX_P20592-1 :swissprotDisplayed "true"^^xsd:boolean .
isoform:NX_P20592-1 :canonicalIsoform "true"^^xsd:boolean .
isoform:NX_P20592-2 :swissprotDisplayed "false"^^xsd:boolean .
isoform:NX_P20592-2 :canonicalIsoform "false"^^xsd:boolean .
```

- Collections of defined **relationships** and **classes** of resources.
 - Classes group together similar resources.
- Terms from well-known vocabularies should be **reused** wherever possible.
- New terms should be defined only if you can not find required terms in existing vocabularies.

A set of well-known vocabularies has evolved in the Semantic Web community. Some of them are:

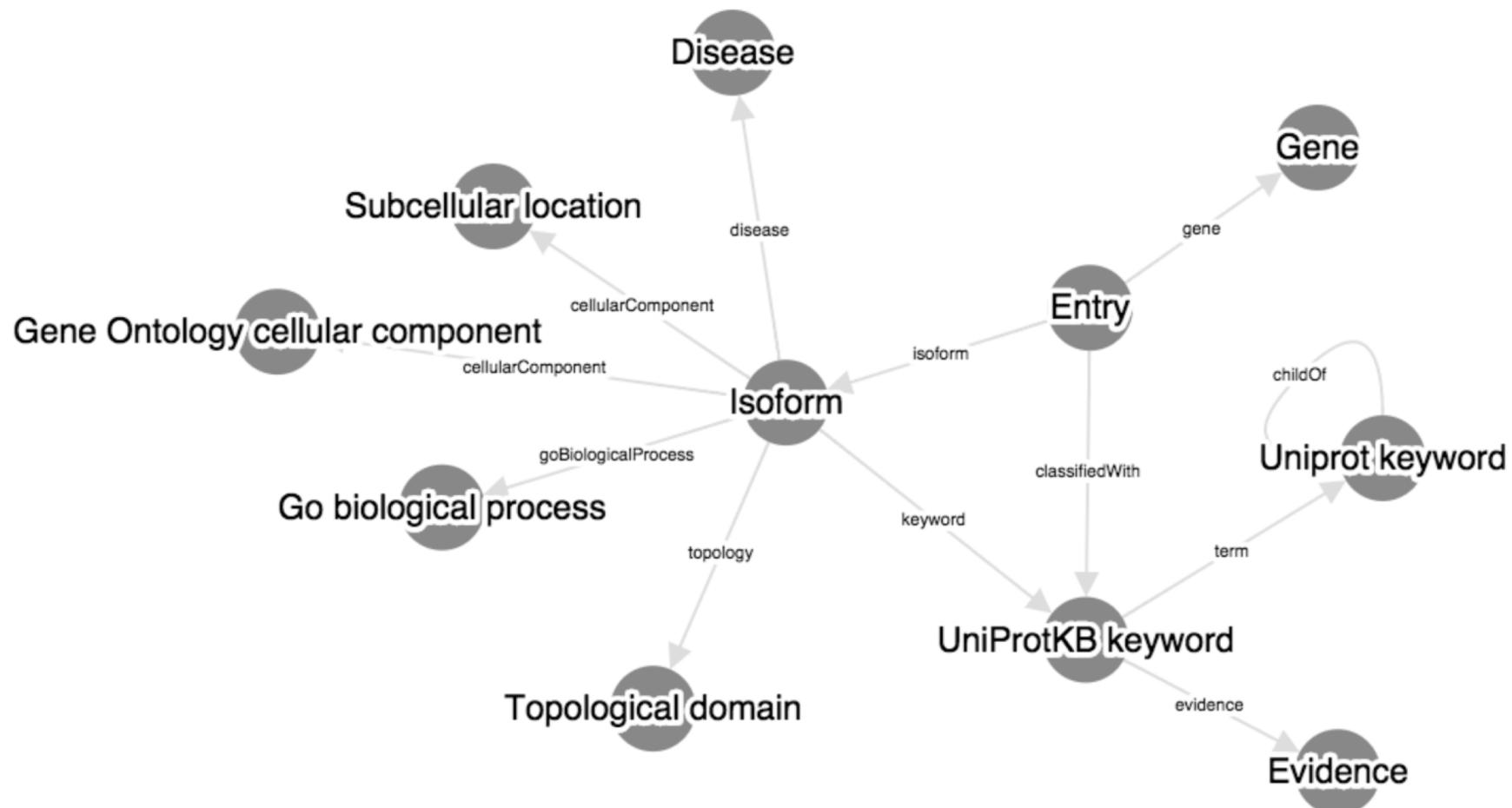
Vocabulary	Description	Classes and Relationships
Friend-of-a-Friend (FOAF)	Vocabulary for describing people.	foaf:Person, foaf:Agent, foaf:name, foaf:knows, foaf:member
Dublin Core (DC)	Defines general metadata attributes.	dc:FileFormat, dc:MediaType, dc:creator, dc:description
Organization Ontology (org)	Publishing of organization information.	org:Organisation, org:Site, org:Role, org:member, org:hasSite
schema.org	Cross-domain vocabulary for annotation of Web pages.	schema:Event, schema:Product, schema:location, schema:image

- Language for two tasks w.r.t. the RDF data model:
 - Definition of vocabulary – nominate:
 - the ‘types’, i.e., classes, of things we might make assertions about, and
 - the properties we might apply, as predicates in these assertions, to capture their relationships.
 - Inference – given a set of assertions, using these classes and properties, specify what should be inferred about assertions that are implicitly made.

- **rdfs:Class** – Example:
`:Entry – Represents the class of proteins`
- **rdfs:subClassOf** – Example:
`:CatalyticActivity rdfs:subClassOf :Function .`

- **rdf:Property** - Class of RDF properties. Example:
`:isofrom` - *Indicates the relationship with an Isoform.*
- **rdfs:domain** - States that any resource that has a given property is an instance of one or more classes.
`:isofrom rdfs:domain :Entry.`
- **rdfs:range** - States that the values of a property are instances of one or more classes.
`:isofrom rdfs:range :Isoform .`

Extract from the NextProt Schema



```
Schema :CatalyticActivity  
rdfs:subClassOf  
:Function .
```

```
Existing fact annotation:NX_000219-1-AN_000219_0291  
rdf:type  
:CatalyticActivity .
```

Inferred fact

```
annotation:NX_000219-1-AN_000219_0291  
rdf:type  
:Function .
```

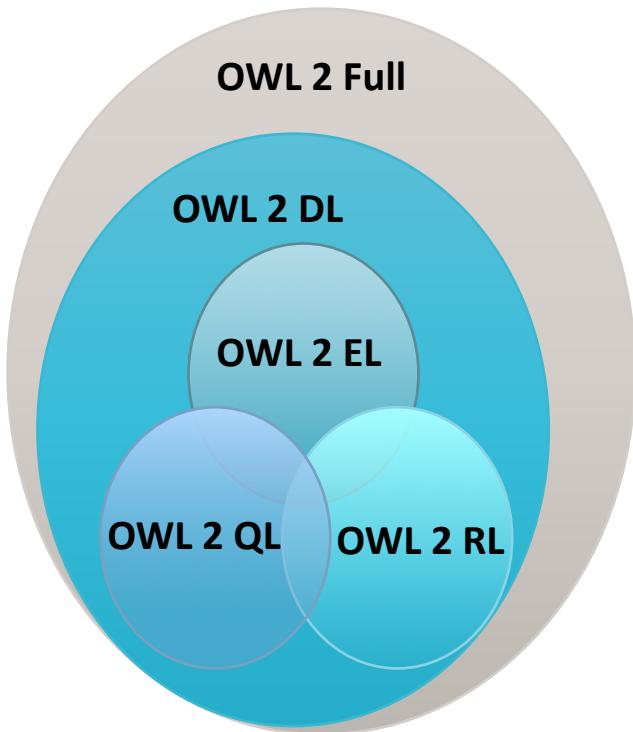
We *expect* to use this vocabulary to make assertions about catalytic activities.

Having made such an assertion...

Inferences can be drawn that we did not explicitly make

- RDFS provides a simplified ontological language for defining vocabularies about specific domains.
- OWL provides more ontological constructs for knowledge representation.
- Semantics grounded in **Description Logics**.
- Most graph databases concentrate on the use of RDFS with a subset of OWL features.

Extends the DL further, but has three more computable fragments (profiles).



OWL 2 Full

- Used informally to refer to RDF graphs considered as OWL 2 ontologies and interpreted using the RDF-Based Semantics.

OWL 2 DL

- Used informally to refer to OWL 2 DL ontologies interpreted using the Direct Semantics.

OWL 2 EL

- Limited to basic classification, but with polynomial-time reasoning.

OWL 2 QL

- Designed to be translatable to relational database querying.

OWL 2 RL

- Designed to be efficiently implementable in rule-based systems.

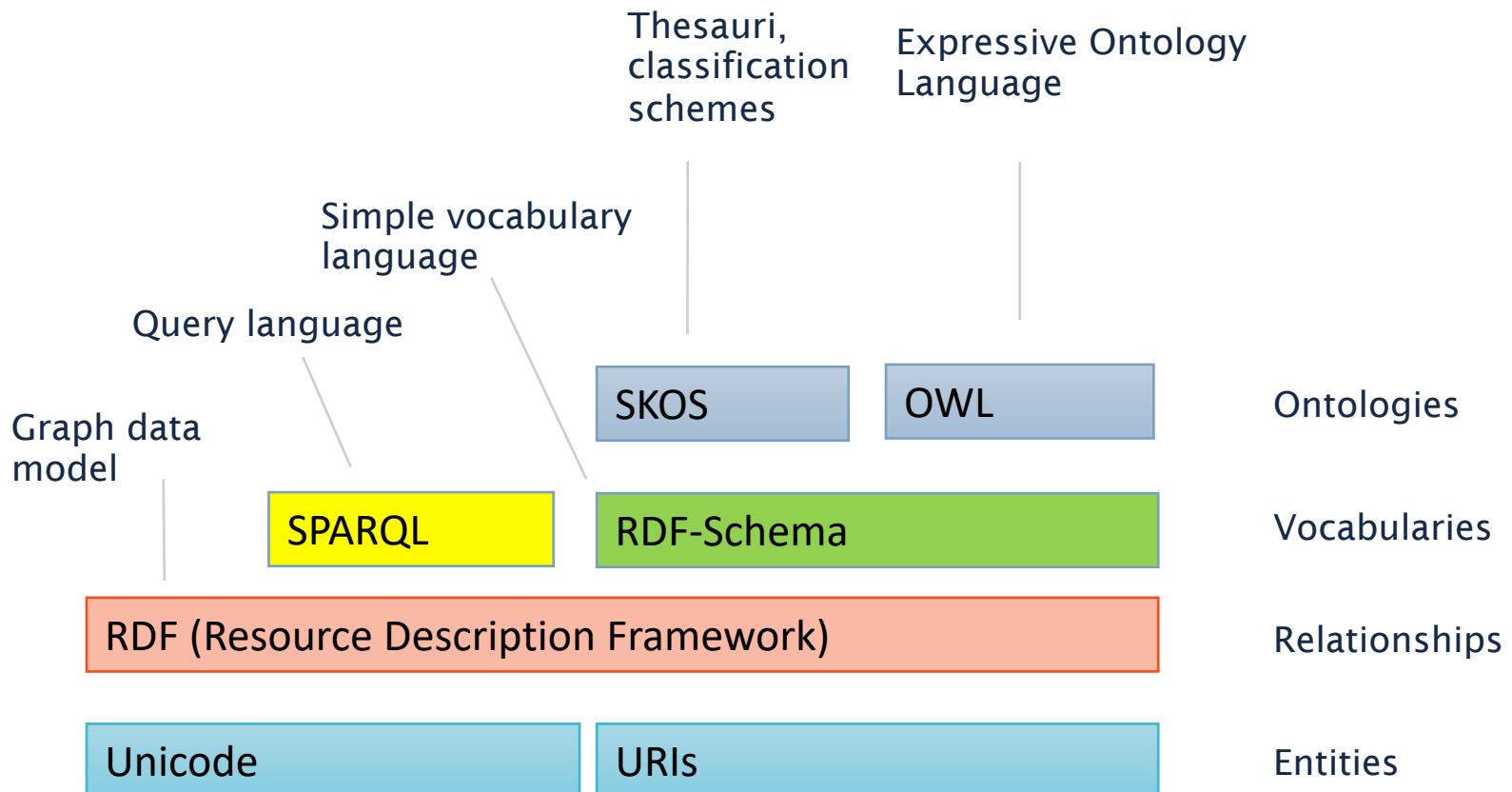
OWL is made up of **terms** which provide for:

- **Class construction:** forming new classes from membership of existing ones (e.g., unionOf, intersectionOf, etc.).
- **Property construction:** distinction between OWL ObjectProperties (resources as values) and OWL DatatypeProperties (literals as values).
- **Class axioms:** sub-class, equivalence and disjointness relationships.
- **Property axioms:** sub-property relationship, equivalence and disjointness, and relationships between properties.
- **Individual axioms:** statements about individuals (sameIndividual, differentIndividuals).

- **owl:equivalentClass** - Example:

```
:Entry owl:equivalentClass uniprot:Protein
```

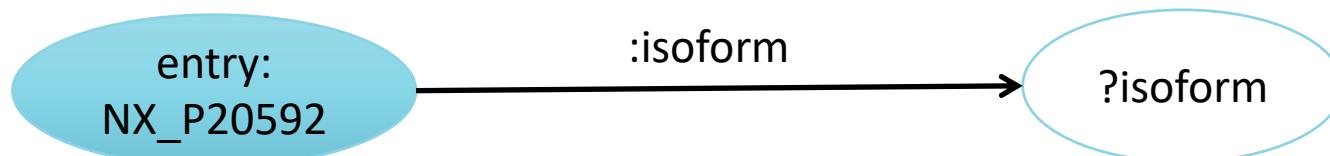
Knowledge Graphs Built on the Semantic Web Layer Cake



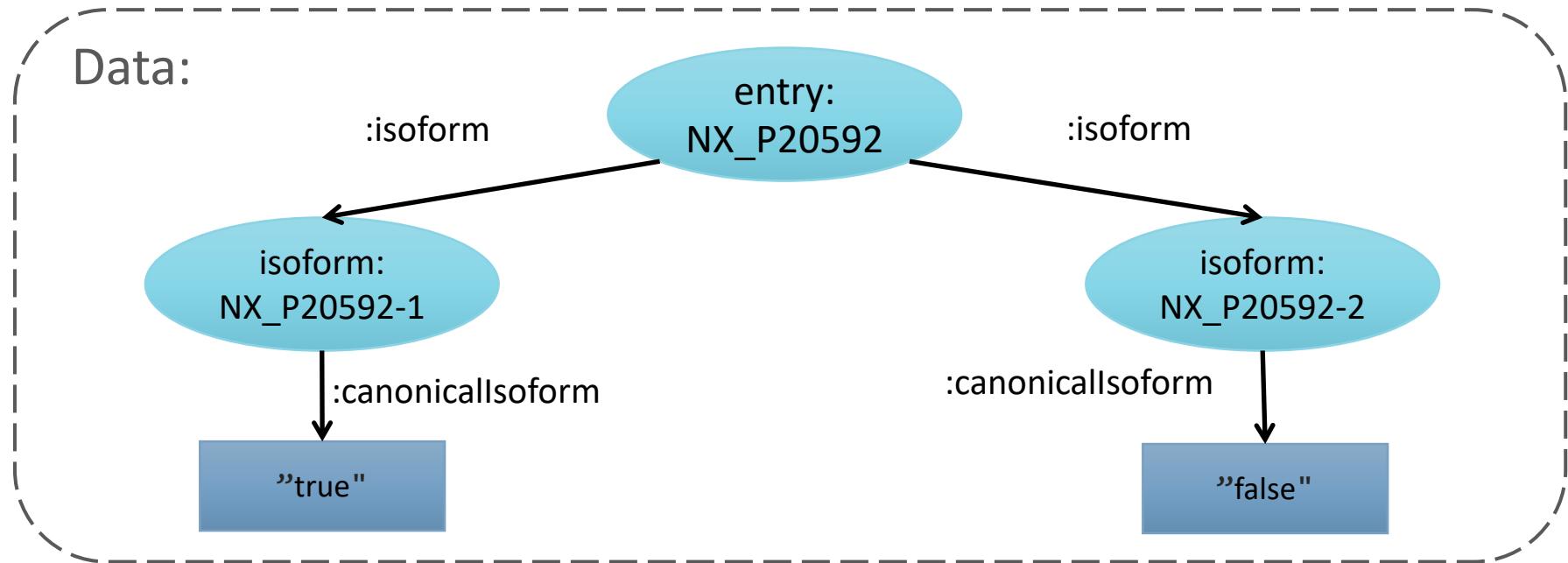
- Query language for RDF-based knowledge graphs.
- Designed to use a syntax similar to SQL for retrieving data from relational databases.
- Different query forms:
 - **SELECT** returns variables and their bindings directly.
 - **CONSTRUCT** returns a single RDF graph specified by a graph template.
 - **ASK** test whether or not a query pattern has a solution. Returns yes/no.
 - **DESCRIBE** returns a single RDF graph containing RDF data about resources.

Main idea: **Pattern matching**

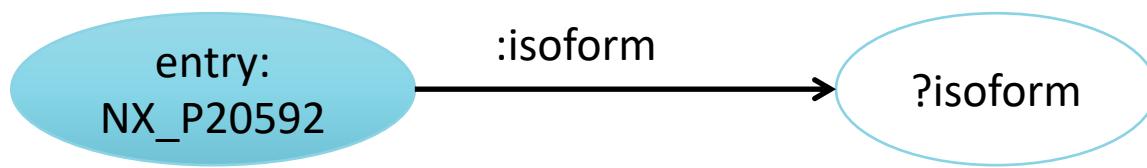
- Queries describe sub-graphs of the queried graph
- **Graph patterns** are RDF graphs specified in Turtle syntax, which contain variables (prefixed by “?”)



- Sub-graphs that match the graph patterns yield a **result**



Graph pattern:



Results:

?isoform
isoform:NX_P20592-1
isoform:NX_P20592-2

```
PREFIX : <http://nextprot.org/rdf#>
PREFIX entry: <http://nextprot.org/rdf/entry>
```

```
SELECT ?isoform ?canonical
WHERE {
    entry:NX_P20592 :isoform ?isoform .
    ?isoform :canonicalIsoform ?canonical
}
ORDER BY ?isoform
```

Prologue:

- Prefix definitions
- Subtly different from Turtle syntax - the final period is not used

```
PREFIX : <http://nextprot.org/rdf#>
PREFIX entry: <http://nextprot.org/rdf/entry>
```

```
SELECT ?isoform ?canonical
WHERE {
    entry:NX_P20592 :isoform ?isoform .
    ?isoform :canonicalIsoform ?canonical
}
ORDER BY ?isoform
```

Query form:

- ASK, SELECT, DESCRIBE or CONSTRUCT
- SELECT retrieves variables and their bindings as a table

```
PREFIX : <http://nextprot.org/rdf#>
PREFIX entry: <http://nextprot.org/rdf/entry>
```

```
SELECT ?isoform ?canonical
WHERE {
    entry:NX_P20592 :isoform ?isoform .
    ?isoform :canonicalIsoform ?canonical
}
ORDER BY ?isoform
```

Query pattern:

- Defines patterns to match against the data
- Generalises Turtle with variables and keywords – N.B. final period optional

```
PREFIX : <http://nextprot.org/rdf#>
PREFIX entry: <http://nextprot.org/rdf/entry/>
```

```
SELECT ?isoform ?canonical
WHERE {
    entry:NX_P20592 :isoform ?isoform .
    ?isoform :canonicalIsoform ?canonical
}
ORDER BY ?isoform
```

Solution modifier:

- Modify the result set
- ORDER BY, LIMIT or OFFSET re-organise rows;
- GROUP BY combines them

SPARQL supports different query forms:

- **ASK** tests whether or not a query pattern has a solution.
Returns yes/no
- **SELECT** returns variables and their bindings directly
- **CONSTRUCT** returns a single RDF graph specified by a graph template
- **DESCRIBE** returns a single RDF graph containing RDF data about resource

- Namespaces are added with the ‘PREFIX’ directive
- Statement patterns that make up the graph are specified between brackets (“{}”)

Query: Is isoform:NX_P20592-1 canonical?

```
PREFIX : <http://nextprot.org/rdf#>
PREFIX isoform: <http://nextprot.org/rdf/isoform>
ASK WHERE { isoform:NX_P20592-1 :canonicalIsoform
  "true"^^xsd:boolean }
```

Results:

true

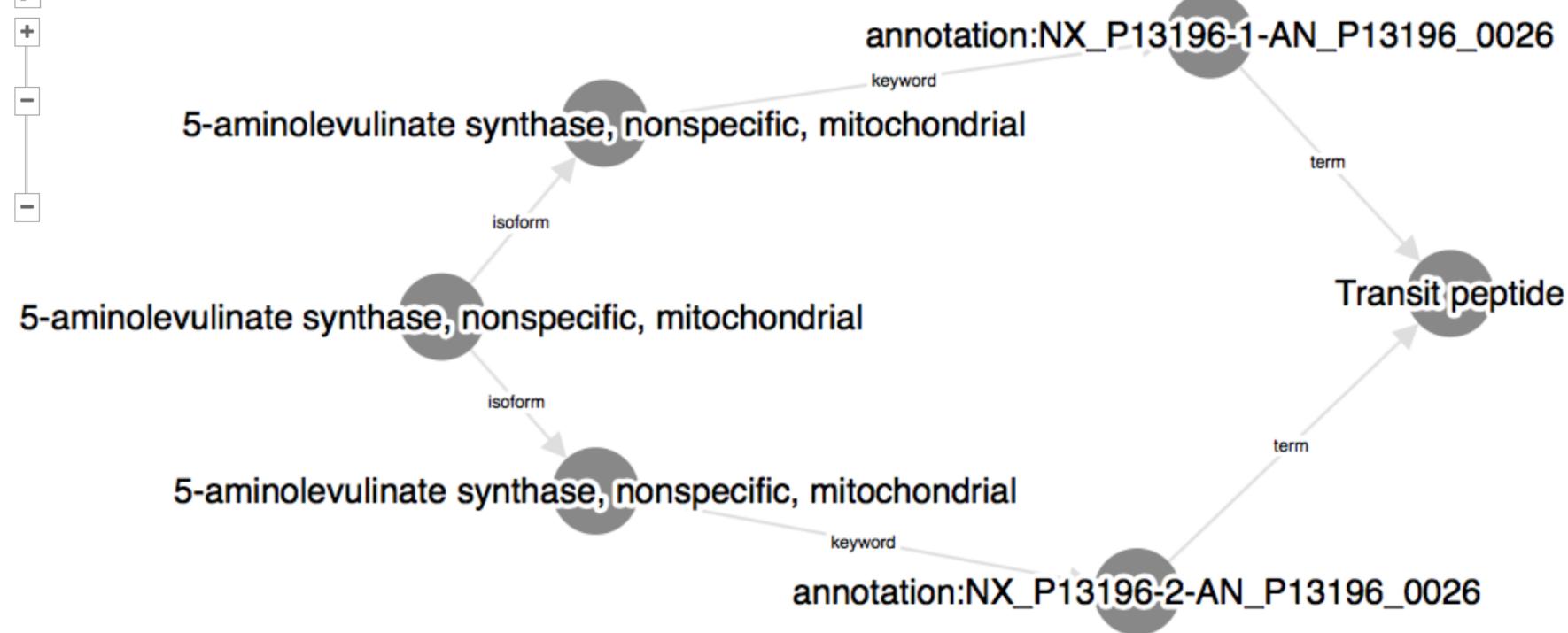
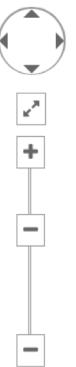
Query: Is isoform:NX_P20592-2 canonical?

```
PREFIX : <http://nextprot.org/rdf#>
PREFIX isoform: <http://nextprot.org/rdf/isoform>
ASK WHERE { isoform:NX_P20592-2 :canonicalIsoform
  "true"^^xsd:boolean }
```

Results:

false

Sample Path in the Data



- The solution modifier **projection** nominates which components of the matches should be returned
- “*” means all components should be returned

Query: *Proteins that have a Transit peptide?*

```
PREFIX : <http://nextprot.org/rdf#>
PREFIX cv: <http://nextprot.org/rdf/terminology/>
SELECT ?entry where {
    ?entry :isoform ?isoform .
    ?isoform :keyword ?keyword .
    ?keyword :term cv:KW-0809
}
```

- A property path is a possible route through a graph between two graph nodes
- Property paths allow for more concise expression of some SPARQL basic graph patterns and also add the ability to match arbitrary length paths

Query: *Proteins that have a Transit peptide?*

```
PREFIX : <http://nextprot.org/rdf#>
PREFIX cv: <http://nextprot.org/rdf/terminology/>
SELECT ?entry WHERE{
    ?entry :isoform / :keyword / :term cv:KW-0809
}
```

Filter expressions

- Different types of filters and functions may be used

Filter: Comparison and logical operators

Query: *Proteins that do not have a Transit peptide?*

```
PREFIX : <http://nextprot.org/rdf#>
PREFIX cv: <http://nextprot.org/rdf/terminology/>
```

```
SELECT ?entry WHERE {
  ?entry :isoform ?iso.
```

```
    FILTER NOT EXISTS { ?iso :keyword/:term cv:KW-0809 }
}
```

Elimination of duplicates

Query: *Distinct proteins that have a Transit peptide?*

```
PREFIX : <http://nextprot.org/rdf#>
PREFIX cv: <http://nextprot.org/rdf/terminology/>
```

```
SELECT DISTINCT ?entry WHERE{
  ?entry :isoform / :keyword / :term cv:KW-0809
}
```

Union Graph Pattern

- Allows the specification of alternatives (disjunctions)

Query: *Proteins that have a Transit peptide or Translation regulation?*

```
PREFIX : <http://nextprot.org/rdf#>
PREFIX cv: <http://nextprot.org/rdf/terminology/>

SELECT DISTINCT ?entry WHERE {

  { ?entry :isoform / :keyword / :term cv:KW-0809 }
  UNION
  { ?entry :isoform / :keyword / :term cv:KW-0810 }

}
```

Aggregates

- Calculate aggregate values: COUNT, SUM, MIN, MAX, AVG, GROUP_CONCAT and SAMPLE
- Built around the GROUP BY operator
- Prune at group level (cf. FILTER) using HAVING

Query: *Proteins with 7 transmembrane regions? (cf. NXQ_00003)*

```
SELECT DISTINCT ?entry WHERE {  
  ?entry :isoform ?iso.  
  ?iso :topology ?statement.  
  ?statement a :TransmembraneRegion.  
}  
GROUP BY ?entry ?iso  
HAVING (COUNT (?statement)=7)
```

Subsets of results

- It is possible to combine the query with **solution modifiers** (ORDER BY, LIMIT, OFFSET)

Query: *Just 10 proteins, please.*

```
prefix : <http://nextprot.org/rdf#>

SELECT *
WHERE {
  ?entry a :Entry
}
LIMIT 10
```

Takes the resources within the solution, and provides information about them as RDF statements. They can be identified by:

- **Specifying explicit IRIs**

```
prefix entry: <http://nextprot.org/rdf/entry/>
DESCRIBE entry:NX_P20592
```

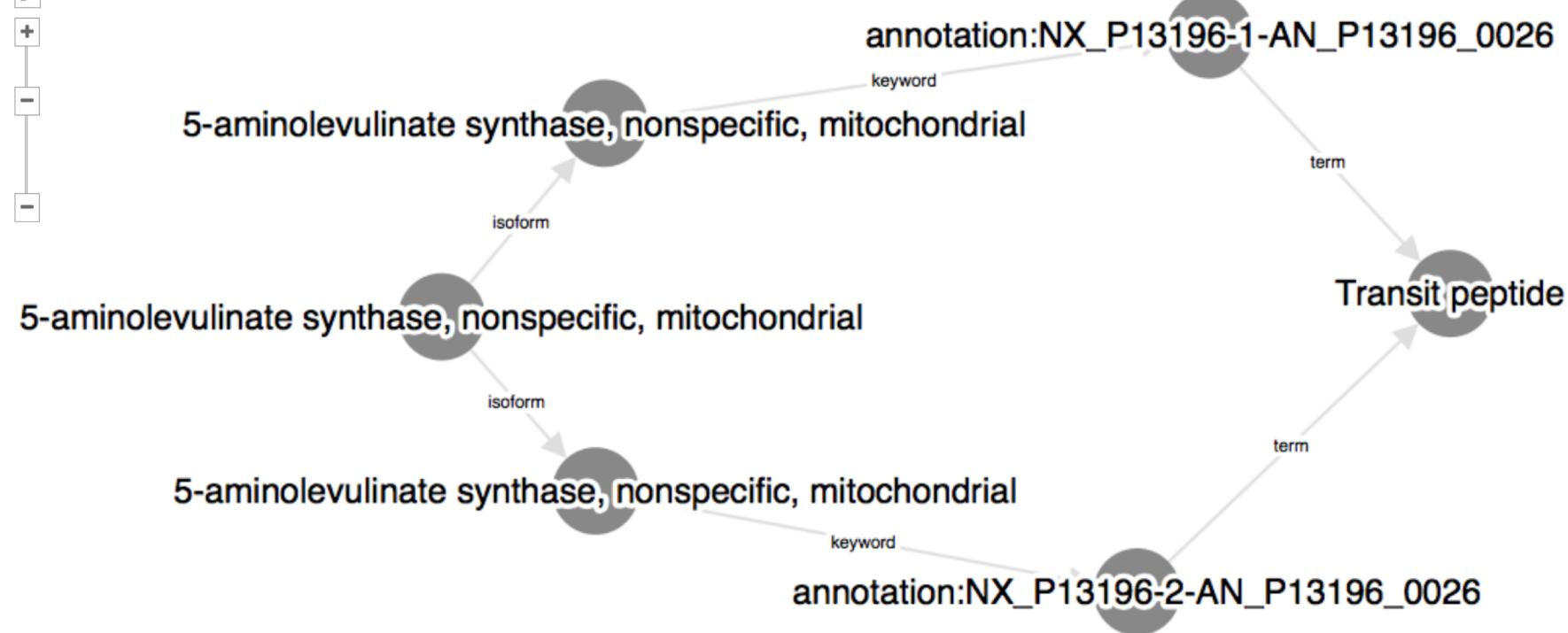
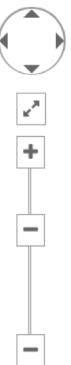
- **Bindings of variables in the WHERE clause**

```
prefix entry: <http://nextprot.org/rdf/entry/>
DESCRIBE ?isoform
WHERE {
  entry:NX_P20592 :isoform ?isoform
}
```

- **CONSTRUCT WHERE:** In order to query for a subgraph, without change, it is no longer necessary to repeat the graph pattern in the template

Example:

```
PREFIX : <http://nextprot.org/rdf#>
PREFIX cv: <http://nextprot.org/rdf/terminology/>
CONSTRUCT WHERE {
    entry:NX_P13196 :isoform ?isoform .
    ?isoform :keyword ?keyword .
    ?keyword :term cv:KW-0809
}
```



- Returns RDF statements created from variable bindings
- Template: graph pattern with variables from the query pattern

Query: *Relationship between proteins and terms*

```
prefix : <http://nextprot.org/rdf#>
prefix entry: <http://nextprot.org/rdf/entry/>

CONSTRUCT { entry:NX_P13196 :FRkeyword ?term }
WHERE {
  entry:NX_P13196 :isoform / :keyword / :term ?term
}
```

<https://nextprot.metaphacts.cloud/resource/training:Q4>

Type of function	Function	Result type
Functional Forms	bound IF COALESCE NOT EXISTS, EXISTS or, and RDFTerm-equal (=), sameTerm IN, NOT IN	xsd:boolean rdfTerm rdfTerm xsd:boolean xsd:boolean xsd:boolean boolean
Functions on RDF Terms	isIRI, isBlank, isLiteral, isNumeric str, lang, datatype IRI BNODE	xsd:boolean simple literal iri iri blank node
Functions on Numerics	ABS, ROUND, CEIL, FLOOR RAND	numeric xsd:double

Type of function	Function	Result type
Functions on Strings	STRLEN SUBSTR, UCASE, LCASE STRSTARTS, STRENDs, CONTAINS STRBEFORE, STRAFTER ENCODE_FOR_URI CONCAT langMatches REGEX REPLACE	xsd:integer string literal xsd:boolean literal simple literal string literal xsd:boolean xsd:boolean string literal
Functions on Dates and Times	now year, month, day, hours, minutes seconds timezone tz	xsd:dateTime xsd:integer xsd:decimal xsd:dayTimeDuration simple literal

- Write a SPARQL query that ...
 1. Returns isoforms of Hyaluronan synthase 3
 2. Finds proteins phosphorylated and located in the cell membrane
 3. Returns proteins with more than 2 intramembrane regions
 4. Returns proteins located in mitochondrion and are associated with a disease
 5. Returns proteins associated with heart diseases

SPARQL 1.1 provides data update operations:

- **INSERT data:** adds some triples, given inline in the request, into a graph
- **DELETE data:** removes some triples, given inline in the request, if the respective graphs contains those
- **DELETE/INSERT data:** uses in parallel INSERT and DELETE

```
BASE <http://example.org/>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

WITH <http://www.metaphacts.com/tutorial/>
INSERT {
    <peter#me>      a foaf:Person ;
                    rdfs:label "Peter";
                    foaf:knows <alice#me> , <bob#me> ;
                    foaf:topic_interest wd:Q12418
}
WHERE {}
```

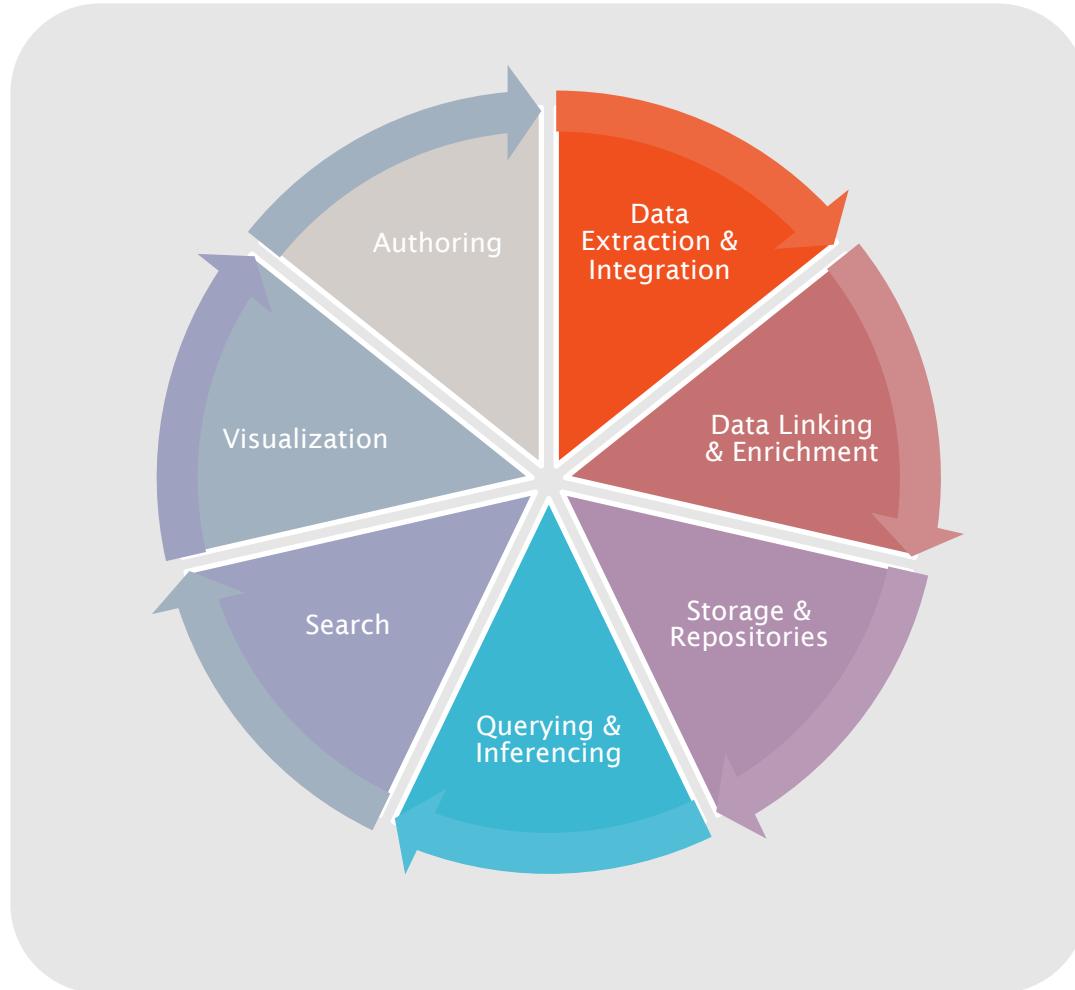


metaphacts

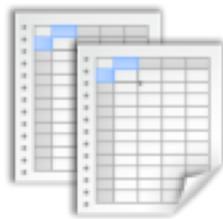
Smart Data. Smart Apps. Smart Decisions.

Knowledge Graph Management

Knowledge Graph Lifecycle: Data Extraction



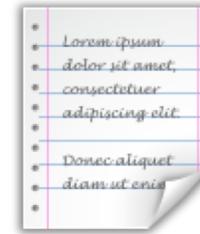
- The data of interest may be stored in a wide range of formats:



Spreadsheets
or tabular data



Databases

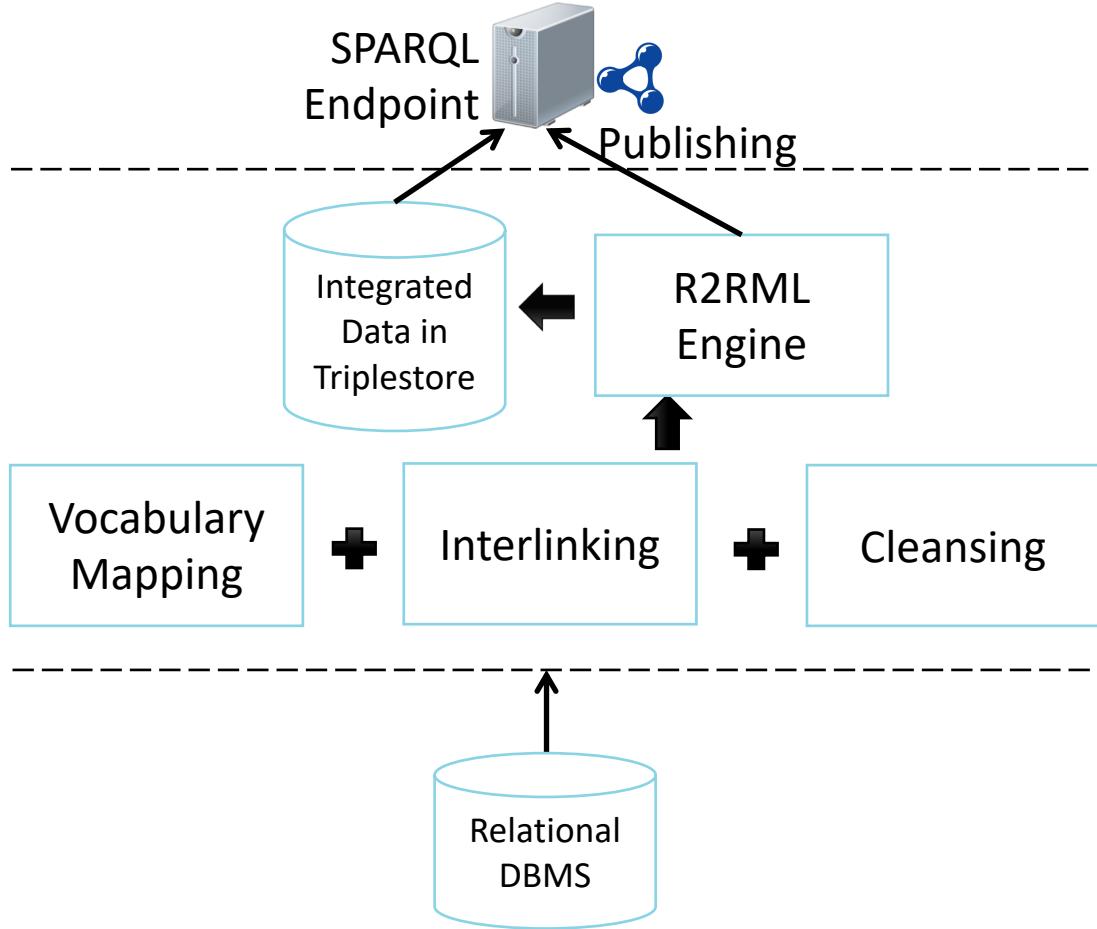


Text

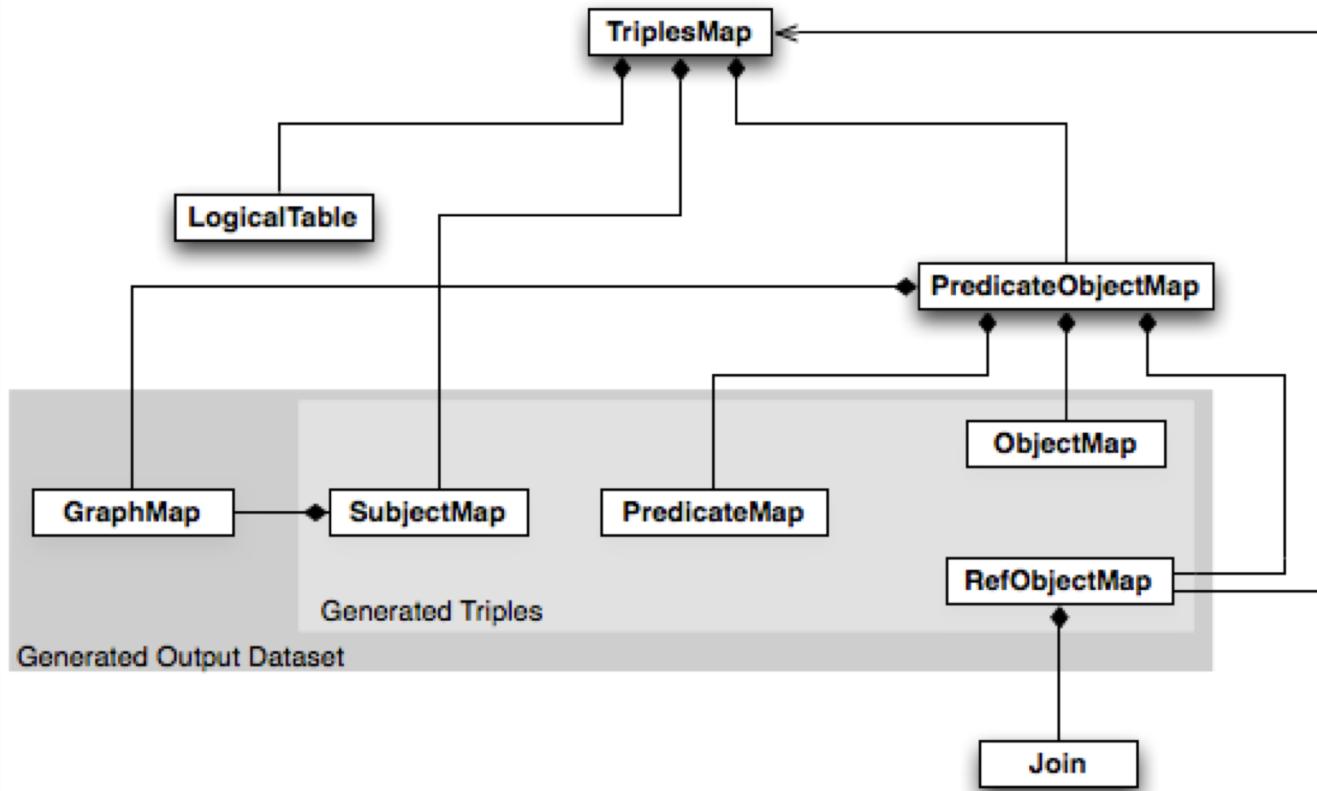
- Several tools support the process of mining data from different repositories, for example:



Access
LD Data set
Data acquisition

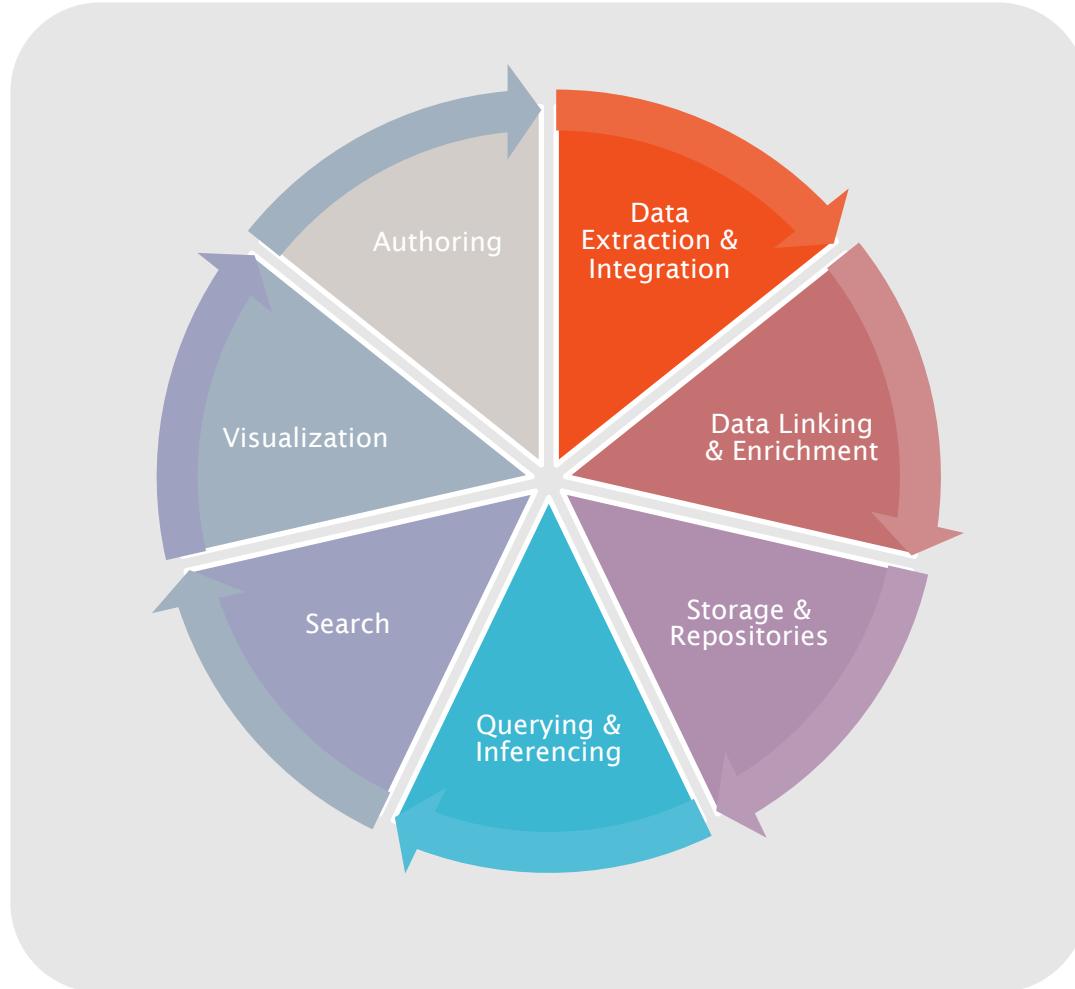


- **Task:** Integrate data from relational DBMS with Linked Data
- **Approach:** map from relational schema to semantic vocabulary with R2RML
- **Publishing:** two alternatives –
 - Translate SPARQL into SQL on the fly
 - Batch transform data into RDF, index and provide SPARQL access in a triplestore



R2RML: RDB to RDF Mapping Language, W3C Recommendation 27 September 2012
 Editors: Souripriya Das, Seema Sundara, Richard Cyganiak
<http://www.w3.org/TR/r2rml/>

Knowledge Graph Lifecycle: Search



Using semantic models for the search process

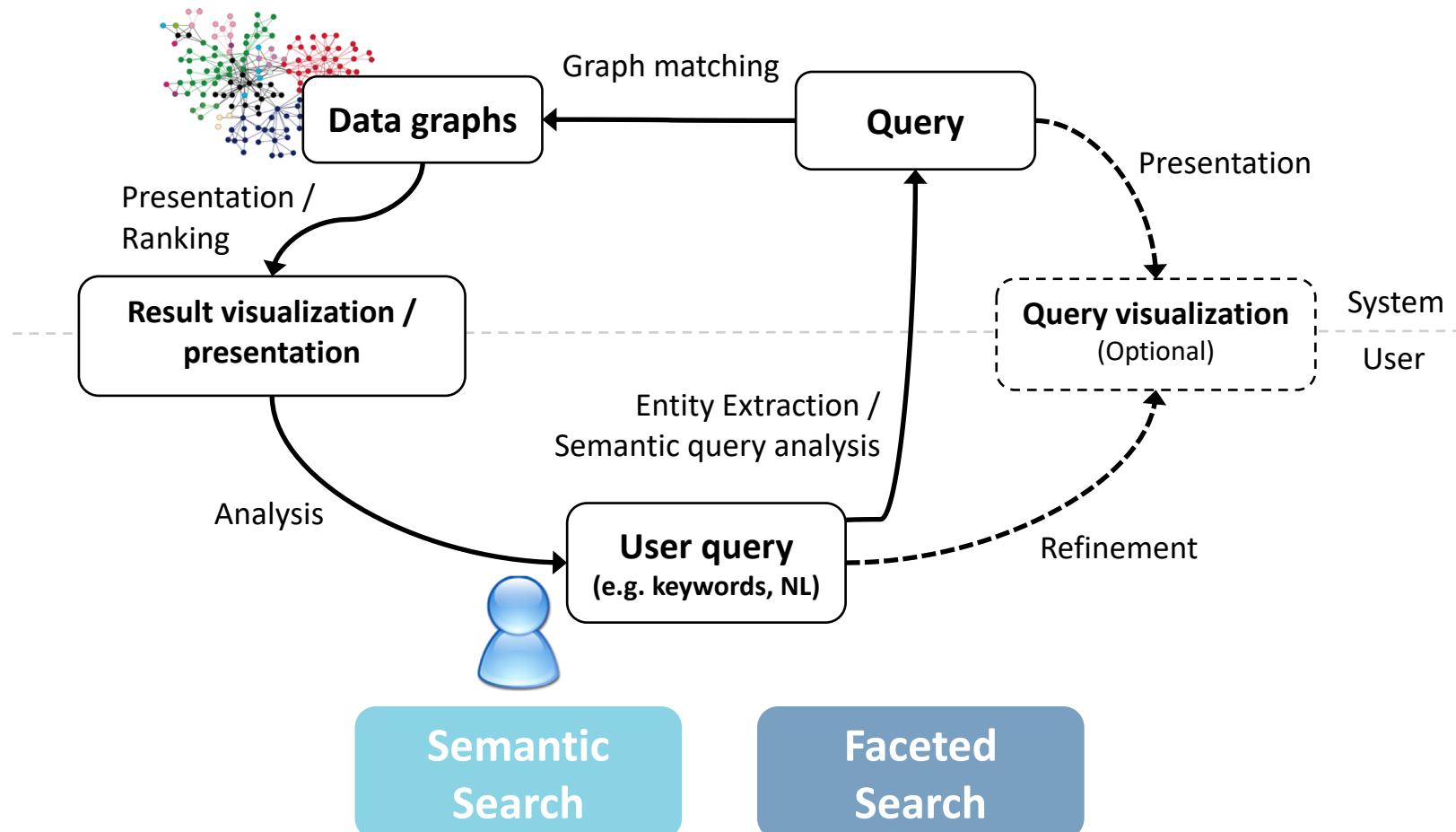


Image based on: Tran, T., Herzig, D., Ladwig, G. SemSearchPro- Using semantics through the search process

- Aims at understanding the **meaning** of the resources specified in the query
- Different approaches to exploit semantics:

- **Query expansion using ontologies**

Since ontologies represent knowledge about specific domains, they can be used to expand the query by incorporating related ontology terms into the query.

- **Contextual analysis**

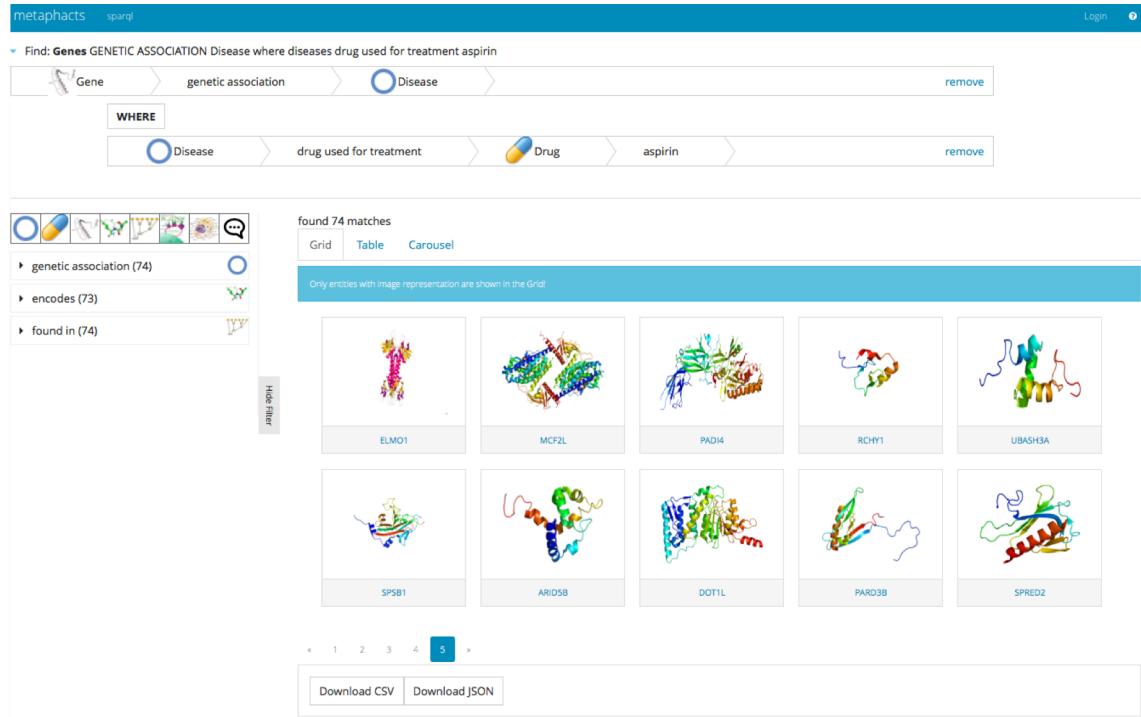
In LD, this approach may explore the resources specified in the query and their adjacent nodes in the RDF graph. Mainly applied to disambiguate query terms.

- **Reasoning**

In some cases, the answer to a specific query is not explicitly contained in the data, but it can be computed by using reasoning methods.

Example: Structured Search Generation

- Domain independent, fully customizable search widget
- Satisfy complex information needs without learning SPARQL
- Search functionalities
 - Graphical query construction
 - End user friendly search interfaces for building and sharing complex queries
 - Semantic auto suggestion
 - Interactive result visualization
 - Faceted search and exploration of item collections
- Ability to invoke external full text search indices such as Solr including the possibility to score, rank and limit the results for responsive autosuggestion
- Saving and sharing of queries and search results



The screenshot shows the metaphacts search interface. At the top, there is a search bar with the query: "Find: Genes GENETIC ASSOCIATION Disease where diseases drug used for treatment aspirin". Below the search bar, the query is visualized as a graph: Gene → genetic association → Disease. This is followed by a WHERE clause: Disease → drug used for treatment → Drug → aspirin. There are "remove" buttons for each step in the WHERE clause.

On the left, there is a sidebar with filters:

- genetic association (74)
- encodes (73)
- found in (74)

Below the filters, it says "found 74 matches" and provides options to switch between Grid, Table, and Carousel. A note states: "Only entities with image representation are shown in the Grid".

The main area displays a grid of 14 results, each with a small image and the gene name below it:

ELMO1	MCF2L	PAD14	RCHY1	UBASH3A
SPSB1	ARID5B	DOT1L	PARD3B	SPRED2

At the bottom, there are navigation arrows (left, right), a page number (5), and download buttons for "Download CSV" and "Download JSON".

GraphScope combines the **simplicity of keyword search** with the **precision of database queries**



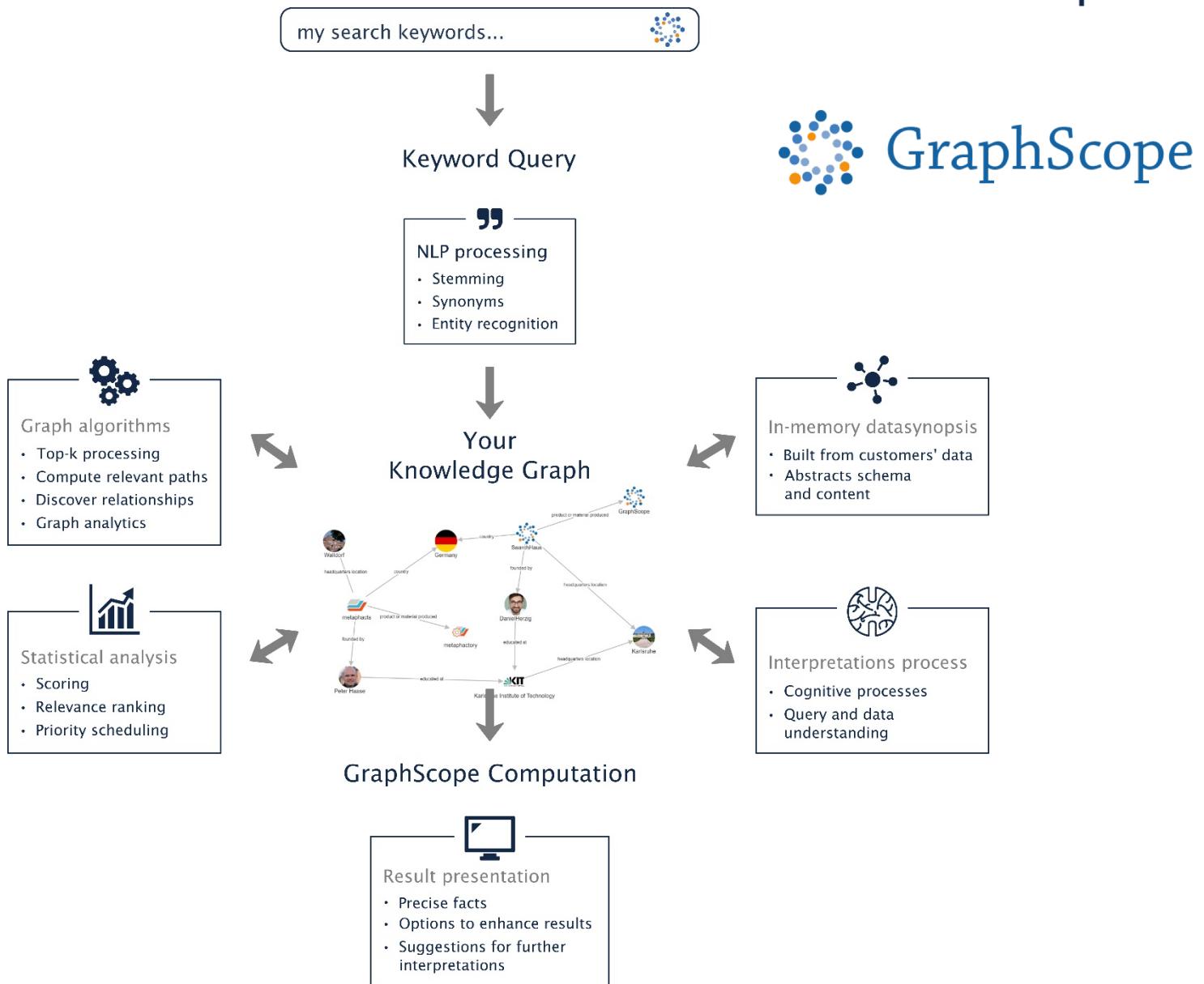
 GraphScope

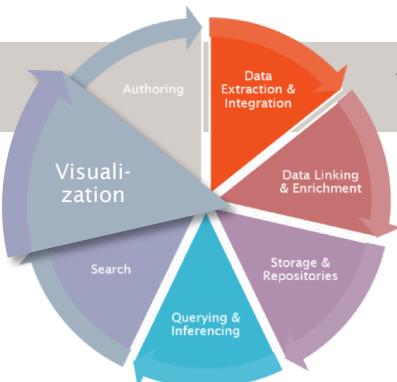
- **Just keywords, no more SPARQL**
 - Users don't need to be SPARQL experts to leverage the power of graph data. Just searching by keywords is enough to take advantage of the relationships in their data and find relevant data.
- **Simple interface**
 - A search box, known to every user, is the interface to start a search. GraphScope understands users' requests and provides feedback on how their search was interpreted. Users can alter their request easily.
- **Engages everybody**
 - GraphScope enables everybody to search the data graph, not just IT-experts. More insights to the knowledge captured by your data will be gained.
- **Exploration**
 - GraphScope allows you to discover data related to your search result and explore relevant data.



<https://wikidata.metaphacts.com>

GraphScope Technology



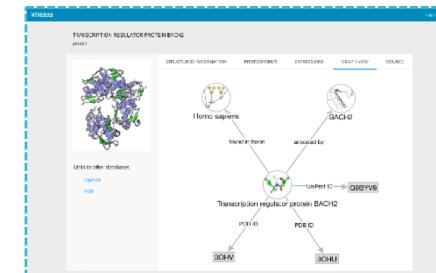


Visualization using metaphactory



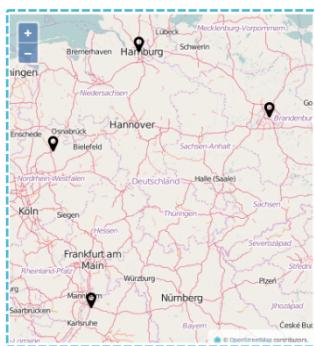
Carousel

Animated browsing through a list of result items



Graph

Visualize and explore connections in a graph view
Custom styling of the graph
Variety of graph layouts



Map

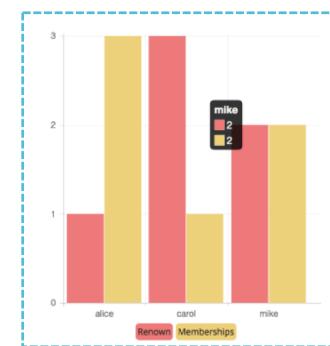
Displaying spatial data on a geographic map

Registered Events

- ▼ Hazards
 - Environmental_Hazards
 - Human_Induced_Hazards
 - Communication
 - Infrastructure
 - Transportation
 - War
 - Natural_Hazards
 - Technological_Hazards

Tree Table

Tree-based visualization, navigation and browsing through sub-tree structures



Chart

Visualize trends and relationships between numbers, ratios, or proportions

Filter Results		
Subject	Predicate	Object
alice	foafknows	bob
alice	foafknows	carol
carol	foafknows	mike
mike	foafknows	carol
bob	foafknows	carol

Table

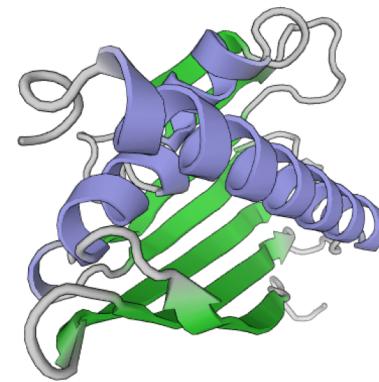
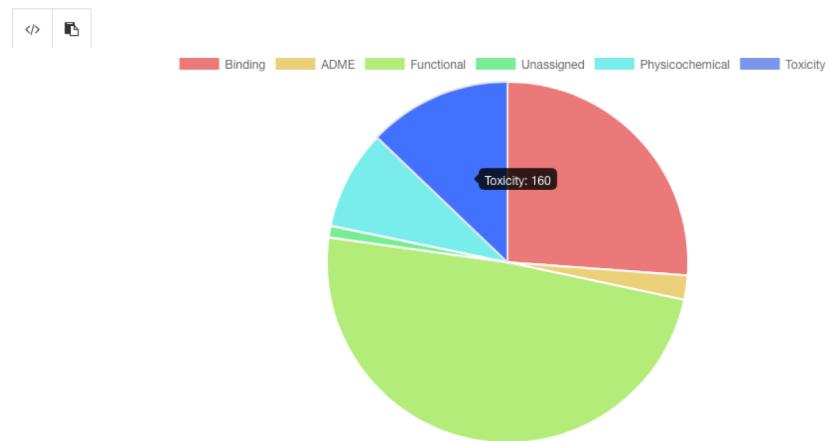
Transform your queries into durable, interactive tables
Many customization possibilities, e.g. pagination, filters and cell templates

Visualization Components in metaphactory

activity	assay	type	target
CHEMBL_ACT_7614455	CHEMBL1909178	Binding	Nitric-oxide synthase, brain
CHEMBL_ACT_7614456	CHEMBL1909178	Binding	Nitric-oxide synthase, brain
CHEMBL_ACT_7614459	CHEMBL1909180	Binding	Delta opioid receptor
CHEMBL_ACT_7614460	CHEMBL1909180	Binding	Delta opioid receptor
CHEMBL_ACT_7614461	CHEMBL1909181	Binding	Kappa opioid receptor
CHEMBL_ACT_7614463	CHEMBL1909182	Binding	Mu opioid receptor
CHEMBL_ACT_7614464	CHEMBL1909182	Binding	Mu opioid receptor
CHEMBL_ACT_7615912	CHEMBL1909118	Binding	Vascular endothelial growth factor receptor 1
CHEMBL_ACT_7615913	CHEMBL1909118	Binding	Vascular endothelial growth factor receptor 1
CHEMBL_ACT_7615916	CHEMBL1909120	Binding	Vasopressin V1a receptor

« 1 2 3 4 5 6 7 8 9 10 11 »

Simple pie chart: Assay types for Vitamin C



Tutorial - Getting Started

Taking the first steps with the metaphacts platform? Take this introductory tutorial to learn:

- How to load data into the system
- How to query the data using SPARQL
- How to visualize the data with rich HTML5 components
- How to create a simple search interface
- How to create simple templates to display data for a certain type
- How to update and insert new data

Prerequisites:

This tutorial assumes that you have installed your own instance of the metaphacts platform. If this is not the case, read the [installation manual](#).



DATA LOADING & QUERYING

[Loading data](#)
[Querying data](#)



VISUALIZATION

[Visualizing results in a table](#)
[Visualizing results in a graph](#)



SEARCH

[Creating a simple search interface](#)



AUTHORING

[Creating a template](#)
[Inserting and updating data](#)

<https://help.metaphacts.com/resource/Help:Tutorial>
<http://knowledgegraph.info>

- Some slides build on the Euclid training material



<http://www.euclid-project.eu>



- Thanks to the team at metaphacts

Get in Touch!



metaphacts GmbH

Daimlerstraße 36

69190 Walldorf

Germany

p +49 6227 6989965

m +49 157 50152441

e info@metaphacts.com

 [@metaphacts](https://twitter.com/metaphacts)

www.metaphacts.com