

KNOWLEDGE GRAPH COMPLETION

PART 1: INTRODUCTION

FATIHA SAÏS ⁽¹⁾

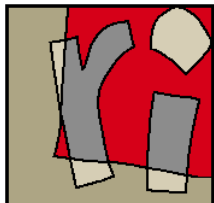
NATHALIE PERNELLE⁽¹⁾

DANAI SYMEONIDOU⁽²⁾



⁽¹⁾ LRI, PARIS SUD UNIVERSITY, CNRS, PARIS SACLAY UNIVERSITY

⁽²⁾ INRA, GAMMA TEAM

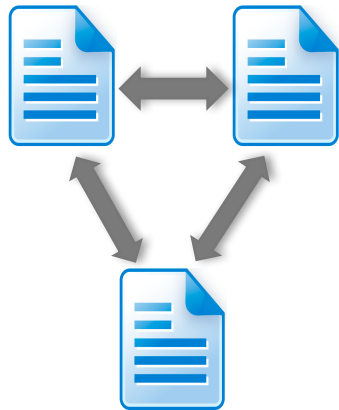


Comprendre le monde,
construire l'avenir®



FROM THE WWW TO THE WEB OF DATA

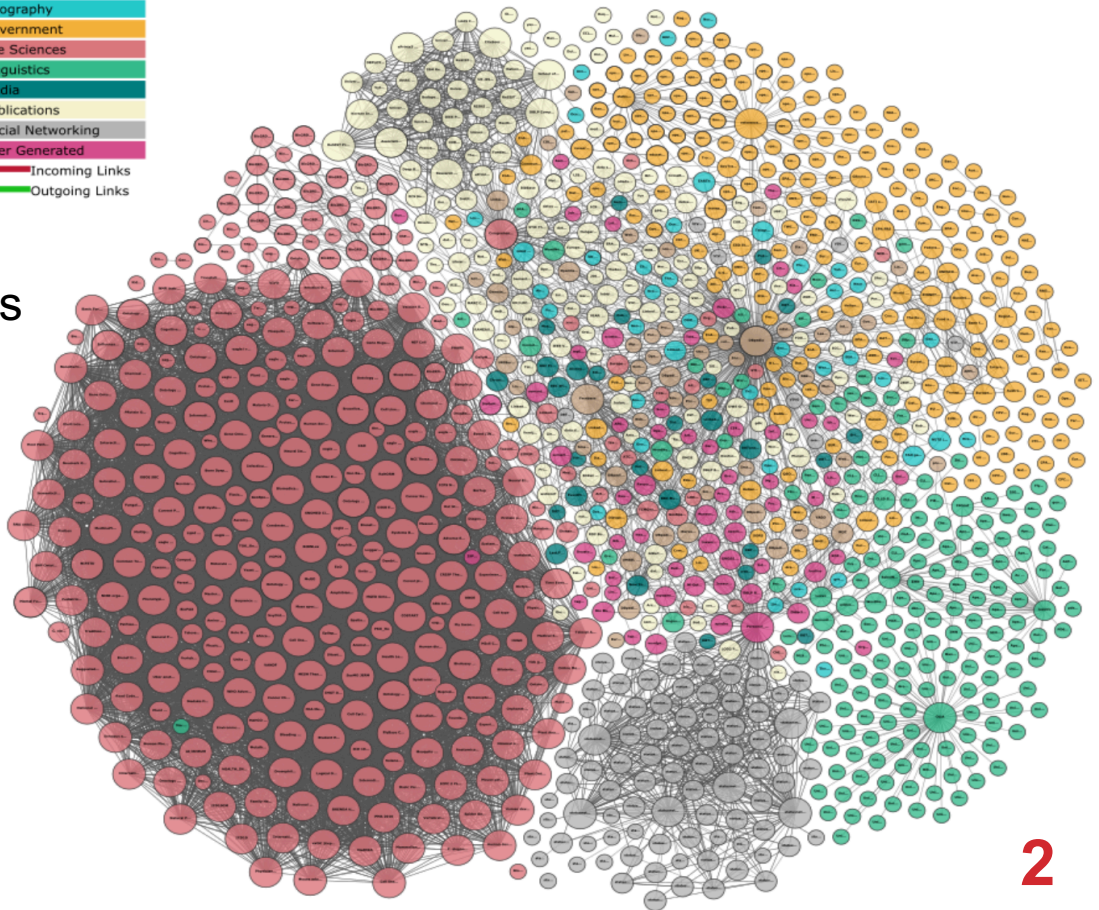
- applying the principles of the WWW to data



data is links,
not only properties



Linked Open Data



LINKED DATA PRINCIPLES

① Use HTTP URIs as identifiers for resources

→ so people can look up the data

② Provide data at the location of URIs

→ to provide data for interested parties

③ Include links to other resources

→ so people can discover more information

→ bridging disciplines and domains

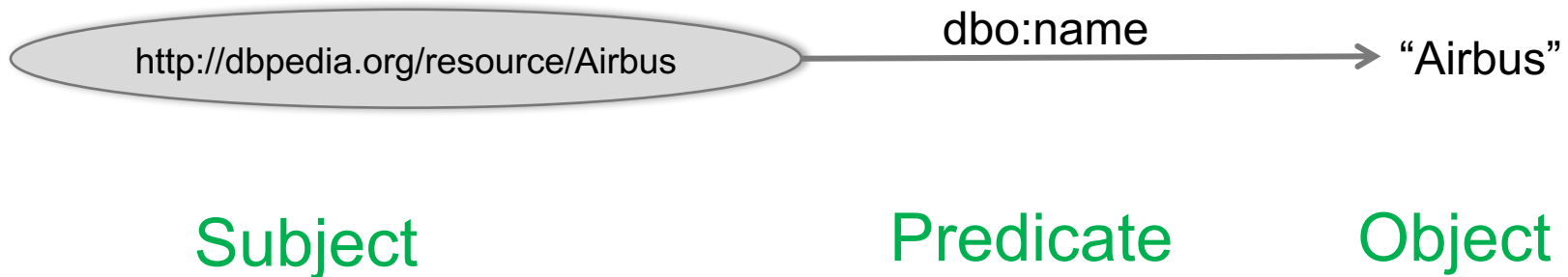
➔ Unlock the potential of isolated repositories (islands)



Tim Berners Lee, 2006

RDF – RESOURCE DESCRIPTION FRAMEWORK

- Statements of < subject predicate object >



... is called a triple

LINKED OPEN DATA

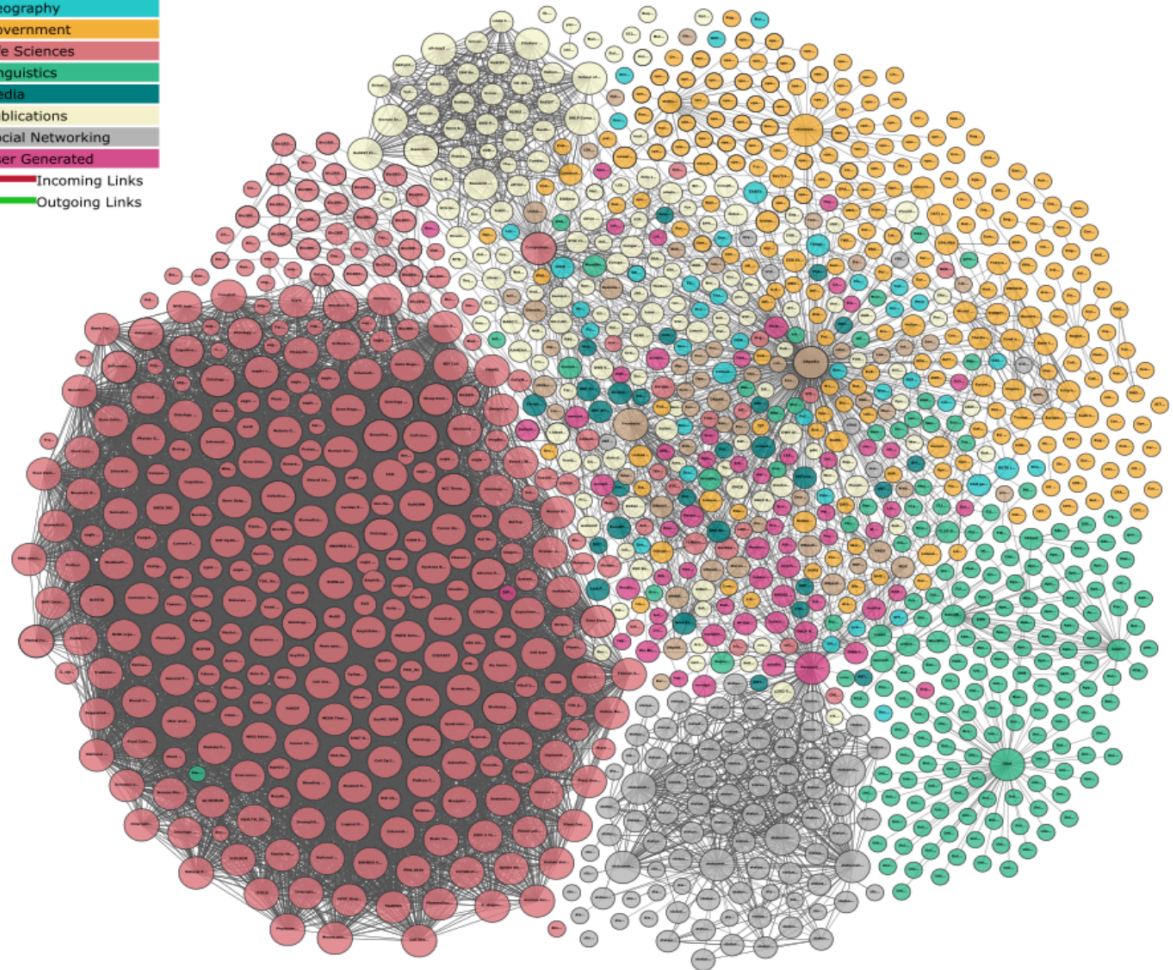
Linked Data - Datasets under an open access

- 1,139 datasets
- over 100B triples
- about 500M links
- several domains

Ex. DBPedia : 1.5 B triples



Linked Open Data (LOD)



"Linking Open Data cloud diagram 2017, by Andrejs Abele, John P. McCrae, Paul Buitelaar, Anja Jentzsch and Richard Cyganiak.
<http://lod-cloud.net/>"

NEED OF KNOWLEDGE

THE ROLE OF KNOWLEDGE IN AI

[Artificial Intelligence 47 (1991)]

ON THE THRESHOLDS OF KNOWLEDGE

Douglas B. Lenat

MCC
3500 W. Balcones Center
Austin, TX 78759

Edward A. Feigenbaum

Computer Science Department
Stanford University
Stanford, CA 94305

Abstract

We articulate the three major findings of AI to date:
(1) The Knowledge Principle: if a program is to perform a complex task well, it must know a great deal about the world in which it operates. (2) A plausible extension of that principle, called the Breadth Hypothesis: there are two additional abilities necessary for intelligent behavior in unexpected situations: falling back on increasingly general knowledge, and analogizing to specific but far-flung knowledge. (3) AI as Empirical Inquiry: we must test our ideas experimentally, on large problems. Each of these three hypotheses proposes a particular threshold to cross, which leads to a qualitative change in emergent intelligence. Together, they determine a direction for future AI research.

opponent is Castling.) Even in the case of having to search for a solution, the *method* to carry out the search may be

*The knowledge principle: “if a program is to perform a complex task well, it **must know a great deal about the world in which it operates.**”*

there is some minimum knowledge needed for one to even formulate it.

ONTOLOGY, A DEFINITION

“An ontology is an **explicit, formal specification** of a **shared conceptualization**.”

[Thomas R. Gruber, 1993]

Conceptualization: abstract model of domain related expressions

Specification: domain related

Explicit: semantics of all expressions is clear

Formal: machine-readable

Shared: consensus (different people have different perceptions)

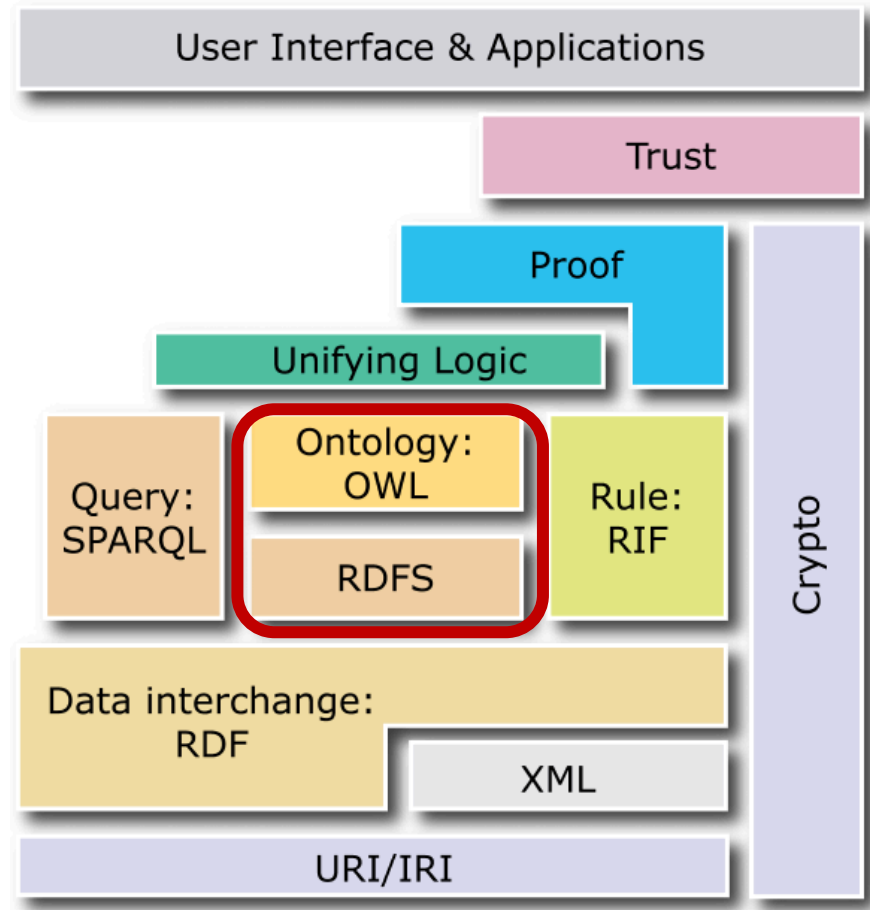
SEMANTIC WEB: ONTOLOGIES

RDFS – Resource Description Framework Schema

- Lightweight ontologies

OWL – Web Ontology Language

- Expressive ontologies



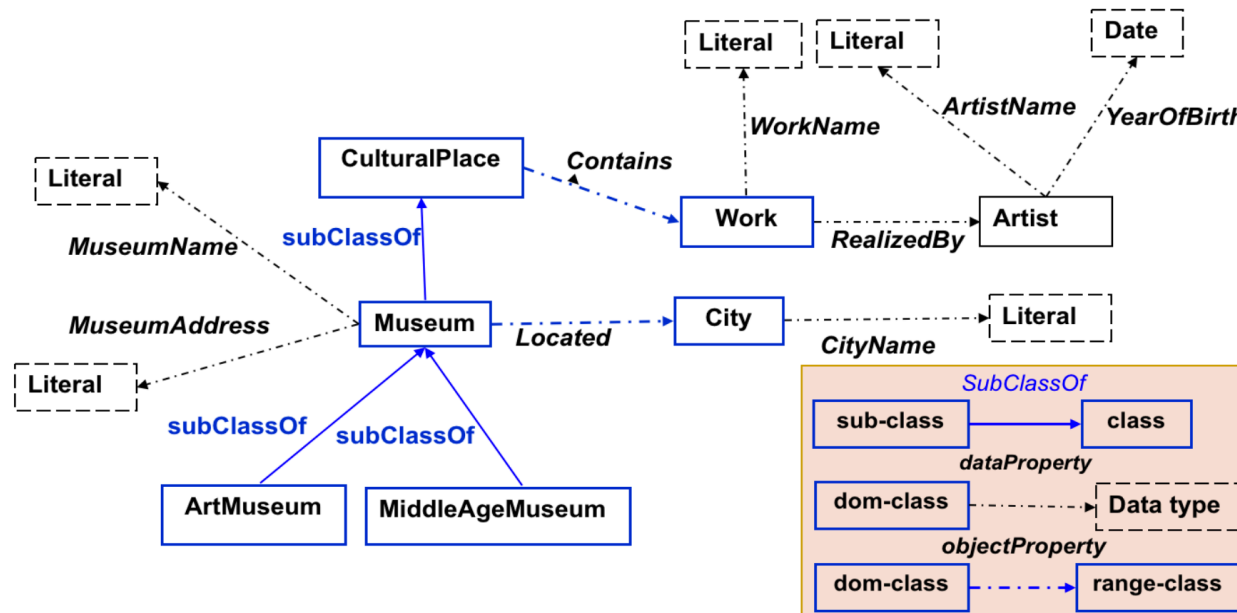
Source: https://it.wikipedia.org/wiki/File:W3C-Semantic_Web_layerCake.png

OWL ONTOLOGY

OWL – Web Ontology Language

- Represents rich and complex knowledge about things
- Based on Description Logic
- Can be used to verify the consistency of knowledge
- Can make implicit knowledge explicit

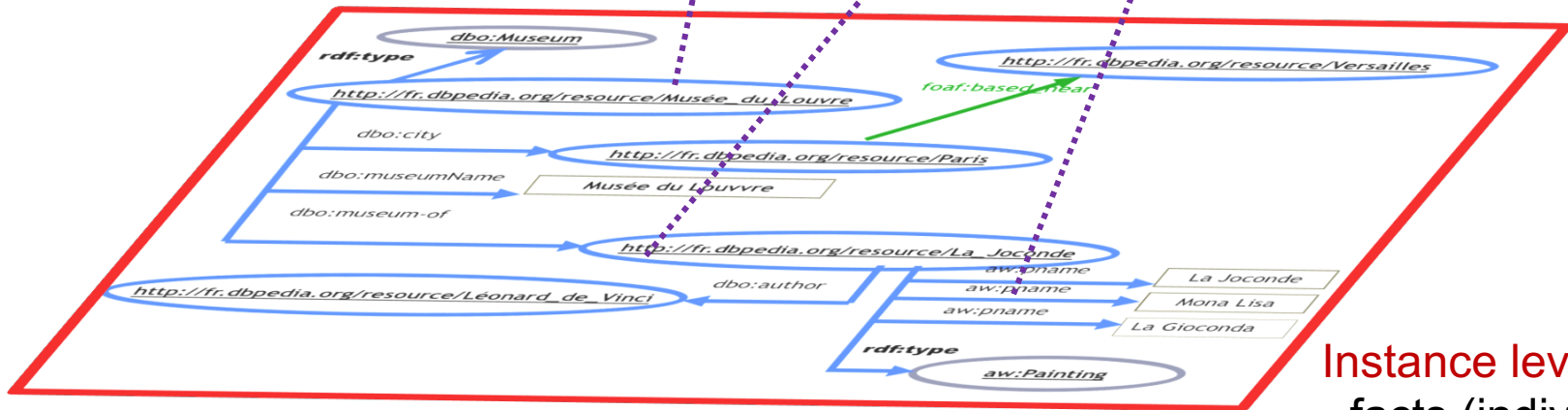
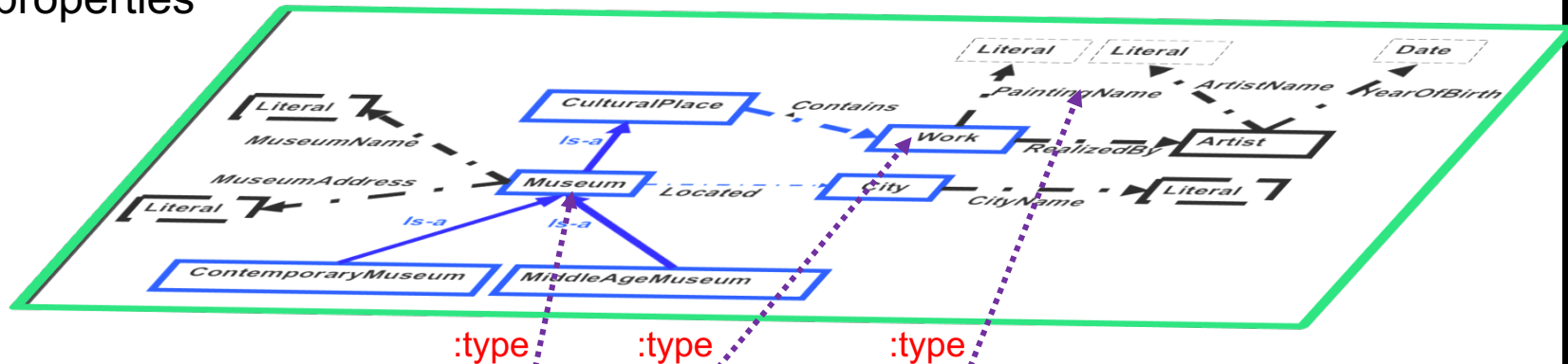
- **Classes:** concepts or collections of objects (individuals)
- **Properties:**
 - owl:DataTypeProperty (attribute)
 - owl:ObjectProperty (relation)
- **Hierarchy:**
 - owl:subClassOf
 - owl:subPropertyOf
- **Individuals:** ground-level of the ontology (instances)



ONTOLOGY LEVELS

Conceptual level:

- classes, properties
(relations)



Instance level:
- facts (individuals)

OWL ONTOLOGY - AXIOMS

- **Axioms:** knowledge definitions in the ontology that were **explicitly defined** and have not been proven true.
 - Reasoning over an ontology
 - Implicit knowledge can be made explicit by logical reasoning

- **Example:**

Pompidou museum is an **Art Museum**

< Pompidou_museum rdf:type ArtMuseum> .

Pompidou museum contains ***Hallucination partielle***

< Pompidou_museum ao:contains ***Hallucination partielle***> .

- **Infer that:**

→ Pompidou museum is a CulturalPlace

< Pompidou_museum rdf:type CulturalPlace> .

Because: **Museum** subsumes **ArtMuseum** and **CulturalPlace** subsumes **Museum**

→ ***Hallucination partielle*** is a Work

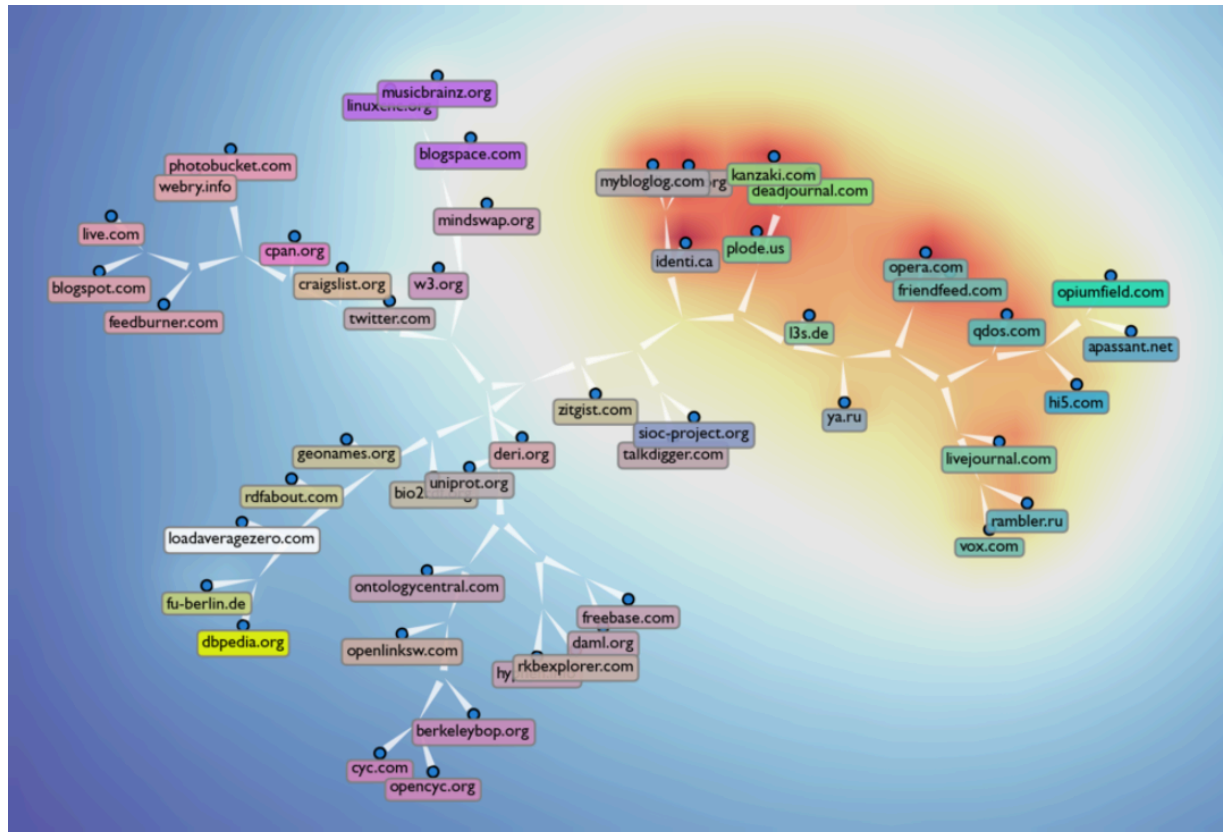
< ***Hallucination partielle*** rdf:type ao:Work> .

Because: the **range** of the object property **contains** is the class **Work**.

ONTOLOGY PREDICATES SPREAD ON THE SEMANTIC WEB

- RDFa (or Resource Description Framework in Attributes)
- Top 50 web sites publishing Semantic Web data, clustered by predicates used.

FOAF



KNOWLEDGE GRAPHS

WEB SEARCH WITHOUT KNOWLEDGE GRAPHS

+Myles Search Images Mail Drive Calendar Sites Groups Admin More -

Google buy olive oil

Web Images Maps News Videos More Search tools






About 51,700,000 results (0.32 seconds)

Ads related to **buy olive oil**

Buy Olive Oil Online - OliveOilLovers.com
www.oliveoillovers.com/
Buy Olive Oil Online For The Best Quality & Best Brands At Low Prices
Infused - Gifts

Buy Olive Oil - igourmet.com
www.igourmet.com/
★★★★★ 688 reviews for igourmet.com
Top Selection of Gourmet Olive Oil Gourmet Foods, Cheese & Gift Ideas

Shop for buy olive oil on Google

Sponsored				
				
Basil Specialty Olive Oil \$34.00 O&CO.	Flora Olive Oil 17 Fluid Oun... \$16.99 Vitamin Shop...	Filippo Berio Extra Virgin O... \$8.75 Soap.com	Extra Virgin Olive Oil 3 Lit... \$14.99 WEBstaurant...	Williams-Sonoma Extra Virgin O... \$59.95 Williams-Son...

Oliver Oil: Buy Gourmet Olive Oil Online. Italian Spanish French...
www.igourmet.com/olive-oil.asp
Olive Oil: Shop the widest selection of gourmet Olive Oil, plus thousands of other gourmet foods from over 100 countries, online exclusively at igourmet.com.

Ads

Pure Italian Olive Oils
www.cybercucina.com/ItalianOliveOils
★★★★★ 166 seller reviews
Buy Now & Save Big! Browse Our Catalog See Our Specials. Free S&H.

Shop O&CO.
www.oliviersandco.com/
Big selection of oils, vinegars, tapenades and other gourmet foods.

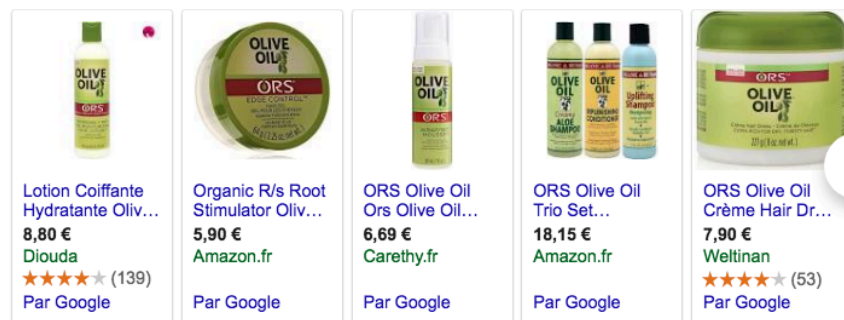
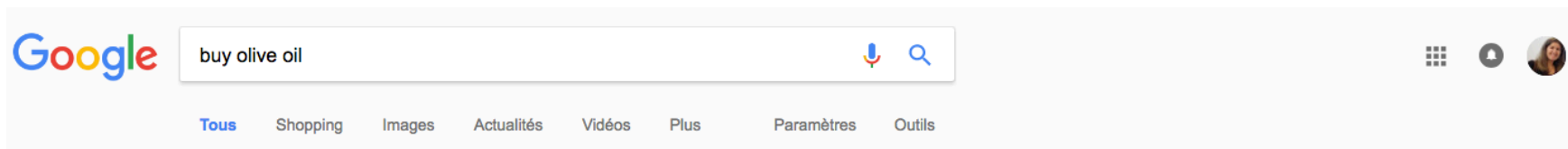
Olive Oil for Soap Making
www.bulkapothecary.com/
1 (800) 396 8740
Extra Virgin Olive Oil & 1000's of Wholesale Soap Making Supplies

Save \$1.00 On Olive Oil
www.pompeian.com/
The Only USDA Quality Monitored Extra Virgin Olive Oil, Get It Now

Eliki Olive Oil at Amazon
www.amazon.com/grocery
Buy Groceries at Amazon & Save. Qualified orders over \$25 ship free

Old Town Olive Oil

WEB SEARCH WITH KNOWLEDGE GRAPHS



Olive oil - Wikipedia

https://en.wikipedia.org/wiki/Olive_oil ▼ Traduire cette page

Olive oil is a liquid fat obtained from olives a traditional tree crop of the Mediterranean Basin. The oil is produced by pressing whole olives. It is commonly used ...

[Olive oil acidity](#) · [Olive oil extraction](#) · [Olive oil regulation and ...](#) · [Oleic acid](#)

OIL BY OLIVE

oilbyolive.com/ ▼ Traduire cette page

OIL BY OLIVE. collection 3 · contact · about · press · past · OIL BY OLIVE · Frontpage made with Lay Theme OIL BY OLIVE C3 made with Lay Theme.

Traduction olive oil français | Dictionnaire anglais | Reverso

dictionnaire.reverso.net/anglais-francais/olive%20oil ▼

traduction olive oil francais, dictionnaire Anglais - Français, définition, voir aussi 'virgin olive oil', 'olive', 'olive branch', 'olive grove', conjugaison, expression, ...

All About Olive Oil - Olive Oil Times

<https://www.oliveoiltimes.com/olive-oil> ▼ Traduire cette page

"Olive oil" is how we refer to the oil obtained from the fruit of olive trees. People have been eating olive oil for thousands of years and it is now more popular than ...

Huile d'olive

L'huile d'olive est la matière grasse extraite des olives lors de la trituration dans un moulin à huile. Elle est un des fondements de la cuisine méditerranéenne et est, sous certaines conditions, bénéfique pour la santé. [Wikipédia](#)

Informations nutritionnelles

Huile d'olive

Valeur pour 100 grammes

Calories 884

Lipides 100 g

Acides gras saturés 14 g

Acides gras poly-insaturés 11 g

Acides gras mono-insaturés 73 g

Cholestérol 0 mg

Sodium 2 mg

Potassium 1 mg

Glucides 0 g

Fibres alimentaires 0 g

Sucres 0 g

Protéines 0 g

Vitamine A	0 IU	Vitamine C	0 mg
Calcium	1 mg	Fer	0,6 mg
Vitamine D	0 IU	Vitamine B6	0 mg
Vitamine B ₁₂	0 µg	Magnésium	0 mg

Recherches associées

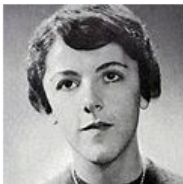
Voir d'autres éléments (plus de 15)

QUESTION ANSWERING WITH KNOWLEDGE GRAPHS



All Images Videos Maps News | My saves

15 900 000 Results Date ▾ Language ▾ Region ▾



Barack Obama · Mother

Ann Dunham

Ann Dunham - Wikipedia

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

Stanley Ann Dunham (November 29, 1942 – November 7, 1995) was an American anthropologist who specialized in the economic anthropology and rural development of ...

Barack Obama Sr · Zarai Taraqiati Bank Limited · Lolo Soetoro · Wikipedia:Good Articles

Family of Barack Obama - Wikipedia

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

The family of **Barack Obama**, the 44th President of the United States, and his wife Michelle **Obama** is made up of people of Kenyan (Luo), African-American, and Old Stock ...

United States Citizen · Craig Robinson · Barack Obama Sr · Jonathan Singletary Dunham

Ann Dunham

Anthropologue



Stanley Ann Dunham, née le 29 novembre 1942 à Wichita et morte le 7 novembre 1995 à Honolulu, est une anthropologue américaine spécialisée dans l'anthropologie économique et le développement rural. Elle est la mère de Barack Obama, le 44^e ... +

W Wikipedia

Parents: Madelyn Dunham (Mother) · Stanley Armour Dunham (Father)

Spouse: Lolo Soetoro (m. 1965 - 1980) · Barack Obama, Sr. (m. 1961 - 1964)

Children: Barack Obama (Son) · Maya Soetoro-Ng (Daughter)

Lived: 29 nov. 1942 - 7 nov. 1995 (age 52)

Education: Mercer Island High School · Université d'Hawaï à Mānoa · Université de Washington

Buried: Océan Pacifique

CONNECTING EVENTS AND PEOPLE WITH KNOWLEDGE GRAPHS

The screenshot shows the LinkedIn interface with a search bar at the top containing the text 'ISWC 2017 vienna'. Below the search bar, there are navigation tabs for 'All', 'People', 'Jobs', 'Content', 'Companies', 'Groups', and 'Schools'. The 'People' tab is selected. Below the tabs, there is a banner with the text 'Rencontrez-nous - Envie de changement professionnel ? Rencontrez un consultant'. Below the banner, it says 'Showing 12 results'. There are five profile cards visible, each with a profile picture, name, title, location, current role, and a 'Connect' button. The profiles are: Axel Polleres (Head of Institute at WU Wien), Marco Fossati (Project Leader at Wikimedia Foundation), Adrian M.P. Brasoveanu (Researcher at MODUL Technology), Antoine Zimmermann (Associate professor in Semantic Web technologies), and Ioannis Chrysakis (Research and Development Engineer, ICT Expert, Senior Developer).

Showing 12 results

Axel Polleres • 2nd
Head of Institute (Institutsvorstand) - Institute for Information Business, WU Wien
Austria area
Current: Faculty at Complexity Science Hub Vienna
15 shared connections

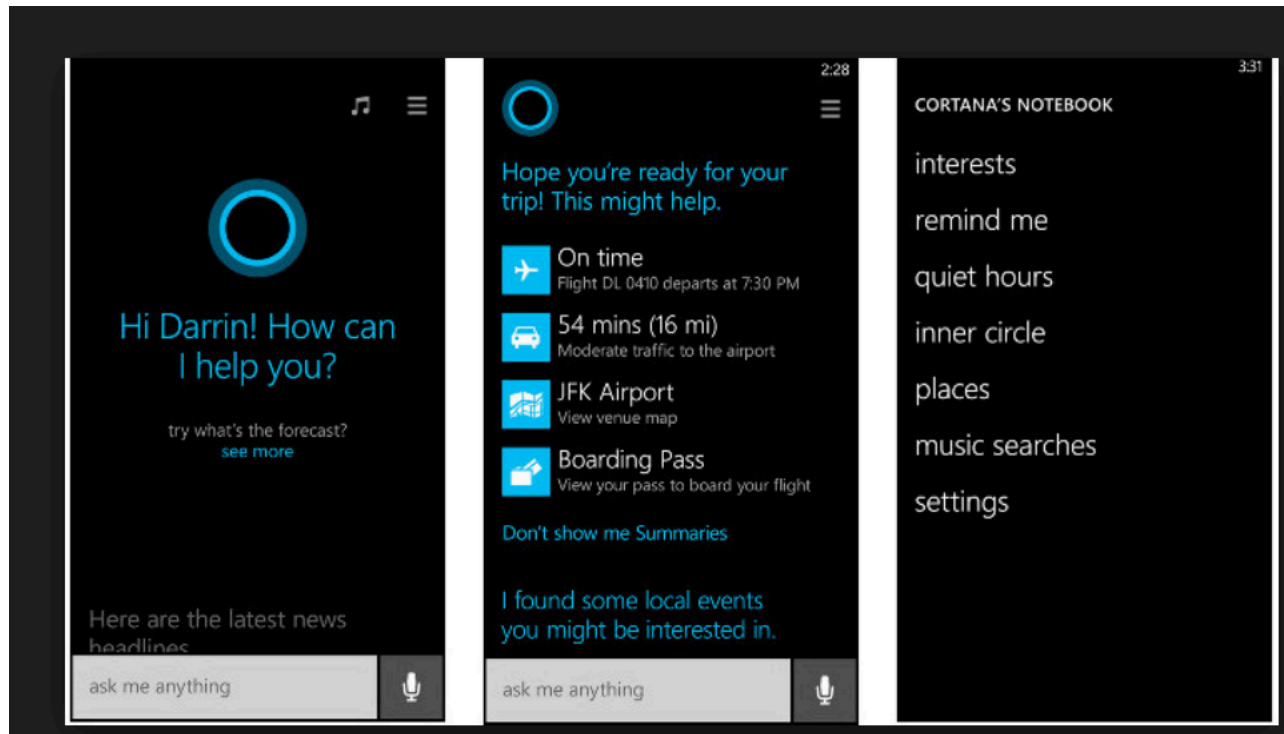
Marco Fossati • 2nd
Project Leader at Wikimedia Foundation IEG program
Trento Area, Italy
Current: Project Leader at Wikimedia Foundation
4 shared connections

Adrian M.P. Brasoveanu • 2nd
Researcher at MODUL Technology
Austria area
Past: Researcher at MODUL University Vienna
1 shared connection

Antoine Zimmermann • 2nd
Associate professor in Semantic Web technologies
Lyon Area, France
8 shared connections

Ioannis Chrysakis • 2nd
Research and Development Engineer, ICT Expert, Senior Developer
Greece
Summary: ...Journal (SWJ), Semantics 2017, SEMAPRO 2017, CSICT 2017...
7 shared connections

TOWARDS A KNOWLEDGE-POWERED DIGITAL ASSISTANT



Cortana (Microsoft)

- Natural access and storage of knowledge
- Chat bots
- Personalization
- Emotion

KNOWLEDGE GRAPH: A DEFINITION ...

The **Knowledge Graph** is a [knowledge base](#) used by [Google](#) to enhance its [search engine](#)'s search results with [semantic-search](#) information gathered from a wide variety of sources. Knowledge Graph display was added to Google's search engine in 2012, starting in the United States, having been announced on May 16, 2012.^[1] It uses a [graph database](#) to provide structured and detailed information about the topic in addition to a list of links to other sites. The goal is that users would be able to use this information to resolve their query without having to navigate to other sites and assemble the information themselves.^[2] The short summary provided in the knowledge graph is often used as a spoken answer in [Google Assistant](#) searches.^[3]

Wikipedia (en)

This is not a formal definition!

KNOWLEDGE GRAPH: A DEFINITION ...

[L. Ehrlinger and W. Wöß, SEMANTICS'2016]

Definition	Source	
"A knowledge graph (i) mainly describes real world entities and their interrelations, organized in a graph, (ii) defines possible classes and relations of entities in a schema, (iii) allows for potentially interrelating arbitrary entities with each other and (iv) covers various topical domains."	Paulheim [16]	→ Populated Ontology
"Knowledge graphs are large networks of entities, their semantic types, properties, and relationships between entities."	Journal of Web Semantics [12]	→ RDF Graph
"Knowledge graphs could be envisaged as a network of all kind things which are relevant to a specific domain or to an organization. They are not limited to abstract concepts and relations but can also contain instances of things like documents and datasets."	Semantic Web Company [3]	→ Populated Ontology
"We define a Knowledge Graph as an RDF graph. An RDF graph consists of a set of RDF triples where each RDF triple (s, p, o) is an ordered set of the following RDF terms: a subject $s \in U \cup B$, a predicate $p \in U$, and an object $U \cup B \cup L$. An RDF term is either a URI $u \in U$, a blank node $b \in B$, or a literal $l \in L$."	Färber et al. [7]	→ RDF Graph
"[...] systems exist, [...], which use a variety of techniques to extract new knowledge, in the form of facts, from the web. These facts are interrelated, and hence, recently this extracted knowledge has been referred to as a knowledge graph."	Pujara et al. [17]	→ Extracted RDF Graph

[3] A. Blumauer. From Taxonomies over Ontologies to Knowledge Graphs, July 2014. <https://blog.semanticweb.at/2014/07/15/from-taxonomies-over-ontologiesto-knowledge-graphs> [August, 2016].

[7] M. Farber, B. Ell, C. Menne, A. Rettinger, and F. Bartscherer. Linked Data Quality of DBpedia, Freebase, OpenCyc, Wikidata, and YAGO. Semantic Web Journal, 2016. <http://www.semantic-web-journal.net/content/linked-data-quality-dbpedia-freebaseopencyc-wikidata-and-yago> [August, 2016] (revised version, under review).

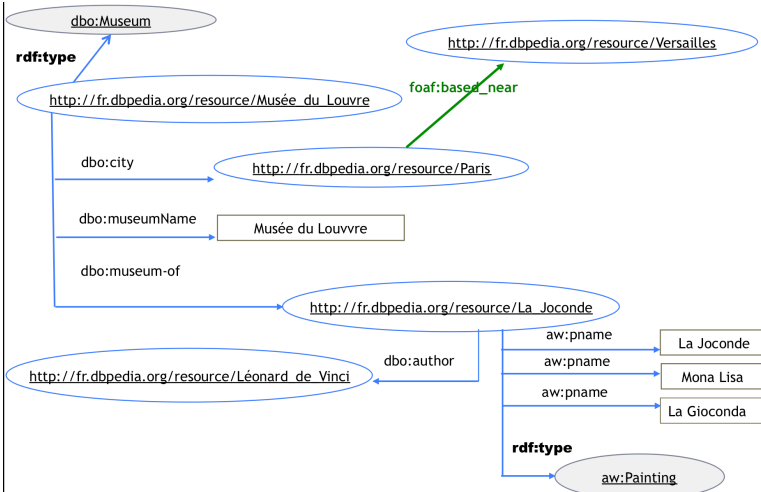
[12] M. Kroetsch and G. Weikum. Journal of Web Semantics: Special Issue on Knowledge Graphs. <http://www.websemanticsjournal.org/index.php/ps/announcement/view/19> [August, 2016].

[16] H. Paulheim. Knowledge Graph Refinement: A Survey of Approaches and Evaluation Methods. Semantic Web Journal, (Preprint):1–20, 2016.

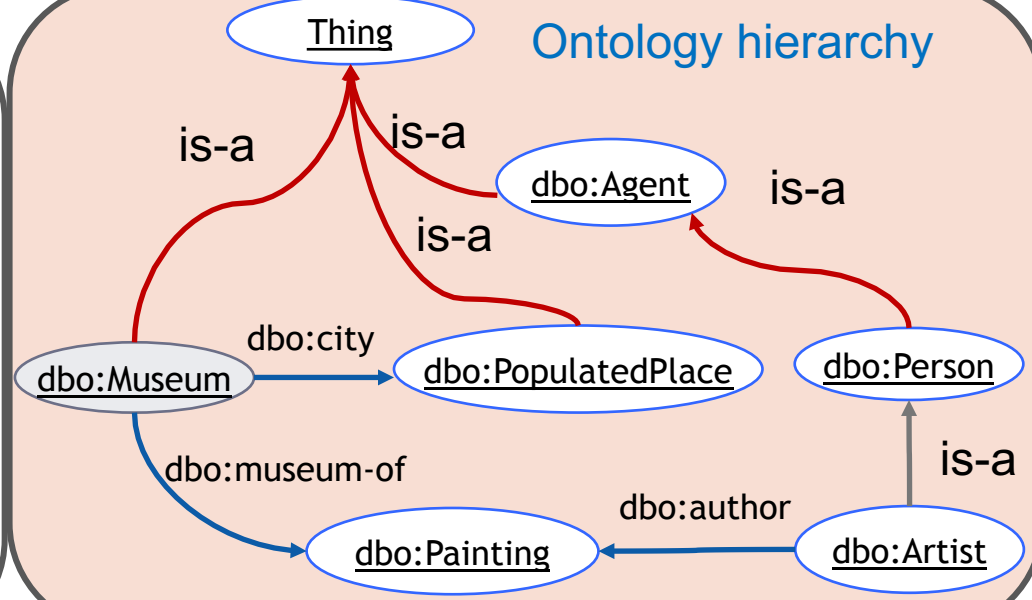
[17] J. Pujara, H. Miao, L. Getoor, and W. Cohen. Knowledge Graph Identification. In Proceedings of the 12th International Semantic Web Conference - Part I, ISWC '13, pages 542–557, New York, USA, 2013. Springer.

KNOWLEDGE GRAPH (KG)

RDF Graphs



Ontology hierarchy



Querying (SPARQL)

```
PREFIX dbo: <http://dbpedia.org/ontology#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
SELECT ?m ?p
WHERE { ?m rdf:type dbo:Museum . ?m dbo:museum-of ?p . }
```

Reasoners: (Pellet, Fact++, Hermit, etc.)

- KG saturation: infer whatever can be inferred from the KG.
- KG consistency checking: no contradictions
- KG repairing
- ...

Ontology axioms and rules

```
owl:equivalentClass(dbo:Municipality, dbo:Place)
owl:equivalentClass(dbo:Place, dbo:Wikidata:Q532)
owl:equivalentClass(dbo:Village, dbo:PopulatedPlace)
owl:equivalentClass(dbo:PopulatedPlace, dbo:Municipality)
owl:disjointClass(dbo:PopulatedPlace, dbo:Artist)
owl:disjointClass(dbo:PopulatedPlace, dbo:Painting)
owl:FunctionalProperty(dbo:city)
owl:InverseFunctionalProperty(dbo:museum-of)
```

```
dbo:birthPlace(X, Y) => dbo:citizensOf(X, Y)
dbo:parentOf(X, Y) => dbo:child(Y, X)
```

KNOWLEDGE GRAPH EXPANSION AND ENRICHMENT

- **Expansion: knowledge graphs are incomplete**
 - Data linking (entity resolution, reference reconciliation, ...)
 - Link prediction: add relations
 - Ontology matching: connect graphs
 - Missing values prediction/inference
- **Enrichment: can new knowledge emerge from knowledge graphs?**
 - Knowledge discovery: Key discovery
 - Automatic reasoning and planning
- **Validation: knowledge graphs may contain errors**
 - Link Validation
 - Dealing with errors and ambiguities

KNOWLEDGE GRAPH EXPANSION AND ENRICHMENT

- **Expansion: knowledge graphs are incomplete**
 - **Data linking (entity resolution, reference reconciliation, ...)**
 - Link prediction: add relations
 - Ontology matching: connect graphs
 - Missing values prediction/inference
- **Enrichment: can new knowledge emerge from knowledge graphs?**
 - **Knowledge discovery: Key discovery**
 - Automatic reasoning and planning
- **Validation: knowledge graphs may contain errors**
 - **Link Validation**
 - Dealing with errors and ambiguities

OUTLINE

- **Introduction** (F. Saïs)
- **Technical Part**
 1. Data Linking (N. Pernelle)
 2. Key Discovery (D. Symeonidou)
 3. Link Validation (F. Saïs)
- **Future Challenges** (F. Saïs)