

Knowledge Graphs as a backbone for Bilateral AI and Intelligent Data Management

Axel Polleres

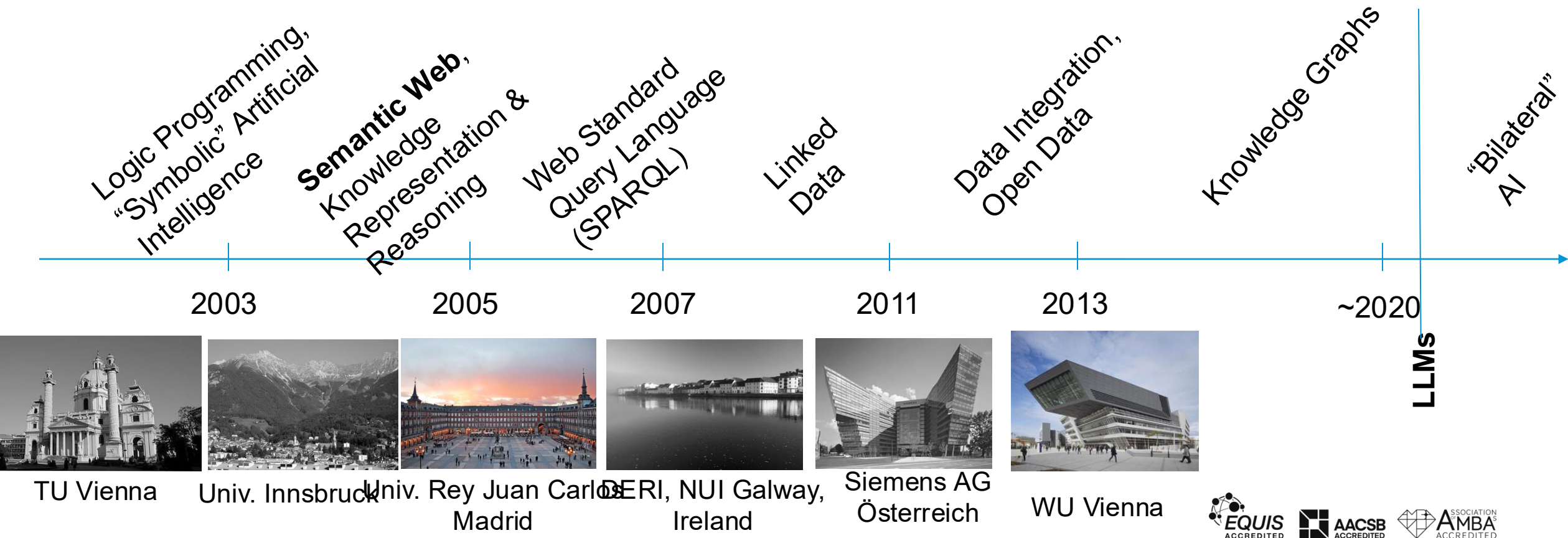
WU
VIENNA

EQUIS
ACCREDITED

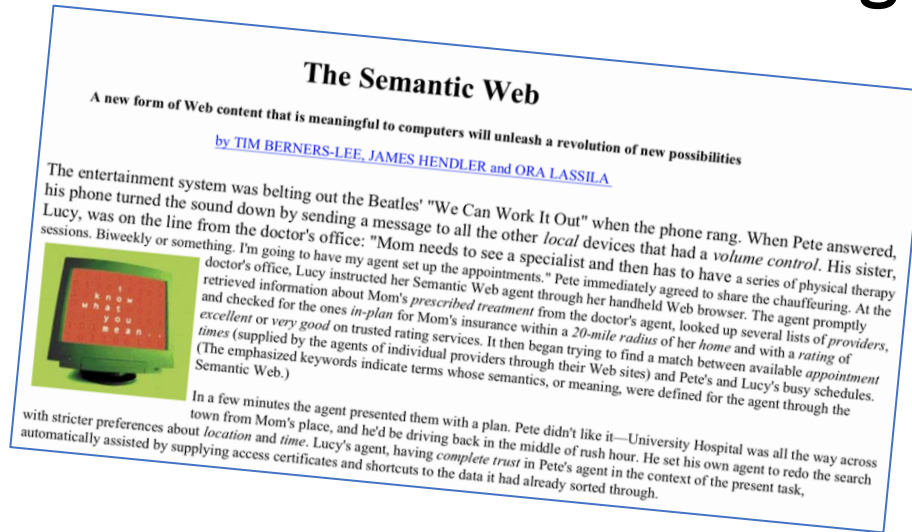
AACSB
ACCREDITED

**ASSOCIATION
AMBA**
ACCREDITED

My background...



Instead of its initial focus on **agents** the “Semantic Web” ...



*“[...] The **agent** promptly retrieved information about Mom's prescribed treatment from the doctor's agent, looked up several lists of providers, and checked for the ones in-plan for Mom's insurance within a 20-mile radius of her home and with a rating of excellent or very good on trusted rating services[...]*”

- *Appointment detection in emails*
- *Semantic Search*
- *Ratings of products/services*

... has then mostly become the basis for the "Web of Data"...

*"If HTML and the Web made all the online documents look like one huge **book**, RDF, schema and inference languages will make all the data in the world look like **one huge database**"*

Tim Berners-Lee, Weaving the Web, 1999

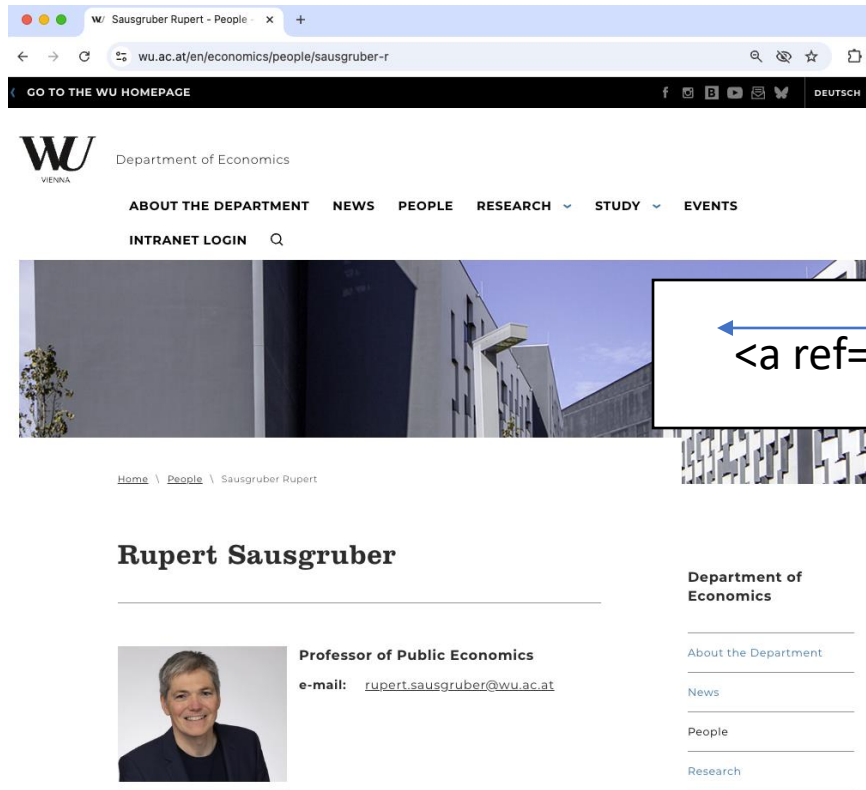
... and its more recent focus on Open Knowledge Graphs...

*" huge knowledge bases, also known as **knowledge graphs**, have been automatically constructed from web data, and have become a key asset for search engines and other use cases.*

Gerhard Weikum, Knowledge Graphs 2021: A Data Odyssey, VDLB 2021

Phase 1: RDF – Linked “Metadata”

- “Typed links” describing Ressources (such as Web pages)



The screenshot shows a web browser window with the URL `wu.ac.at/en/economics/people/sausgruber-r`. The page header includes the WU logo and navigation links. The main content area features a profile for Rupert Sausgruber, identified as a Professor of Public Economics, with his contact information and a list of links to department resources.

WU Department of Economics

ABOUT THE DEPARTMENT NEWS PEOPLE RESEARCH STUDY EVENTS

INTRANET LOGIN

Home \ People \ Sausgruber Rupert

Rupert Sausgruber

Professor of Public Economics

e-mail: rupert.sausgruber@wu.ac.at

Department of Economics

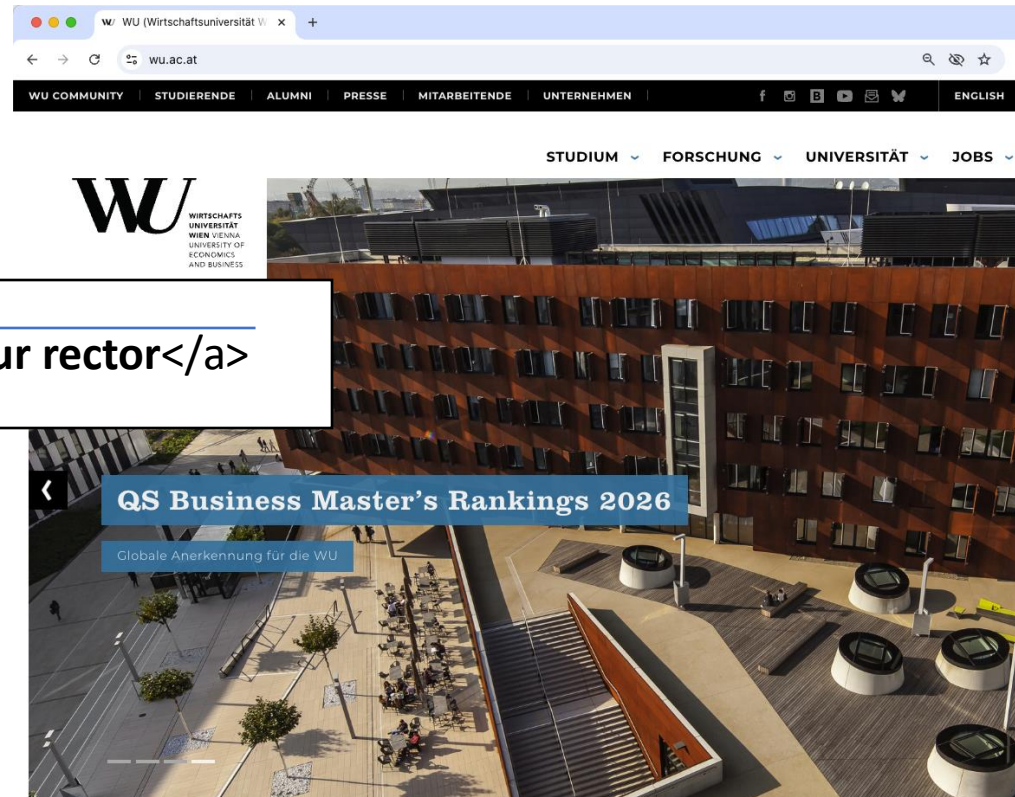
About the Department

News

People

Research

`our rector`



The screenshot shows the WU website with a prominent banner for the QS Business Master's Rankings 2026. The banner includes the text 'Globale Anerkennung für die WU' and a background image of the WU building.

WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

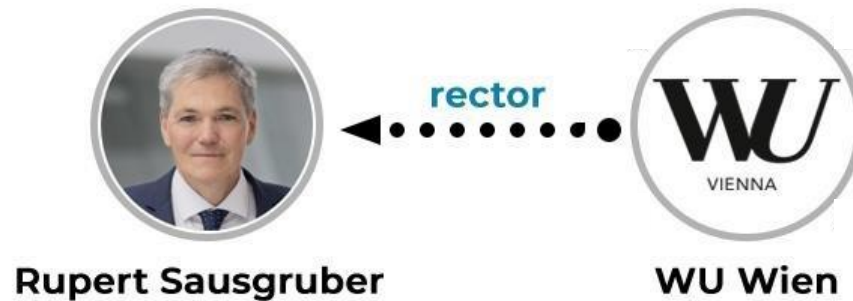
STUDIUM FORSCHUNG UNIVERSITÄT JOBS

QS Business Master's Rankings 2026

Globale Anerkennung für die WU

Phase 1: RDF (Ressource Description framework) – Linked “Metadata”

- Typed links “triples” describing Ressources (such as Web pages)

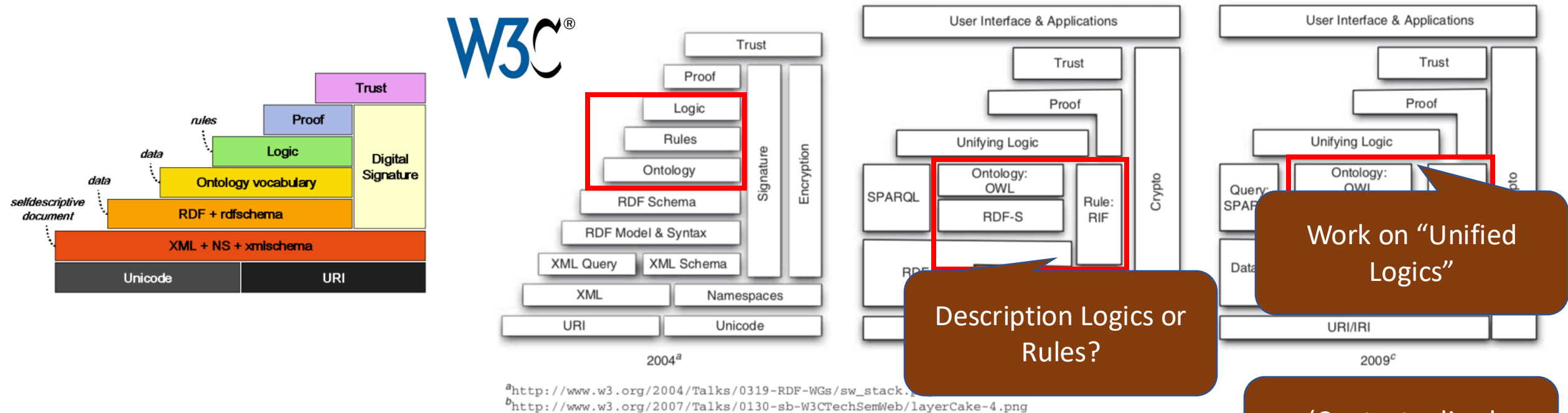


`<http://www.wu.ac.at#wuwien>` `<http://www.wu.ac.at#hasRector>` `<http://www.wu.ac.at/Rupert_Sausgruber>`

Semantic Web: Standard formats, Reasoning & Logics



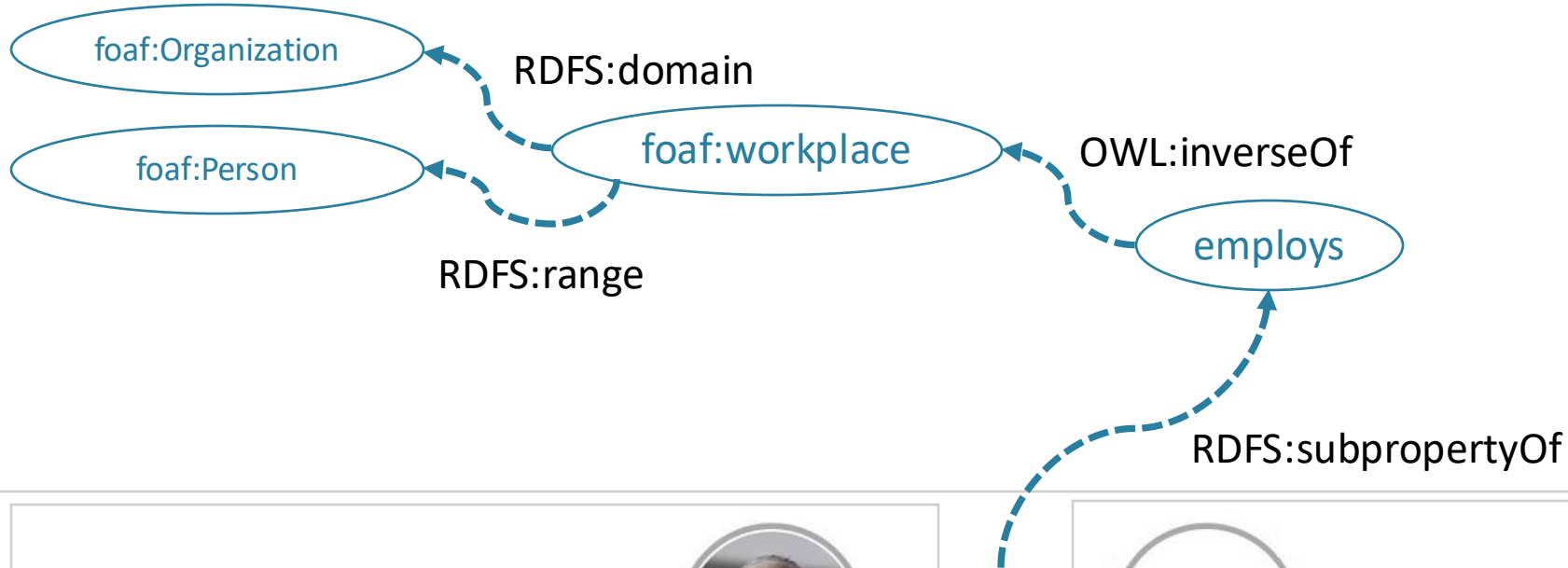
- (2000s - ca. 2009)



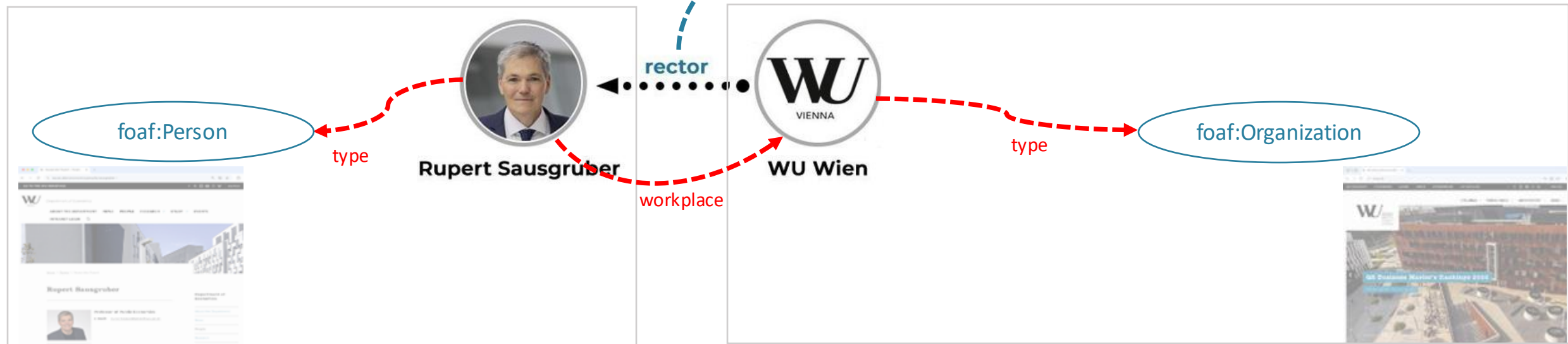
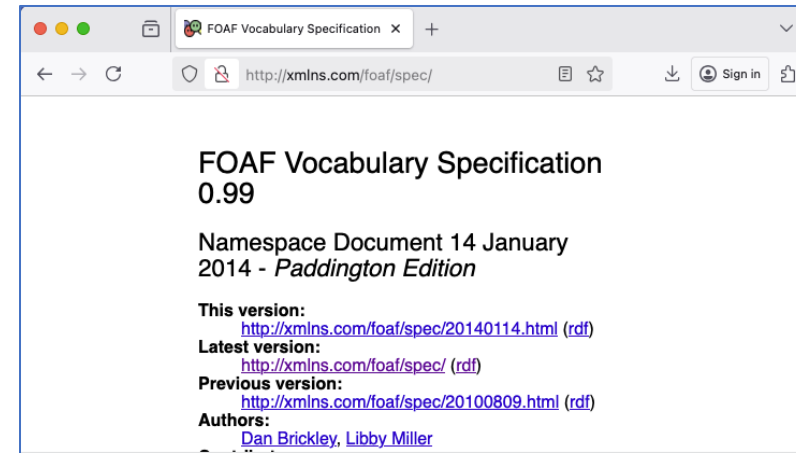
*Good news! Standards meant a Boost in KR/AI research:
We know very well which ontological reasoning
approaches are decidable and how they scale
→ RDF, OWL, SPARQL Ontology-based Data Access but also:
constraint checking (SHACL)*

SELECT ?P
{ ?P a foaf:Person .
?P foaf:workPlace wu:WUWien. }

Phase 2: RDF, RDF Schema, OWL Linked “Metadata” with Rules and Ontologies.



FOAF “Ontology”



Phase 2: RDF, RDF Schema, OWL, *SPARQL* Linked “Metadata” with Rules and Ontologies.

*"If HTML and the Web made all the online documents look like one huge **book**, RDF, schema and inference languages will make all the data in the world look like **one huge database**"*

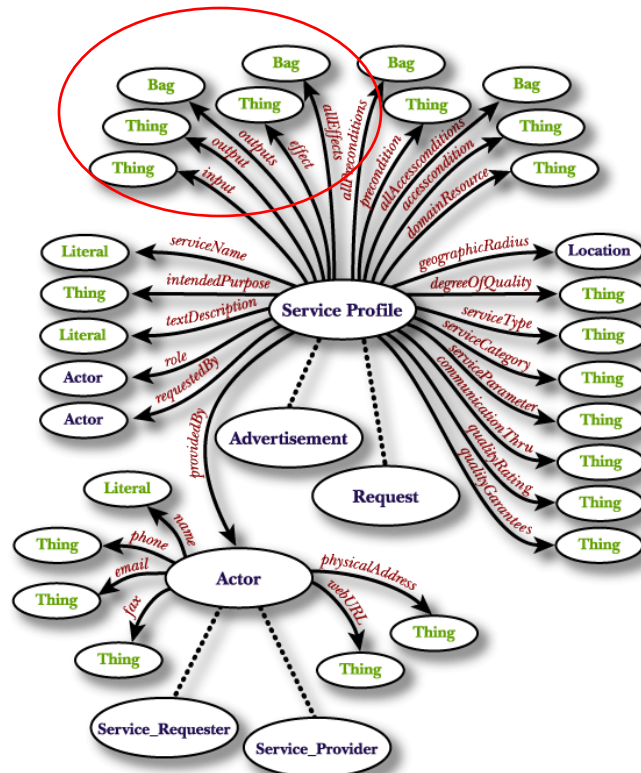
Tim Berners-Lee, Weaving the Web, 1999

```
SELECT ?P
{ ?P a foaf:Person .
  ?P foaf:workPlace wu:WUWien. }
```



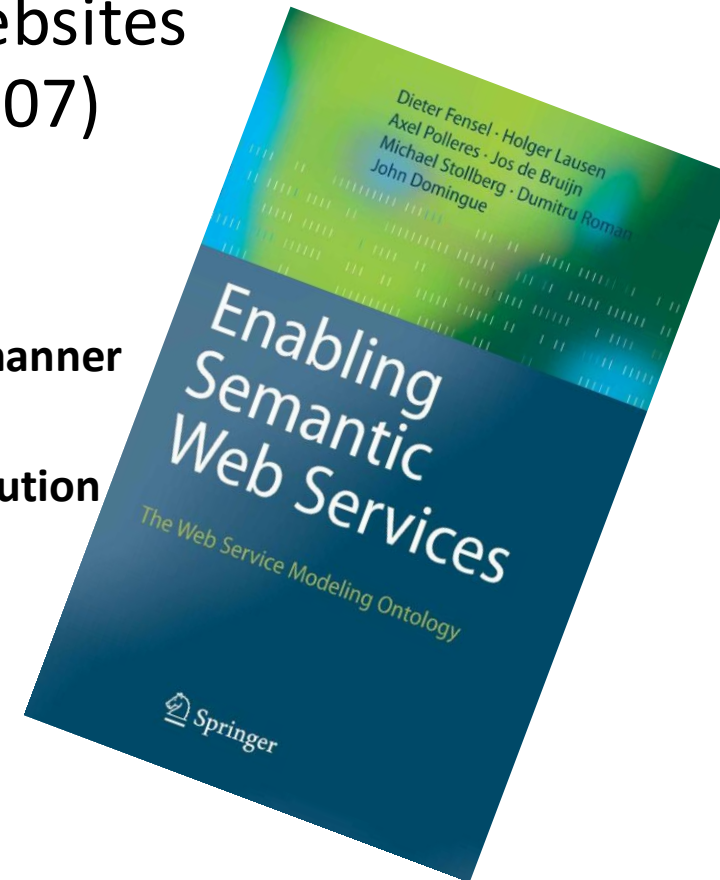
Phase 2 “interlude”

- **Semantic Web Services:** We cannot only describe Websites and Metadata, but also Services and APIs! (~2002-2007)



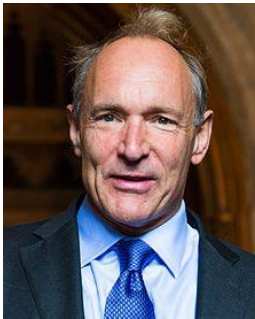
- Main Idea:
 - Describing **Services & Agents** in a **declarative manner**
 - using ontologies...
 - ... should enable **automated composition, execution**

On hindsight: automated composition by purely symbolic inference ... was maybe maybe ahead of its time...



Focus on Data: Linked Data

- (ca. 2006/7 – ca. 2013)
 - Main question: How can I **publish** “Knowledge on the Web” ...



Linked Data Principles

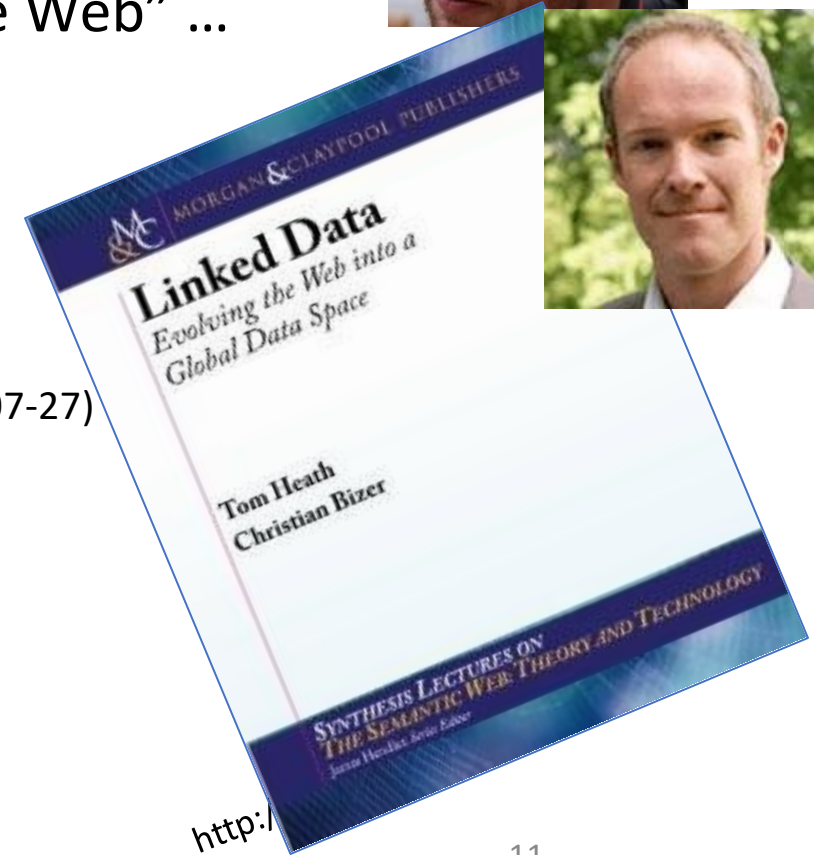
- **LDP1**: use URIs as names for things
- **LDP2**: use HTTP URIs so those names can be dereferenced
- **LDP3**: return useful – RDF? – information upon dereferencing those URIs
- **LDP4**: include links using externally dereferenceable URIs.

<https://www.w3.org/DesignIssues/LinkedData.html> (originally published 2006-07-27)



“A Little Semantics Goes a Long Way” (Jim Hendler)

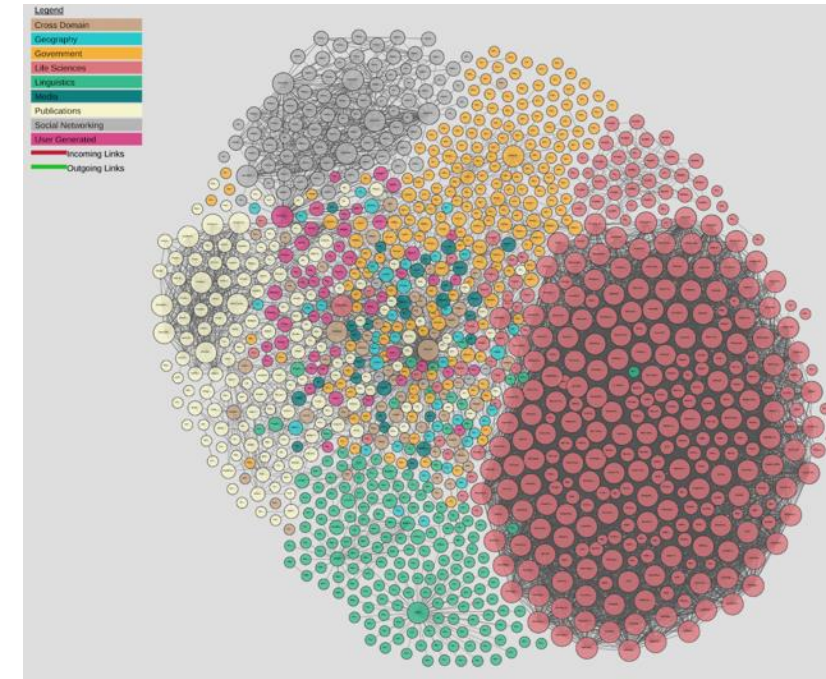
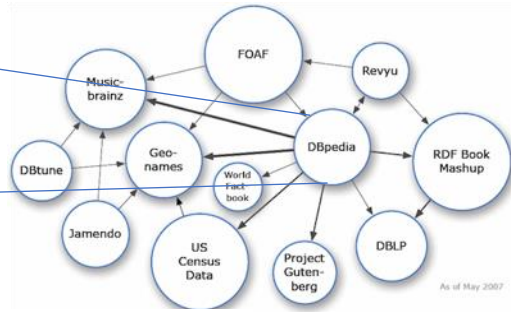
<https://www.cs.rpi.edu/~hendler/LittleSemanticsWeb.html>





Phase3: From Semantic Web to Linked (Open) Data

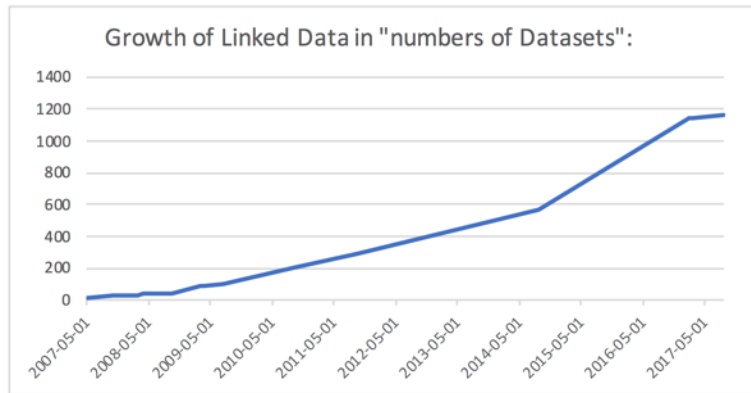
- (ca. 2006/7 – 2012)
 - Main question: How can I **publish** “Knowledge on the Web” ...
 - Linked **Open** Data... growth slowed down a bit
 - A lot of active developments to publish and link RDF Data
 - also in Enterprises (“Enterprise Linked Data”)



<http://lod-cloud.net/>

2017-08-22
2017-02-20
2017-01-26
2014-08-30
2011-09-19
2010-09-22
2009-07-14
2009-03-27
2009-03-05
2008-09-18
2008-03-31
2008-02-28
2007-11-10
2007-11-07
2007-10-08
2007-05-01

1163
1139
1146
570
295
203
95
93
89
45
34
32
28
28
25
12



Axel Polleres, Maulik R. Kamdar, Javier D. Fernández, Tania Tudorache, and Mark A. Musen. [A more decentralized vision for linked data](#). In *Decentralizing the Semantic Web (Workshop of ISWC2018)*.



Phase3: From Semantic Web to Linked (Open) Data

- (ca. 2006/7 – 2013)
 - Main question: How can I **publish** “Knowledge on the Web” ...
- *Side question(s):*
Can deductive symbolic inference and queries scale to a decentralized Web?

Reasoning Web. Semantic Technologies for Intelligent Data Access (Reasoning Web 2013)

International Semantic Web Conference (2014)

Strategies for executing federated queries in SPARQL1.1

Carlos Buil-Aranda^{1*}, Axel Polleres^{2**}, and Jürgen Umbrich²

¹ Department of Computer Science, Pontificia Universidad Católica, Chile
cbuil@ing.puc.cl

² Vienna University of Economics and Business (WU)
{first.last}@wu.ac.at

RDFS & OWL Reasoning for Linked Data

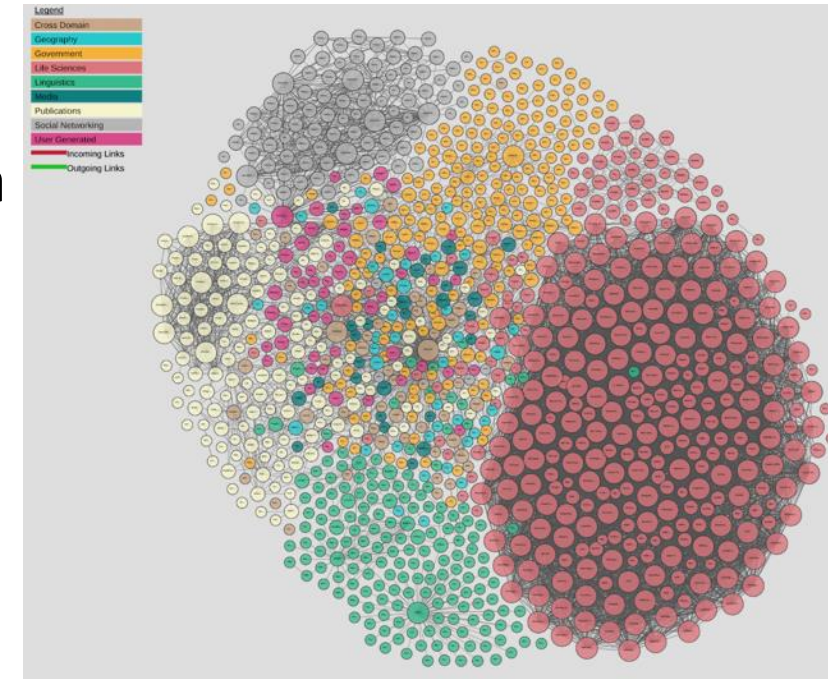
Axel Polleres¹, Aidan Hogan², Renaud Delbru², and Jürgen Umbrich^{2,3}

¹ Siemens AG Österreich, Siemensstraße 90, 1210 Vienna, Austria

² Digital Enterprise Research Institute, National University of Ireland, Galway

³ Fujitsu (Ireland) Limited, Swords, Co. Dublin, Ireland

Abstract. Linked Data promises that a large portion of Web Data will be available as one big interlinked RDF database against which structured queries can be answered. In this lecture we will show how reasoning – using the RDF Schema (RDFS) and the Web Ontology Language (OWL) – can be used to obtain more complete answers for such queries over Linked Data. We then introduce two high-level architectures for answering over Linked Data and outline how these can be used to address the challenges faced and discussing theoretical and practical issues.



From Linked Open Data to Knowledge Graphs:



2013: Google adopts Semantic Web ideas under a new name

- Jamie Taylor, Google, Inc., Keynote [ISWC2017](#)

The Power of Knowledge Graph: Interlocking data

Destination suggestions: museum, city, airport

Reservation: airport, airline, flight

Flight Status: terminal, gate, status, leg

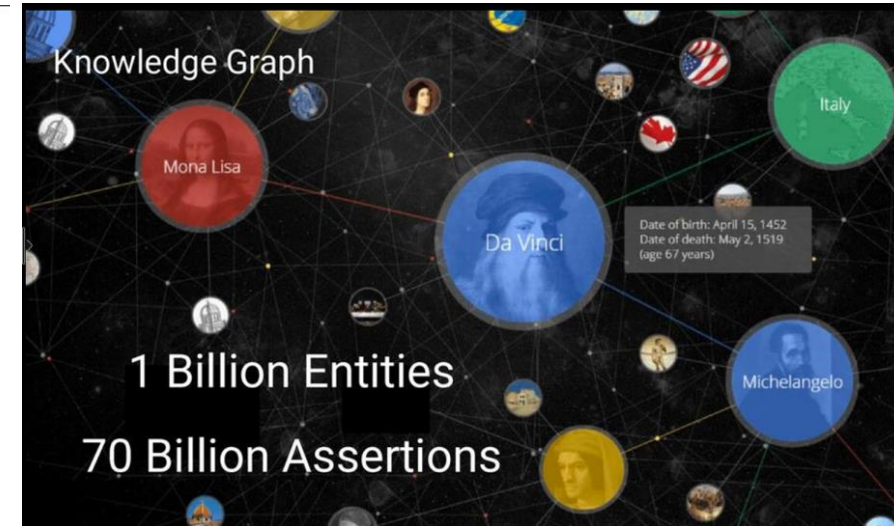
Google

ISWC 2017

Discuss live!

iswc2017.semanticweb.org/live/

<http://videolectures.net>

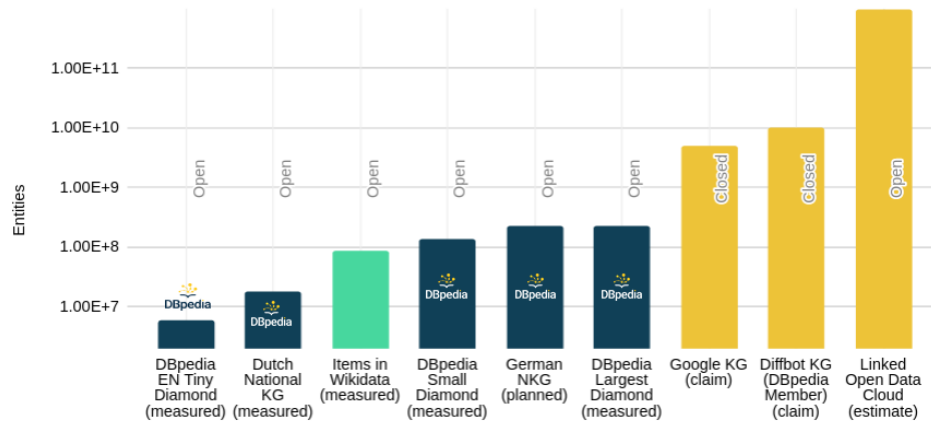


- Actors, Directors, Movies
- Art Works & Museums
- Cities & Countries
- Islands, Lakes, Lighthouses
- Music Albums & Music Groups
- Planets & Spacecraft
- Roller Coasters & Skyscrapers
- Sports Teams
- [...]

From Linked Open Data to Knowledge Graphs:

Success stories of mainly monolithic (but huge) Knowledge Graphs rather than a network of Linked small KGs:

<https://www.slideshare.net/Frank.van.Harmelen/adoption-of-knowledge-graphs-late-2019>



<https://www.dbpedia.org> 2021



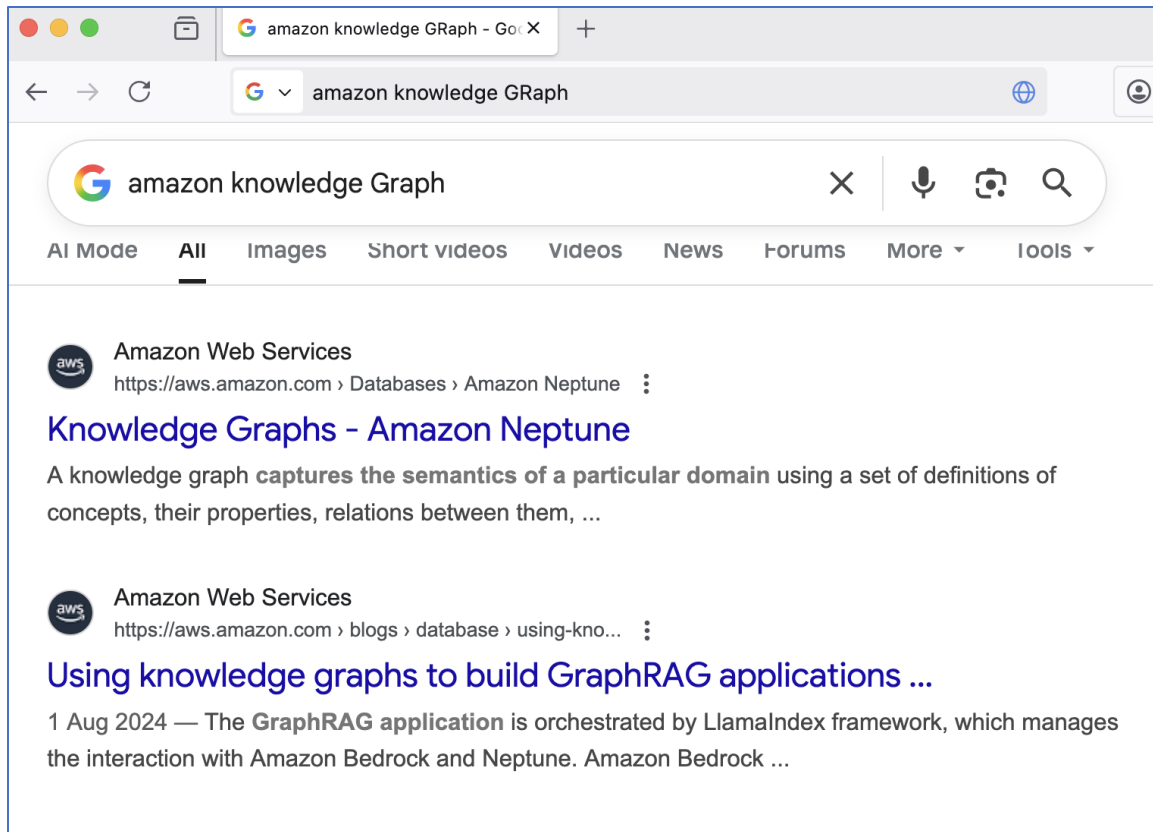
Open KGs (April 2021)	
DBpedia	~4.58m entities
Yago4	~50m entities
Wikidata	~93m entities

N. Noy, Y. Gao, A. Jain, A. Narayanan, A. Hogan et al.: Knowledge Graphs. Co.



It's not only Search – Obviously... User Recommendation, Language Technologies, ...

Since then, industry has become a main driver in scaling and using KG Technologies...



Collaborative, Open Knowledge Graphs:



DBpedia generates a graph from links and facts in Wikipedia's Infoboxes:

http://wikipedia.org/wiki/Zaha_Hadid

WIKIPEDIA
The Free Encyclopedia

Not logged in | Talk | Contributions | Create account | Log in

Article | Talk

Read | Edit | View history | Search Wikipedia

Zaha Hadid

From Wikipedia, the free encyclopedia

For the architectural firm, see *Zaha Hadid Architects*.

Dame Zaha Mohammad Hadid DBE RA (Arabic: زها حديد; born 31 October 1950 – 31 March 2016) was a *British-Iraqi* architect, artist and designer, recognised as a major figure in architecture of the late 20th and early 21st centuries. Born in *Baghdad, Iraq*, Hadid studied mathematics as an undergraduate and then enrolled at the *Architectural Association School of Architecture* in 1972. In search of an alternative system to traditional architectural drawing, and influenced by *Suprematism* and the *Russian avant-garde*, Hadid adopted painting as a design tool and abstraction as an investigative principle to "reinvestigate the aborted and untested experiments of Modernism [...] to unveil new fields of building".

She was described by *The Guardian* as the "Queen of the curve",^[1] who "liberated architectural geometry, giving it a whole new expressive identity".^[2] Her major works include the London Aquatics Centre for the 2012 Olympics, the Broad Art Museum, Rome's MAXXI Museum, and the Guangzhou Opera House.^[3] Some of her awards have been presented posthumously, including the statuette for the 2017 *Brit Awards*. Several of her buildings were still under construction at the time of her death, including the *Daxing International Airport* in Beijing, and the *Al Wakra Stadium* in Qatar, a venue for the 2022 FIFA World Cup.^[4] [5]

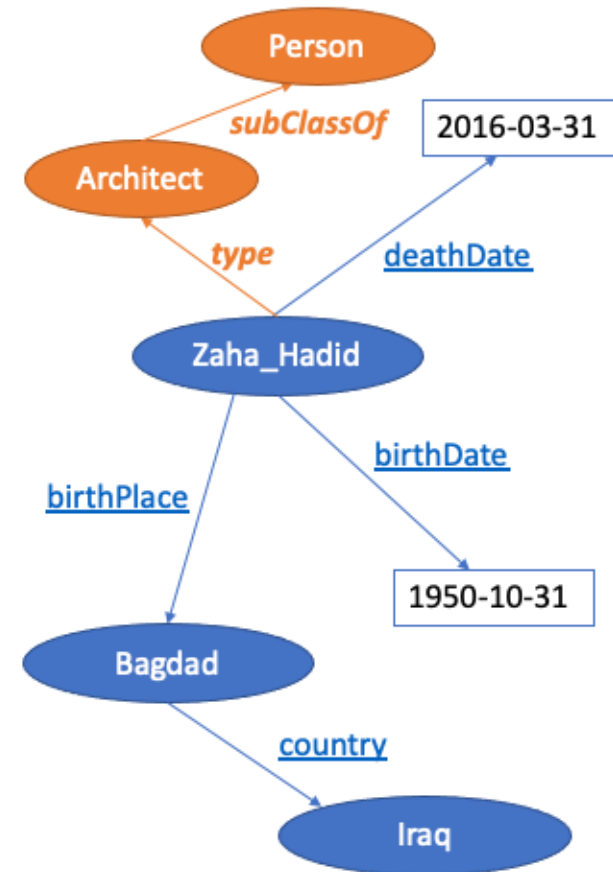
Hadid was the first woman to receive the *Pritzker Architecture Prize*, in 2004.^[6] She received the UK's most prestigious architectural award, the *Stirling Prize*, in 2010 and 2011. In 2012, she was made a *Dame of Elizabeth II* for services to architecture, and in February, 2016, the month preceding her death,^[7] she became the first woman

Dame Zaha Hadid DBE RA

Hadid at the Heydar Aliyev Cultural Center, November 2013

Born	Zaha Mohammad Hadid <div>31 October 1950</div> <div>Baghdad, Kingdom of Iraq</div>
Died	31 March 2016 (aged 65) <div>31 March 2016 (aged 65)</div> <div>Miami, Florida, U.S.</div>
Nationality	 Iraq, United Kingdom
Alma mater	American University of Beirut <div>Architectural Association School of Architecture</div>
Occupation	Architect
Parent(s)	Mohammed Hadid <div>Wajeeha Sabonji</div>
Practice	Zaha Hadid Architects
Buildings	<i>Vitra Fire Station</i>, <i>MAXXI</i>, <i>Bridge Pavilion</i>, <i>Contemporary Arts Center</i>, <i>Heydar Aliyev Center</i>, <i>Riverside Museum</i>
Website	www.zaha-hadid.com

http://dbpedia.org/resource/Zaha_Hadid




Collaborative, Open Knowledge Graphs:



Lionel Messi (Q615)

Argentine association football player

image	
occupation	<div>association football player</div> <div>▶ 2 references</div>
FIFA player ID (archived)	<div>229397</div> <div>▶ 1 reference</div>
country of citizenship	<div>Argentina</div> <div>start time1987</div> <div>▼ 0 references</div>
	<div>Spain</div> <div>start time2005</div> <div>▶ 1 reference</div>

Revision history of "Lionel Messi" (Q615)

[View logs for this item](#) ([view abuse log](#))

▼ Filter revisions

Diff selection: Mark the radio buttons of the revisions to compare and hit enter or the button at the bottom

Legend: **(cur)** = difference with latest revision, **(prev)** = difference with preceding revision, **m** = minor edit

(latest | [earliest](#)) View (newer 50 | older 50) (20 | 50 | 100 | 250 | 500)

Compare selected revisions

(cur | prev)

☒

08:56, 8 December 2024

দ্রাফ্‌ন গাজা (talk | contribs)

.. (537,664 bytes)

(+92)

..

(cur | prev)

☒

20:39, 7 December 2024

Sanremofilo (talk | contribs)

.. (537,572 bytes)

(+363)

messi/5663) (Tag: Wikidata user interface)

(cur | prev)

☐

15:21, 2 December 2024

Ytterbyz (talk | contribs)

.. (537,209 bytes)

(+349)

.. (Tag: Wikidata user interface)

(cur | prev)

☐

23:45, 29 November 2024

Mickey Đại Phát (talk | contribs)

.. (536,860 bytes)

(-)

Wikidata user interface, Mobile termbox)

(cur | prev)

☐

19:47, 27 November 2024

KrBot (talk | contribs)

.. (536,877 bytes)

(-14)

.. (See autofix на / on Property talk:P12924)

From Linked Open Data to Knowledge Graphs: What's the state of affairs?

- Jamie Taylor, <http://www.kylemiller.com/2017/02/20/linked-open-data-to-knowledge-graphs/>
- Large-scale, still data-focused (rather than schema-focused)

- Often monolithic, rather than linked/decentralised
- Knowledge extraction rather than Knowledge engineering

- Collaborative large-scale KGs:

- Collectively created (automated or curated)

- Notoriously incomplete

- (Logical) **consistency** not a must

- Actors, Directors, Movies

- Art Works & Museums

- Cities & Countries

- Islands, Lakes, Lighthouses

- Music Albums & Music Groups

- Planets & Spacecraft

- Roller Coasters & Skyscrapers

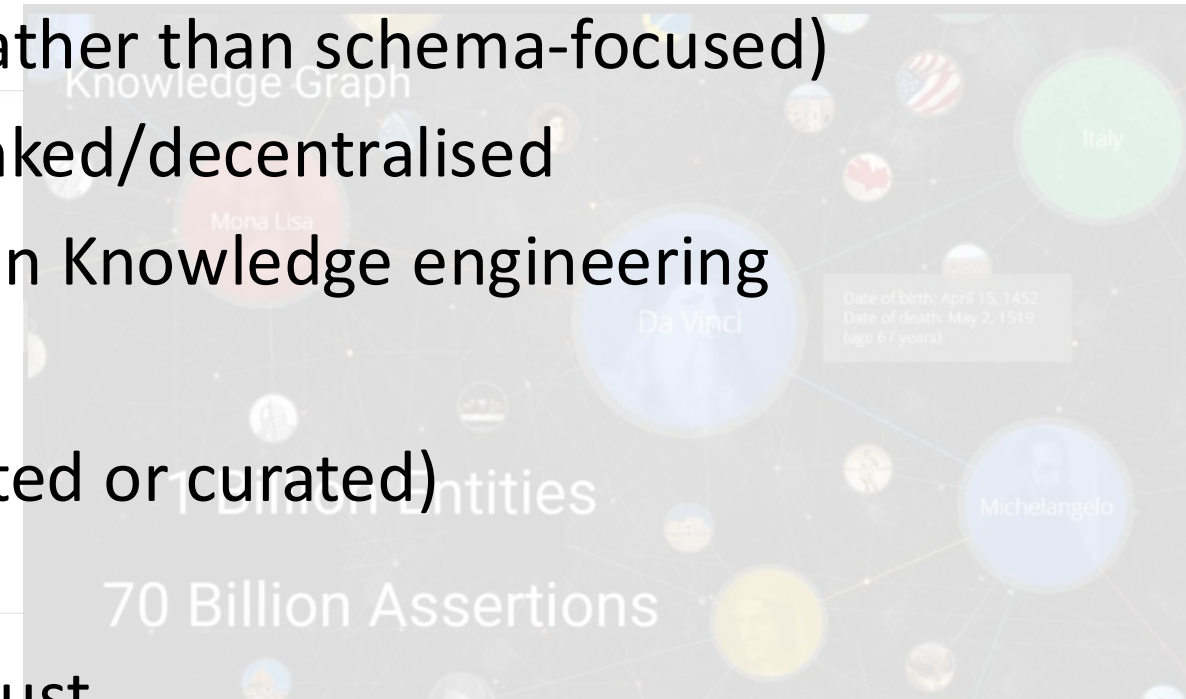
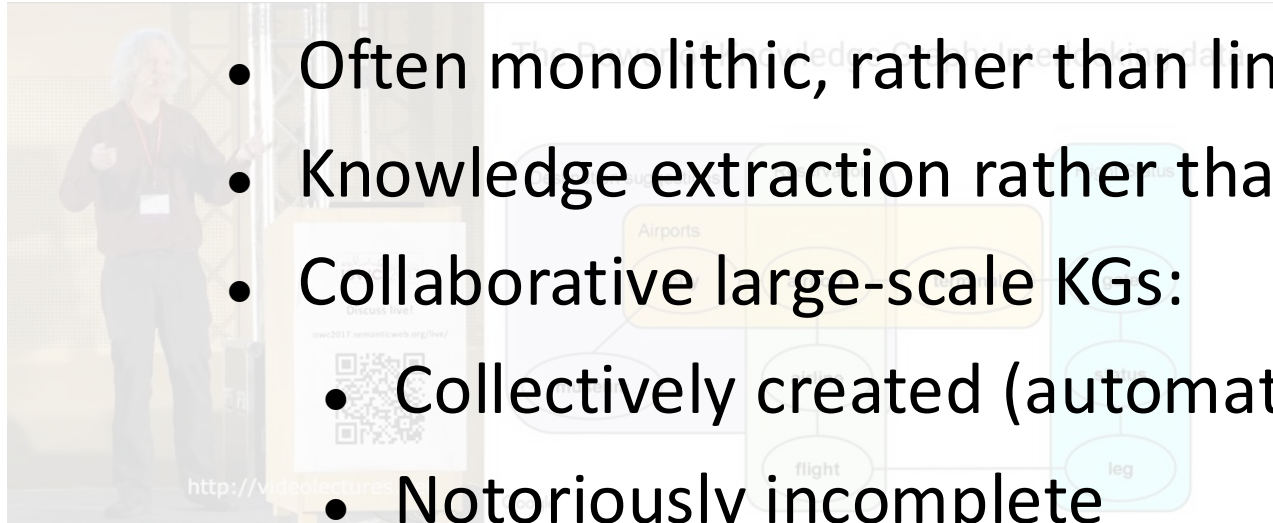
- Sports Teams

[...]

- Enterprise KGs: knowledge necessary to power applications
- *Ontological expressivity* not central – BUT: **Expressing context** is!

For instance:

- Provenance
- Temporal context



Let's have a look at practical examples of such collaboratively curated Knowledge Graphs:

• DBpedia (since 2007)

vs.

Wikidata (since 2012)

DBpedia



Developer(s)	Leipzig University University of Mannheim
Initial release	10 January 2007 (17 years ago)
Stable release	DBpedia 2016-10 / 4 July 2017
Repository	github.com/dbpedia/
Written in	Scala · Java
Type	Semantic Web · Linked Data
License	GNU General Public License
Website	dbpedia.org

- ✓ • RDF
- ✓ • Standard ontology language (OWL)
- ✓ • SPARQL endpoint
- ✗ • Consistent
- ✗ • Context

Wikidata

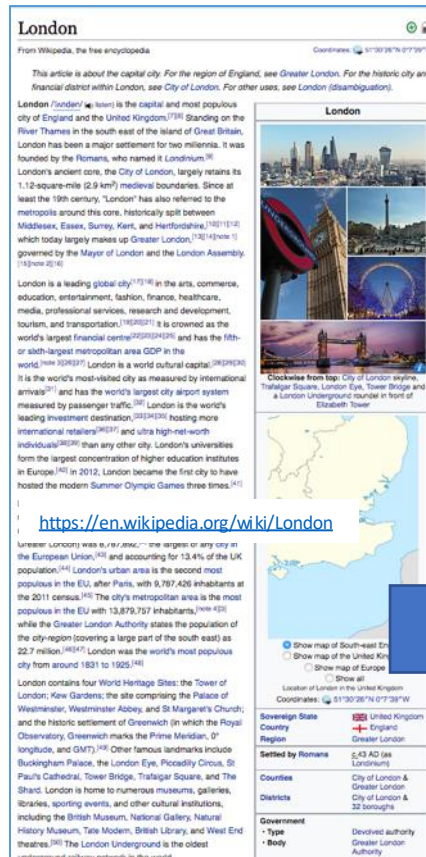


	Screenshot [show]
Type of site	Knowledge base · Wiki
Available in	Multiple languages
Owner	Wikimedia Foundation
Editor	Wikimedia community
URL	www.wikidata.org/wiki/Wikidata:Main_Page
Commercial	No
Registration	Optional
Launched	29 October 2012; 12 years ago ^[1]

- ✓
- ✗
- ✓
- ✗
- ✓

SPARQL: Using KGs to answer questions:

- E.g. from



Automatic
Extractors

- One of the central datasets of the Linked Open Data-Cloud
- RDF extracted from Wikipedia-Infoboxes
- You can use a language called SPARQL endpoint (roughly: SQL for RDF) to do **structured queries** over RDF:
 - „Cities in the UK with more than 1M population“:

Structured queries (SPARQL):

<https://api.triplydb.com/s/gZZskqRpQ>


```
PREFIX : <http://dbpedia.org/resource/>
PREFIX dbo: <http://dbpedia.org/ontology/>
PREFIX yago: <http://dbpedia.org/class/yago/>
```

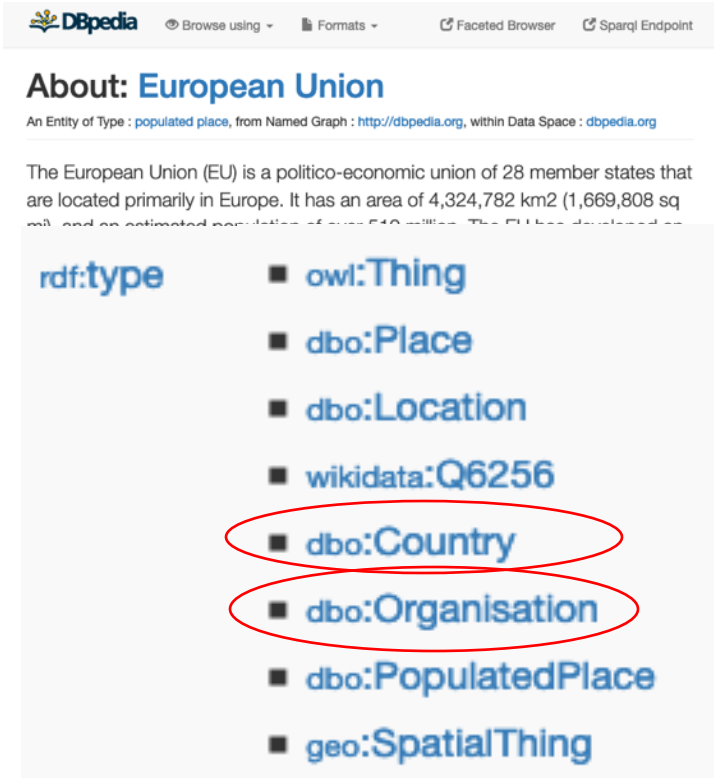
```
SELECT DISTINCT ?city ?pop WHERE {
  ?city a schema:City .
  ?city dbo:country :United_Kingdom.
  ?city dbo:populationTotal ?pop
```

```
  FILTER ( ?pop > 1000000 )
```

```
}
```


Knowledge Graphs like Dbpedia are not logically consistent! ☹ [1]

- E.g. 



DBpedia

About: **European Union**

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

The European Union (EU) is a politico-economic union of 28 member states that are located primarily in Europe. It has an area of 4,324,782 km² (1,669,808 sq mi) and an estimated population of over 540 million. The EU has developed a

rdf:type

- [owl:Thing](#)
- [dbo:Place](#)
- [dbo:Location](#)
- [wikidata:Q6256](#)
- [dbo:Country](#)
- [dbo:Organisation](#)
- [dbo:PopulatedPlace](#)
- [geo:SpatialThing](#)

Dbpedia Ontology:

`dbo:Agent owl:disjointWith dbo:Place.`


`dbo:Country rdfs:subClassOf dbo:Place.`

`dbo:Organisation rdfs:subClassOf dbo:Agent.`



1. Stefan Bischof, Markus Krötzsch, Axel Polleres, and Sebastian Rudolph. Schema-agnostic query rewriting in SPARQL 1.1. In *Proceedings of the 13th International Semantic Web Conference (ISWC 2014)*, Lecture Notes in Computer Science (LNCS). Springer, October 2014. [[.pdf](#)]

Wikidata is also not “consistent”, but doesn’t use OWL

image	
occupation	<div>association football player</div> <div>▸ 2 references</div>
FIFA player ID (archived)	<div>229397</div> <div>▸ 1 reference</div>
country of citizenship	<div>Argentina</div> <div>start time1987</div> <div>▼ 0 references</div>
	<div>Spain</div> <div>start time2005</div> <div>▸ 1 reference</div>

The same question as before in Wikidata:

Note: Wikidata does not even use standard OWL

- “Simple” surface [query](#):

Which cities in the UK have more than 1M people?

```
SELECT DISTINCT ?city WHERE {  
  ?city wdt:P31/wdt:P279* wd:Q515.  
  ?city wdt:P1082 ?population .  
  ?city wdt:P17 wd:Q38 .  
  FILTER (?population > 1000000) }
```

Note: Wikidata uses numeric IDs

city (Q515)
large and permanent human
settlement

population (P1082)
number of people inhabiting the
place; number of people of
subject

country (P17)
sovereign state of this item

United Kingdom (Q145)
country in Europe

instance of (P31)
that class of which this subject is
a particular example and
member. (Subject typically an
individual member with Proper
Name label.) Different from P279
(subclass of).

subclass of (P279)
all instances of these items are
instances of those items; this
item is a class (subset) of that
item. Not to be confused with
Property:P31 (instance of).

- What's this?

The same question as before in Wikidata:

<https://w.wiki/BqRX>

Which cities in the Austria have more than 1M/2M people?

```
SELECT DISTINCT ?City ?Pop
{
  ?City wdt:P17 wd:Q40;
        wdt:P31/wdt:P279* wd:Q515;
        wdt:P1082 ?Pop.
  FILTER (?Pop > 1000000)
  # note: Vienna historically had more than 2M inhabitants!
  # FILTER (?Pop > 2000000)
}
```

Note: Wikidata also has such contextual information!!!!

So, WHEN did Vienna have 2M inhabitants?

https://www.wikidata.org/wiki/Q1741

Item

Discussion

Vienna

(Q1741)

capital of and state in Austria

Wien | Vienna, Austria

population	<div><div></div><div>1,973,403</div></div>	
	point in time	1 October 2022
	determination method or standard	demographics
	<div>> 1 reference</div>	
	<div><div></div><div>2,083,630</div></div>	
	point in time	1910
	<div>< 0 references</div>	

The same question as before in Wikidata:

<https://w.wiki/BqRj>

Which cities in the Austria have more than 1M/2M people?

```
SELECT DISTINCT ?City ?Pop ?Timepoint
{
  ?City wdt:P17 wd:Q40;
        wdt:P31/wdt:P279* wd:Q515;
        p:P1082 ?Stmnt.
  ?Stmnt ps:P1082 ?Pop;
        pq:P585 ?Timepoint.
  # FILTER (?Pop > 1000000)
  # note: Vienna historically had more than 2M inhabitants!
  FILTER (?Pop > 2000000)
}
```

So, WHEN did Vienna have 2M inhabitants? Works!!!!

But needs an understanding of **Wikidata's proprietary RDF reification model** to model context!

See our recent ISWC2024 tutorial: <https://ww101.ai.wu.ac.at/>

Phase 4: Knowledge Graphs: “(Less Linked) Graph data with Metadata”

- Aka Labelled property graphs



Messi
(wd:Q615)



country of citizenship (wdt:P27)

{startTime: 1987}



Spain
(wd:Q29)



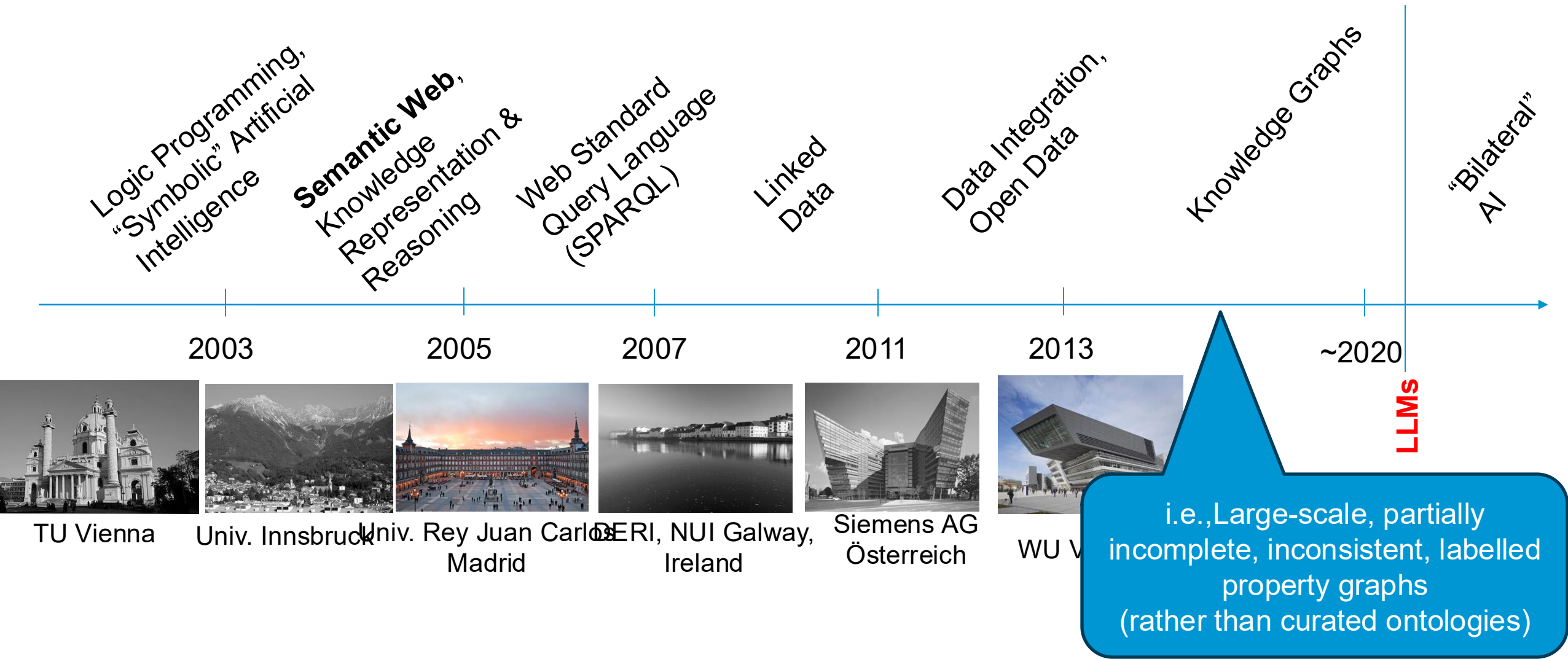
Argentina
(wd:Q414)

{startTime: 2005,
reference: https://cadenaser.com/ser/2005/09/26/deportes/1127690892_850215.html}

country of citizenship (wdt:P27)

country of citizenship	Argentina	
	start time	1987
	0 references	
Spain		
	start time	2005
	1 reference	

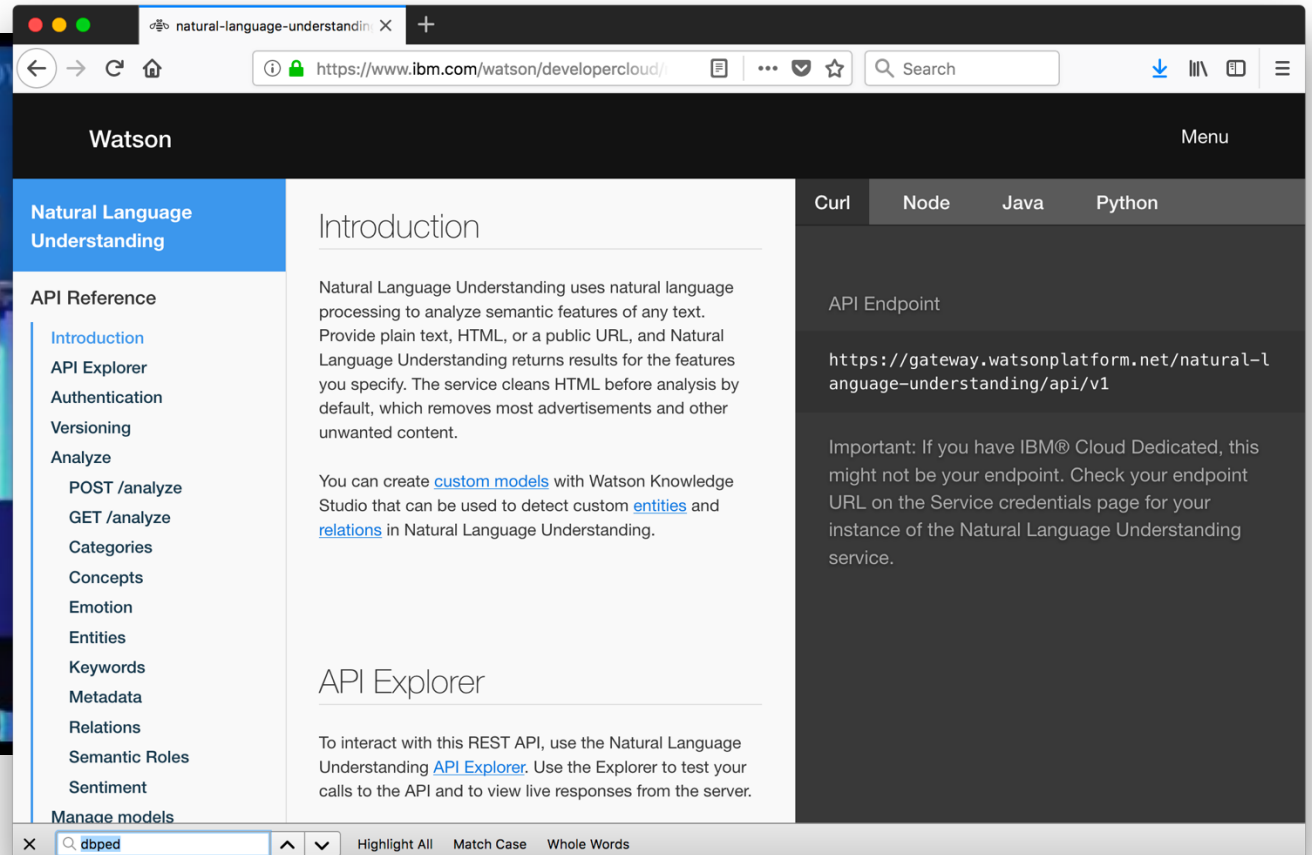
So, what are these KGs actually good for in the age of LLMs



So, for what are these KGs actually good for in the age of LLMs and AI?

- Interesting Note - IBM Watson - Jeopardy! (2011)

“Super-human” Question Answering was achieved by Knowledge Graphs before the LLM hype!



<https://youtu.be/P0Obm0DBvwl?t=951>

Some of our own research in this area:



- How good or bad can KGs deal with Question answering?
 - *By essentially translating Questions to SPARQL queries (Watson)?*
 - *By other bespoke techniques such as Message passing:*

Svitlana Vakulenko, Javier Fernández, Axel Polleres, Maarten de Rijke, and Michael Cochez. *Message passing for complex question answering over knowledge graphs. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM2019, pages 1431--1440, Beijing, China, November 2019. ACM.*

Some of our own research in this area:

- How good or bad can KGs deal with Question answering?
 - *By essentially translating Questions to SPARQL queries (Watson)?*
 - *By other bespoke techniques such as Message passing:*

Svitlana Vakulenko, Javier Fernández, Axel Polleres, Maarten de Rijke, and Michael Cochez. *Message passing for complex question answering over knowledge graphs. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM2019, pages 1431--1440, Beijing, China, November 2019. ACM.*


- How good or bad are LLMs with Question answering compared with KGs?

Gerhard Georg Klager and Axel Polleres. *Is GPT fit for KGQA? -- preliminary results. In Proceedings of the International Workshop on Knowledge Graph Generation from Text (Text2KG2023), co-located with Extended Semantic Web Conference 2023 (ESWC 2023), May 2023.*


Admittedly things have improved dramatically, since then!

So, for what are these KGs actually good for in the age of LLMs and AI?

ChatGPT ▾

Share 


Which cities in Austria have more than 2M inhabitants?



No city in Austria has more than 2 million inhabitants. Austria's capital and largest city, **Vienna (Wien)**, has a population of approximately **1.9 million people** as of recent estimates. It is the only city in Austria with a population exceeding 1 million. Other cities, such as Graz, Linz, and Salzburg, have much smaller populations, typically below 300,000.






If you're interested in any specific details about Austrian cities, feel free to ask!

Did Vienna ever have more than 2M inhabitants?



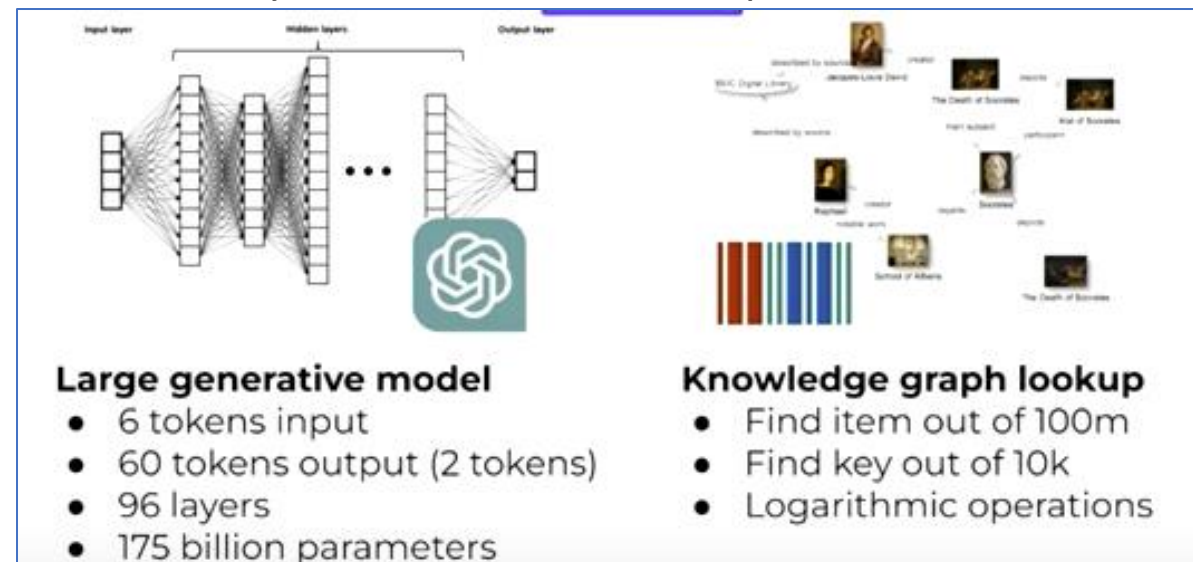
Yes, Vienna once had a population exceeding 2 million. During the early 20th century, particularly around the year **1910**, Vienna's population peaked at approximately **2.1 million people**. This was during the Austro-Hungarian Empire, when Vienna was a major imperial capital and one of the largest cities in Europe.

After World War I and the dissolution of the Austro-Hungarian Empire in 1918, Vienna's population began to decline. This was due to several factors, including changes in national borders, economic challenges, and shifts in urbanization trends. The city's population has been growing again in recent decades but has not yet surpassed 2 million.

For instance:

- Explainable answers, fact-checking against hallucinating
- Potentially less resource consumption!



- KGC23 Keynote: “The Future of Knowledge Graphs in a V LLMs — Denny Vrandečić, Wikimedia”

<https://www.youtube.com/watch?v=ww99npDh4cg>



What's good for what?

LLMs, Search Engines, KGs



Large Language Models, Knowledge Graphs and Search Engines: A Crossroads for Answering Users' Questions

Aidan Hogan
ahogan@dcc.uchile.cl
IMFD; DCC, Universidad de Chile
Santiago, Chile

Denny Vrandečić
denny@wikimedia.org
Wikimedia Foundation
San Francisco, California, USA

Xin Luna Dong
lunadong@meta.com
Meta Reality Labs
Redmond, Washington, USA

Gerhard Weikum
weikum@mpi-inf.mpg.de
Max Planck Institute for Informatics
Saarbrücken, Germany

SEARCH ENGINES ASSUME BOTH HUMANS AND MACHINES ARE STUPID

∴ WIDELY USED AND WIDELY USEFUL



KNOWLEDGE GRAPHS ASSUME MACHINES ARE STUPID AND HUMANS SMART

∴ BARELY USED AND BARELY USEFUL



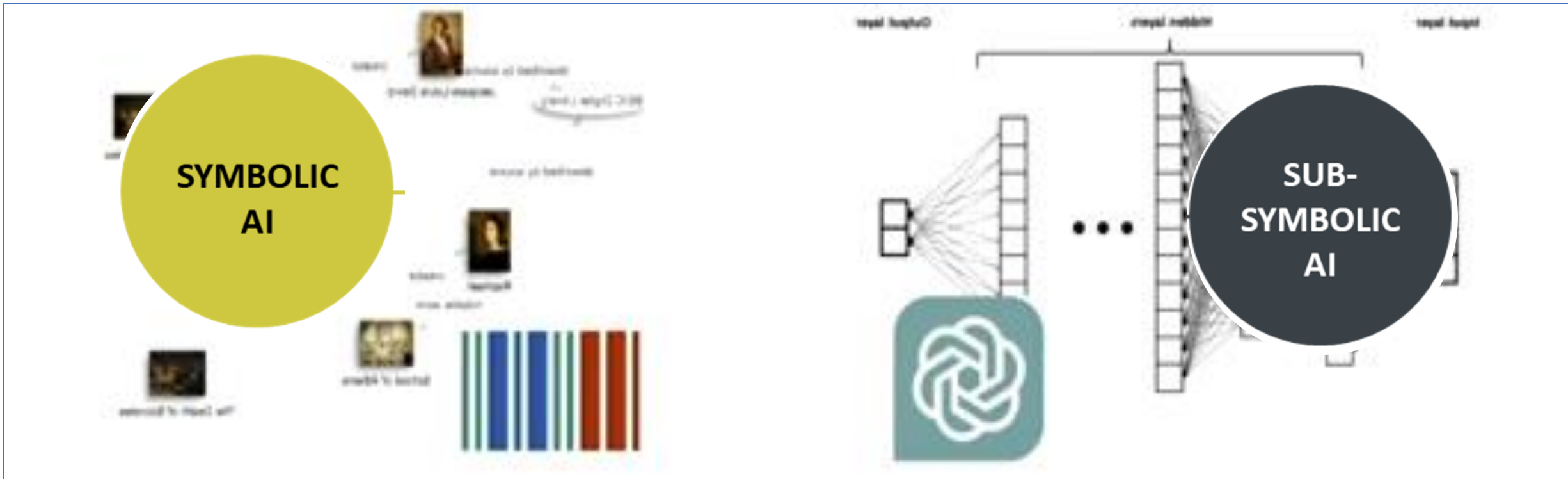
LARGE LANGUAGE MODELS ASSUME MACHINES ARE SMART AND HUMANS STUPID

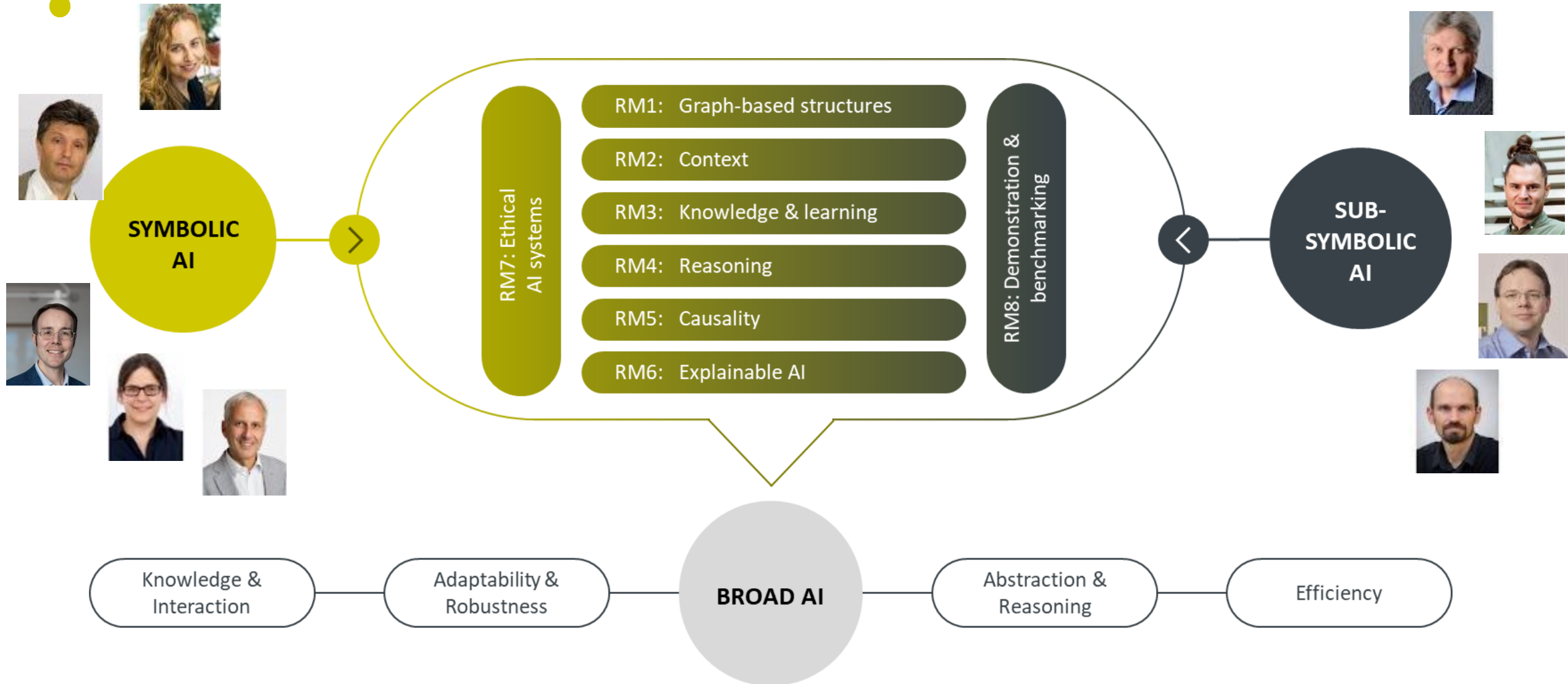
∴ WIDELY USED AND BARELY USEFUL



Dimension	SE	KG	LM
Precision	+ stores corpus - noisy content	+ stores corpus + precise operators	- abstracts corpus - hallucinations
Coverage	+ broad coverage	+ domain specific - patchy coverage	+ broad coverage - poor long tail
Freshness	+ quick updates + news often text	+ deprecation - structure lag	- slow updates - cold start
Generation	- no generation	+ ontologies/rule + graph learning	+ text generation
Synthesis	- no integration - no synthesis	+ data integration + synthesis	+ text integration + synthesis
Transparency	+ provenance - opaque ranks	+ algorithmic + provenance	- black box - no provenance
Determinism	+ deterministic	+ deterministic	- unstable results - randomness
Curation	+ curatable - opaque ranks	+ curatable	- indirect curation - unpredictable
Fairness	+ content as-is - bias in ranks	+ content as-is - bias in coverage	- generative biases - needs safeguards
Usability	+ natural language + simple queries	- structured - complex queries	+ natural language + conversational
Expressivity	- simple queries - ambiguity	+ complex queries - lacks nuance	+ complex queries - ambiguity
Efficiency	+ retrieval-based + simple queries	+ retrieval-based - complex queries	- inference-based - costly training
Multilingual	- lang. dependent	+ lang. agnostic - manual labels	+ multilinguality - variable results
Context	- limited context - not interactive	- limited context - not personalizable	+ in-context learn. + personalizable

So, let us better combine KGs and LLMs!





Bilateral AI

JKU



Martina Seidl
Symbolic AI
SAT Solving
Formal methods



Sepp Hochreiter
Machine Learning
LSTM
Vanishing gradient

- Institute for Machine Learning
- ELLIS Unit Linz
- LIT AI Lab
- Institute for Symbolic Artificial Intelligence

UNIVERSITÄT KLAGENFURT



Gerhard Friedrich
Symbolic AI
Model-based reasoning

- Institute for Artificial Intelligence and Cybersecurity

TU Graz



Robert Legenstein
Machine Learning
Computational Neuroscience

- Institute of Machine Learning and Neural Computation

WU VIENNA



Axel Polleres
Knowledge Graphs

- Institute for Data Process and Knowledge Management

ISTA

Institute of
Science and
Technology
Austria



Christoph Lampert
Machine Learning
Trustworthy Learning

- Machine Learning and Computer Vision group
- ELLIS Unit ISTA

TU WIEN

TECHNISCHE
UNIVERSITÄT
WIEN



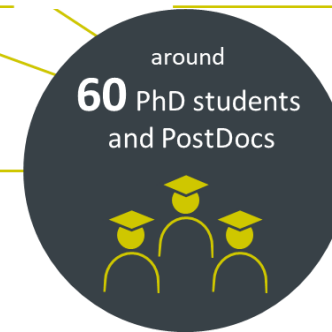
Agata Ciabattoni
Logic Reasoning



Thomas Eiter
Symbolic AI
Knowledge representation

- Institute for Logic and Computation

FWF Österreichischer
Wissenschaftsfonds



BILATERAL AI – First results and plans from our Institute and other BILAI partners...



What's next?



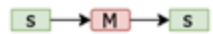
Neurosymbolic AI Systems
Prof. Marta Sabou



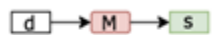
Ingredient1: We have a pretty good understanding of Engineering **Neuro-Symbolic Systems...**

Atomic-Patterns

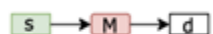
A1: s-M-s



A2: d-M-s

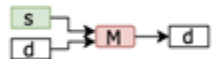


A3: s-M-d

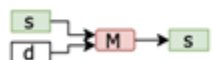


Fusion-Patterns

F1: d/s-M-d



F2: d/s-M-s

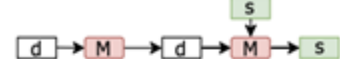


F3: d/s-M-d/s

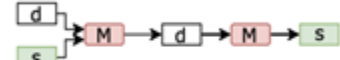


T-Patterns

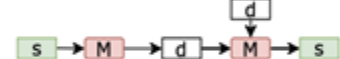
T1: {d-M-d/s}-M-s



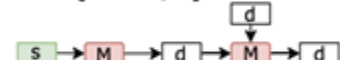
T2: {d/s}-M-d-M-s



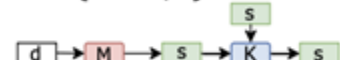
T3: {s-M-d/d}-M-s



T4: {s-M-d/d}-M-d



T5: {d-M-s/s}-K-s

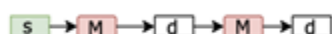


I-Patterns

I1: s-M-d-M-s

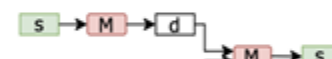


I2: s-M-d-M-d

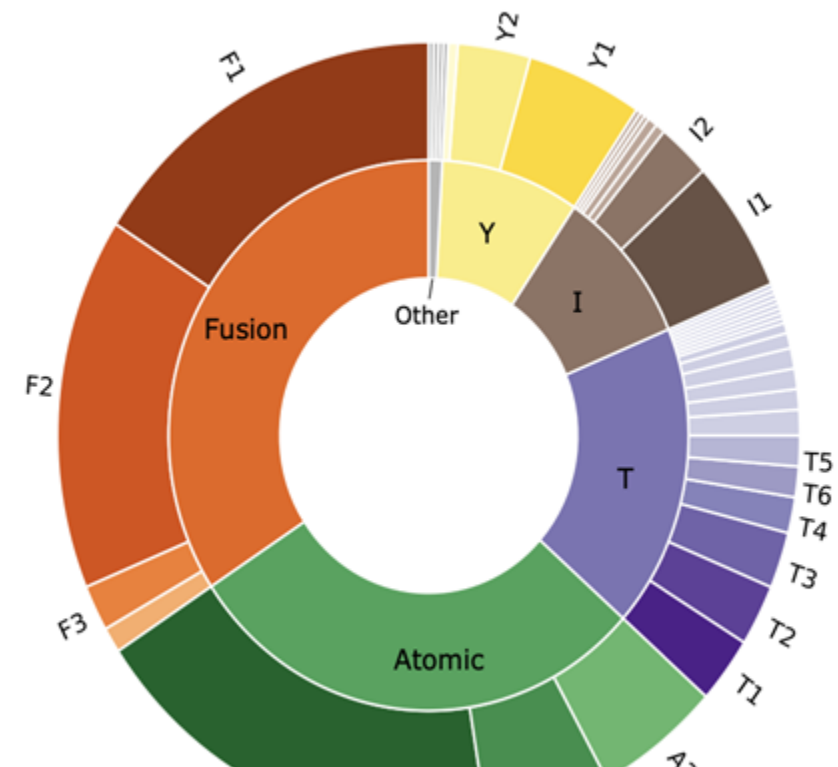
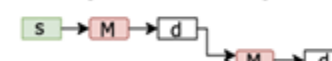


Y-Patterns

Y1: {s-M-d/d-M-d}-M-s

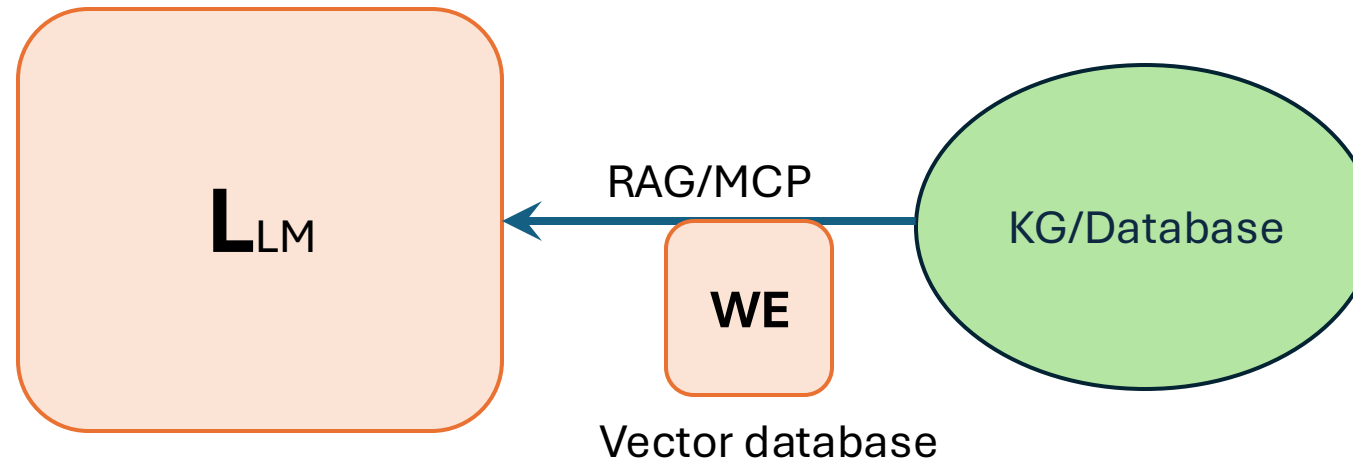


Y2: {s-M-d/d-M-d}-M-d



A. Breit, L. Waltersdorfer, F.J. Ekaputra, M. Sabou, A. Ekelhart, A. Iana, H. Paulheim, J. Portisch, A. Revenko, A. ten Teije, and F. van Harmelen. 2023. Combining Machine Learning and Semantic Web: A Systematic Mapping Study. ACM Computing Survey. March 2023.

One such architecture... For instance:



LLMs get better ... but still (plausibly) hallucinate “are supposed to be PhD level”



AP Please give me the top 10 publications by Axel Polleres

I'll search for information about Axel Polleres' top publications.

Axel Polleres top publications 10 results

Let me fetch his Google Scholar page to see his most cited publications.

Failed to fetch <https://scholar.google.com/citations?user=R-SmVOYAAAAJ&hl=en>

Axel Polleres most cited papers SPARQL RDF 10 results

Let me fetch his DBLP page which should have a comprehensive list of his publications.

Fetches Axel Polleres dblp.org

Based on Axel Polleres' extensive publication record with over 18,778 citations, here are his top 10 most influential publications:

Top 10 Publications by Axel Polleres

1. **"Knowledge Graphs"** (2021/2022) Co-authored with Aidan Hogan, Eva Blomqvist, and others, published in ACM Computing Surveys and as a book in the Synthesis Lectures series [dblp: Axel Polleres](#) ok
2. **"A survey of RDF stores & SPARQL engines for querying knowledge graphs"** (2022) Published in The VLDB Journal [A survey of RDF stores & S...](#) !
3. **"From SPARQL to rules (and back)"** (2007) A highly cited paper on the relationship between SPARQL and rule systems [ResearchGate](#) [ACM Digital Library](#) ok
4. **"Evaluating query and storage strategies for RDF archives"** (2019) Co-authored with Javier D. Fernández, Jürgen Umbrich, and Magnus Knuth [dblp: Axel Polleres](#) ok
5. **"Binary RDF representation for publication and exchange (HDT)"** (2013) Published in Journal of Web Semantics, introducing the HDT format for RDF compression [dblp: Axel Polleres](#) ok
6. **"A fine-grained evaluation of SPARQL endpoint federation systems"** (2016) Co-authored with Muhammad Saleem and others, evaluating federated SPARQL query systems [A fine-grained evaluation o...](#) !
7. **"Federating queries in SPARQL 1.1: Syntax, semantics and evaluation"** (2013) Co-authored with Carlos Buil-Aranda, Marcelo Arenas, and Óscar Corcho [dblp: Axel Polleres](#) ok
8. **"RDFS and OWL Reasoning for Linked Data"** (2013) Published in Reasoning Web, co-authored with Aidan Hogan, Renaud Delbru, and Jürgen Umbrich [dblp: Axel Polleres](#) ok
9. **"Mapping between RDF and XML with XSPARQL"** (2012) Co-authored with Stefan Bischof, Stefan Decker, Thomas Krennwallner, and Nuno Lopes [dblp: Axel Polleres](#) ok
10. **"Everything you always wanted to know about blank nodes"** (2014) Published in Journal of Web Semantics with Aidan Hogan, Marcelo Arenas, and Alejandro Mallea ok

Here are 10 influential publications by Axel Polleres, including papers, book-chapters, and working documents. These are among his most cited or otherwise important contributions. If you want, I can also fetch full download links or more recent ones.

Top 10 Publications by Axel Polleres

#	Title	Co-authors / Where / When	
1	Web Service Modeling Ontology (WSMO)	Jos de Bruijn, Dieter Fensel, Uwe Keller, Holger Lausen, Rubén Lara, Michael Kifer, Martin Hepp, Birgitta König-Ries, Jacek Kopecky, Dumitru Roman, John Domingue, Eyal Oren + Axel Polleres etc. (W3C Member Submission, 2005) W3C +3	ok
2	Binary RDF representation for publication and exchange (HDT)	Javier D. Fernández, Miguel A. Martínez-Prieto, Claudio Gutiérrez, Axel Polleres, Mario Arias (Journal of Web Semantics, 2013) repositorio.latinno... +4	ok
3	From SPARQL to Rules (and back)	Axel Polleres (WWW 2007) dl.acm.org +2	ok
4	SPARQL1.1: New Features and Friends (OWL2, RIF)	Axel Polleres (in Web Reasoning and Rule Systems (RR 2010), Lecture Notes in Computer Science) link.springer.com +2	ok
5	A General Framework for Representing, Reasoning and Querying with Annotated Semantic Web Data	Antoine Zimmermann, Nuno Lopes, Axel Polleres, Umberto Straccia (2011) arxiv.org +1	ok
6	Are We Better Off With Just One Ontology on the Web?	Armin Haller, Axel Polleres (Semantic Web Journal, 2020) WU Wirtschafts... +2	ok
7	Evaluating Query and Storage Strategies for RDF Archives	Javier David Fernández García, Jürgen Umbrich, Axel Polleres, Magnus Knuth (Date ~ 2015) WU Wirtschafts... +1	ok
8	Serializing RDF in Compressed Space (HDT++)	Antonio Hernández-Ilerra, Miguel A. Martínez-Prieto, Javier David Fernández, Axel Polleres WU Wirtschafts... +1	!
9	How (Well) Do Datalog, SPARQL and RIF Interplay?	Axel Polleres (in Datalog 2.0 2012, LNCS) link.springer.com +1	ok
10	Semantic Web Languages and Semantic Web Services as Application Areas for Answer Set Programming	Axel Polleres (Dagstuhl Seminar Proceedings, 2005) drops.dagstuhl.de +1	ok



Inspired by Georg Gottlob's SEMANTiCS2025 keynote

Heng Ji's SEMANTiCS2025 keynote

“Larger models are not necessarily better!”

But are KGs (& Databases) actually better?

Google asks it's
users to manually
correct/merge
results

DBLP or scopus
attempt to curate
ambiguities by
constraints (not always
successful..)

Axel Polleres

Vienna University of Economics and Business (WU - Wirtschaftsuniversität Wien)
Verified email at wu.ac.at - [Homepage](#)

[Data & Knowledge Manage...](#) [Artificial Intelligence](#) [Logic Programming](#) [Semantic Web](#)
[Knowledge Graphs](#)

FOLLOW

- ☐ **TITLE**
- ☐ **Knowledge graphs**
A Hogan, E Blomqvist, M Cochez, C d'Amato
ACM Computing Surveys (CSUR) 54 (4), 1-37, 2021
- ☐ **Web service modeling ontology**
D Roman, U Keller, H Lausen, J De Bruijn, M Stollberg
Applied ontology 1 (1), 77-106
- ☐ **Enabling semantic web services: the**
D Fensel, H Lausen, J de Bruijn, M Stollberg
Springer Berlin Heidelberg
- ☐ **Web service modeling ontology (wsr)**
J De Bruijn, C Bussler, J Domingue, D Fensel
W3C member submission
- ☐ **The web service modeling language**
J De Bruijn, H Lausen, A Polleres, D Fensel
European semantic web conference, 590-604

Merged citations		
This "Cited by" count includes citations to the following articles in Scholar. The ones marked * may be different from the article in the profile.		
Knowledge graphs		2591
A Hogan, E Blomqvist, M Cochez, C d'Amato, GD Melo, C Gutierrez, ... ACM Computing Surveys (Csur) 54 (4), 1-37, 2021		
Anisa Rula, Lukas Schmelzeisen, Juan F. Sequeda, Steffen Staab, and Antoine Zimmermann		344 *
A Hogan, E Blomqvist, M Cochez, C d'Amato, G De Melo, C Gutierrez, ... Knowledge graphs, 2020		
Knowledge graphs		21 *
H Aidan, E Blomqvist, M Cochez, C d'Amato, G de Melo, C Gutiérrez, ... Mines Saint-Etienne, 2020		
Knowledge graphs. CoRR abs/2003.02320 (2020)		13 *
A Hogan, E Blomqvist, M Cochez, C d'Amato, G de Melo, C Gutiérrez, ... URL: https://arxiv.org/abs/2003.02320 , 2003		

Search dblp

powered by CompleteSearch, courtesy of Hannah Bast, University of Frankfurt

> Home

Author search results

Exact matches

- **Thomas Müller** — [disambiguation page](#)
- **Thomas Müller** 0001
TU Munich, Faculty of Computer Science, Germany
- **Thomas Müller** 0002
Eastphalia University of Applied Sciences, Braunschweig/Wolfsbüttel, Germany
- **Thomas Müller** 0003
Friedrich Schiller University Jena, Institute of Computer Science, Germany
- **Thomas Müller** 0004
Fraunhofer Institute for Optoelectronics, System Technologies and Image Exploitation (IOSB), Karlsruhe, Germany
- [show all](#)

Likely matches

- **Thomas Müller** 0013
NVIDIA Research, Zürich, Switzerland
- **Thomas Müller-Gronbach**
University of Passau, Germany
- **Thomas Müller** 0009
Google Research
- **Thomas Müller** 0007
Department of Philosophy, University of Konstanz, Germany



But are KGs actually better?

Example from Wikidata...



Tilikum (Q127236)

[Item](#) [Discussion](#)

[Read](#) [View history](#) [☆](#)

dugout canoe

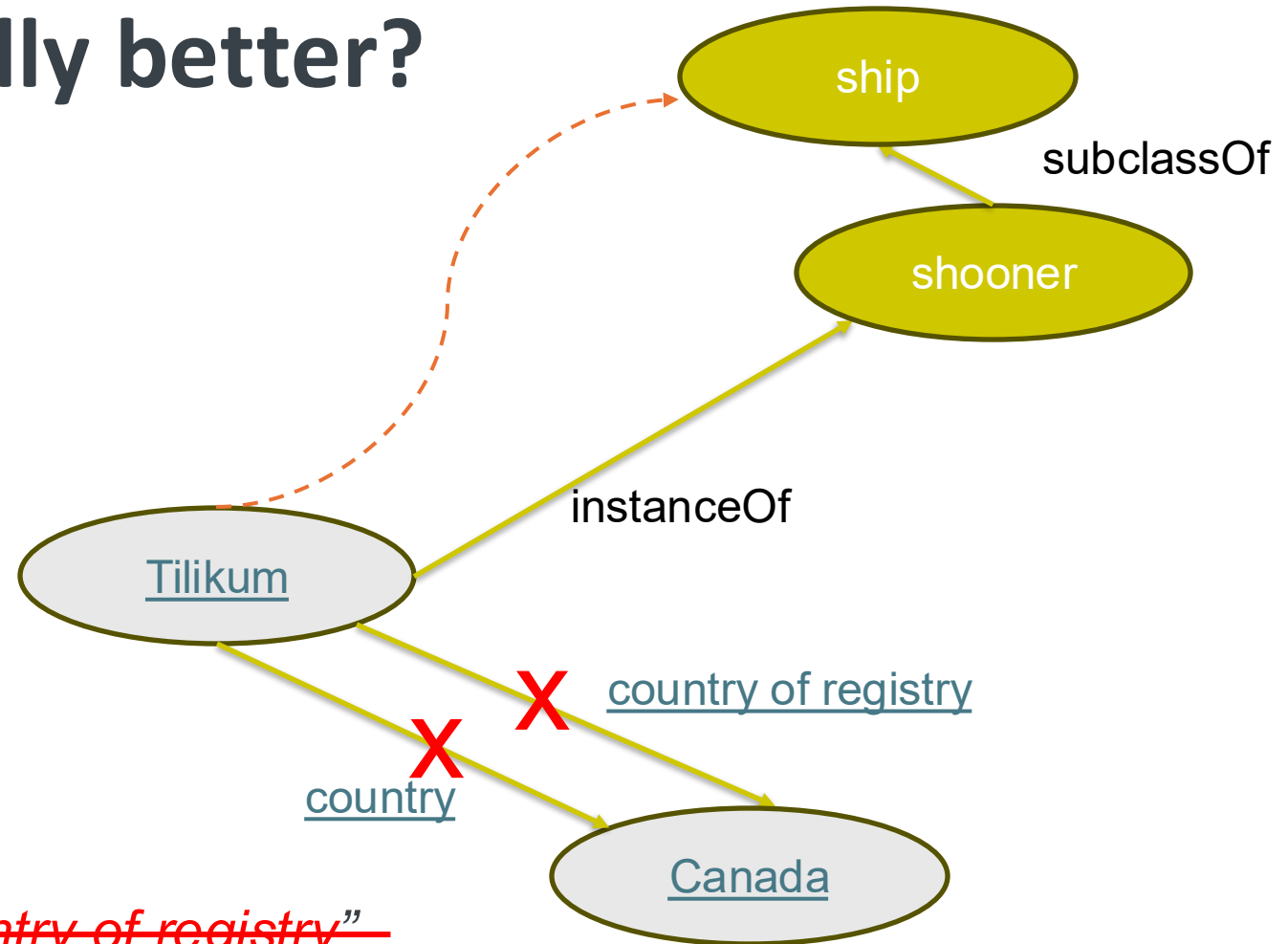
[edit](#)

▼ [In more languages](#)

[Configure](#)

Language	Label	Description	Also known as
default for all languages	No label defined	—	
English	Tilikum	dugout canoe	
British English	Tilikum	No description defined	
American English	No label defined	No description defined	
German	Tilikum	Segelboot	

But are KGs actually better?



Ontologies + Constraints,
such as:

~~“country conflicts with country of registry”~~

- “i.e., an entity should not have both a “country” and a “country of registry”

Why KGs can't (scalably) be correct:

Collaborative Editing

Multi-sourced Automatic Extraction

Knowledge is distributed and evolving!

- *Remote changes affect consistency*
- *Ontologies & Constraints evolve as well*

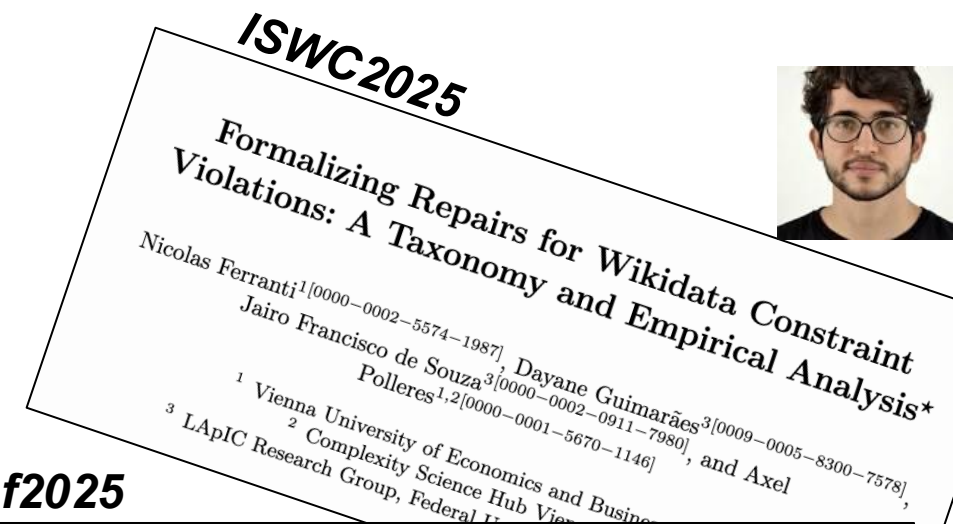
What can we do about it?

- Best of both worlds needs **Bilateral AI** approaches!

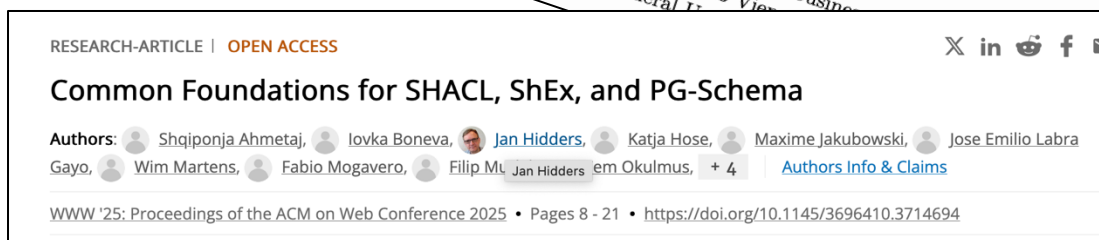
- **Symbolic** methods to repair inconsistencies

Forthcoming...

- **Subsymbolic** methods to resolve inconsistencies
 - **LLMs & KG embeddings** for resolving inconsistencies
 - **GNNs** for *learning* repairs (from historical data, from users' repairs)
 - Iterative **RAG** and **Agent-based** pipelines to fix and construct KGs from text



Webconf2025

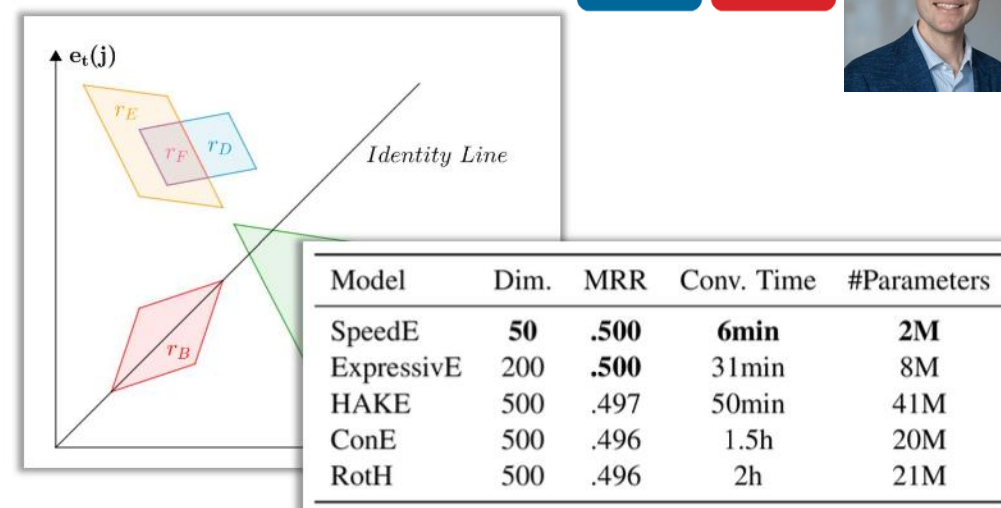


Further starting points in BILAI:

Notably, in BILAI, colleagues from TU Vienna (Sallinger, Pavlovic) work on KG Embeddings that can capture (certain) rules and constraints already:

- Similar to Word embeddings and LLMs, **Knowledge Graph Embeddings** allow to
 - predict missing edges in incomplete KGs
 - predict inconsistencies & possible repairs

Slide: Emanuel Sallinger

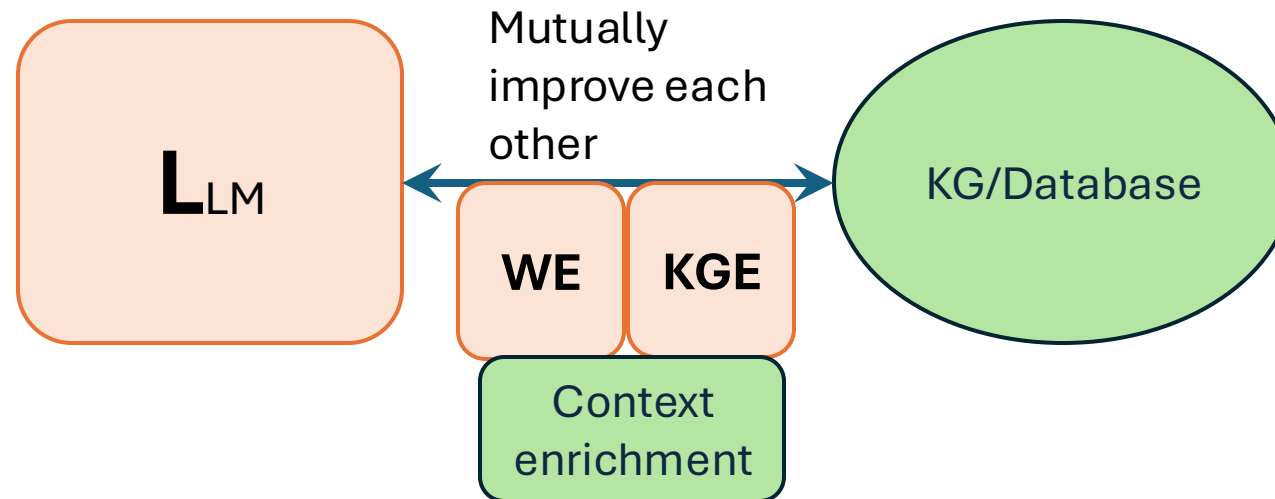


Open Problems:

- Scaling KG Embeddings to full, decentralized KGs ...
- ... but (1) **modularization** might help here, relation to the
- (2) corresponding trend to LLMs-based “multi-agent frameworks”

Logical Rule	ExpressivE	RotatE	TransE	BoxE	ComplEx	DistMult
Symmetry: $r_1(X, Y) \Rightarrow r_1(Y, X)$	✓	✓	✗	✓	✓	✓
Anti-symmetry: $r_1(X, Y) \Rightarrow \neg r_1(Y, X)$	✓	✓	✓	✓	✓	✗
Inversion: $r_1(X, Y) \Leftrightarrow r_2(Y, X)$	✓	✓	✓	✓	✓	✗
Comp. def.: $r_1(X, Y) \wedge r_2(Y, Z) \Leftrightarrow r_3(X, Z)$	✓	✓	✓	✗	✗	✗
Gen. comp.: $r_1(X, Y) \wedge r_2(Y, Z) \Rightarrow r_3(X, Z)$	✓	✗	✗	✗	✗	✗
Hierarchy: $r_1(X, Y) \Rightarrow r_2(X, Y)$	✓	✗	✗	✓	✓	✓
Intersection: $r_1(X, Y) \wedge r_2(X, Y) \Rightarrow r_3(X, Y)$	✓	✓	✓	✓	✗	✗
Mutual exclusion: $r_1(X, Y) \wedge r_2(X, Y) \Rightarrow \perp$	✓	✓	✓	✓	✓	✓

So we actually need this?



...

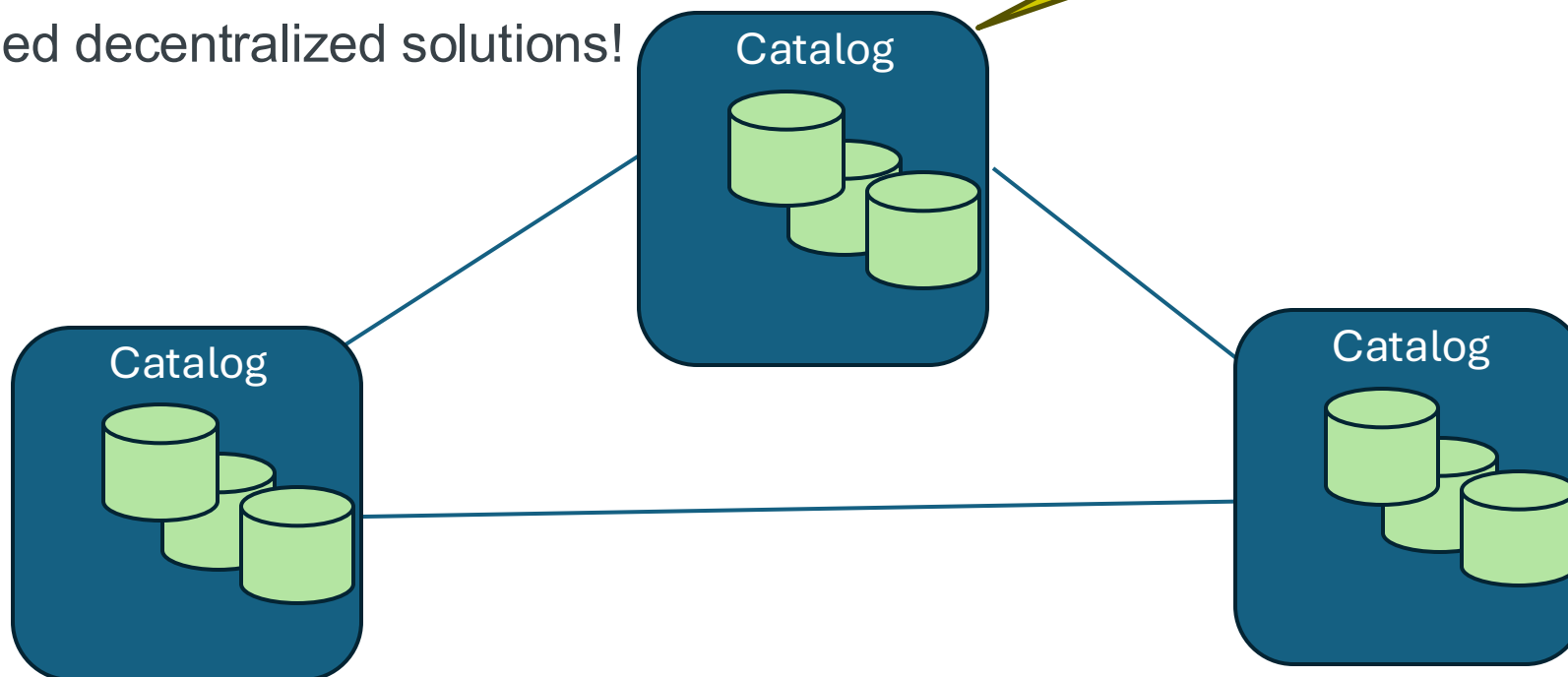
Quite some work at current ISWC2025, e.g. RAGE-KG workshop is looking into such extensions...

What's next? Where should this go?

Caveat: Our data and knowledge is typically:

- **Distributed/decentralized** and
- **not always sharable** with a central (API) model
- → we need decentralized solutions!

“Our knowledge will never go into a central LLM” – BILAI industry network keynote Stefan Rohringer (Infineon)



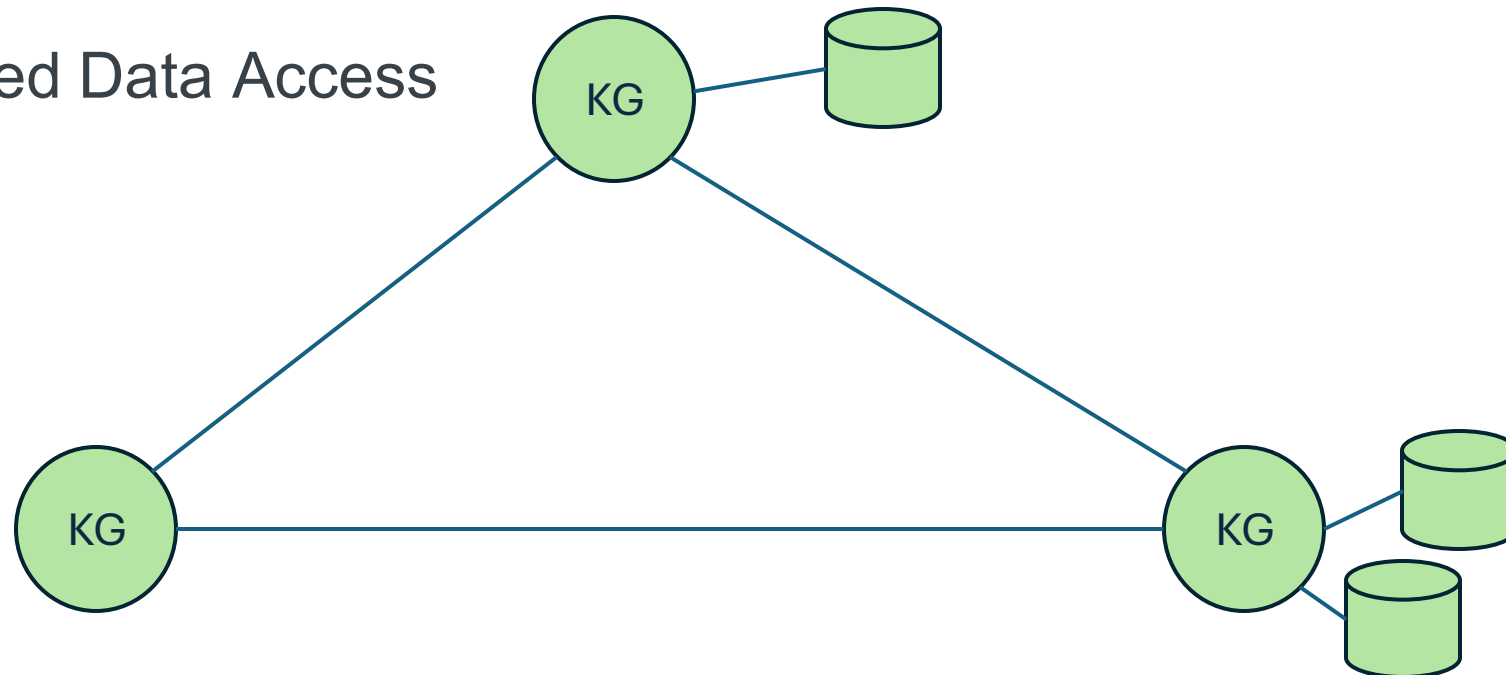
Ingredient 4: “Data Spaces”



What's next? Where should this go?

Ingredient 2: Expertise on (Symbolic) Data Integration and Repair. Before the LLM hype, we long worked on *purely symbolic* solutions to integrating *decentralised* Knowledge

- Linked Data
- Ontology-based Data Access



Scalable data access to decentralized KGs:



- Starting Points:
 - graph partitioning and splitting processing between clients and SPARQL endpoints the performance, the problems of central query endpoints can be significantly alleviated
 - Investigating how the *asynchronous evolution* of KGs affects consistency
- Interesting Open Problems:
 - E.g. Synchronisation/Updates/Repairs in Polyglott Databases



SWJ (2024) <https://doi.org/10.3233/SW-243571>

Wikidata Query Service

Examples Help More tools

```

1 # Which settlements had ever over 1M population?
2
3 SELECT DISTINCT ?City ?Country ?Pop (Min(?Timepoint) AS ?FirstTime)
4 {
5   ?City wdt:P17 ?Country;
6       wdt:P31/wdt:P279* wd:Q486972;
7       p:P1082 ?Stmnt.
8   ?Stmnt ps:P1082 ?Pop;
9       pq:P585 ?Timepoint.
10  FILTER (?Pop > 1000000)
11
12 } GROUP BY ?City ?Country ?Pop
  
```

Query timeout limit reached

SHORT-PAPER | OPEN ACCESS

The Massive Problem of Remote Changes in Ontology Reuse

Authors: Romana Pernisch, Daniil Dobryi, Axel Polleres | [Authors Info & Claims](#)

WWW '25: Companion Proceedings of the ACM on Web Conference 2025 • Pages 1254 - 1258 • <https://doi.org/10.1145/3701716.3715478>



Isn't there something missing?



Ora Lassila (Keynote ISWC2024)

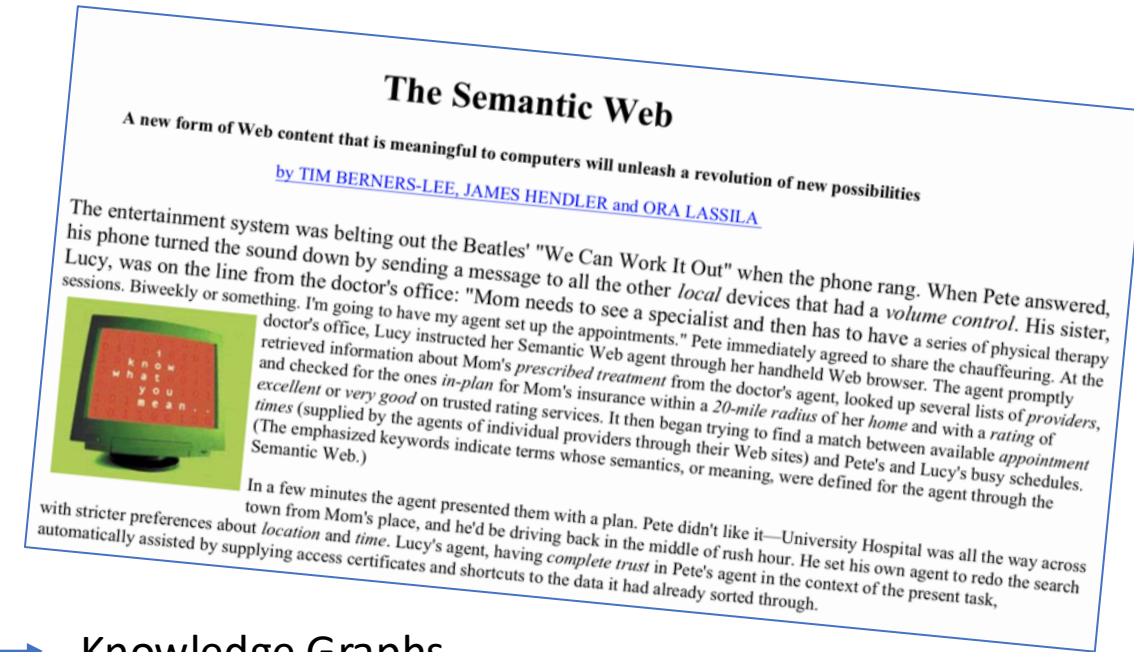
<https://www.lassila.org/publications/2024/lassila-iswc2024-keynote.pdf>

The realization of autonomous agents minimally requires these:

1. KR & reasoning
2. Planning
3. Ability to converse with the agents

LLMs will give you #3, but not #1 or #2

- (despite what you hear)
- "agentic", "agentive", ... huh?
- good news: we already have #1 and #2



Knowledge Graphs

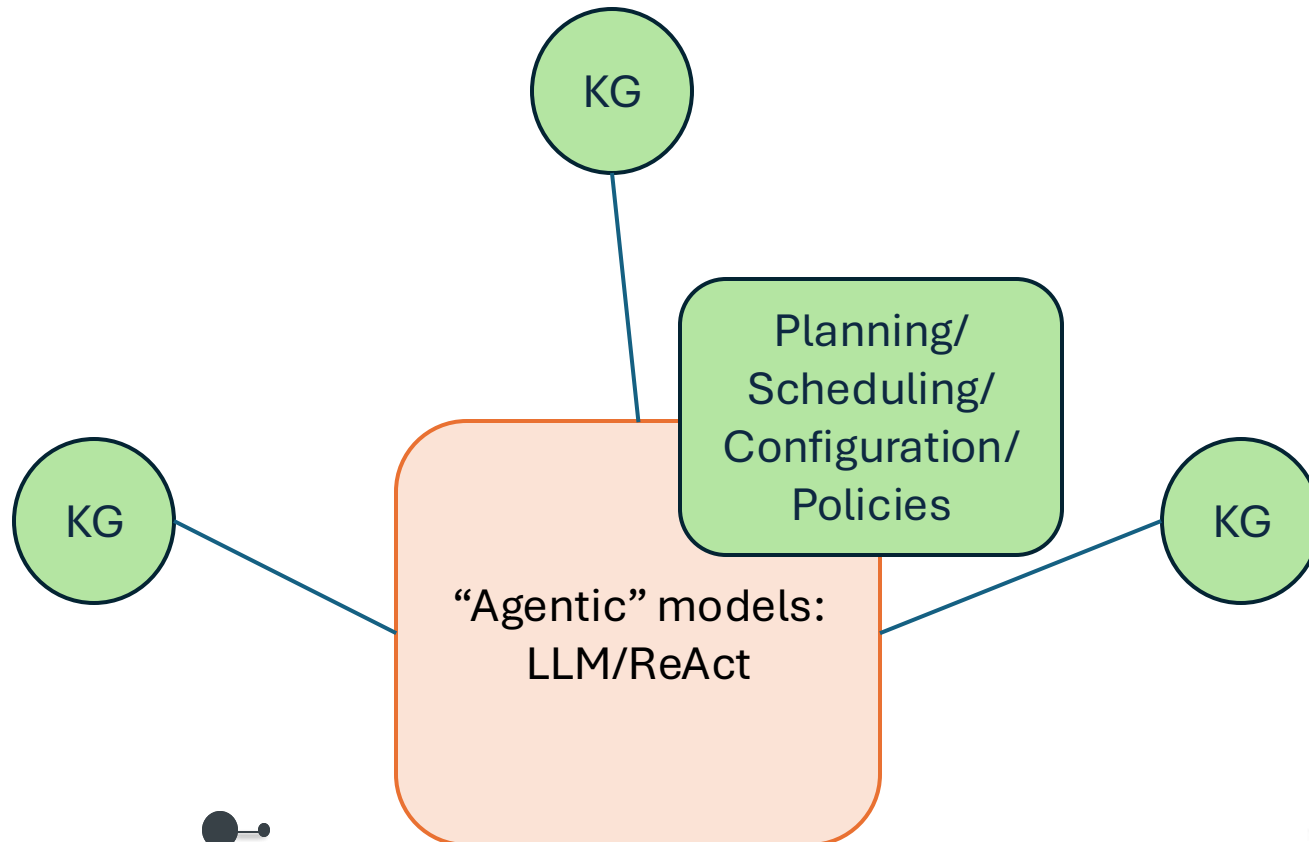
"Agentic" Models ... and maybe a revival?

LLMs



What's next? Where should this go?

Ingredient3: Combinations of LLMs and decentral KGs via **RAG** and **Agent frameworks** are one example of such Neuro-symbolic architectures:



Example: Federated querying of KGs via agents (work in progress):

Prompt: “Write a federated SPARQL query that queries all of Tim Berners-Lee’s publications from DBLP the DOIs of which also appear on Wikidata”

```

PREFIX dblp:    < http://dblp.org/rdf/schema#>
PREFIX bibo:    <http://purl.org/ontology/bibo/>
PREFIX wdt:     <http://www.wikidata.org/prop/direct/>
SELECT DISTINCT ?dblpPub ?title ?doi ?wikidataItem
WHERE {
  SERVICE < https://dblp.org/sparql > {
    ?dblpPub a bibo:Article ;
    dblp:author ?author ;
    dblp:title ?title ;
    bibo:doi ?doi .
    ?author dblp:name "Tim Berners-Lee" .}
  SERVICE <https://query.wikidata.org/sparql> {
    ?wikidataItem wdt:P356 ?doi . } }
  
```

What's next? Where should this go?

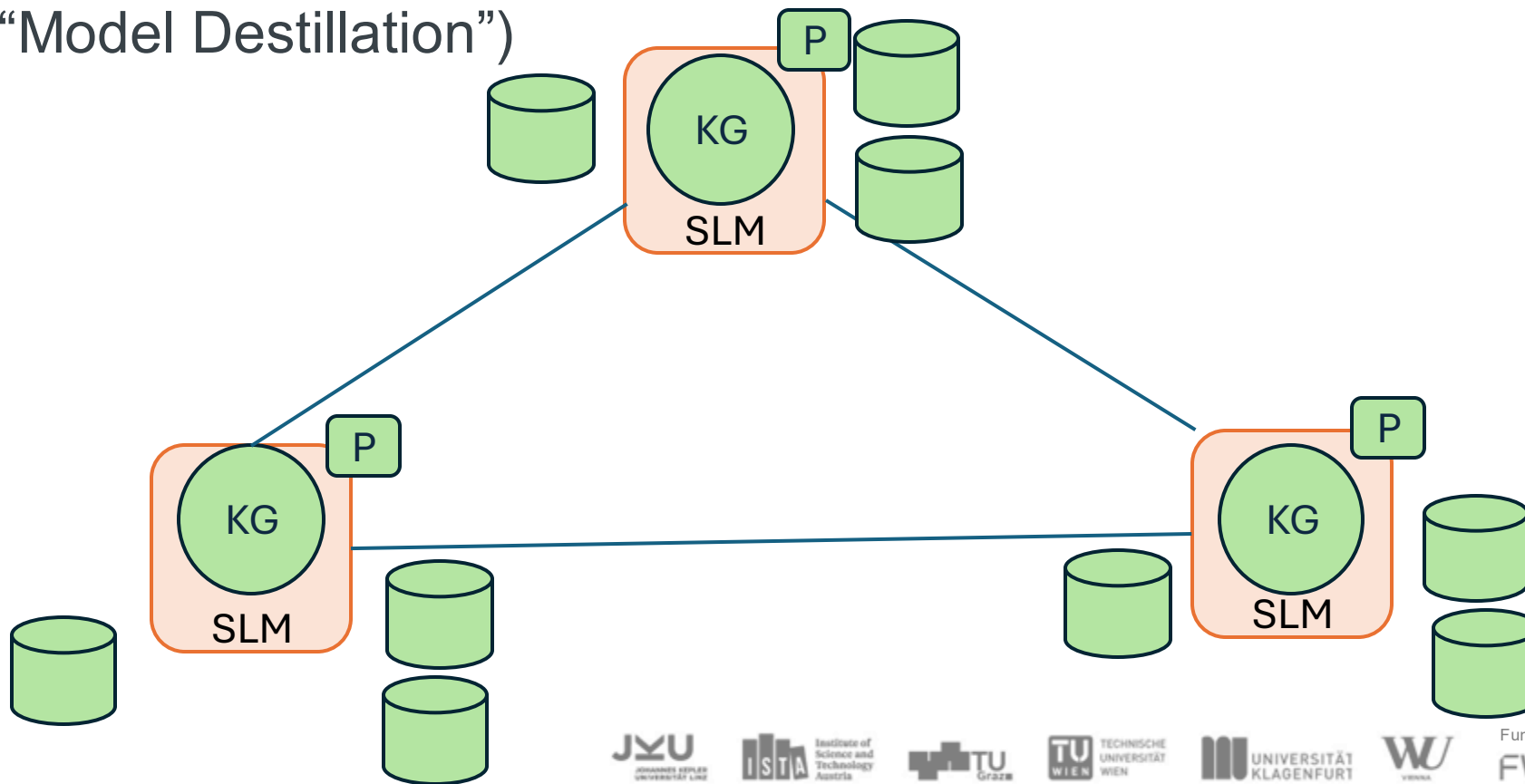
What is a solution we should jointly work towards?

Neuro-symbolic agents

Trustable Decentralized Agents

Small models (e.g. “Model Destillation”)

“Agent Spaces”



Thanks!

Austrian National “Cluster of Excellence” BILAI:

- Broad AI
- (Knowledge) Graph-Based AI plays a key role in BILAI

Knowledge Graphs are symbolic AI with “features” of LLMs:

- Large-scale
- Not always consistent

(**Dezentralized**) Agents, KGs and LLM combined, could realize BILAI’s vision of a broad, **robust** AI, and change the way we access and manage Data

- **Ingredient1:**
Neurosymbolic Architectures
- **Ingredient2:**
scalable
distributed, fed
erated data
symbolic
processing



Currently running!