

Wissensgraphen – eine kurze Einführung

Prof. Dr. Birgitta König-Ries

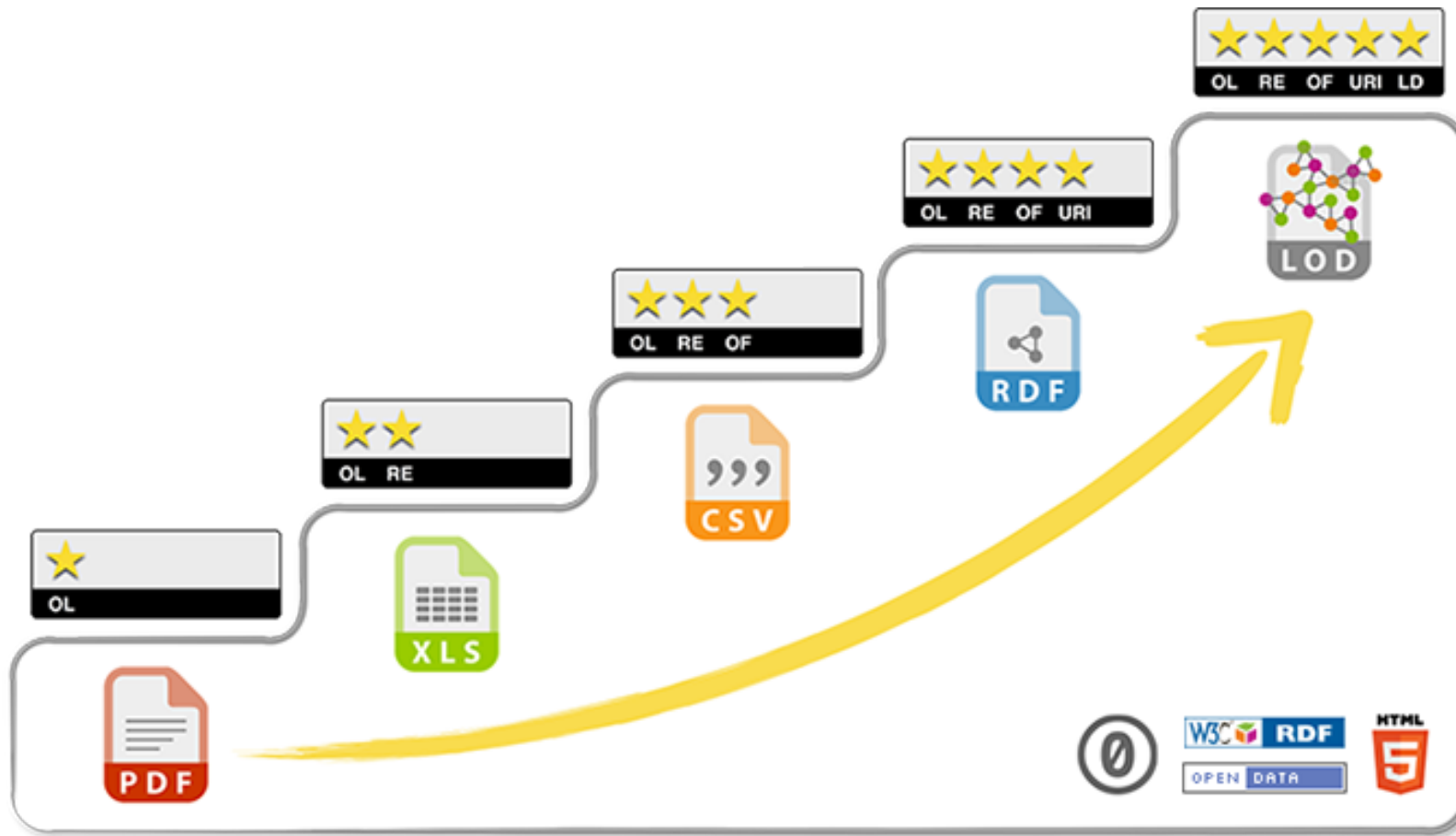
Heinz-Nixdorf-Proessur für verteilte
Informationssysteme

Friedrich-Schiller-Universität Jena

OpenDVA Kongress
Dornburger Schlösser 25. - 26. Mai 2023



5 Sterne Daten (nach Tim Berners-Lee)



Drei Problemstellungen – eine
Antwort



Suchmaschinen: Antworten statt Links

Früher

Uni Jena Studierendenzahl

Lorem ipsum dolor sit amet Uni Jena, consectetur adipiscing elit. Aenean commodo ligula eget dolor.

Aenean massa. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur Uni Jena ridiculus mus.

Donec pede Uni justo, fringilla vel, aliquet nec, vulputate eget, Jena arcu.

In enim justo, Jena rhoncus ut, imperdiet a, venenatis

1 2 3 4 5 ... 119





Suchmaschinen: Antworten statt Links

Früher

Uni Jena Studierendenzahl

Lorem ipsum dolor sit amet Uni Jena, consectetur adipiscing elit. Aenean commodo ligula eget dolor.

Aenean massa. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur Uni Jena ridiculus mus.

Donec pede Uni justo, fringilla vel, aliquet nec, vulputate eget, Jena arcu.

In enim justo, Jena rhoncus ut, imperdiet a, venenatis

1 2 3 4 5 ... 119

Heute

Uni Jena Studierendenzahl

18.048

Lorem ipsum dolor sit amet Uni Jena, consectetur adipiscing elit. Aenean commodo ligula eget dolor.




Aenean massa. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur Uni Jena ridiculus mus.

1 2 3 4 5 ... 119






Wikipedia: zentrale Datenquelle statt verteilter Updates

Früher




	Jena Einwohner: 111.343
	Jena Population: 107.200
	Jena Nüfus: 110.321

Wikipedia : zentrale Datenquelle statt verteilter Updates

Früher

	Jena Einwohner: 111.343
	Jena Population: 107.200
	Jena Nüfus: 110.321

Heute (oder morgen)

	Jena Einwohner: 111.343
	Jena Population: 111.343
	Jena Nüfus: 111.343

Organisationen: Expertise erfassen, erhalten und kommunizieren

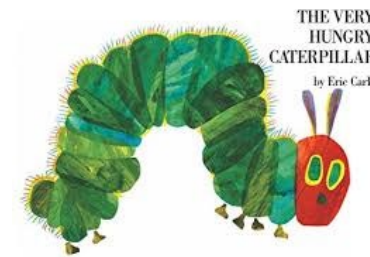
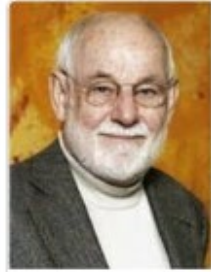


Gemeinsame Lösung: Wissensgraphen

- Grundidee: Ein Modell von Wissen über die Welt entwickeln, das
 - für Menschen und Maschinen verständlich ist
 - durchsucht werden kann
 - abgefragt werden kann
 - Schlussfolgerungen ermöglicht
 - schrittweise erweitert und angepasst werden kann
 - dezentral aufgebaut werden kann
 - unabhängig entwickelte Teile miteinander verbinden kann



Was ist ein Wissensgraph?



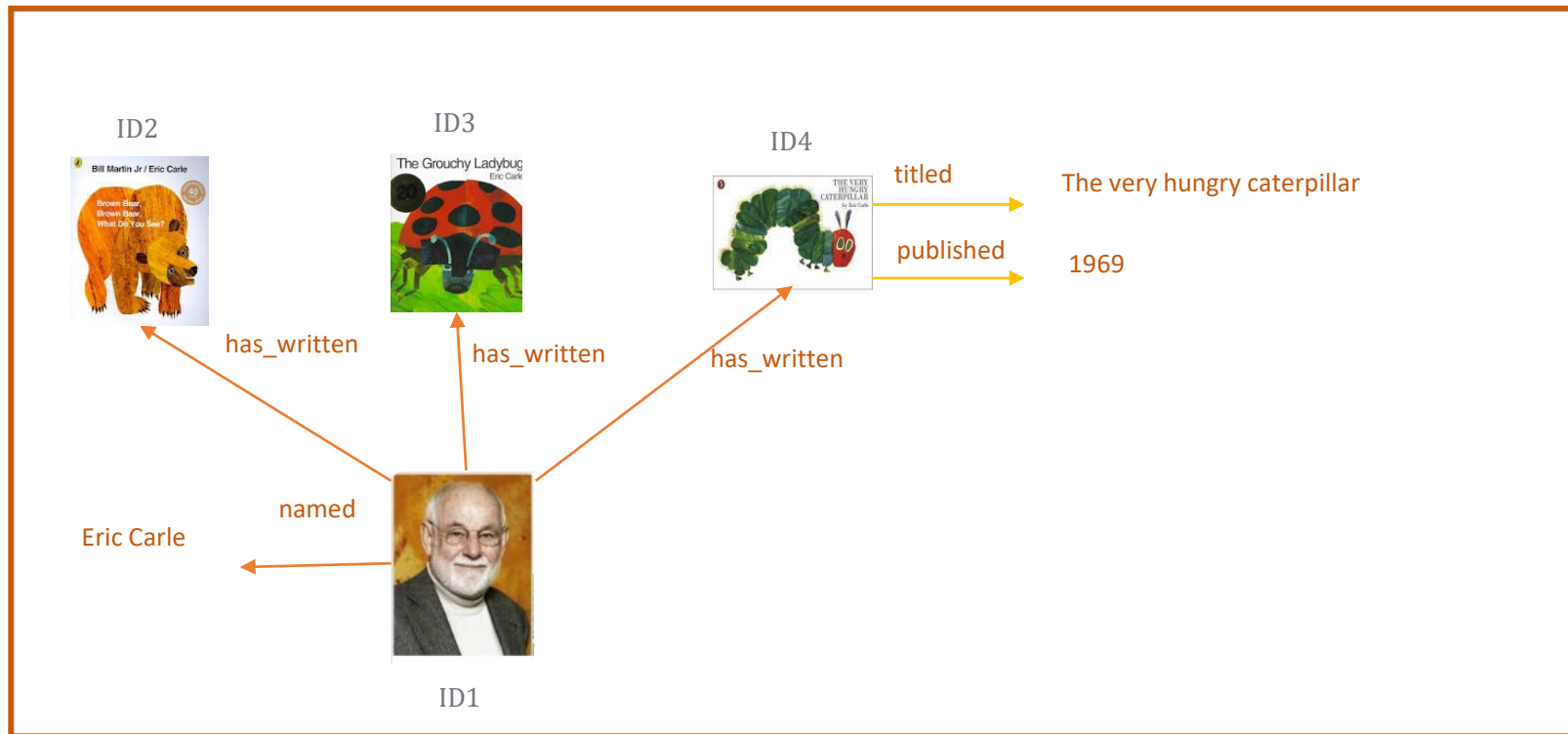
Was ist ein Wissensgraph?

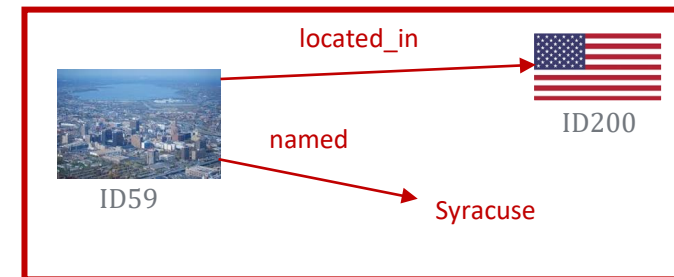
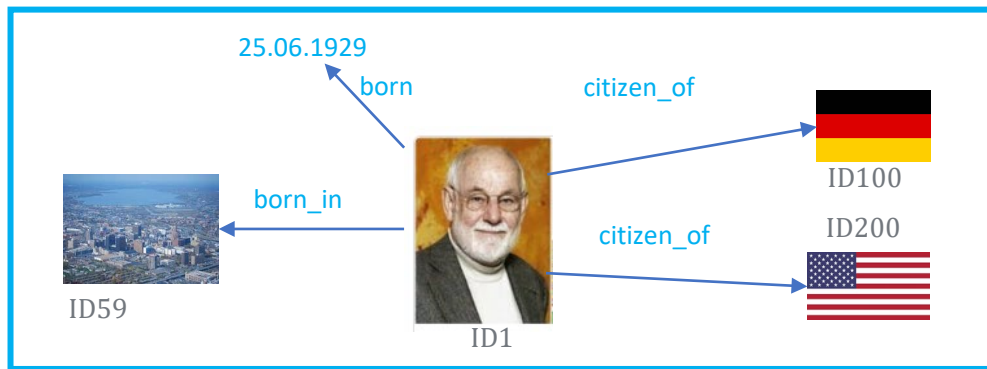
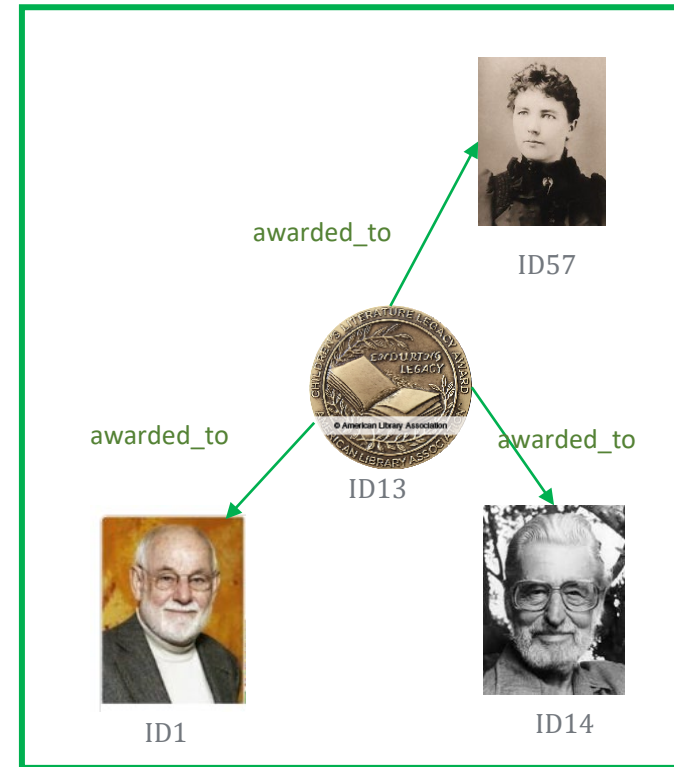
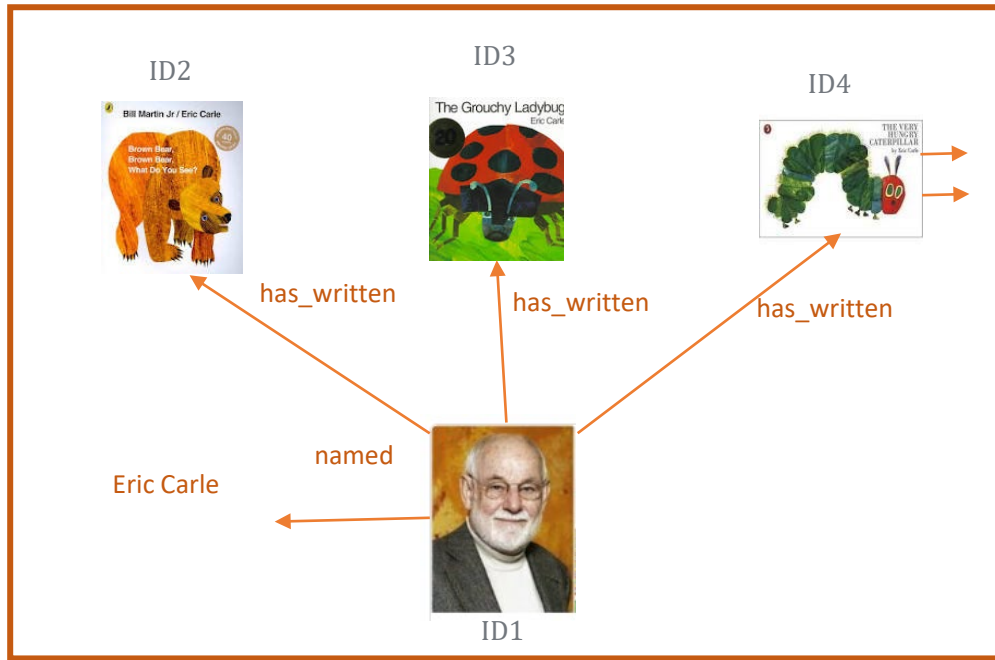


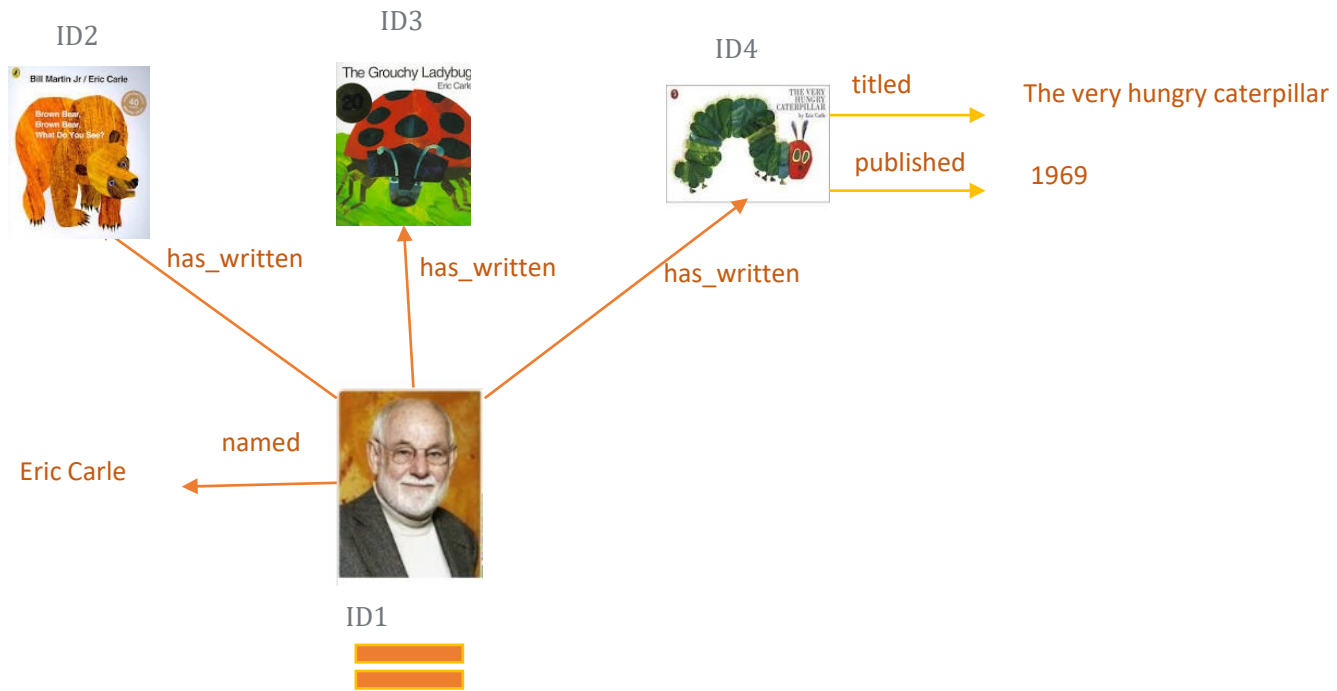
Was ist ein Wissensgraph?



Was ist ein Wissensgraph?







```

{(ID1 named „Eric Carle“)
 (ID1 has_written ID2)
 (ID1 has_written ID3)
 (ID1 has_written ID4)
 (ID4 titled „The very hungry caterpillar“)
 (ID4 published „1969“)}

```

{{(ID1 named „Eric Carle“)
(ID1 has_written ID2)
(ID1 has_written ID3)
(ID1 has_written ID4)
(ID4 titled „The very hungry caterpillar“)
(ID4 published „1969“}}

{{(ID1 born „25.06.1929“)
(ID1 citizen_of ID100)
(ID1 citizen_of ID200)
(ID1 born_in ID59)}}

{{(ID13 awarded_to ID1)
(ID13 awarded_to ID57)
(ID13 awarded_to ID14)}}

{{(ID59 located_in ID100)
(ID59 called „Syracuse“}}


```
{(ID1 named „Eric Carle“)  
(ID1 has_written ID2)  
(ID1 has_written ID3)  
(ID1 has_written ID4)  
(ID4 titled „The very hungry caterpillar“)  
(ID4 published „1969“)}
```

```
{(ID1 born „25.06.1929“)  
(ID1 citizen_of ID100)  
(ID1 citizen_of ID200)  
(ID1 born_in ID59)}
```

```
{(ID13 awarded_to ID1)  
(ID13 awarded_to ID57)  
(ID13 awarded_to ID14)}
```

```
{(ID59 located_in ID100)  
(ID59 called „Syracuse“)}
```

Mit diesem Datenmodell kann man

- Aus einem Wissensgraphen auf einen anderen verweisen
- Daten über Datenquellen hinweg integrieren
- Einfache und komplexe Anfragen beantworten
- Implizites Wissen schlussfolgern

{{(ID1 named „Eric Carle“)
(ID1 has_written ID2)
(ID1 has_written ID3)
(ID1 has_written ID4)
(ID4 titled „The very hungry caterpillar“)
(ID4 published „1969“}}

{{(ID1 born „25.06.1929“)
(ID1 citizen_of ID100)
(ID1 citizen_of ID200)
(ID1 born_in ID59)}}

{{(ID13 awarded_to ID1)
(ID13 awarded_to ID57)
(ID13 awarded_to ID14)}}

{{(ID59 located_in ID100)
(ID59 called „Syracuse“}}



Eric Carle
American designer

eric-carle.com

Eric Carle is an American designer, illustrator, and writer of children's books. He is most noted for *The Very Hungry Caterpillar*, a picture book that has been translated into more than 66 languages and sold more than 50 million copies, equivalent to 1.8 copies sold every minute since it was published in 1969. [Wikipedia](#)

Born: June 25, 1929 (age 91 years), Syracuse, New York, United States

Children: 2

Residence: Key West, Florida, U.S

Spouse: Barbara Morrison

Awards: Regina Medal, Children's Literature Legacy Award

Books [View 45+ more](#)

 The Very Hungry Caterpillar 1969	 Brown Bear, Brown B... 1967	 The Grouchy Ladybug 1977	 The Very Busy Spider 1984	 From Head to Toe 1997
---	--	---	--	--

Profiles

[Facebook](#) [Instagram](#)

People also search for [View 10+ more](#)

 Bill Martin, Jr.	 Dr. Seuss	 Maurice Sendak	 Roald Dahl	 Shel Silverstein
--	---	--	--	--

● RDF Wissensgraph

● Wikidata

@prefix rdf: <<http://www.w3.org/1999/02/22-rdf-syntax-ns#>> .

@prefix wd: <<http://www.wikidata.org/entity/>> .

@prefix wdt: <<http://www.wikidata.org/prop/direct/>> .

Resource: Eric Carle (Q606557)

wd:Q606557 rdf:type wd:Q5 ;

wdt:P1559 "Eric Carle" ;

wdt:P800 wd:Q491657, wd:Q7738358, wd:Q4976047 .

Resource: Book 1 - The Very Hungry Caterpillar (Q491657)

wd:Q491657 rdf:type wd:Q571 ;

wdt:P1476 "The very hungry caterpillar" ;

wdt:P577 "1969" .

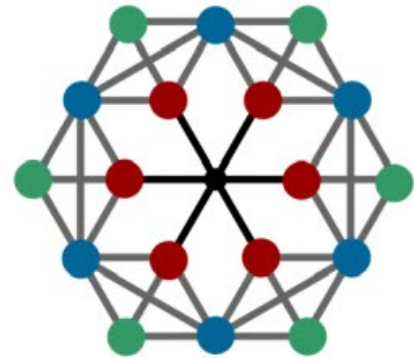


Sind die irgendwie wichtig?

– Wahrscheinlich schon

Google

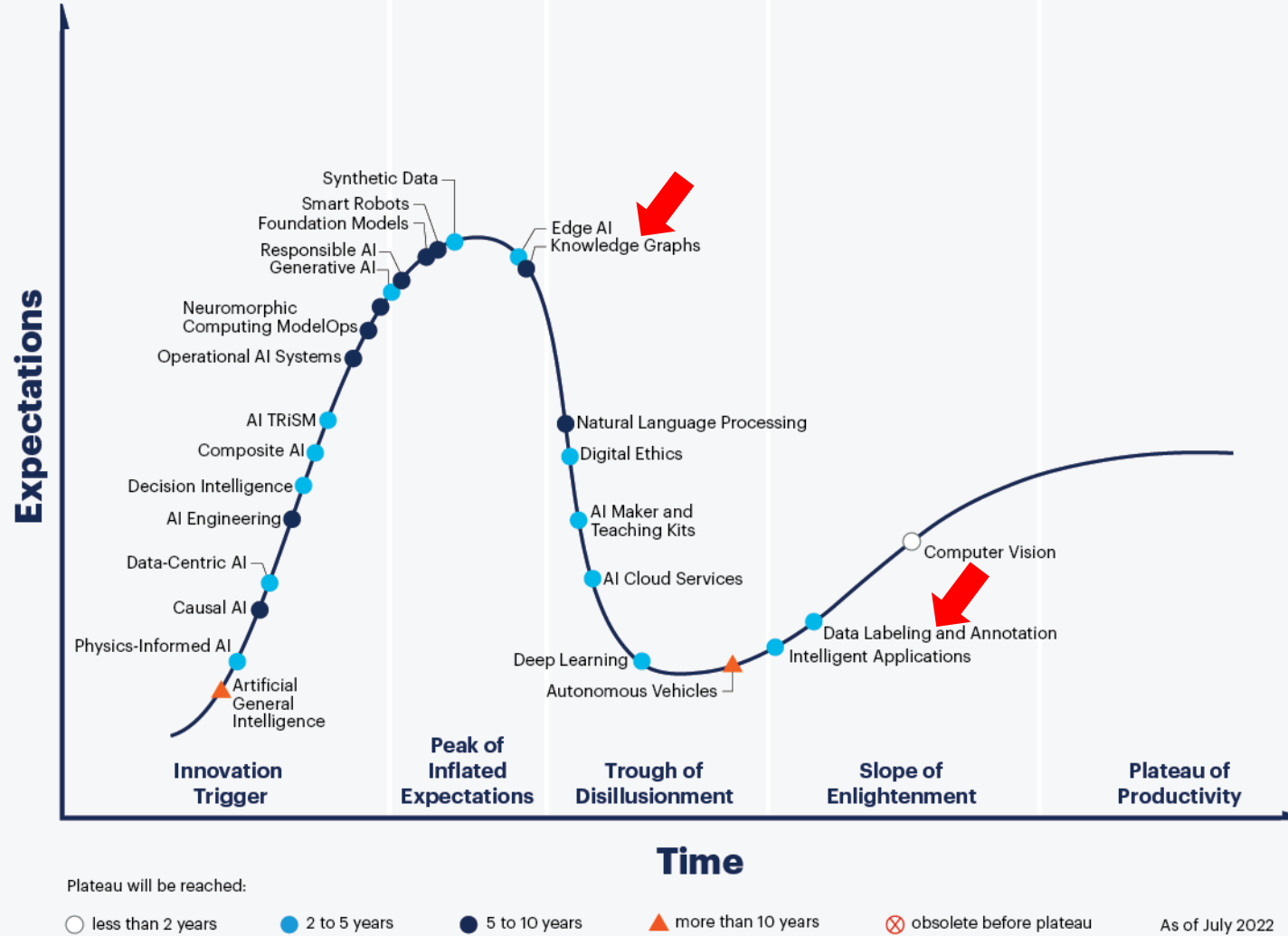
amazon



WIKIDATA

EA
ELECTRONIC ARTS™

Hype Cycle for Artificial Intelligence, 2022



[gartner.com](https://www.gartner.com)

Source: Gartner
 © 2022 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner and Hype Cycle are registered trademarks of Gartner, Inc. and its affiliates in the U.S. 1957302

Gartner

● Nutzen von Wissensgraphen: ● Verständnisbildung

- Klare Definition von Konzepten und Beziehungen
- Erkennen von Ungenauigkeiten und Ungereimtheiten
- Förderung von Einigungsprozessen oder Explizitmachen von unterschiedlichen Interpretationen



● Nutzen von Wissensgraphen: ● Zusammenführen von Informationen

- eindeutige Identifikatoren erlauben es, unterschiedliche Aspekte des selben Dings in unterschiedlichen Wissensgraphen zu beschreiben
- Verlinkung von bislang getrennter Information

● Nutzen von Wissensgraphen: Erschließen ● impliziter Information

- Logische Schlußfolgerungen
- Identifikation wahrscheinlicher Information



Nutzen von Wissensgraphen: Provenance

- Wissensgraphen können (ACHTUNG: müssen aber nicht!) für jeden Datenpunkt Herkunftsinformation oder Belege enthalten.

Nutzen von Wissensgraphen: Anfragen

- SPARQL ist eine standardisierte Anfragesprache, mit der Informationen aus einem oder mehreren Wissensgraphen gezielt abgefragt werden können.
 - Für viele Anwendungen sind instanziierbare Schablonen ausreichend
 - Für komplexere Anforderungen: erste Werkzeuge da, die bei Anfrageformulierung unterstützen
 - Auswertung von Anfragen mittlerweile (mit der richtigen Plattform) effizient auch auf (sehr) großen Wissensgraphen möglich.

SPARQL (created with a bit of help from ChatGPT)

- PREFIX dbo: <http://dbpedia.org/ontology/>
- PREFIX dbp: <http://dbpedia.org/property/>
- PREFIX dbr: <http://dbpedia.org/resource/>

```
SELECT ?person ?yearOfBirth ?birthplaceName ?inhabitantCount
WHERE {
    ?person dbo:birthPlace ?birthplace ;
            dbo:birthDate ?yearOfBirth .

    ?birthplace dbo:country dbr:Germany ;
                rdfs:label ?birthplaceName ;
                dbp:populationTotal ?inhabitantCount .

    FILTER(?inhabitantCount > 100000)
}
LIMIT 10
```





SPARQL (created with a bit of help from ChatGPT)

PREFIX wd: <http://www.wikidata.org/entity/>

PREFIX wdt: <http://www.wikidata.org/prop/direct/>

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

```
SELECT ?person ?yearOfBirth ?birthplaceLabel ?inhabitantCount
```

```
WHERE {
```

```
  ?person wdt:P19 ?birthplace ;
```

```
    wdt:P569 ?birthDate .
```

```
  ?birthplace wdt:P17 wd:Q183 ;
```

```
    rdfs:label ?birthplaceLabel ;
```

```
    wdt:P1082 ?inhabitantCount .
```

```
  FILTER(langMatches(lang(?birthplaceLabel), "en"))
```

```
  FILTER(?inhabitantCount > 100000)
```

```
}
```

```
LIMIT 10
```



Oops....

- PREFIX wd: <http://www.wikidata.org/entity/>
- PREFIX wdt: <http://www.wikidata.org/prop/direct/>
- PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

```
SELECT ?person ?yearOfBirth ?birthplaceLabel ?inhabitantCount
WHERE {
  ?person wdt:P19 ?birthplace ;
    wdt:P569 ?birthDate .

  ?birthplace wdt:P131 wd:Q1619 ;
    rdfs:label ?birthplaceLabel ;
    wdt:P1082 ?inhabitantCount .

  FILTER(langMatches(lang(?birthplaceLabel), "en"))
  FILTER(?inhabitantCount > 100000)
}
LIMIT 10
```



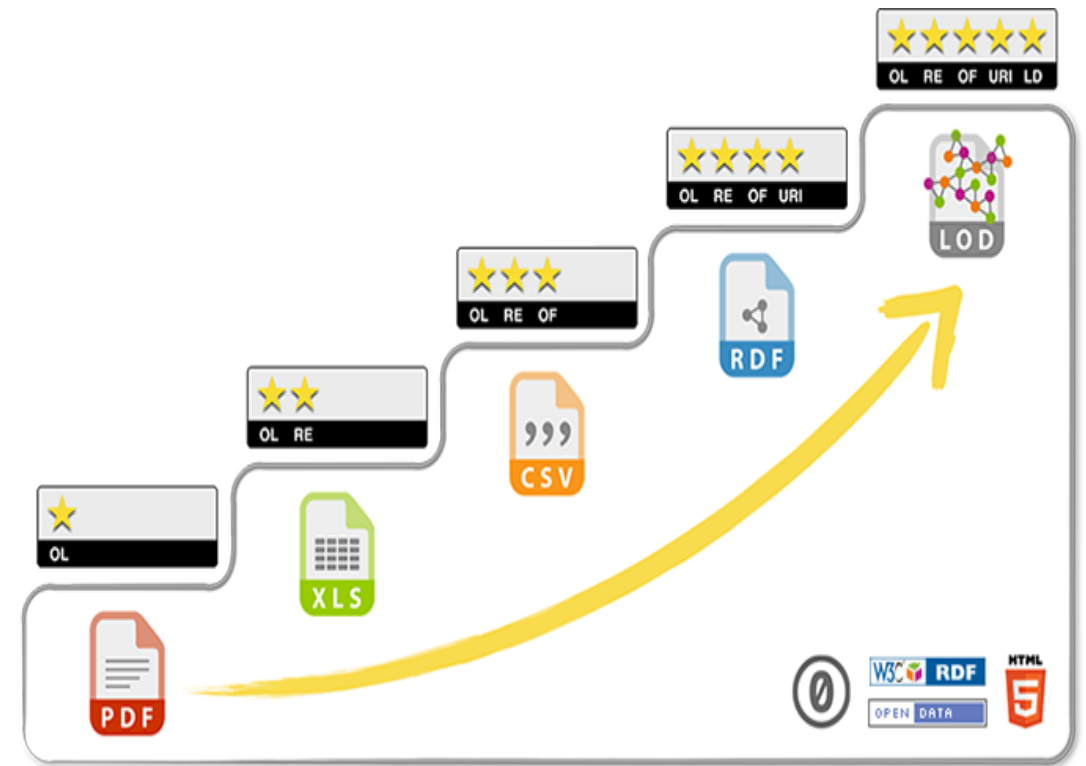
Unterstützung

- Wachsende Zahl von Plattformen zur Erstellung, Verwaltung, Nutzung
- Forschungsarbeiten zur automatisierten Erstellung und Verlinkung
- Werkzeuge zur Unterstützung bei der Formulierung von Anfragen



Fazit

- Wissensgraphen werden weithin als vielversprechender Ansatz gesehen um
 - über Organisationen verteilt relevante Ausschnitte der Welt zu modellieren (auf Schema- und Instanzebene)
 - Unabhängig voneinander entstandene Modelle zu integrieren
 - Expert*innen und Laien einfachen Zugriff auf Modelle zu gewähren
 - 5-Sterne-Daten im Sinne von Tim Berners-Lee bereitzustellen





**Vielen Dank
für Ihre Aufmerksamkeit!**

Kollaboration:



Gefördert durch:



aufgrund eines Beschlusses
des Deutschen Bundestages

