

1 4 6 4 1 4 3 9 6 5 7  
3 6 3 1 4 7 9 4 8 3 1  
5 2 7 3 4 1 9 7 7 3 4  
1 5 2 7 8 4 0 3 2 6 3  
3 0 6 7 0 9 3 7 1 3 2



# Einführung in Wissensgraphen

---

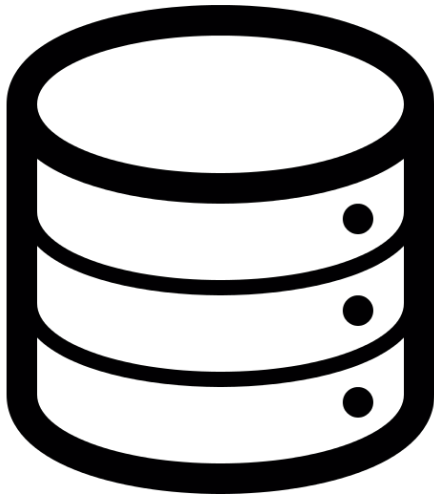
Wissensgraphen und große Sprachmodelle in den Digital Humanities  
Workshop @ DHd 2025

*Prof. Dr. Ralf Krestel*  
Christian-Albrechts-Universität zu Kiel &  
ZBW – Leibniz-Informationzentrum Wirtschaft

# Wie speichert man Wissen?



- Grundlegendes Problem der Menschheit seit Anbeginn
  - Höhle von Lascaux (16000 v. Chr.)
- Wie speichert man Daten?



# Vom Datenspeicher zum Wissensspeicher

---



- **Daten** in einer **Datenbank (DB)** (+ Schema)
  - Relationale / Graph- / NoSQL-DB

# Beispiel: Datenbank



- Informationen über
  - Bücher
  - Verleger
  - Autoren

**Books**

Title	Author	Publisher	Year Published	Followed By
To Kill a Mockingbird	Harper Lee	J. B. Lippincott Company	1960	Go Set a Watchman
Go Set a Watchman	Harper Lee	HarperCollins, LLC; Heinemann	2015	
The Picture of Dorian Gray	Oscar Wilde	J. B. Lippincott & Co.	1890	
2001: A Space Odyssey	Arthur C. Clarke	New American Library, Hutchinson	1968	

**Publishers**

Name	City	Country
J. B. Lippincott & Company	Philadelphia	United States
HarperCollins, LLC	New York City	United States
Heinemann	Portsmouth	United States
New American Library	New York City	United States
Hutchinson	London	United Kingdom

**Authors**

Name	Country of Birth
Harper Lee	United States
Oscar Wilde	Ireland
Arthur C. Clarke	United Kingdom

# Vom Datenspeicher zum Wissensspeicher

---

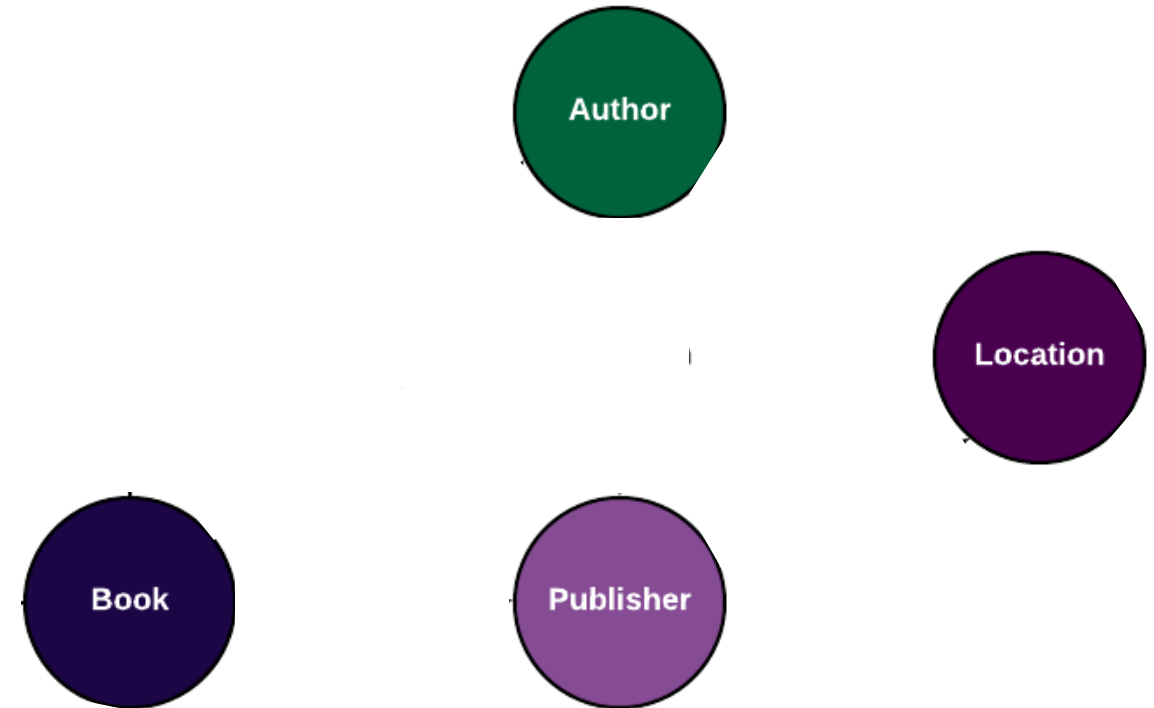


- **Daten** in einer **Datenbank (DB)** (+ Schema)
  - Relationale / Graph- / NoSQL-DB
- **Metadaten** in einer **Ontologie**
  - Informationen über Daten und wie diese zusammenhängen (üblicherweise in Form von Tripeln)

# Beispiel: Ontologie



- Klassen, z.B.
  - Bücher
  - Autoren
  - Verleger
  - Orte
- Attribute, z.B.
  - Bücher werden zu einem Datum veröffentlicht
- Relationen, z.B.
  - Bücher haben Autoren
  - Bücher haben Verleger
  - Bücher können einen “Nachfolger” haben



# Vom Datenspeicher zum Wissensspeicher

---



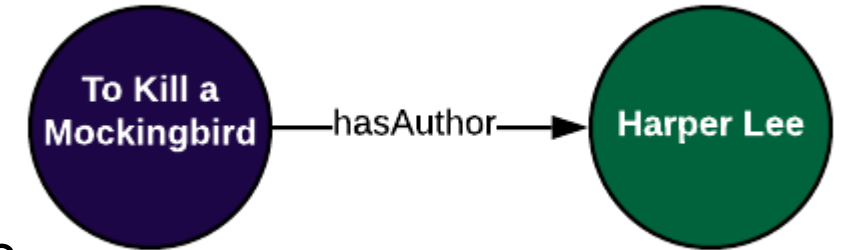
- **Daten** in einer **Datenbank (DB)** (+ Schema)
  - Relationale / Graph- / NoSQL-DB
- **Metadaten** in einer **Ontologie**
  - Informationen über Daten und wie diese zusammenhängen (üblicherweise in Form von Tripeln)
- **Fakten** in einer **Wissensbasis (Knowledge Base, KB)**
  - **Fakten** in Form von Fakten-Tripeln



# Beispiel: Wissensbasis



- Sammlung von Fakten:
  - *Harper Lee* → *arbeitet\_mit* → *JBL&C*
  - *JBL&C* → *gelegen\_in* → *Philadelphia*
  - *To Kill a Mockingbird* → *hat\_Autor* → *Harper Lee*
  - *To Kill a Mockingbird* → *hat\_Verleger* → *JBL&C*
  - *Philadelphia* → *gelegen\_in* → *United States of America*
  - ...
- Zusätzliche Relation „ist\_ein/e“ oder „ist\_vom\_Typ“ können Fakten mit Konzepten aus einer Ontologie verbinden
  - *Harper Lee* → *ist\_eine* → *Autorin*
  - *To Kill a Mockingbird* → *ist\_ein* → *Buch*
  - *Philadelphia* → *ist\_eine* → *Stadt*
  - ....



Darstellung als Tripel:  
Subjekt-Prädikat-Objekt



# Vom Datenspeicher zum Wissensspeicher

---

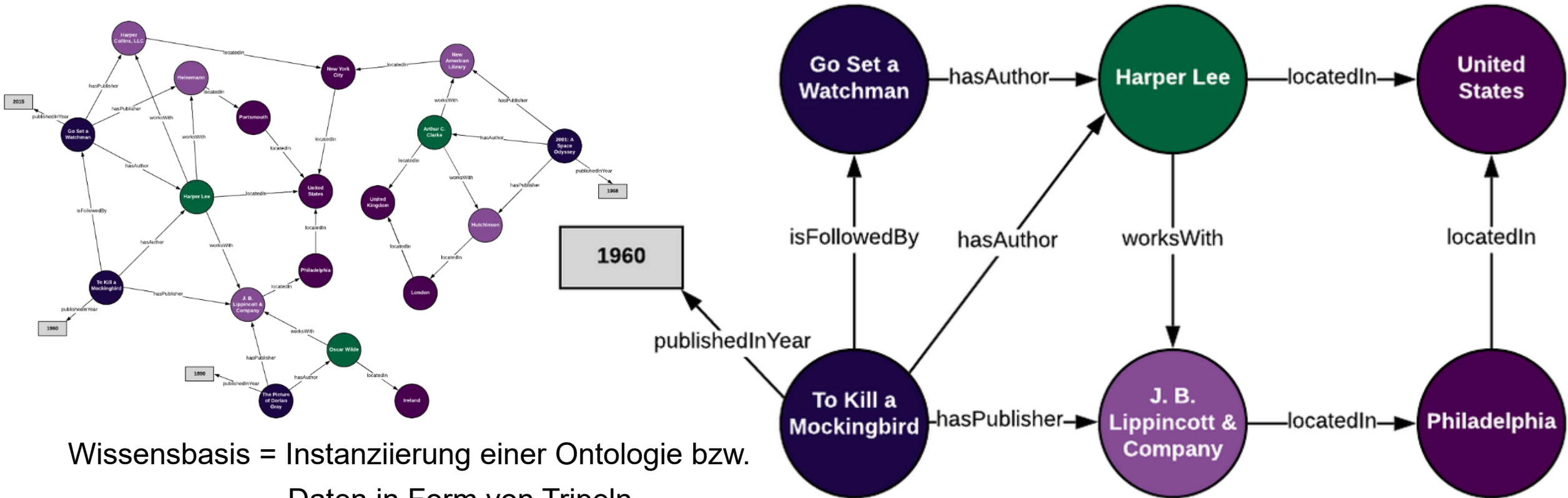


- **Daten** in einer **Datenbank (DB)** (+ Schema)
  - Relationale / Graph- / NoSQL-DB
- **Metadaten** in einer **Ontologie**
  - Informationen über Daten und wie diese zusammenhängen (üblicherweise in Form von Tripeln)
- **Fakten** in einer **Wissensbasis (Knowledge Base, KB)**
  - **Fakten** in Form von Fakten-Tripeln)
- **Verbundene Fakten** in einem **Wissensgraphen (Knowledge Graph, KG)**
  - Die Fakten-Tripel werden zu einem Netzwerk verbunden

# Beispiel: Wissensgraph



- Ein Graph mit allen Tripeln aus der Wissensbasis miteinander verbunden



Wissensbasis = Instanziierung einer Ontologie bzw.

Daten in Form von Tripeln

Wissensgraph = Wissensbasis + Graph

# Definition: Wissensgraph

---

- Ein Wissensgraph ist
  - **strukturiert** (nutzt eine bestimmte Datenstruktur)
  - **normalisiert** (besteht aus Knoten (Entitäten) und Kanten (Relationen))
  - **vernetzt** (mit Verbindungen zwischen Objekten)
- Wissensgraphen sind typischerweise
  - **explizit** (bewusst erstellt mit dem klaren Ziel Wissen zu speichern)
  - **deklarativ** (selbsterklärend und unabhängig von individuellen Implementierungen)
  - **annotiert** (angereichert mit Kontextinformationen, Metadaten, Quellen)
  - **nicht-hierarchisch** (keine einfache Baum-Struktur)
  - **sehr groß** (Millionen von Elementen)

# Unternehmenswissensgraphen



	Data model	Size of the graph	Development stage
Microsoft	The types of entities, relations, and attributes in the graph are defined in an ontology.	~2 billion primary entities, ~55 billion facts	Actively used in products
Google	Strongly typed entities, relations with domain and range inference	1 billion entities, 70 billion assertions	Actively used in products
Facebook	All of the attributes and relations are structured and strongly typed, and optionally indexed to enable efficient retrieval, search, and traversal.	~50 million primary entities, ~500 million assertions	Actively used in products
eBay	Entities and relation, well-structured and strongly typed	Expect around 100 million products, >1 billion triples	Early stages of development and deployment
IBM	Entities and relations with evidence information associated with them.	Various sizes. Proven on scales documents >100 million, relationships >5 billion, entities >100 million	Actively used in products and by clients

Noy, N., Gao, Y., Jain, A., Narayanan, A., Patterson, A., & Taylor, J. (2019). Industry-scale knowledge graphs: lessons and challenges. *Communications of the ACM*, 62(8), 36-43.

# Offene Wissensgraphen



- Bekannte offene KGs

- Dbpedia
- YAGO
- Freebase
- Wikidata



- Große, allgemeine Wissensgraphen
- Beinhalten vielfältige Entitäten und Relationen
- Sind in RDF modelliert
- Veröffentlicht nach den Linked Data Prinzipien
- Ermöglichen Zugriff via
  - download (RDF)
  - Knotenabfrage (Linked Data)
  - Graphmuster/Anfragen (SPARQL)



# Wikidata



- Kollaborativer Wissensgraph
- Wikimedia-Projekt (2012)
  - Federführend ist Wikimedia Deutschland
- Ursprüngliche Idee:
  - Sprachunabhängige Zentralisierung von Fakten/Wissen
  - Bietet einen zentralen Ort für Infobox-Daten für alle Wikipedias
  - Erzeugen und Updaten von List-Artikeln basierend auf Wikidata-Daten
- Wikidata query service (2015)
- Statistiken
  - >100m items, 11k properties, >2b edits, 25k active users
  - <https://www.wikidata.org/wiki/Wikidata:Statistics>



# Wissensgraph nutzen



- Zum Beispiel um Antworten auf komplexe Fragen zu bekommen
- Beispiel: **Wikidata**-Wissensgraph
  - <https://query.wikidata.org/>
  - SPARQL Anfragesprache





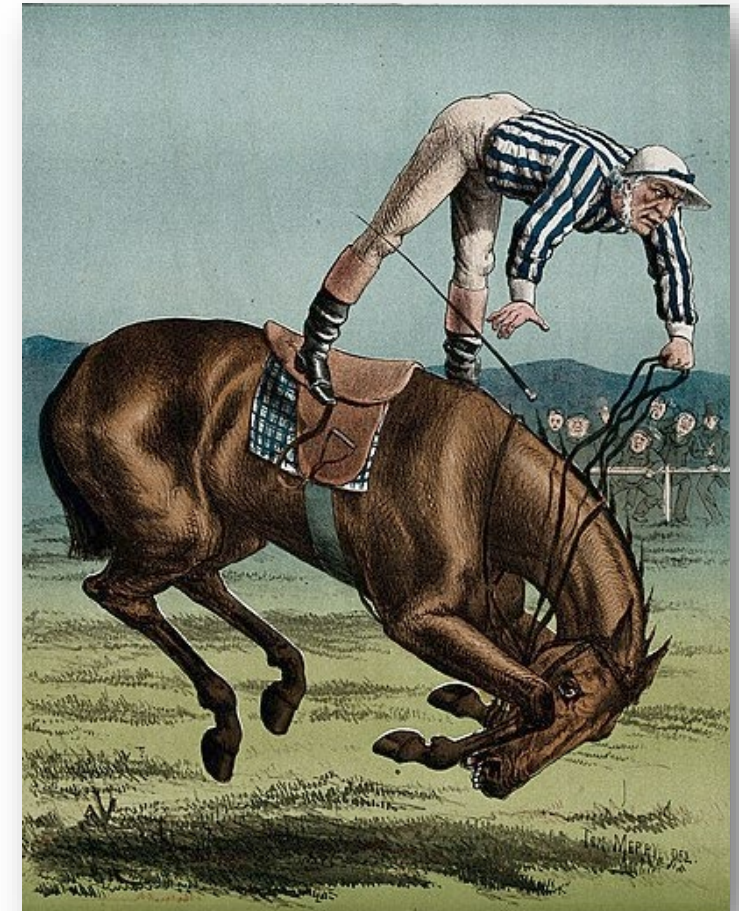
# Anfragen beantworten



- Woran sind die meisten adeligen Menschen gestorben?

Ursache	Anzahl	Ursache	Anzahl
im Einsatz getötet	423	Schlaganfall	141
Enthauptung	295	Herzinfarkt	137
Krankheit	266	Ertrinken	118
Tuberkulose	232	Gift	109
Kindbettfieber	198	Krebs	97
Pest	197	Sturz vom Pferd	91
Lungenentzündung	179	Exekution	90
Schusswunde	144	Hängen	81
Pocken	144	Straßenverkehrsunfall	61

```
SELECT ?cause ?causeLabel (COUNT(?person) AS ?count) WHERE {  
  ?person wdt:P31 wd:Q5;  
          wdt:P509 ?cause;  
          wdt:P53 [].  
  SERVICE wikibase:label { bd:serviceParam wikibase:language "de". }  
GROUP BY ?cause ?causeLabel  
HAVING(?count > 1)  
ORDER BY DESC(?count)
```



# Anfragen beantworten



- Nach welchen Frauen wurden die meisten Straßen in den Niederlanden benannt?

```
#defaultView:ImageGrid
SELECT ?item ?name ?image (COUNT(DISTINCT ?street)
    AS ?streetcount) WHERE {
    ?item wdt:P31 wd:Q5 .
    ?item wdt:P21 wd:Q6581072 .
    ?item wdt:P18 ?image .
    ?street wdt:P138 ?item .
    ?street wdt:P31 wd:Q79007 .
    ?street wdt:P17 wd:Q55 .
    SERVICE wikibase:label {
        bd:serviceParam wikibase:language "en" .
        ?item rdfs:label ?name .
    }
}
GROUP BY ?item ?name ?image
ORDER BY DESC(?streetcount)
LIMIT 10
```



591  
[commons:Jacoba Maria Wilhelmina van der Meulen](#)  
[Q:Q150747](#)



515  
[commons:Prinses Beatrix.jpg](#)  
[Beatrix of the Netherlands](#)  
[Q:Q29574](#)



370  
[commons:Prinses Irene 1978.jpg](#)  
[Princess Irene of the Netherlands](#)  
[Q:Q263623](#)



370  
[commons:Prinses Irene van Bourbon-Parme](#)  
[Princess Irene of the Netherlands](#)  
[Q:Q263623](#)



255  
[commons:Emma van Waldeck-Pyrmont](#)  
[Emma of Waldeck and Pyrmont](#)  
[Q:Q57664](#)



255  
[commons:Prinses Christina eer...](#)  
[Princess Christina of the Netherlands](#)  
[Q:Q435324](#)



190  
[commons:Queen Juliana of the Netherlands](#)  
[Juliana of the Netherlands](#)  
[Q:Q154946](#)



160  
[commons:Margriet der Nederlanden](#)  
[Princess Margriet of the Netherlands](#)  
[Q:Q268821](#)



67  
[commons:Aletta Jacobs, 1895-1905.jpg](#)  
[Aletta Jacobs](#)  
[Q:Q463478](#)



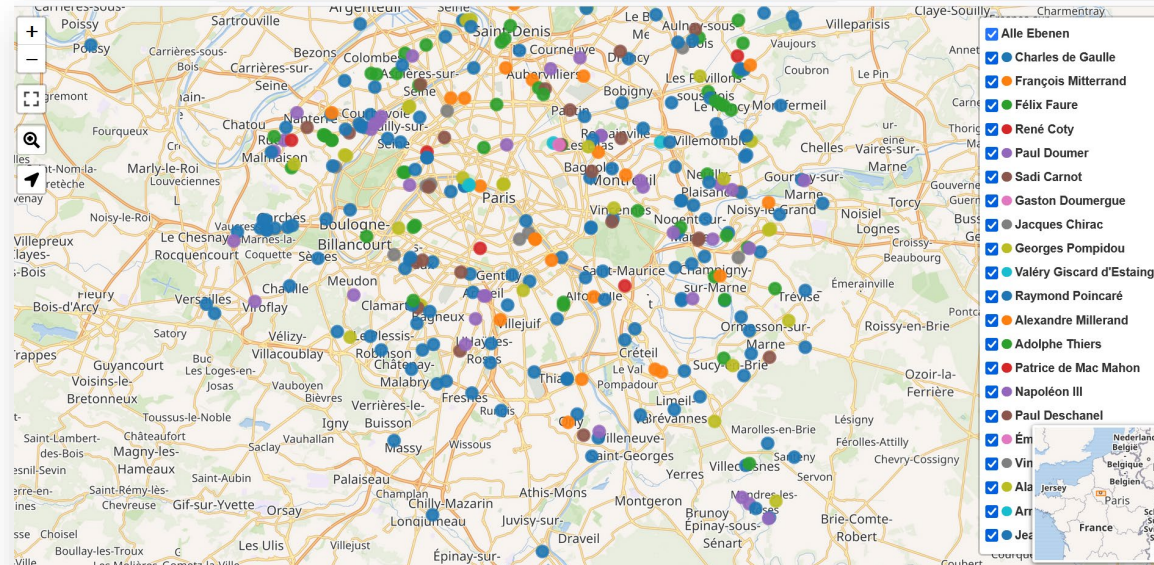
60  
[commons:Marie Curie c. 1920...](#)  
[Marie Curie](#)  
[Q:Q7186](#)



# Antwort als Karte



- Was und wo auf der Welt wurde nach früheren französischen Präsidenten benannt?



<https://t.co/LHkfuHLiq2>

```
#defaultView:Map
```

```
SELECT ?truc ?presLabel ?trucLabel ?coord ?player WHERE {{  
  SELECT DISTINCT ?truc (SAMPLE(?coord) AS ?coord) (SAMPLE(?player) AS ?player) WHERE {  
    ?pres wdt:P39 wd:Q191954 ; rdfs:label ?player.    FILTER((LANG(?player)) = "fr")  
    ?truc wdt:P138 ?pres ; wdt:P625 ?coord.}  
  GROUP BY ?truc ?trucLabel}  
SERVICE wikibase:label { bd:serviceParam wikibase:language "fr".}}
```

# Komplexe Anfragen



- Welche Gemälde von Johannes Vermeer zeigen eine Landkarte?

```
#defaultView:ImageGrid
SELECT ?painting ?paintingLabel ?object ?objectlabel ?image
where {
  ?painting wdt:P31/wdt:P279* wd:Q3305213 .
  ?painting wdt:P180 ?object .
  ?painting wdt:P170 wd:Q41264 .
  ?object wdt:P31/wdt:P279* wd:Q4006 .
  ?painting wdt:P18 ?image .
  SERVICE wikibase:label {
    bd:serviceParam
    wikibase:language "en" }}

```



commons:Vermeer - Woman with a Lute near...  
 wd:Q28873626  
 Woman with a Lute



commons:Johannes Vermeer - 'De liefdesbri...'  
 wd:Q28873420  
 The Love Letter



commons:Jan Vermeer van Delft 019.jpg  
 wd:Q28873546  
 Woman with a Water Jug



commons:Vermeer, Johannes - Woman readin...  
 wd:Q28873420  
 Woman in Blue Reading a Letter



commons:Johannes Vermeer - De Soldaat en het L...  
 wd:Q28873420  
 Officer and Laughing Girl



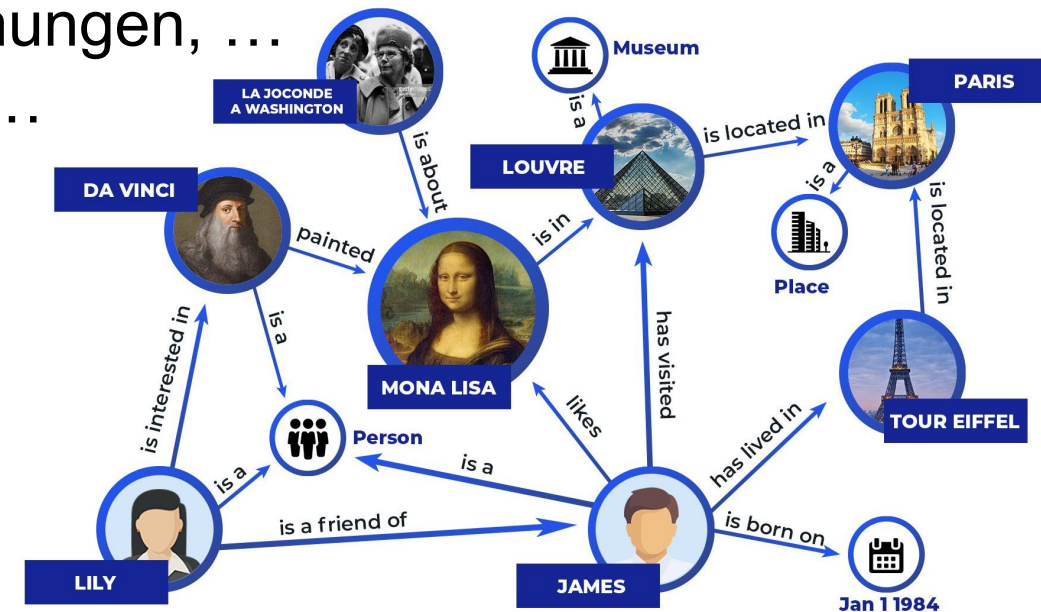
commons:Johannes Vermeer - The Astronomer - ...  
 wd:Q22999223  
 The Astronomer

# Domänenspezifischer Wissensgraph



## Am Beispiel Kunstgeschichte

- Kunsthistoriker/innen, sowie Laien können Wissensgraphen explorieren und neue, nicht-offensichtliche Muster und Beziehungen entdecken
- Spezifische Entitätstypen und Relationen
  - Kunstwerke, Gallerien, Museen, Kunstströmungen, ...
  - Beeinflusst\_von, Teil\_der\_Kunstströmung, ...
- Existierende, allgemeine Kunst-Ontologie
  - Passend? Verständlich? Vollständig?
- Maßgeschneiderte Ontologie?
  - Nicht-Trivial und teuer!





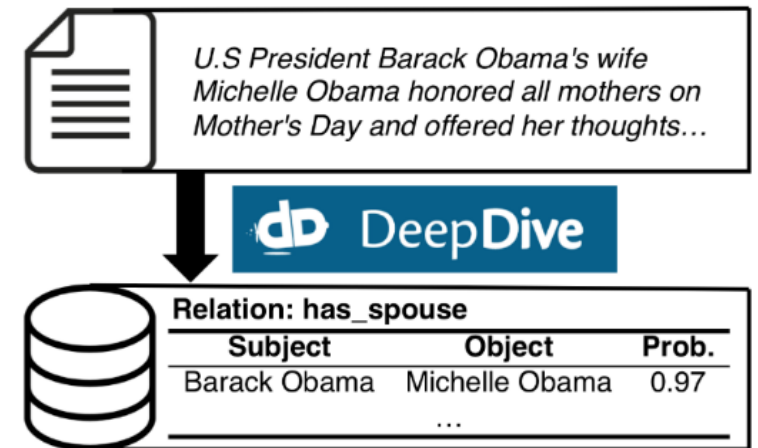
- # Ontologien

- The diagram illustrates the relationships between various entities and vocabularies in the context of the Rembrandt Harmensz. van Rijn painting "Jeremiah Lamenting the Destruction of Jerusalem".
- Entities and Relationships:**
- edm:WebResource** (Image of the painting) is linked to **edm:Agent** (Rijksmuseum) via **edm:isShownBy**.
  - edm:Agent** (Rijksmuseum) is linked to **edm:ProvidedCHO** (Rembrandt Harmensz. van Rijn) via **edm:aggregatedCHO**.
  - edm:ProvidedCHO** is linked to **edm:Agent** (Rembrandt Harmensz. van Rijn) via **dc:creator**.
  - edm:ProvidedCHO** is linked to **edm:Agent** (Rijksmuseum) via **edm:isShownBy**.
  - edm:ProvidedCHO** is linked to **edm:Agent** (Rembrandt Harmensz. van Rijn) via **dc:title**.
  - edm:ProvidedCHO** is linked to **edm:Agent** (Rijksmuseum) via **dc:subject**.
  - edm:ProvidedCHO** is linked to **edm:Agent** (Rembrandt Harmensz. van Rijn) via **dc:format**.
- Vocabularies:**
- skos:Concept** (aat:300015050) is linked to **skos:Concept** (aat:1000014078-en) via **skos:prefLabel**.
  - skos:Concept** (ic:71077) is linked to **skos:Concept** (ic:71) via **skos:broader**.
  - skos:Concept** (ic:71077) is linked to **skos:Concept** (ic:71) via **skos:prefLabel**.
  - skos:Concept** (ic:71) is linked to **skos:Concept** (ic:71) via **skos:prefLabel**.
  - skos:Concept** (ic:71) is linked to **skos:Concept** (ic:71) via **skos:prefLabel**.
- The diagram shows how the painting is linked to the Rijksmuseum, the artist, and the provided CHO, and how these entities are linked to various vocabularies.

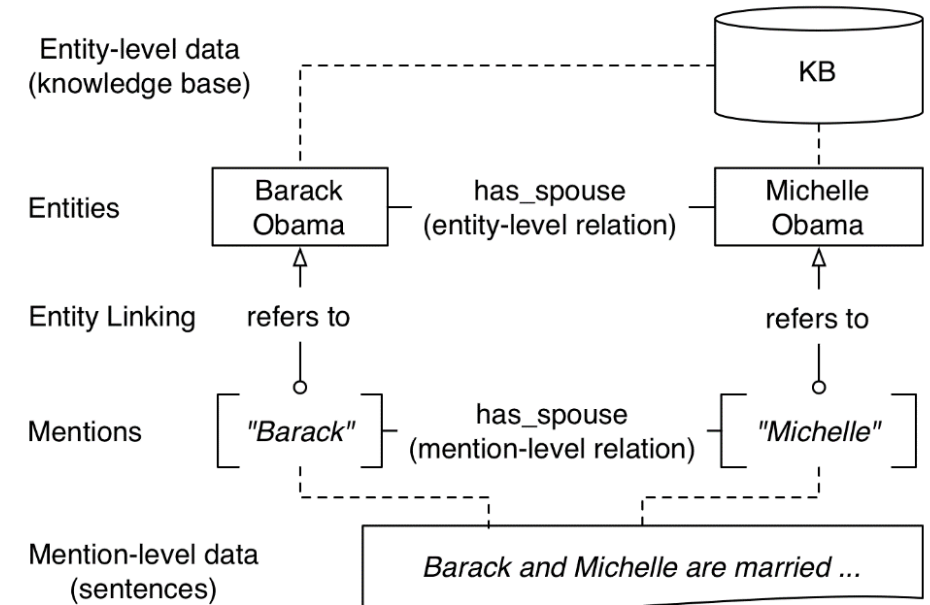


# Aufgaben / Anwendungen

- Wissensgraphkonstruktion
  - Aufbau von Grund auf
  - Existierender Wissensgraph vervollständigen
  - Benannte Entitäten
    - Erkennen (NER)
    - Typisieren
    - Disambiguieren (NED) / Verlinken (NEL)
  - Relationen
    - Extrahieren (RE)
    - Verfeinern, Zusammenführen,...
- Reasoning / Inference
  - QA, Chatbots, IR
- Wissensrepräsentation / KG Embeddings



<http://deepdive.stanford.edu/kbc>





# Potenzielle Anwendungesgebiete für LLMs



- Unterstützung beim Benutzen von KGs
  - Automatisch SPARQL-Anfragen aus natürlicher Sprache generieren
  - Explorieren und Zusammenfassen existierender KGs
  - Kompetenzfragen in SPARQL-Anfragen übersetzen (zur Evaluierung von KGs)
  - Code-Generierung oder Konfiguration von Verarbeitungspipelines
- Unterstützung bei KG-Konstruktion
  - KGs bevölkern
  - Erzeugen oder Erweitern von KG-Schemas/Ontologien
  - Erkennen von KG-Design-Problemen durch Probleme des LLM
  - Suche nach Konzepten/Entitäten in anderen, existierenden KGs
  - Verbessern von KGs durch Kompetenzfragen

Meyer, L. P., Stadler, C., Frey, J., Radtke, N., Junghanns, K., ... & Martin, M. (2023). LLM-assisted knowledge graph engineering: Experiments with chatGPT. In *Working conference on AI Development for a Resilient and Sustainable Tomorrow* (pp. 103-115).

- Wissensgraphen
  - Speichern Faktenwissen
  - Klare Struktur und klar definierte Semantik
  - Erlaubt strukturierte (komplexe) Anfragen
  - Erlaubt manuelles korrigieren
  - Selbsterklärend, transparent, verständlich
- Wissensgraphkonstruktion
  - Allgemein oder domänenspezifisch
  - Automatische Erstellung & Erweiterung / Vervollständigung
  - Klassische Methoden für jeden Schritt (NER, RE, ...)
- LLMs können unterstützen bei
  - Wissensgraphkonstruktion (einzelne Schritte, komplett?)
  - Wissensgraphbenutzung (Anfragen, Explorieren, Konfigurieren)

Aktuelle Forschung:  
Kombination von  
KGs und LLMs

- Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., Melo, G. D., ... & Zimmermann, A. (2021). **Knowledge graphs**. *ACM Computing Surveys (Csur)*, 54(4), 1-37.
- Zhong, L., Wu, J., Li, Q., Peng, H., & Wu, X. (2023). **A comprehensive survey on automatic knowledge graph construction**. *ACM Computing Surveys*, 56(4), 1-62.
- Meyer, L. P., Stadler, C., Frey, J., Radtke, N., Junghanns, K., ... & Martin, M. (2023). **LLM-assisted knowledge graph engineering: Experiments with chatGPT**. In *Working conference on AI Development for a Resilient and Sustainable Tomorrow* (pp. 103-115).
- Veseli, B.; Singhanian, S.; Razniewski, S.; Weikum, G. (2023). **Evaluating Language Models for Knowledge Base Completion**. In *Proceedings of the Semantic Web—20th International Conference, ESWC 2023; Volume 13870* (pp. 227–243).

## Webseiten zur (praktischen) Einführung

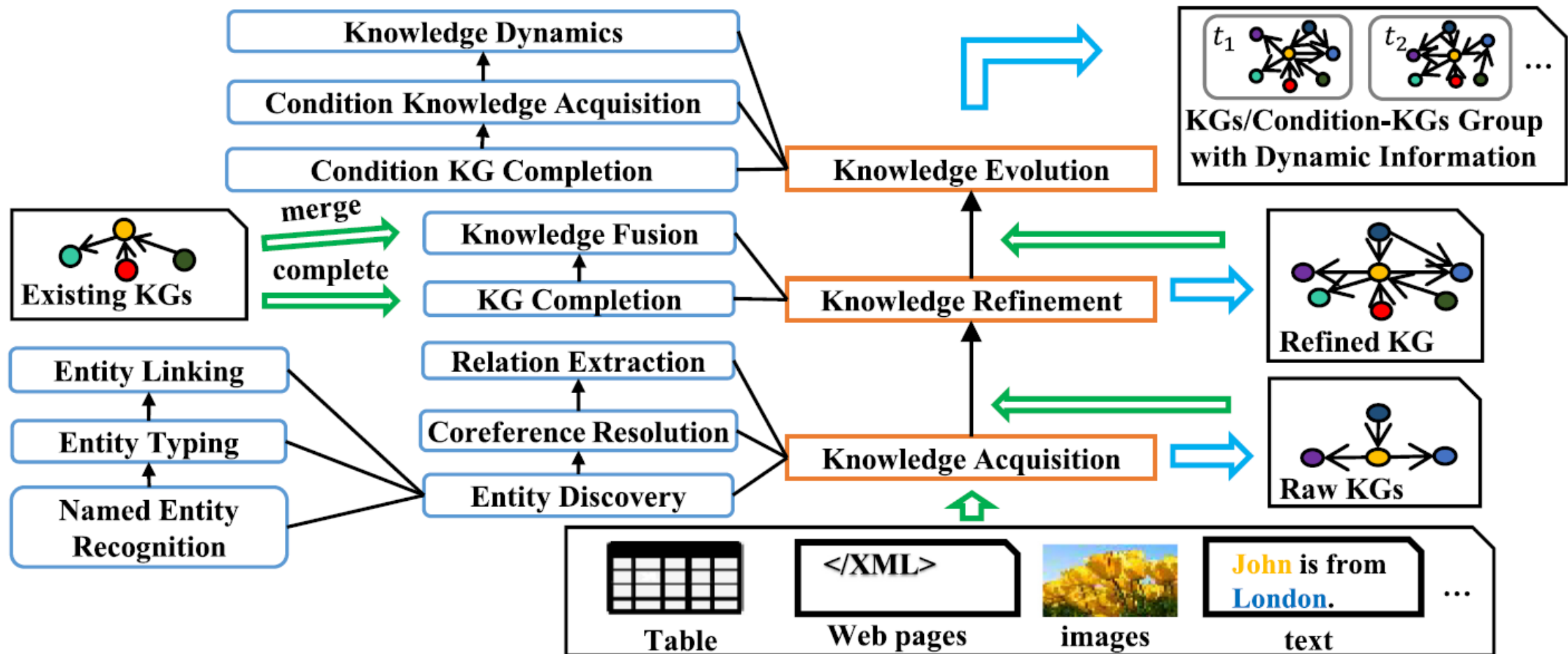
- <https://www.dataversity.net/what-is-a-knowledge-graph/>
- <https://towardsdatascience.com/a-guide-to-the-knowledge-graphs-bfb5c40272f1>
- <https://medium.com/@ayushlall/how-to-build-your-own-knowledge-graph-using-janusgraph-and-gremlin-1a0c69524cb>
- <https://medium.com/@jenlindadsouza/how-to-prompt-an-llm-to-generate-a-research-knowledge-graph-kg-for-complex-scientific-domains-2447fe8ff081>

# Bildernachweis



- <https://pxhere.com/en/photo/1452649>
- [https://commons.wikimedia.org/wiki/File:Lascaux\\_04\\_%28with\\_circle%29.jpg](https://commons.wikimedia.org/wiki/File:Lascaux_04_%28with_circle%29.jpg)
- [https://commons.wikimedia.org/wiki/File:Linecons\\_database.svg](https://commons.wikimedia.org/wiki/File:Linecons_database.svg)
- [https://commons.wikimedia.org/wiki/File:Bildschirmfoto\\_2021-02-14\\_um\\_15.59.17.png](https://commons.wikimedia.org/wiki/File:Bildschirmfoto_2021-02-14_um_15.59.17.png)
- [https://commons.wikimedia.org/wiki/File:William\\_Gladstone\\_as\\_a\\_jockey\\_falling\\_from\\_his\\_horse\\_Wellcome\\_V0050380.jpg](https://commons.wikimedia.org/wiki/File:William_Gladstone_as_a_jockey_falling_from_his_horse_Wellcome_V0050380.jpg)
- <https://community.atlassian.com/t5/image/serverpage/image-id/123489i7BBAEFC53942889D/image-dimensions/2000>

# Anhang: Wissensgraphkonstruktion



Zhong, L., Wu, J., Li, Q., Peng, H., & Wu, X. (2023). A comprehensive survey on automatic knowledge graph construction. *ACM Computing Surveys*, 56(4), 1-62.