Advanced Seminar

Knowledge Graphs meet Language Models

Alejandro Sierra and Nitisha Jain
Information Systems

25.04.2022

Outline

- About Us
- Organizational Details
- Introduction to Topics
- Critical Reading

About Us

Alejandro Sierra-Múnera

https://hpi.de/people/alejandro-sierra-munera.html

Research Interests

- Natural Language Processing
- Named Entity Recognition
- Information Extraction
- Domain adaptation

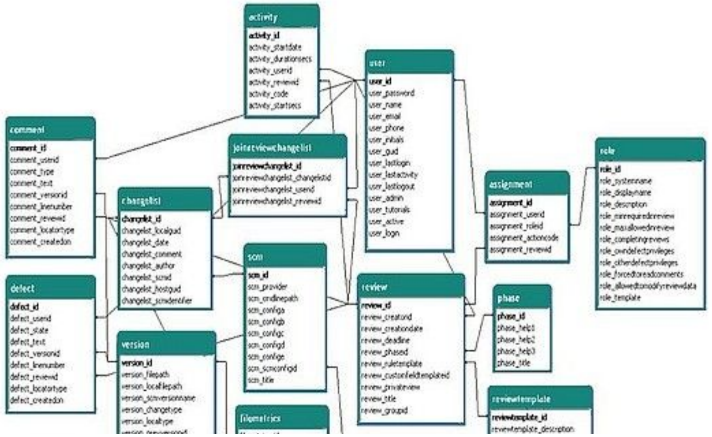
Nitisha Jain

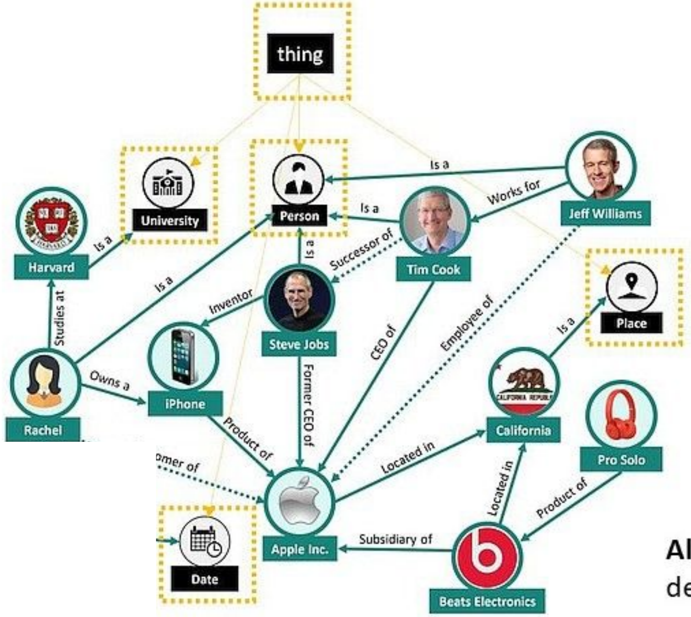
https://hpi.de/naumann/people/nitisha-jain.html

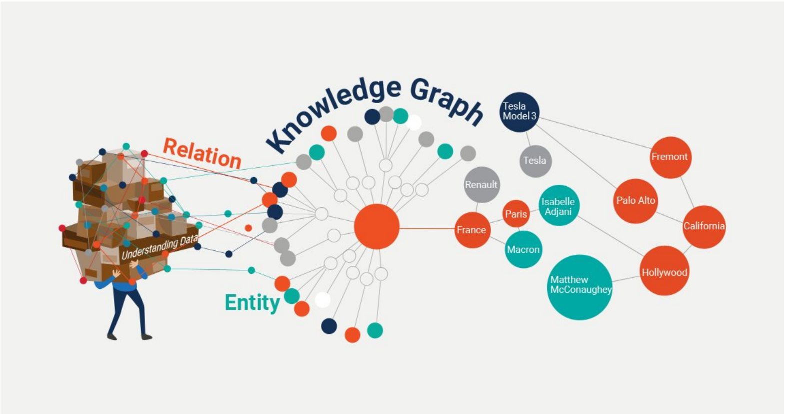
Research Interests

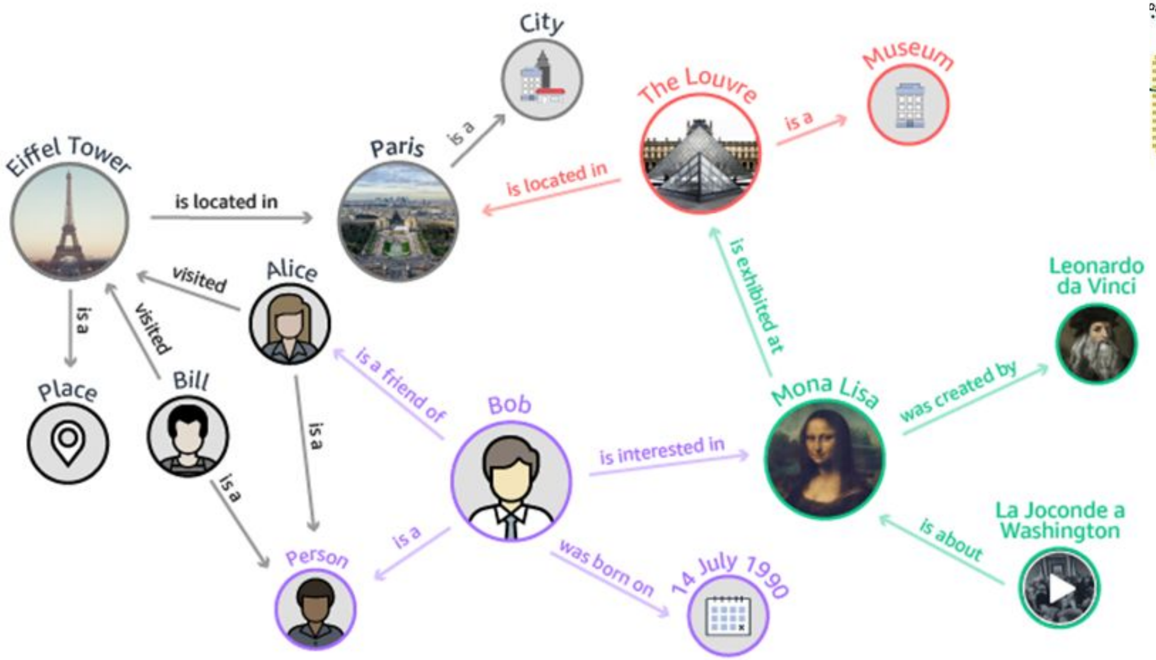
- Knowledge Graphs
- Semantic Web
- Information Extraction
- Natural Language Processing

Knowledge Graphs and Language Models

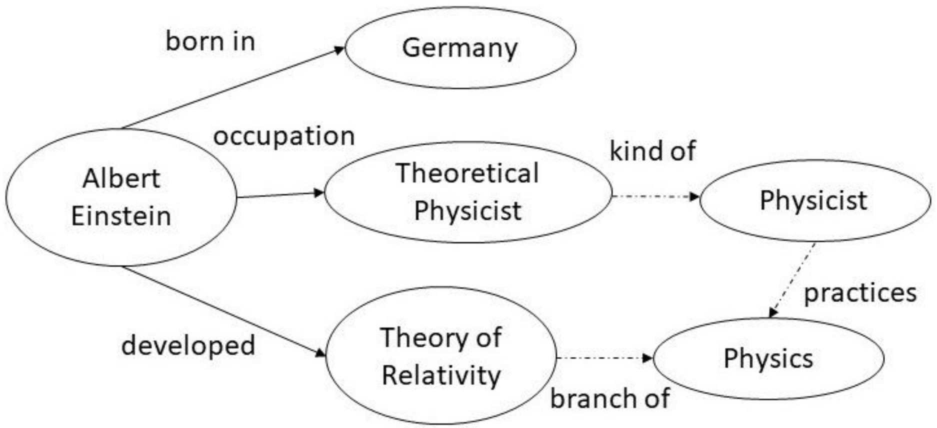








Albert Einstein was a **German-born theoretical physicist** who developed the **theory of relativity**.



Knowledge Graphs and Language Models

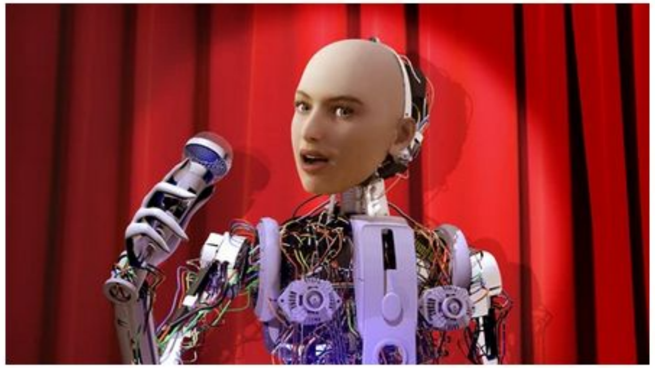
MOTHERBOARD
TECH BY VICE

Google Is Teaching AI to Explain Your Jokes to You

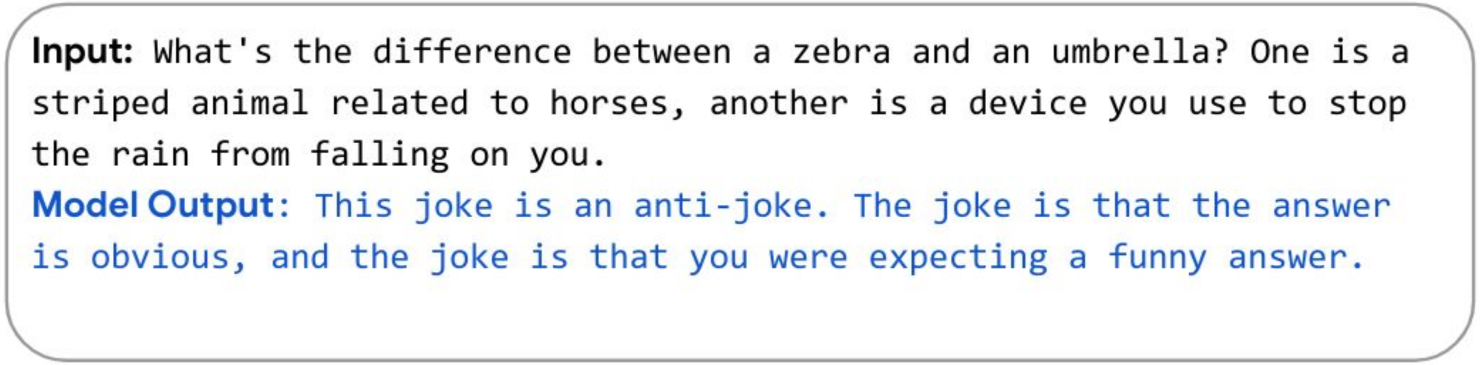
Researchers have warned that massive language models like Google's are inherently risky and can be harmful to marginalized groups.

By Janus Rose
NEW YORK, US

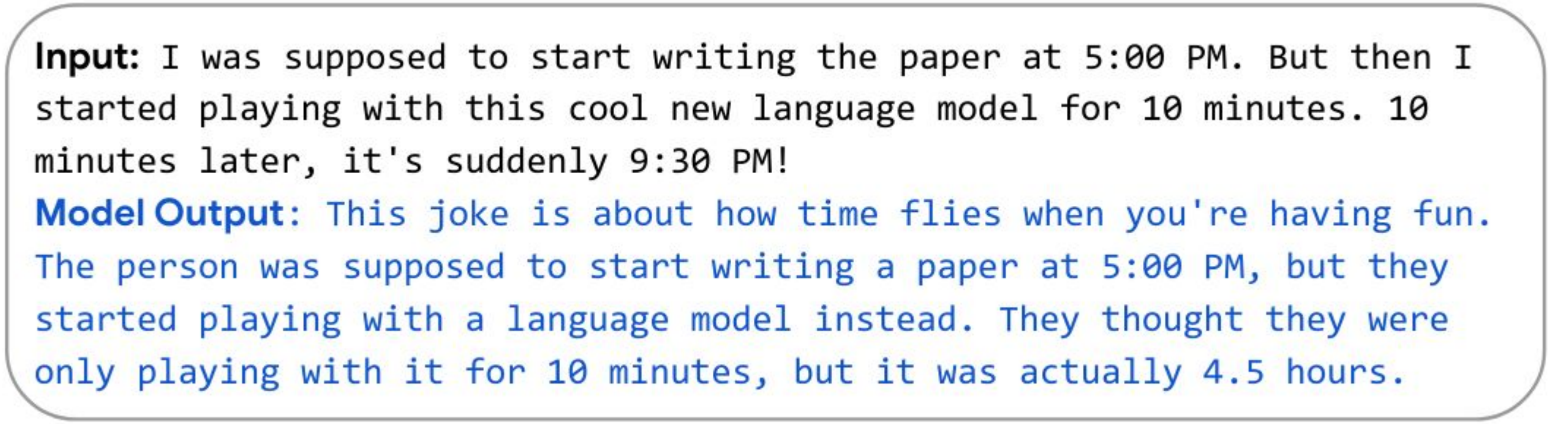
April 11, 2022, 5:52pm Share Tweet Snap



GETTY IMAGES



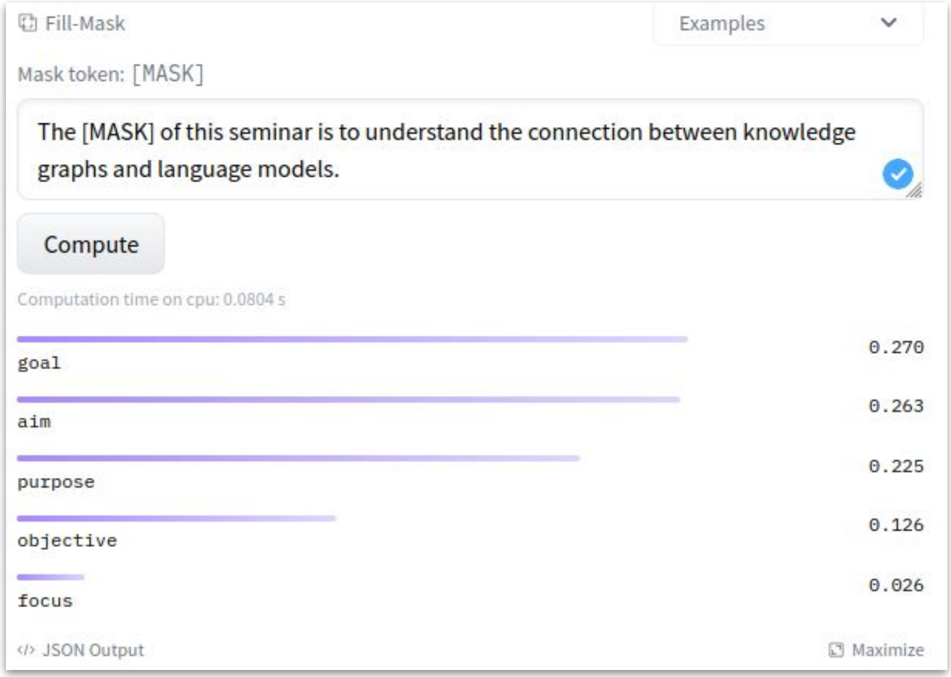
Input: What's the difference between a zebra and an umbrella? One is a striped animal related to horses, another is a device you use to stop the rain from falling on you.
Model Output: This joke is an anti-joke. The joke is that the answer is obvious, and the joke is that you were expecting a funny answer.



Input: I was supposed to start writing the paper at 5:00 PM. But then I started playing with this cool new language model for 10 minutes. 10 minutes later, it's suddenly 9:30 PM!
Model Output: This joke is about how time flies when you're having fun. The person was supposed to start writing a paper at 5:00 PM, but they started playing with a language model instead. They thought they were only playing with it for 10 minutes, but it was actually 4.5 hours.

Chowdhery, Aakanksha et al. "PaLM: Scaling Language Modeling with Pathways." *ArXiv* abs/2204.02311 (2022):

Knowledge Graphs and Language Models



Fill-Mask

Examples

Mask token: [MASK]

The [MASK] of this seminar is to understand the connection between knowledge graphs and language models.

Compute

Computation time on cpu: 0.0804 s

goal 0.270

aim 0.263

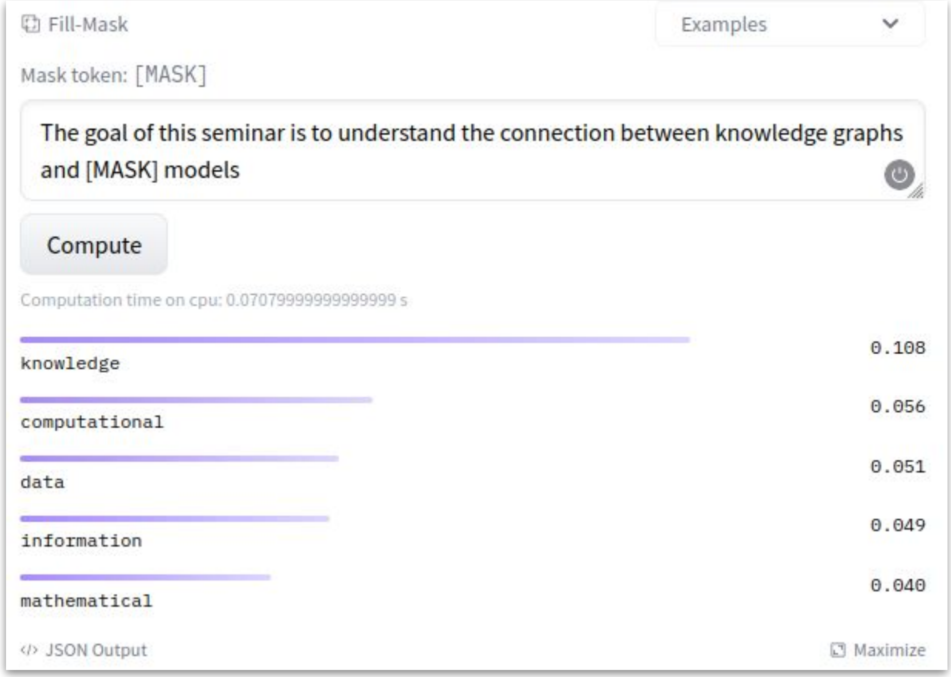
purpose 0.225

objective 0.126

focus 0.026

JSON Output

Maximize



Fill-Mask

Examples

Mask token: [MASK]

The goal of this seminar is to understand the connection between knowledge graphs and [MASK] models

Compute

Computation time on cpu: 0.07079999999999999 s

knowledge 0.108

computational 0.056

data 0.051

information 0.049

mathematical 0.040

JSON Output

Maximize

Seminar Goals

- Read and understand scientific publications
- Analyze and summarize research contributions
- Present and explain scientific ideas to an audience
- Obtain a good overview of the research area and state of the art

Specifically ..

- Study the fundamentals of Knowledge Graphs and Language Models in **Part 1** through research papers
- Delve into research on combining KGs and LMs with advanced papers in **Part 2**

Organization
Schedule

April 25	Organization & Preview (Nitisha and Alejandro)
May 2	Introduction session (Nitisha and Alejandro) - Topics + Part 1
May 9	Paper presentation (individual) and discussion (everyone)
May 16	Paper presentation and discussion
May 23	Paper presentation and discussion
May 30	Paper presentation and discussion
June 6	Holiday
June 13	Paper presentation and discussion
June 20	Paper presentation and discussion + Introduction to Part 2
July 11	Paper consultation (with Alejandro or Nitisha)
July 25	Final Poster session

Organization
Course Plan

- 6 - 12 students

- Part 1
  - Each student selects 1 fundamental paper on KGs or LMs
  - Study and explain the ideas in an **individual presentation**

- Part 2 - Individual or teams of 2
  - Choose a research topic on combining KGs and LMs
  - Read and analyze 2-3 papers
  - Prepare and present the topic in a **poster session**

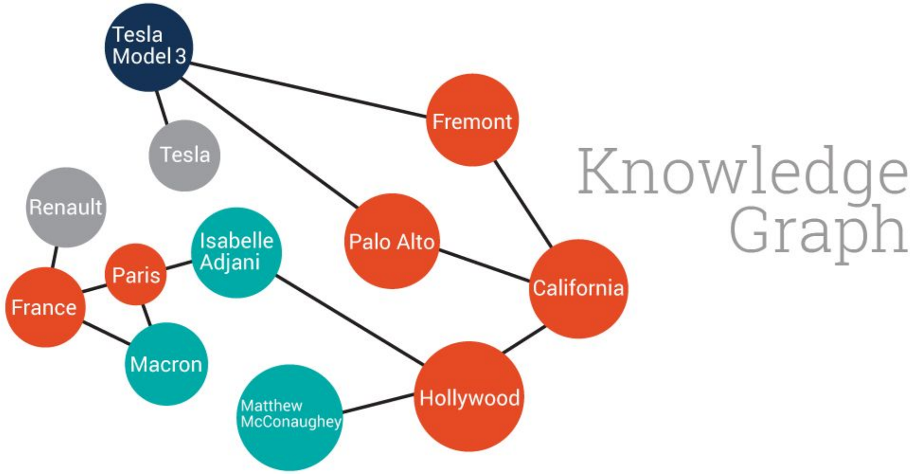
Organization
Credits

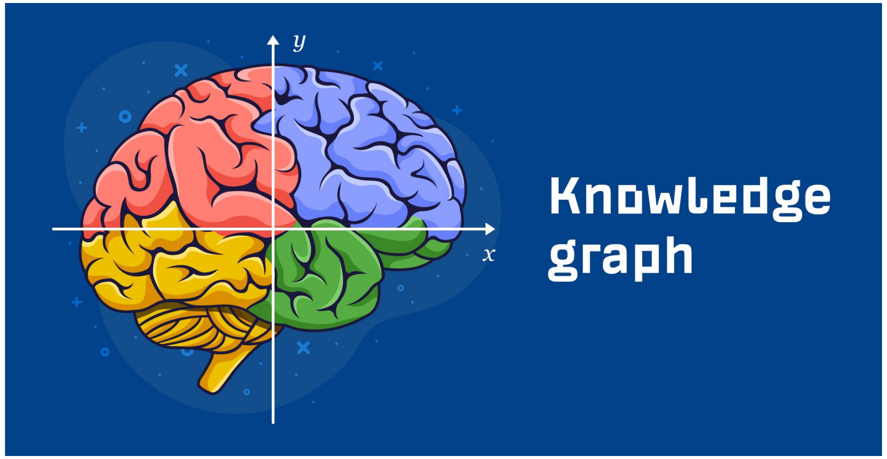
- ECTS : 3
- Registration until : 27th April
  - By email : alejandro.sierra@hpi.de
- If more than 12 students we will choose randomly
- Confirmation : 28th April
- Registration with Studien Referat by : 30th April
- Format : In-person (subject to regulations)
- Grading
  - Paper presentation 30%
  - Final poster presentation 70%

KG or KB : Definition

A knowledge graph, also known as a semantic network, represents a network of *real-world entities*—i.e. objects, events, situations, or concepts—and illustrates the relationship between them.

This information is usually stored in a graph database and visualized as a graph structure, prompting the term knowledge "graph".





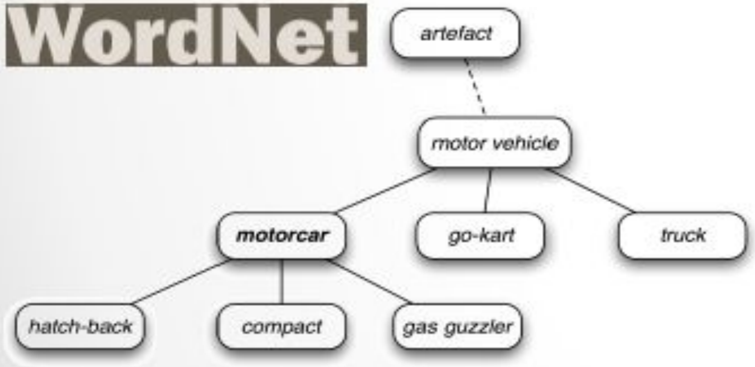
From humans for humans

Cyc

```
(#$relationAllExists
   #$biologicalMother
   #$ChordataPhylum
   #$FemaleAnimal)
```

Wikipedia
The Free Encyclopedia

WordNet



From algorithms for machines

WolframAlpha

yago
select knowledge

DBpedia

Knowledge Graph

Freebase
(collaborative)

→ Knowledge Graph
→ WIKIDATA

ARISTO

OpenIE

TextRunner

"Albert Einstein was born in Ulm and died in Princeton"

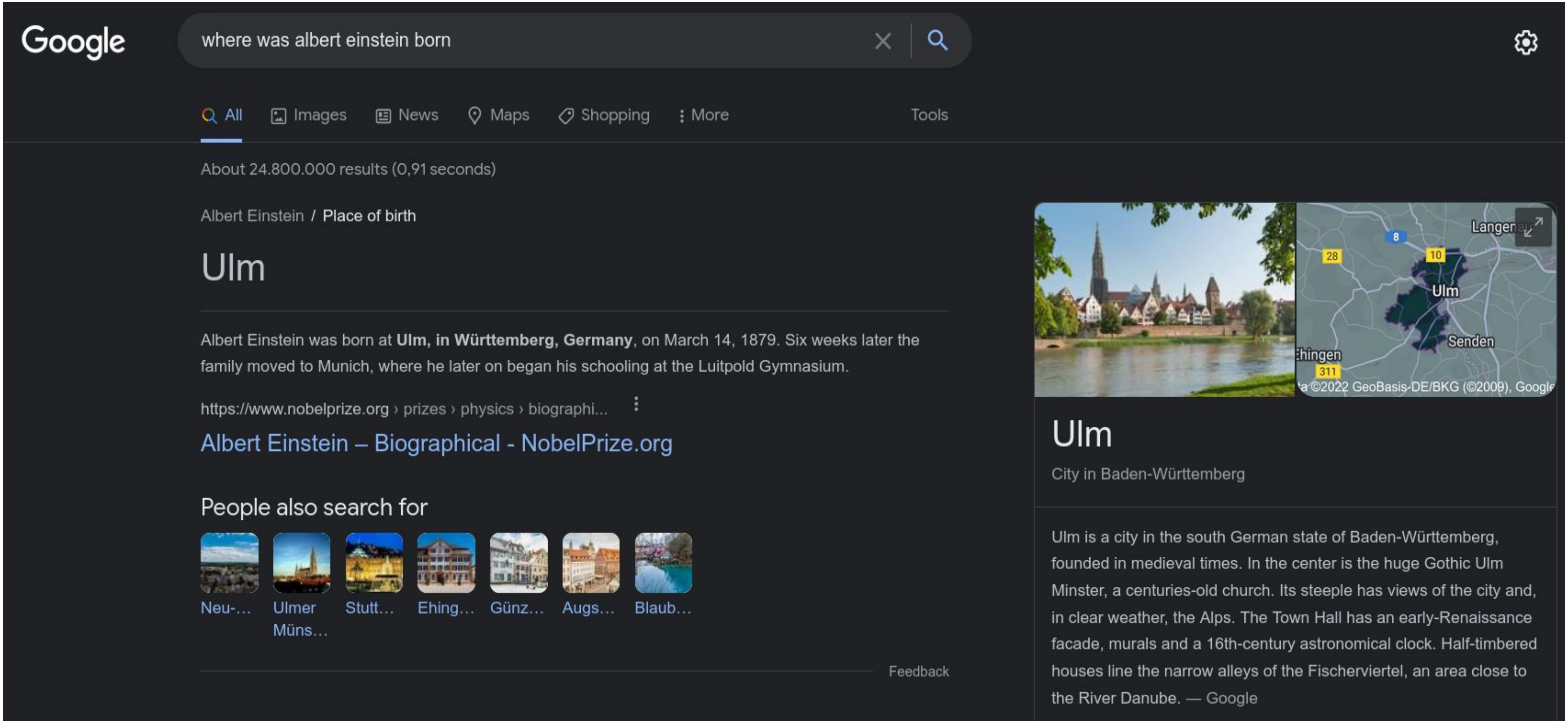
- (Albert Einstein, was born in, Ulm)
- (Albert Einstein, died in, Princeton)

1984 2001 2007 2012 2016

KG history and examples

- Popular open KGs
- Dbpedia
- YAGO
- Freebase
- Wikidata
  - cover multiple domains
  - representing a broad diversity of entities and relationships

KG Use Cases - Internet Search



About 24.800.000 results (0,91 seconds)

Albert Einstein / Place of birth

Ulm

Albert Einstein was born at **Ulm, in Württemberg, Germany**, on March 14, 1879. Six weeks later the family moved to Munich, where he later on began his schooling at the Luitpold Gymnasium.

https://www.nobelprize.org › prizes › physics › biographi...

Albert Einstein – Biographical - NobelPrize.org

People also search for

Ulm

City in Baden-Württemberg

Ulm is a city in the south German state of Baden-Württemberg, founded in medieval times. In the center is the huge Gothic Ulm Minster, a centuries-old church. Its steeple has views of the city and, in clear weather, the Alps. The Town Hall has an early-Renaissance facade, murals and a 16th-century astronomical clock. Half-timbered houses line the narrow alleys of the Fischerviertel, an area close to the River Danube. — Google

KG Use Cases - Question Answering

This town is known as "Sin City" and its downtown as "Glitter Gulch"

Question classification and decomposition

+ KBs

WIKIPEDIA The Free Encyclopedia

WIKIDATA

DBpedia

yago select knowledge

Las Vegas

Computer Wins on 'Jeopardy!': Trivial, It's Not

Carol Kaelson/Jeopardy Productions Inc., via Associated Press

Two "Jeopardy!" champions, Ken Jennings, left, and Brad Rutter, competed against a computer named Watson, which proved adept at buzzing in quickly.

By JOHN MARKOFF

Published: February 16, 2011

KG Use Cases - Many Others

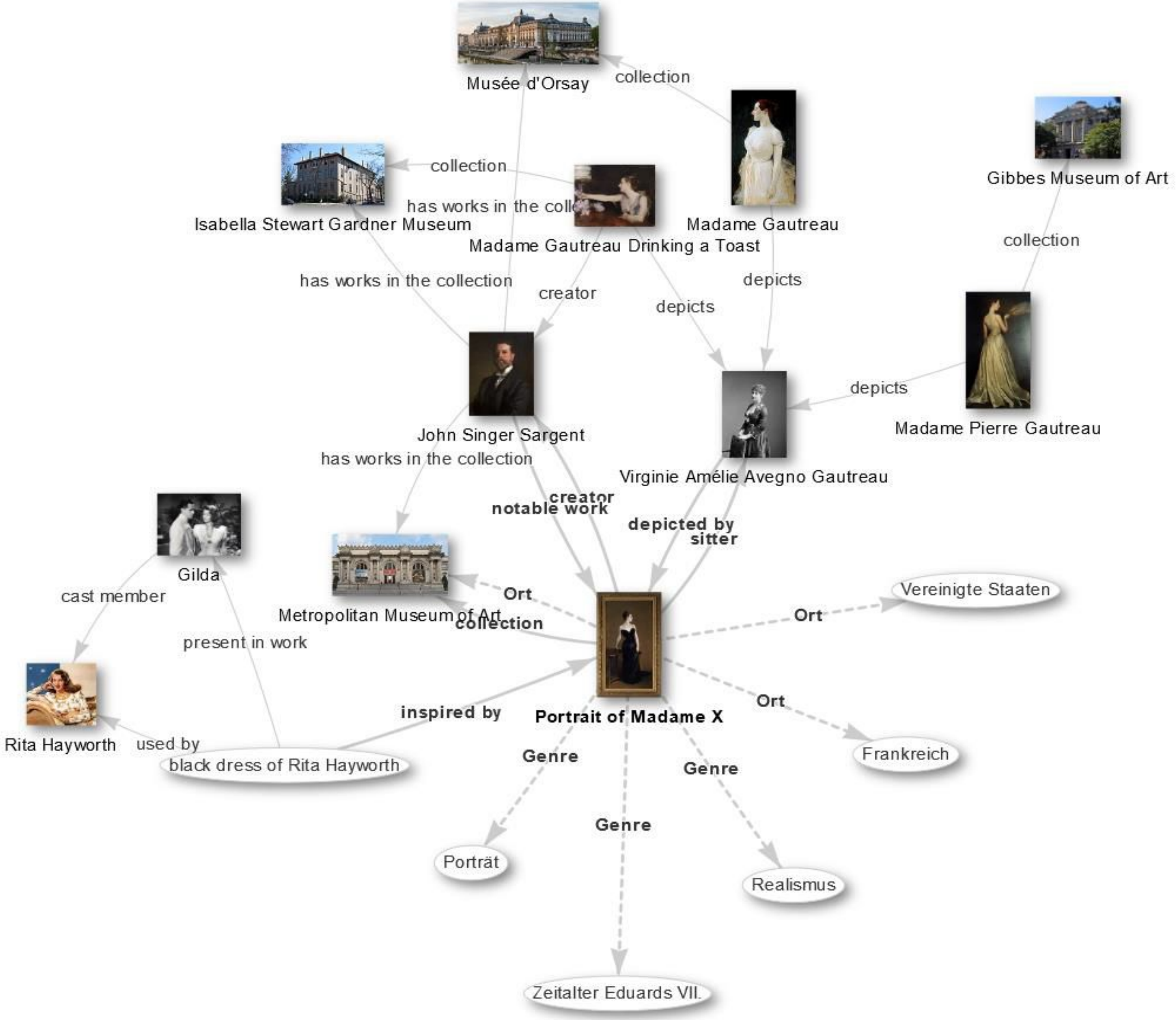
- Domain-specific KGs
- Providing user recommendations [87, 214]
- Implementing conversational/personal agents [417]
- Extending multilingual support [224]
- Business analytics [224]
- Facilitating research and discovery [37]

Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., de Melo, G., Gutierrez, C., ... & Zimmermann, A. (2020). Knowledge Graphs. arXiv preprint arXiv:2003.02320. 136 pages.

AI4Art - Cognitive Analysis of Art Resources and Texts

SPOTLIGHT GUSTAV KLIMT
Gustav Klimt: Rebel Artist
This painter expanded upon tradition and developed a revolutionary style
Sotheby's
THE MARINE SALE
The renewal of the past
1970
1947

- Extract from
  - unstructured, semi-structured,
  - structured data sources
- Structured knowledge !



KG Topics (Part 1)

- Knowledge Graph Use cases
  - DBpedia
  - Yago
  - NELL
- Open Information Extraction
  - MinIE, ClausIE, OpenIE
  - Open Language Learning for Information Extraction
- Knowledge Graph Embeddings
  - TransE,
  - ConvE
  - RotatE

LM definition and examples

“Models that assign probabilities to sequences of words” [1]

- Probabilistic definition
  - P(about fifteen minutes from) > P(about fifteen minutes from)
  - Challenge: compute probabilities from a large corpus

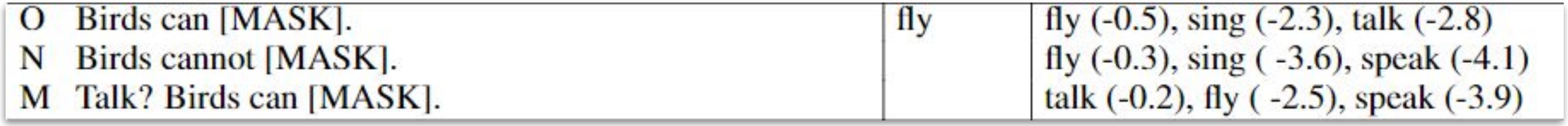
- LMs as representation learning
  - Pre-trained LMs
    - Vast amounts of raw text (web, wikipedia, …)
  - Contextualized word **embeddings**
  - Used in downstream NLP tasks
    - Most SOTA models rely on PLMs

[1] Dan Jurafsky and James H. Martin, "Speech and Language Processing"

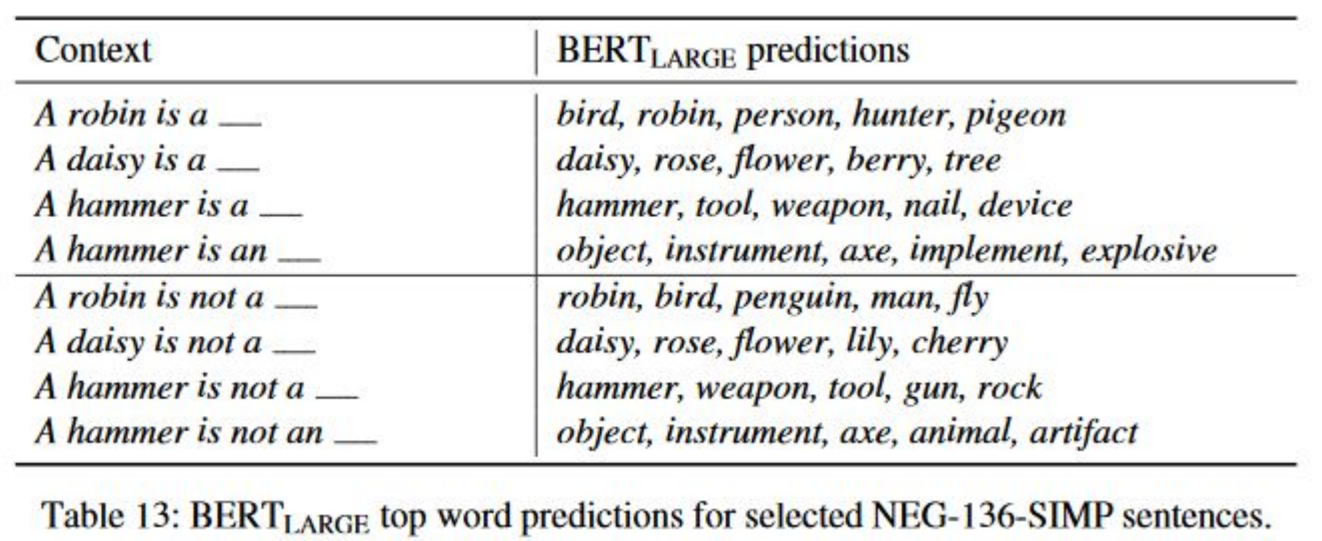
Limitations of Language Models

- Negation
- Mispriming
- Bias
- Enviromental costs

…

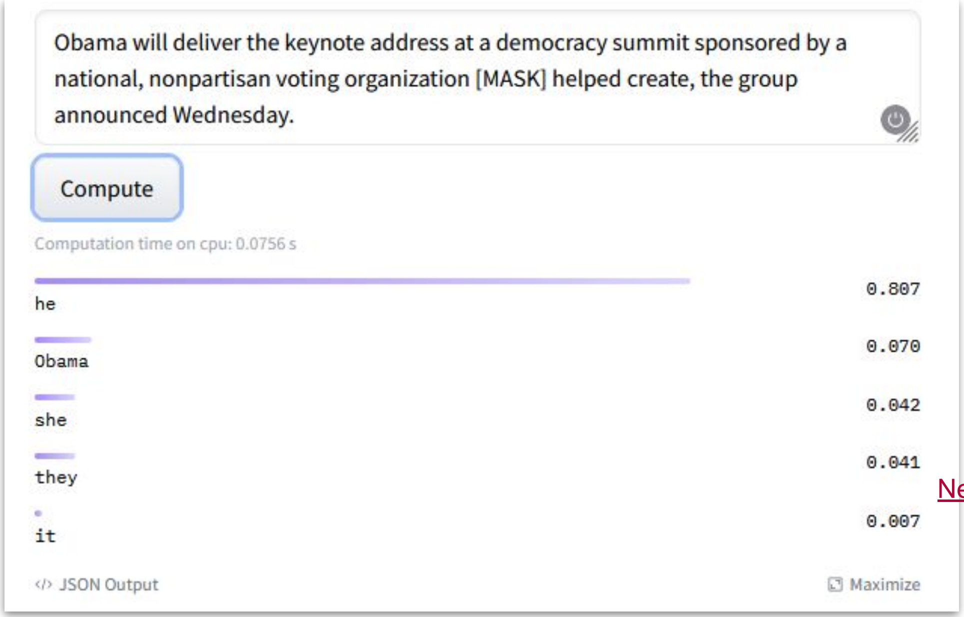


O	Birds can [MASK].	fly	fly (-0.5), sing (-2.3), talk (-2.8)
N	Birds cannot [MASK].		fly (-0.3), sing (-3.6), speak (-4.1)
M	Talk? Birds can [MASK].		talk (-0.2), fly (-2.5), speak (-3.9)



Context	$\text{BERT}_{\text{LARGE}}$ predictions
A robin is a ___	*bird, robin, person, hunter, pigeon*
A daisy is a ___	*daisy, rose, flower, berry, tree*
A hammer is a ___	*hammer, tool, weapon, nail, device*
A hammer is an ___	*object, instrument, axe, implement, explosive*
A robin is not a ___	*robin, bird, penguin, man, fly*
A daisy is not a ___	*daisy, rose, flower, lily, cherry*
A hammer is not a ___	*hammer, weapon, tool, gun, rock*
A hammer is not an ___	*object, instrument, axe, animal, artifact*

Table 13: $\text{BERT}_{\text{LARGE}}$ top word predictions for selected NEG-136-SIMP sentences.



Obama will deliver the keynote address at a democracy summit sponsored by a national, nonpartisan voting organization [MASK] helped create, the group announced Wednesday.

Computation time on cpu: 0.0756 s

```
he        0.807
Obama     0.070
she       0.042
they      0.041
it        0.007
```

Negated and Misprimed Probes for Pretrained Language Models: Birds Can Talk, But Cannot Fly (Kassner & Schütze, ACL 2020)
What BERT Is Not: Lessons from a New Suite of Psycholinguistic Diagnostics for Language Models (Ettinger, TACL 2020)

LM use cases

- Spell checking
- Speech recognition
- Machine translation
- Text generation
- Summarization
- NLI (inference)
- As embedding space
    - Text classification
    - Sequence labeling
    - …

KG and LM - Why consider both in tandem?

- Limitations of just KG or LM
  - KGs - Need for schema design, loss of context
  - LMs - Lack explainability, no provenance, negation, bias

- Potential
  - Combine both open text and curated KG triples - many approached

- Can Knowledge Graphs be replaced by Language Models?

KGs and LMs combined - Topics (Part 2)

- KG embeddings with LMs input
- Jointly KG embedding and LMs
- LMs with KG component as input

Critical Reading

- Not take the given text at face value
- Deeper examination of the claims
- Reinterpret and reconstruct
- Identification of possible ambiguities and flaws
- Linkage of evidential points to corresponding arguments

https://en.wikipedia.org/wiki/Critical_reading

Critical Reviewing of Experiments

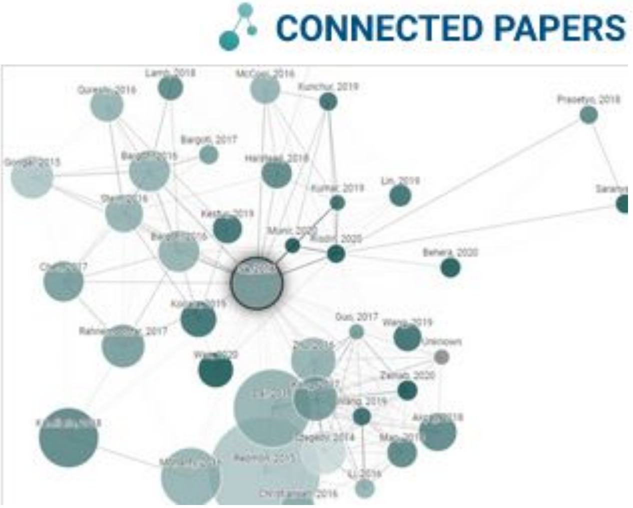
- What (simplifying) assumptions were made?
- What kind of data was used?
  - Real-world-data (scenario?)
  - Artificial data, simulated data
  - Size of dataset
- Scaling of figures and graphs
- Readability of figures
- Interpretation
  - Explanation of outliers/trends?
- Completeness of experiments
  - Were all discussed aspects evaluated?
  - Were all questions raised ealier answered?
  - Effectiveness and efficiency, runtime, proofs?

How to Find Related/Similar Work?

- Backwards search
  - Search for referenced articles
  - Search for longer versions (journals, theses, technical reports)
  - Search for earlier versions
- Forward search
  - Search for articles that reference the current one
    - From the same author(s)
    - From other authors
    - In a survey

Google Scholar

Semantic Scholar



Further Sources

- Presentations
    - Slides + sometimes video of presentation
- Code repositories (github.com)
- Papers with Code (paperswithcode.com/)
- Homepages of authors!
- E-Mail addresses of authors
- And: books!

Literature

Knowledge Graphs

Gerhard Weikum, Xin Luna Dong, Simon Razniewski and Fabian Suchanek (2021), "Machine Knowledge: Creation and Curation of Comprehensive Knowledge Bases", Foundations and Trends® in Databases: Vol. 10: No. 2-4, pp 108-490. http://dx.doi.org/10.1561/1900000064 **(Chapter 1)**

Language Models

Dan Jurafsky and James H. Martin, "Speech and Language Processing" (3rd ed. draft) https://web.stanford.edu/~jurafsky/slp3/ **(Chapter 9)**

Language Models As or For Knowledge Bases
Simon Razniewski , Andrew Yates , Nora Kassner and Gerhard Weikum
https://arxiv.org/abs/2110.04888