

Knowledge Graphs in the age of Large Language Models

Nuremberg 2023, Invited Talk

Aidan Hogan

DCC, Universidad de Chile

Instituto Milenio Fundamentos de los Datos



UNIVERSIDAD
DE CHILE



Instituto Milenio
Fundamentos
de los datos

The problem with the Web

Example 1.1

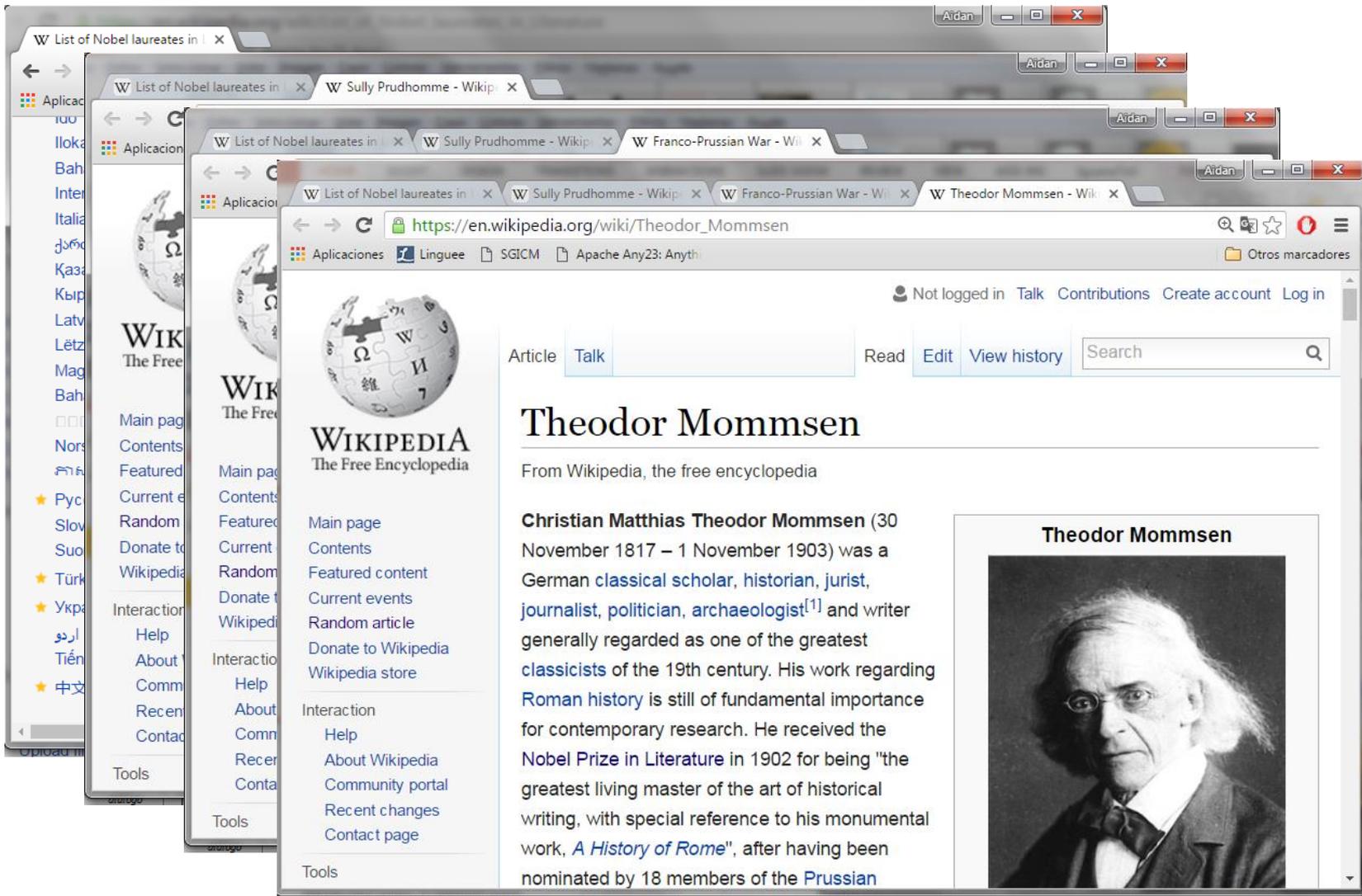


Julie is a literature student. For her thesis, she needs to find a list of Nobel Laureates in Literature who fought in a war, the year they were awarded the Nobel prize, and the name of the war(s) they fought in.



Aidan Hogan

The Web of Data



Web
Data

Nobel Laureates

We list Nobel laureates ordered by category then year.

Literature

Year	Awardee
1901	Sully Prudhomme
...	...
1953	Winston Churchill
1954	Ernest Hemingway
...	...
1964	Jean-Paul Sartre [†]
...	...
1983	William G. Golding
...	...
1985	Claude Simon
...	...

[†]Sartre declined.

Peace

Year	Awardee
1901	Henry Dunant
1901	Frédéric Passy
...	...

...

William G. Golding

Sir William Gerald Golding CBE (1911–1993) was a British novelist, playwright and poet. His most celebrated works include the novels "[Lord of the Flies](#)", and "[Rites of Passage](#)" for which he was awarded the [Booker Prize](#).

...

He joined the [Royal Navy](#) in 1940, commanding a landing ship at the [Normandy Invasion](#) on [D-Day](#).

...

Invasion of Normandy

On June 6th 1944 ([D-Day](#)), the [Western Allies](#) of [World War II](#) assaulted [Nazi](#) occupation forces from the sea at [Normandy](#), as part of [Operation Overlord](#).

...

What Nobel Laureates fought in WWII?

William G. Golding, ...

Aidan Hogan

The Web of Data

Lucys Sedgyefyz

Ny sozf Lucys sedgyefyz ugwygyw ci befyjugi fhyl iyeg.

Sofygefdgy

Iyeg	Enegwyw
1901	Zdssi Rgdwhummy
...	...
1953	Nolzful Bhdgbhoss
1954	Yglyzf Hymoljnei
...	...
1964	Vyel-Reds Zegfgy [†]
...	...
1983	Nossoem Juswolj
...	...
1985	Bsedwy Zomul
...	...

[†]Zegfgy wybsolyw.

Ryeby

Iyeg	Enegwyw
1901	Hylgi Wdlelf
1901	Pgýwýgob Rezzi
...	...

...

Nossoem J. Juswolj

Zog Nossoem Jygesw Juswolj BCY (1911–1993) nez e Cgofozh lutysozf, rseigojhf elw ruyf. Hoz muzf bysycgefyw nugqz olbsdwy fhy lutysz "[Sugw up fhy Psoyz](#)", elw "[Gofyz up Rezzejy](#)" pug nhobh hy nez enegwyw fhy "[Cuuqyg Rgoay](#)".

...

Hy vuolyw fhy [Guies Leti](#) ol 1940, bummelwolj e selwolj zhor ef fhy [Lugmelwi Oltezoul](#) ul [W-Wei](#).

...

Oltezoul up Lugmelwi

Ul Vdly 6th 1944 ([W-Wei](#)), fhy [Nyzfygl Essoyz](#) up [Nugsw Neg OO](#) ezzedsfyw [Leao](#) ubbdrefoul pugbyz pgum fhy zye ef [Lugmelwi](#), ez regf up [Urygefoul Utygsugw](#).

...

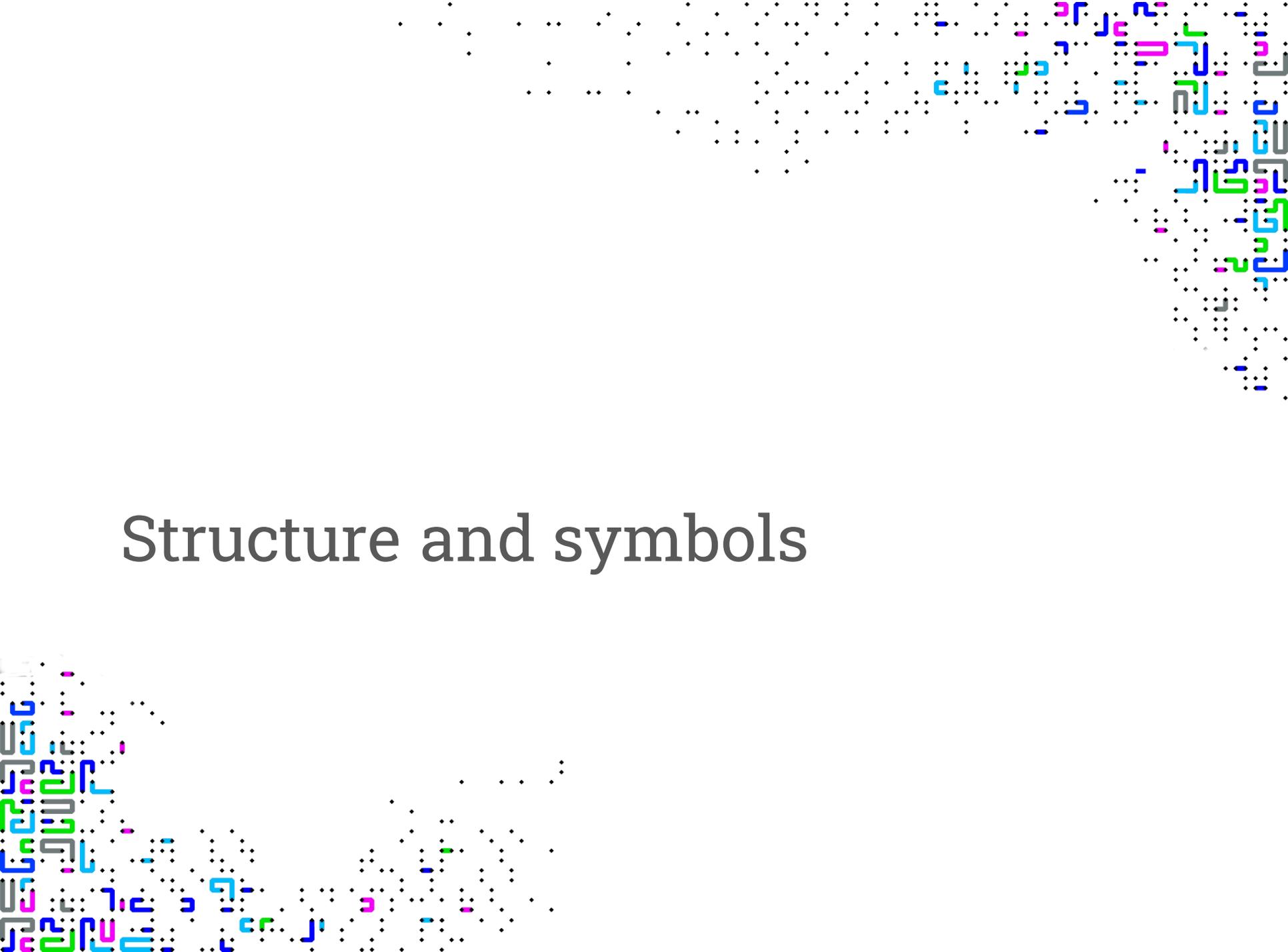
Ol nhobh negz hety Lucys Sedgyefyz pudjhf?

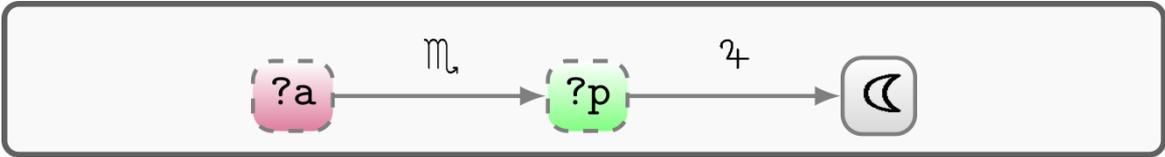
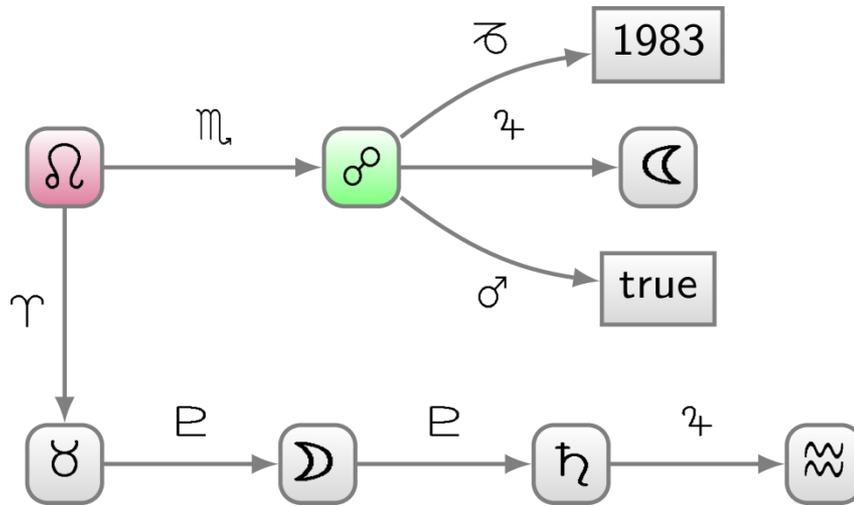
Nugsw Neg OO, ...

Aidan Hogan

The Web of Data

Structure and symbols

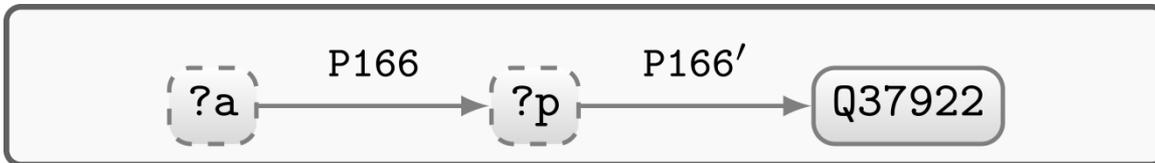
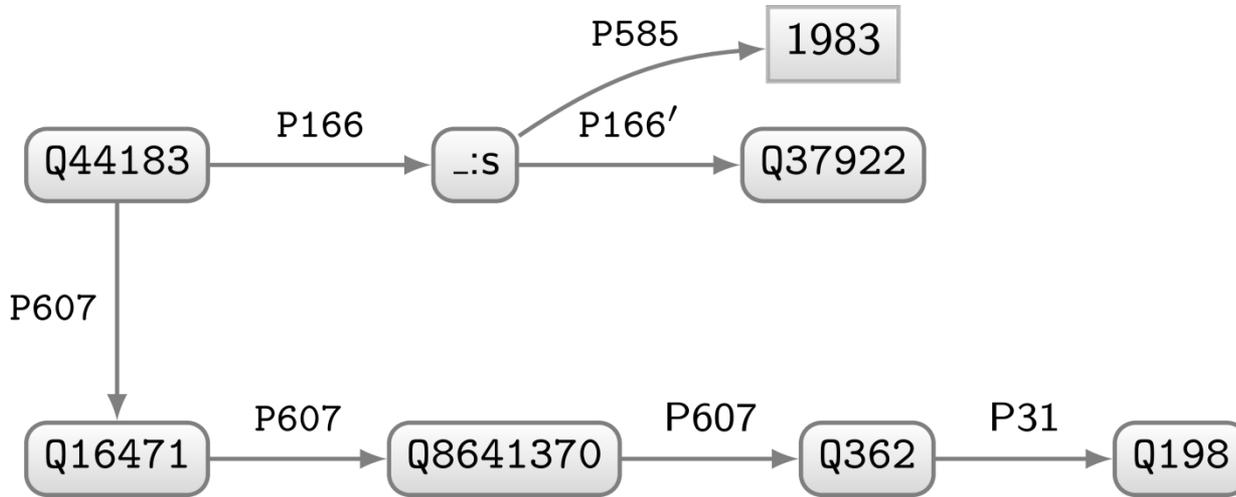




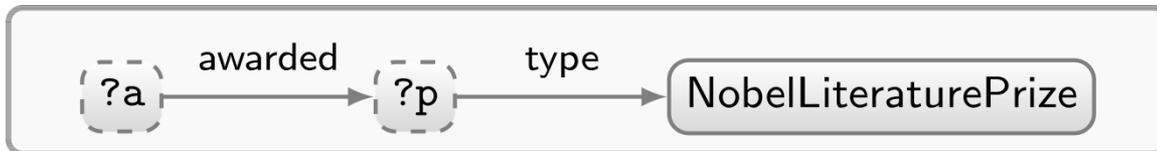
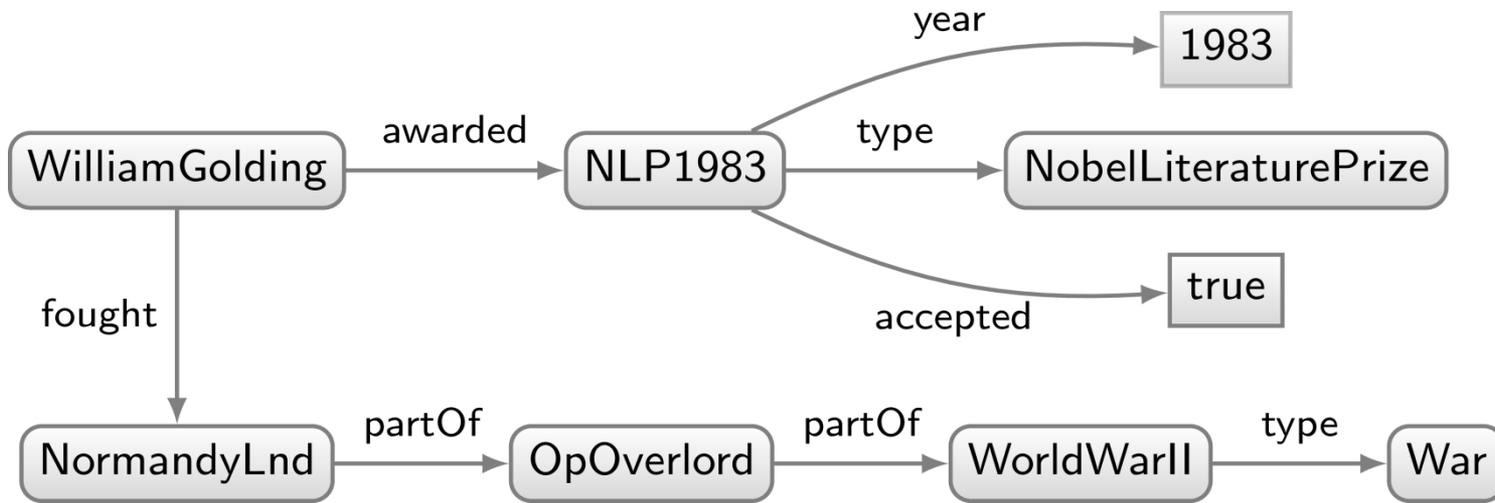
?a	?p
Ω	♂

Aidan Hogan

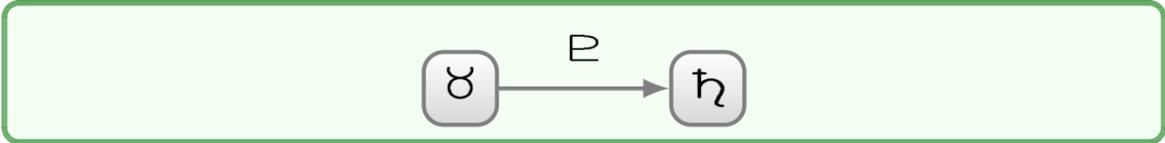
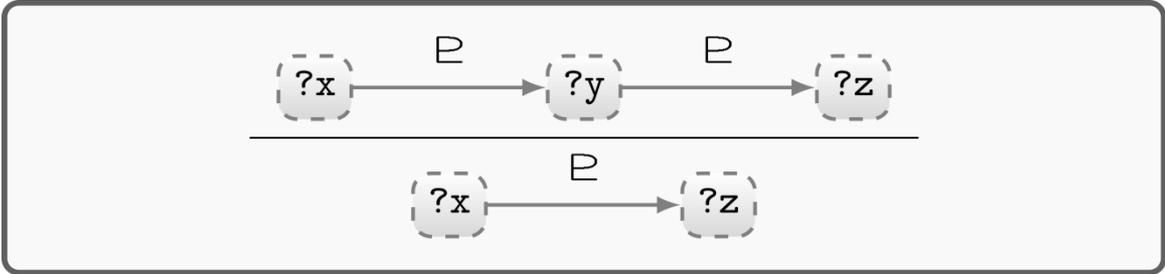
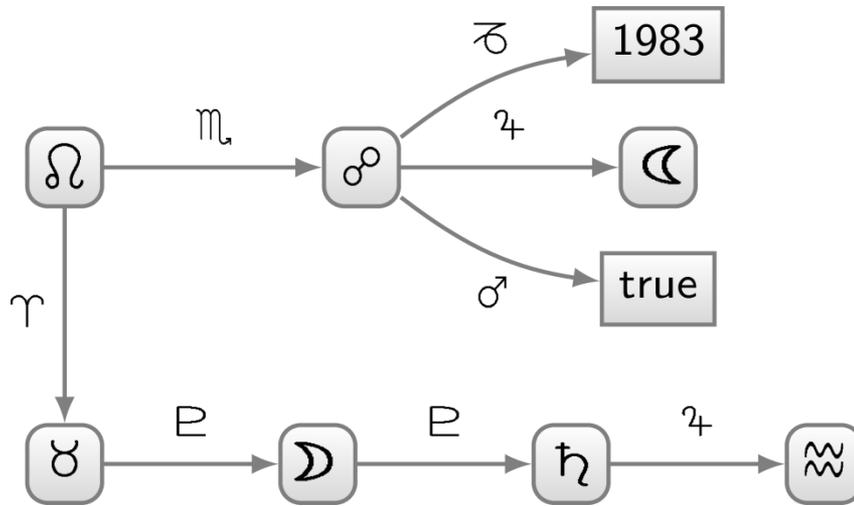
The Web of Data



?a	?p
Q44183	_:s

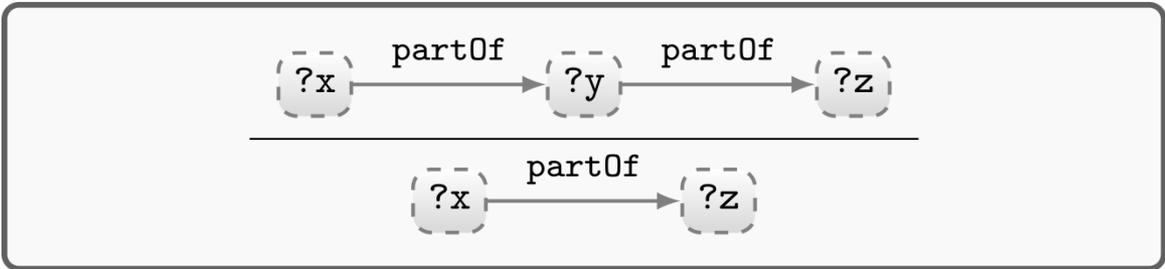
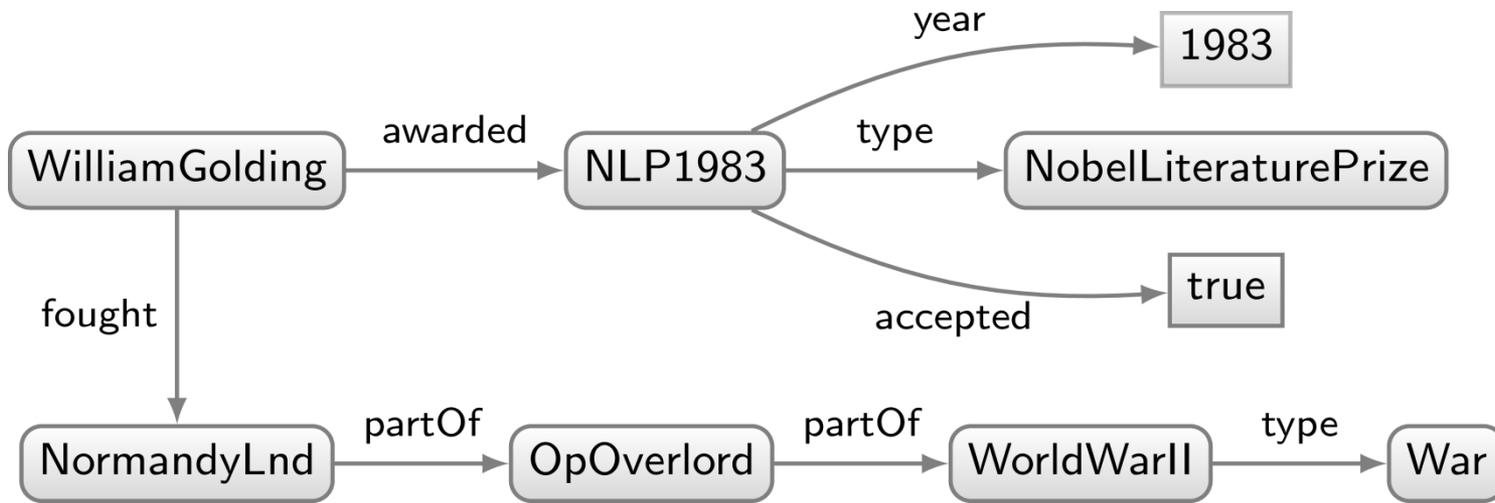


?a	?p
WilliamGolding	NLP1983



Aidan Hogan

The Web of Data



The image features a white background with decorative elements in the corners. In the top right and bottom left, there are clusters of small black dots and colorful lines (blue, green, pink, cyan) that resemble a stylized network or data visualization. The text "The problem with the Web?" is centered in a dark gray, sans-serif font.

The problem with the Web?

Lucys Sedgyefyz

Ny sozf Lucys sedgyefyz ugwygyw ci befjyugi fhyl iyeg.

Sofygefdgy

Iyeg	Enegwyw
1901	Zdssi Rgdwhummy
...	...
1953	Nolzful Bhdgbhoss
1954	Yglyzf Hymoljnei
...	...
1964	Vyel-Reds Zegfgy [†]
...	...
1983	Nossoem Juswolj
...	...
1985	Bsedwy Zomul
...	...

[†]Zegfgy wybsolyw.

Ryeby

Iyeg	Enegwyw
1901	Hylgi Wdlelf
1901	Pgýwýgob Rezzi
...	...

...

Nossoem J. Juswolj

Zog Nossoem Jygesw Juswolj BCY (1911–1993) nez e Cgofozh lutysozf, rseigojhf elw ruyf. Hoz muzf bysycgefyw nugqz olbsdwy fhy lutysz "[Sugw up fhy Psoyz](#)", elw "[Gofyz up Rezzejy](#)" pug nhobh hy nez enegwyw fhy "[Cuuqyg Rgoay](#)".

...

Hy vuolyw fhy [Guies Leti](#) ol 1940, bummelwolj e selwolj zhor ef fhy [Lugmelwi Oltezoul](#) ul [W-Wei](#).

...

Oltezoul up Lugmelwi

Ul Vdly 6th 1944 ([W-Wei](#)), fhy [Nyzfygl Essoyz](#) up [Nugsw Neg OO](#) ezzedsfyw [Leao](#) ubbdrefoul pugbyz pgum fhy zye ef [Lugmelwi](#), ez regf up [Urygefoul Utygsugw](#).

...

Ol nhobh negz hety Lucys Sedgyefyz pudjhf?

Nugsw Neg OO, ...

Aidan Hogan

The Web of Data

Information may be split over multiple clauses that use references such as pronouns that may be difficult to resolve; for example, compare the “*she*” in both:

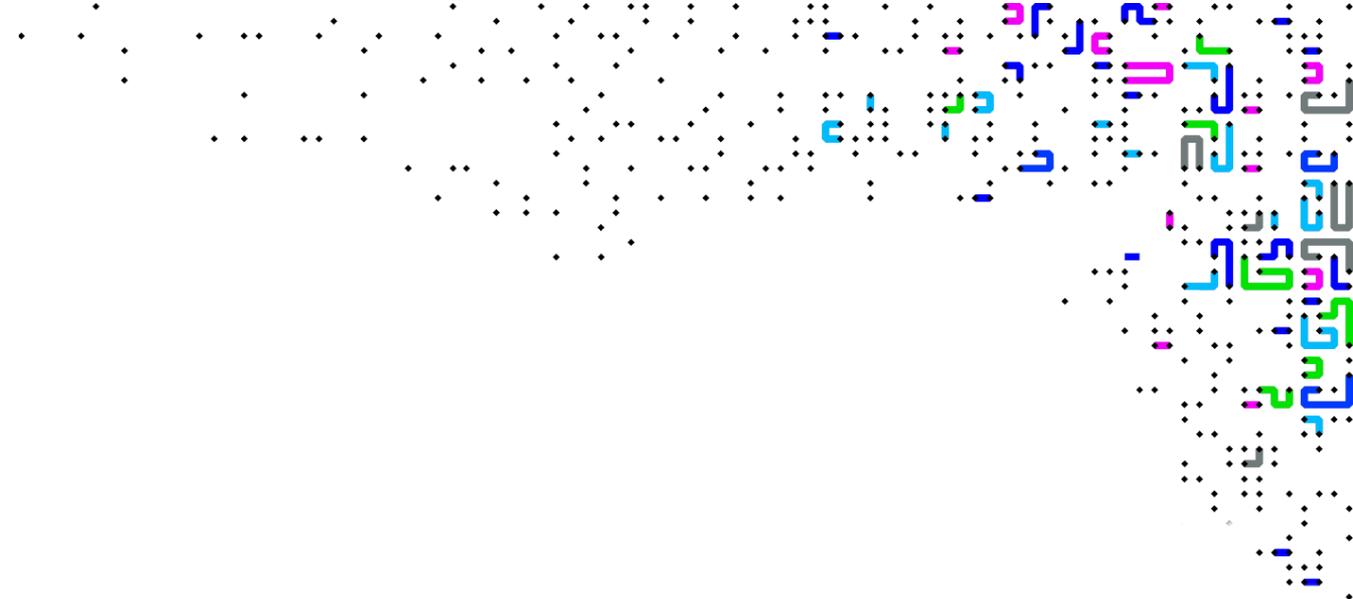
“Jane tried to follow her late grandmother’s recipe, but she had not included details of the cooking times.”

and

“Jane tried to follow her late grandmother’s recipe, but she had forgotten to buy lemons”.

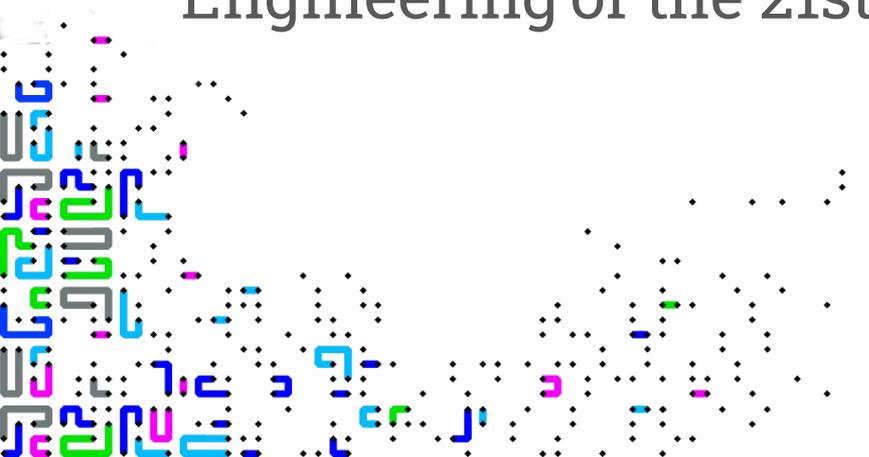
Aidan Hogan

The Web of Data



Dagstuhl Seminar 22372

Knowledge Graphs and Their Role in the Knowledge Engineering of the 21st Century





© SCHLOSS DAGSTUHL – LZI GMBH
licensed under Creative Commons License CC BY-NC-ND

KGC

KGC23 Keynote: The Future of Knowledge Graphs in a World of LLMs



Denny Vrandečić

Head of Special Projects

Wikimedia





A

Jane tried to follow her late grandmother's recipe, but she had not included details of the cooking times. Who forgot to include the cooking times: Jane or her grandmother?



Jane's late grandmother forgot to include the cooking times in the recipe.

A

Jane tried to follow her late grandmother's recipe, but she had forgotten to buy lemons. Who forgot to buy lemons: Jane or her grandmother?



Jane forgot to buy lemons.

A

Why was Jane's grandmother late?



The information provided in the previous questions does not mention why Jane's grandmother was late. If you could provide more context or details, I'd be happy to help you with a specific answer.

A

Jane's grandmother had passed away. Why was she late?



I apologize for any confusion, but based on the information provided, there is no indication or context given as to why Jane's grandmother might have been late. The information provided in the previous questions only mentions Jane trying to follow her late grandmother's recipe and Jane forgetting to buy lemons, but it doesn't provide any information about her grandmother being late or the reasons behind it. If you have more details or a specific scenario in mind, please provide additional information, and I'll do my best to assist you.

A

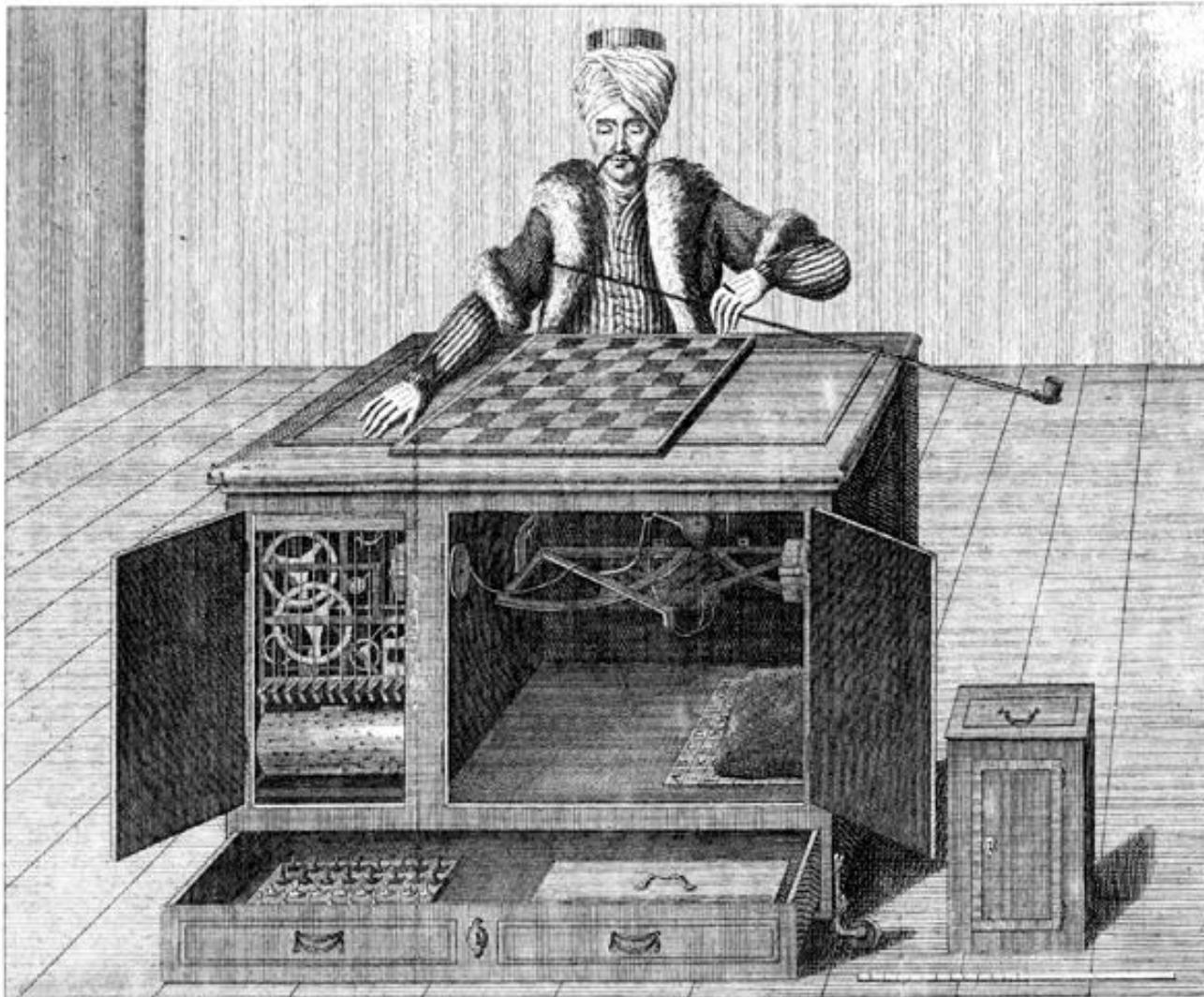
Why was Jane's grandmother late?



The phrase "late grandmother" typically means that Jane's grandmother has passed away, rather than being late in the sense of tardiness. In this context, "late" is used to indicate that the person is deceased. Therefore, the sentence doesn't provide information about Jane's grandmother being late for an event; it indicates that she is no longer alive.

Back to the future



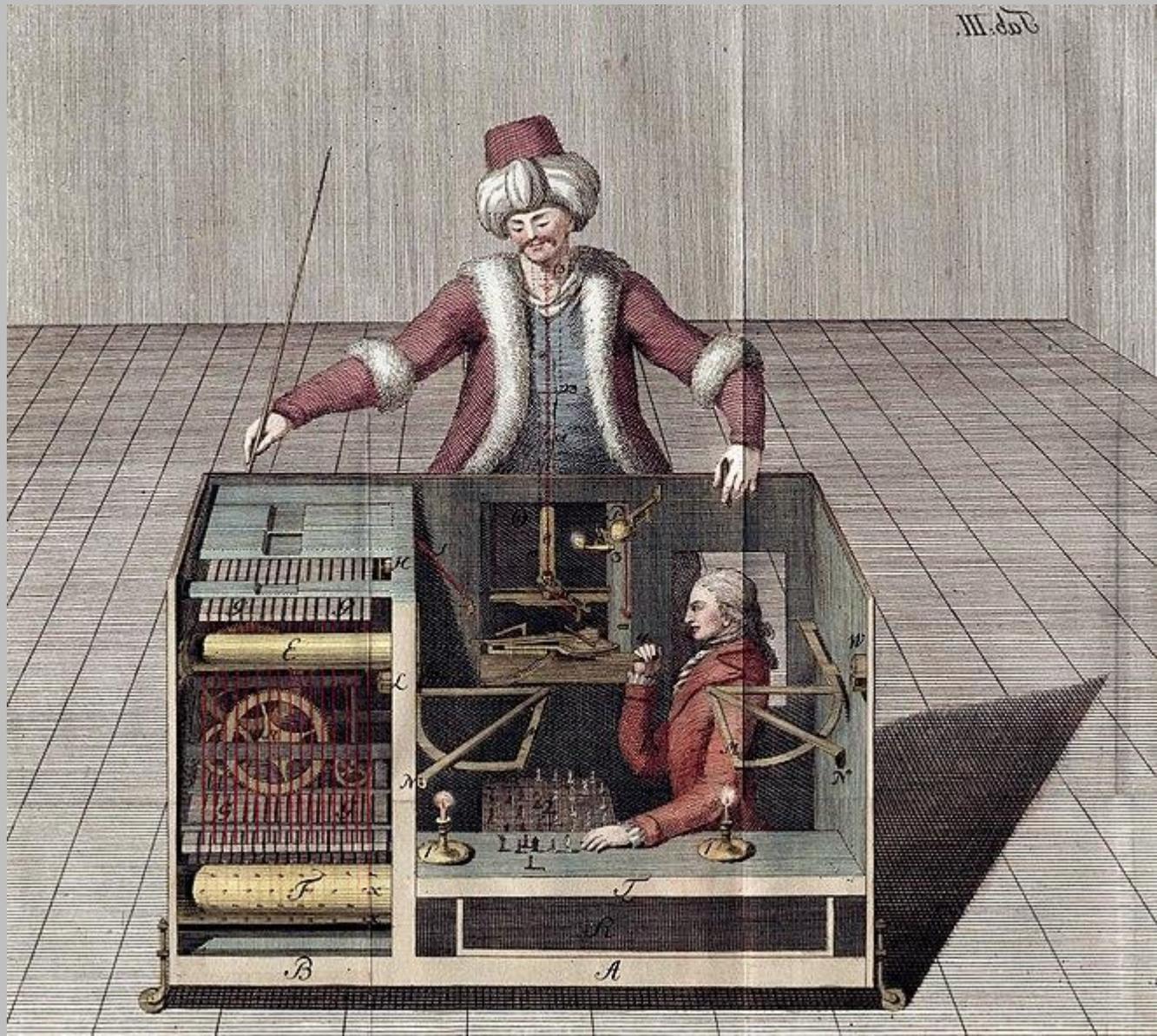


W. de Kempelen del.

Che a Mechel, exco'd. Basilea.

P. G. Piaty, sc.

Der Schach-Spieler, wie er vor dem Spiel gezeigt wird von Herrn L. Foucar & Checs, tel qu'on le montre avant le jeu, par decaant.



A PROPOSAL FOR THE
DARTMOUTH SUMMER RESEARCH PROJECT
ON ARTIFICIAL INTELLIGENCE

J. McCarthy, Dartmouth College
M. L. Minsky, Harvard University
N. Rochester, I. B. M. Corporation
C. E. Shannon, Bell Telephone Laboratories

August 31, 1955

A Proposal for the
DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

August 31, 1955

A PROPOSAL FOR THE

1) Automatic Computers

If a machine can do a job, then an automatic calculator can be programmed to simulate the machine. The speeds and memory capacities of present computers may be insufficient to simulate many of the higher functions of the human brain, but the major obstacle is not lack of machine capacity, but our inability to write programs taking full advantage of what we have.

August 31, 1955

2) How Can a Computer be Programmed to Use a Language

It may be speculated that a large part of human thought consists of manipulating words according to rules of reasoning and rules of conjecture. From this point of view, forming a generalization consists of admitting a new word and some rules whereby sentences containing it imply and are implied by others. This idea has never been very precisely formulated nor have examples been worked out.

I could also
write a lot
this.

3. Neuron Nets

How can a set of (hypothetical) neurons be arranged so as to form concepts. Considerable theoretical and experimental work has been done on this problem by Uttley, Rashevsky and his group, Farley and Clark, Pitts and McCulloch, Minsky, Rochester and Holland, and others. Partial results have been obtained but the problem needs more theoretical work.

5) Self-Improvement

Probably a truly intelligent machine will carry out activities which may best be described as self-improvement. Some schemes for doing this have been proposed and are worth further study. It seems likely that this question can be studied abstractly as well.

7) Randomness and Creativity

Randomness is an artifact appear
in the H.C. method of solving

A fairly attractive and yet clearly incomplete conjecture
is that the difference between creative thinking and un-
imaginative competent thinking lies in the injection of a
some randomness. The randomness must be guided by
intuition to be efficient. In other words, the educated
guess on the hunch include controlled randomness in
otherwise orderly thinking.

prob
the H.C.
prob.
that is
"H.C."
method
is a

Mechanical Translation

New Challenge to Communication

Jacob Ornstein

One of the thorniest problems in present-day communication is the translation of important writings, especially in the fields of science and technology, from one language to another. The machine age has done little to alter the translation process, which today, as it was centuries ago, must be laboriously performed at considerable expense by human beings. Moreover, growing concern over the dangerous lag between the appearance of important works in other languages and their translation into English and other tongues has within the past decade focused serious attention on the problem of mechanical translation. Until recently, however, mechanical translation has remained one of those human dreams the realization of which was relegated to some point in the unpredictable future.

It was with the development of the electronic computer that real hope came to be felt concerning the possibility of mechanical translation. The man who first envisioned the possibility of translation by electronic means was apparently Warren Weaver, director of the Natural Sciences Division of the Rockefeller Foundation. First of all in 1945, and more concretely in a memorandum dated 15 July 1949, Weaver raised the question of the feasibility of designing a computer-like machine capable of translating from one language to another. This memorandum was circulated among a number of

linguists and scientists; their reactions ranged from high optimism to complete skepticism. Weaver's concepts, nevertheless, aroused considerable interest and stimulated preliminary research in the field.

The problem of mechanical translation is being approached by a growing number of scholars, although much of the research is in the realm of speculation and theory. Among mathematicians, one may mention such figures as Yehoshua Bar-Hillel, formerly of Massachusetts Institute of Technology and now of Hebrew University, whose articles on mechanical translation were accorded considerable space in the 5 April 1954 issue of *Time*, the weekly news magazine. Important research is being carried on at Massachusetts Institute of Technology by Victor H. Yngve of the Research Laboratory of Electronics and William N. Locke, professor of modern languages. Anthony Oettinger of Harvard University's Computation Laboratory has also been conducting experiments. Erwin Reifer of the University of Washington has concerned himself with the elaboration of a mechanical dictionary. An active group, headed by William Bull and Victor Oswald, is working on problems of mechanical translation at the University of California at Los Angeles.

The first successful experiment in performing mechanical translation was the

result of a joint project undertaken by Georgetown University's Institute of Languages and Linguistics, Washington, D.C., and the International Business Machines Corporation. Considerable publicity followed their demonstration at I.B.M. headquarters in New York on 7 January 1954 of the translation of more than 60 sentences from Russian into English. An invited group of government officials, linguists, and scientists watched a typist who knew no Russian type the sentences, which had been transliterated into the Roman alphabet, on an electric card punch and feed them into the "electronic translator," which produced accurate English translations. The sentences were from the workaday fields of science, technology, communications, and international affairs. The following are a few examples of the transliterated Russian sentences and the English equivalents:

Myezhdunarodnoye ponyimaniye vyayetsya vazhnim faktorom v ryeshenyiy polytyichyestykh voprosov.

International understanding constitutes an important factor in decision of political questions.

Dorogi stroyatsya yiz byetona.

Roads are constructed from concrete.

Komandyr poluchayet svyedyeniya po tyelyegrafu.

A commander gets information over a telegraph.

Vyelyichyina ugla opryedyelyetsya otnoshenyem dlyiny dugi k radiusu.

Magnitude of angle is determined by the relation of length of arc to radius.

Obrabotka povyshayet kachestvo nyeftyi.

Processing improves the quality of crude oil.

It is revealing to consider in some detail the background of the Georgetown-

The author is on the staff of the U.S. Department of Agriculture Graduate School.

Welcome to

```
EEEEEE LL      IIII ZZZZZZZZ  AAAAA
EE      LL      II      ZZ      AA  AA
EEEEEE LL      II      ZZZ      AAAAAA
EE      LL      II      ZZ      AA  AA
EEEEEE LLLLLL IIII ZZZZZZZZ  AA  AA
```

Eliza is a mock Rogerian psychotherapist.

The original program was described by Joseph Weizenbaum in 1966.

This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?

YOU: Men are all alike.

ELIZA: What is the connection, do you suppose ?

YOU: They're always bugging us about something or other.

ELIZA: Can you think of a specific example ?

YOU: Well, my boyfriend made me come here.

ELIZA: Is it important to you that your boyfriend made you come here ?

YOU: He says I'm depressed much of the time.

ELIZA: I am sorry to hear that you are depressed.

YOU: It's true. I am unhappy.

ELIZA: Can you explain what made you unhappy ?

YOU: █

Pandora's box?

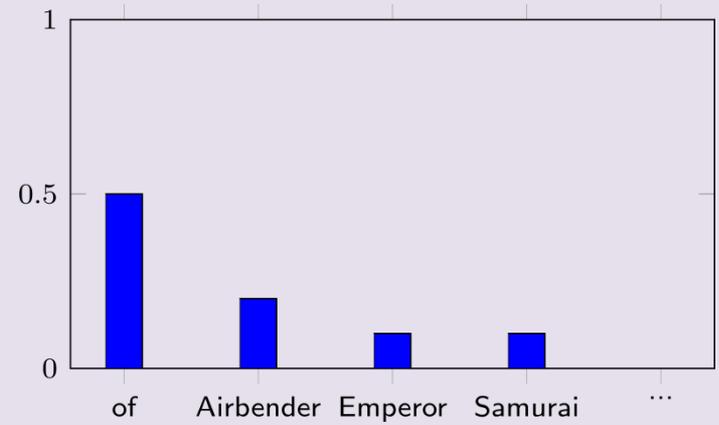


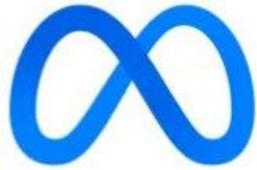


 COMMON
CRAWL



My favourite movie is The Last _____



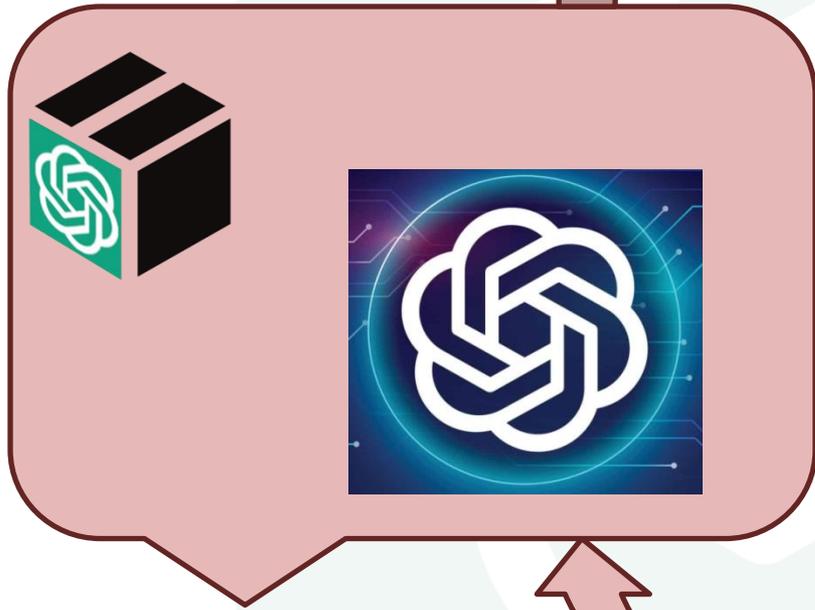


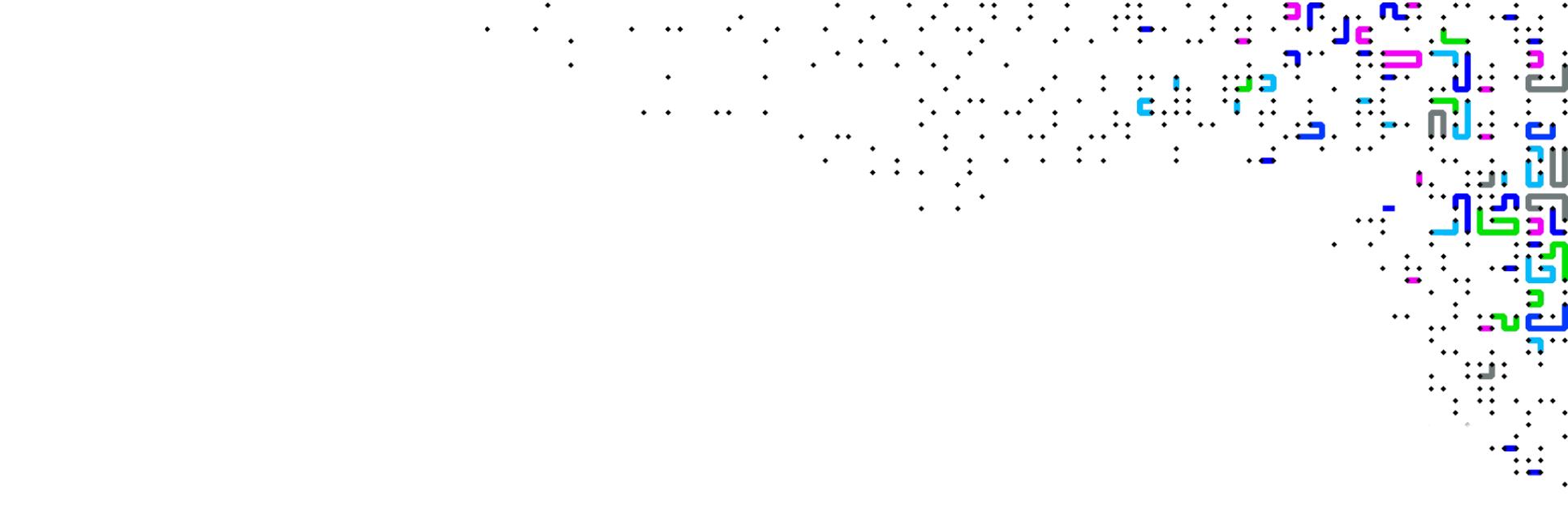
LLAMA 2



PanGu- Σ







The KG is dead, long live the LLM?



Language Models sounds the Death Knell of Knowledge Graphs

Kunal Suri
Optum, India
kunal_suri@optum.com

Atul Singh
Optum, India
atul_singh18@optum.com

Prakhar Mishra
Optum, India
prakhar_mishra29@optum.com

Swapna Sourav Rout
Optum, India
rout.swapnasourav@optum.com

Rajesh Sabapathy
Optum, India
rajesh_sabapathy@uhc.com

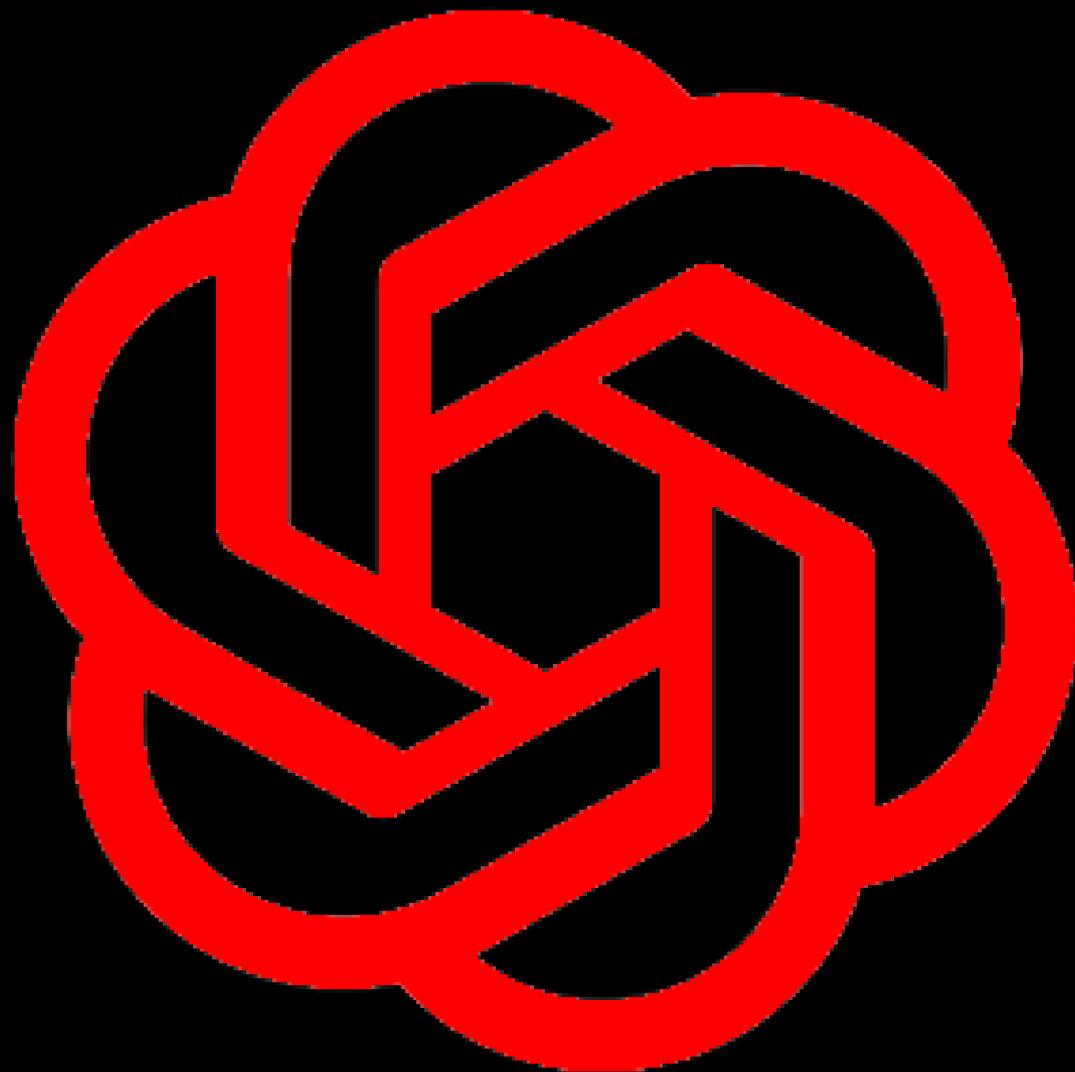
Abstract—Healthcare domain generates a lot of unstructured and semi-structured text. Natural Language processing (NLP) has been used extensively to process this data. Deep Learning based NLP especially Large Language Models (LLMs) such as BERT have found broad acceptance and are used extensively for many applications. A Language Model is a probability distribution over a word sequence. Self-supervised Learning on a large corpus of data automatically generates deep learning-based language models. BioBERT and Med-BERT are language models pre-trained for the healthcare domain. Healthcare uses typical NLP tasks such as question answering, information extraction, named entity recognition, and search to simplify and improve processes. However, to ensure robust application of the results, NLP practitioners need to normalize and standardize them. One of the main ways of achieving normalization and standardization is the use of Knowledge Graphs. A Knowledge Graph captures concepts and their relationships for a specific domain, but their creation is time-consuming and requires manual intervention from domain experts, which can prove expensive. SNOMED CT (Systematized Nomenclature of Medicine - Clinical Terms), Unified Medical Language System (UMLS), and Gene Ontology (GO) are popular ontologies from the healthcare domain. SNOMED CT and UMLS capture concepts such as disease,

relationships among different entities makes knowledge graphs redundant for many applications.

This paper uses similar terms from SNOMED-CT KG and passes them through a language model for the healthcare domain BioRedditBERT to get a 768-dimensional dense vector representation. The paper presents the results for analyzing these embeddings. The experiments presented in the paper validate that similar terms cluster together. The paper uses simple heuristics to assign names to clusters. The results show that the cluster names match the names in the KG. Finally, the experiments demonstrate that the cosine similarity of vector representation of similar terms is high and vice versa.

Our contributions include: (i) We propose a study to demonstrate the value and application of Large Language Models (LLMs) in comparison to Knowledge Graph-based approaches for the task of synonym extraction. (ii) We extensively evaluate our approach on a standard, widely accepted dataset, and the results are encouraging.

Will design
ontologies
for food



A



Which Turing Award winners were born in Latin America?

As of my last update in September 2021, Adolfo Guzmán-Arenas is a Turing Award winner born in Latin America. He was born in Guatemala. He received the Turing Award in 2002 along with his colleague Manuel Blum for their work on computational complexity theory and the development of the concept of a "proof system."

Please note that my information might not be current, so I recommend verifying with the latest sources to confirm the most up-to-date information about Turing Award winners from Latin America.

A



Which Turing Award winners were born in Latin America?

As of my last update in September 2021, there are no Turing Award winners who were born in Latin America. The Turing Award is often considered the “Nobel Prize of Computer Science” and is awarded annually by the Association for Computing Machinery (ACM). The recipients have primarily been from the United States, Europe, and Asia, reflecting in part the historical investment and focus on computer science research in these regions.

However, it’s important to note that the field of computer science is increasingly global, and outstanding contributions are being made by researchers from all around the world, including Latin America. While there have not been Turing Award winners born in Latin America as of my last update, this could change in the future as the field continues to evolve and diversify.

Please verify this information for the most current data, as it might have changed after my last training data in September 2021.

A

Can you give me a SPARQL query for Wikidata to find Turing Award winners who were born in Latin America?



Certainly!

```
# Turing Award winners born in Latin America
SELECT ?person ?personLabel ?birthPlace ?birthPlaceLabel ?award ?awardLabel WHERE {
  ?person p:P166 ?statement.
  ?statement ps:P166 wd:Q7251; pq:P582 ?awardDate. # Turing Award
  ?person wdt:P19 ?birthPlace. # Place of birth
  ?birthPlace wdt:P30 wd:Q12521. # Latin America
  SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE]". }
}
ORDER BY ?awardDate
```

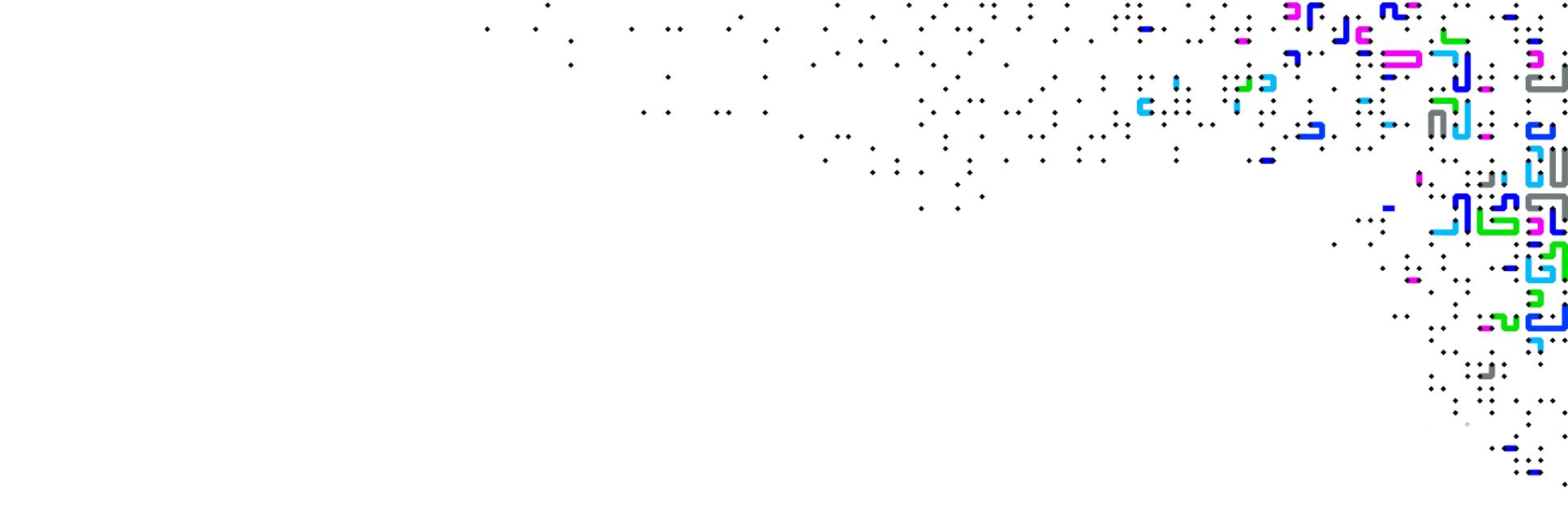
A

Can you give me a SPARQL query for Wikidata to find Turing Award winners who were born in Latin America?

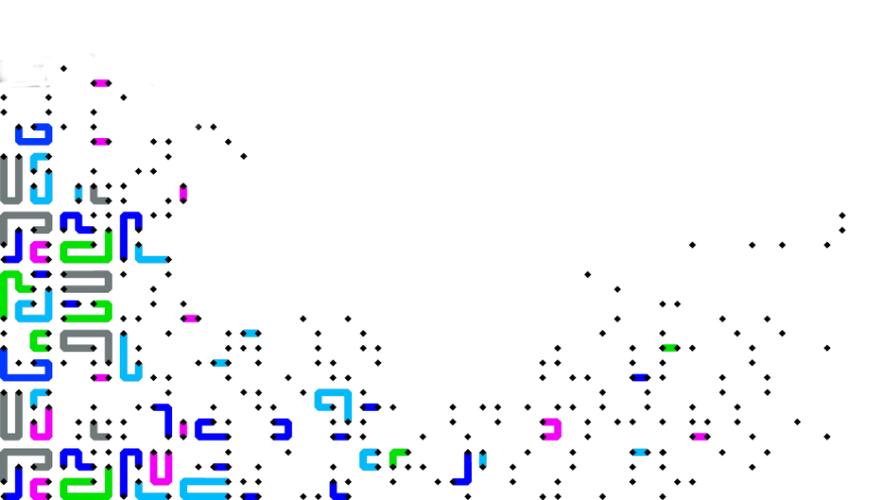


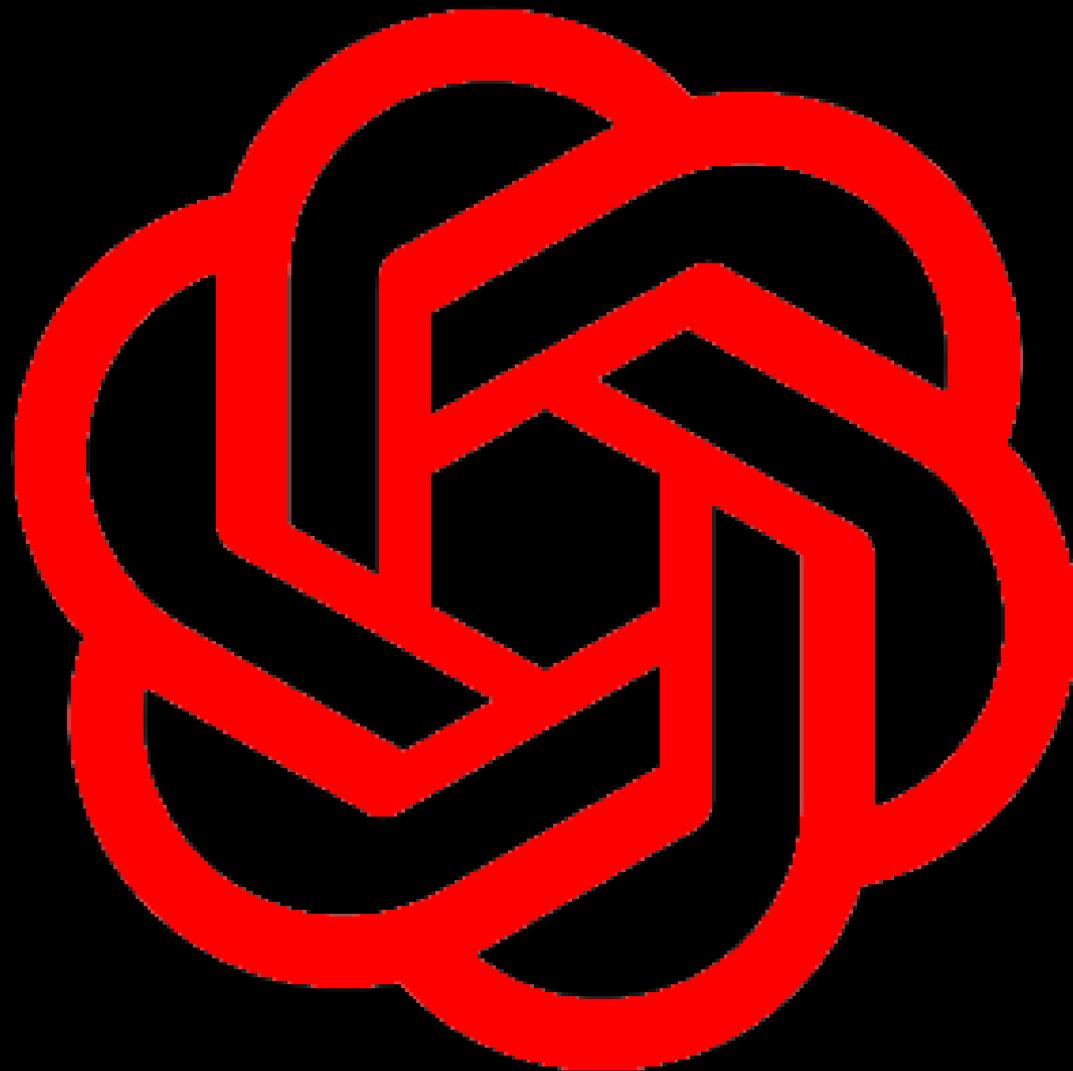
Certainly!

```
# Turing Award winners born in Latin America
SELECT ?person ?personLabel ?birthPlace ?birthPlaceLabel ?award ?awardLabel WHERE {
  ?person p:P166 ?statement.
  ?statement ps:P166 wd:Q7251; pq:P582 ?awardDate. # Turing Award
  ?person wdt:P19 ?birthPlace. # Place of birth
  ?birthPlace wdt:P30 wd:Q12521. # Latin America
  SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE]". }
}
ORDER BY ?awardDate
```



LLMs + KGs: A dream team?







LLM

KG

Hallucinations



Opaqueness:



Intransigence:



Staleness:



Bias:



Costs:



Short tail:



Sanitised:



Non-determinism:



Usability:



Factual only:



LLM KG KG + LLM

Hallucinations



Opaqueness:



Intransigence:



Staleness:



Bias:



Costs:



Short tail:



Sanitised:



Non-determinism:

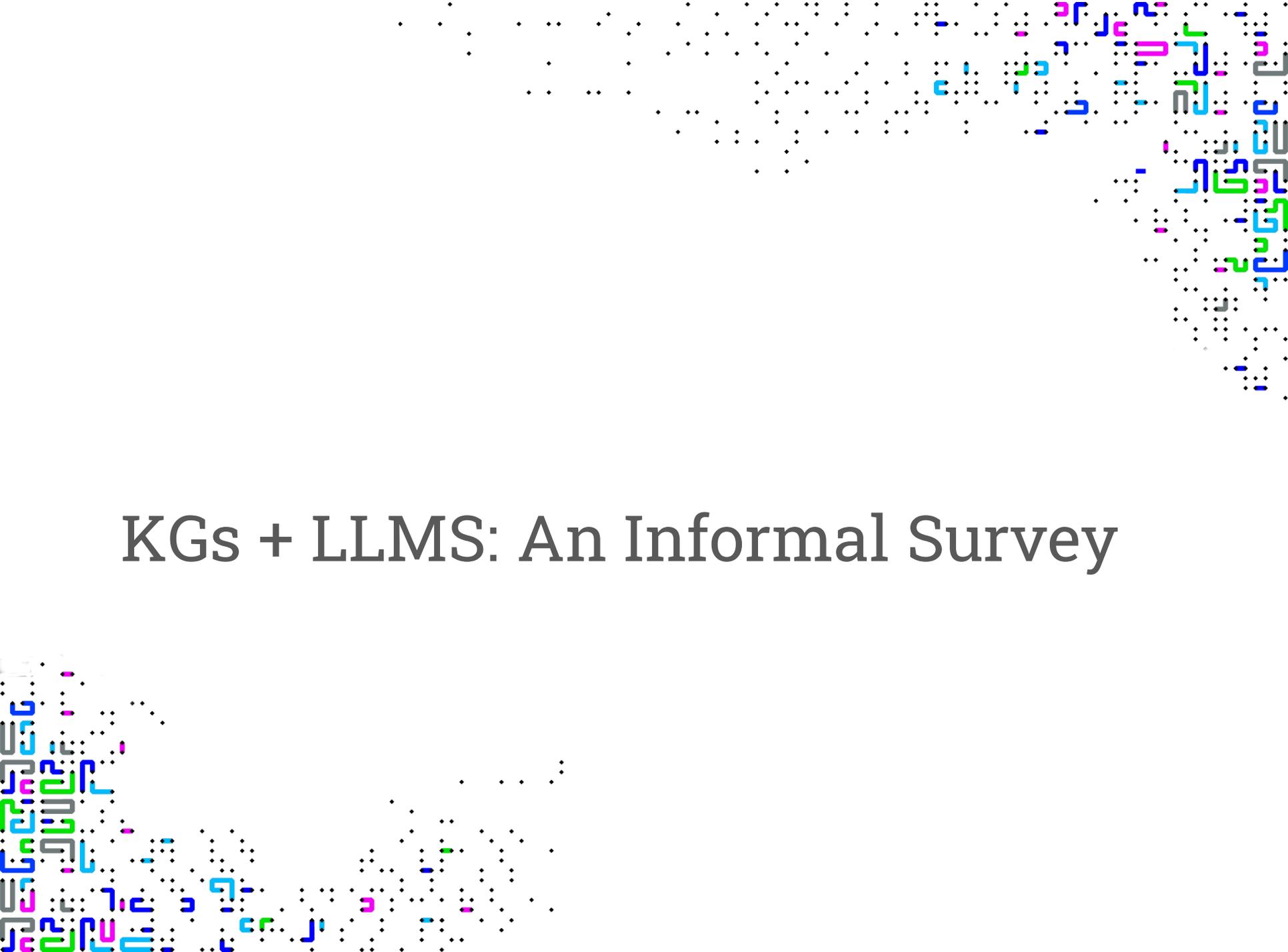


Usability:



Factual only:





KGs + LLMS: An Informal Survey

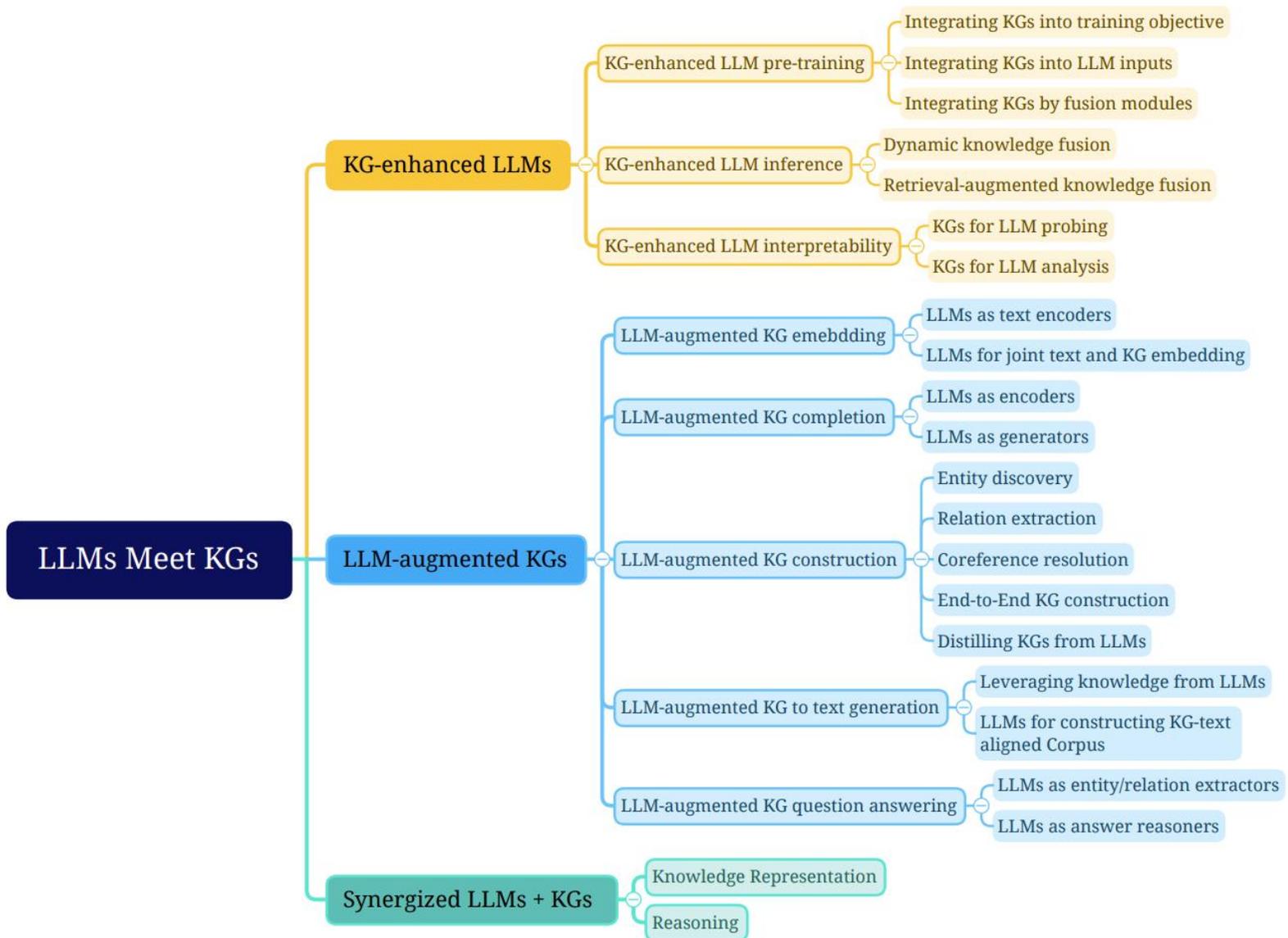
Google Scholar

(llm OR "language model" OR gpt OR chatgpt) AND "knowledge graph"

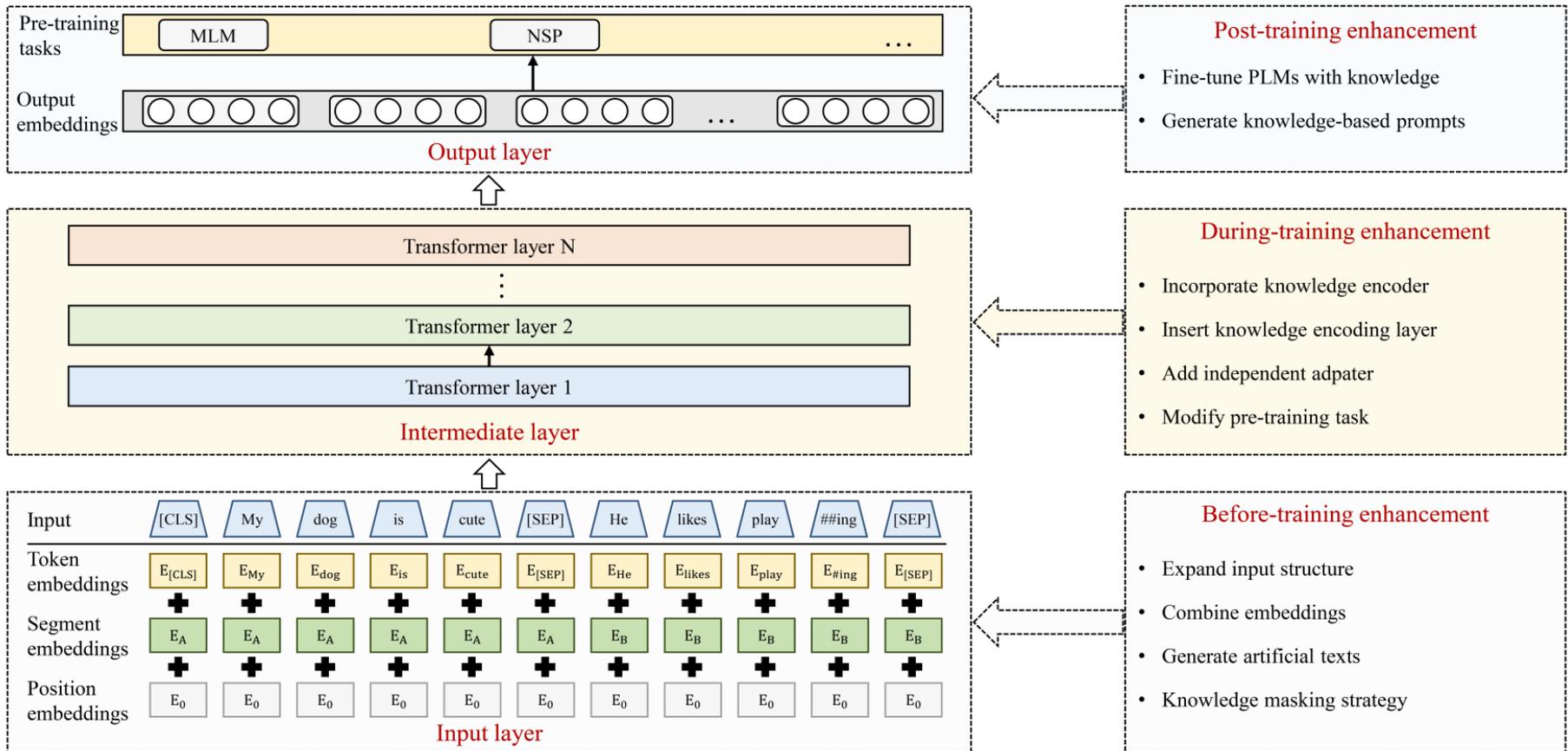


Articles Case law

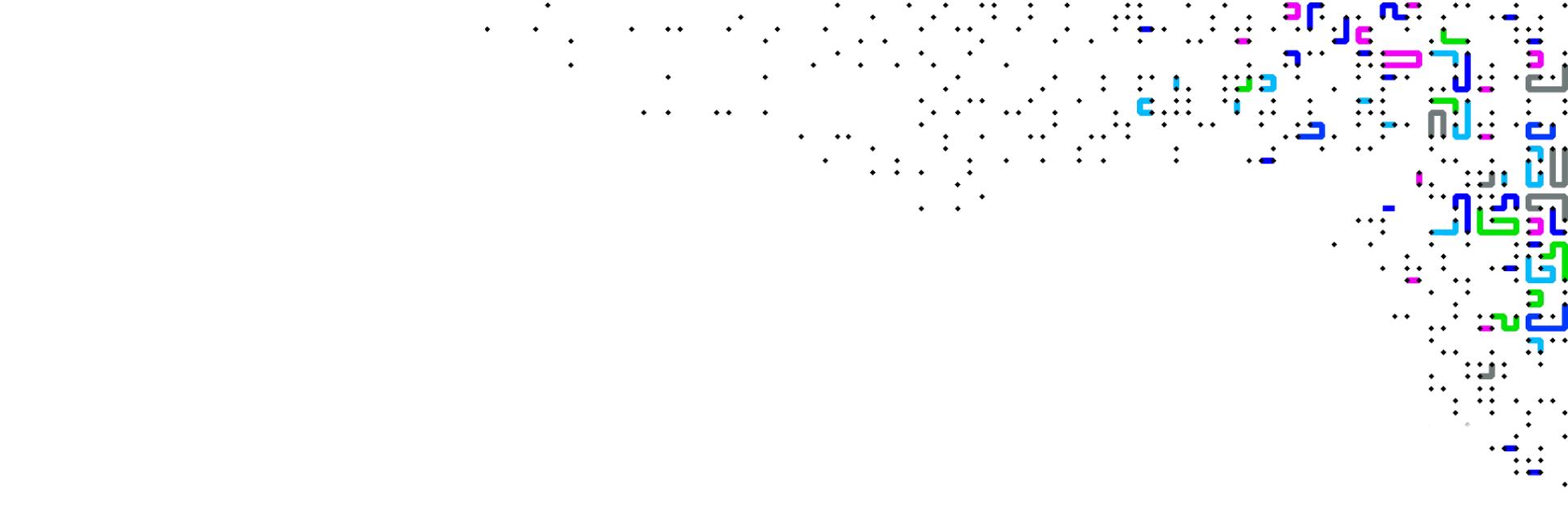
- Training models that combine KGs and (L)LMs
[Logan+19, Fei+20, Verga+21, Hu+22, Yasunaga+22, Yu+22, Colon+21*]
- Decoding (L)LM “knowledge” as KGs
[Hao+23, Wang+20*]
- Verbalising KGs using (L)LMs
[Agarwal+21, Li+21, Ahmed+23, AxelssonS23*]
- Hybrid KG+LLM Tasks
 - Query/Question Answering
[Yasunaga+21, ChoudharyR23*, JSun+23*, Lehmann+23* Omar+23*]
- LLM-enriched KG Tasks
 - KG Completion:
[Kin+20, Yao+23*]
 - KG Construction
[Fei+20, Cenikj+23, Trajanoska+23*, Zhu+23*]
 - KG Summarization
[Vassiliou+23]
 - Multi-task
[Meyer+23*]
- Surveys, Discussions, Experiments
[YangaT20, Pan+23, JSun+23*, KSun+23*, LiuW23*, Mruthyunjaya+23*, Pan+23*, Yang+23*]



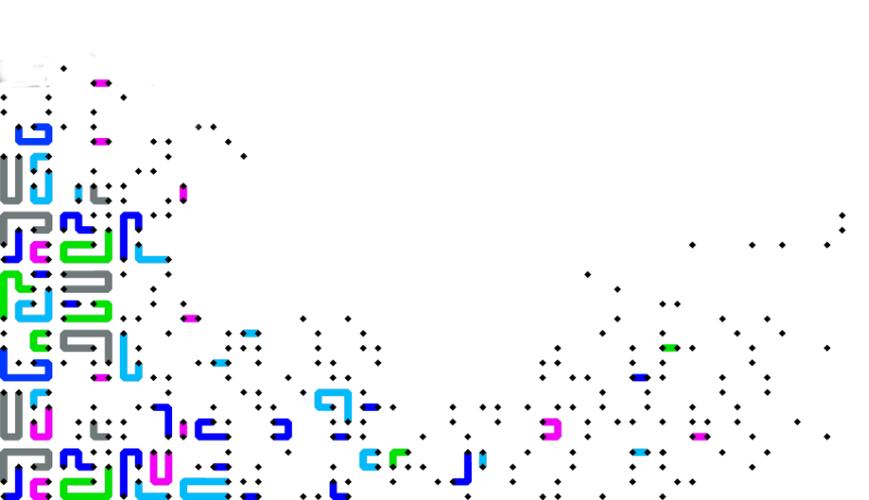
Pan+23* Shirui Pan, Linhao Luo, Yufei Wang, Chen Chen, Jiapu Wang, Xindong Wu: Unifying Large Language Models and Knowledge Graphs: A Roadmap. CoRR abs/2306.08302 (2023).



Yang+23* Linyao Yang, Hongyang Chen, Zhao Li, Xiao Ding, Xindong Wu: ChatGPT is not Enough: Enhancing Large Language Models with Knowledge Graphs for Fact-aware Language Modeling. CoRR abs/2306.11489 (2023).



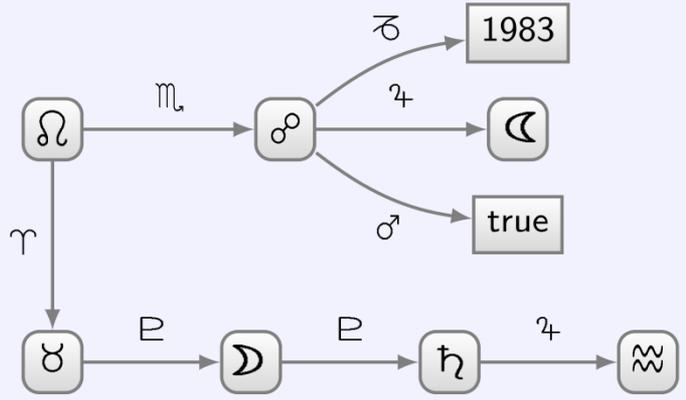
LLMs can make KGs more usable



Niche Problem (1)	6	26	33	35	13	$\bar{x} \approx 3.204, p \approx 0.047$
Redundant w/ML (2)	3	14	23	31	42	$\bar{x} \approx 3.841, p \approx 0.000$
Unreliable Publishers (3)	20	42	25	20	6	$\bar{x} \approx 2.558, p \approx 0.000$
Ontological Disagreement (4)	14	30	22	33	14	$\bar{x} \approx 3.027, p \approx 0.821$
Publishing Costly (5)	14	21	20	32	26	$\bar{x} \approx 3.310, p \approx 0.014$
Lacks Incentives (6)	25	49	19	14	6	$\bar{x} \approx 2.354, p \approx 0.000$
Verbose Standards (7)	22	22	21	27	19	$\bar{x} \approx 2.991, p \approx 0.946$
Won't Scale (8)	13	12	38	29	21	$\bar{x} \approx 3.292, p \approx 0.011$
Lacks Usable Tools (9)	32	51	14	11	4	$\bar{x} \approx 2.143, p \approx 0.000$
Decentralisation Costly (10)	11	17	23	25	36	$\bar{x} \approx 3.518, p \approx 0.000$
	1	2	3	4	5	



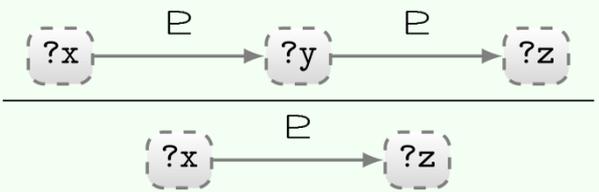
DATA



QUERY



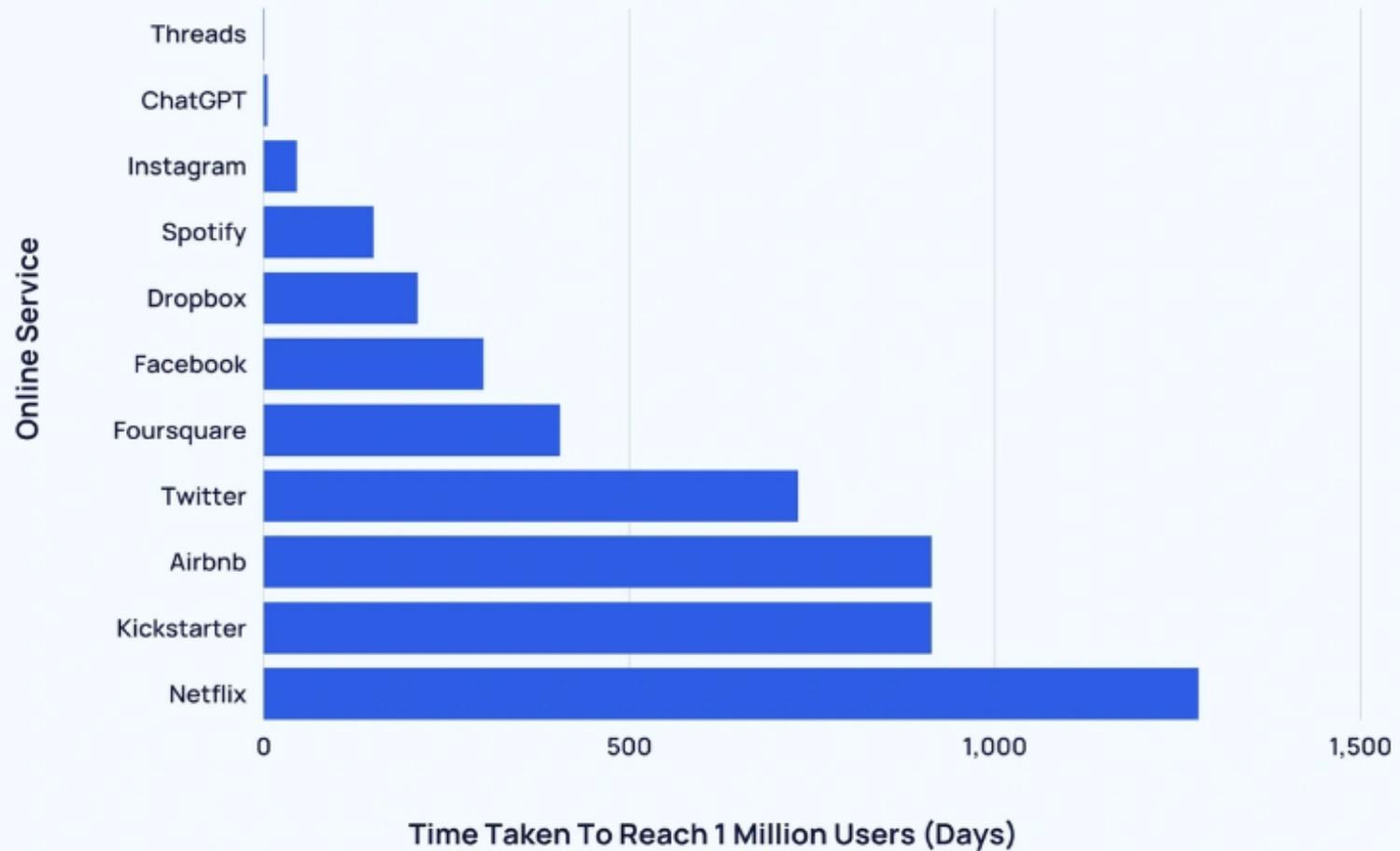
LOGIC



LANGUAGE



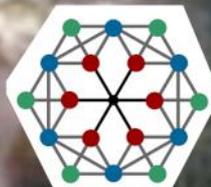
Time taken to reach 1 million users



Conclusions

- LLMs do solve some tasks that KGs do
- LLMs do not replace KGs
- LLMs like likely never replace KGs
- LLMs + KGs = An interesting research topic
- LLMs can make KGs more usable and bring them to the masses
(LLMs will likely take the credit, but that's okay)

I'M NOT DEAD YET



THINK I'LL GO FOR A WALK

References

- Agarwal+21:** Oshin Agarwal, Heming Ge, Siamak Shakeri, Rami Al-Rfou: Knowledge Graph Based Synthetic Corpus Generation for Knowledge-Enhanced Language Model Pre-training. NAACL-HLT 2021: 3554-3565
- Ahmed+23** Abdullah Fathi Ahmed, Asep Fajar Firmansyah, Mohamed Ahmed Sherif, Diego Moussallem, and Axel-Cyrille Ngonga Ngomo Explainable Integration of Knowledge Graphs using Large Language Models. NLDB 2023: 124-139
- AxelssonS2023*** Agnes Axelsson, Gabriel Skantze Using Large Language Models for Zero-Shot Natural Language Generation from Knowledge Graphs. CoRR abs/2307.07312 (2023).
- Cenikj+23** Gjorgjina Cenikj, Lidija Strojnik, Risto Angelski, Nives Ogrinc, Barbara Koroušić Seljak, Tome Eftimov: From language models to large-scale food and biomedical knowledge graphs. Scientific Reports, 2023.
- ChoudharyR2023*** Nurendra Choudhary, Chandan K. Reddy: Complex Logical Reasoning over Knowledge Graphs using Large Language Models. CoRR abs/2305.01157 (2023).
- Colon+21*** Pedro Colon-Hernandez, Catherine Havasi, Jason B. Alonso, Matthew Huggins, Cynthia Breazeal: Combining pre-trained language models and structured knowledge. CoRR abs/2101.12294 (2021).
- Fei+20** Hao Fei, Yafeng Ren, Yue Zhang, Donghong Ji and Xiaohui Liang: Enriching contextualized language model from knowledge graph for biomedical information extraction. Briefings Bioinform. 22(3) (2021)
- Hao+23** Shibo Hao, Bowen Tan, Kaiwen Tang, Bin Ni, Xiyan Shao, Hengzhe Zhang, Eric P. Xing, Zhiting Hu: BertNet: Harvesting Knowledge Graphs with Arbitrary Relations from Pretrained Language Models. ACL (Findings) 2023: 5000-5015

- Hu+22** Ziniu Hu, Yichong Xu, Wenhao Yu, Shuohang Wang, Ziyi Yang, Chenguang Zhu, Kai-Wei Chang, Yizhou Sun Empowering Language Models with Knowledge Graph Reasoning for Question Answering. EMNLP 2022: 9562-9581
- JSun+23*** Jiashuo Sun, Chengjin Xu, Lumingyuan Tang, Saizhuo Wang, Chen Lin, Yeyun Gong, Heung-Yeung Shum, Jian Guo Think-on-Graph: Deep and Responsible Reasoning of Large Language Model with Knowledge Graph. CoRR abs/2307.07697 (2023).
- Kin+20** Bosung Kim, Taesuk Hong, Youngjoong Ko, Jungyun Seo: Multi-Task Learning for Knowledge Graph Completion with Pre-trained Language Models. COLING 2020: 1737-1743
- KSun+23*** Kai Sun, Yifan Ethan Xu, Hanwen Zha, Yue Liu, Xin Luna Dong: Head-to-Tail: How Knowledgeable are Large Language Models (LLM)? A.K.A. Will LLMs Replace Knowledge Graphs? CoRR abs/2308.10168 (2023).
- Lehmann+23*** Jens Lehmann, Preetam Gattogi, Dhananjay Bhandiwad, Sébastien Ferré, Sahar Vahdati: Language Models as Controlled Natural Language Semantic Parsers for Knowledge Graph Question Answering. Amazon TR, 2023.
- Li+21** Junyi Li, Tianyi Tang, Wayne Xin Zhao, Zhicheng Wei, Nicholas Jing Yuan, Ji-Rong Wen: Few-shot Knowledge Graph-to-Text Generation with Pretrained Language Models. ACL/IJCNLP (Findings) 2021: 1558-1568
- LiuW2023*** Chang Liu, Bo Wu: Evaluating large language models on graphs: performance insights and comparative analysis. CoRR abs/2308.11224 (2023).
- Logan+19** Robert L. Logan IV, Nelson F. Liu, Matthew E. Peters, Matt Gardner, Sameer Singh: Barack's Wife Hillary: Using Knowledge Graphs for Fact-Aware Language Modeling. ACL (1) 2019: 5962-5971.

Meyer+23* Lars-Peter Meyer, Claus Stadler, Johannes Frey, Norman Radtke, Kurt Junghanns, Roy Meissner, Gordian Dziwis, Kirill Bulert, and Michael Martin LLM-assisted Knowledge Graph Engineering: Experiments with ChatGPT. CoRR abs/2307.06917 (2023)

Mruthyunjaya+23* Vishwas Mruthyunjaya, Pouya Pezeshkpour, Estevam Hruschka, Nikita Bhutani: Rethinking Language Models as Symbolic Knowledge Graph. CoRR abs/2308.13676 (2023).

Omar+23* Reham Omar, Omij Mangukiya, Panos Kalnis, Essam Mansour: ChatGPT versus Traditional Question Answering for Knowledge Graphs: Current Status and Future Directions Towards Knowledge Graph Chatbots. CoRR abs/2302.06466 (2023).

Pan+23 Jeff Z. Pan, Simon Razniewski, Jan-Christoph Kalo, Sneha Singhanian, Jiaoyan Chen, Stefan Dietze, Hajira Jabeen, Janna Omeliyanenko, Wen Zhang, Matteo Lissandrini, Russa Biswas, Gerard de Melo, Angela Bonifati, Edlira Vakaj, Mauro Dragoni, Damien Graux Large Language Models and Knowledge Graphs: Opportunities and Challenges. TGDK 2023.

Pan+23* Shirui Pan, Linhao Luo, Yufei Wang, Chen Chen, Jiapu Wang, Xindong Wu: Unifying Large Language Models and Knowledge Graphs: A Roadmap. CoRR abs/2306.08302 (2023).

Trajanoska+23* Milena Trajanoska, Riste Stojanov, Dimitar Trajanov: Enhancing Knowledge Graph Construction Using Large Language Models. CoRR abs/2305.04676 (2023).

Vassiliou+23 Giannis Vassiliou, Nikolaos Papadakis, Haridimos Kondylakis: SummaryGPT: Leveraging ChatGPT for Summarizing Knowledge Graphs, ESWC Demo 2023.

- Verga+21** Pat Verga, Haitian Sun, Livio Baldini Soares, William W. Cohen: Adaptable and Interpretable Neural Memory Over Symbolic Knowledge. NAACL-HLT 2021: 3678-3691
- Wang+20*** Chenguang Wang, Xiao Liu, Dawn Song: Language models are open knowledge graphs. CoRR abs/2010.11967 (2020).
- YangaT20** Xinran Yanga, Ilaria Tiddi Creative Storytelling with Language Models and Knowledge Graphs, CIKM (Workshops) 2020.
- Yang+23*** Linyao Yang, Hongyang Chen, Zhao Li, Xiao Ding, Xindong Wu: ChatGPT is not Enough: Enhancing Large Language Models with Knowledge Graphs for Fact-aware Language Modeling. CoRR abs/2306.11489 (2023).
- Yao+23*** Liang Yao, Jiazhen Peng, Chengsheng Mao, Yuan Luo: Exploring Large Language Models for Knowledge Graph Completion. CoRR abs/2308.13916 (2023).
- Yasunaga+21** Michihiro Yasunaga, Hongyu Ren, Antoine Bosselut, Percy Liang, Jure Leskovec: QA-GNN: Reasoning with Language Models and Knowledge Graphs for Question Answering. NAACL-HLT 2021: 535-546
- Yasunaga+22** Michihiro Yasunaga, Antoine Bosselut, Hongyu Ren, Xikun Zhang, Christopher D Manning, Percy Liang, Jure Leskovec: Deep Bidirectional Language-Knowledge Graph Pretraining, NeurIPS 2022.
- Yu+22** Donghan Yu, Chenguang Zhu, Yiming Yang, Michael Zeng: JAKET: Joint Pre-training of Knowledge Graph and Language Understanding. AAAI 2022: 11630-11638.
- Zhu+23*** Yuqi Zhu, Xiaohan Wang, Jing Chen, Shuofei Qiao, Yixin Ou, Yunzhi Yao, Shumin Deng, Huajun Chen, Ningyu Zhang: LLMs for Knowledge Graph Construction and Reasoning: Recent Capabilities and Future Opportunities. CoRR abs/2305.13168 (2023).