



STRUCTURED QUERY VERBALIZATION WITH LARGE LANGUAGE MODELS



By Patrik Valach & Louisa Siebel

What is the goal?

We need training data to train LLM's

Who does work at TUM
Campus Heilbronn ?



LLM



```
SELECT P.Name  
FROM Persons P  
WHERE P.work='TUM'
```

this is easier in reverse

Who does work at TUM
Campus Heilbronn ?




LLM



```
SELECT P.Name  
FROM Persons P  
WHERE P.work='TUM'
```



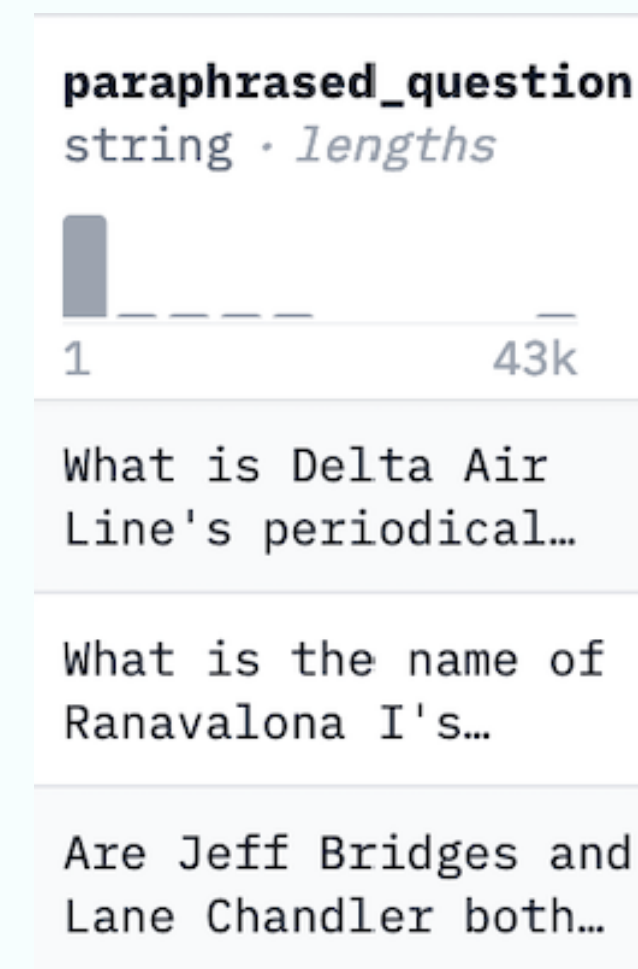
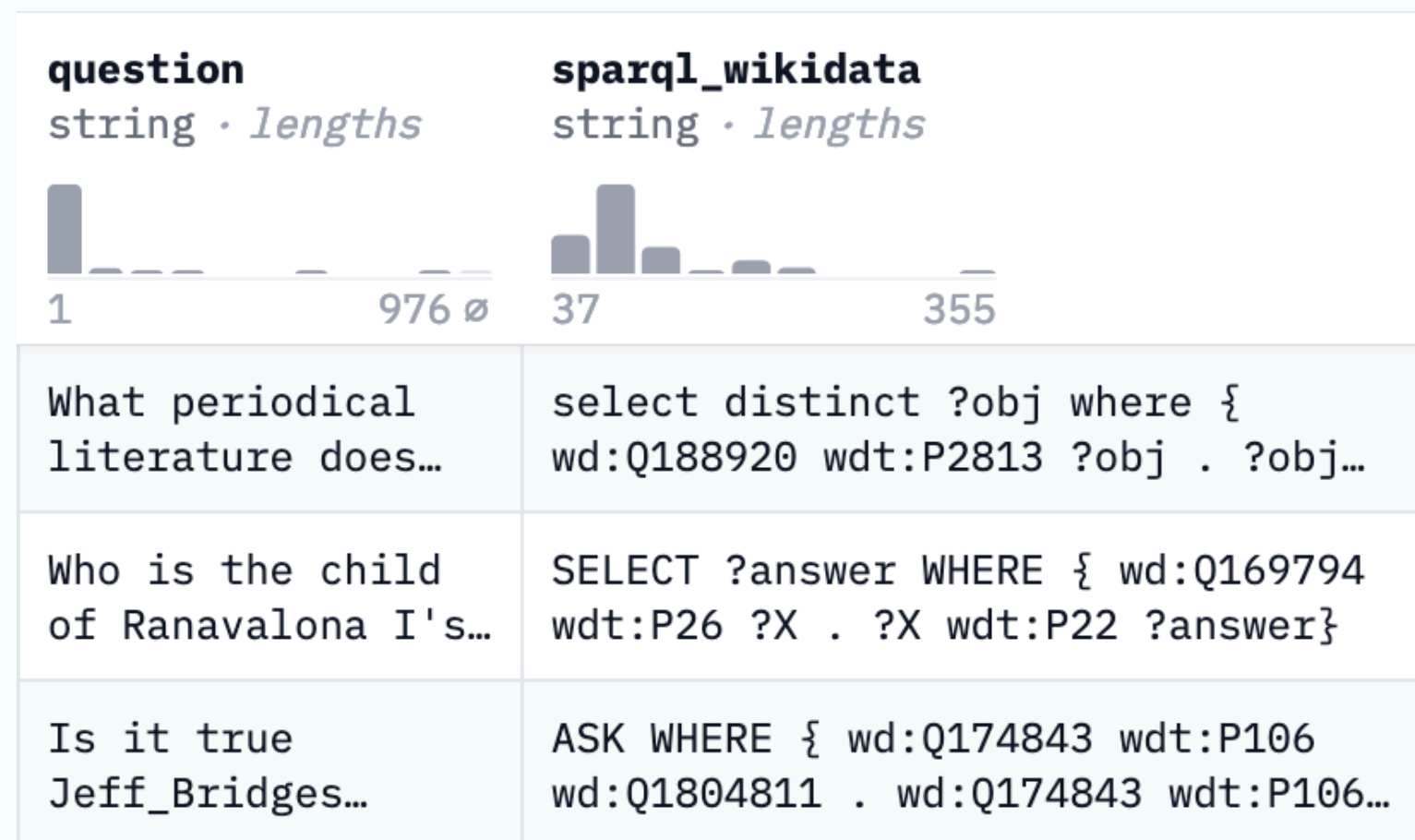
Questions to Consider

- What is the best way to create the training data?
 - What are the best Models and Methods to use?
 - How well can large language models already perform this transformation?
 - What resources do we have at our disposal?
- 

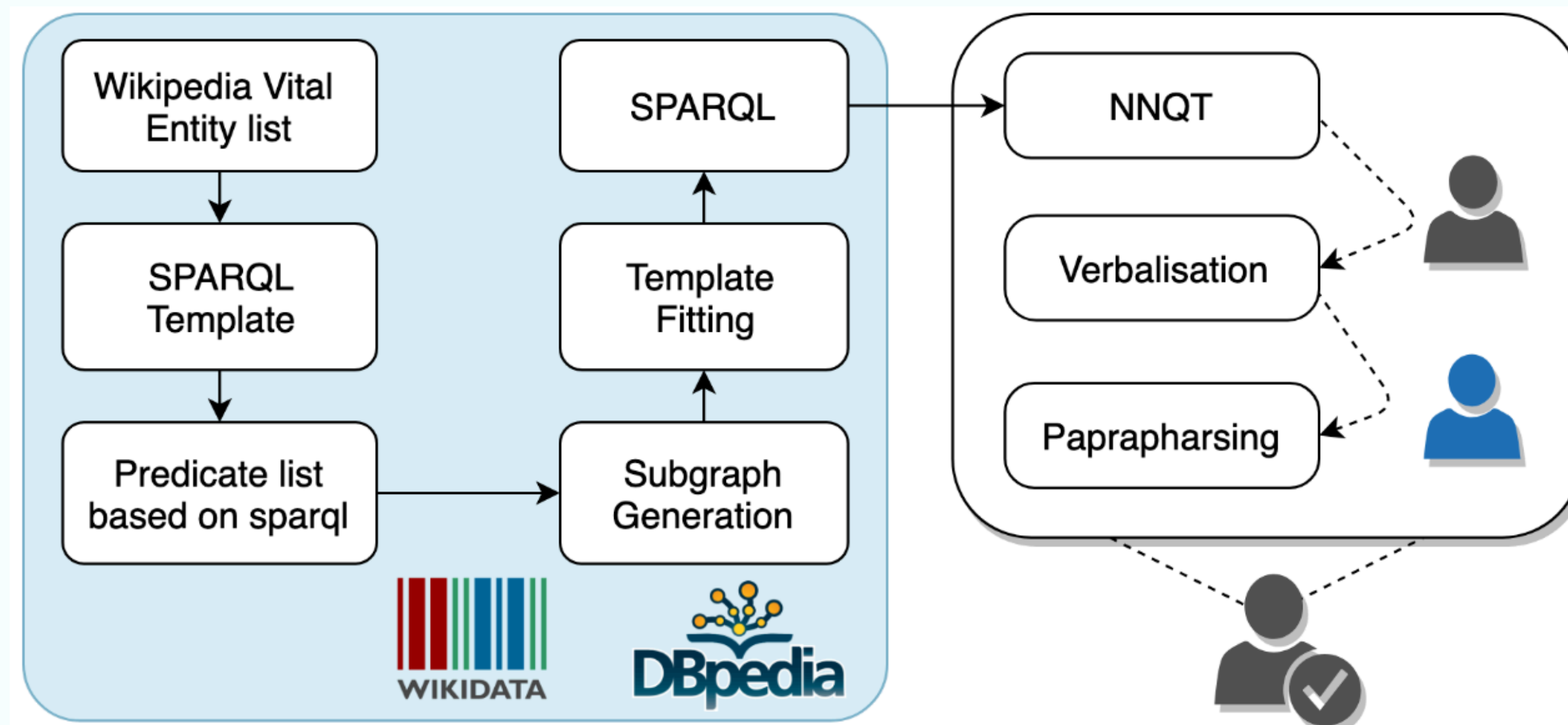
The Dataset

lc_quad

- Created to answer natural language questions
- 30 000 sets of Questions and their corresponding SPARQL queries
- target knowledge base is Wikidata and DBpedia
- We are most interested in the ‘Question’, ‘SPARQLWikidata’ and the ‘Paraphrase question’

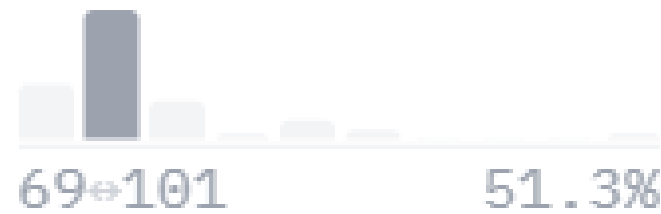


Dataset Generation and Population



Additional Considerations

sparql_wikidata
string · lengths



```
select distinct ?obj
where { wd:Q188920
wdt:P2813 ?obj . ?obj
wdt:P31 wd:Q1002697 }
```

- What is wd:Q188920
- What is wdt:P2813
- etc...

- we would need to query the knowledge graph of the dataset alongside the dataset, for this information



What are LLMs?

“a deep learning model that has been given examples to recognize and interpret human language or other types of complex data.”

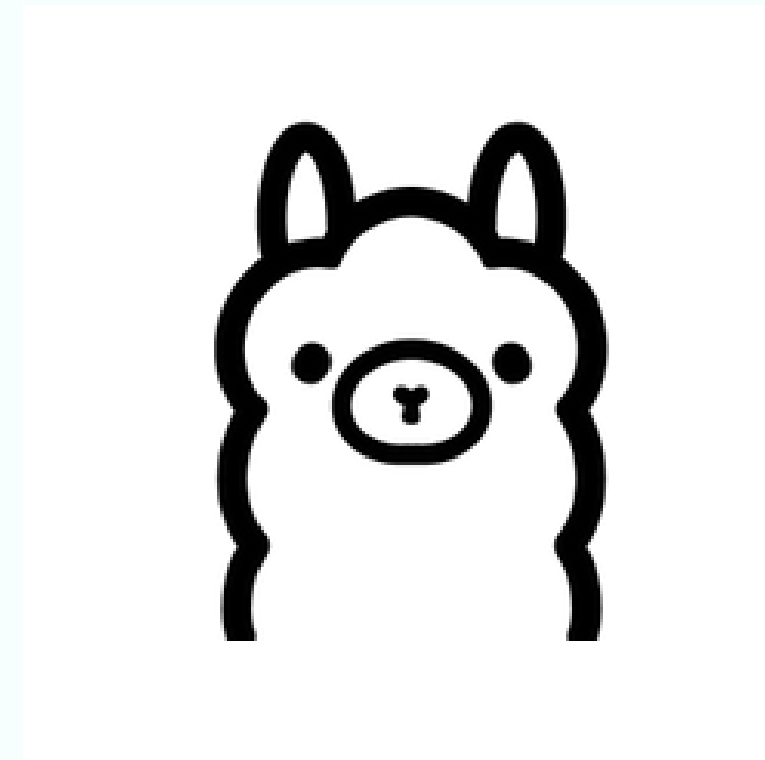
Related concepts:

- Autoregressive decoder
- Learning the probability distribution of the next token


Models to Consider

Llama3

- Developed by Meta
- **8** and 70B sizes
- Optimized for dialogue
- Text only



Potentially other ... GPT-3.5



How do we evaluate how well the model works?


Using 3 different methods:



Manual inspection



Using LLM



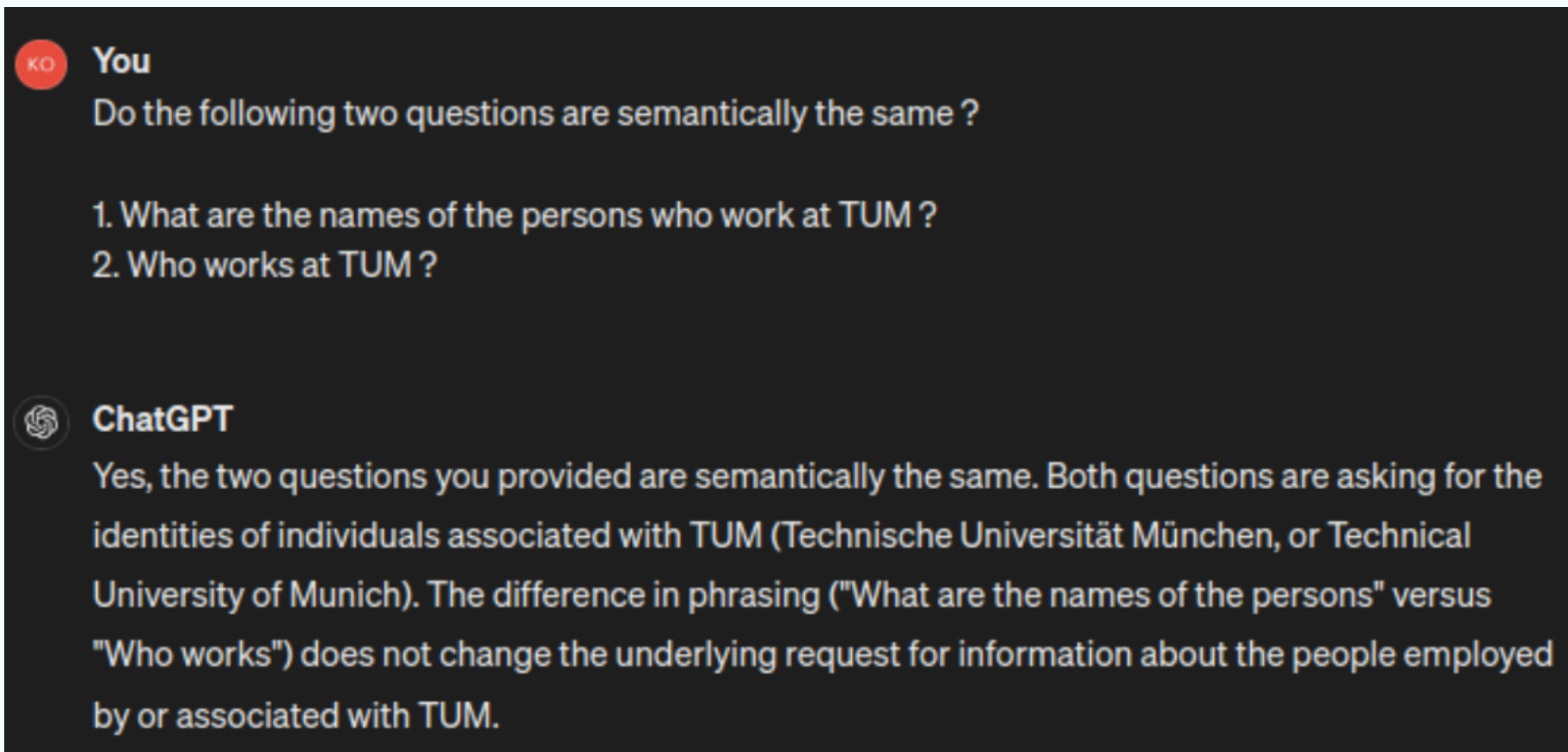
Using normal MLP
matrix

Manual inspection

- *0/1 matrix*
- *Either the SQL is converted to the NL query correctly or not*

LLM

- *Let an LLM compare the true and predicted NLP sentence*



The screenshot shows a chat interface with a dark background. On the left, there are two circular avatars: a red one with 'KO' and a blue one with the OpenAI logo. The text is in a light blue font.

You
Do the following two questions are semantically the same ?

1. What are the names of the persons who work at TUM ?
2. Who works at TUM ?

ChatGPT
Yes, the two questions you provided are semantically the same. Both questions are asking for the identities of individuals associated with TUM (Technische Universität München, or Technical University of Munich). The difference in phrasing ("What are the names of the persons" versus "Who works") does not change the underlying request for information about the people employed by or associated with TUM.

MLP matrix

BERT score

- encoder language model
- converts text to vector
- semantically indifferent text will give vectors that are close to each other

BLEU score

- how many transformations are needed to do to get the exact same text

Questions?