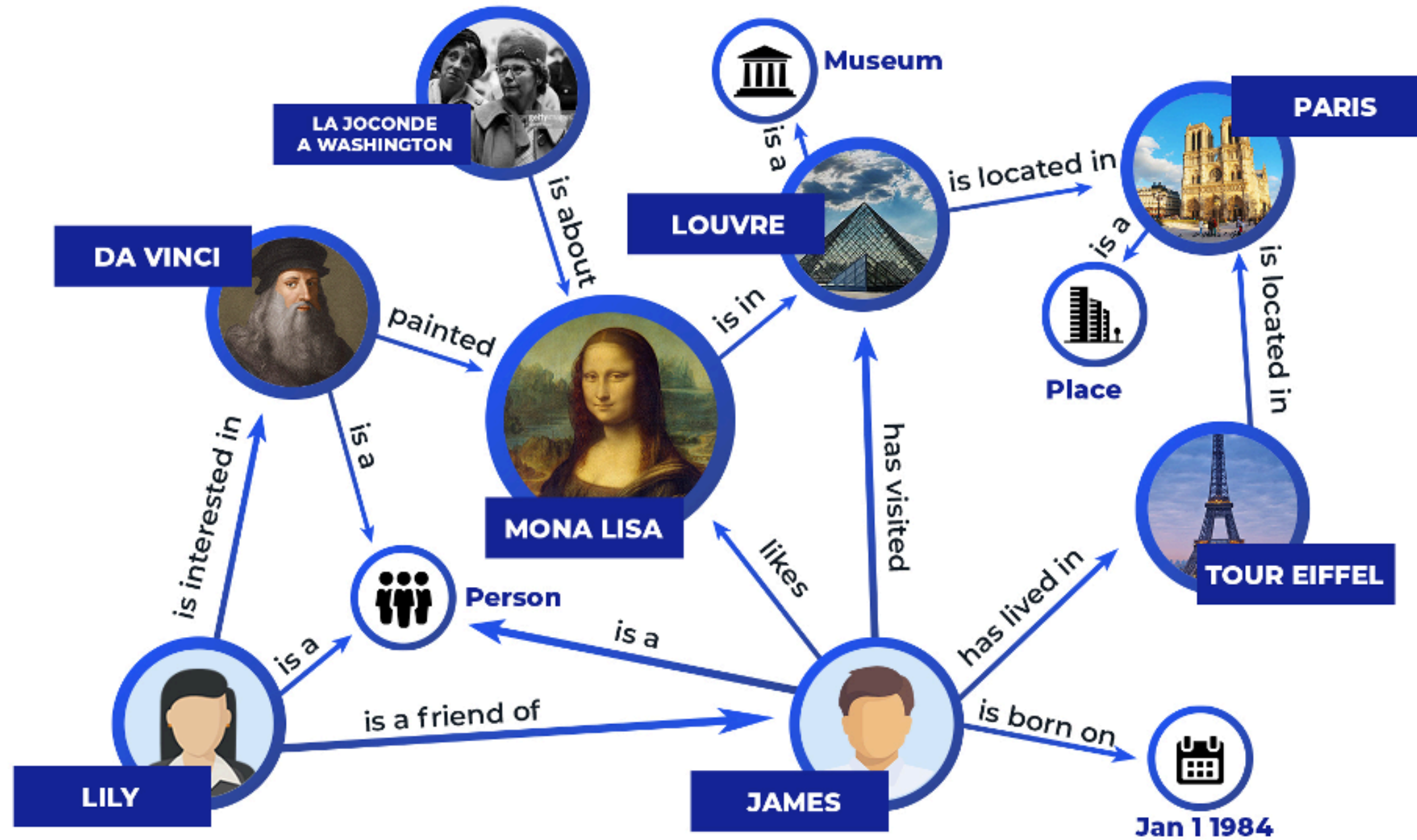# From Top-1 to Top-10: A Two-Stage Entity Alignment Framework for Knowledge Graphs

Status update

Knowledge graphs

Task: KG Entity Matching

- Given two or more knowledge graphs (KGs)

- Find node pairs, which refer to the same entity

- Selection: both nodes from different graphs

- Evaluation: gold entity pairs are given

Current solution

- Graph embedding with "Dogtags"

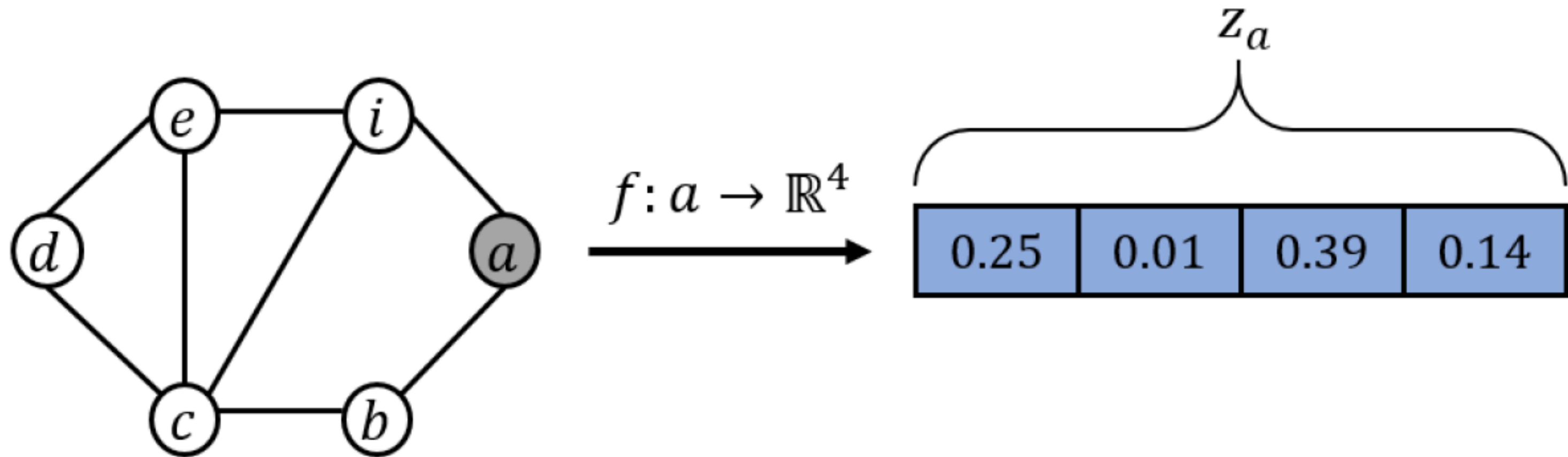- Cosine similarity: top-1

Proposed solution

- Graph embedding with "Dogtags"

- Cosine similarity: top-10

- Use LLMs to pick an entity from top-10
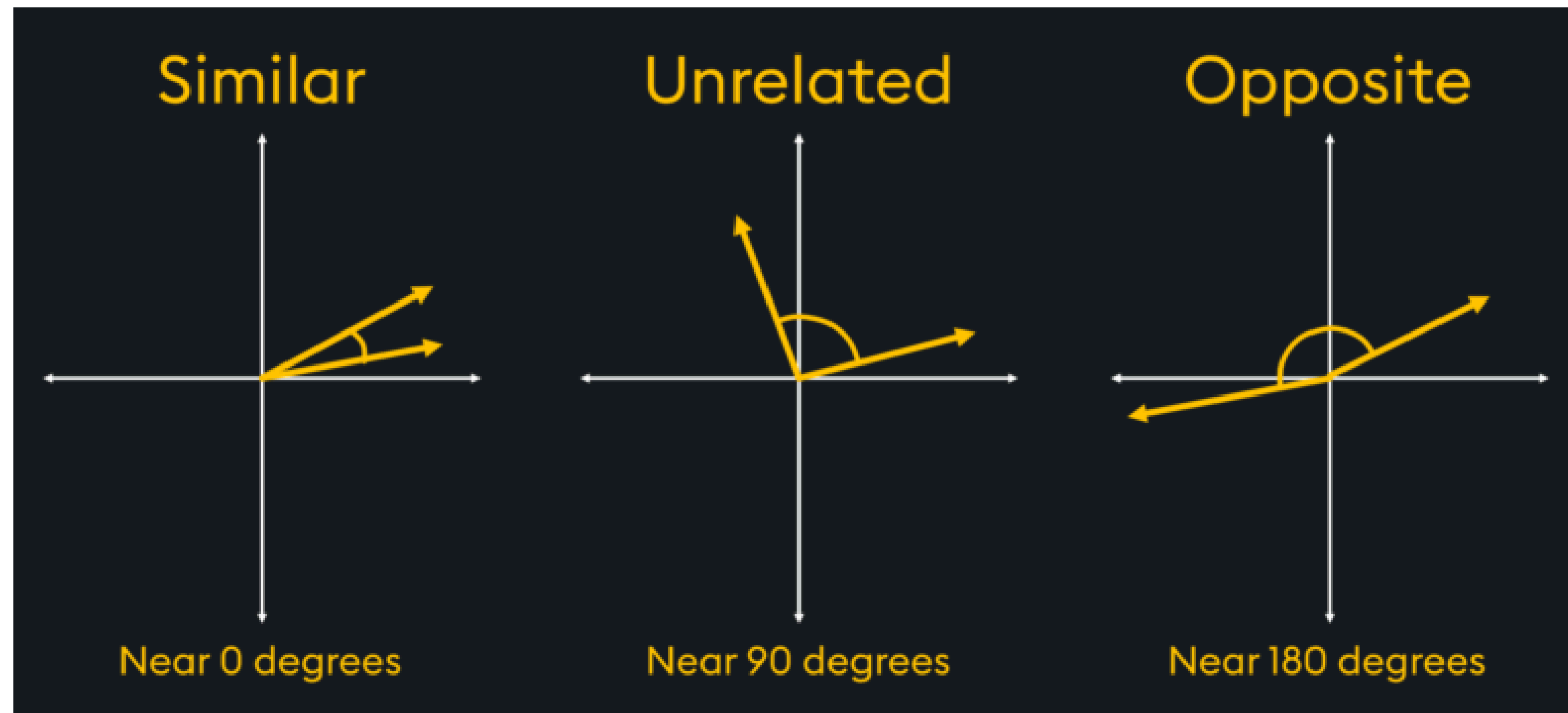
Dogtags

```
{
    "2352": {
        "Label": "Anakin Skywalker"
        "Alternative Label": "Darth Vader/Anakin Skywalker"
        "Type": "interlang"
        "Abstract": "Anakin Skywalker, a Force-sensitive human male, was a Jedi Knight of the Galactic Republic, a hero of the Clone Wars, and the Chosen One
        of the Force. During the reign of the Galactic Empire, he was known as Darth Vader, Sith Lord and apprentice to Emperor Darth Sidious. Skywalker was
        born in 41 BBY, forty-one years before the Battle of Yavin, on the planet Tatooine in the Outer Rim Territories of the galaxy. Although he had a
        mother, Shmi Skywalker, there was no father involved in his conception. This revelation caused Jedi Master Qui-Gon Jinn to suspect that Skywalker had
        been conceived by the midi-chlorians. After liberating Skywalker from slavery, Jinn brought him to the Jedi Temple on Coruscant for training but the
        Jedi High Council deemed the boy too old and ruled by his emotions to become a Jedi. Following the Battle of Naboo, however, the High Council
        reversed their decision in light of the revelation that their ancient foes, the Sith, had returned. Despite the reservations of Grand Master Yoda,
        Skywalker became the Padawan of Jinn's former apprentice, the newly-promoted Jedi Knight Obi-Wan Kenobi ... "
        "Comment": "Anakin Skywalker, a Force-sensitive human male, was a Jedi Knight of the Galactic Republic, a hero of the Clone Wars, and the Chosen One
        of the Force. During the reign of the Galactic Empire, he was known as Darth Vader, Sith Lord and apprentice to Emperor Darth Sidious. Skywalker was
        born in 41 BBY, forty-one years before the Battle of Yavin, on the planet Tatooine in the Outer Rim Territories of the galaxy. Although he had a
        mother, Shmi Skywalker, there was no father involved in his conception. This revelation caused Jedi Master Qui-Gon Jinn to suspect that Skywalker had
        been conceived by the midi-chlorians."
    "2353":
        ...
        ...
        ...
}
```

Graph Embeddings

Cosine similarity

Measures the similarity between two vectors by calculating the cosine of the angle between the two vectors.

Current solution

- Graph embedding with "Dogtags"

- Cosine similarity: top-1

Proposed solution

- Graph embedding with "Dogtags"

- Cosine similarity: top-10

- Use LLMs to pick an entity from top-10

AI integration: LLM usage

- System prompt: "You will get 10 entities from a KG as a string and your task is to select the most similar entity to the one given in the user prompt."

- User prompt: "<GOLD_ENTITY>"

- Output: "<FOUND_PAIR>"

Results so far...

- Current solution (for a certain KG-pair):

  ◦ top 1 → 81% of gold pairs found

- Proposed solution:

  ◦ gold pair is in top 10 → 91%

  ◦ after pairing → ???

Thank you for your attention!

Questions?