

Improving Book Search with AI

Author(s)

Bašić, Suzana; Fuckner, Marcio; Wiggers, Pascal

Publication date

2024

Document Version

Final published version

[Link to publication](#)

Citation for published version (APA):

Bašić, S., Fuckner, M., & Wiggers, P. (2024). *Improving Book Search with AI*. Poster session presented at ICT Open 2024, Utrecht, Netherlands.

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please contact the library: <https://www.amsterdamuas.com/library/contact/questions>, or send a letter to: University Library (Library of the University of Amsterdam and Amsterdam University of Applied Sciences), Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Improving Book Search with AI

Suzana Bašić, Marcio Fuckner, Pascal Wiggers
Responsible AI Lab HvA

BACKGROUND

GOAL: enable smart natural-language search in Dutch on the Amsterdam public library (ObA) website
e.g., “Books about a boy who takes the train to wizarding school” → Harry Potter

DATA

- Sample from the library database with basic information about the books: author, genre, summary, etc.
- The creation of a test dataset is planned for later stages; so far only qualitative evaluation

PHASE 1: KNOWLEDGE GRAPH WITH NAMED ENTITIES

- Create a knowledge graph (KG) from tabular data
- Property graph (Neo4j) for flexibility
- Enrich the KG with named entities extracted from book summaries using a multilingual NER model from the Flair library

An example of book summary



Een fantasie op het historisch verhaal van de tocht van de **Perzische koning Xerxes** tegen de **Grieken, 480 v. Chr.**, met bedekte toespelingen op en parallellen tussen de aanvallen der **Duitsers op België en Frankrijk**.

(Entity labels: Perzische koning Xerxes - person, Grieken - geo-political entity, 480 v. Chr. - date, Duitsers - geo-political entity, België - geo-political entity, Frankrijk - geo-political entity)

Searching by named entities

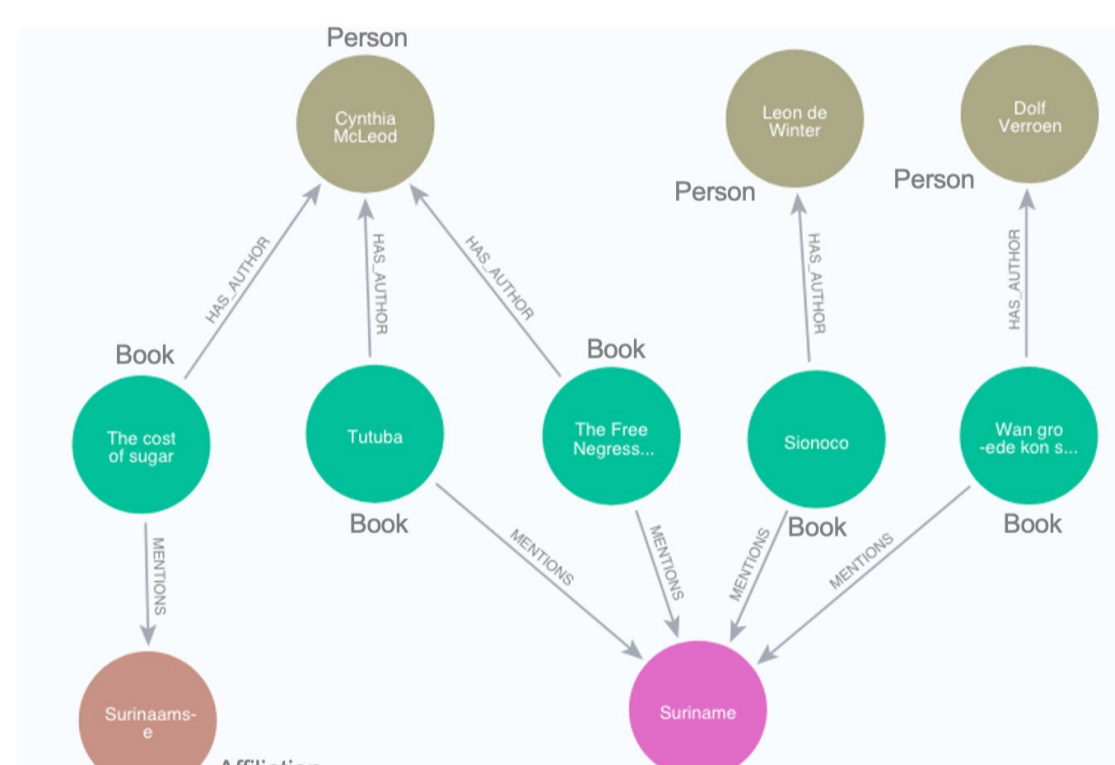
- Authors who write about Suriname
- A book mentioning Orhan Pamuk (autobiography)

Schrijvers die over **Suriname** schrijven

(geo-political entity)

```

MATCH (author:Person)-[:HAS_AUTHOR]-(item:Item),
      (item)-[:MENTIONS]->(entity)
WHERE apoc.text.fuzzyMatch(entity.name,"Suriname")
RETURN author
    
```

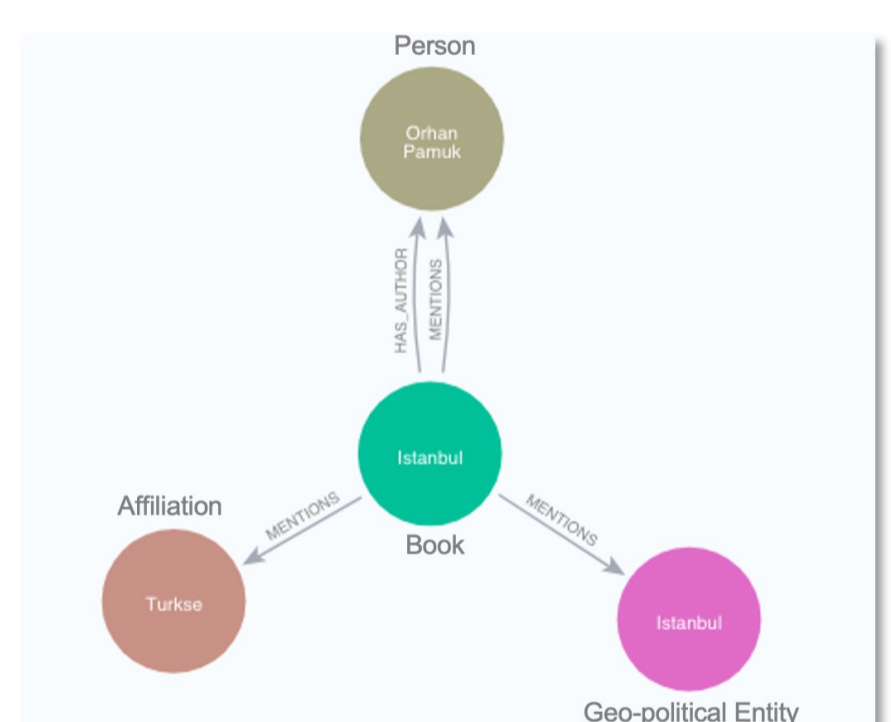


Iets met **Orhan Pamuk**?

```

MATCH (item:Item)-[:MENTIONS]->(p:Person)
WHERE p.name = "Orhan Pamuk"
RETURN item

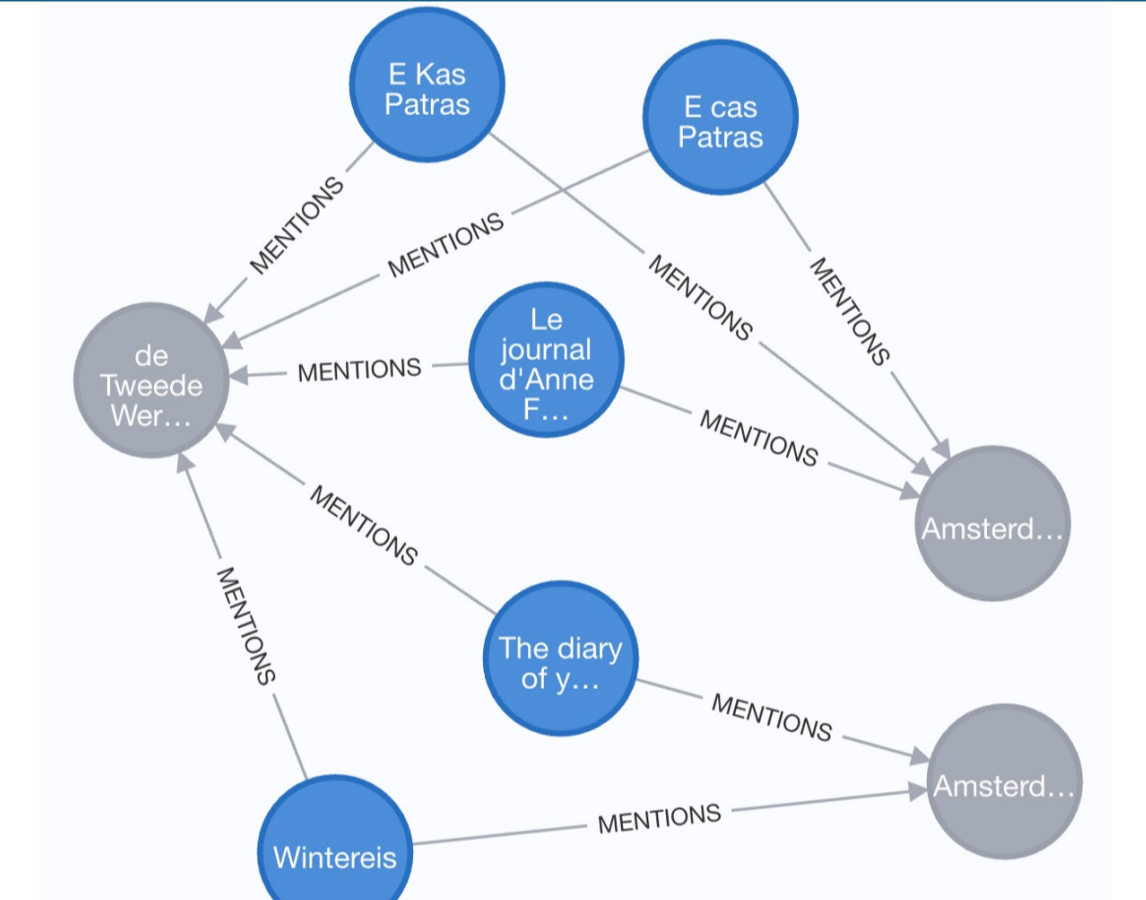
MATCH (item:Item)-[:HAS_AUTHOR]->(author:Person)
WHERE author.name = "Orhan Pamuk"
RETURN item
    
```



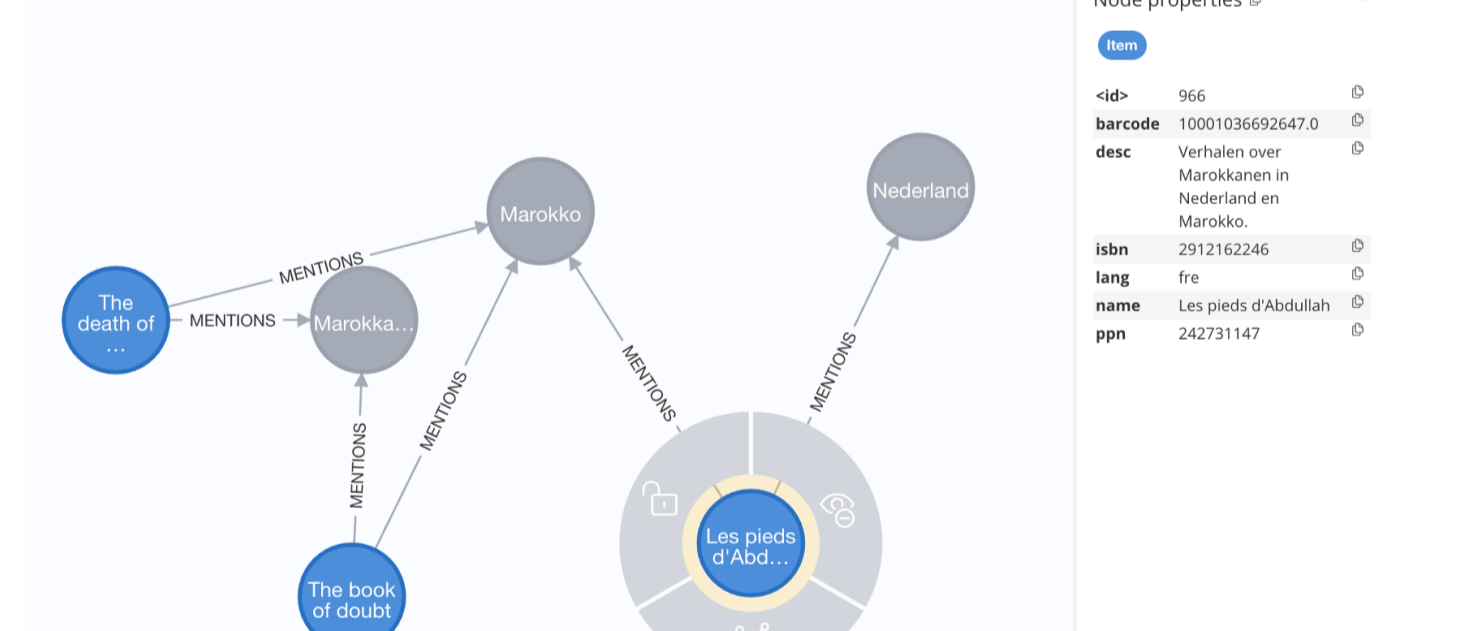
Finding combinations of entities

- “De Tweede Wereldoorlog in Amsterdam”
- “Verhalen over Marokkanen in Nederland”

• “Amsterdam” + “Tweede Wereldoorlog”



• “Nederland” + “Marokko”



PHASE 2: EMBEDDINGS

How to incorporate embeddings into the knowledge graph?

- Embed individual graph entities
- Embed book summaries to find similar books
- Other approaches

PHASE 3: INTEGRATING KG+LLM

Aim: utilize the power of LLMs while retaining a robust fallback strategy

MSc thesis project 1:
Translating natural language queries to Cypher queries

- Use an LLM to automatically turn user searches into formal graph queries

MSc thesis project 2:
Fine-tuning an LLM on the knowledge graph entities

- A closed LLM trained specifically to answer questions about the library collection based on the information from the KG