

Make your company data SPARQL

Integrating a Graph Database with Large Language Models for Natural Language Indexing & Retrieval

Team Project FSS 2025

sovanta AG, 12.02.2025

Dr. Tommi Kramer, Niklas Frühauf

In cooperation with the
Chair of General Management and Information Systems (Prof. Dr. A. Heinzl)



Our **locations** worldwide

Germany | Heidelberg

Germany | Hamburg

Germany | Cologne

Spain | Valencia

Argentina | Buenos Aires

150+ Customers

500+ Projects

2 Mio. User

200 Employees

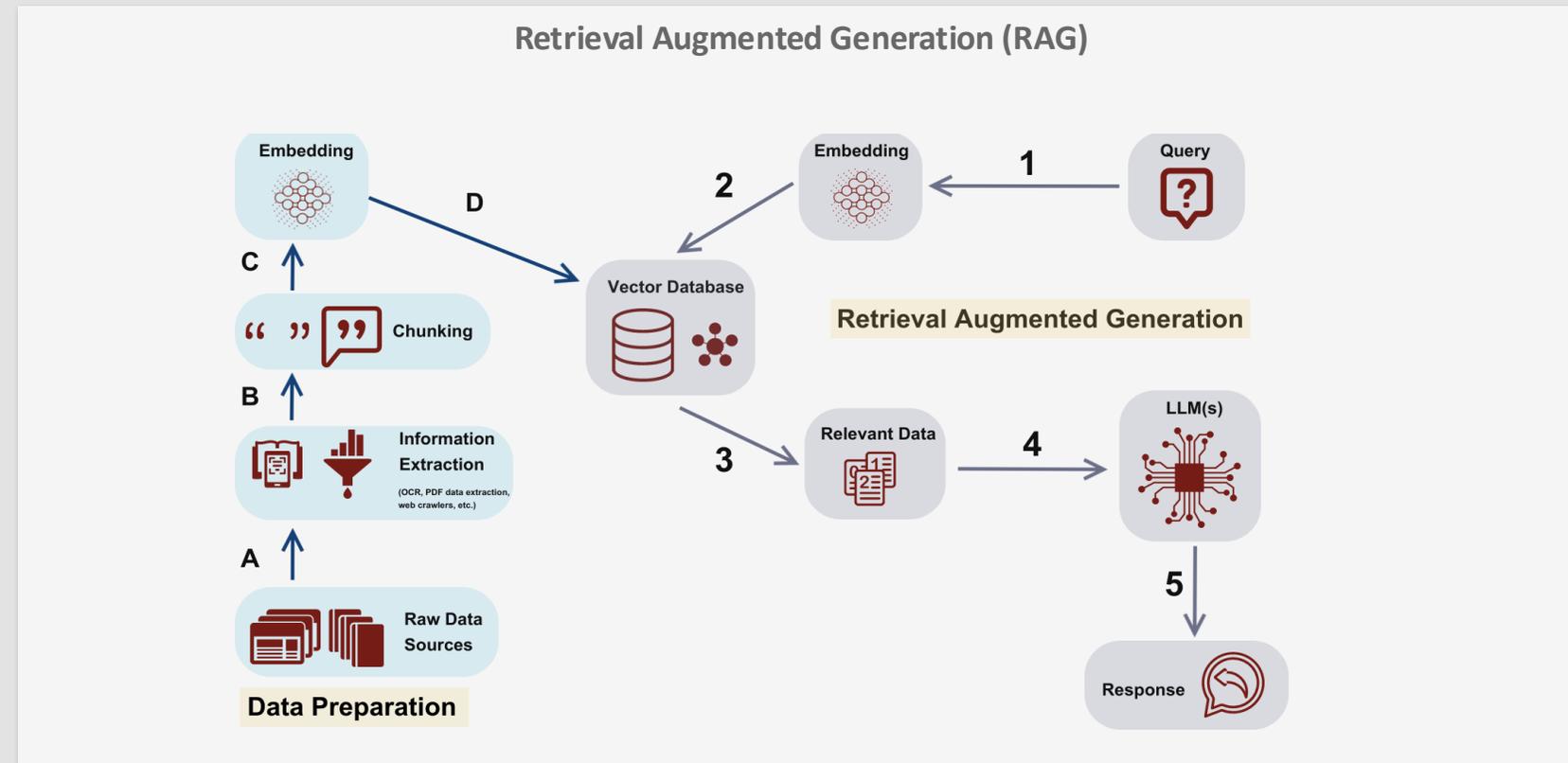
Project Background: Retrieval-Augmented Generation

Works for:

- ✓ Questions that do not require more than 10-15 chunks of text info
- ✓ Simple translations / lookups
- ✓ Simple FAQ

Fails for:

- ✗ Questions that would require a lot of chunks as input
- ✗ Questions that aggregate information
- ✗ Questions that need to relate different parts of different documents to each other



Solution: LLMs + Knowledge Graphs + SPARQL

Indexing (ABC.pdf)

Section 2 Duration of Tenancy The tenancy commences on: 01.01.2024; it is concluded for 5 years. The contracting parties intend to enter into a long-term tenancy. The landlord waives his right to statutory notice of termination (termination for personal use, for use as a granny flat, partial termination and termination for commercial exploitation of the flat pursuant to sections 573, 573a, 573 b of the German Civil Code (BGB)) for a period of 3 years and 9 months from the conclusion of this contract. Therefore, notice cannot be given before this period has elapsed. In other respects, the prerequisites for a termination are governed by statutory provisions and contractual agreements (cf. section 8, 17-22 of this contract). Note: Pursuant to section 22 of this tenancy agreement, the parties to this tenancy agreement may also agree a permanent or long-term waiver of the landlord's right to terminate the tenancy.

LLM Extracts Data

```
ABC.pdf, is, HOUSING_CONTRACT
ABC.pdf, fileName, ABC.pdf
ABC.pdf/Section_2, belongsTo, ABC.pdf
ABC.pdf/Section_2, refersToInternal, ABC.pdf/Section_22
ABC.pdf/Section_2, refersToInternal, ABC.pdf/Section_8
...
ABC.pdf/Section_2, refersToExternal, BGB_573_a
ABC.pdf/Section_2, refersToExternal, BGB_573_b
ABC.pdf, startsOn, 01.01.2024
ABC.pdf, endsOn, 01.01.2029
...
BGB_573_a, belongsTo, BGB_573
BGB_573_b, belongsTo, BGB_573
```

Querying

Q: Which of our open (as of 2025-01-01) housing contracts may directly be impacted by changes to §575b BGB?

LLM constructs query
(e.g. SPARQL)

```
SELECT ?s
WHERE {
  ?s is HOUSING_CONTRACT ;
  fileName ?f ;
  startsOn ?sD ;
  endsOn ?eD .
  FILTER (?sD <= "2025-01-01" && ?eD >= "2025-01-01")

  ?sec belongsTo ?f ;
  refersToExternal ?bgbSec .

  ?bgbSec belongsTo BGB_573 .
}
```

Execute query against KG

```
A:
ABC.pdf
DEF.pdf
...
```

Project Scope & Goals



Construct Knowledge Graph

Using sovanta Confluence, extract past projects, team members, roles, skills and others*



Create Ground Truth & Evaluate

Invent natural-language queries and expected answers, iteratively prompt engineer and evaluate performance of different LLMs on data.



Productize on SAP BTP

Include the code into our existing RAG code-base, add tests, deploy on SAP BTP and verify adaptivity.

Technologies:

- SAP Cloud Foundry
- SAP HANA Database (or RDFLib for EV)
- SAP Gen AI Hub

✓ Indexing & Retrieval should eventually allow dynamic/customer-specified domain models!

Prerequisites

Must Have:

- ✓ Have fun while solving practical problems
- ✓ Knowledge about Generative Models, optimally w/ RAG setup
- ✓ Knowledge Databases & Data Structures
- ✓ Python skills
- ✓ High level of proactivity

Preferred:

- ✓ Knowledge about SAP cloud technologies
- ✓ Knowledge about Graph Databases
- ✓ Semantic Web technologies knowledge

Participants

3 – 5 Students

Language

English/German (for project)

Duration

6 months / 1 semester

Suitability

Wifo, MMDS very welcome

Work Mode

Agile setup, interim presentations



Contacts

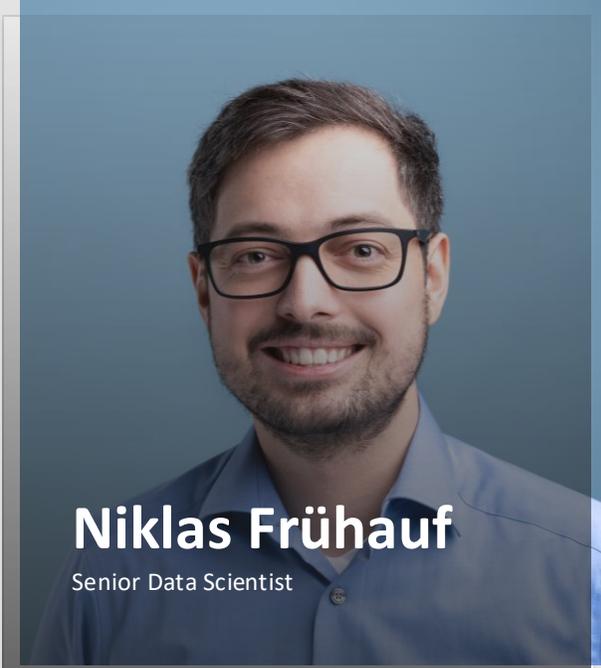


Tommi Kramer

Head of Data Science

P +49 (0)6221 18733-127

tommi.kramer@sovanta.com



Niklas Frühauf

Senior Data Scientist

P +49 (0)175 253-5056

niklas.fruehauf@sovanta.com



Thank you

 +49 (0)6221 18733-0

 info@sovanta.com

 sovanta AG, Mittermaierstraße 31, 69115 Heidelberg