

Trustworthy Semi-Agentic Retrieval-Augmented Generation

Győző Csóka
/dʒø:zø: tʃo:kɒ/

HUN-REN Alfréd Rényi Institute of Mathematics

5th November 2025
DLM Forum Members' Meeting Copenhagen



Historical
Archives of
the Hungarian
State Security



Agenda

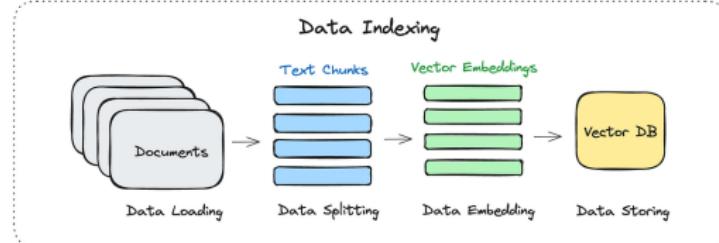
- introduction to Retrieval Augmented Generation
- technical details of our implementation
 - backend
 - agentic workflow
 - UI
- ongoing projects

Problems with plain llm calls

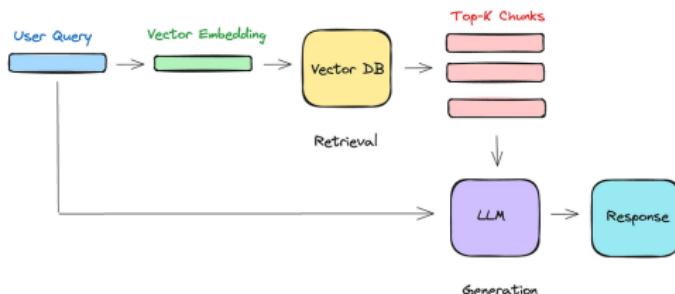
- Lack of factual grounding
- May generate plausible but hallucinated answers.
- Context length & cost limitations
- Can't handle the problem of rapidly changing data

RAG architecture

Basic RAG Pipeline



Data Retrieval & Generation



source: <https://medium.com/@drjulija>

Baseline RAG applications

- NotebookLM
- simple python implementation:

```
from llama_index.core import
    VectorStoreIndex, SimpleDirectoryReader
docs = SimpleDirectoryReader("./docs").load_data()
index = VectorStoreIndex.from_documents(docs)
print(index.as_query_engine()
      .query("Who were the members of the council?"))
```

Imre Nagy



Imre Nagy

- communist reformist leader
- Prime Minister of Hungary (1953–1955, 1956)
- became the face of the 1956 Hungarian Revolution, supporting neutrality and withdrawal from the Warsaw Pact
- arrested after the Soviet invasion that crushed the revolution, executed in 1958 for treason
- posthumously rehabilitated in 1989, his reburial became a symbol of Hungary's democratic turn



Our use case

- Council of Ministers' meeting minutes of the first government of Imre Nagy from 1953-1954.
- original scans of typewritten records
- high-quality human transcriptions in PDF format
- tables of agenda items, laws
- long footnotes



Tables, footnotes

2

[500]

Szigorúan titkos!
Készült a példányban
Kapja: Rákosi Mátyás elvtárs

Jegyzőkönyv

Jelen voltak:

Nagy Imre, a Minisztertanács elnöke, Gerő Ernő, a Minisztertanács elnökjel helyettese, Hegedűs András, a Minisztertanács elnökök helyettese, Hidas Imre vezető miniszter, Szatmirecz Miklós köhly- és építári miniszter, Szalai Balázs, az Országos Tervhivatal elnöke, Boldoczki János külügyminiszter, Okt Károly pénzügyminiszter, Erdéi Ferenc igazságügy-miniszter, Kiss Árpád könyvvilági miniszter, Bognár József bel- és külkereskedelmi miniszter, Almáhori Iván elémérési miniszter, Tisza József bejegyzési miniszter, Bebrits Lajos közelgazdasági miniszter, Darvas József neművelősi miniszter, Edey Gyula Tibor oktatási miniszter, Nográdi Sándor, a honvédelmi miniszter elő helyettese, Klíni József, az építésügyi miniszter elő helyettese, Simon Lajos, az egészségügyi miniszter elő helyettese, Hái Árpád, az Államellenőrző Központi elnöke, Szeitz János, az Országos Tervhivatal elnöknél helyettese, Sára Béla, a Minisztertanács Titkárságának elnöke.

Egyes napirendi pontokhoz mint meghívottak:

- 2.) napirendi ponton: Szász Márton elvtárs
1.) napirendi pont: Tausz György elvtárs
2.) napirendi pont: Tausz György elvtárs és Antos István elvtárs.
3.) napirendi pont: Szász Márton elvtárs, Zentai Ernő elvtárs és Dénes István elvtárs
4.) napirendi pont: Tausz György elvtárs, Antos István elvtárs és Szász Márton elvtárs
5.) napirendi pont: Szász Márton elvtárs és Pócsay Tibor elvtárs

Különfélék: 11.) a.) napirendi pont: Osztromszky György elvtárs; b.) napirendi pont:

10

1955. AUGUSTUS 7.

25.)	503/25/1953.	Kormányellegcím kiküldése a Lippai Várára.
26.)	—	Külföldön vásárolt alkalmazottak.
27.)	—	Szabóly János nekibírált miniszterhelyettes kinevű/s.

Napirenden körül:

Napirendi pont	Hagyomány náma:	A hagyomány címe:
1.)	503/nk/1/1953.	A szerepjátékokról leírószabály utolgozta önmagához tartozó funkcióhalmat szájban rendelkezni kérő hozzájárulás.
2.)	503/nk/2/1953.	Kapitál Pál kinevezése.
3.)	503/nk/3/1953.	Kortárs Béla kinevezése.
4.)	—	Varga András kinevezése.

Napirend előir.

Nagy Imre elvtárs: Bejelenti, hogy Bebrits Lajos elvtárs, kötelezettségszolgáltató miniszter, Altomári Iván élelmiszeripari miniszter és Olt Károly pénzügyminiszter kiértek rendes évi szabadságukat.

Kéri a Minisztertanács hozzájárulását, hogy az említett miniszter elvtársak megkezdhessék szabadiságukat.

[Napirendi]

1.) A termelőszövetkezetek és termelőszövetkezeti csoportok hiteltartósainak rendezése¹

Document structure & chunking

chunking → user query → retrieval → reranking → tool calls → llm call

- extracted data tables
 - government members and their title/role
 - table of contents
- semantic chunking
 - following agenda items
 - footnotes connected to chunks
- metadata:
 - agenda_item_identifier
 - pdf_path
 - pages
 - date
 - minutes_id

Preprocess user query

chunking → user query → retrieval → reranking → tool calls → llm call

- extract keywords
- add synonyms
- find other wordings

chunking → user query → retrieval → reranking → tool calls → llm call

- hybrid retrieval
 - semantic (dense)
 - lexical (sparse)

Reranking

chunking → user query → retrieval → reranking → tool calls → llm call

- vectorised cosine-similarity

Agentic workflow

chunking → user query → retrieval → reranking → tool calls → llm call

- tool calls
 - resolve abbreviations
 - fetch government members
 - fetch name index
 - fetch agenda items

Model choices

chunking → user query → retrieval → reranking → tool calls → llm call

- Local/API
- embedding models
 - bge-m3
 - text-embedding-3-large
- LLM model
 - llama3.1
 - gpt-4o_mini
 - gpt-4o

Model choices

chunking → user query → retrieval → reranking → tool calls → llm call

- Local/API
- embedding models
 - bge-m3
 - **text-embedding-3-large**
- LLM model
 - llama3.1
 - **gpt-4o_mini**
 - gpt-4o

Evaluation via human feedback

- domain knowledge
- human feedback
- retrieval evaluation
- user behaviour & guardrails

User interface

demo

Features

- openwebUI
- streaming
- linking sources
- source highlighting
- chat title generation
- follow-up questions
- user feedback
- admin page

Ongoing projects: Sensitive data detection

Werner Ugyével kapcsolatban ~~nagyon gyakran néztem~~ meg kell jegyezni, hogy az egykori budapesti Gestapo-Ugynökök tevékenysége meg lehetősen hosszúban van a hatóságok előtt. Igen sok ilyen Ugynök ma is szabadon sétál. Igen, hogy sokan ezzel tudnak védekezni, hogy a német bevonulás után ildöntésekkel váltak, de a már említett rendszeráltosás figyelembe-vételével ez a védekeésés elensik. /Ugy láttszik, hogy a német megszállók nem akartak többé miadó megbízottakkal dolgozni./

Rov Krdna-Kinfald nevű zandarmonik törz /Galadoss nevű törz



Werner Ugyével kapcsolatban ~~nagyon gyakran néztem~~ meg kell jegyezni, hogy az egykori budapesti ~~Gestapo-Ugynökök~~ tevékenysége meg lehetősen hosszúban van a hatóságok előtt. Igen sok ilyen ~~Ugynök~~ ma is ~~szabadon~~ sétál. Igen, hogy sokan ezzel tudnak védekezni, hogy a német bevonulás után ildöntésekkel váltak, de a már említett rendszeráltosás figyelembe-vételével ez a védekeésés elensik. /Ugy láttszik, hogy a német megszállók nem akartak többé miadó megbízottakkal dolgozni./

Rov Krdna-Kinfald nevű zandarmonik törz /Galadoss nevű törz

Ongoing projects: Sensitive data detection

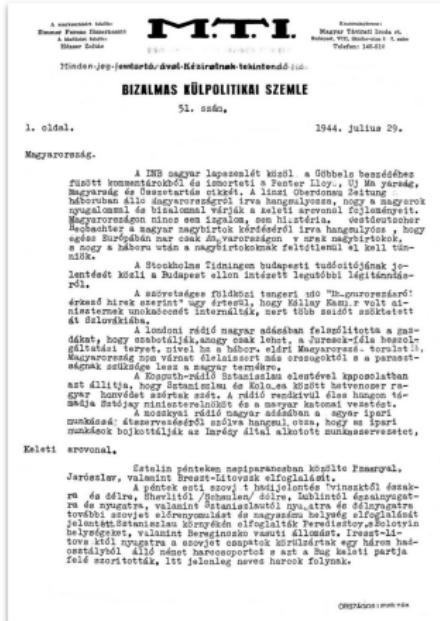
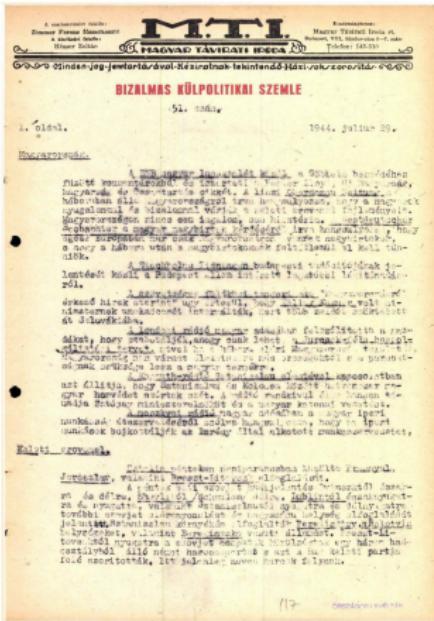
- goal is different per use case
- noisy text
- gpt-oss is big leap for text in Hungarian with OCR errors

Ongoing projects: Summarization



source: <https://abtl.hu>

Ongoing projects: Document image enhancement



Ongoing projects: Document image enhancement

Synthetic Data Generator framework: Stages



Text:

- language
- layout
- fonts

Background:

- paper type
- overbleed

Perturbations:

- 30 noise types
- aged document

Geometric:

- rotation
- warp

Ongoing projects: Document image enhancement



https://huggingface.co/spaces/renyi-ai/die_demo

Partners



Zoltán Lux



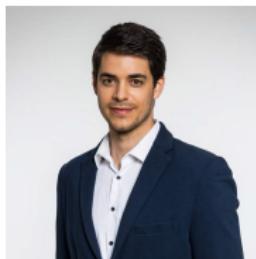
Team



HUN-REN Alfréd Rényi Institute of Mathematics
Artificial Intelligence Group

<https://ai.renyi.hu/>

[in/renci-ai](#)



Gábor Kovács



Domonkos Czifra



Péter Körösi-Szabó



Győző Csóka

Thank you for your attention!
Now it is time for questions.

