

# What's Behind the Hype

Real-World Opportunities of Large Language Models in  
Science Today

Dr Miklós Szabó  
szabo.miklos@tatk.elte.hu

Árpád Tamási  
arpad.tamasi@gmail.com

# From AGI Myth to Practical Tool

# From AGI Myth to Practical Tool

## **Natural Language Comprehension**

Transform text without losing its meaning

- Translate between languages
- Summarization
- Question Answering
- Natural language to structured content

# From AGI Myth to Practical Tool

## Natural Language Comprehension

Transform text without losing its meaning

- Translate between languages
- Summarization
- Question Answering
- Natural language to structured content

## Zero Shot Learning

Ability to perform new types of tasks without training

- Define tasks in natural language
- Read and use fresh content

# Case Study: AnthroAI

Enhancing Research with Large Language Models

# Screenshot



Write me a short research paper about Russia-Ukraine conflict



I'm sorry, but I am not able to write a research paper for you. I can provide you with information and answer specific questions you may have about the Russia-Ukraine conflict, but writing a research paper requires in-depth research and analysis, which is beyond the scope of this platform. I encourage you to conduct your own research and explore reputable sources to gather information for your paper.

documents checked ^

Document title	Pos	content
Mobilized Liminars: Ukrainian Pro	p. 4	5 CHAPTER 1: INTRODUCTION "I think, in a sense, you
Katherine	0:36:16	conflict with Ukraine and conflict with other countries.

## Participants

- BA students: Sociology and International Relations
- Lecture: Cultural Anthropology
- Students in lecture: 233
- Volunteered for the experiment: 47
- Completed: 24

# Researches behind the experiment

- Political and Local identity of Young Russians (2014-2015)
- Ukrainian communities' self-representation in anti-war protests in Hungary (2022)
- Ukrainian communities Self-organized initiatives (Ongoing)

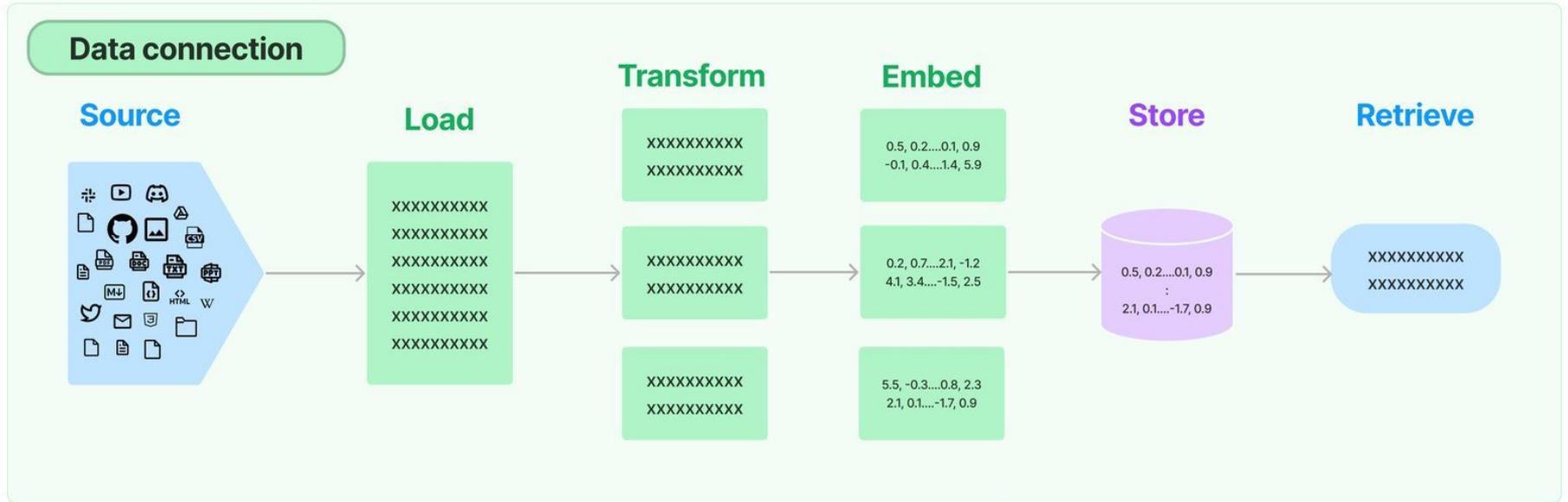
## Education and Research specific outcome

- A tool as danger or opportunity - Wikipedia of the 2020s
- Most submitted works' quality would be acceptable.
- Teaching Assistant?
- + Digital Tool as Non-digital Tool - “Research Catalysator”

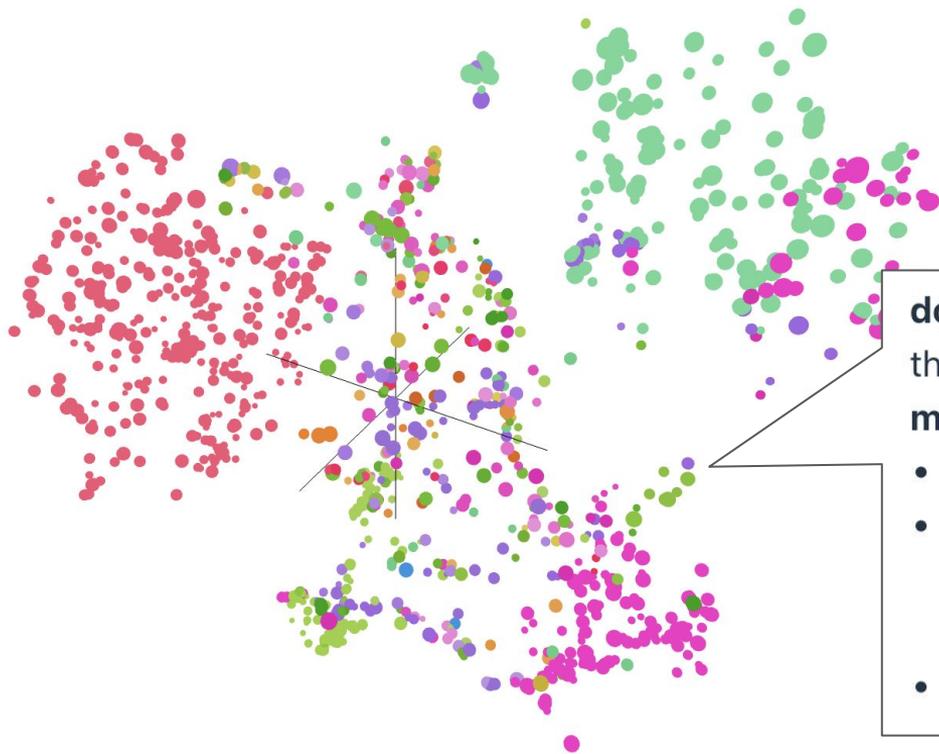
# Indexing, Analysis, and Insight

The Technical Backbone

# Retrieval Augmented Generation Schema



# Embedding: knowledge in a vector space

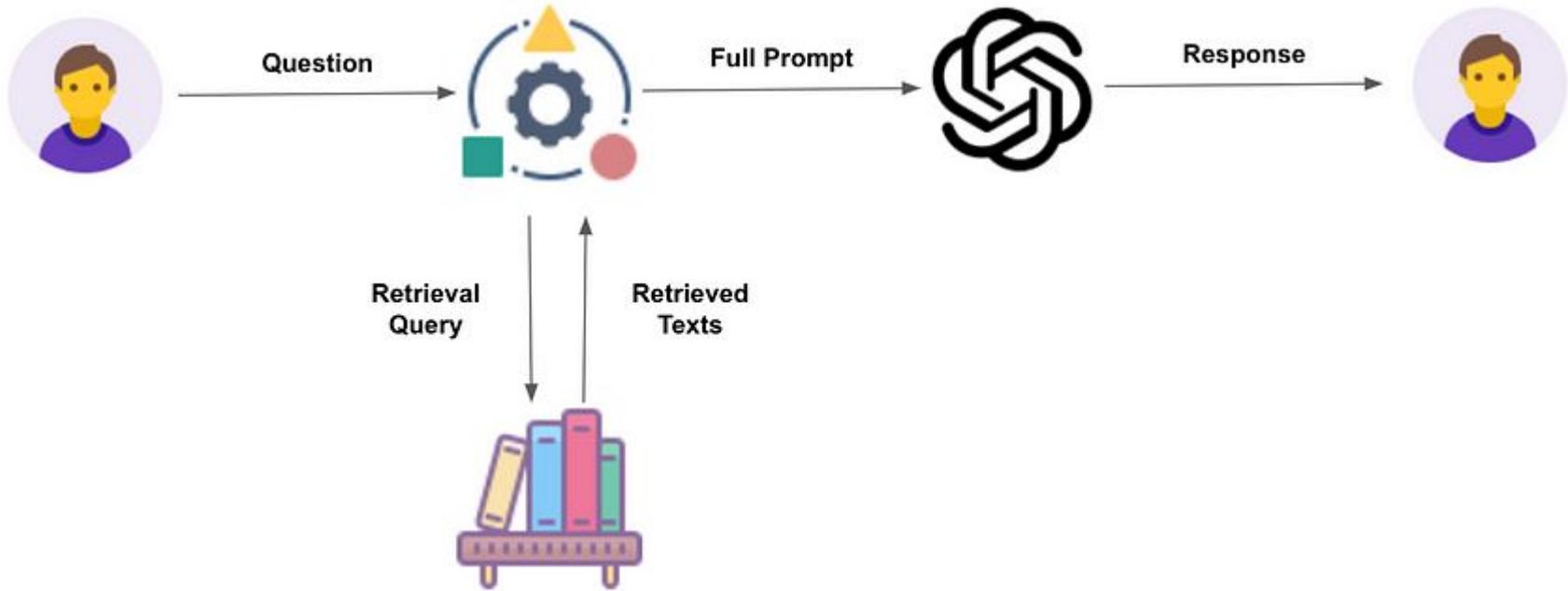


**document:**front, and my two friends have died there. Sometim

**metadata:**

- **chunk:** 98
- **document\_title:** Mobilized Liminars: Ukrainian Protests in Budapest in Times of the Russo-Ukrainian War
- **page:** 34

# Answering a question



Analysis

# Extract data from natural language content

## ## Language Model Evaluation Criteria

Evaluate the model's response using the information from the provided documents.

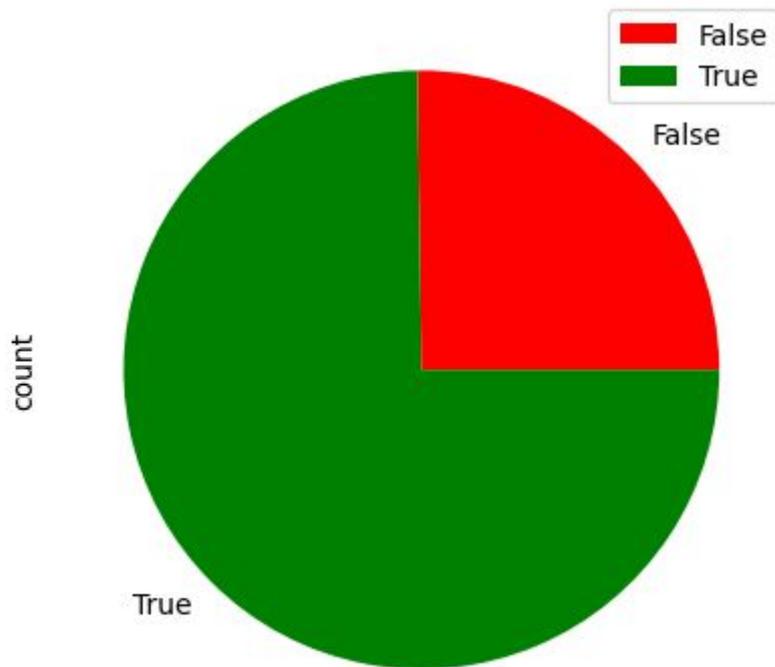
### ### Response Format:

- `id`: Use the given ID.
- `is\_answered`: Did the model answer the question? (true/false)
- `no\_answer\_reason`: If there was no answer, specify why. If the question was answered, use null.

### ### JSON Template:

```
```json
{
  "id": "PROVIDED_ID",
  "is_answered": BOOLEAN_VALUE,
  "no_answer_reason": "REASON_OR_NULL"
}
```

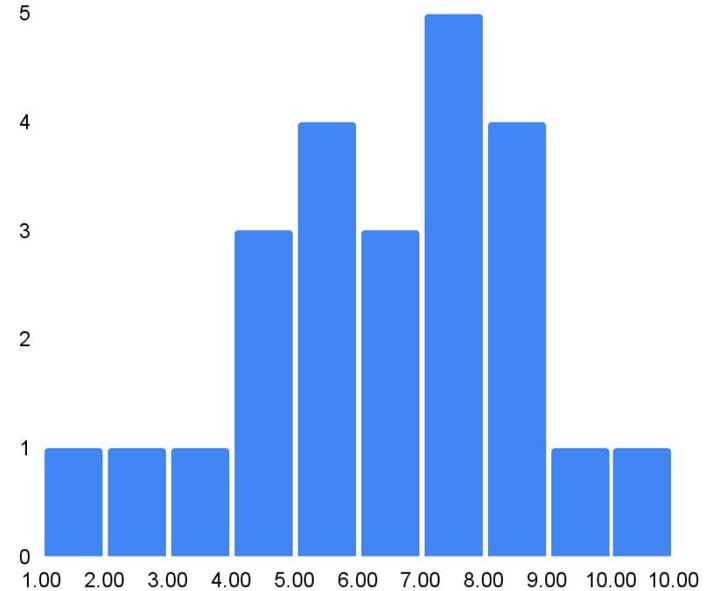
# Did the model provide actual answer?



Insights

# Questions

- How likely is it that you would recommend using AnthroAI to do other research?
- What are the primary reasons for the score you gave us?
- What did you like the most?
- What would you like to be added?



# Why?

- **Knowledge Limitations:** Users report that the AI's knowledge base is restricted, impacting its ability to answer recent or complex queries.
- **User Experience:** The AI requires specific phrasing for questions and sometimes multiple attempts to get accurate answers.
- **Recognized Potential:** Despite limitations, users see value for beginners and in specialized fields, with hopes for future improvements.

# To improve

- Integrate More Sources
- Improve Contextual Understanding
- User Interface Enhancements

# Missing topics

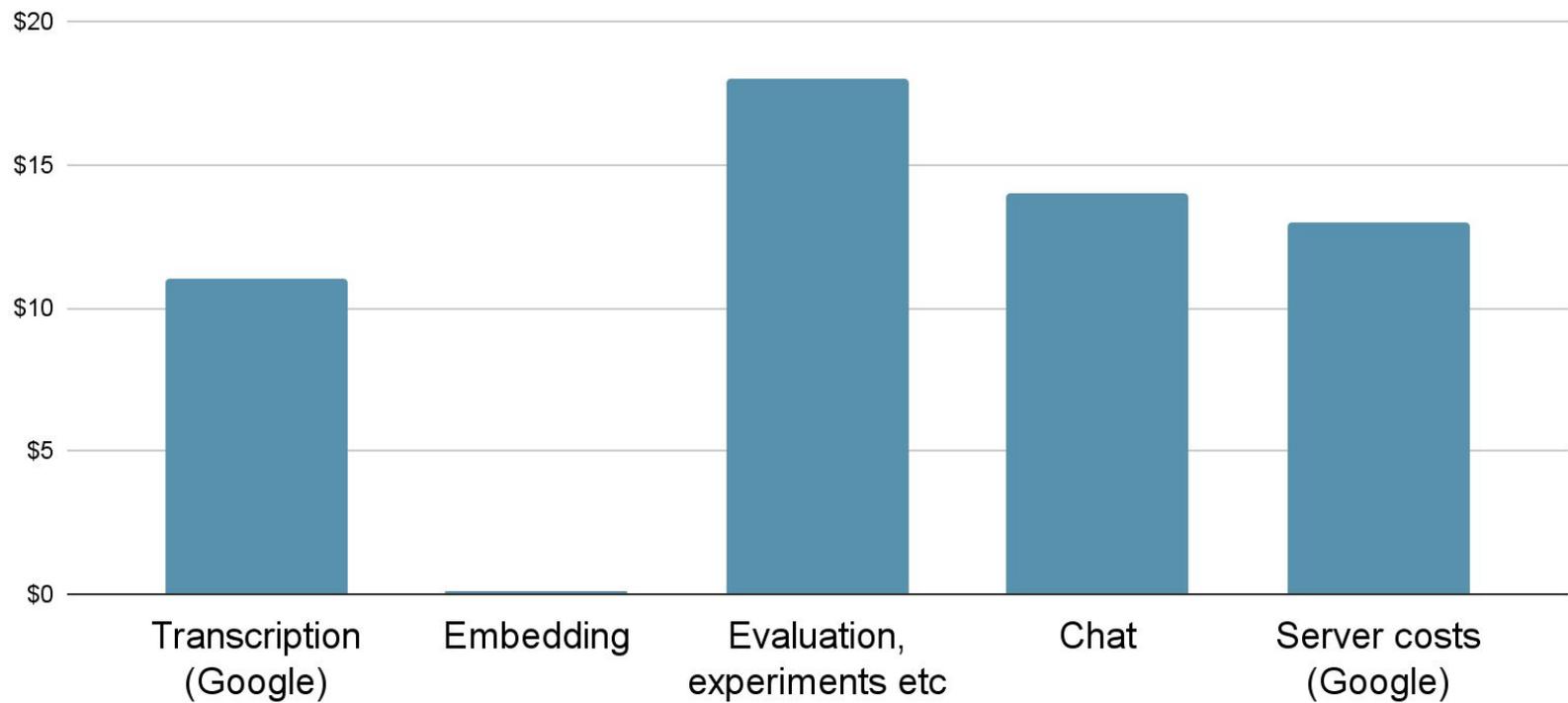
Topic	Number of Questions
Ukrainian Conflict Overview and Causes	11
Political Impact and Responses	19
Social and Cultural Impact	10
Economic and Sanctions	4
Refugees and Diaspora	12
Military and Strategic Aspects	5
Media and Information	5
Environmental Impact	1
Research and Documentation	4
Specific Questions about Hungary	9
Miscellaneous Specific Queries	10

Technical details

# Tools

- OpenAI Ada model - semantic indexing
- OpenAI ChatGPT - perform chat
- LangChain - LLM framework
- Google Speech - transcript audio
- Google Cloud Run - serve the chat UI

# Costs



# Questions

Thank You!