

RETRIEVAL AUGMENTED GENERATION

András Kornai

RAG 2024/09/11

ADMIN STUFF

- Course website <https://nessie.ilab.sztaki.hu/~kornai/2024/RAG>
- If you are not on Slack you don't exist
- What do you want to do with RAG?
- Homework, grading

TOPICS THAT WILL BE DISCUSSED (NOT NECESSARILY IN THIS ORDER)

- 1 Ancient history (IR)
- 2 Hallucination
- 3 Implementing (some flavor of) RAG
- 4 Measuring the effectiveness of the proposed techniques
- 5 Understanding LLM storage and deduction better

CONCLUSIONS REACHED

- We discussed the origins of RAG in blind relevance feedback
- We will concentrate on the task of mitigating source hallucination
- We will probably use old TREC data to test this
- We will move the physical class to SZTAKI
- Read https://nessie.ilab.sztaki.hu/~kornai/2024/RAG/Resources/xu_1996.pdf
- Next week: guest talk, zoom link will be different (will be posted on slack)