QUT Digital Observatory

RAG and AI: Ethics advisor

AI is creating fake legal cases and making its way into real courtrooms, with disastrous results

Airline held liable for its chatbot giving passenger bad advice - what this means for travellers

23 February 2024

Share Save +

Maria Yagoda
Features correspondent

Retrieval Augmented Generation

- RAG introduces a information retrieval process

- This enhances the content generation process by presenting relevant data as context

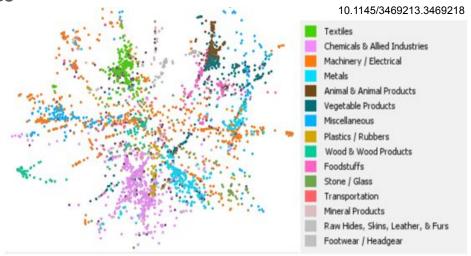# Lots of documents; but human memory

The tasks:

- Create an index of all of our documents

- Search for documents (semantic search)

- Ask questions, receive answers based on document text

- Discuss documents*

https://tinyurl.com/resbaz-ragai

Embeddings - what are they?

- Vector representations of data

- Aim is to simplify complex data while maintaining meaningful relationships and similarities

- Similar items end up near each other in vector space

10.1145/3469213.3469218

- Textiles
- Chemicals & Allied Industries
- Machinery / Electrical
- Metals
- Animal & Animal Products
- Vegetable Products
- Miscellaneous
- Plastics / Rubbers
- Wood & Wood Products
- Foodstuffs
- Stone / Glass
- Transportation
- Mineral Products
- Raw Hides, Skins, Leather, & Furs
- Footwear / Headgear

Chunking and token counts

- We need to maintain a context window 'balance'
    - Too short of a chunk and we lose necessary context
    - Too long of a chunk and we dilute meaning

- Small chunks suit specific answers; Large suit broader themes

- Should we overlap chunks?

Australian Digital Observatory

digitalobservatory.net.au

AusReddit beta

Archive of Australian Reddit

NEWS TALK

youte+

Beta Version 0.0.1

Collect video metadata and comments for research