council_analytics

Christoph Blattgerste

Dec 20, 2024

CONTENTS:

Add your content using `reStructuredText` syntax. See the reStructuredText documentation for details.

CHAPTER

ONE

LLM_RAG PACKAGE

1.1 Submodules

1.2 llm_rag.query module

class llm_rag.query.**RAG_LLM**

Bases: object

configure_query_engine(*index*) → RetrieverQueryEngine

display_prompt_dict(*prompts_dict*)

huggingface_login()

init_embedding_model(*embed_name*)

init_llm_model(*llm_name*, *token*)

load_index_storage(*index_dir*)

query_rag_llm(*user_query*)

1.3 llm_rag.web_app module

1.4 Module contents

CHAPTER

TWO

PREPROCESSING PACKAGE

2.1 Submodules

2.2 preprocessing.download module

`preprocessing.download.`**`download_pdf`**`(`*idx*, *verbose=False*`)`

`preprocessing.download.`**`request_pdf`**`(`*idx*, *verbose=False*`)`

Request the PDF file from the municipal council website. Returns the content of the file if it's a PDF, otherwise None.

2.3 preprocessing.embedding module

`preprocessing.embedding.`**`initVectorStore`**`(`*embedding_model*`)`

`preprocessing.embedding.`**`initialize_embedding_model`**`()`

`preprocessing.embedding.`**`load_txt_files`**`(`*directory*`)`

`preprocessing.embedding.`**`load_txt_files_nextcloud`**`(`*start_idx*, *end_idx*`)`

2.4 preprocessing.extractor module

`preprocessing.extractor.`**`extract_text`**`(`*doc*`)`

`preprocessing.extractor.`**`save_text`**`(`*text: str*, *fout: str*`)`

2.5 preprocessing.preprocessing module

`preprocessing.preprocessing.`**`download_from_nextcloud`**`(`*folder*, *filename*`)`

`preprocessing.preprocessing.`**`parallel_process_pdf`**`(`*args*`)`

A helper function to process the PDFs in parallel.

`preprocessing.preprocessing.`**`process_pdf`**`(`*idx*, *verbose=False*`)`

Process the PDF by downloading, uploading to Nextcloud, extracting text, and uploading the text file to Nextcloud.

`preprocessing.preprocessing.`**`upload_to_nextcloud`**(*folder*, *filename*, *content*, *content_type='binary'*, *verbose=False*)

Upload the given content (PDF or text data) to Nextcloud. If content_type is 'text', it will upload the content as text (encoded in utf-8), otherwise it will treat it as binary content.

2.6 Module contents

PYTHON MODULE INDEX

INDEX