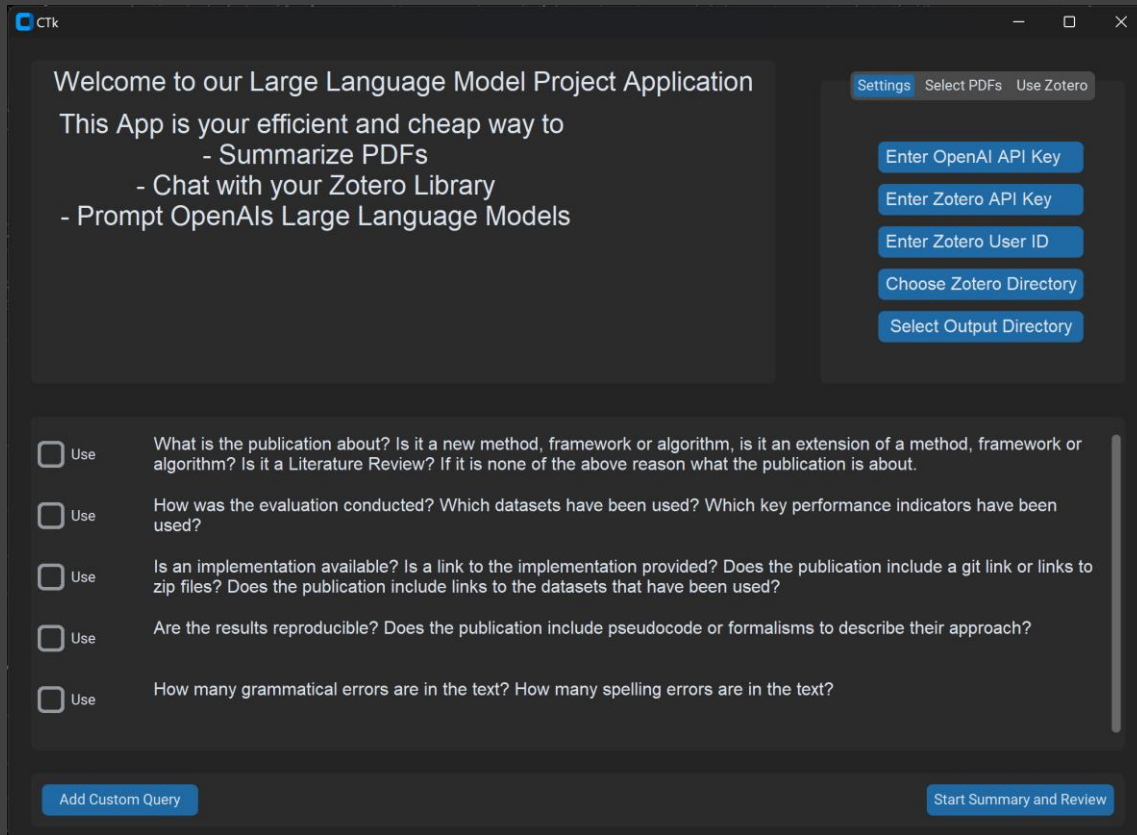


# Automation Helper for Literature Reviews

Kaan Apaydin, Stefan David Horak, Yorck Zisgen

# Automation Helper for Literature Reviews



The screenshot shows a web application titled 'Automation Helper for Literature Reviews'. The interface is dark-themed. At the top, there's a welcome message: 'Welcome to our Large Language Model Project Application'. Below this, it states 'This App is your efficient and cheap way to' followed by a list of features: '- Summarize PDFs', '- Chat with your Zotero Library', and '- Prompt OpenAIs Large Language Models'. On the right side, there are tabs for 'Settings', 'Select PDFs', and 'Use Zotero'. Under the 'Settings' tab, there are five input fields with corresponding buttons: 'Enter OpenAI API Key', 'Enter Zotero API Key', 'Enter Zotero User ID', 'Choose Zotero Directory', and 'Select Output Directory'. Below the settings, there are five checkboxes, each labeled 'Use', followed by a list of prompts: 'What is the publication about? Is it a new method, framework or algorithm, is it an extension of a method, framework or algorithm? Is it a Literature Review? If it is none of the above reason what the publication is about.', 'How was the evaluation conducted? Which datasets have been used? Which key performance indicators have been used?', 'Is an implementation available? Is a link to the implementation provided? Does the publication include a git link or links to zip files? Does the publication include links to the datasets that have been used?', 'Are the results reproducible? Does the publication include pseudocode or formalisms to describe their approach?', and 'How many grammatical errors are in the text? How many spelling errors are in the text?'. At the bottom, there are two buttons: 'Add Custom Query' and 'Start Summary and Review'.

Welcome to our Large Language Model Project Application

This App is your efficient and cheap way to

- Summarize PDFs
- Chat with your Zotero Library
- Prompt OpenAIs Large Language Models

Settings Select PDFs Use Zotero

Enter OpenAI API Key

Enter Zotero API Key

Enter Zotero User ID

Choose Zotero Directory

Select Output Directory

☐ Use What is the publication about? Is it a new method, framework or algorithm, is it an extension of a method, framework or algorithm? Is it a Literature Review? If it is none of the above reason what the publication is about.

☐ Use How was the evaluation conducted? Which datasets have been used? Which key performance indicators have been used?

☐ Use Is an implementation available? Is a link to the implementation provided? Does the publication include a git link or links to zip files? Does the publication include links to the datasets that have been used?

☐ Use Are the results reproducible? Does the publication include pseudocode or formalisms to describe their approach?

☐ Use How many grammatical errors are in the text? How many spelling errors are in the text?

Add Custom Query

Start Summary and Review



Kaan Apaydin

M. Sc. Business Information Systems, CAU  
Ph.D. Student at the University of Kiel  
at the Chair of B.I.S. and Process Analytics

Stefan David Horak

B.A. Multimedia Production, FH Kiel  
10 years in video production  
Founder MALA Yoga, SmarterDoc



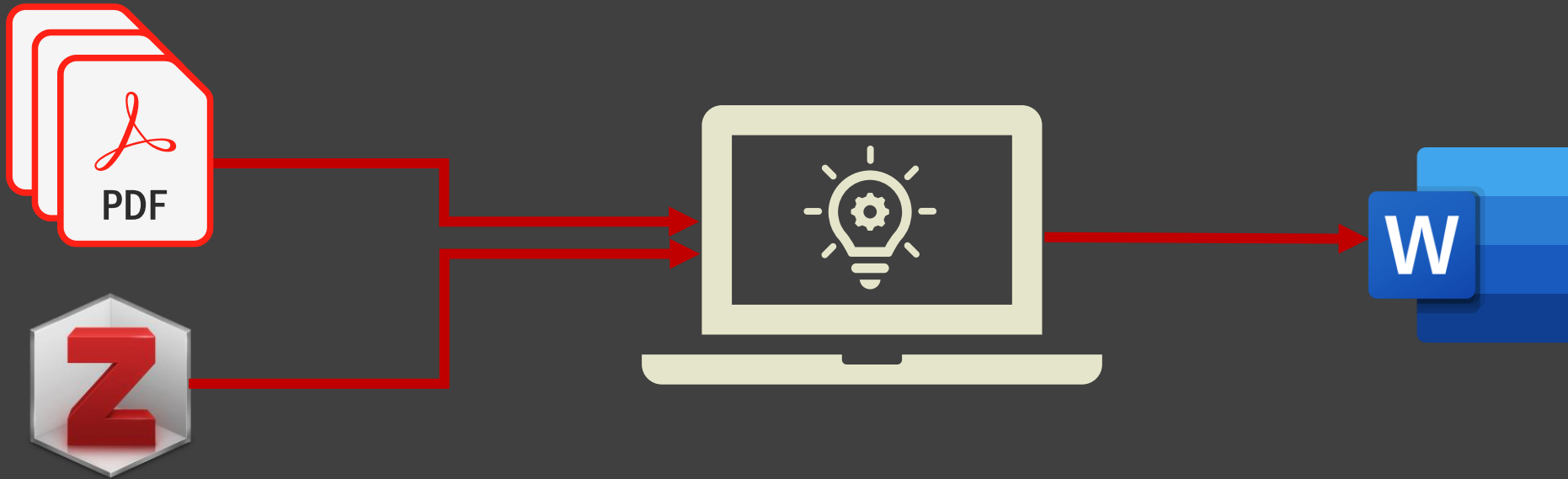
Yorck Zisgen

M. Sc. Business Information Systems, CAU  
Ph.D. Student at the University of Bayreuth  
at the Chair of B.I.S. and Process Analytics



# Automation Helper for Literature Reviews

- The Automation Helper takes a set of PDF files as input ...



- ... and generates a single output file (.docx format).

# Automation Helper for Literature Reviews

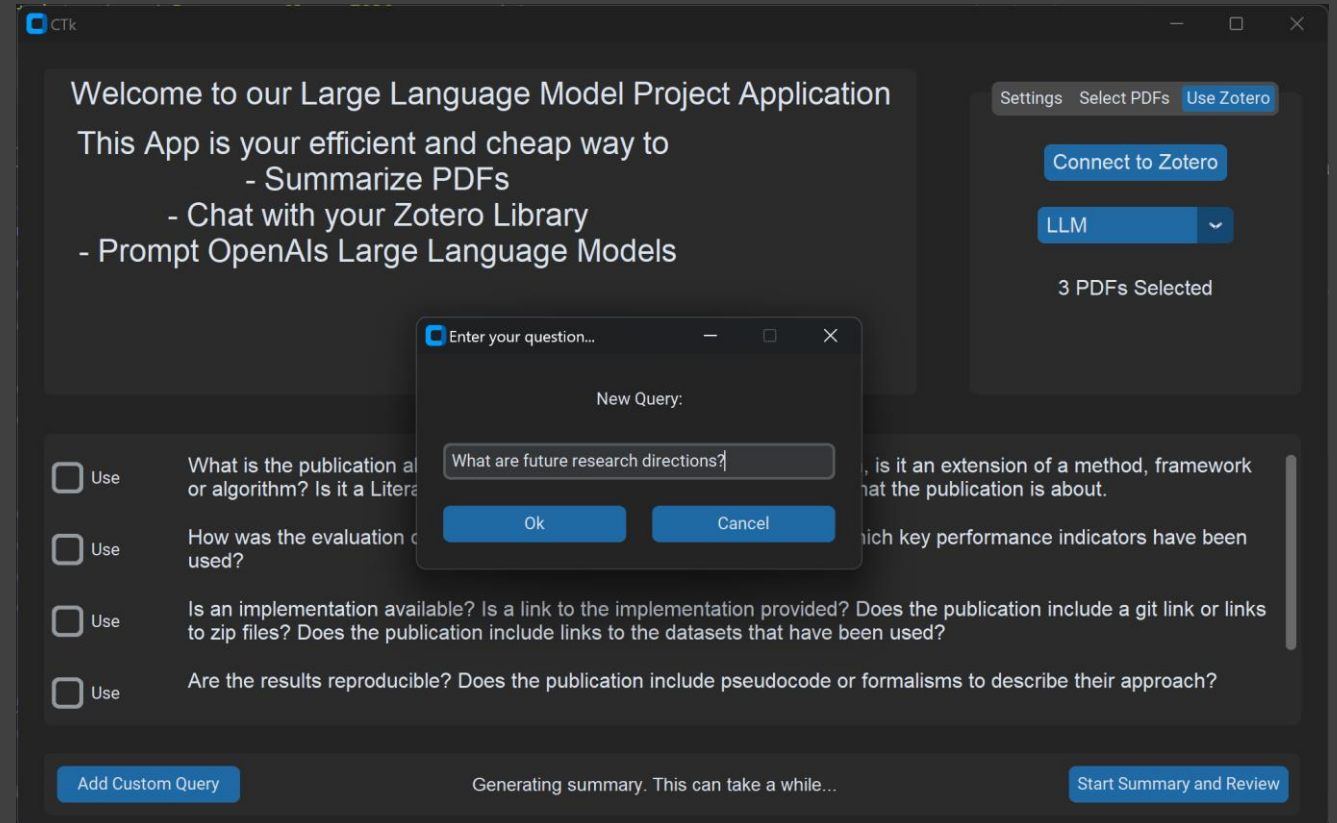
- The output file contains:
  - Title, authors, publication date,
  - a set of related keywords,
  - word count and estimated read length,
  - short and long summaries,
  - answers to all queries given (incl. references),
  - all images contained in the original PDF file
  - ... of every PDF file passed into it.
- Useful for researchers in all domains.

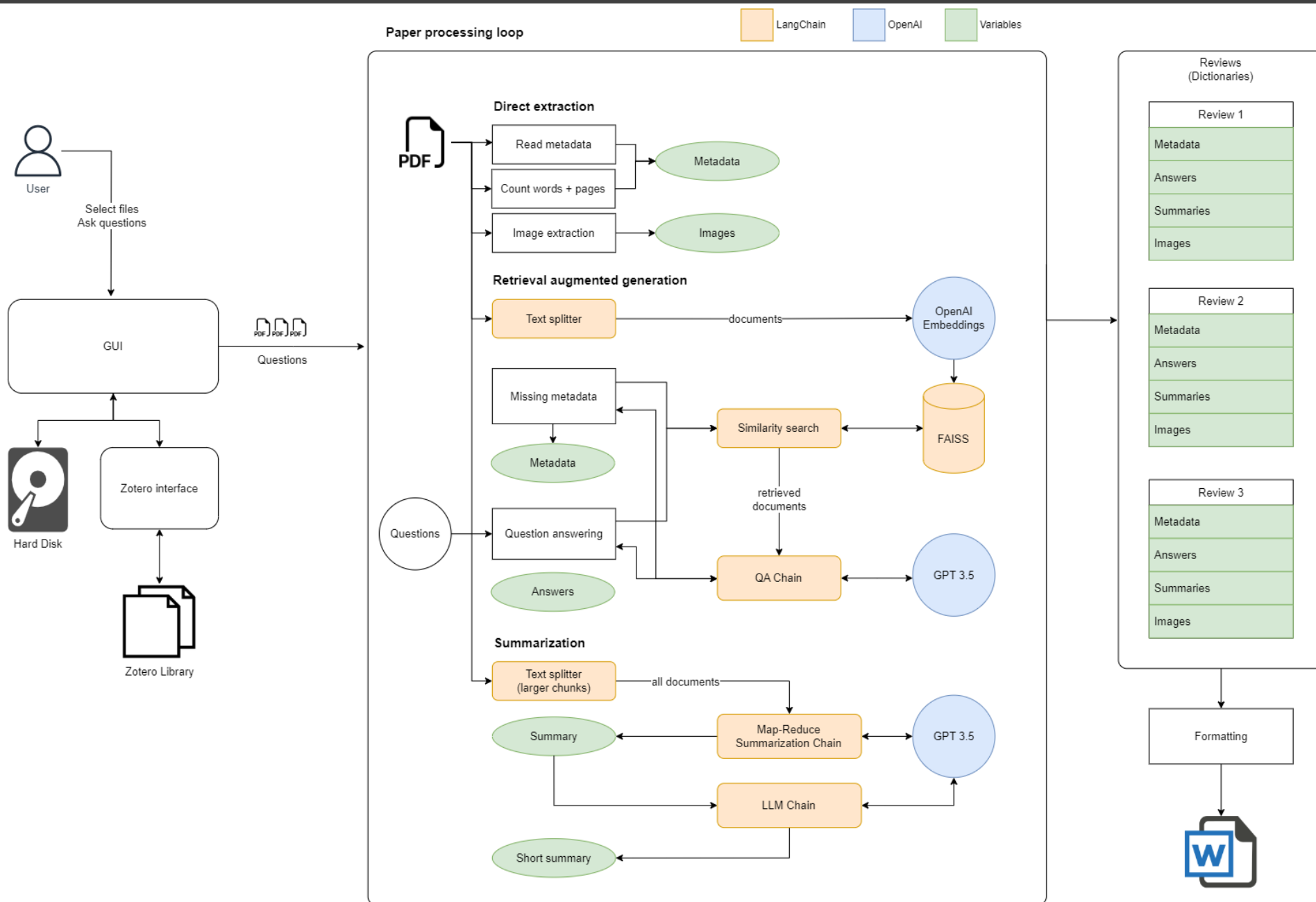


# Automation Helper for Literature Reviews

## GUI

- Started with Tkinter
  - python based GUI for small prototypes
- Transferred to CustomTkinter
  - more modern and consistent design
  - still lightweight with similar ease of implementation
- Outlook: Transfer to Flet
  - Python-based cross-platform supporting App Development Framework powered by Flutter
  - Truly modern and responsive design





## Data Handling

- pypdf2
- fitz
- Pillow
- docx

## Embeddings

- OpenAI

## LLM

- OpenAI GPT 3.5 Turbo
- LangChain

## Database

- FAISS

## Challenges and Solutions

- Q&A vs. summarization
- Metadata
- Images

# Automation Helper for Literature Reviews

- Future Enhancements I (until release)
  - Successfully extracting header data (authors, title)
  - Improving image extraction
  - Improving GUI
  - Cleaning up code, adding a license, and **publishing** the project as open source

# Automation Helper for Literature Reviews

- Future Enhancements II (after release)
  - Connecting to Zotero group libraries and potentially Citavi as well
  - Moving queries from the GUI into a loadable file (e.g. XML)
  - Working on run-time-per-paper
  - Adding singular analysis documents to each Zotero entry



# Automation Helper for Literature Reviews

- Closing Remarks
  - Decreases cognitive load of assessing the relevance of an unknown paper
  - Decreases evaluation time of papers
  - Adds automated insight such as word count, estimated read time, categorization of the paper, or applied evaluation methods
  - Supports and speeds up elementary scientific work

# Automation Helper for Literature Reviews

- Discussion -

# Automation Helper for Literature Reviews

- The Automation Helper takes a set of PDF files as input ...
  - from an 'Open File Dialogue'
  - or from a connected Zotero library
- It generates a single output file (.docx format) that contains:
  - title, authors, publication date,
  - a set of related keywords,
  - word count and estimated read length,
  - short and long summaries,
  - all responses to the prompts asked via the GUI
  - all images contained in the original PDF file
  - ... of every PDF file passed into it.
- Useful for Researchers in all domains.