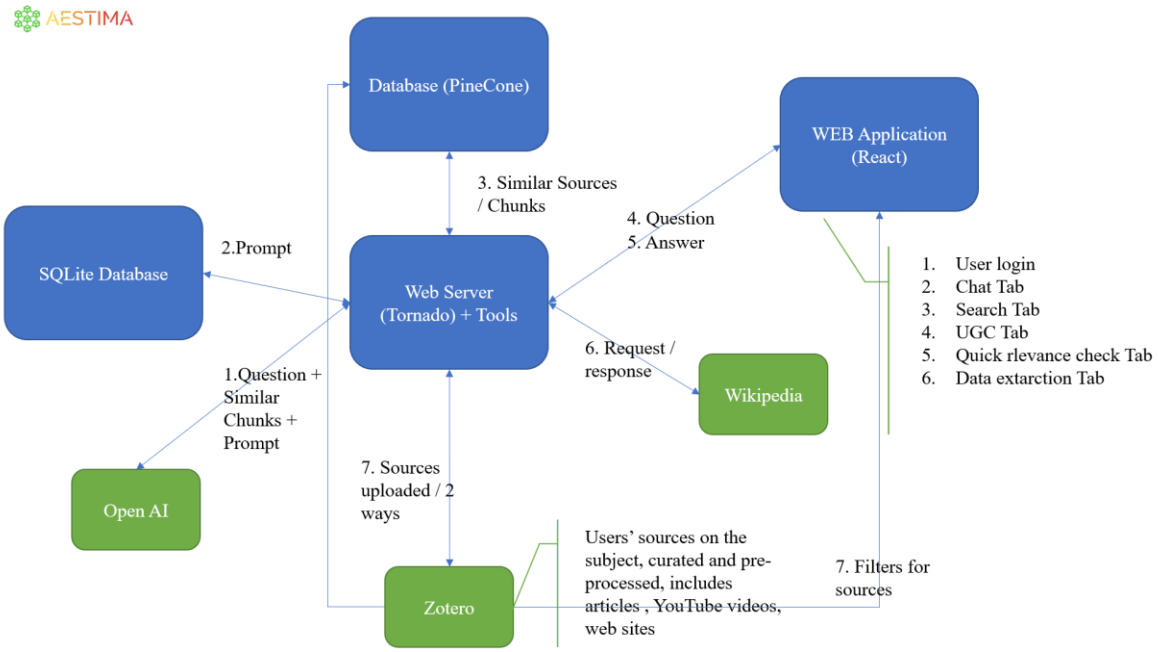
# The Aestima SuperBot features. 22.03.2024 Release.

## Current Architecture




The Aestima SuperBot is designed to work with user sources stored in Zotero. Those sources can include such as pdfs, and/or notations, and or other items such as links to YouTube or WEB sites.

Information is then taken from pdfs/notes in Zotero and cut in chunks (pieces of the pdfs/notes) which are stored in Pinecone vector database and embedded. Some of the metadata, such as TAGS , is taken directly from Zotero.

## Search tab

1. Searching for sources[1] which are stored in Zotero in a user's database.
    a. Searching in articles, from user's knowledge base only
    b. Searching in YouTube, from user's knowledge base only (item needs to be created in Zotero, with a link to the video)
    c. Searching in Web Pages, from user's knowledge base only (item needs to be created in Zotero, with a link to the webpage)

[1] every time a reference is made to … it applies to the user's Zotero collection.

2. Filtering by time of publication/creation
3. Filtering by Zotero tags (AND/OR)
4. Filtering by ABS journal rating
5. A reference (a link) to a pdf chunk the LLM's response is based on is provided as well as full chunk
6. Display of ZOTER ID – unique coding of items in Zotero
7. A link to the original source (a doi link) if available
8. A link to the source (article) detailed info and data extraction summary is available for each source.
9. Sources quick summary = insight on the user's query
10. Automated papers relevance scoring (from 1 to 10) by hard-coded prompting and relevance scoring system
11. Sorting sources by their relevance
12. Creation of users preferred sources set (a starred collection)
13. Selecting sources for further analysis/extraction by LLM (check the box)
14. Customized columns - extracting additional data from sources by users' choice (simplified data extraction)
15. Editing prompts for data extraction
16. Recalculation of customized columns and cells with a choice of LLM model
17. Count display (sources checked, starred and total)
18. A copy button for the content of specific data extraction cells
19. Uploading search results and all extracted data to excel
20. Multiple searches

    **Advanced automated data extraction (**8 hard-coded questions for every article: Definitions of the key terms, Concepts, Research limitations, Sample, Data analysis methods, Results, Further research areas, and Strengths)
21. Summary of data extraction from all searches (article card, same as in the Advanced data extraction Tab)
22. Creation of customized data extraction categories
23. Searching articles from which data was extracted by the title

**Chat tab**

1. Searching for sources[2] which are stored in Zotero in a user's database
    a. Searching/querying in articles, from the user's knowledge base only
    b. Searching/querying in YouTube, from the user's knowledge base only (item needs to be created in Zotero, with a link to the video)

[2] every time a reference is made to … it applies to the user's Zotero collection.

c. Searching/querying in Web Pages, from user's knowledge base only (item needs to be created in Zotero, with a link to the webpage)

2. Filtering by time of publication/creation
3. Filtering by Zotero tags (AND/OR)
4. Filtering by ABS journal rating
5. A reference (a link) to a pdf chunk the LLM's response is based on is provided as well as a full chunk.
6. Display of ZOTER ID – unique coding of items in Zotero
7. A link to the original source (a doi link) if available.
8. A link to the source (article) detailed info and data extraction summary is available for each source.
9. Selection of different analytic prompting scenarios ( 4 scenarios currently available) and detailed scenario description.
10. Selection of LLM models (3 models currently available)
11. Selection of sources from the user's starred collection
12. Selection of the number of incoming chunks of information to be processed by the tool and LLM.
13. Possibility to attach a file to the prompt straight in the query field, including pictures and automated selection of picture processing LLM model.
14. Queries based on LLM's own knowledge base (same as asking ChatGPT).
15. Evaluating LLM responses to improve reasoning (selected advanced prompting scenario).
16. Fact (hallucination) checking responses using text similarity statistics (Jaccard coefficient, Euclidian distance, Cosine similarity).
17. Fact (hallucination) checking responses using LLM reversed sequential check.
18. Reviewing LLM reasoning and steps behind the responses for advanced (React) scenarios.
19. Saving prompts and responses in the UGC tab.
20. Copying prompts and responses from the UGC tab.
21. Display parameters (limiters and filters) used for question/prompt generation.
22. Multiple chats.

**Quick relevance check Tab**

1. Step-by-step process to conduct a quick relevance check of selected ZOTERO collection – no PDFs needed, but TITLE AND ABSTRACTS are required
2. Pre-written prompt
3. Export to Excel with relevance decision.
4. Creation of RIS for re-export to Zotero.
5. Creation of statistic graphs – journal and authors.

**UGC tab**

1. Saving external information to be included in the prompts
2. Creating digital persona
3. Saving prompts
4. Saving responses
5. Creating folder system

**Data Extraction Table Tab (under development)**

Articles from which data was extracted ( advanced extraction or user-defined extraction methods) will be listed here and available for search and further upload to Excel.

**FAQs**