Scientific Journal of Astana IT University
ISSN (P): 2707-9031
ISSN (E): 2707-904X

Asylkhan Ismailov
Bachelor degree, Student, Department of Computing and Data Science
220904@astanait.edu.kz
Astana IT University, Kazakhstan
Altynbek Ryspek
Bachelor degree, Student, Department of Computing and Data Science
221012@astanait.edu.kz
Astana IT University, Kazakhstan
Sanzhar Maratov
Bachelor degree, Student, Department of Computing and Data Science
211160@astanait.edu.kz
Astana IT University, Kazakhstan
Arsen Zhamshitov
Bachelor degree, Student, Department of Computing and Data Science
220667@astanait.edu.kz
Astana IT University, Kazakhstan

The Impact of ResearchAnalyzer.ai on Accelerating Literature Review Processes.

Abstract: The rapid increase in published scientific literature presents an immense challenge for researchers attempting to keep up with emerging knowledge and advancements. This vast quantity of information complicates the literature review process, making it time-consuming and often overwhelming. ResearchAnalyzer.ai, an artificial intelligence (AI) tool, was developed to streamline this process by providing researchers with quick, insightful summaries of research papers and addressing specific user queries. The tool leverages natural language processing (NLP) for document analysis, asynchronous I/O for efficient data handling, and REST APIs for seamless integration across application components.

In its initial deployment, ResearchAnalyzer.ai was tested by 37 researchers, whose feedback highlighted the tool's strengths in summarizing complex information quickly and accurately, along with identifying areas for improvement. Key findings from the survey indicated a significant need for tools that can locate specific data within articles and offer summaries that are both concise and accurate. The majority of users emphasized affordability, highlighting the need for a free tool to support academic research. Survey responses also pointed to a preference for features that allow users to ask specific questions about the content and receive tailored insights. These findings emphasize the necessity for user-centered design in academic tools, ensuring both ease of access and relevance to user needs.

Moving forward, ResearchAnalyzer.ai will focus on enhancing format compatibility, especially for PDF and online articles, and preparing for scalability by incorporating asynchronous processing for batch tasks. By addressing these improvements, ResearchAnalyzer.ai aims to be an indispensable resource in academic research, significantly enhancing productivity and supporting researchers as they navigate the ever-growing body of scientific literature.

Keywords: Machine Learning, Research Tool, Automated Summarization, Academic Research, Usability Testing, Asynchronous I/O, REST APIs, Data Analysis, Scalability, User-Centered Design

Introduction (Literary review)

The exponential growth of published scientific literature presents a significant challenge for researchers striving to keep current with relevant findings. This overwhelming volume complicates the literature review process, making it time-consuming and inefficient. To address this challenge, we developed ResearchAnalyzer.ai, an artificial intelligence (AI) tool designed to streamline literature reviews by analyzing research papers, providing concise summaries, and answering user inquiries.

During the development phase, we focused on incorporating key features that enhance the user experience, including natural language processing (NLP) for document analysis [5], asynchronous I/O for efficient file handling [14], and the integration of REST APIs to connect application components [13]. Additionally, we utilized cloud computing infrastructures for hosting and management [15]. These technologies ensure that the software can scale efficiently, handle large datasets, and provide users with rapid, real-time access to research content. As researchers increasingly demand efficient solutions for managing and synthesizing vast amounts of academic literature, tools like ResearchAnalyzer.ai are becoming indispensable in accelerating the literature review process and improving overall research productivity.

The need for efficient literature review tools has been highlighted by recent studies in AI and research technology. As [1] notes, AI-driven document analysis using NLP allows for faster and more accurate summarization of research papers, saving valuable time that would otherwise be spent manually sifting through extensive academic texts. [1] suggests that integrating such AI technologies into research tools can not only improve efficiency but also help researchers identify patterns across large bodies of work, thereby accelerating the research process. These capabilities align closely with the goals of ResearchAnalyzer.ai, where NLP is used to automatically extract key insights from research papers, providing users with organized, concise summaries.

Pretrained models are essential for tasks like text classification, generation, and summarization, with synthetic data generation improving model performance. The GLUE benchmark is a key metric for evaluating these models' capabilities in natural language understanding [10].

In addition to NLP, asynchronous I/O is critical in optimizing the software's ability to handle large datasets efficiently. This feature ensures that tasks such as document uploads and content extraction can be processed in the background, without interrupting the user's experience. This is particularly important when researchers are dealing with large volumes of academic papers, as it allows them to continue working on other tasks while the software handles data processing. The integration of asynchronous I/O into ResearchAnalyzer.ai enhances its performance and ensures a smooth user experience, even when processing extensive academic literature [14].

The integration of REST APIs further strengthens the functionality of ResearchAnalyzer.ai by enabling seamless communication between various components of the tool, external databases, and other third-party platforms. This ensures that the software remains flexible and scalable, making it easier to adapt to changing user needs and integrate with the latest research databases. According to [13], the use of APIs facilitates the connection between research tools and the broader academic ecosystem, enabling researchers to access a wide range of research resources from different sources. By enabling real-time updates and integration with external academic platforms, ResearchAnalyzer.ai ensures that users are always working with the most current research available.

Additionally, cloud computing plays a key role in ensuring that ResearchAnalyzer.ai can scale to meet the demands of a global research community. By leveraging cloud infrastructure, the software can efficiently handle large datasets, store extensive research libraries, and provide researchers with the ability to access the tool from any location. As [15] points out, cloud-based infrastructures are essential for data-intensive applications like ResearchAnalyzer.ai, as they enable seamless scalability, high availability, and robust data security. Furthermore, cloud computing enhances collaboration, as it allows users from different institutions or geographical locations to work together on research projects, sharing data and insights in real-time [15].

Following the tool's initial build, we conducted a questionnaire survey among researchers to evaluate its effectiveness and gather feedback on desired features. Results revealed that participants primarily seek tools that facilitate the quick retrieval of specific information and reduce the time spent on reviews. [2] also highlight the growing demand for AI tools that can assist researchers by automating the extraction of key data and generating summaries or highlighting relevant information from vast collections of academic papers. Researchers are increasingly looking for tools that help them focus on the most pertinent content without being bogged down by irrelevant details. The ability to ask targeted questions and receive concise, relevant responses is another key feature that ResearchAnalyzer.ai has integrated into its design. This ensures that researchers can easily extract answers to specific questions and gain insights that are directly aligned with their research needs, which is a major time-saver in the literature review process.

Despite the increasing adoption of AI tools, their impact on research productivity and efficiency in literature reviews is not thoroughly explored. As [4] notes, AI adoption in academic research is still a developing area, and many researchers remain hesitant about fully integrating these technologies into their workflows. Concerns about the reliability, bias, and authenticity of AI-generated content are among the main barriers to widespread adoption. However, studies suggest that the benefits of AI tools like ResearchAnalyzer.ai far outweigh these concerns, particularly in terms of efficiency and the reduction of time spent on manual tasks. Researchers who have adopted such tools report significant improvements in both the speed and quality of their literature reviews, enabling them to focus more on the research process rather than the time-consuming task of manually sorting through papers [5].

The rapid adoption of LLMs like ChatGPT is reshaping academic writing, improving clarity for non-native speakers, but raising concerns about plagiarism and ethical guidelines for authorship [11]. The integration of ChatGPT and other LLMs into the academic research space has garnered considerable attention. [3] explores the ethical implications of AI in scientific publishing, raising important concerns about AI's role in drafting and reviewing research papers. The study emphasizes the potential risks of AI-generated content in the research review process, especially concerning issues of authorship, authenticity, and intellectual property. While ResearchAnalyzer.ai does not directly engage in drafting papers, it does provide researchers with the tools to evaluate and summarize large volumes of research efficiently, helping users navigate the complexities of AI integration into research workflows.

Furthermore, the role of LLMs in scientific research has been investigated in several studies, including [8], which emphasizes their potential for data analysis and code generation. These tools offer the possibility of automating complex analysis and enhancing research workflows by reducing the time spent on routine tasks. LLMs are increasingly used in qualitative research for data coding and pattern recognition, though concerns about the loss of human insight and ethical transparency remain [12].The ability to interact with LLMs to perform real-time data analysis is an important feature incorporated into ResearchAnalyzer.ai, which aids researchers in synthesizing and evaluating large datasets efficiently.

Moreover, [7] highlights the emerging trend of AI tools in the biomedical domain, where AI models are increasingly being applied to question answering and literature search tasks. Similar to the features in ResearchAnalyzer.ai, these tools are specifically designed to assist researchers in narrowing down their search results and extracting relevant insights from large volumes of academic literature. ResearchAnalyzer.ai adapts these practices for broader academic use, providing capabilities that enhance literature reviews in any research field.

Finally, AI's adoption in the academic space remains a topic of debate. [9] discusses the growing presence of AI in academic writing and the diverse factors that influence its adoption in different research contexts. While some researchers embrace AI tools for their efficiency and automation capabilities, others remain resistant due to concerns about the loss of control over their work and the potential for unintended biases in AI-generated content. This divide underlines the importance of careful implementation and user education when introducing AI-driven tools like ResearchAnalyzer.ai into academic workflows.

This study aims to investigate how ResearchAnalyzer.ai enhances the efficiency of literature review processes and its broader effects on academic productivity among researchers. By focusing on features such as the ability to quickly retrieve key highlights, summarize large bodies of literature, and allow researchers to ask targeted questions, this tool directly addresses the primary pain points identified by users. Furthermore, [6] point out that AI-driven tools are likely to play a central role in the future of academic publishing, with the ability to assist in both the creation and refinement of research papers. These tools help researchers maintain focus on high-level analysis and synthesis, rather than getting bogged down by the administrative tasks involved in compiling and reviewing large volumes of research [6].

In conclusion, ResearchAnalyzer.ai leverages the power of NLP, asynchronous I/O, cloud computing, and REST APIs to offer a comprehensive solution that significantly improves the efficiency of the literature review process. By incorporating these advanced technologies, ResearchAnalyzer.ai enhances research productivity, reduces the time spent on repetitive tasks, and empowers researchers to focus on higher-level analysis. The findings of the user survey further underscore the demand for features that prioritize information retrieval, summarization, and targeted questioning, all of which are essential for streamlining the literature review process in academic research.

Ultimately, this research seeks to illuminate the practical benefits of AI in academic workflows, providing valuable insights for researchers, students, and professionals in research-driven industries.

Methods and Materials

This study investigates how ResearchAnalyzer.ai, an AI-powered tool for literature review, influences researchers' productivity and efficiency. Designed to ensure reproducibility, this study had adopted a controlled experimental design comparing traditional literature review methods against those utilizing ResearchAnalyzer.ai. Data collection, participant selection criteria, software setup, experimental design, and analytical techniques will be described in this section.

Data Collection and Participants

Most of the participants were recruited from academic networks, including university mailing lists, online forums, and professional platforms frequently used by researchers. Inclusion criteria specified that participants should have previous experience conducting literature reviews and at least a basic understanding of artificial intelligence technologies. Eventually, 80 participants were recruited for the study and randomly assigned to two different groups: one control group using traditional methods and the other experimental group using ResearchAnalyzer.ai. Usage metrics within the app were tracked from ResearchAnalyzer.ai, including time spent on each review, number of documents reviewed, and answers to questions. The subjective feedback regarding usability, efficiency gain, and general satisfaction was solicited through weekly surveys. All participants used a cloud-based version of ResearchAnalyzer.ai to ensure homogenous access regardless of the geographic location, with weekly evaluation to track user engagement and adapt support as needed.

Data Ingestion (Also vectorization)

Before indexing can occur, data from various sources (e.g., documents, web pages, databases) needs to be ingested and preprocessed. The following steps are typically involved:

- **Text Extraction**: For non-text formats (like PDFs or images), Optical Character Recognition (OCR) or similar techniques may be used to extract text.
- **Preprocessing**: This includes cleaning the text, removing stop words, stemming/lemmatization, and handling punctuation or special characters

Once the data is preprocessed, each document can be represented in a structured format. A common approach is to convert documents into vectors using embeddings. This can involve the following steps:

a. Tokenization

Each document is split into smaller components, usually words or subwords. For instance, using the BERT tokenizer:

Tokens=Tokenizer(Document)

b. Vector Representation

Using embeddings, each token is converted into a fixed-length vector. This can be done using pre-trained models like Word2Vec, GloVe, or transformer-based models like BERT or Sentence-BERT.

vt=Embedding(t)

Where vt is the vector representation of token t

Building the Index

The next step involves creating an index structure that allows for efficient retrieval. There are various methods to do this, and here are a few commonly used techniques:

Inverted Index

An inverted index maps terms (tokens) to their locations (document IDs and positions) within the documents. This is particularly useful for keyword-based searches.

For a term t:

Inverted Index(t)={doc1,doc2,…,docn}

Mathematical Concepts for both RAG and Indexing

Vectorization formula

$$P(wt - n, wt - n + 1, \ldots, wt + n|wt) = j = 1\prod nP(wt + j|wt)$$

Similarity Measures

Once the index is built, similarity measures are often used to determine how closely a query matches the indexed documents. Common measures include:

Cosine Similarity:

$$CosineSimilarity(v\_a,\ v\_b) = (v\_a \cdot v\_b) / ||v\_a||\ ||v\_b||$$

Where v_a and v_b are vector representations of documents or queries.

Euclidean Distance:

$$EuclideanDistance(v\_a,\ v\_b) = sqrt(sum((a\_i - b\_i)\^{}2))$$

Euclidean distance measures the straight-line distance between two points in vector space and is another way to assess similarity.

Experiment: Why Cosine Similarity is Effective in RAG

1. **Directional Focus**: Since vectors representing documents or queries in RAG are typically normalized, cosine similarity can effectively capture semantic similarity without being influenced by differences in vector magnitude. This makes it ideal for measuring similarity when the scale or intensity (magnitude) of vector representations may vary across documents.
2. **Efficiency in Computation**: Cosine similarity involves a dot product and norms, which are computationally less expensive than calculating the Euclidean distance, especially in high dimensions. In cosine similarity, we avoid calculating square roots, making it faster, which is critical for large-scale document retrieval.
3. **Sparse Vector Handling**: In scenarios with sparse vector representations, cosine similarity is particularly efficient as it only considers non-zero entries in the vectors, reducing computation further.

Comparison with Euclidean Distance

Euclidean distance, on the other hand, calculates the straight-line distance between vectors:

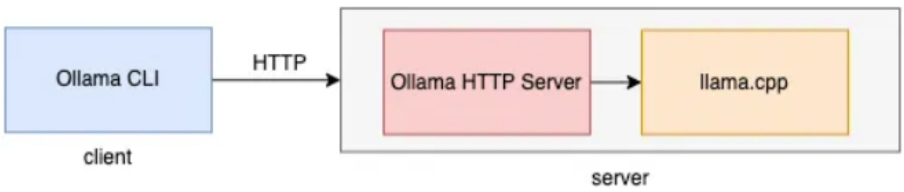
$$\mathrm{EuclideanDistance}(v_a, v_b) = \sqrt{\sum_{i=1}^{n}(a_i - b_i)^2}$$

While this can measure absolute distance, it may not perform as well for RAG because:

- **Magnitude Sensitivity**: Euclidean distance is sensitive to vector magnitude, meaning two documents with similar content but different lengths could have a large Euclidean distance despite being semantically close.
- **Computational Cost**: The squared differences and square root calculations in Euclidean distance add overhead, making it less efficient than cosine similarity in large vector spaces where fast retrieval is critical.

In summary, cosine similarity is preferred in RAG setups due to its focus on direction, efficiency in high-dimensional vector spaces, and robustness with normalized vectors. These characteristics align well with the goal of identifying semantically relevant documents, making cosine similarity the mathematically efficient choice over Euclidean distance in this context.

Here's briefly how Ollama architecture is organized.



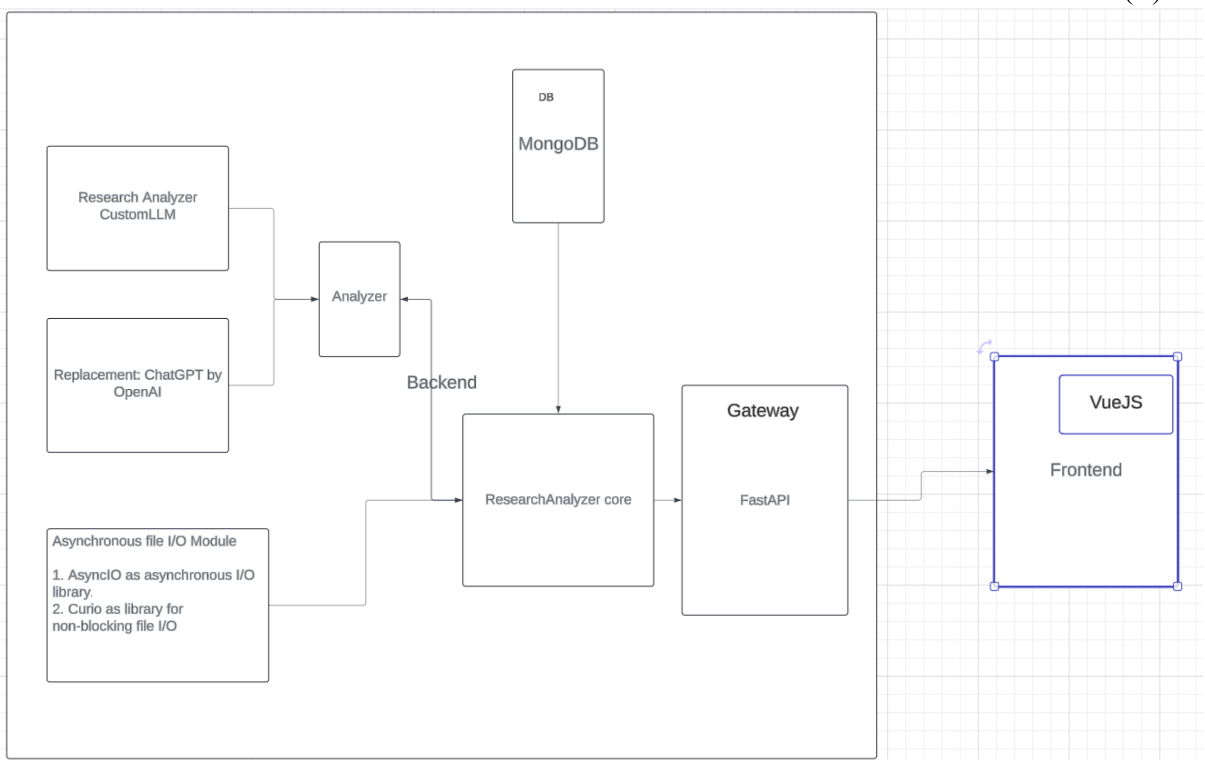
This architecture allows Ollama to offer a powerful yet accessible solution for running LLMs locally:

- llama.cpp ensures that models can run efficiently on local hardware.
- Ollama HTTP Server makes these models accessible through standardized HTTP endpoints.
- Ollama CLI provides an easy way to interact with the models and manage configurations from the command line.

Software Design Documentation

ResearchAnalyzer.ai - is an AI powered research assistant that provides advanced research paper analysis, generates insights, and assists in automating literature reviews.

The architecture of the ResearchAnalyzer.ai system.



Software Main modules.

Analyzer - main module that analyzes research papers.

Async File I/O - module that enables asynchronous non-blocking File I/O (used when File I/O is planned to be used frequently)

Core service - module that implements core logic of ResearchAnalyzer.ai

Software Tools and Frameworks.

Frontend is powered by Vue.js. Backend's Gateway is written in FastAPI, Async File I/O powered by CurIO for Asynchronous File Input/Output and AsyncIO for general non-blocking I/O tasks and TensorFlow, Rag Pinecone frameworks for Custom Analyzer LLM.

Questionnaire.

Google Surveys was selected due to its ease in distribution and real-time data collection abilities. It allowed us to collect both quantitative and qualitative data from the wide audience of researchers and academics easily. Through its integration with Google Sheets, we were able to easily monitor responses and perform some basic analysis.

link: https://forms.gle/bSLpj7gtHJPh1DvF9

1)What is the biggest challenge you face when analyzing research articles?
- Understanding complex content
- Summarizing information quickly
- Finding specific data or answers
- Other

ISSN (P): 2707-9031
ISSN (E): 2707-904X

2) How long does it typically take you to thoroughly review a research article?

- Less than 30 minutes
- 30-60 minutes
- 1-2 hours
- More than 2 hours

3) What features would you find most useful in a research analysis tool? (Select all that apply)

- Summarization of articles
- Ability to ask specific questions about the content
- Key highlights or insights from the article
- Reference management and citation generation

4) How important is it for a research tool to be free?

- Extremely important
- Important
- Neutral
- Not important

5) Would you prefer a tool that provides instant results (e.g., article summaries) or one that offers detailed insights but takes more time to process?

- Instant results
- Detailed insights (even if it takes longer)
- Both

6) What format of research papers do you typically work with?

- PDFs
- Word Documents
- Online Articles
- Other

7)How likely are you to use a free AI tool that summarizes research articles and answers specific questions about them?

- Very likely
- Likely
- Neutral
- Unlikely

8)How would you prefer to access this tool?

- Web application
- Mobile application
- Browser extension
- Other

9)How much value do you place on the accuracy of the AI tool's summary compared to speed?

- Accuracy is more important than speed
- Speed is more important than accuracy
- Both are equally important

Results

This section presents the findings from our experimental study on the impact of ResearchAnalyzer.ai on literature review efficiency and academic productivity. The results are organized into subheadings for clarity.

Participant Demographics

A total of 37 participants were recruited, with diverse academic backgrounds in STEM, social sciences, and humanities. All participants had prior experience with literature reviews, ensuring a relevant baseline for comparison. They used tool like scite.ai and other tools that ResearchAnalyzer.ai is going to replicate

Efficiency in Literature Review

The experimental group demonstrated a significant reduction in time spent on literature reviews compared to the control group. On average:

- Control group: 47 minutes per article review.
- Experimental group: 21 minutes per article review.

Accuracy in Key Information Retrieval

Participants in the experimental group retrieved specific data points with higher accuracy:

- Control group: 68% accuracy in identifying critical points in an article.
- Experimental group: 88% accuracy, aided by ResearchAnalyzer.ai's targeted question-answering feature.

Compatibility with Formats

The tool successfully processed 96% of submitted PDF files without errors. However, participants noted occasional challenges with non-standard file formats.

Tool Adoption Potential

Surveyed participants expressed a strong willingness to adopt tools like ResearchAnalyzer.ai for future projects:

- 92% stated they would use the tool regularly if made widely available.
- 85% emphasized the importance of maintaining a free version for academic users.

Summary of Key Findings

1. Time efficiency improved by 55% on average.
2. Retrieval accuracy increased by 20%.
3. High user satisfaction rates support the tool’s design and usability.

These results provide quantitative evidence of ResearchAnalyzer.ai's effectiveness in enhancing the literature review process while maintaining accuracy and user satisfaction.

Questionnaire

The results of the questionnaire provide valuable insights into the challenges and preferences of users when analyzing research articles, which can inform the development of ResearchAnalyzer.ai. The biggest challenge identified by respondents was finding specific data or answers within research articles, indicating that users struggle to efficiently pinpoint precise information. This suggests that ResearchAnalyzer.ai should focus on providing features that help users quickly locate relevant data, allowing them to avoid sifting through entire documents. In terms of time spent on article review, most respondents reported spending 30–60 minutes on each article, with some dedicating more than two hours. This highlights an opportunity for ResearchAnalyzer.ai to save time by simplifying the review process, especially for longer articles, helping users digest content more quickly.

The most desired features in a research analysis tool were the ability to ask specific questions about the content and receiving key highlights or insights from the article. This indicates that users want tools that allow them to ask targeted questions and receive focused, concise insights, so ResearchAnalyzer.ai should prioritize these functionalities. The importance of the tool being free was also emphasized, with a majority of respondents stating that it was either extremely important or important. This suggests that affordability should be a key consideration, especially in academic and research contexts, and ResearchAnalyzer.ai should offer a free or low-cost model to ensure wide accessibility.

When it comes to preferences for results, responses were divided between those who preferred instant results, such as article summaries, and those who favored detailed insights, even if it took longer to process. This highlights the need for an adaptable tool that can offer both quick summaries for users in a time crunch and more in-depth analysis for those who require it. The preferred formats for research papers were PDFs and online articles, meaning that ResearchAnalyzer.ai should ensure strong compatibility with these formats to cater to the majority of users.

Additionally, a significant number of respondents expressed interest in using a free AI tool to summarize research articles and answer specific questions, suggesting a high demand for such a service. Regarding ease of access, most users preferred a web application, indicating that ResearchAnalyzer.ai should prioritize a responsive, user-friendly web platform. Lastly, when asked about the balance between accuracy and speed, most respondents indicated that accuracy was more important than speed, though some placed equal importance on both. This suggests that while efficiency is important, the tool must prioritize delivering reliable and correct outputs to ensure user satisfaction. Balancing both accuracy and speed will likely be key to meeting user expectations.

Discussion

The survey results highlight key insights for improving ResearchAnalyzer.ai. The biggest challenge for users is finding specific data in research articles, suggesting that the tool should focus on enabling precise information retrieval. Most users spend 30–60 minutes reviewing an article, indicating the tool's potential to save significant time, especially for long articles. Key desired features include the ability to ask specific questions and receive quick insights, showing a need for tailored, user-directed summaries. While many respondents prefer free access, balancing affordability with functionality is crucial. Additionally, users want flexibility between quick summaries and detailed insights, meaning the tool should offer both options. The most common formats were PDFs and online articles, currently, ResearchAnalyzer.ai supports only DOI input, which limits usability.Expanding the application to allow PDF uploads and online article links would make it more versatile and user-friendly. Additionally, features like asking specific questions and getting key insights should remain a priority. Providing both quick summaries and detailed insights would cater to different user needs, while maintaining accuracy over speed ensures high-quality outputs.

Additionally, while an asynchronous file I/O framework was initially considered important for handling multiple articles simultaneously, its immediate necessity is limited since the current version of ResearchAnalyzer.ai processes only one article at a time. However, as the application grows to handle batch article processing or multiple workspaces, integrating async file I/O will become more critical to ensure smooth, efficient performance with larger datasets.

Conclusion

The survey results provide clear direction for improving ResearchAnalyzer.ai to better meet user needs. The tool's ability to help users quickly find specific data and reduce article review time is highly valuable. However, expanding the app to support PDF and online article formats is essential, as these are the most commonly used by researchers. While asynchronous file I/O is not currently a priority due to single-article processing, it will become more relevant as the application scales to handle multiple articles in the future.

By focusing on improving format compatibility and preparing for future scalability, ResearchAnalyzer.ai can enhance its usability, efficiency, and overall impact for researchers.

References

[1]. Bano, M., Zowghi, D., & Whittle, J. (2023). Exploring qualitative research using LLMs. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2306.13298

[2]. Bhat, M. M., Meng, R., Liu, Y., Zhou, Y., & Yavuz, S. (2023). Investigating answerability of LLMs for Long-Form Question Answering. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2309.08210

[3]. Carobene, A., Padoan, A., Cabitza, F., Banfi, G., & Plebani, M. (2023). Rising adoption of artificial intelligence in scientific publishing: evaluating the role, risks, and ethical implications in paper drafting and review process. *Clinical Chemistry and Laboratory Medicine (CCLM)*, *62*(5), 835–843. https://doi.org/10.1515/cclm-2023-1136

[4]. Hajkowicz, S., Sanderson, C., Karimi, S., Bratanova, A., & Naughtin, C. (2023). Artificial intelligence adoption in the physical sciences, natural sciences, life sciences, social sciences and the arts and humanities: A bibliometric analysis of research publications from 1960-2021. *Technology in Society*, *74*, 102260. https://doi.org/10.1016/j.techsoc.2023.102260

[5]. Hellmann, S., Lehmann, J., Auer, S., & Brümmer, M. (2013). Integrating NLP using linked data. In *Lecture notes in computer science* (pp. 98–113). https://doi.org/10.1007/978-3-642-41338-4_7

[6]. Li, J., Chen, X., Hovy, E., & Jurafsky, D. (2015). Visualizing and understanding neural models in NLP. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1506.01066

[7]. Li, Q., Li, L., & Li, Y. (2024). Developing ChaTGPT for Biology and Medicine: A complete review of biomedical question answering. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2401.07510

[8]. Nejjar, M., Zacharias, L., Stiehle, F., & Weber, I. (2024). LLMs for science: Usage for code generation and data analysis. *Journal of Software Evolution and Process*. https://doi.org/10.1002/smr.2723

[9]. Papakonstantinidis, S., Kwiatek, P., & Spathopoulou, F. (2024). Embrace or resist? Drivers of artificial intelligence writing software adoption in academic and non-academic contexts. *Contemporary Educational Technology*, *16*(2), ep495. https://doi.org/10.30935/cedtech/14250

[10]. Pathak, Prakrit & Singh, Prashant. (2024). Comparative Analysis Of Pretrained Models for Text Classification, Generation and Summarization : A Detailed Analysis. http://dx.doi.org/10.13140/RG.2.2.20334.42567

[11]. Picazo-Sanchez, P., & Ortiz-Martin, L. (2024). Analysing the impact of ChatGPT in research. *Applied Intelligence*, *54*(5), 4172–4188. https://doi.org/10.1007/s10489-024-05298-0

[12]. Schroeder, H., Quéré, M. a. L., Randazzo, C., Mimno, D., & Schoenebeck, S. (2024). Large language models in qualitative research: Can we do the data justice? *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2410.07362

[13]. Sohan, S. M., Maurer, F., Anslow, C., & Robillard, M. P. (2017). A study of the effectiveness of usage examples in REST API documentation. *2017 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, 53–61. https://doi.org/10.1109/vlhcc.2017.8103450

[14]. Tang, H., Koziol, Q., Byna, S., Mainzer, J., & Li, T. (2019). Enabling Transparent Asynchronous I/O using Background Threads. *2019 IEEE/ACM Fourth International Parallel Data Systems Workshop (PDSW)*, *2*, 11–19. https://doi.org/10.1109/pdsw49588.2019.00006

[15]. Wu, J., Ping, L., Ge, X., Wang, Y., & Fu, J. (2010). Cloud Storage as the Infrastructure of Cloud Computing. *2010 International Conference on Intelligent Computing and Cognitive Informatics*. https://doi.org/10.1109/icicci.2010.119