

# **NASApalooza: Revolutionizing Research for NASA with LLM's**

Aman Ulla, Srinivas Alva

## **Introduction:**

Our innovative AI-driven research assistant is tailored to provide researchers with powerful and efficient resources for effectively exploring the constantly evolving realm of scientific literature. Researchers often face the daunting task of staying up-to-date with emerging breakthroughs across diverse fields. This challenge can be overwhelming as they sift through numerous papers, each filled with distinct terminology and formatting conventions.

Conducting a comprehensive literature review and analysis becomes particularly challenging when dealing with a large volume of papers. Identifying seminal papers within a specific domain is equally arduous. Drawing meaningful conclusions from this extensive ocean of papers and identifying research gaps is a time-consuming process that demands creativity and expertise.

As the NASApalooza team, we address these challenges with a multifaceted approach. At its core, we've developed Palooza, an interactive web application that serves as the researcher's command center. This intuitive platform provides researchers with quick access to comprehensive insights and summaries of research papers relevant to their chosen domains. By utilizing Palooza, researchers can efficiently identify whitespace in research, draw conclusions from a vast ocean of papers, and explore research topics in less time.

Our platform also enables researchers to effortlessly navigate mind maps connecting related research topics and delve into nested citations, providing a deeper understanding of the interconnectedness of ideas. Additionally, we offer Q&A sessions on papers and quick analysis of datasets, further enhancing the research experience.

## **Problem Statement:**

Researchers face significant challenges in maintaining up-to-date knowledge of emerging scientific and technological advancements and applying them to their missions. These challenges stem from the need to comprehend extensive research literature, both within and outside their areas of expertise. The current workflow involves running multiple queries on public and commercial databases, resulting in the retrieval of hundreds of papers. Researchers must then assess the relevance of these papers based on their titles, abstracts, and metadata.

To synthesize relevant information, researchers are required to read or at least walkthrough hundreds of papers, navigating diverse reporting standards and disciplinary jargon. Consequently, a substantial amount of time and effort is invested in reviewing information that often proves irrelevant or non-useful. This inefficiency hinders their ability to harness the vast and continually expanding body of scientific literature.

In summary, the key pain points for researchers include:

1. **Information Overload:** Researchers are bombarded with an excessive amount of literature from various sources, making it challenging to keep pace with the latest developments.
2. **Time-Consuming Review:** The process of reviewing and assessing the relevance of numerous papers is time-consuming, diverting valuable time away from research and analysis.
3. **Interdisciplinary Challenges:** Researchers encounter difficulties in deciphering different reporting standards and disciplinary terminology, particularly when exploring literature outside their areas of expertise.
4. **Relevance Filtering:** Researchers often spend substantial effort on papers that do not contribute to their research goals.
5. **Efficiency Gap:** There is a need for more powerful tools to streamline the literature review process, enabling researchers to efficiently identify and leverage relevant knowledge while potentially discovering novel insights within the ever-expanding scientific literature.

### Proposed Approach:

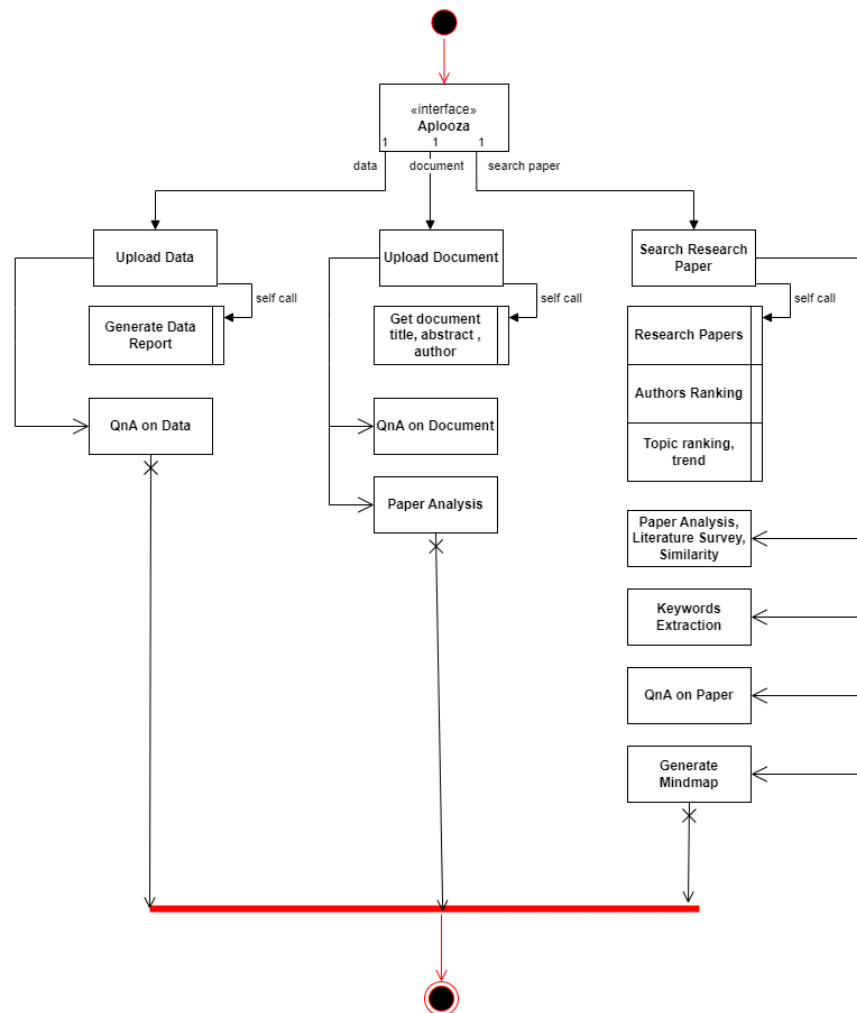
To address the challenges faced by researchers in managing and extracting insights from the vast scientific literature, we propose the development of Palooza, an interactive web application with a comprehensive set of features designed to streamline the research process and enhance efficiency. Palooza's features are specifically tailored to cater to the pain points outlined in the problem statement.

1. **Paper Retrieval:** Palooza will integrate with multiple data source engines, including Arxiv, Google Scholar, and Pubmed, allowing researchers to retrieve relevant papers effortlessly by providing keywords as input.
2. **Paper Analysis:** Researchers can leverage Palooza for in-depth paper analysis, including summarization, dataset identification, conclusion extraction, and assessment of paper quality. The platform will also assist in identifying research gaps within papers. Furthermore, Q&A sessions can be conducted on specific papers to gain further insights.
3. **Questioning & Answering (Q&A):** Palooza is powered by LLM (Language Model), currently embedded with GPT-3.5, enabling researchers to ask questions related to their documents or specific topics. This functionality facilitates quick access to key information within the research literature.
4. **Document Analysis:** Researchers can upload their own documents, including private or unpublished papers, for analysis and Q&A. This feature significantly saves time by providing insightful information on uploaded papers.
5. **Dataset Analysis:** Palooza offers dataset analysis capabilities, generating overview reports for uploaded datasets. Researchers can also download analysis reports for further examination.
6. **Author Analysis:** Researchers can explore author analytics, including citation counts and index values, to identify influential authors in their fields of interest.
7. **Mindmap:** Palooza's mindmap feature enables researchers to navigate interconnected research topics, fostering a deeper understanding of the relationships between ideas and citations.
8. **Keyword Extraction and Annotation:** Researchers can use this feature to identify critical keywords and domain-specific entities within papers, facilitating specialized research in various fields.

9. **Literature Review:** Palooza compiles summaries of papers, streamlining the literature review process and providing a consolidated view of research in a given domain.
10. **Topic Trend Analysis:** Researchers can analyze the global trend of specific topics, gaining insights into how subjects evolve and their dominance over time. This feature aids in staying current with emerging trends.
11. **Relevant Searches:** Palooza offer a feature that suggests relevant searches based on the topics researchers are exploring. This function will help researchers discover related research areas and expand their understanding of the subject matter.
12. **Bookmark and Favorite List:** Palooza also allows researchers to add papers to their favorite list, functioning as a bookmark feature. This enables researchers to save and review papers of interest at a later time, enhancing organization and accessibility.

Palooza aims to revolutionize the research experience by offering a holistic platform that addresses the challenges researchers face in accessing, analyzing, and making sense of scientific literature. With its powerful features, Palooza empowers researchers to efficiently navigate the ever-expanding world of knowledge and unlock new avenues of exploration within their domains.

The following diagram represents the behavior of the Palooza application:



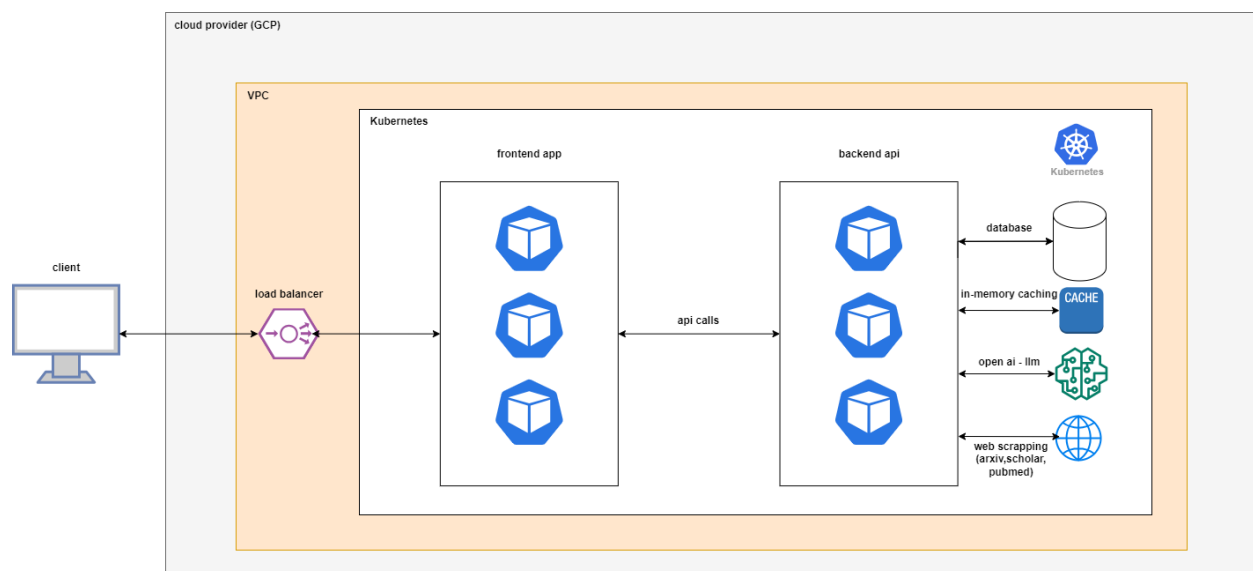
**Palooza Relevance for defined problem Statment:** Palooza directly addresses the pressing challenges faced by researchers in today's information-rich world, offering a solution that streamlines the literature review process, saving valuable time and resources while ensuring that only the most pertinent research is considered.

**Effectiveness of Palooza application for NASA researchers:** With its wide array of features, from automated paper retrieval to advanced paper analysis and Q&A capabilities, Palooza empowers researchers to efficiently and effectively extract meaningful insights from the vast scientific literature, ultimately enhancing their ability to leverage the latest advancements for their missions.

**Impact of Palooza Application on NASA Researchers:** By reducing information overload, facilitating interdisciplinary research, and offering trend analysis, Palooza has the potential to significantly impact the research landscape, fostering innovation, and accelerating the pace of scientific discovery across a multitude of domains.

### Deployment Features and Consideration for Palooza Application:

**Cloud Agnostic and easy deployment:** The Palooza application embodies a cloud-agnostic architecture that prioritizes adaptability and scalability across diverse cloud providers. To attain this goal, our initial step involved conceptualizing the application using a microservices-oriented methodology, dividing it into discrete, autonomous components or containers. We harnessed containerization tools such as Docker to package each service together with its requisite dependencies. Subsequently, we adopted container orchestration frameworks like Kubernetes, which provide abstraction layers seamlessly aligning with multiple cloud providers. This abstraction empowered us to consistently deploy and administer the application without being constrained by the underlying infrastructure.



**Can cater up to hundreds of thousands of users?** The deployment approach utilized plays a pivotal role in ensuring our application's readiness to cater to millions of users. By embracing a cloud-agnostic architecture founded on microservices and containerization, we've laid the groundwork for remarkable scalability. This strategy allows us to efficiently scale individual components of the application to meet increased demand, ensuring optimal performance and responsiveness as our user base expands.

**Cost to Run App:** Below is the cost breakdown to run this app at your own infrastructure.

Item	Pricing (per month)
SERP API Cost	\$250
Cloud (GKE) (includes CPU allocation time, memory allocation time, requests, Idle min-instance CPU allocation time, Idle min-instance memory allocation time) For 20 concurrent users at any given point	\$105
LLM (OpenAI) (0.004\$ - 1k tokens), considering 20 concurrent users querying (10k context per user)	\$25
<b>Total*</b> (Excluded domain hosting charges)	~ \$380

### Comparing Palooza to Other Available Applications

Features	Palooza	ELICIT	Connected Papers	ResearchRabbit
Multiple sources (such as arXiv, Pubmed, etc.)	✓	✓	✓	✓
Provides Relevant Searches	✓	✓	✓	✓
Paper Analysis	✓	✓	✗	✗
Document/Paper analysis with QnA powered by LLM's	✓	✗	✗	✗
Dataset Analysis with QnA functionality	✓	✗	✗	✗
Download Literature Review	✓	✗	✗	✗
Author Analysis	✓	✗	✗	✓
MindMap	✓	✗	✓	✗
Download Literature Review	✓	✗	✗	✗
Keyword Extraction and Annotation with Trend analysis	✓	✗	✗	✗

### Future Enhancement:

In addition to the existing features, we plan to implement several advanced enhancements in current Palooza application:

- **Complex and Domain-Specific Retrieval:** Enhancing search capabilities to support complex and domain-specific queries, ensuring precise paper retrieval.
- **Advanced Mind Maps:** Expanding the mind map feature to offer more sophisticated visualization and interconnectivity of research topics.
- **Universal Retrieval Engine Support:** Enabling Palooza to seamlessly access and retrieve papers from a wide range of data source engines, ensuring comprehensive coverage.

- **UI Enhancement:** Continuously improving the user interface for a more intuitive and user-friendly experience, prioritizing ease of navigation.
- **Diverse Data Source Integration:** Integrating additional data sources for comprehensive data analysis, allowing researchers to draw insights from a wider range of datasets.
- **Advanced Table Comparison View:** Offering an advanced table comparison view for researchers to analyze and analyse across multiple research papers efficiently.
- **Downloading Reports:** Facilitating the bulk downloading of analysis reports across multiple research papers, simplifying data aggregation.
- **Database Integration for User Login:** Implementing user login functionality and session management, enabling users to store recent searches and maintain a list of favorite papers for future reference.

These enhancements will make Palooza an even more powerful tool for researchers, providing a seamless and comprehensive research experience across diverse domains.

## References:

- <https://elicit.org/>
- <https://www.researchrabbit.ai/>
- <https://www.connectedpapers.com/>
- <https://custom-writing.org/blog/research-papers-analysis>
- <https://cloud.google.com/products/calculator>

## Contributors

### 1. Aman Ulla

- LinkedIn: <https://in.linkedin.com/in/connectaman>
- Email: [connectamanulla@gmail.com](mailto:connectamanulla@gmail.com)

### 2. Srinivas Alva

- LinkedIn: <https://in.linkedin.com/in/srinivas-alva>
- Email: [alva.srinu@gmail.com](mailto:alva.srinu@gmail.com)

**DRIVEN DATA**

