



Proxima: Enhanced Search Engine for Research Articles



Ceydanur Şen - Nurullah Uçan

Boğaziçi University - CMPE 492 Project - Advisor: H. Birkan Yılmaz

Motivation and Problem Definition

- Finding relevant academic papers is time-consuming and often overwhelming for researchers.
- Traditional keyword-based search engines frequently return many irrelevant or low-quality results.
- Existing academic search tools do not fully exploit modern semantic search and rich metadata.
- Researchers need a system that surfaces recent, high-quality papers tailored to their specific interests.

Objectives and Contributions

- Design and implement an enhanced academic search engine on top of the OpenAlex dataset.
- Combine multi-abstract input, semantic embeddings, and keyword-based retrieval to improve result relevance.
- Provide a web interface for exploring, saving, and managing topic-focused **paper recommendations**.
- Support long-term topic tracking through **email subscriptions** and user feedback collection.

Methodology

- The user submits keywords and/or one or more seed abstracts, optionally with a publication year range.
- The backend retrieves candidate papers from OpenAlex using keyword-based filtering and metadata constraints.
- We encode seed abstracts and candidate papers with a pretrained sentence embedding model (Alibaba-NLP/gte-large-en-v1.5 via Sentence-Transformers, 1024-D), and re-rank OpenAlex candidates by semantic similarity.
- The top-ranked papers are returned to the frontend, where users can inspect results, create subscriptions, and provide feedback.

[Discover the Most Relevant Research Articles Instantly](#)

Paste an abstract or enter keywords to explore top matches from millions of academic papers.

Keywords

- Channel Modeling
- Molecular communication

Type keywords and press Enter...

Abstracts

Nanotechnologies promise new solutions for several applications in the biomedical, industrial and military fields. At the nanoscale, a nanomachine is considered as the most basic functional unit which is able to perform very simple tasks. Communication among nanomachines will allow them to accomplish more complex functions in a distributed manner. In this paper, the state of the art in molecular electronics is reviewed to motivate the study of the Terahertz Band (0.1-10.0 THz) for electromagnetic (EM) communication among nano-devices. A new propagation model for EM communications in the Terahertz Band is developed

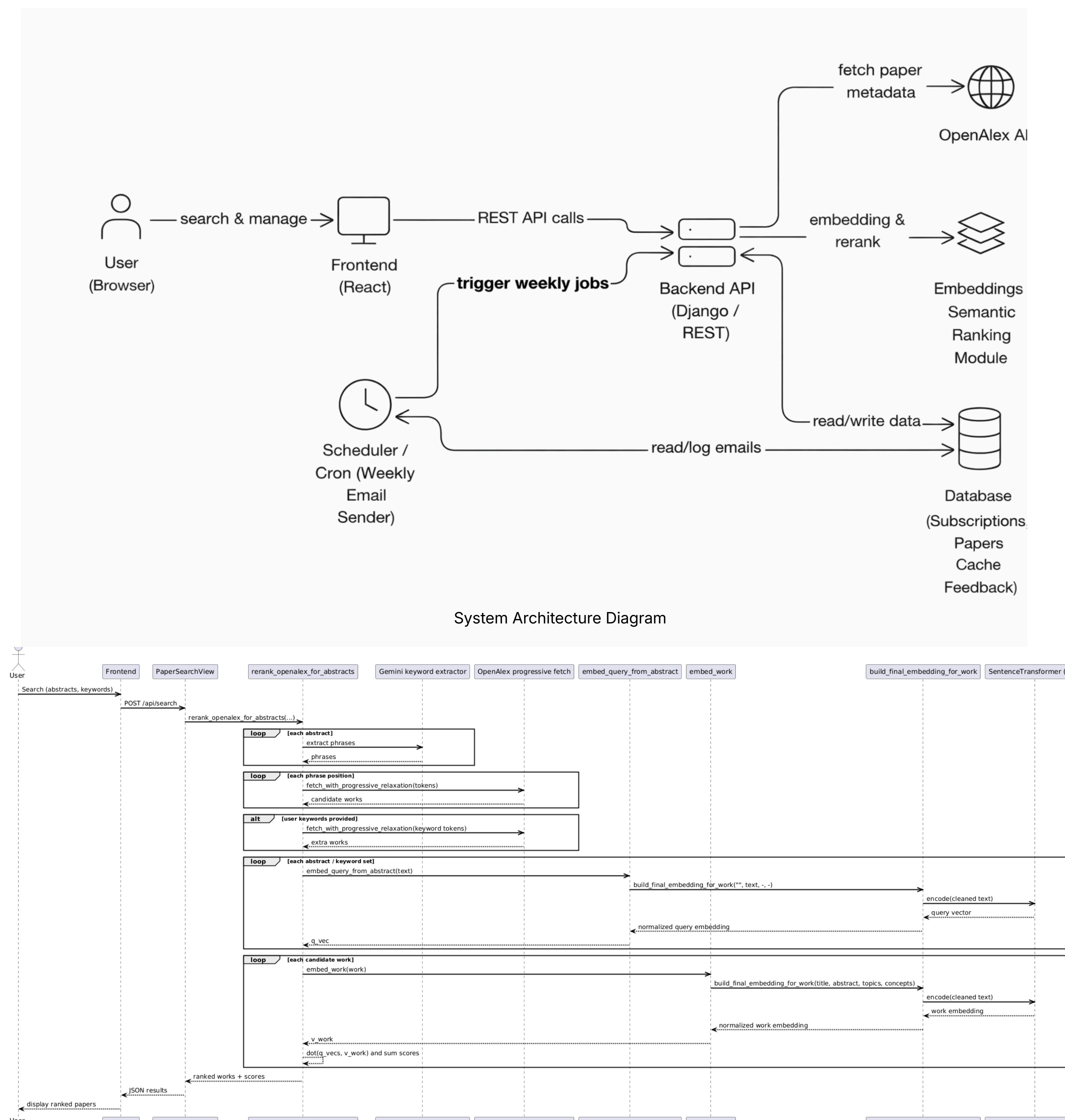
Molecular communication is a promising paradigm for nanoscale networks. The end-to-end (including the channel) models developed for classical wireless communication networks need to undergo a profound revision so that they can be applied for nanonetworks. Consequently, there is a need to develop new end-to-end (including the channel) models which can give new insights into the design of these nanoscale networks. The objective of this paper is to introduce a new physical end-to-end (including the channel) model for molecular communication. The new model is investigated by means of three modules, i.e., the

Year filter

min — max Clear

Search Research Papers

Home Page View of Proxima



Sequence Diagram of Embedding-based multi-abstract search pipeline

Experiments and Results

- We compare three search methods: embedding-based search, raw OpenAlex keyword-only search, and Gemini-augmented OpenAlex search.
- Users evaluate randomized result layouts, select their preferred list, and optionally rank all three pipelines.
- Feedback is stored anonymously (without paper IDs) and qualitatively shows clear differences between semantic and keyword-only retrieval, guiding future ranking and query-expansion improvements.



Conclusion

- This system combines multi-abstract input, semantic embeddings, and OpenAlex data API to make academic paper discovery more effective and personalized.

References

- J. Priem et al., OpenAlex: A Fully Open Index of Scholarly Works, 2022. <https://openalex.org>
- Alibaba-NLP, GTE-large-en-v1.5, Hugging Face, 2024. <https://huggingface.co/Alibaba-NLP/gte-large-en-v1.5>