AI-Assisted Patent Search & Analysis for Redox Flow Batteries

Authors: Georg Niess, Stefan Spirk, Roman Kern
Contact: georg.niess@tugraz.at

Abstract

Patent search is essential to identify and avoid patent infringement. By January 2025, there are more than 26,000 patents which contain the term redox-flow battery. The sheer volume makes it very challenging to identify relevant technical or legal information. Language barriers and the use of obscure or vague terminology further reduce the effectiveness of traditional keyword-based searches.

Our project: Over the past 18 months, we have collected information on more than 26.000 patents related to redox flow batteries and used this dataset to develop an AI-powered patent-search assistant built on large language models (LLMs).

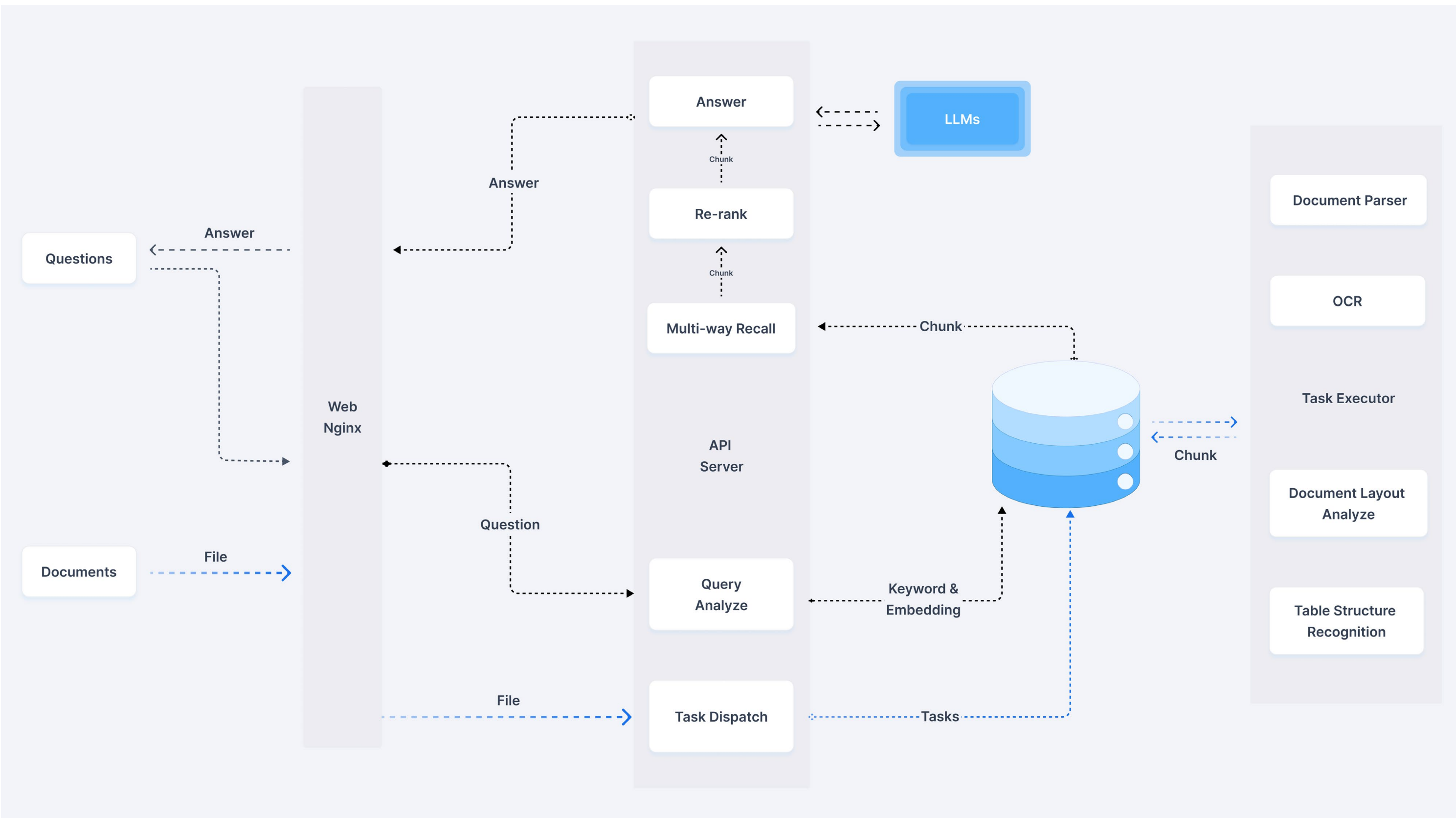
Key features:

1. **Embeddings-based Knowledge Base**: We embed all collected patents into a vector store, enabling semantic retrieval based on user queries.
2. **Retrieval-Augmented Generation (RAG):** Retrieved documents serve as context for the LLM, improving answer accuracy and minimizing hallucinations.
3. **Agentic Research Reports:** A series of specialized LLM agents can generate higher-level research reports (e.g., market trends) by leveraging the patent knowledge base.

Outcome: Our prototype shows promise in enhancing patent search and classification for redox flow battery R&D; early feedback indicates it helps project members more quickly identify relevant patents, gain insight into technology trends, and generate concise research summaries.

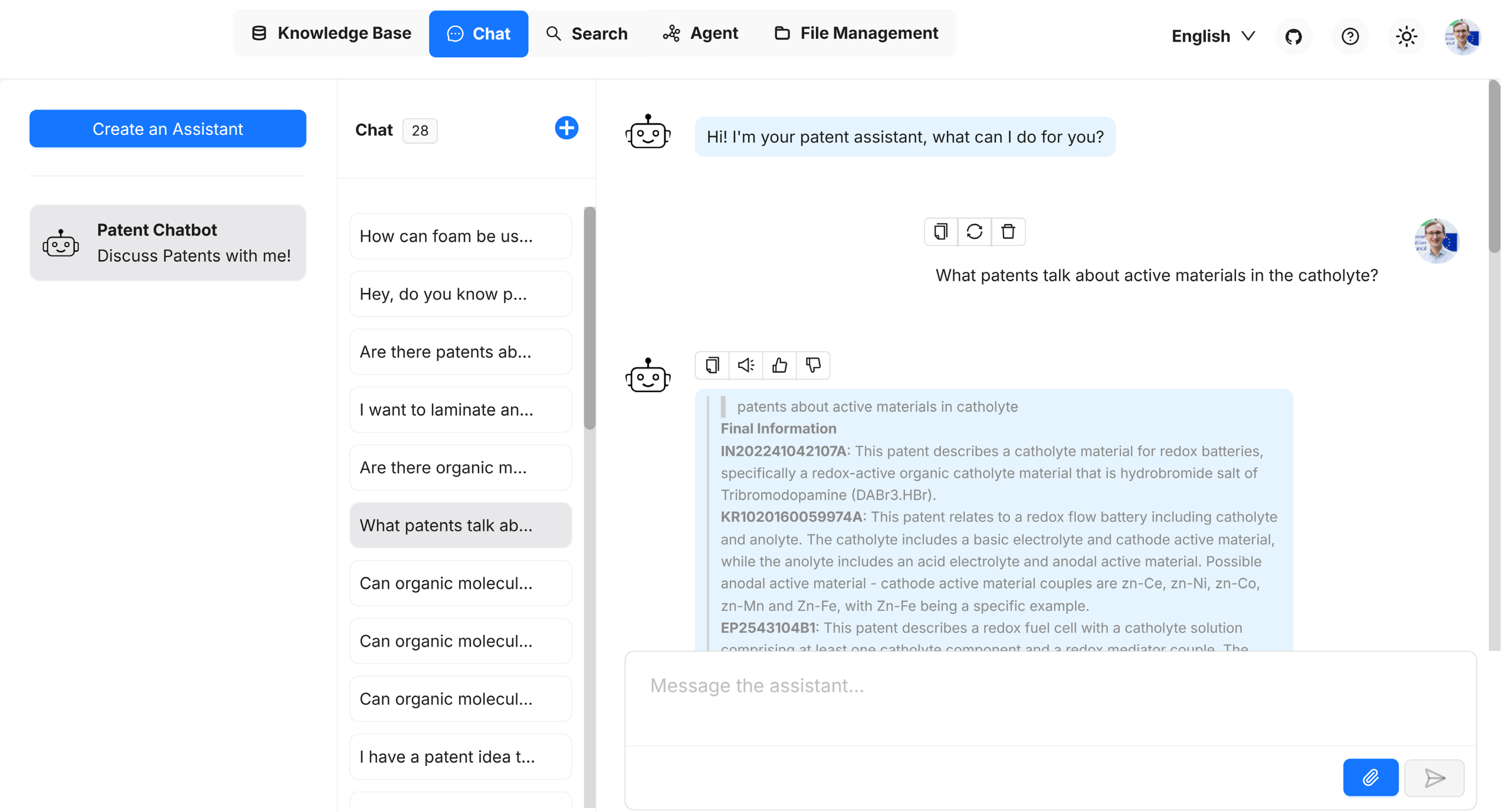
Architecture

In our system, user-submitted questions first pass through a reverse proxy (Web/Nginx) and are routed to the API server's "Query Analyze" component, which converts each question into both keyword queries and dense embeddings. These embeddings retrieve semantically relevant "chunks" from the central vector-indexed knowledge base, while traditional keyword queries can run in parallel for a hybrid recall approach. Retrieved chunks are then re-ranked, and the top passages are passed (with the original question) to a LLM to generate a precise answer, which is sent back through the proxy to the user. When a user uploads a document, Nginx forwards the file to the "Task Dispatch" component, which creates OCR, document parsing, layout analysis, and table-structure recognition jobs. The Task Executor processes each job, splits the document into smaller text or table chunks, and indexes those chunks (with both keyword indices and embeddings) into the same central knowledge base. This unified design allows the LLM to answer queries using both previously ingested and freshly uploaded files, ensuring that all information is semantically searchable and available for RAG-based retrieval.



System Architecture: End-to-end workflow illustrating how user questions and uploaded documents are handled by the API server, knowledge base, and task executor modules.

Examples



Hi! I'm your patent assistant, what can I do for you?

What patents talk about active materials in the catholyte?

patents about active materials in catholyte

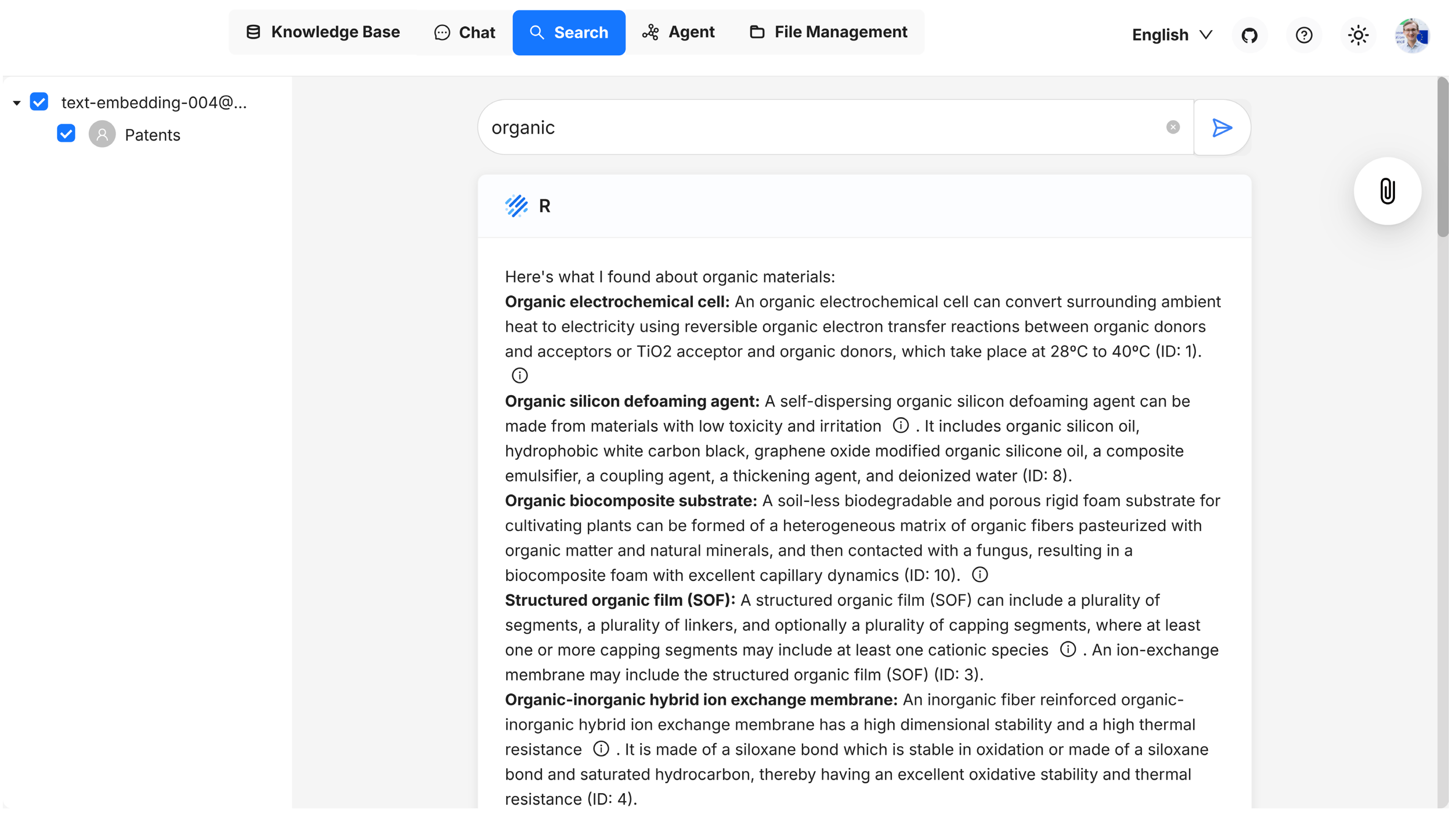
Final Information

IN202241042107A: This patent describes a catholyte material for redox batteries, specifically a redox-active organic catholyte material that is hydrobromide salt of Tribromodopamine (DABr3.HBr).

KR1020160059974A: This patent relates to a redox flow battery including catholyte and anolyte. The catholyte includes a basic electrolyte and cathode active material, while the anolyte includes an acid electrolyte and anodal active material. Possible anodal active material - cathode active material couples are zn-Ce, zn-Ni, zn-Co, zn-Mn and Zn-Fe, with Zn-Fe being a specific example.

EP2543104B1: This patent describes a redox fuel cell with a catholyte solution comprising at least one catholyte component and a redox mediator couple. The

***Chat Mode**: Interactive Q&A interface over the entire patent corpus.*



Here's what I found about organic materials:

Organic electrochemical cell: An organic electrochemical cell can convert surrounding ambient heat to electricity using reversible organic electron transfer reactions between organic donors and acceptors or TiO2 acceptor and organic donors, which take place at 28°C to 40°C (ID: 1). ⓘ

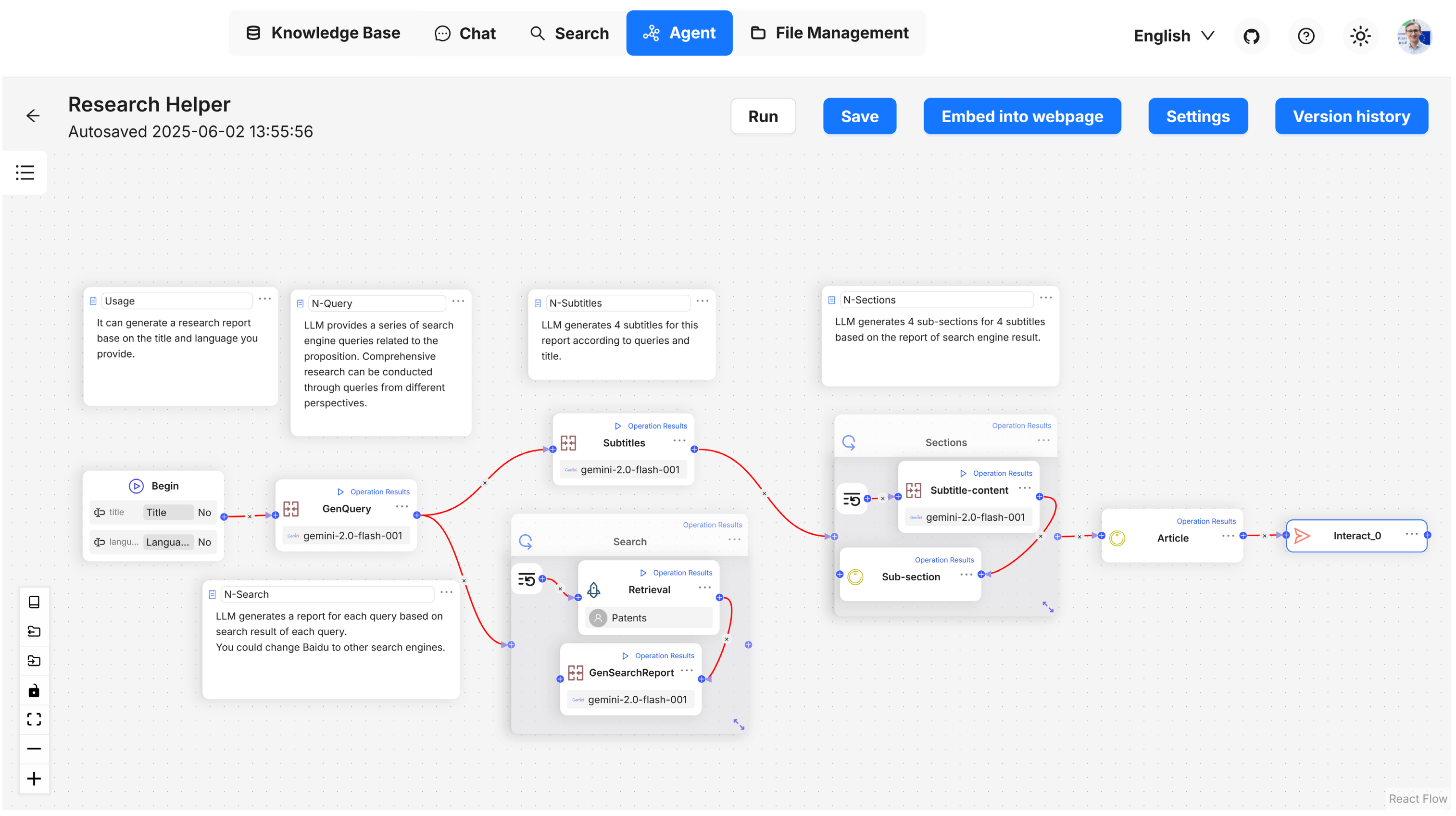
Organic silicon defoaming agent: A self-dispersing organic silicon defoaming agent can be made from materials with low toxicity and irritation ⓘ . It includes organic silicon oil, hydrophobic white carbon black, graphene oxide modified organic silicone oil, a composite emulsifier, a coupling agent, a thickening agent, and deionized water (ID: 8).

Organic biocomposite substrate: A soil-less biodegradable and porous rigid foam substrate for cultivating plants can be formed of a heterogeneous matrix of organic fibers pasteurized with organic matter and natural minerals, and then contacted with a fungus, resulting in a biocomposite foam with excellent capillary dynamics (ID: 10). ⓘ

Structured organic film (SOF): A structured organic film (SOF) can include a plurality of segments, a plurality of linkers, and optionally a plurality of capping segments, where at least one or more capping segments may include at least one cationic species ⓘ . An ion-exchange membrane may include the structured organic film (SOF) (ID: 3).

Organic-inorganic hybrid ion exchange membrane: An inorganic fiber reinforced organic-inorganic hybrid ion exchange membrane has a high dimensional stability and a high thermal resistance ⓘ . It is made of a siloxane bond which is stable in oxidation or made of a siloxane bond and saturated hydrocarbon, thereby having an excellent oxidative stability and thermal resistance (ID: 4).

***Search Mode:** Keyword/phrase search augmented with semantic ranking.*



Autosaved 2025-06-02 13:55:56

Usage: It can generate a research report base on the title and language you provide.

N-Query: LLM provides a series of search engine queries related to the proposition. Comprehensive research can be conducted through queries from different perspectives.

N-Subtitles: LLM generates 4 subtitles for this report according to queries and title.

N-Sections: LLM generates 4 sub-sections for 4 subtitles based on the report of search engine result.

N-Search: LLM generates a report for each query based on search result of each query.
You could change Baidu to other search engines.

***Research-Agent Mode:** Automated agents that can process user-level "research questions" (e.g., "What are the recent material innovations in redox flow electrolytes?") and generate structured reports.*

Want to see some example reports?
Scan me!

VANILLA FLOW
Funded by the European Union
VanillaFlow has received funding from the European Innovation Council (EIC) under grant agreement No 101115293.

TU Graz – Institute of Machine Learning and Neural Computation
8010 Graz, Inffeldgasse 16b/I, Austria, Tel.: +43 316 873-5811
georg.niess@tugraz.at