CORTEVA FORMULATION SCIENCE CHATBOT

CORTEVA agriscience

The Data Mine

Team: Diya Patel, Julian Yoon, Lewis Le, Tharun Senthilkumar, Victor Popescu, Vijaysimha Naidu, Xena Nguyen, Yog Trivedi

01 Introduction

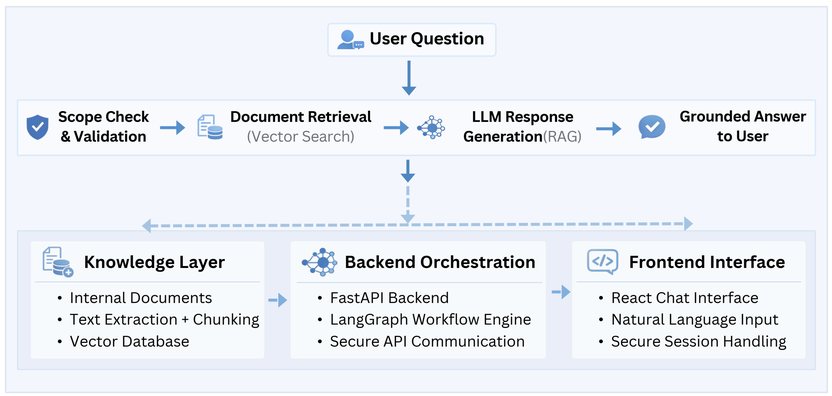
Formulation scientists at Corteva work with **large volumes of internal technical documents**, making it **time-consuming** to locate accurate and relevant information. To streamline this process, we developed a **secure, AI-powered Formulation Science Chatbot** that enables researchers to ask **natural-language questions** and receive **source-grounded answers** from internal documents. This system **improves efficiency** and makes trusted knowledge **more accessible**.

- Secure
- Domain-Specific
- Source-Grounded
- Efficient Retrieval

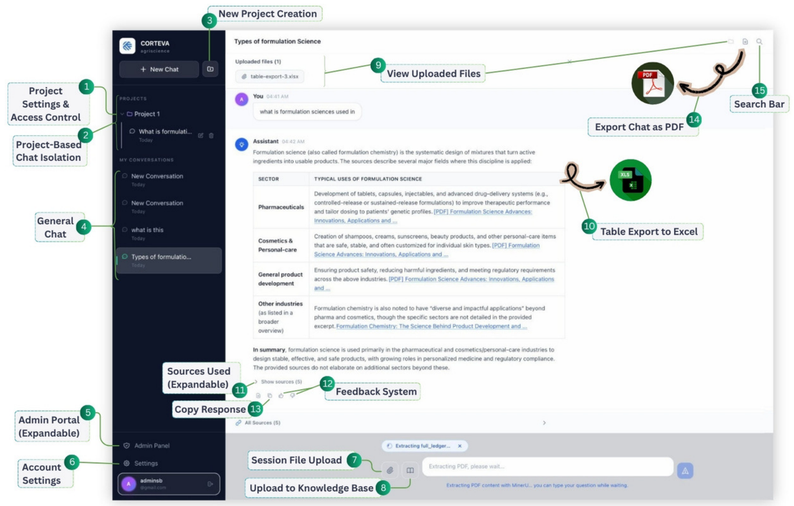
02 Project Overview

The Formulation Science Chatbot is a **secure, domain-specific AI system** built for Corteva's research teams. It uses **Retrieval-Augmented Generation (RAG)** to retrieve relevant internal document content and generate **concise, source-backed responses**. By combining semantic search with language generation, the system **reduces manual document searching** while maintaining **data security** and minimizing hallucinations.

03 Methodology



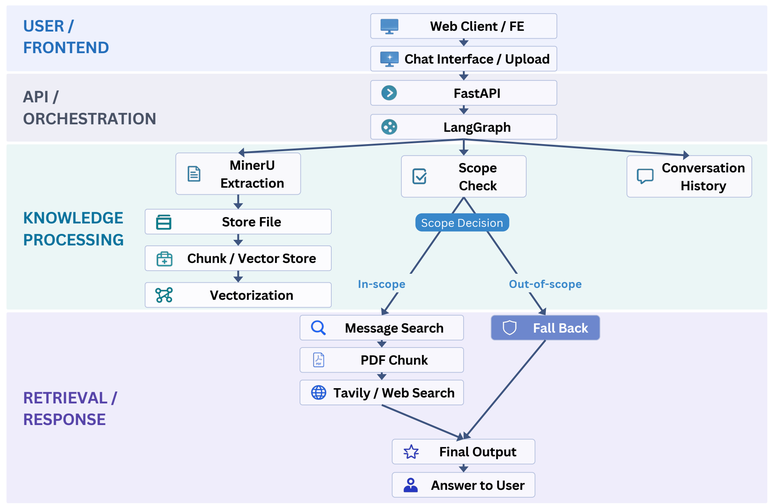
04 Frontend Interface & Features



1.1 Frontend Interface

Frontend user interface demonstrating key chatbot features and user workflows

05 System Architecture



06 Results & Current Capabilities

Traditional Workflow vs AI-Assisted Knowledge Retrieval

Manual Document Search	Formulation Science Chatbot
Time-consuming document review	Instant retrieval
Scattered PDFs	Centralized, searchable knowledge base
Manual scanning	Automatic section retrival
Required technical familiarity with documents	Conversational interace
Risk of missing info	Grounded, concise responses
No scope filtering	Upload + Indexing support

AI-Powered Retrieval

07 Conclusion & Future Work

IMPACT

- Reduces time spent searching formulation documents
- Enables faster, more informed research decisions
- Improves access to reliable, source-backed information

FUTURE IMPROVEMENTS

- Expand knowledge sources
- Improve retrieval accuracy
- Add long-term conversation memory
- Enhance citation visibility & confidence

Future work focuses on enhancing system performance, scalability, and overall user experience.

Acknowledgements & Key Sources

Purdue University - The Data Mine
Corteva Agriscience

The Data Mine — Staff & Teaching Team (incl. Maggie Betz & Bryce Castle) | **Special Thanks:** Kali Lacy (with interpreter Naomi) & Doug Crabill | **TAs:** Maninder Kaur & Clay Hampton

Mentors: Sam Zukowski & Jason Sun

Scan for References