Sandia National Laboratories

Exceptional service in the national interest

INFERNO

Leveraging Large Language Models with Sandia's Atlas to augment developer understanding of Sandia's Dante Simulation Codebase

Christopher Symonds – 06535

Catherine Appleby - 05575

ML/DL Workshop – Sandia National Laboratories

September 9-12 2024

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

BACKGROUND

LLMs are powerful helpers!

- Writing prose, code, Q&A

New techniques make them even more helpful

- Retrieval-Augmented Generation (RAG) Vector DBs

Software engineering can be helped in many ways

- Writing + Analyzing code
  - Inefficiencies and performance, security flaws, etc.
- Writing documentation
- Code Q&A – a guide for developers new to your code
  - "Where in the code are we doing path planning?"
  - "What does foo.cpp do?"

BRINGING LLMS TO SANDIA CODE

- Sandia – National Security Mission
- Includes protection of critical national assets
- How to ensure protection systems safekeep assets?
- Live exercises are best, but costly

DANTE™

BRINGING LLMS TO SANDIA CODE

- Dante – a force-on-force simulation system

- Mature codebase (>15 years!)

- Rotation of developers over the years

- Documentation, training materials, etc.

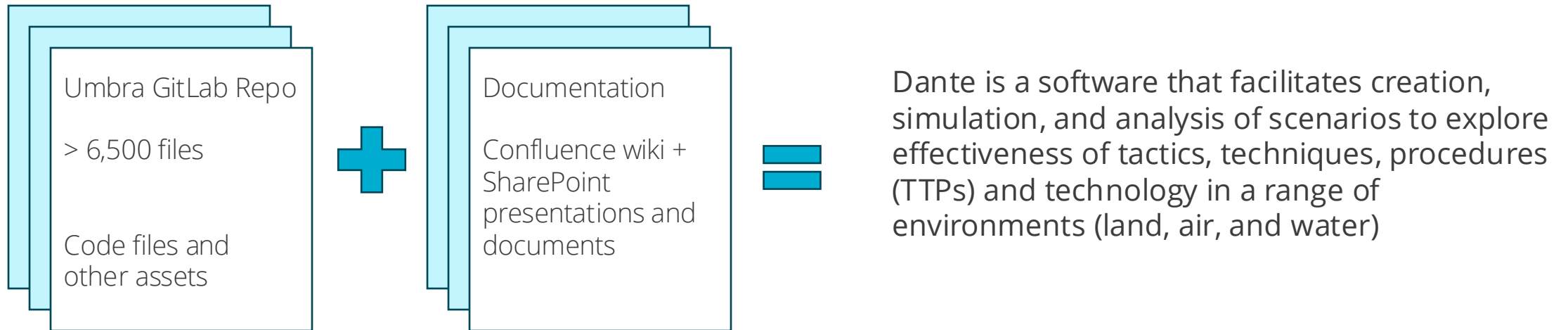# LARGE LANGUAGE MODELS FOR AIDING DEVELOPERS

1. Code completion
   - Generates code that the developer is likely to want next
   - GitHub copilot

2. Code documentation
   - Summarize what is happening in code to augment or create documentation

3. Code translation
   - Turn a codebase in one language (ex. MatLab) to another language (ex. Python)

4. Code Q&A
   - Interactively ask questions and receive answers about a codebase

WHY CAN'T WE USE AN LLM OUT OF THE BOX?

You 09:29 AM

what is Dante?

ATLAS 09:29 AM

Dante is a software package used for handling and analyzing data in computational fluid dynamics (CFD) simulations. It is developed and maintained by the Center for Exascale Technology (CET) at Sandia National Laboratories.

Umbra GitLab Repo

> 6,500 files

Code files and other assets

+

Documentation

Confluence wiki + SharePoint presentations and documents

=

Dante is a software that facilitates creation, simulation, and analysis of scenarios to explore effectiveness of tactics, techniques, procedures (TTPs) and technology in a range of environments (land, air, and water)

PROMPT ENGINEERING

System prompt – a set of instructions for the model to follow when answering questions

- You are an AI assistant who explains a codebase to programmers.
- Answer any questions relating to the codebase with thorough explanations.
- Reference files from the codebase to support your answers.

Initial messages – a set of questions and answers added to the model's memory, so it has context for new queries
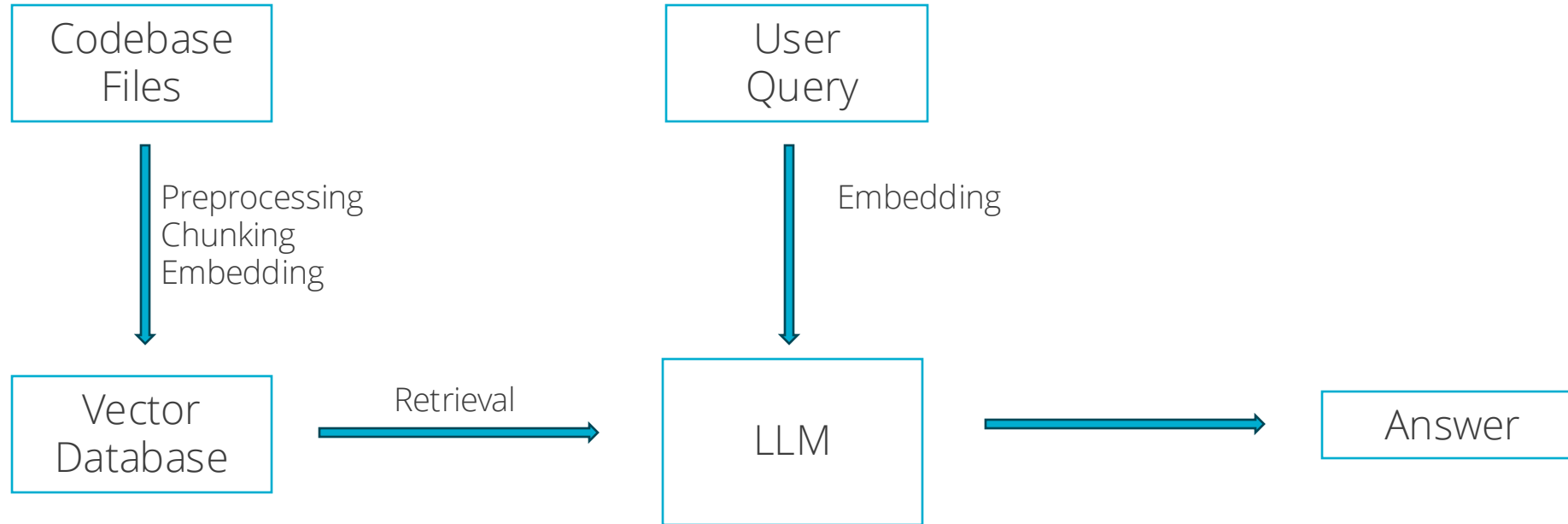
> Role: assistant, Content: How may I help you?
> Role: user, Content: What is the purpose of this codebase?,
> Role: assistant, Content: This codebase is a force simulation software

Preprompting and postprompting – adding additional text before and after each of the user's prompts to further guide answers

INITIAL SETUP

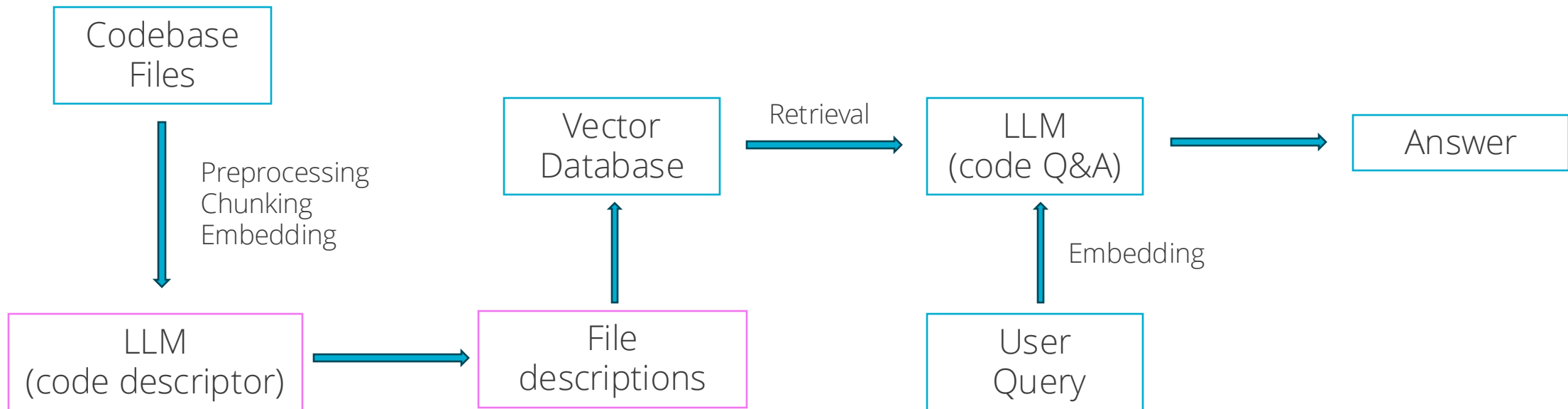Retrieval Augmented Generation using embeddings of the codebase

Performance

✓ Questions about individual code files

X Questions about overall codebase, requiring knowledge of hierarchical class structures

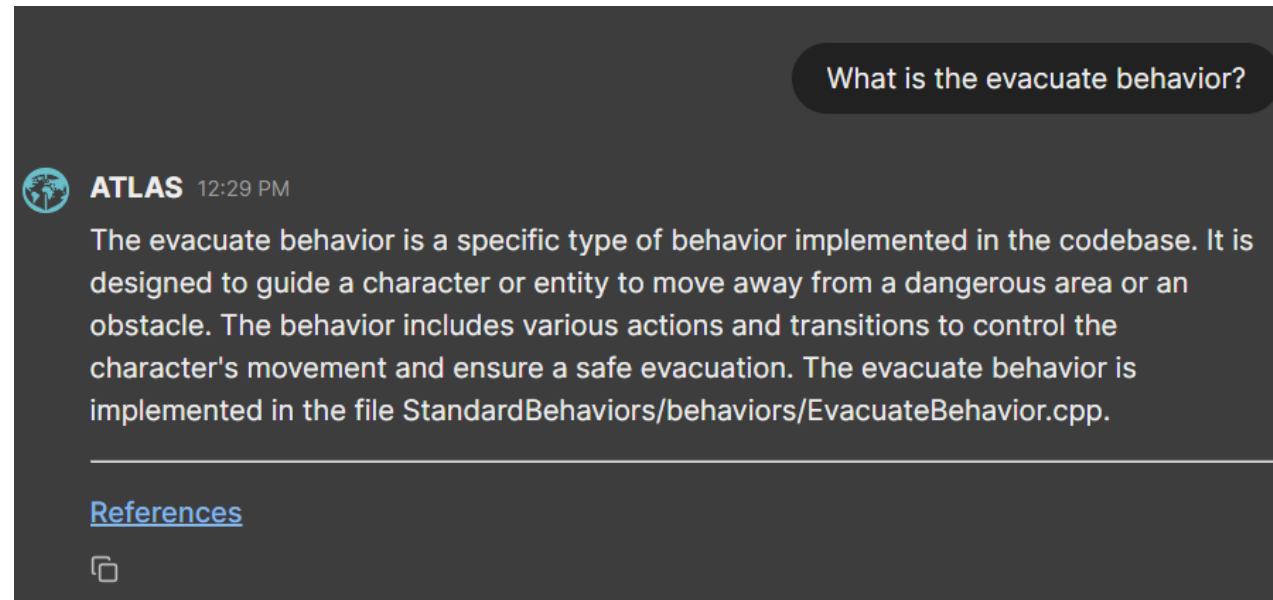Retrieval Augmented Generation using generated documentation of the codebase

Performance

✓ Questions about individual code files

✓ Questions about overall codebase, requiring knowledge of hierarchical class structures

RESULTS AND NEXT STEPS

Able to accurately answer questions on the codebase

Next steps
- Integrate with GitLab CI/CD
- Productionize on Atlas
- User testing

THANK YOU!