

Introduction

Background (Revisiting SC2)

- Currently, CAT engineers fetch machine data from SuperComm2, Caterpillar machine data interface. However, manually querying and analyzing data can be tedious and confusing.

Project Goals

- Intelligent Data Exploration:** Develop an LLM-integrated chatbot that summarizes machine data channels without requiring manual user queries.
- Automated Analytical Insights:** Enable the system to perform high-level statistical analysis, such as calculating sensor ranges, to provide immediate diagnostic summaries.

Research Methodology

Exploratory Data Analysis

- Familiarized the team with Caterpillar data through time-series data exploration and visualization of machine data

Data Visualization Tools

- Programmed functions to build histograms, line charts, and geospatial maps from user inputs

Model Context Protocol Integration

- Implemented a REST API backend and MCP Client-Server system connected to an LLM for natural language query processing and appropriate tool selection

UI Design and User Experience

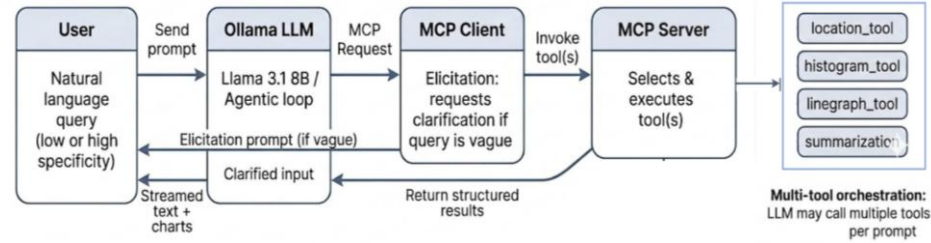
- Designed a user-friendly app UI to interface with the LLM and display generated visualizations.

Tools for Data Exploration

- Programmed and implemented additional tools for exploring data availability and types of data channels.



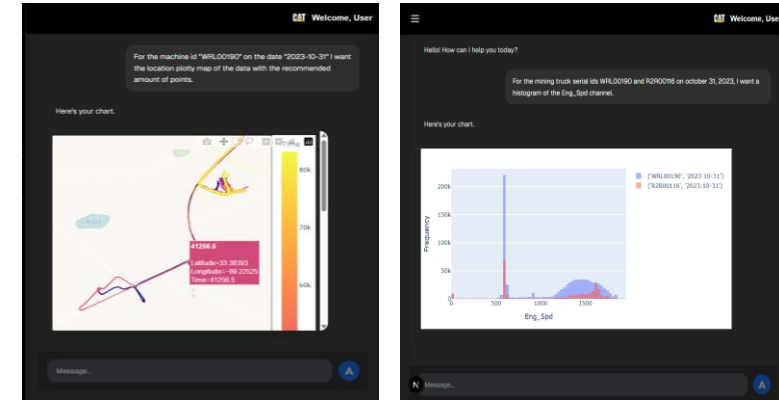
Application Workflow



Our application flows from the User to LLM (Ollama) to MCP Client and then to the MCP server which calls what specialized tool to use. Each layer communicates with each other to ensure the correct end result.

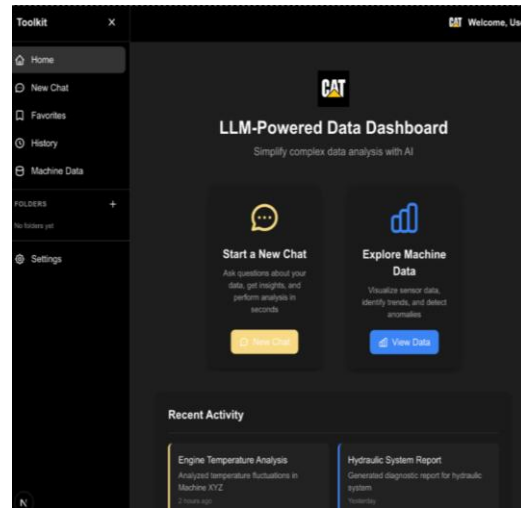
Figure 1: MCP tool calling application workflow

User Chat Examples



Figures 3/4: chat examples querying for geospatial location & multiple trucks

UI and Features



Intuitive Chatbot Interface:

- Create new chats and communicate with the app in natural language

Exploratory & Analytical Queries

- Ask about serial numbers, machine families, and channels; get summaries, min/max, etc.

Visualization Tools

- Generate line graphs, histograms, and geolocation maps across single or multiple machines and columns

SC2 Access & Authentication

- Fetches data from a live CAT SC2 API data endpoint

Figure 2: UI and features

Conclusion and Future Goals

Conclusion

- Developed an LLM chatbot and MCP server to automatically query, analyze, and visualize product sensor data from natural language inputs.
- Created a full-stack application that converts raw data into interactive 2D/3D maps and graphs to summarize machine data channels.

Future Goals

- Secure Access:** Integrate CAT Microsoft authentication for single sign-on access across the application
- AWS Integration:** Migrate to AWS Bedrock to leverage enterprise-grade LLMs within the CAT cloud environment

Acknowledgements and References

Thank you to Caterpillar Inc., our mentor Jim Katter, all the Data Mine staff, Dr. Ward, and our TA Sid Sancheti.

References: Model Context Protocol. (2024, November 25)