



Everyday AI for Statistical Programmers: Global Implementation of a Clinical Analysis Assistant

Mathieu Cayssol, Senior Data Scientist
Christoph Centner, Analytical Data Scientist

Table of contents

Introduction

Challenge

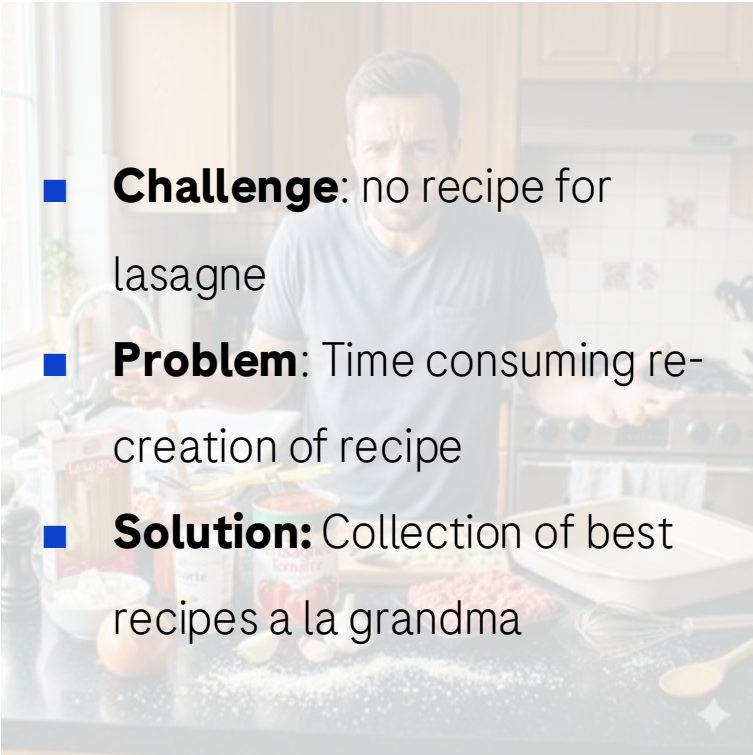
Solution

Technical Implementation

Adoption

Future Work

Introduction

- 
- **Challenge:** no recipe for lasagne
 - **Problem:** Time consuming re-creation of recipe
 - **Solution:** Collection of best recipes a la grandma

- 
- **Challenge:** Tight timelines for programming new outputs
 - **Problem:** Fragmented codebases lead to redundant programming
 - **Solution:** Code reusability supported by AI

Introduction

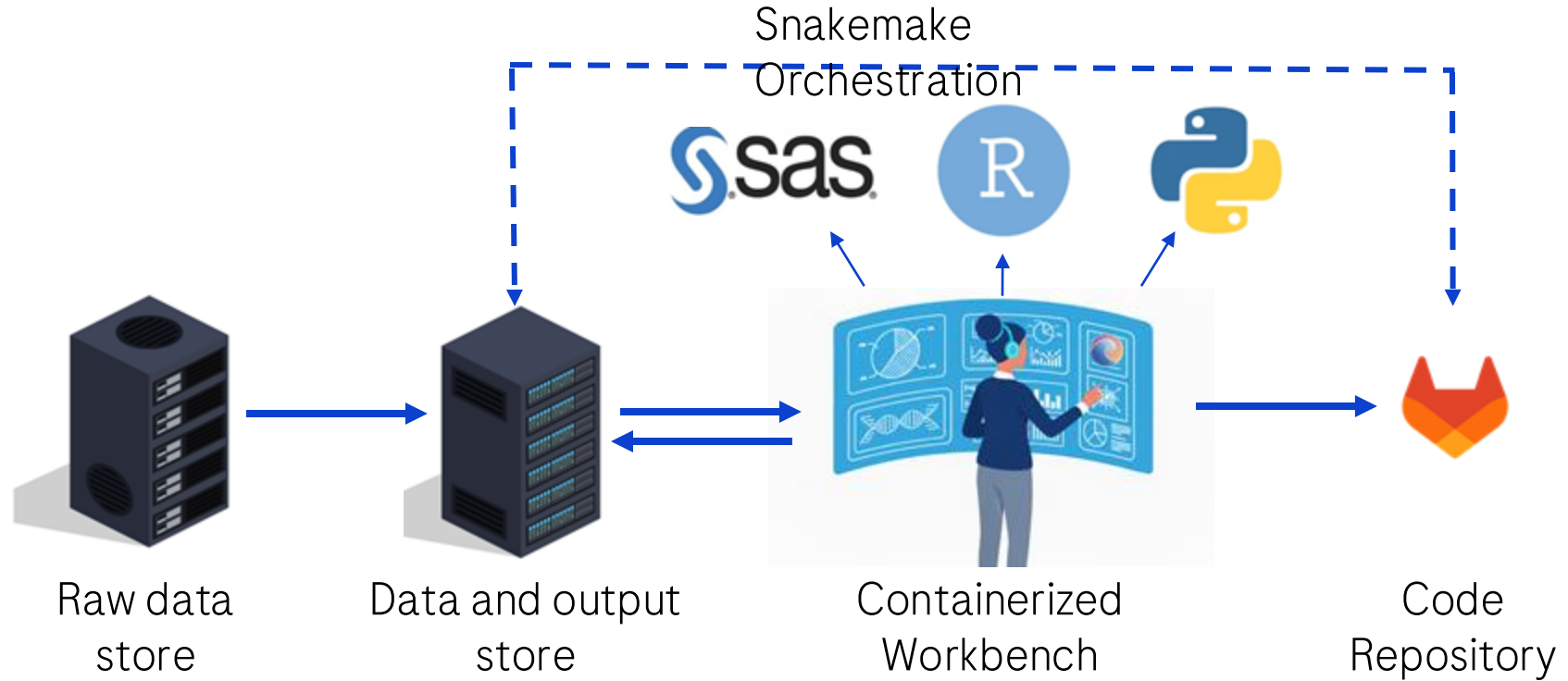
- **The Code Search Agent (CSA):** A novel, AI-powered solution built on top of our existing infrastructure.
- Core Capability: Combines **Generative AI with a large-scale, metadata-driven code search.**
- How it Works: Indexes **over 70,000 programs**, enriching code chunks with LLM-generated summaries.

Less duplicated effort + facilitating **reuse**

How is the situation at Roche?

Challenge

One CEntralised ANalytics platform (OCEAN)



Challenge

OCEAN Challenges



Steep Learning Curve: New skills needed (Git, containers, orchestration).



Ecosystem Transition: Moving from single-tool to an integrated toolchain



Fragmented Information: Guidance and validated code are scattered across repositories and wikis.



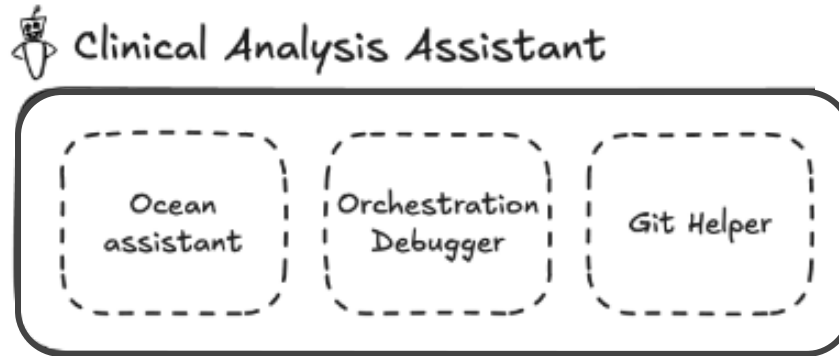
Code Discoverability and Reuse: finding, comparing, and adapting validated code remains manual and time-consuming.

Solution

The Clinical Analysis Assistant (CAA)

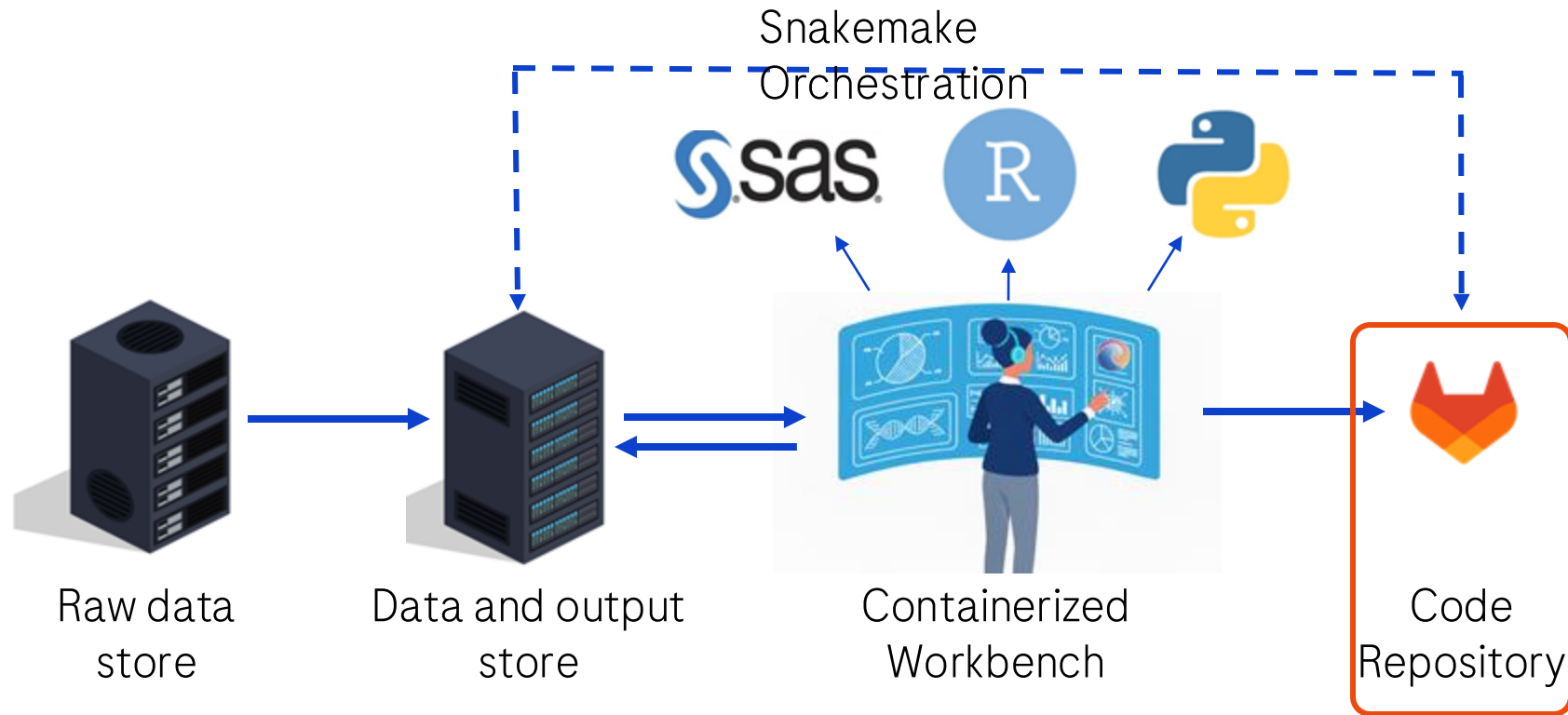
Clinical Analysis Assistant: An AI-powered conversational layer built on OCEAN to bridge the transition gaps.

Initial Agents:



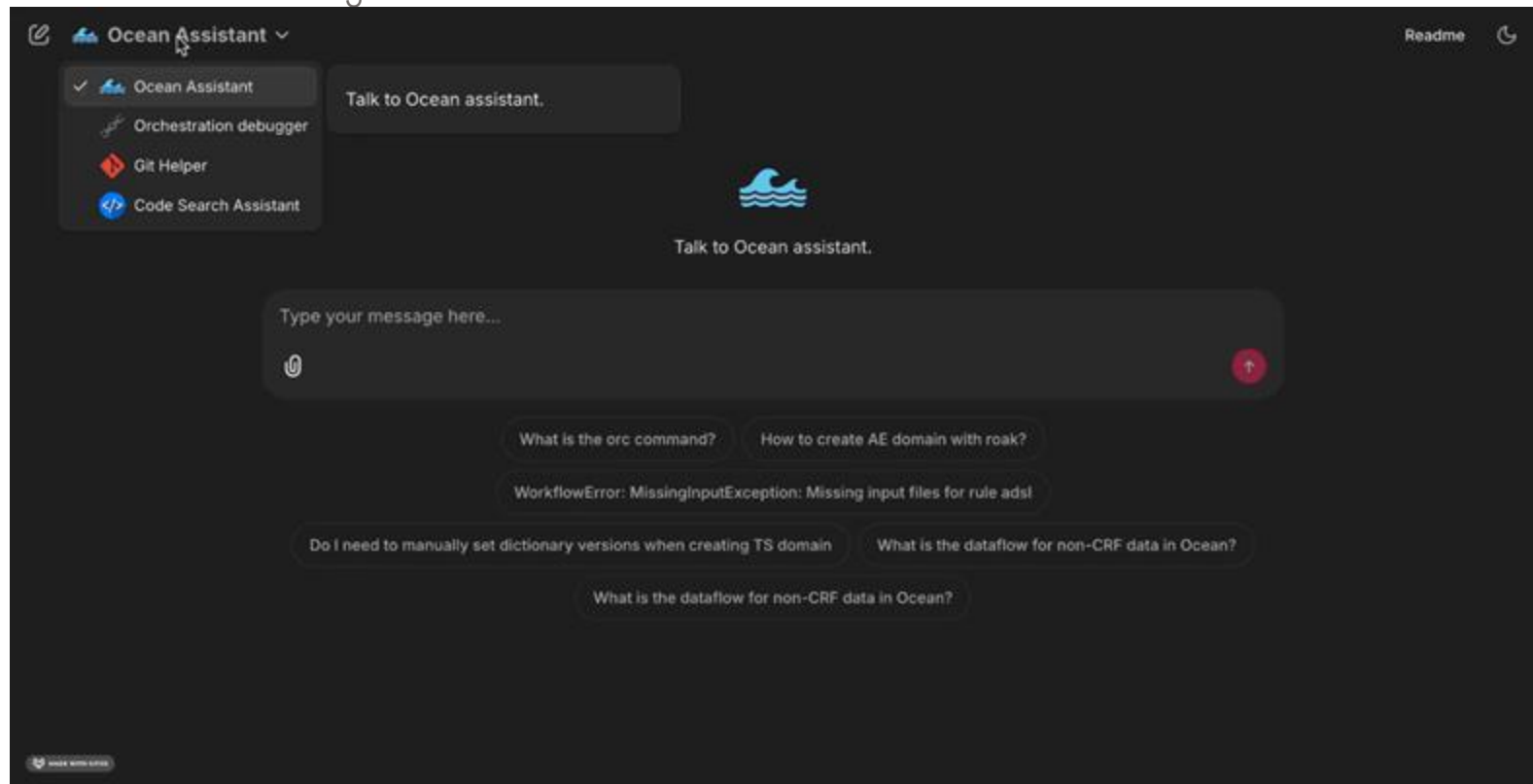
Challenge

One CEntralised ANalytics platform (OCEAN)



Solution

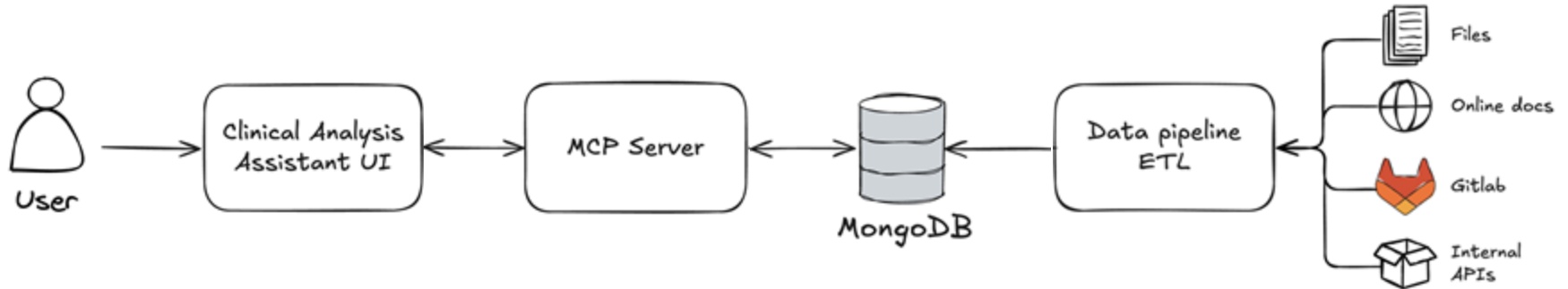
Demo of Code Search Agent



Technical Implementation

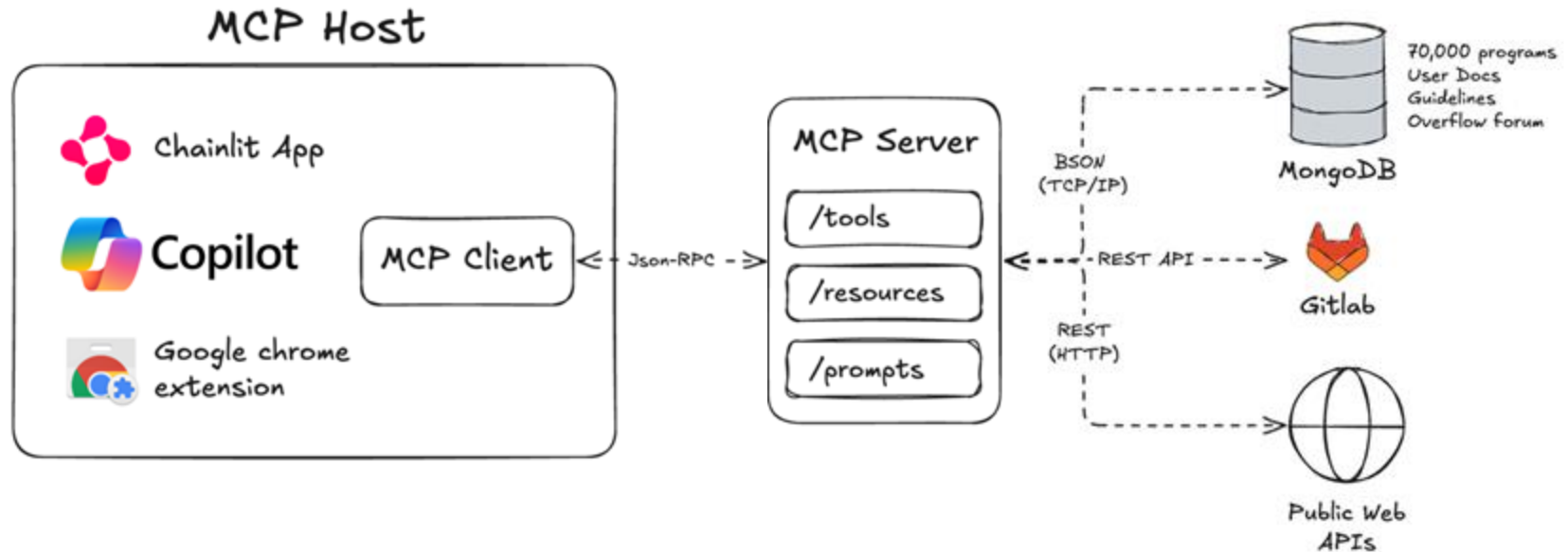
High level overview of Clinical Analysis Assistant

Clinical Analysis Assistant: Purpose: AI-powered chatbot to support clinical programmers in the OCEAN platform - improving code reuse, documentation access, and debugging.



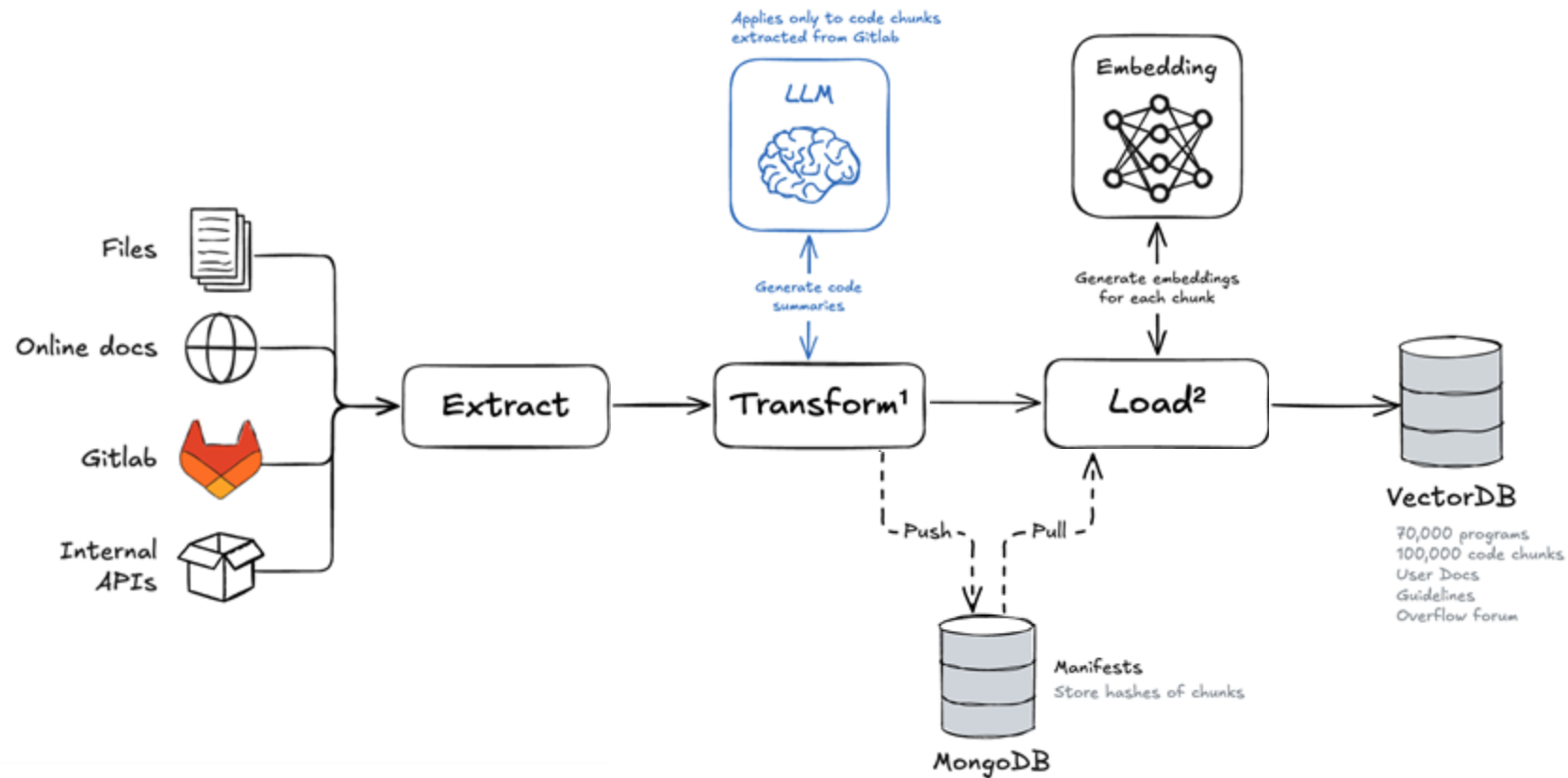
Technical Implementation

Clinical Analysis Assistant MCP



Technical Implementation

Data Pipeline - ETL

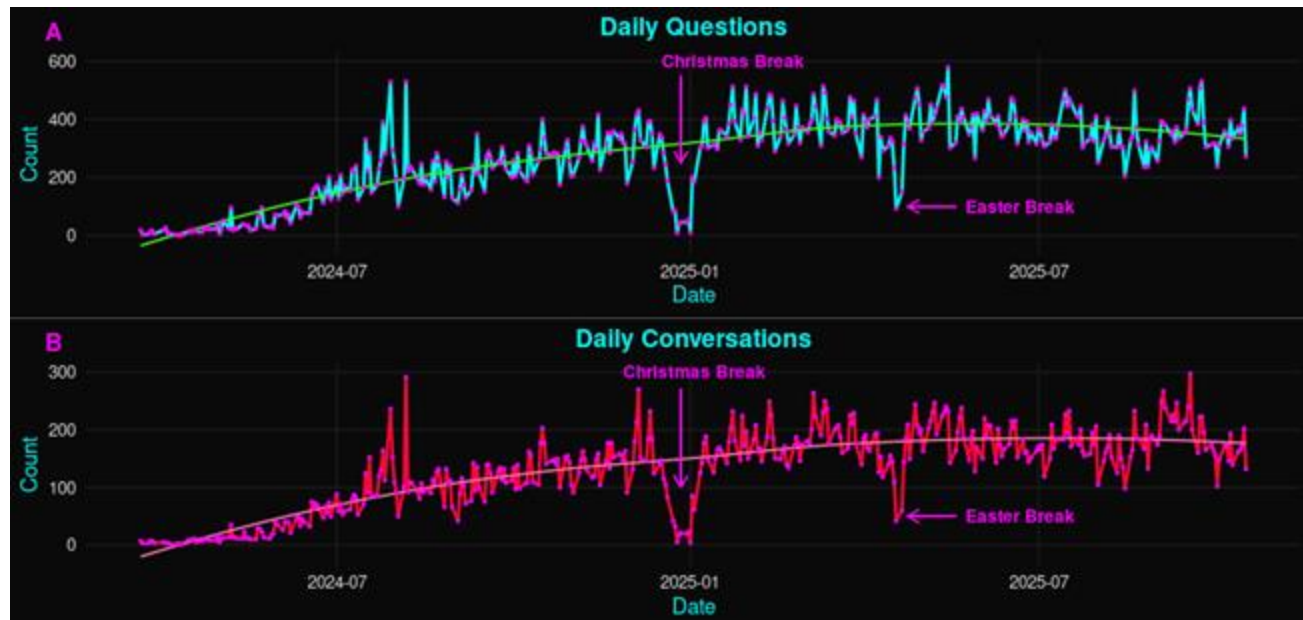


Adoption

Clinical Analysis Assistant

+120K
Questions

~600
Unique users



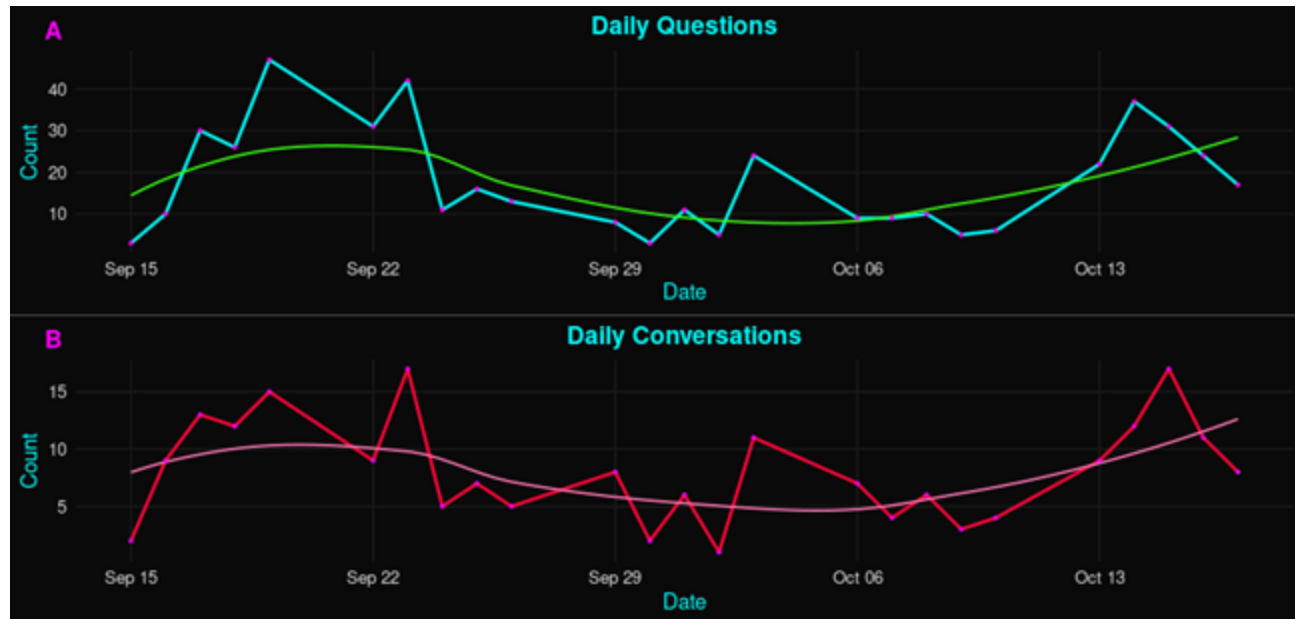
Clinical Analysis Assistant

Adoption

Code Search Agent

+1K
Questions

~45
Unique users



Code Search Agent

Summary

Current Limitations

- **Lacks context** of the study and working environment from the statistical programmer
- No capability for **live execution** and **validation** (no sandbox environment).
- Web-based interface requires an unwanted **context switch** (not in-line).



Summary

Next Milestones

- **Deep IDE Integration**

Embed CSA into RStudio and VS Code via MCP for seamless, in-context access.

- **Integration with AiR (AI for R)**

Combine CSA with AiR to generate study-specific R code.

- **Study aware multi-agents**

Build intelligent agents that understand study context and autonomously produce compliant code.



Summary

*“Spend less time “**reinventing the wheel**” and more time on meaningful analysis.”*

- Code Search Agent reduces **inconsistencies** and **duplication**.
- Demonstrated **high user adoption** and improvements in **efficiency**

Key Takeaways



Are there any questions?

Mathieu Cayssol

mathieu.cayssol@roche.com

Christoph Centner

christoph.centner@roche.com



Thank you very much
for your attention