

COMPILER RESEARCH TEAM
Princeton University

DATE
16/10/2024



INTEGRATE A LARGE LANGUAGE MODEL WITH THE XEUS-CPP JUPYTER KERNEL

Google Summer of Code 2024 - Final Presentation

MENTORS

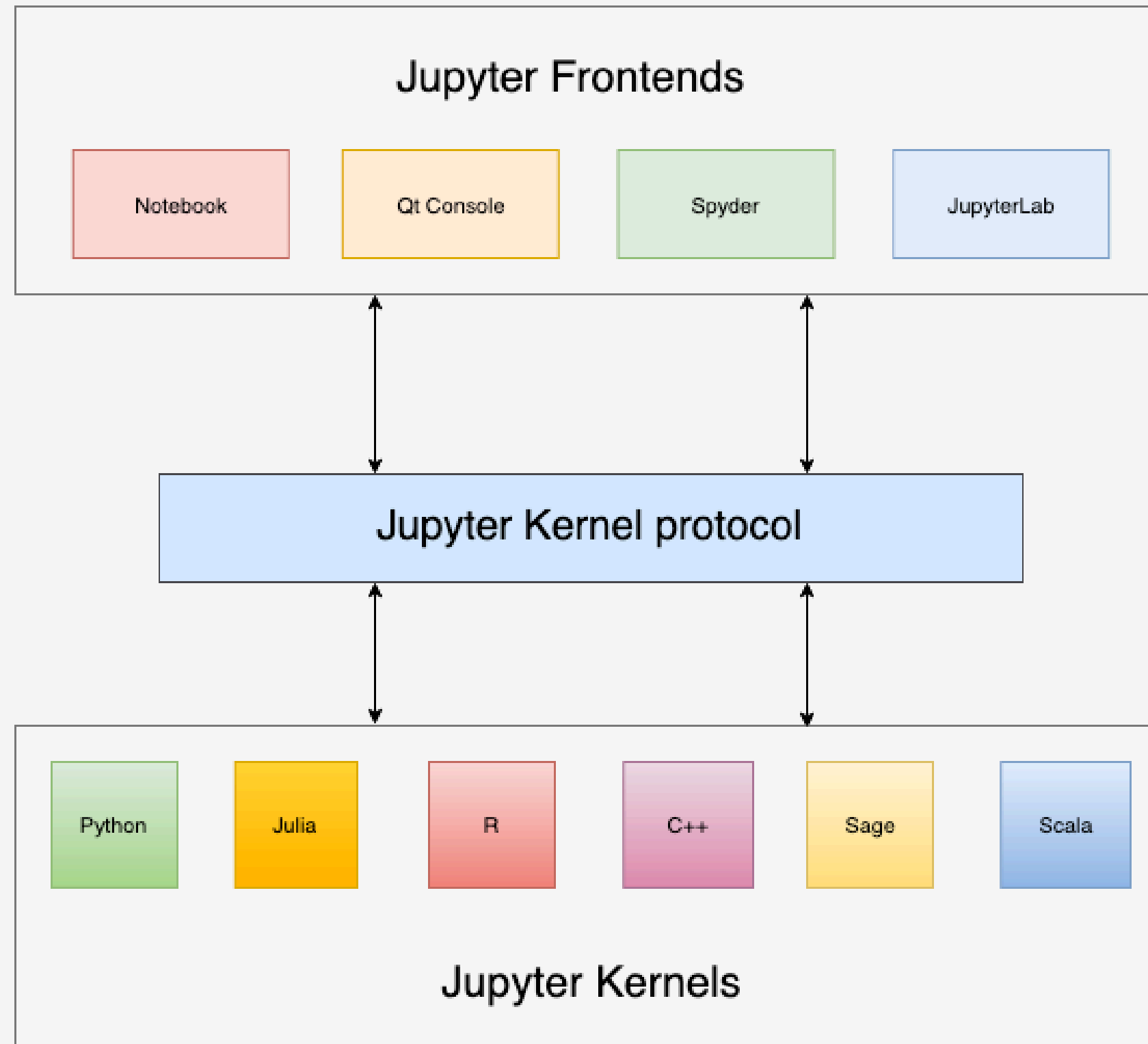
Anutosh Bhat, Johan Mabille, Aaron Jomy,
David Lange, Vassil Vassilev

MENTEE

Tharun Anandh

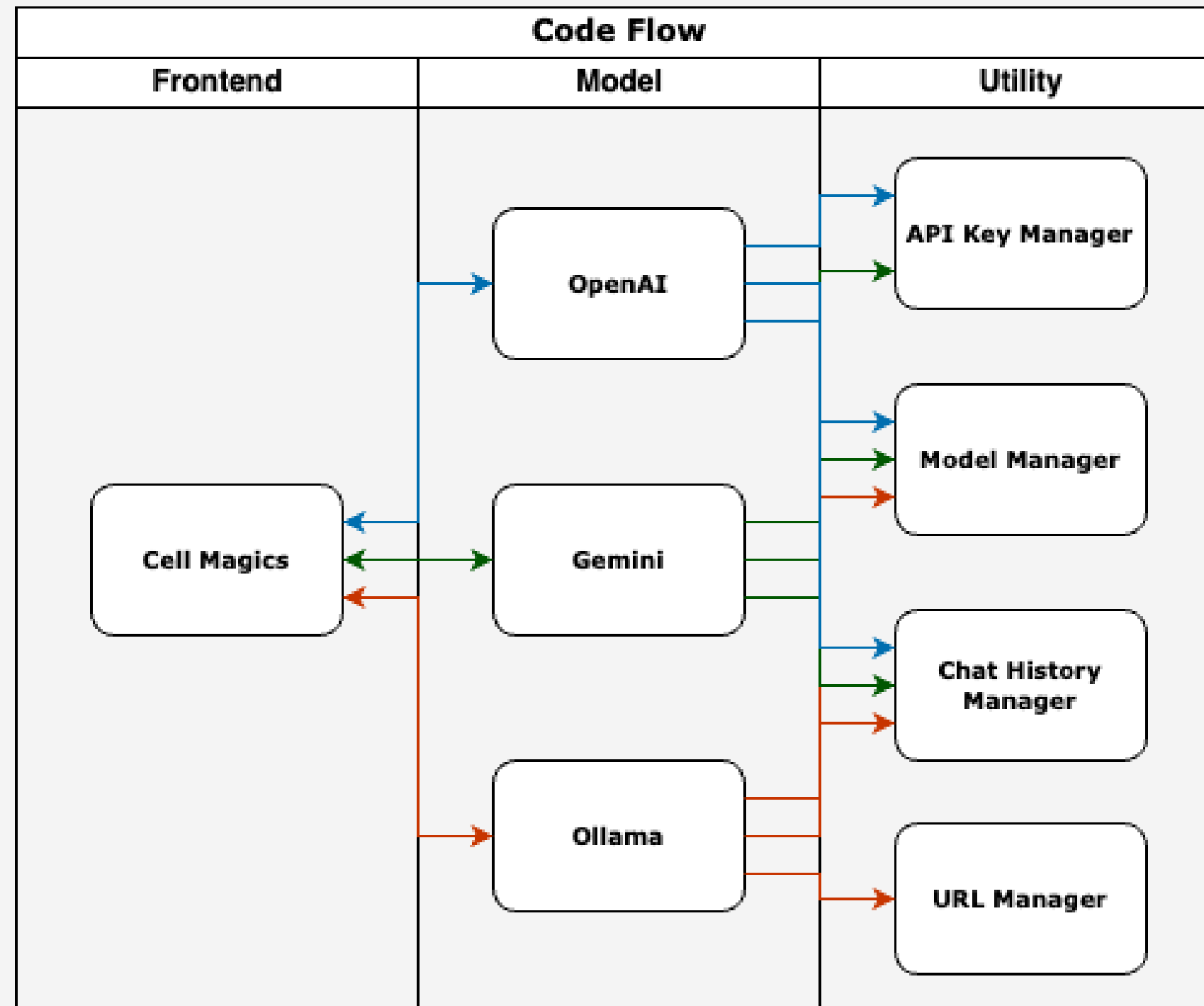
ABOUT XEUS

- Xeus is a framework meant to facilitate the implementation of kernels for Project Jupyter. It takes the burden of implementing the Jupyter Kernel protocol so developers can focus on implementing the interpreter part of the kernel.



WHY THIS PROJECT?

- With the ever growing popularity of large language models, this project aims to integrate a large language model with the xeus-cpp Jupyter kernel. This integration will enable users to interactively generate and execute code in C++ leveraging the assistance of the language model.



CODING PHASE 1

- Test coverage.
- Testing framework.
- Understanding XMagics.
- OpenLLM research.
- Using cURL library in cpp.
- Saving API keys in text files.
- Customized JSON calls for each model.

CODING PHASE 2

- Fix issues with initial PR.
- Debug windows issue.
- Add context (chat history) to LLM calls.
- Give user more freedom on the model used.
- Support for Ollama.
- Documentation.

COMPILER RESEARCH TEAM
Princeton University

INTEGRATE A LARGE LANGUAGE MODEL WITH THE
XEUS-CPP JUPYTER KERNEL

DATE
16/10/2024

DEMO

WHAT'S NEXT?

- Explore other projects in the org
- Xeus-cpp clad issue
- Plugin framework for xeus cpp
- Testing framework for clad

COMPILER RESEARCH TEAM
Princeton University

INTEGRATE A LARGE LANGUAGE MODEL WITH THE
XEUS-CPP JUPYTER KERNEL

DATE
16/10/2024

THANKS FOR LISTENING

MENTORS

Anutosh Bhat, Johan Mabilie, Aaron Jomy,
David Lange, Vassil Vassilev

MENTEE

Tharun Anandh