

Launch Your Own AI Assistant

Zichao Yang

Zhongnan University of Economics & Law

Date: November 9, 2025

Roadmap

In this session, we will discuss three ways to launch your own AI assistant:

- Using your large language models (LLMs) via a browser
- Hosting your LLMs locally
- Accessing your LLMs via APIs

LLM via browser

All popular LLMs can be accessed through a browser on your computer. (*Note: not every company provides service in every region of the world.*)

- DeepSeek
- Kimi
- ChatGPT
- Grok
- Claude
- Perplexity

Host Local LLMs

You can also host small size LLMs on your personal computers.

Choice 1: LM Studio

Choice 2: Ollama + AnythingLLM

Large LLMs via APIs

To utilize cutting-edge LLMs, the most effective approach is to access them through their provided APIs.

In this guide, I will demonstrate how to use **Cherry Studio** to access the **DeepSeek models** hosted on ByteDance's **Volcano Engine** servers.