

WIMI CHAT v4.0

AI 驱动的下一代智能对话平台产品介绍

版本演进与定位

演进历程

版本	核心特性
v1.x - v2.x	基础对话 & CLI 优化
v3.x	Web 界面引入
v4.0	工具 + 技能 + MCP

v4.0 核心更新

工具系统 (Tool System)

集成内置与动态工具，赋予 AI 结构化执行能力。

技能系统 (Skill System)

预定义领域知识库，LLM 自动感知并加载上下文。

双端架构 (Dual-mode)

CLI 与 Web 模式深度融合，支持 SSE 实时流式响应。

MCP 集成 (MCP Standard)

全面支持 Model Context Protocol，接入全球工具生态。

核心架构概览



后端 Backend

FastAPI Python SQLite

- REST API 服务提供
- LLM 适配与对话管理
- 工具注册与执行引擎
- 技能加载与解析



前端 Frontend

Vue 3 TypeScript Vite

- 响应式聊天 UI 界面
- 工具管理与配置工作台
- Pinia 状态管理逻辑
- Element Plus 组件库



通信与体验

SSE Hooks Stream

- SSE 流式实时响应
- 思考过程可视化展示
- CLI/Web 双端同步
- 打字机效果流畅交互

工具与技能系统

工具系统 (Tools)

内置工具池

`bash` `calculator` `datetime` `http` `read_file`

动态扩展能力

- 支持通过数据库动态加载 JavaScript 工具
- LLM 自动识别参数 Schema 并生成调用指令
- 沙箱化执行环境，确保系统命令调用安全

技能系统 (Skills)

领域知识包

`pdf-processing` `xlsx-expert` `order-manager`

核心逻辑

- 预定义上下文：动态注入领域操作指南
- 参数占位符：支持灵活的变量替换与渲染
- 自动感知：LLM 根据意图自动触发技能加载
- 工具授权：技能可限定允许调用的工具集

MCP Server 集成（新特性）

MODEL CONTEXT PROTOCOL

WIMI CHAT v4.0 深度集成 MCP 标准协议，通过统一的接口规范，实现 LLM 与外部工具服务器的无缝连接。

`stdio`

标准输入输出

适用于本地进程通信 (uvx/npx)

`http`

远程 HTTP 服务

支持远程 MCP 服务器集群接入

已集成服务生态

`my_sqlite_db`

基于 mcp-alchemy 的 SQLite 数据库自动化查询与分析。

`fetch`

高性能网络请求工具，支持网页内容抓取与解析。

`tavily-mcp`

集成 Tavily AI 搜索，提供精准的实时网络信息检索。

`context7`

文档向量化查询，支持大规模知识库的语义检索。

核心价值与未来规划

★ 核心价值维度

交互体验

CLI 与 Web 双模，SSE 实时流式交互，过程透明。

扩展能力

内置与动态工具，MCP 插件化接入全球生态。

技术领先

FastAPI + Vue3，模块化设计，便于维护。

📍 未来路线图

多用户认证系统

Security & Multi-tenancy

对话历史云同步

Data Persistence

支持更多 LLM 提供商

Model Diversity