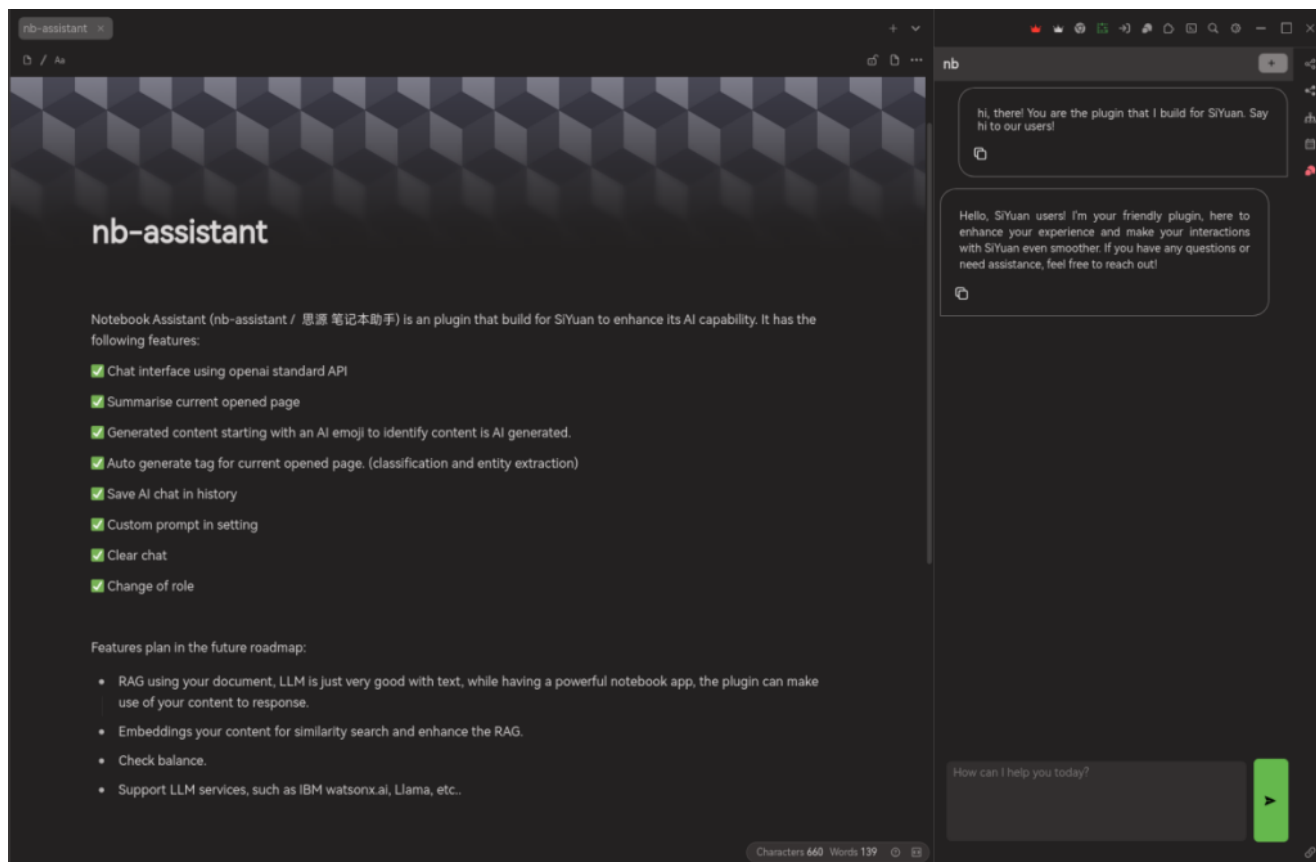


思源笔记本助手

English

思源笔记本助手 (nb-assistant) 是一个为 SiYuan 构建的插件，旨在增强其 AI 能力。



它具有以下功能：

- ✅ 使用 OpenAI 标准 API 的聊天界面
- ✅ 总结当前打开的页面
- ✅ 生成的内容以 AI 表情符号开头，以识别内容是 AI 生成的
- ✅ 自动为当前打开的页面生成标签（分类和实体提取）
- ✅ 保存 AI 聊天记录
- ✅ 在设置中自定义提示
- ✅ 清除聊天记录
- ✅ 角色切换让大模型更好的回答你的问题

注意！这个插件使用思源原生AI设置，你需要完成AI设置才能使用这个插件。

设置

搜索

编辑器

文档树

闪卡

AI

资源

导出

外观

集市

搜索

快捷键

账号

云端

关于

OpenAI

API 提供商

选择后将使用该提供商的 API 服务实现 AI 相关功能

OpenAI

超时时间

发起请求的超时时间，单位：秒

30

最大 Token 数

请求 API 时传入的 `max_tokens` 参数，用于控制生成的文本长度

2048

温度

请求 API 时传入的 `temperature` 参数，用于控制生成的文本随机性

0.4

最大上下文数

请求 API 时传入的最大上下文数

7

模型

请求 API 时传入的 `model` 参数，用于控制生成的文本风格（使用 Azure OpenAI 服务时需填入 Deployment ID）

LLaMA_CPP

API Key

建议为思源单独分配 API Key 以便于后续管理。该项留空则禁用 AI 相关功能

.....

网络代理

发起请求的网络代理，如 `socks://127.0.0.1:1080`

API 基础地址

发起请求的基础地址，如 `https://api.openai.com/v1`

http://localhost:8080/v1

以上设置同时也适用于本地LLM的配置，当然任何支持OpenAI API spec的服务也可以尝试配置。

未来路线图中的功能计划：

- 使用您的文档进行 RAG，LLM 在文本处理方面非常出色，而拥有一个强大的笔记本应用程序，插件可以利用您的内容进行响应。
- 嵌入您的内容以进行相似性搜索并增强 RAG。
- 检查余额。
- 支持 LLM 服务，例如 IBM watsonx.ai、Llama 等。

当前版本限制

- Zero shot。您只是提示 LLM 进行响应，因此结果完全取决于 LLM 的训练程度。
- 该插件仍在开发中，预计需要相当长的时间进行预览，以使其表现良好。

