langchain官网位于[python.langchain.com/en/latest/i…](https://link.juejin.cn?target=https%3A%2F%2Fpython.langchain.com%2Fen%2Flatest%2Findex.html" \o "https://python.langchain.com/en/latest/index.html" \t "_blank) ，它可以简化对各种大语言模型的使用，比如内置有OpenAI等。

pdfplumber顾名思义，是用于读取和处理PDF文件的，选择这库是因为今年还在更新，并且对中文的支持还不错。本次项目demo会导入自己的PDF文件，并将其作为知识库，回答你的提问。

python-dotenv用于读取.env文件，本例在该文件中放入Open AI平台的key。

streamlit用于绘制 UI界面，当前大多数ChatGPT应用都使用它，如大名鼎鼎的[gpt4free](https://link.juejin.cn?target=https%3A%2F%2Fgithub.com%2Fxtekky%2Fgpt4free" \o "https://github.com/xtekky/gpt4free" \t "_blank)，streamlit默认会收集信息进行分析，可通过配置文件关闭，macOS和Linux位于~/.streamlit/config.toml，Windows位于%userprofile%/.streamlit/config.toml，添加如下内容即可：

**[browser]**

gatherUsageStats = false

其它配置项可通过streamlit config show可查看。

faiss-cpu是facebook开源用于相似搜索的库，[github.com/facebookres…](https://link.juejin.cn?target=https%3A%2F%2Fgithub.com%2Ffacebookresearch%2Ffaiss" \o "https://github.com/facebookresearch/faiss" \t "_blank) ，GPU版本请使用faiss-gpu。

openai和tiktoken都是调用ChatGPT接口时使用的。

保留当前使用版本请使用pip freeze > requirements.txt。

OpenAI的注册方式这里就不介绍了，最常用的是在[sms-activate.org/](https://link.juejin.cn?target=https%3A%2F%2Fsms-activate.org%2F) 上购买服务获取短信验证码完成注册。

提取文本：

接下对文本进行分片，这里每个分片长充为1000字符，为保留上下文选择了重叠200字符：

接下来配置 embedding，也即将离散值转化为连续向量：

为界面添加一个输入框：

最后完成回复的逻辑：