**Graph Rag 本地部署与基础使用**

**本地部署**

**安装官方package**

|  |
| --- |
| PowerShell pip install graphrag |

**修改Tokenizer**

注意：因为graphrag默认使用ck100l\_base，在处理中文语料时会在chunck开头产生乱码。所以我们需要替换成langchain.text\_splitter里的RecursiveCharacterTextSplitter

* 在graphrag里找到tokens.py
* 默认路径为~\Users\AppData\Local\Programs\Python\{your\_Python\_version}\Lib\site-packages\graphrag\index\verbs\text\chunk\strategies\tokens.py
* 将tokens.py替换成如下文件并安装langchain
* [tokens.py](https://zhipu-ai.feishu.cn/file/AlWYbMPZto0X62x4qOvccIOenFe)
* 注意：目前Graphrag v0.3.0 依赖 tenacity>=9.0.0，langchain 依赖tenecaity<9.0.0, 如果遇到依赖冲突需要进行如下操作
* 下载text\_splitter.py并保存到同一目录

**[该类型的内容暂不支持下载]**

* 更改替换后的tokens.py

|  |
| --- |
| Python #删除 # from langchain.text\_splitter import RecursiveCharacterTextSplitter  #替换为 from .text\_splitter import RecursiveCharacterTextSplitter |

**基础使用**

注意：详细使用及参数设置请参考官方文档（[Get Started](https://microsoft.github.io/graphrag/posts/get_started/)）。下列运行代码只适用于graphrag 0.3.0 和0.3.1,如遇报错，请参考官方文档（[Get Started](https://microsoft.github.io/graphrag/posts/get_started/)）和报错信息

**创建工作目录**

|  |
| --- |
| PowerShell mkdir -p ./ragtest/input |

* 将文本复制到input文件夹下，目前仅支持编码为UTF-8的txt文件

**初始化**

|  |
| --- |
| PowerShell python -m graphrag.index --init --root ./ragtest |

初始化后将创建.env和settings.yaml 两个参数设置文件

**修改参数设置**

* 修改.env
* 将GRAPHRAG\_API\_KEY=<API\_KEY>修改为OPENAI\_API\_KEY=<你的API KEY>
* 修改settings.yaml
* 根据你使用的模型修改llm：和 embbeding： 两个部分
* 在claim\_extraction: 部分，设置 enabled: true

**INDEXING**

**Prompt Tune**

|  |
| --- |
| PowerShell python -m graphrag.prompt\_tune --root ./ragtest --config settings --output ./ragtest/prompts |

这里列出的为必填参数

--root: 工作目录

--config：.yaml 参数设置文件

--output: prompt保存路径

**Start Indexing**

|  |
| --- |
| PowerShell  python -m graphrag.index --root ./ragtest |

**Query**

详细query方法请参考官方notebook [Query Engine Notebooks](https://microsoft.github.io/graphrag/posts/query/notebooks/overview/)

**Global Search**

|  |
| --- |
| PowerShell python -m graphrag.query --root ./ragtest --method global "替换为具体Query语句" |

**Local Search**

|  |
| --- |
| PowerShell python -m graphrag.query --root ./ragtest --method local "替换为具体Query语句" |