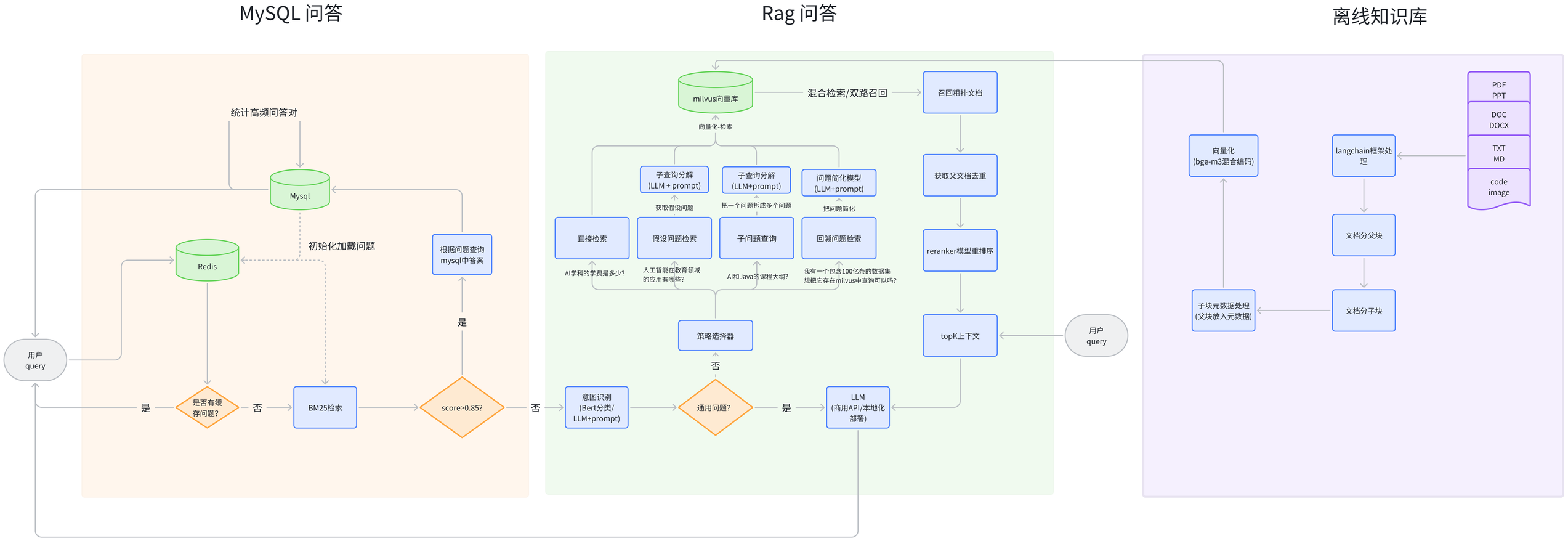

MySQL 问答
Rag 问答
离线知识库
统计高频问答对
Mysql
初始化加载问题
Redis
根据问题查询mysql中答案
是
用户query
是否有缓存问题?
是
否
BM25检索
score>0.85?
否
milvus向量库
混合检索/双路召回
召回粗排文档
向量化-检索
子查询分解(LLM + prompt)
子查询分解(LLM+prompt)
问题简化模型(LLM+prompt)
获取假设问题
把一个问题拆成多个问题
把问题简化
直接检索
假设问题检索
子问题查询
回溯问题检索
AI学科的学费是多少?
人工智能在教育领域的应用有哪些?
AI和Java的课程大纲?
我有一个包含100亿条的数据集想把它存在milvus中查询可以吗?
获取父文档去重
reranker模型重排序
topK上下文
用户query
策略选择器
否
意图识别(Bert分类/LLM+prompt)
通用问题?
是
LLM(商用API/本地化部署)
向量化(bge-m3混合编码)
langchain框架处理
PDF
PPT
DOC
DOCX
TXT
MD
code
image
文档分父块
文档分子块
子块元数据处理(父块放入元数据)