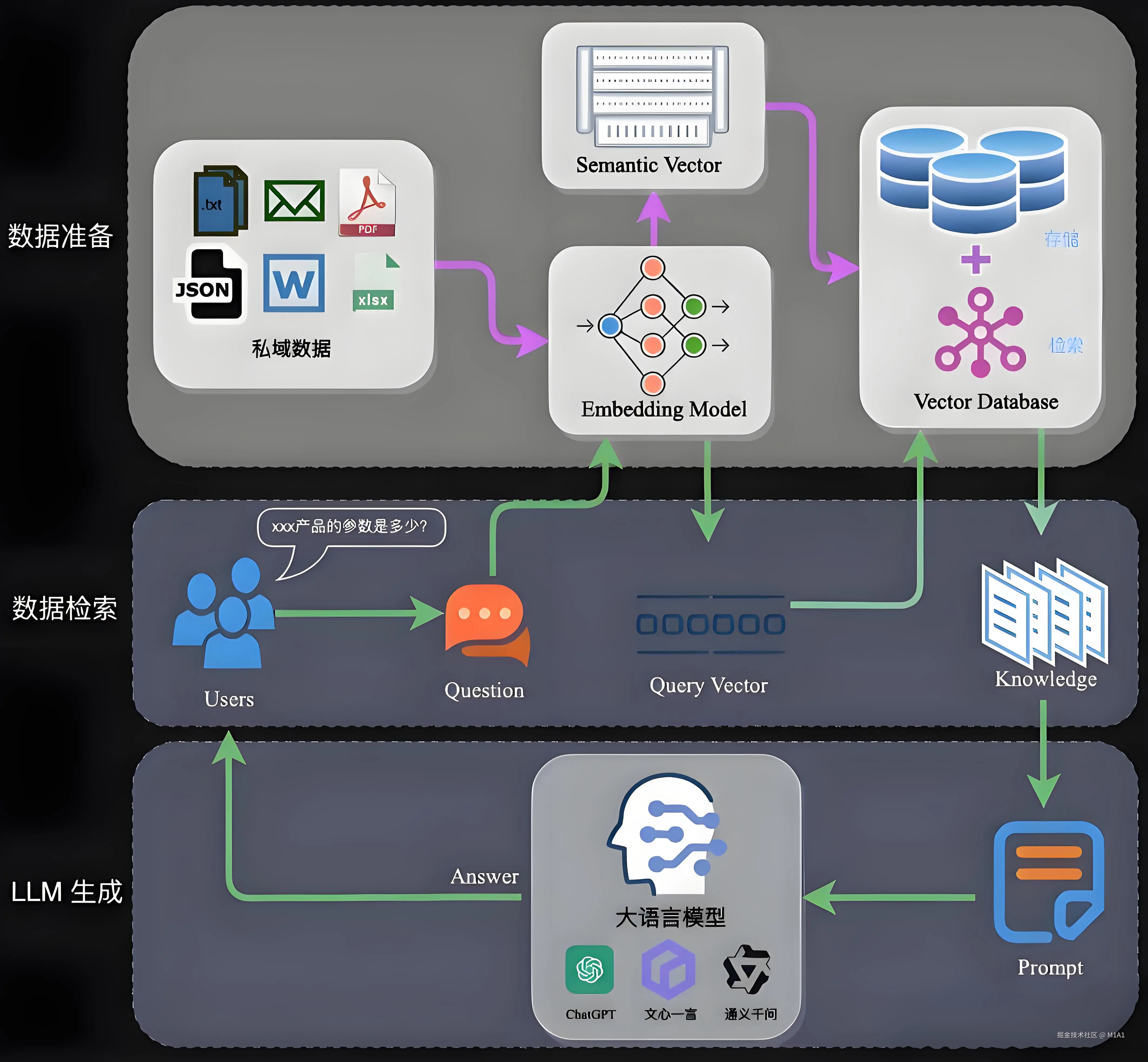
## RAG相关记录

### RAG工作原理



**RAG** 的主要流程主要包含以下 2 个阶段：

1. **数据准备阶段：** 管理员将内部私有数据向量化后入库的过程，**向量化**是一个将文本数据转化为向量矩阵的过程，该过程会直接影响到后续检索的效果；**入库**即将向量数据构建索引，并存储到向量数据库的过程。
2. **用户应用阶段：** 根据用户的 Prompt 提示词，通过检索召回与 Prompt 提示词相关联的知识，并融入到原 Prompt 提示词中，作为大模型的输入 Prompt 提示词，通用大模型因此生成相应的输出。

### ****Docling**** 是一个开源的以AI驱动的文档解析和处理工具

#### 主要功能

* **多格式支持**：支持读取和转换多种文档格式，让你无需担心不同格式的兼容性问题。
* **PDF 文档理解**：不仅能够提取 PDF 文档的文字内容，还能解析页面布局、阅读顺序和表格结构，提供深入的信息，完美还原数据。
* **统一格式**：使用统一的 DoclingDocument 格式来表示文档内容，方便后续处理和分析。
* **RAG / QA 应用**：可以轻松集成 LlamaIndex 和 LangChain 等工具，实现强大的 RAG / QA 应用。
* **OCR 支持**：支持 OCR 技术，可以处理扫描的 PDF 文件。
* **简单易用**：提供简单的 CLI 命令行界面，用户快速使用。