

[Cover](#)[Contents](#)[Transition](#)[Ending](#)

基于DoRA+GraphRAG的 公共艺术设计领域学习助手

问答系统实现方案





THE CO 目录 CONTENTS

01

研究路线



02

微调方案



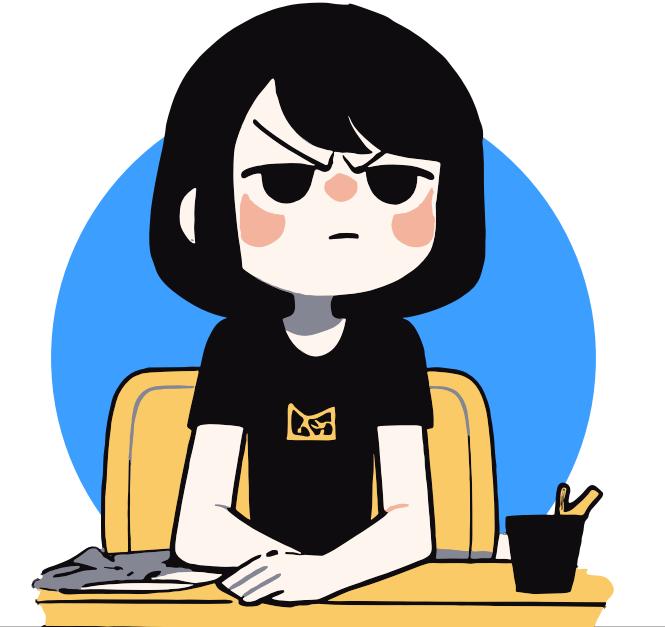
03

检索增强



04

项目管理





01 研究路线

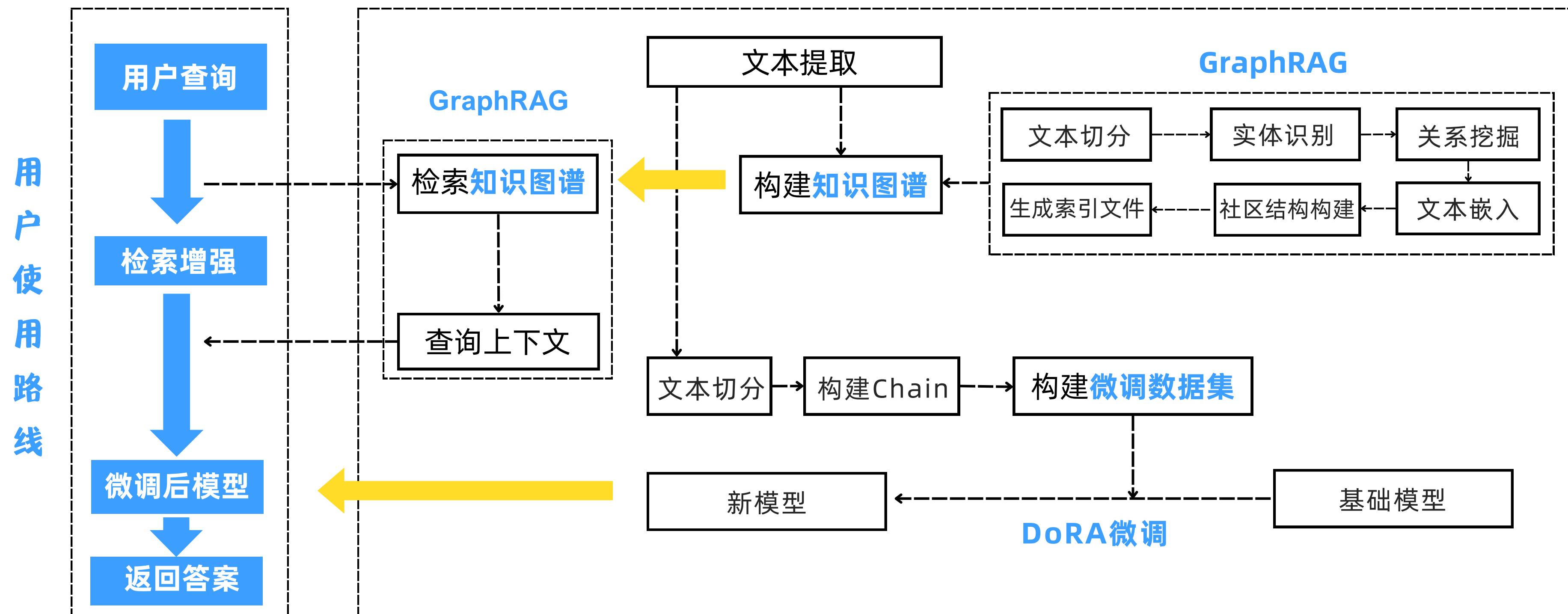
RESEARCH ROUTE





研究路线

技术实现路线





预期成果

需求分析

REQUIREMENT ANALYSIS

系统希望满足端到端的使用需求，拟开发桌面应用程序，主要分为**三个功能模块**，围绕**公共艺术问答**展开，配置问答记录导出功能；同时增加可视化设计，将推理过程通过知识图谱的形式进行展现，为用户提供公共艺术知识问答过程中的**内在逻辑展示**。



大模型问答模块

本模块将依托于经过检索增强及微调的本地大模型，专注于公共艺术垂类领域。用户可以输入自然语言形式的问题，系统将智能化地提供精准的答案。

知识图谱可视化模块

当系统响应用户提问时，会同时生成与问题相关联的知识图谱。用户可以选择查看这些关系图，以更加直观和深入地理解话题内容及其背后的关联结构。

问答记录导出模块

此模块旨在帮助用户管理和回顾其查询历史，主要功能包含生成问答记录和下载问答记录。归档用户的问答历程，并支持便捷的下载选项以便后续查阅。



预期成果

01 大模型问答界面

The screenshot shows the Yuxi-Know AI interface. On the left, there's a sidebar with icons for Conversation, Intelligent Body, Graph, and Knowledge Base. The main area has a search bar at the top with placeholder text "输入问题...". Below it are three buttons: "联网搜索" (Network Search), "知识图谱" (Knowledge Graph), and "食品营养知识库" (Food Nutrition Knowledge Base). A list of questions is on the left, and a detailed answer is shown on the right. The answer states: "根据查询到的资料, 可以抗辐射的水果有: 苹果、猕猴桃、胡萝卜、香蕉、黑莓、蓝莓、草莓、樱桃等。这些水果富含多种抗氧化剂和营养物质, 可以帮助减少辐射对机体的损害, 并促进放射性物质从体内排出。" A blue button at the bottom right says "哪些水果可以抗辐射?".

02 关系图展示界面

The screenshot shows the Yuxi-Know AI interface with a knowledge graph visualization. On the left, there's a sidebar with icons for Conversation, Intelligent Body, Graph, and Knowledge Base. The main area has a search bar at the top with placeholder text "输入问题...". Below it are three buttons: "联网搜索" (Network Search), "知识图谱" (Knowledge Graph), and "食品营养知识库" (Food Nutrition Knowledge Base). A list of questions is on the left. A large, colorful network graph is displayed on the right, representing relationships between various entities. A tab labeled "关系图" (Relationship Graph) is selected at the top right. A modal window titled "参考信息" (Reference Information) is open, showing a snippet of text from the graph: "根据查询到的资料, 可以抗辐射的水果有: 苹果、猕猴桃、胡萝卜、香蕉、黑莓、蓝莓、草莓、樱桃等。这些水果富含多种抗氧化剂和营养物质, 可以帮助减少辐射对机体的损害, 并促进放射性物质从体内排出。"



02

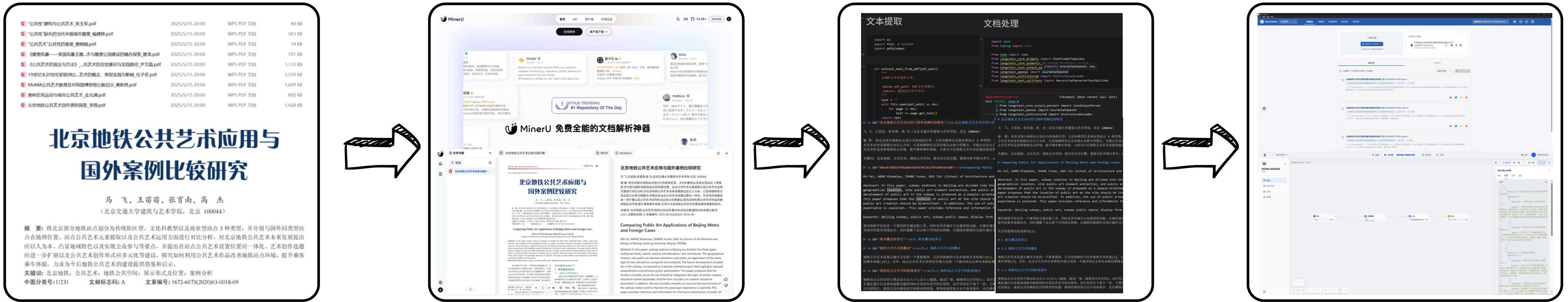
微调方案

FINE-TUNING





数据集构建



明 确 目 标

基于公共艺术设计领域相关的文本资料，通过大模型构造结构化的数据集，保存为 json 文件。

文本提取

利用 MinerU 工具 提取文
本内 容， 将 PDF 文件 转为
markdown 格式

文本切片

对比code、MinerU和Easy
Dataset工具
进行文本切片的效果
code：按照字段长度切片
MinerU&Easy Dataset：
按照章节切片

构建Chain

借助大模型的能力，通过**Prompt**工程根据文本形成问答对，对比**Easy Dataset**工具、**Dify**工作流的效果



构建Chain方法

01 Easy Dataset

北京地铁公共艺术应用与国外案例比较研究

马飞, 王苗苗, 张育南, 高杰
(北京交通大学建筑与艺术学院, 北京 100044)

摘要: 将北京部分地铁站点划分为传统街型、文化科教型以及商业型站点 3 种类型, 并分别与国外同类型站点在地理位置、站台公共艺术采取以及公共艺术运用等方面进行对比分析, 对北京地铁公共艺术未来发展提出以人为本、凸显地域特色以及实现公众参与等要素, 并提出在站点公共艺术设置位置一体化、艺术创作进站进一步扩展以及公共艺术创作形式多元等建议。研究如何利用公共艺术改善地铁站点环境, 提升乘客体验, 为今后地铁公共空间、展示形式以及艺术表现形式提供参考。

Comparing Public Art Applications of Beijing Metro and Foreign Cases
MA Fei, WANG Miaomiao, ZHANG Yunnan, GAO Jie
(School of Architecture and Design of Beijing Jiaotong University, Beijing 100044)

Abstract: In this paper, subway stations in Beijing are divided into three types: traditional block, culture and science, and commerce. The geographical location, station platform public art adoption, and public art application are compared and analyzed between domestic and foreign counterparts. This paper proposes that the location of public art at the site should be integrated, the topic of artistic creation should be further expanded, and the form of public art creation should be diversified. In addition, the use of public artworks to improve the environment of the subway station and to improve the passenger experience is explored. This paper provides reference and information for the future construction of public art in the subway.

Keywords: Beijing subway; public art; subway space; display form and location; case study

现代地铁不仅是一个便捷的交通工具, 同时也是城市文化的重要组成部分。
对也具有城市文化象征意义。在城市建设中, 地铁站空间已成为展示城市文化的重要场所¹。地铁站空间除了为乘客提供便利外, 还能通过其独特的展示形式来提升乘客的乘车体验, 同时也丰富了乘客的乘车体验²。另外, 站点公共艺术多样化的呈现方式将一个城市的过去和未来形象展现给大众面前, 在唤起人们城市记忆的同时, 也增强了人们所在城市现代性的认识。

发表日期: 2019-04-21 | 收录日期: 2019-03-29
第一作者: 马飞, 男, 硕士研究生, 从事轨道交通与艺术设计研究与应用的研究, 1610213321@qq.com
基金项目: 北京市社科基金项目(14GZ20006)

(C)1994-2023 China Academic Journal Electronic Publishing House, All rights reserved. http://www.cnki.net

选取示例

文本较长
有图有表

上传md文件
进行切块

文本块
根据问题
生成问答对

数据集展示

```
[{"instruction": "北京地铁站点被划分为哪三种类型进行公共艺术对比研究?", "input": "", "output": "<think>北京地铁站点公共艺术对比研究的分类依据主要基于站点周边区城功能属性, 首先分析城市空间结构中地铁站点", "system": ""}, {"instruction": "在比较北京与国外同类型地铁站点时主要分析了哪些方面?", "input": "", "output": "<think>在比较北京与国外同类型地铁站点时, 分析过程遵循系统性分层框架: \n\n首先聚焦**地理位置功能维度**", "system": ""}, {"instruction": "公共艺术在地铁站点的作用可从空间功能和文化表达两个维度展开分析。首先从物理环境角度思考, 地下空间天", "input": "", "output": "<think>公共艺术在地铁站点的作用可从空间功能和文化表达两个维度展开分析。首先从物理环境角度思考, 地下空间天", "system": ""}, {"instruction": "未来北京地铁公共艺术发展需要重点遵循哪些核心原则?", "input": "", "output": "<think>未来北京地铁公共艺术发展的核心原则可通过系统性分析得出: \n\n首先需要深入理解地铁作为城市重要公共", "system": ""}, {"instruction": "提升乘客乘车体验可以通过哪些具体的公共艺术手段实现?", "input": "", "output": "<think>提升乘客乘车体验的公共艺术手段可从空间构、内容延展、形式创新三个维度系统推进: \n\n首先需要解决地", "system": ""}, {"instruction": "站点公共艺术设置一体化建议的具体内涵是什么?", "input": "", "output": "<think>站点公共艺术设置一体化的核心内涵, 需要从空间规划与艺术设计的耦合关系切入分析, 首先需明确该建议", "system": ""}],
```

便捷

耗时、费钱
生成问答对
质量一般

02 Dify工作流

大模型
提示词

选取示例

文本较长
有图有表

搭建工作流

数据集展示

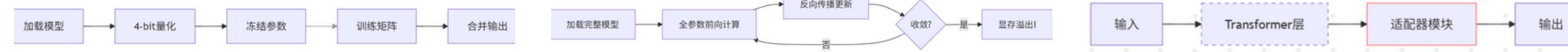
```
[{"block_id": "北京地铁公共艺术应用与国外案例比较研究", "questions": [ {"question": "该研究的核心目标是什么?", "answer": "研究旨在通过比较北京地铁三种类型站点(传统街区型、文化科教型、商业型)与国外同类型站点在公共艺术方面的差异, 探讨其背后的原因, 为今后地铁公共空间、展示形式以及艺术表现形式提供参考。"}, {"question": "研究中对比分析的主要维度有哪些?", "answer": "研究从三个主要维度进行对比: ① 地理位置与区域文化特征; ② 站点公共艺术元素的提取方式; ③ 公共艺术在地铁站点中的具体应用与效果。"}], {"block_id": "1 相关概念的界定", "questions": [ {"question": "地铁公共艺术的核心功能是什么?", "answer": "地铁公共艺术具有双重功能: 文化功能(展示城市文化、传承历史记忆、增强城市认同感)和空间功能(提升乘客乘车体验、美化环境、增加趣味性)。"}, {"question": "地铁公共艺术的主要表现形式有哪些? 请举例说明。", "answer": "根据原文, 地铁站点公共空间分为6个部分: ① 出入口及临近广场; ② 需具备识别性与过渡功能; ③ 入口广场及附属设施; ④ 站厅; ⑤ 站台; ⑥ 出入口通道。主要表现形式包括: ① 壁画、雕塑(如北京南锣鼓巷站的琉璃墙画); ② 现代形"}, {"question": "地铁站点公共空间由哪些主要部分组成? 各部分的艺术设计重点有何不同?", "answer": "根据原文, 地铁站点公共空间分为6个部分: ① 出入口及临近广场; ② 需具备识别性与过渡功能; ③ 入口广场及附属设施; ④ 站厅; ⑤ 站台; ⑥ 出入口通道。各部分的艺术设计重点不同, 如站厅可能强调空间的开阔感和视觉引导, 而出入口则可能强调识别性和安全性。"}], {"question": "地铁公共艺术的主要表现形式有哪些? 请举例说明。", "answer": "根据原文, 地铁站点公共空间分为6个部分: ① 出入口及临近广场; ② 需具备识别性与过渡功能; ③ 入口广场及附属设施; ④ 站厅; ⑤ 站台; ⑥ 出入口通道。主要表现形式包括: ① 壁画、雕塑(如北京南锣鼓巷站的琉璃墙画); ② 现代形"}, {"question": "地铁公共艺术的表现形式如何与周围环境融合?", "answer": "根据原文, 地铁站点公共空间分为6个部分: ① 出入口及临近广场; ② 需具备识别性与过渡功能; ③ 入口广场及附属设施; ④ 站厅; ⑤ 站台; ⑥ 出入口通道。主要表现形式包括: ① 壁画、雕塑(如北京南锣鼓巷站的琉璃墙画); ② 现代形"}, {"question": "地铁公共艺术的表现形式如何与周围环境融合?", "answer": "根据原文, 地铁站点公共空间分为6个部分: ① 出入口及临近广场; ② 需具备识别性与过渡功能; ③ 入口广场及附属设施; ④ 站厅; ⑤ 站台; ⑥ 出入口通道。主要表现形式包括: ① 壁画、雕塑(如北京南锣鼓巷站的琉璃墙画); ② 现代形"}], {"question": "地铁公共艺术的表现形式如何与周围环境融合?", "answer": "根据原文, 地铁站点公共空间分为6个部分: ① 出入口及临近广场; ② 需具备识别性与过渡功能; ③ 入口广场及附属设施; ④ 站厅; ⑤ 站台; ⑥ 出入口通道。主要表现形式包括: ① 壁画、雕塑(如北京南锣鼓巷站的琉璃墙画); ② 现代形"}]
```

问答对
质量可控
耗时、费钱
生成问答对少



微调方法对比

图示



概述

4-bit量化的LoRA变种，冻结原模型参数仅训练低秩矩阵

完全微调所有模型参数

在Transformer层间插入轻量适配模块，支持多任务插拔

基于方向和幅度分解的LoRA改进，适合垂直领域知识注入

优势

1. 显存需求最低 (8-10GB)
2. 训练速度快
3. 易与HuggingFace集成

1. 理论性能最优
2. 无需架构修改
3. 支持任意领域适配

1. 理论性能最优
2. 无需架构修改
3. 支持任意领域适配

1. 术语准确率最高 (+5%)
2. 参数效率高
3. 支持动态调整知识注入强度

劣势

1. 知识记忆能力较弱
2. 复杂推理准确率下降
3. 无法直接扩展领域词表

1. 显存需求高
2. 训练耗时 (7B模型需7天+)
3. 易过拟合学生提问数据

1. 单任务性能低于LoRA
2. 增加10%推理延迟
3. 对艺术术语敏感度不足

1. 实现复杂度、计算量较高
2. 需要人工选择模块
3. 需调整幅度系数初值



微调策略

01 领域数据定向强化

从现有资料中提炼核心术语、作品关系与概念等数据并格式化，确保模型能够更好地理解。同时加入一些“扰动样本”训练模型区分对错，提升鲁棒性。

03 混合训练优化

通过混合精度和梯度累积将有效批次进行提升，扩展专业术语，确保模型对公关艺术领域关键信息的记
忆强度超过普通词汇



02 DoRA模型配置

定向增强领域知识记忆，在预训练模型的注意力层嵌入DoRA模块，将参数更新分解为方向矩阵和幅度系数——前者学习关联模式，后者控制知识注入强度

04 效果评估和模型部署

训练后，评估模型在术语识别和内容准确性方面的表现，然后对模型进行压缩优化，并部署到高效的推理系统中，确保运行时的快速稳定。



03

检索增强

RETRIEVAL-AUGMENTED

GENERATION





RAG方法对比

	GraphRAG	AgenticRAG	传统RAG
数据结构	知识图谱（节点、边）	向量数据库（嵌入向量）	向量数据库（嵌入向量）
检索方式	图查询（遍历、社区检测）	通过Agent管理检索流程	相似性搜索（余弦相似度）
适用场景	复杂关系查询，长文本	复杂，多步骤任务	简单关键词匹配，短文本
生成优化	单一生成， 但比传统RAG更精准全面	自我反思与迭代优化	单一生成，无反馈循环
支持格式	多模态	多模态	主要文本
计算成本	高	高	低

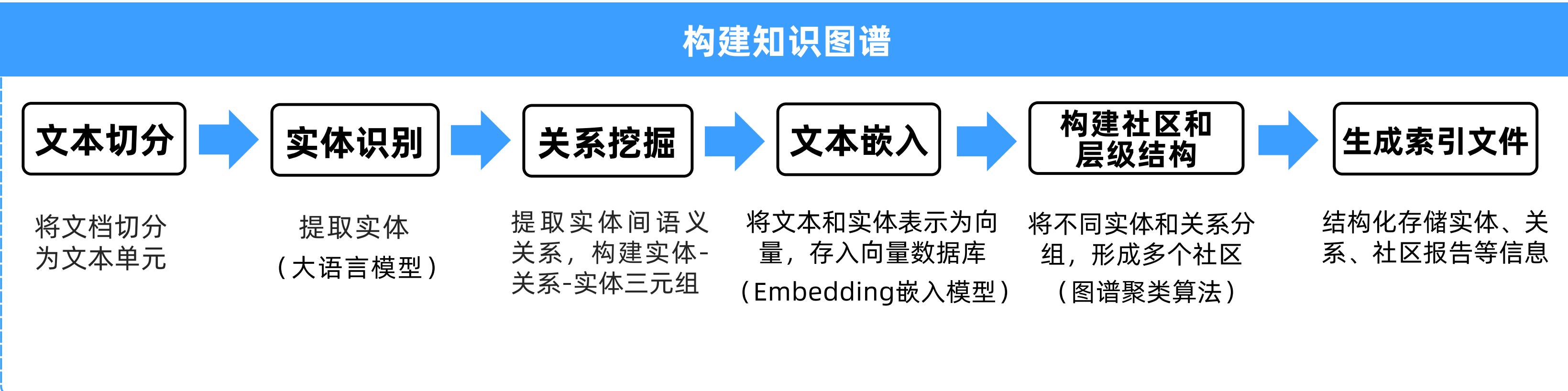


GraphRAG

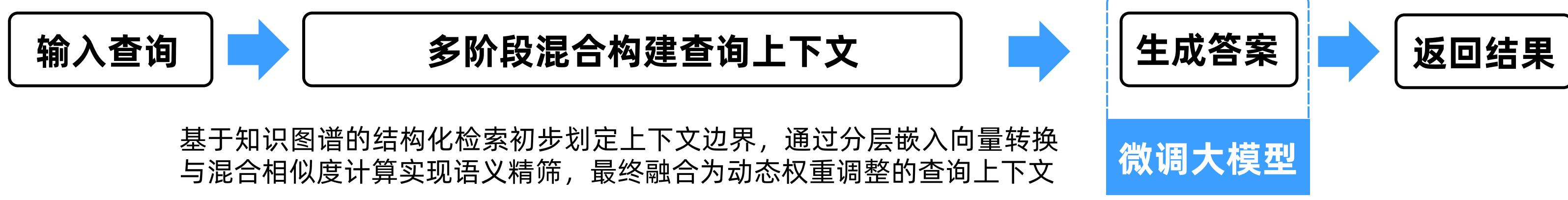
选择[微软开源项目GraphRAG](#)

GraphRAG能够实现全本地化部署
对比阿里KAG无需人工构建Schema

索引



查询





本地化

ollama+

大模型运行环境

- 创建虚拟环境
- 安装魔搭社区
- 安装ollama
- 安装ollama Python SDK
- 启动ollama
- 创建模存放型文件夹

文件名 ^	大小	类型
.autodl		文件夹
graphrag		文件夹
graphrag-ollama		文件夹
kotaemon		文件夹
ml-ollama		文件夹
models		文件夹

```
conda create --name graphrag-ollama python=3.10
conda init
source ~/bashrc
conda activate graphrag-ollama
conda install jupyterlab
conda install ipykernel
python -m ipykernel install --user --name graphrag-ollama --display-name "Python (graphrag-olla
```

修改微软 GraphRAG源码

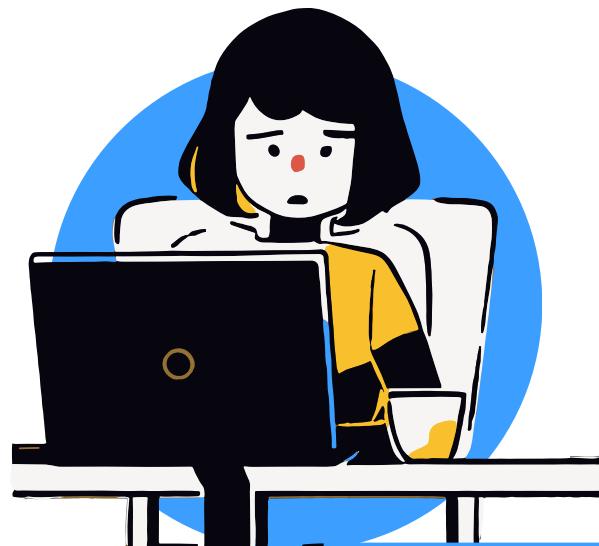
- graphrag快速安装
 - 创建检索项目文件夹
 - 上传数据集[.txt]
 - 初始化项目文件
- ```
graphrag init --root /root/autodl-tmp/graphrag-ollama/ml-ollama
```
- 修改项目配置
    - 模型的上下文长度
    - PARAMETER添加配置
  - 修改项目源码
    - openai\_embedding
    - s\_llm.py
    - embedding.py

## 将Ollama接入 微软GraphRAG

- 基于本地Ollama的GraphRAG Index过程
  - 在构建图谱的时候使用高性能模型[若模型本身性能较弱，则在对文本进行索引时有时会报错]
- 基于本地Ollama的GraphRAG query过程
  - 命令行对话
  - 基于Python SDK的对话



# 效果测试



## 01 模型层 (DoRA微调)

- 核心指标:

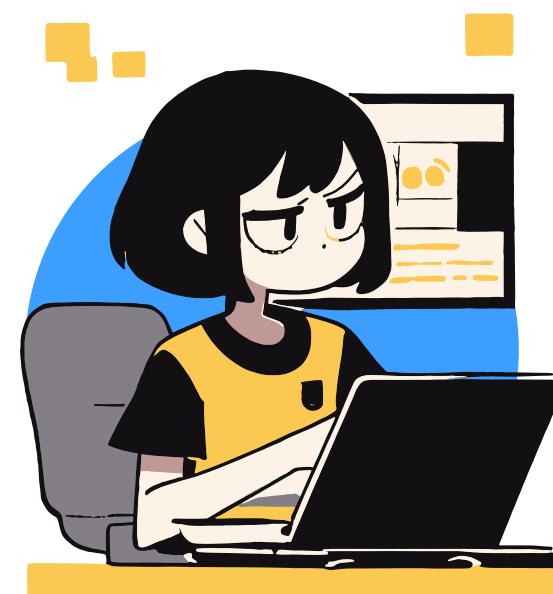
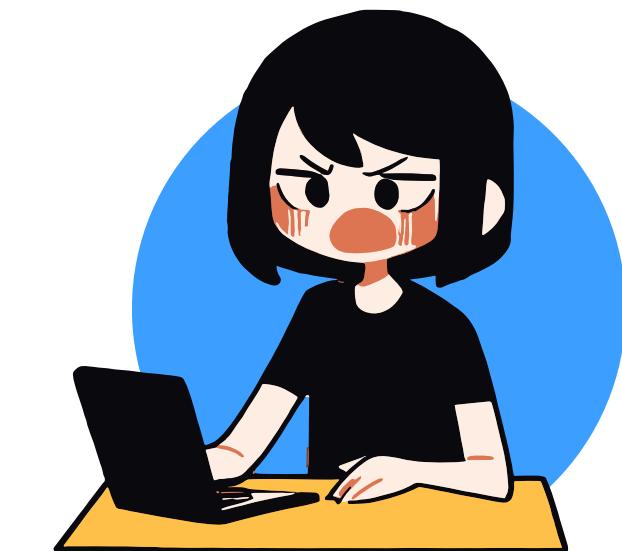
- 冷门知识召回率 - 构建专业测试集-F1混淆矩阵
- 参数效率比 - 寻找参数-精度平衡点



## 02 检索层 (GraphRAG)

- 核心指标:

- 实体覆盖率 - 随机抽取1k个术语，检查图谱节点
- 查询相关性&多跳查询 - 人工核验



## 03 端到端问答层

- 核心指标:

- 事实准确性 - 自动校验+专家打分 (A/B test)
- 响应效率 - 系统测试



04

# 项目管理

PROJECT MANAGEMENT







[Cover](#)[Contents](#)[Transition](#)[Ending](#)

讲完啦  
(:^)

感谢观看呀

