



ModelScope
魔搭社区



Datawhale

AI 2026 Hackathon Tour

矩巢 MatriXNest 小队

爱尔兰科克大学博士后：何彪

浙江大学博士生：张振亮

浙江大学博士生：王坚



01 矩巢项目简介



ModelScope
魔搭社区



Datawhale



MatriXNest (MxN)

Turn Documents into an AI Productivity Engine

RAG-Powered Intelligent Retrieval - Read, Reason, and Answer from Your Documents



Chunking & Embedding

MxN
Semantic
Vector Engine



Engineering Cost
Tunnel Estimation



Medical & Health
Clinical / Pathology QA

What is the quota standard for tunnel lining?

According to Clause 6.3.4 of the Tunnel Design Specification, the lining thickness and reinforcement ratio must comply with Grade II surrounding rock requirements. Relevant tables and calculation references are provided below.

+10x

Retrieval Efficiency

All-In-One Solution to Enhance Retrieval Efficiency and Answer Accuracy

本项目构建了一套 **基于 RAG 架构的智能文档存储与精准检索工具**，利用大模型对行业技术指南、规范与报告文档进行语义向量化表示与高效索引，突破传统关键词检索在 **专业文档场景下精度低、上下文割裂的问题**，实现面向工程与技术应用的高可靠信息检索与问答支持

特点与优势

- ◆ 语义级精准检索
- ◆ 面向专业场景的 RAG 架构
- ◆ 显著提升资料复用效率

4 知识沉淀与传承:

03 AI 创新性说明



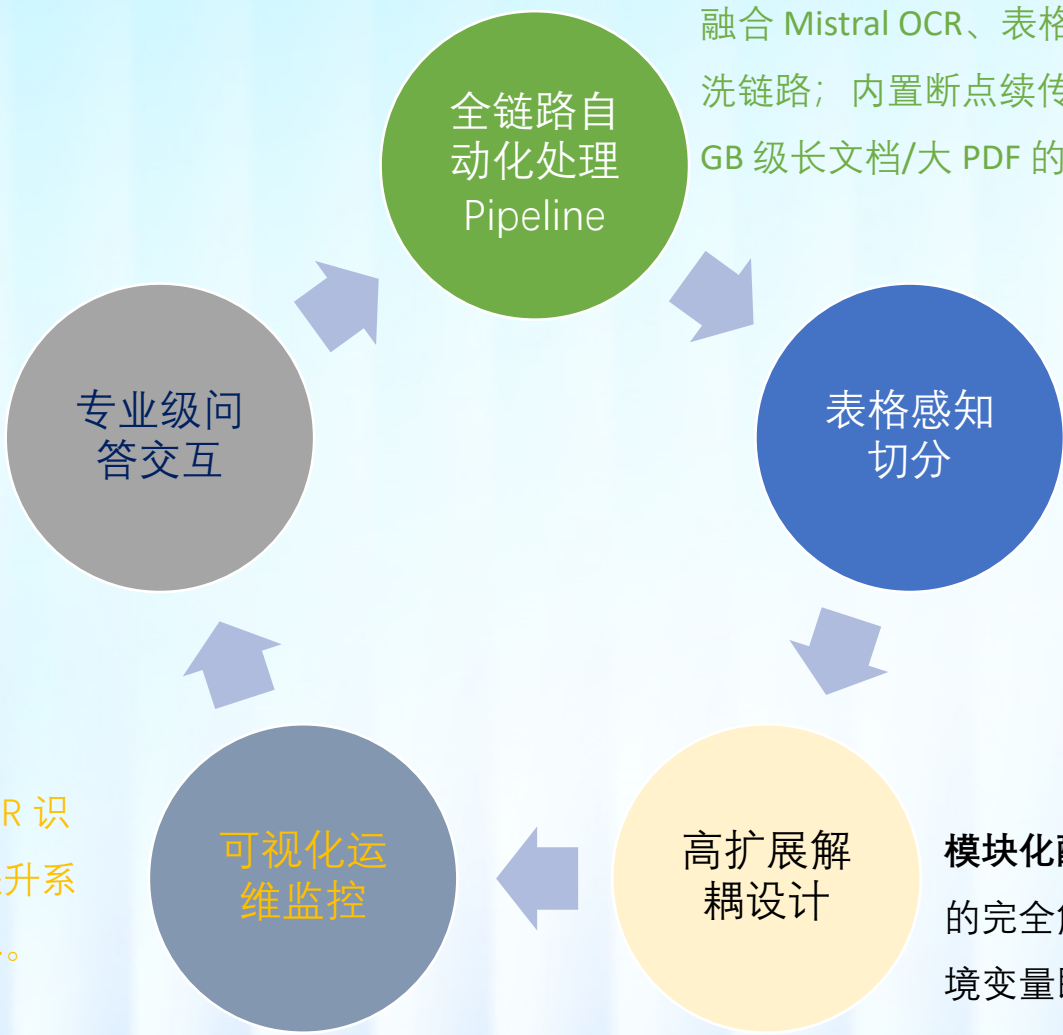
ModelScope
魔搭社区



Datawhale

- ◆ 可信溯源设计：基于 Streamlit 打造沉浸式前端，定制 Prompt 模板，问答注明页码、表格来源，并提供“未找到”时的严谨反馈；
- ◆ 业务探索辅助：侧边栏实时展示索引统计数据，并提供“定额查询”等典型示例问题。

质量速查： 允许开发者快速审查 OCR 识别精度与切块逻辑的合理性，大幅提升系统上线后的可解释性与运维调试效率。



融合 Mistral OCR、表格语义切分与向量化入库，构建一站式数据清洗链路；内置断点续传与批量嵌入 (Batch Embedding) 机制，保障 GB 级长文档/大 PDF 的稳定解析。

- 结构化保留 TableAwareChunker，智能检测跨页长表格并自动合并；
- 上下文增强：在切块时自动注入表头与章节层级信息，配合重叠策略，确保 RAG 检索时信息不断层、逻辑不丢失；
- 精准定位：引入区段标记技术，检索时精准区分“表内数据”与“正文条款”。

模块化配置： 实现 OCR、Embedding、Chat 模型的完全解耦。基于 大模型统一接口，仅需修改环境变量即可无缝切换模型版本。



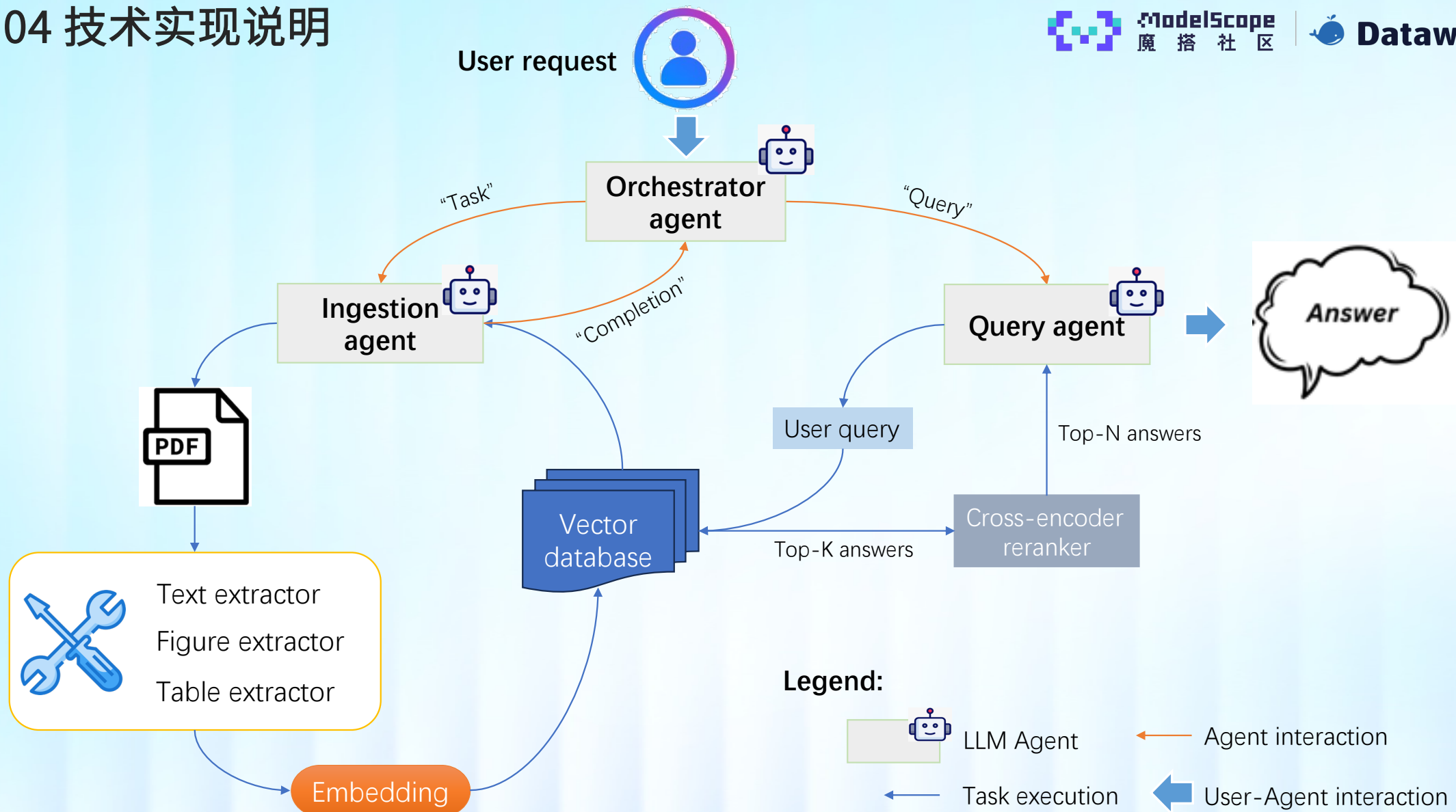
04 技术实现说明



ModelScope
魔搭社区



Datawhale



MatriXNest 矩巢智能体核心架构设计图



05 Demo 视频展示

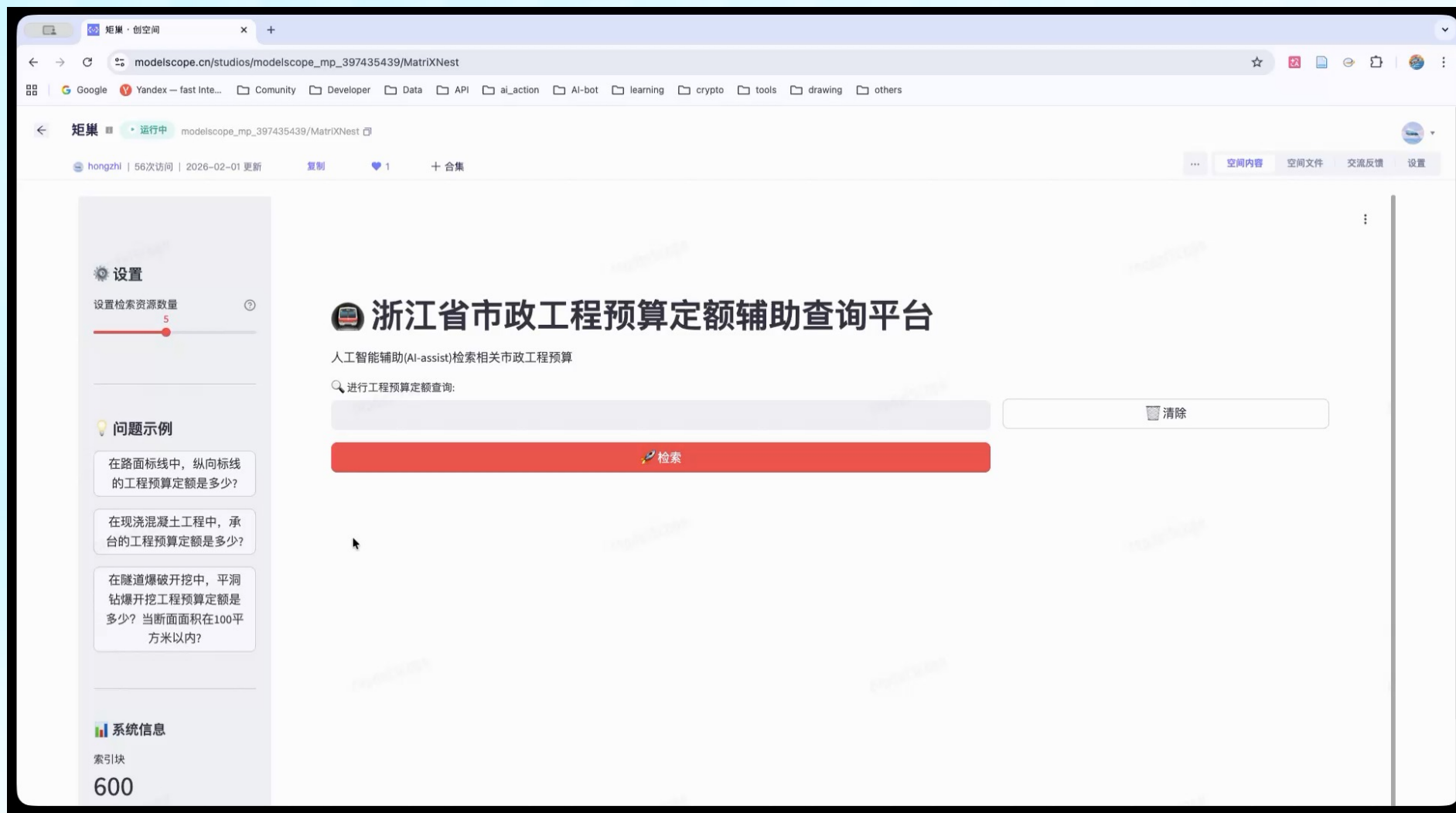


ModelScope
魔搭社区



Datawhale

说明：Demo基于MatriXNest架构以工程造价中市政预算编制智能体为例进行演示



Thank you

AI 2026 Hackathon Tour

矩巢 MatriXNest 小队

爱尔兰科克大学博士后：何彪

浙江大学博士生：张振亮

浙江大学博士生：王坚

