

2024年终工作总结

陈兴

目录

- 工作内容
- 工作总结
- 新年计划

工作内容

活动组：

1. IIm活动：参与机构1300家，参与馆员1700位，宣传了公司大模型平台；
2. 健康之屋：医疗行业的AI版微刊，负责其中的积分商城开发；
3. 公共组件：红包、抽奖组件增加新的配置规则，已接入的活动类型有：小游戏、读书榜V6、元宵元宇宙；
4. 元宵元宇宙V2：在V1的基础上增加了多种游戏和NPC互动，趣味性和可玩性更高；
5. 老项目维护：答题V4、趣味答题、VR入馆培训等等。

数据组：

1. RAG流程：大段落切分小段落，提升召回的准确率；
2. Flink技术预演：sql任务、python任务、java任务、CDC同步；
3. 索引向量化：对比分词搜索，语义相关的检索能力更强；
4. 消费mq入库es：订阅binlog，实现业务表的实时向量化。

工作小结

做的好的：

- 1, robot_qa项目索引别名配置改成使用es别名机制，并无缝替换原有逻辑，使其不再依赖redis服务。
- 2, Flink项目的依赖包dolphinscheduler-datasource-clickhouse的实现。

做的不好的：

吉林师范大学的入馆培训活动，在对接完成后实际使用的过程中，发现经常有用户的读者证号填写错误，老师每次反馈后就需要手动修正。如果在对接时让他们提供一个校验读者证的接口就好了。（反思：外部对接时需要提前做好需求分析，想到各种可能出现的场景。）

新年计划

1. 保证所负责项目的稳定运行和需求迭代;
2. 向量化检索, 实现百万级向量的快速检索;
3. 学习基于大模型的应用层开发, 例如AI代码生成, Claude的Model Context Protocol等等。

谢谢大家