# 赛题10 金融年报问答系统

（本赛题仅限于本科生选择）

**命题企业介绍**

恒生电子是一家以“让金融变简单”为使命的金融科技公司，总部位于中国杭州。1995年成立，2003年在上海证券交易所主板上市（600570.SH）。

恒生聚焦金融行业，致力于为证券、期货、基金、信托、保险、银行、交易所、私募等机构提供整体解决方案和服务。恒生已连续16年入选FinTech100全球金融科技百强榜单，2023 年排名第22位，位列中国上榜企业第一。目前公司拥有超过13000名员工，其中产品技术人员占比约73%。恒生自创立便以技术服务为核心，注重科技创新，每年以35%以上营业收入投入研发，同时重视技术人才培养和知识产权保护，2022年获批设立国家级博士后科研工作站。公司一直专注在金融领域深耕细作，凭借多年金融IT建设经验，以及对金融业务的深刻洞察和理解，并依靠“产品领先”战略，用优质的产品与服务，持续赋能金融机构创新发展。

**1、项目背景说明**

**【项目的行业背景】**

上市公司每年发布的一份企业年报包含了公司的财务状况，业务表现，市场趋势等关键信息。通过分析年报，可以预测公司的未来发展趋势，例如收入增长，利润增长，市场份额等。投资者可以了解公司的财务状况，业务表现，市场趋势等，从而做出更明智的投资决策。通过分析竞争对手的年报，公司可以了解竞争对手的策略，优势，弱点等，从而制定更有效的竞争策略。随着大模型技术的发展，模型问答能力的不断提升，基于大模型技术提升年报问答的自动化分析和问答水平，为投资决策提供智能化，专业化的建议，在智能投研、投顾领域具有非常大的应用前景和价值。

**【项目的客户背景】**

恒生电子通过AI赋能金融客户群体（包括银行、券商、基金等）办公、生产流程智能化升级，为客户提供智能化软件与数据服务平台。基于恒生金融大模型为金融客户提供智能化软件提效、AI软件驾驶舱、金融数据智能问答助手等在内的一体化金融大模型服务。

**【项目的业务背景】**

构建一套完整的上市公司年报知识问答系统，能够对用户关于年报中的任意的内容进行准确快速答复，包括但不限于年报中的报表、文字内容，同时能够支持不同公司、不同年份的财务指标数据进行对比分析，并给出专业性分析。同时针对给出的答复能够提供对应的出处。

**2、项目说明**

**【问题说明】**

年报内容分析是一项费时费力的大工程，一般年报篇幅均有几百页，内容涵盖公司基本信息、财务报告、经营情况、治理报告、社会责任、审计报告以及重要事件等内容，组成上以图表和文本内容为主。结合大模型能力，期望构建一种基于年报知识库问答系统能够对用户关注的年报内容进行精准回答，同时能够结合年报中的内容给出合理的分析建议。

**【用户期望】**

开发一个基于大模型的年报问答助手，能够回答针对年报中的任意

内容（表格、文本）进行精准答复，同时能够对多个公司、多年份的年报内容进行对比分析。金融从业人员以及金融投资者能够基于该系统快速获取上市公司的基本信息、财务、经营、重大事件等内容，**支持不同公司、不同年份的财务、事项内容的分析对比**。同时问答助手需要给出每次生成答案的参考来源。

1. **任务要求**

**【内容要求】**

基于上述期望构建一个年报问答助手，该系统一般包括以下几个核心流程模块：

**1)知识库构建**

针对年报内容进行精准解析和存储。

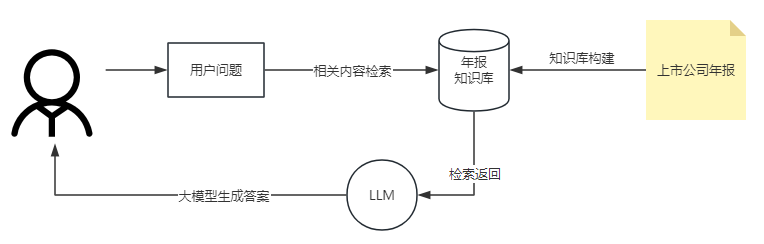
**2)问题检索**

针对用户问题，获取相关的年报内容，包括各种检索策略的设计。

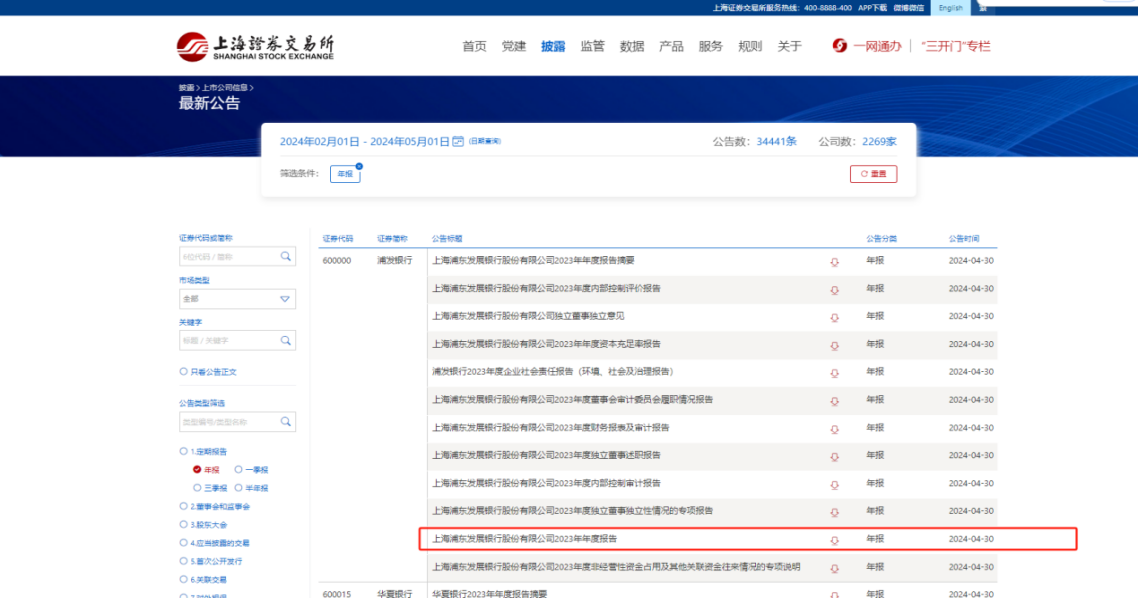
**3)答案合成与参考来源**

针对检索的内容以及问题使用大模型合成最终用户问题的**答案**，并提供答案**参考的原始内容**。

**一般流程 ：**



**年报范围：近5年**（2019-2023年）上海证券交易所上市的公司年报文件，下载地址可访问<https://www.sse.com.cn/disclosure/listedinfo/announcement/>。



下载“XXX公司XX年年度报告”。

**问题示例：**

[上海浦东发展银行](https://static.sse.com.cn/disclosure/listedinfo/announcement/c/new/2024-04-30/600000_20240430_MWRU.pdf)近三年的营收如何？

恒生电子IT人员有多少？

[上海浦东发展银行](https://static.sse.com.cn/disclosure/listedinfo/announcement/c/new/2024-04-30/600000_20240430_MWRU.pdf)与[华夏银行](https://static.sse.com.cn/disclosure/listedinfo/announcement/c/new/2024-04-30/600015_20240430_EJQ1.pdf)去年谁的净利润高？

**【格式要求】**

开发语言：python/Java等

交付结果：产品规划文档、 技术架构文档、源代码、产品原型系统说明文档、演示demo

效果评价： 问答准确率>85%，响应速度30S以内

（但不限以上资料，其他认为有必要的资料都可以递交）

**【其他说明】**

项目中涉及的大模服务可以使用私有化部署的模型服务、SaaS以及其他商业化模型服务，支持私有化场景模型精调（需提供精调方案）。