**重庆市首届高校AI大模型创新应用大赛**

**企业出题赛道组织方案**

**（重庆联通出题：基于联通元景+Deepseek人工智能知识库的自动化构建）**

# 一、比赛背景

当前，人工智能大模型已成为全球科技竞争的战略制高点，正以前所未有的速度重塑全球科技产业格局。以深度求索（DeepSeek）等为代表的中国AI企业强势崛起，持续推动全球人工智能技术向更高能级跃升。高等院校作为科技第一生产力与人才第一资源的交汇点，亟需在人工智能技术突破、产业创新应用及传统产业智能化升级中担当核心引擎。为全面落实国家创新驱动发展战略，精准对接重庆市"416"科技创新战略布局与"33618"现代制造业集群体系建设需求，重点培育具备前沿洞察力与创新领导力的复合型高端人才，为科技强国建设提供核心智力支撑。

# 二、比赛目标

本次比赛围绕大模型时代人工智能理论创新和实践应用主题，旨在为重庆市高校学生搭建基于AI大模型应用的创新实践平台，以此激发青年学子创新活力，推动AI技术的持续进步；促进产学研融合，推动AI技术与产业发展深度融合；培养具有深厚人工智能素养和卓越创新能力的未来人才，为AI产业发展提供智力支持和人才保障。

# 三、竞赛方式

## 3.1队伍设置

参赛对象：全市高校在校大学生、教师。

组队要求：每个团队成员3-10人，允许跨校组建、师生共同组建参赛团队。

指导教师：每个参赛队伍可配备1-3名指导教师。

## 3.2组织机构

1.大赛设置评审委员会，负责项目评审等工作。

2.大赛设置纪律与监督委员会，负责对赛事组织、项目评审、协办单位等进行监督，对违反大赛纪律的行为予以处理。

## 3.3比赛赛制

本赛项采用初赛决赛两轮赛制，参赛队伍需要根据竞赛命题要求完成作品，初赛在规定时间前提交作品，经评选后选出初赛提交作品队伍的前30%进入决赛，进入决赛后可在规定的时间内继续完善作品提交，并在现场演示完善后的决赛作品。

# 四、竞赛内容

本赛题要求参赛团队需基于大小模型协同的新一代人工智能技术，设计并开发一个领域知识库自动化构建平台，形成一套自动化构建知识库的工具；数据来源包括互联网，本地数据库，本地文档，实现多源数据收集，并通过文本清洗（查重、纠错等）、文本分段、向量化等技术完成知识库的构建。该平台能够支持对docx、pdf、doc、md、excel等格式文件的内容进行智能化文档语义分割，实现对一篇文档按照各个标题和段落层次进行切分，可以建立专门的AI模型实现该功能。该平台需具备友好的用户交互界面，支持知识片段的查重，支持网页抓取参数的配置，文本清洗规则的设置，文本分段基础参数的配置，文本向量化过程的展示，实现全自动化知识库构建。

## 4.1 竞赛题目

联通元景+Deepseek人工智能知识库自动化构建

## 4.2 作品要求

**（1）功能模块要求**

文档结构解析：实现对docx/pdf/doc/md/excel等文件的标题层级识别；段落语义切分：基于上下文连贯性分析进行段落智能划分（段落边界识别误差率≤3%）

交互式配置系统：网页抓取参数模组；提供正则表达式编辑器、停用词库管理、敏感词过滤模块；分段参数矩阵：开放段落长度阈值、语义连贯性阈值等可调参数。

可视化工作台：向量化过程监视，动态展示文本向量化进度与预估完成时间；知识查重看板：提供知识片段的查重与溯源功能。

**（2）系统基础要求**

**稳定性**：保证系统基本运行流畅，无明显卡顿或崩溃

**可靠性**：实现关键数据的手动备份与恢复功能

**安全性**：基础数据加密存储  
用户账号权限分级管理

**（3）成果提交要求**

一、技术文档

1. 解决方案技术报告（15-20页）

系统架构设计说明

核心算法实现方案

关键技术难点与创新点

测试验证结果分析

1. 用户使用手册

平台安装配置指南

功能使用说明

常见问题解决方案

二、系统实现成果

1. 可运行程序

完整安装包（支持Windows/Linux）

或可访问的Web系统演示地址

1. 源代码

核心功能模块完整代码

必要的第三方依赖说明

基础注释说明（关键函数/模块）

三、演示材料

1. 系统演示视频（5-8分钟）

主要功能完整演示

核心操作流程展示

典型应用场景示例

1. 测试数据集

用于验证的示例文档集（含各支持格式）

测试结果样例

## 4.3 模型推荐

联通元景

Qwen2.5-14B-Instruct

DeepSeek-R1-Distill-Qwen-14B

# 五、赛题资源

数据支持：政府公开数据，城市管理相关数据，包括但不限于国家部委和省级政府部门公开的涉及市政设施、园林绿化、城市供水供电等领域的行政规范性文件及政策解读；医疗卫生相关数据，包括但不限于国家部委和省级政府部门公开的医疗服务、公共卫生、健康管理与保健、医疗保障与支付、医疗信息化等领域的行政规范性文件及政策解读。

专家支持：提供技术专家支持，参赛团队可以在开发过程中获得技术指导和问题解答。

计算资源和技术支持：相应算力资源(每个参赛团队一个账号)，为期3个月（6-8月），入围决赛团队增加1个月。

联通元景

# 六、竞赛流程

## 6.1 时间安排

具体时间请及时关注大赛官网和QQ群（咨询QQ群：1041935478）通知。

## 6.2 流程安排

1、报名

请登录竞赛官网（https://5cqs3u4l.mh.chaoxing.com）填写报名信息。

2、培训

培训安排将在比赛官方QQ群发布。

3、初赛

参赛队伍报名后，根据赛题要求完成作品，参赛队伍需在截止时间前将作品提交至指定链接，逾期未提交者视为自动放弃参赛。评委评分后根据成绩取排名前30%的参赛队伍晋级决赛。

4、决赛

决赛参赛队伍线下参赛（需自行规划行程，涉及到交通食宿费用自理，自带笔记本电脑）。决赛为路演+答辩形式，每支参赛队伍路演时间控制在8-10分钟，超时可能导致扣分或强制终止，路演包括ppt介绍和真实系统演示，答辩环节5分钟，参赛队伍需根据评委等提出的问题进行回答。决赛参赛队伍路演顺序由现场抽签决定。

5、成绩公示及申诉

决赛成绩公布后进行2小时公示，如参赛队伍对成绩有异议，可向竞赛组委会提出书面申诉书，竞赛仲裁组进行仲裁。确定无异议后递交最终成绩至竞赛组委会，由组委会举行颁奖仪式。

# 七、联系方式

企业出题赛道（重庆联通出题：基于联通元景+Deepseek人工智能知识库的自动化构建）官方QQ群：1041935478，赛题内容、作品要求、提交要求等详询：雷芳（重庆邮电大学）、刘峰（联通）。