

智领新质 共创未来

第二届北京市大学生“人工智能+”创新大赛

赛题规则

赛题

“纸书智伴”——纸质图书多模态
智能升级与智能伴读

北京世纪好未来教育科技有限公司

二、赛题背景

纸质图书是儿童阅读的主要载体，但传统图书缺乏交互与个性化伴读。将静态绘本升级为“会动、会讲、能问答”的交互式读物，是儿童阅读与素质教育领域的真实需求，也是 AIGC 多模态技术落地的重要场景。本赛题以这一真实业务需求为原型，聚焦三大难点：如何降低交互延迟、如何通过语音准确查询视频中的内容、如何保证触控式万物识别的准确度，同时兼顾生成内容与原绘本的风格一致性与页间连贯性。参赛队伍将在真实纸质绘本样本上，完成从 OCR/VLM 内容理解、多 Agent 协同生成到语音交互问答的全链路实践，推动多模态 AI 在儿童阅读场景的落地。

三、赛题考察点及技术创新

本赛题重点考察多模态内容理解、多 Agent 协同生成、风格一致性与内容连贯性、语音交互与视频内检索、触控式万物识别以及实时性评测等能力。参赛方案需结合 OCR 与 VLM，实现绘本图文内容的理解和结构化提取；合理编排“OCR—VLM—故事生成—视频生成”的多 Agent 链路，兼顾生成质量与执行效率；保证生成视频与原绘本在视觉风格上的一致性 & 跨页叙事的连贯性；集成 ASR、TTS 等技术，支持通过语音查询视频内容，并保留包含语音和图片的多模态对话历史；提升视觉触控式识别的准确度；同时围绕交互延迟、绘本生成质量和回答正确性构建三维评测指标，并通过现场演示验证系统的实时交互能力与整体效果。

四、参赛要求

本次大赛面向北京地区普通高等学校在校本科生，不限专业、年级，鼓励计算机 + 设计/教育/文学等跨专业组队；以团队形式参赛，每队 3-5 名队员，指导教师不多于 2 人。

五、赛程赛制和赛事规则

本赛题按高校预选推荐、企业分赛场选拔、北京市线下决赛三个阶段实施，遵循大赛总体时间安排。**高校预选推荐阶段**由各高校结合发布的赛题任务自行组织校内选拔，推荐队伍进入企业选拔阶段；每所高校可推荐不超过 5 支队伍参加企业赛。**企业分赛场比赛形式**为“离线开发 + 成果展示/系统演示 + 答辩”。**线下总决赛**为“现场限时任务 + 答辩”。

赛前培训：统一组织 1 次线上技术说明会，介绍绘本样本、Query 测试音频集、评测接口与基线环境。

（一）企业选拔任务：纸质绘本互动化改造与 AI 伴读

1. 命题组提供

6 本纸质绘本（授权/脱敏样本）、每本 50 条 Query 测试音频、统一输出格式与评测接口。

2. 核心任务

(1) 内容理解：OCR/VLM 提取绘本图文内容与叙事结构；

(2) 互动视频生成：生成多模态互动绘本视频（含文生故事、文生视频、语音），保持风格一致与页间连贯；

(3) 实时交互：实现语音提问、触控式万物识别的实时响应；

(4) 多 Agent 协同：构建可编排、可复用的多 Agent 协同链路。

3. 基础完成门槛

系统可运行；四项核心任务均完成；回答正确率不低于 0.60；交互延迟达到基础门槛（具体数值以定稿为准）；关键生成内容可复现；未达门槛的作品不进入一等奖候选。

4. 答辩环节

每组答辩约 8 分钟，包括陈述 5 分钟及问答 3 分钟。

5. 材料提交

需提交材料内容及要求如下，提交时间与提交方式另行通知：

- a. 可运行的网页/平板 Demo（或演示环境），以及 5 分钟以内演示视频（参考样例见文件夹中“1-参考视频.mp4”）；
- b. 项目说明书（建议不超过 20 页），包括需求分析、系统架构、核心算法设计、数据与模型来源、主要功能、测试结果和局限性等内容。命题组将统一提供项目说明书参考模板，明确各部分的主要内容及编写要求。参赛队伍可根据实际技术方案适当调整，不限定具体技术路线及实现方式（参考模版见文件夹中“2-项目说明书模版”）；

- c. 源代码或可审查的关键流程配置；如使用第三方闭源服务，应说明调用范围（参考样例见文件夹中“3-参考代码.zip”）；
- d. 数据、模型、工具、图片、文本、音频等素材来源清单及知识产权承诺书；

6. 量化评分

评分维度	评价标准	评测方式
绘本生成质量	风格一致性；页间连贯性；图文内容准确度；生成效率	专家盲评 + 指标
回答正确性	语音问答正确率；视频内容检索准确率	隐藏测试集
交互延迟	端到端交互延迟；实时性；稳定性	系统实测
触控式万物识别	识别准确率	隐藏测试集
系统实现与多 Agent 协同	链路完整性；协同编排合理性与稳定性	系统演示、代码
规范合规	材料完整、工具披露、数据与知识产权合规；儿童内容审核机制与复核责任人落实；年龄适配标注完整	成果包

7. 扣分要点与违规分级

关键生成/复现说明无法定位、工具或数据来源披露不完整、使用未授权绘本或赛前预先固定结果冒充为现场实时生成等情况，取消成绩。

儿童内容安全违规按“八、安全及风险提示”分级处置：轻微不适读扣对应分数并责令整改、明显不适读/低俗扣对应分数且相关生成项不得分、严重违规（暴力、色情、歧视、危害未成年人等）取消成绩。

未指定内容安全复核责任人或审核机制缺失，在“规范合规”项扣对应分数。

8. 同分处理

企业分赛场主体赛题同分时，依次比较“绘本生成质量”“回答正确性”“交互延迟”分数；仍相同则由评审组集体复议。

（二）主体赛题与创新赛题计分关系

企业分赛场评选以主体赛题成绩为依据，产生比赛排名及推荐晋级总决赛的队伍名单；总决赛单独计分，用于总决赛现场评奖。

（三）通用规则

允许使用开源代码、预训练模型与第三方服务，但必须合法合规、完整披露并遵守其许可证；禁止使用测试集对应的人工标注答案、未授权资料或赛前人工制作的固定结果冒充现场生成；赛后须提交代码或关键流程配置供查重与复核；相似度异常或无法复现的作品可取消成绩。

六、数据安全、成果权利与使用规则

赛事数据分级管理。训练参考集与公开验证集仅供参赛开发使用；隐藏测试集由命题组保管，不向参赛队伍开放原始文件与标准答案；参赛队伍须签署数据使用承诺，不得转发、公开、出售或用于赛事以外的模型训练/商业用途，比赛结束后按通知删除受限数据及本地副本；绘本、图片、音频、资料及标注权利归原权利人或授权提供方所有，提供比赛使用不构成权利转让；组委会绘本样本授权范围仅限本赛事开发、评测与现场演示，不包括商业发行、公开分发完整衍生作品或对外再授权；参赛队伍不得将样本用于模型预训练或赛事外训练，训练语料须合法合规并完整披露来源与授权范围。

参赛队伍对自主完成的代码、方案与原创表达保留署名权及依法享有的知识产权，同时授权组委会与命题企业为赛事评审、宣传、教学案例和非商业展示目的无偿使用；生成内容公开使用边界：互动视频、问答记录、演示材料仅可用于赛事评审、复盘与组委会授权的非商业展示；未经组委会与原绘本权利人书面同意，不得公开传播完整绘本衍生内容、不得上传公开平台用于引流或商业推广、不得二次授权第三方使用。

涉及成果孵化、商业应用、专利申请或源代码使用的，由相关方另行协商并签署书面协议。

七、设备要求及资源支持

软件： 网页/平板 Demo；可使用开源 OCR/VLM/文生视频/ASR/TTS 方案、多 Agent 编排工具及赛题组提供的统一评测接口；开发语言不限；

硬件： 企业分赛场提供统一云 GPU 算力环境，不强制参赛队伍自备 GPU；

数据集： 提供授权/脱敏纸质绘本样本与 Query 测试音频集。

企业分赛场： 现场配置不少于 2 条可用互联网接入链路或主备网络方案、公共演示大屏、HDMI 及无线投屏、平板/触控演示设备、音频采集/播放设备、计时与录屏设备、评审席、候场区和技术支持席；每队须同时准备本地演示视频与关键页面截图作为故障备用；系统异常时先暂停计时，由技术负责人记录并按统一规则决定重试或启用备用材料。

八、安全及风险提示

版权与知识产权： 仅使用组委会提供的授权绘本样本；禁止使用未授权图书、扫描件或其他品类读物作为输入；成果不得侵犯第三方知识产权；训练语料与第三方模型/素材须合法合规并披露授权边界。

儿童内容安全： 生成内容须符合儿童内容规范，不得生成不当、低俗或违规内容，生成式 AI 输出应进行人工复核并标注不确定性。

数据安全：数据集不得传播、不得用于赛事以外用途；不得采集真实儿童隐私数据；交互日志中如含语音/图像，须脱敏存储并仅用于本队调试与评测，赛后按通知删除。

操作安全： 现场演示注意平板/触控设备与用电安全。

成果公开与合规追责： 生成内容的公开传播、二次分发与商业使用边界见前文“六、数据安全、成果权利与使用规则”；违反者取消成绩并保留追责权利。

九、附加说明

激励奖励： 好未来可为优秀队伍开放多模态/AIGC 相关实习机会、企业专项奖、认证证书，并对具备落地条件的项目开放内部孵化对接通道；

本规则未尽事宜由赛题技术委员会与大赛组委会协商确定。本规则如与大赛组委会的其它规定不一致时，以大赛组委会规定为准。

“纸书智伴”——纸质图书多模态智能升级与智能伴读

赛题答疑群

