

附件：

## 中国科学技术大学第四期“福昕创新创业基金” 申报指南

为进一步发掘和培养创新人才，不断增强学生创新意识、提高创新能力、培植创业精神，根据《中国科学技术大学“福昕创新创业基金”专项捐赠协议》及《中国科学技术大学“福昕创新创业基金”管理办法》，中国科学技术大学创新创业学院福昕创新实验室（以下简称“实验室”）现面向全校学生启动第四期“福昕创新创业基金”申报工作。

### 一、申报类别

本年度共设立“试点项目”及“探索项目”两类课题。其中：

#### （一）“试点项目”课题

课题来源：实验室通过面向福建福昕软件开发股份有限公司（以下简称“福昕软件”）、我校师生征集选题并建立课题库，每年从课题库中遴选产生年度课题，“试点项目”课题由在校学生自由组队申报。

（课题描述详见附件）：

课题 1：AI 驱动的大模型输出质量预测系统

课题 2：AI 驱动的 CAD 电子图纸文档生成与辅助设计

课题 3：文本行连续性判定问题

课题 4：基于视觉语言模型的适用于 PDF 应用场景的 OCR 系统

课题 5：PDF 文件转 Markdown

## （二）“探索项目”课题

课题来源：须聚焦人工智能与智能电子文档技术及应用，围绕可信智能文档处理、文档智能化、信息安全、智慧家居等领域应用场景，应具有一定的产业前沿性及科技创新性，并能将结果运用于实践和应用。“探索项目”课题由在校学生结合所学专业、兴趣爱好等进行自主选题，自由组队申报。

## 二、申报要求

1. 课题组应由不少于 5 名学生组成，原则上须跨学科、跨院系、跨专业组队。在课题申报及执行期间，课题组负责人及成员须为拥有我校学籍的本、硕、博在读学生，课题组全体成员及指导教师应真实参与课题研究，无故不得更换或退出；

2. 我校各类公益性创新创业基金（包括：双创基金、华米基金、雏鹰基金、雄鹰基金）在研项目或两年内立项但未结项的项目不得重复申报或变相重复申报；原则上，课题组负责人及主要成员当年申报的课题数（含“试点项目”和“探索项目”）仅限 1 项；违反上述要求者，一经查实，取消申报或资助资格；

3. 课题指导教师不超过 2 名，第一指导教师应为我校在编教职工，并承诺承担经费监管、资产领用等职责；

4. 课题执行周期原则上不超过 10 个月，课题组应根据课题实际需要做好课题执行计划及经费使用预算。执行周期内，应定期组织召开不少于 3 次课题讨论会，实验室将邀请相关专家参会指导。课题执行中期及结题前，应按时提交中期检查报告、结题报告等。

### 三、申报支持

1. 实验室将组织专家进行课题申报评审，课题支持额度为：“试点项目”30 万元/项，“探索项目”15 万元/项；

2. 课题经费的使用范围含业务费、劳务费等，可主要用于支持开展技术工程化、产品商业化，具体可用于产品设计、产品推广、团队建设、咨询培训等。经费使用按批准的课题任务书和学校、教育基金会及福昕软件财务管理规定执行；

3. 实验室根据课题实际需要，提供相关支持，包括但不限于：提供创新创业课程、讲座、项目辅导会等专题培训和个性化辅导，优先推荐立项项目参加各级各类创新创业赛事、创新创业活动等，优先推荐项目团队入驻校内外实训、实践基地等；

4. 对于结题考核优秀的项目，创新创业学院将优先推荐至“雄鹰基金”及福昕软件相关的产业资源等。

### 四、申报方式

有意愿报名的学生以团队为单位组成课题组，由课题组负责人注册并登录中国科大创新创业项目管理系统（网址：<https://cxcy.ustc.edu.cn/seeding/sign-up-user/>），并于**2026年3月8日18:00**前填写并提交《中国科学技术大学第四期“福昕创新创业基金”课题申报书》。

## 五、联系咨询

王老师, 0551-63607959, [ilab@ustc.edu.cn](mailto:ilab@ustc.edu.cn);

蔡老师, 0551-63602786, [chuying@ustc.edu.cn](mailto:chuying@ustc.edu.cn)。

附：“试点项目”课题描述

<b>课题 1</b>	AI 驱动的大模型输出质量预测系统
<b>课题描述</b>	<p><b>背景：</b></p> <p>随着生成式人工智能技术的迅猛发展，以大语言模型、图像大模型、视觉语言大模型为代表的各类大模型已广泛渗透到内容创作、智能客服、科研辅助、工业检测、医疗诊断等众多领域，成为推动产业升级和社会创新的核心驱动力。</p> <p>然而，大模型在生成内容的过程中，受输入提示工程（上下文）、模型训练数据偏差、复杂任务适配性不足等因素影响，其输出质量存在显著不确定性，常出现事实错误、逻辑矛盾、图像失真、跨模态语义不匹配等问题。</p> <p>当前行业内对大模型输出质量的评估多依赖于事后统计分析，即通过人工审核或后置算法校验已生成的内容，这种方式存在明显弊端：</p> <p>一方面，滞后性较强，若低质量输出已投入应用，可能引发信息误导、业务损失甚至安全风险；</p> <p>另一方面，人工审核成本高昂，难以适配大模型海量输出的应用场景，而传统后置算法又难以全面覆盖不同类型大模型的输出特性。</p> <p>在此背景下，构建一套能够实现大模型输出质量即时预测的系统，在用户获取答案的同时即可知晓其可信程度，进而精准判断是否需要人工介入，成为解决大模型应用痛点、提升应用安全性与可靠性的关键需求。</p> <p>该系统的研发不仅能够填补大模型质量管控的即时预测技术空白，还能为不同类型大模型的规范化应用提供技术支撑，推动生成式 AI 在高要求场景的深度落地。</p> <p><b>目标：</b></p> <p>本课题旨在研发一套 AI 驱动的多类型大模型输出质量预测的原型系统，核心目标是实现大模型输出质量的即时精准预测（输出</p>

内容的同时，而不是事后统计），为用户提供实时可信的质量评估结果，辅助人工介入决策。具体目标如下：

1. 构建适配多类型大模型（大语言模型、图像大模型、视觉语言大模型）的输出质量评估指标体系，明确不同模态输出的质量核心维度（核心是准确性，不同模型可以有不同的特色指标）。

2. 研发基于 AI 的即时预测模型，能够利用大模型的输入信息、模型自身特性等前置信息，以及模型输出等，精准预测输出结果的质量等级与可信度（置信度）分数。

3. 设计并实现大模型输出质量预测系统的完整架构，包含数据输入模块、特征提取模块、预测模型模块、结果展示与决策辅助模块，支持多类型大模型的接入与适配。

完成原型系统的验证与优化，通过多组不同模态、不同任务场景的实验，验证系统预测的准确率、召回率等核心性能指标，验证其在实际应用场景中的可靠性与实用性。

#### **预期工作量：**

本课题预期需 4-5 名大学生协作完成，总工作量约 10-12 个月，具体工作量分配如下：

1. 前期调研与指标体系构建（1-2 个月）：梳理国内外大模型质量评估相关研究成果，调研不同模态大模型的输出特性与质量痛点；设计多维度质量评估指标，完成指标权重确定与量化方法研究。

2. 数据集构建与预处理（1-2 个月）：收集多类型大模型的输入-输出样本（涵盖文本、图像、跨模态数据），标注样本质量等级与关键特征；完成数据清洗、归一化、特征工程等预处理工作。

3. 预测模型设计与训练（2-3 个月）：设计适配多模态场景的 AI 预测模型架构；基于构建的数据集进行模型训练、参数调优与性能验证；完成不同模型结构的对比实验。

4. 系统架构设计与开发（2-3 个月）：完成系统各功能模块的设计；基于 Python、TensorFlow/PyTorch 等工具实现系统开发；实现多类型大模型的接入适配。

5. 系统测试与优化（1-2 个月）：设计多场景测试用例，对系统的预测准确率、实时性、稳定性等性能进行全面测试；针对测试

中发现的问题进行系统优化。

6.成果整理与报告撰写（1个月）：整理实验数据、代码文档；撰写课题研究报告、系统使用说明书等成果材料。

### **难点与创新点：**

#### **难点**

1. 多模态大模型输出质量的统一评估标准构建难点：不同模态大模型（文本、图像、跨模态）的输出特性差异显著，质量评价的核心维度与量化方式存在本质区别，如何构建一套兼具通用性与针对性的统一评估指标体系，实现对不同类型大模型输出质量的精准刻画，是本课题的核心难点之一。

2. 即时预测的特征提取与建模难点：即时预测需仅依赖输入信息、模型特性和模型输出结果完成质量预判，无法获取模型运算的过程信息及得到人工确认，如何精准提取可用信息中的隐含质量关联特征（如提示词的清晰度、完整性、复杂度等），并构建能够有效关联输入与输出质量的预测模型，突破“无人工介入即可精准判断输出质量”的技术瓶颈，有较大难度。

3. 多类型大模型的适配难点：不同大模型的架构设计、训练数据、输出逻辑存在较大差异，如何实现预测系统对多种大模型的灵活接入与高效适配，确保系统在不同模型场景下均能保持稳定的预测性能，面临较大挑战。

4. 预测模型的实时性与准确性平衡难点：实际应用场景对质量预测的实时性要求较高，而复杂的预测模型往往难以满足实时响应需求，如何在保证预测准确率的前提下，优化模型结构与算法，提升系统的实时响应速度，是系统落地的关键难点。

#### **创新点**

1. 技术理念创新：突破传统事后质量评估的局限，提出“输入-模型特性-输出结果-质量预测”的即时评估框架，实现大模型输出质量的实时预判，为大模型应用的质量管控提供前瞻性技术支撑，填补行业内即时预测的技术空白。

2. 评估体系创新：构建多模态统一的质量评估指标体系，引入模态适配因子，针对不同类型大模型的输出特性设计差异化评估维

度与量化方法，实现对文本、图像、跨模态输出质量的全面、精准评估。

3. 模型架构创新：设计基于多特征融合的 AI 预测模型，融合输入提示词特征、模型元数据特征、任务场景特征、输出结果特征等多维度前置信息，采用轻量化模型结构与高效推理算法，在保证预测准确性的同时，满足实时性需求。

4. 应用模式创新：实现多类型大模型的灵活适配，系统支持模块化接入不同架构的大模型，提供“预测结果-人工介入决策”的一体化解决方案，适配多样化的应用场景，提升系统的实用性与推广价值。

#### **预期成果：**

1. 理论成果：1 份《多模态大模型输出质量评估指标体系研究报告》，明确各模态大模型质量评估的核心维度、量化标准与权重分配方法，为相关研究提供理论参考。

2. 技术成果：1 套 AI 驱动的大模型输出质量预测系统，包含完整的软件代码与可执行程序，支持大语言模型、图像大模型、视觉语言大模型的接入与质量预测。

3. 数据成果：1 个多模态大模型输入-输出质量标注数据集，涵盖至少 3 种模态（文本、图像、跨模态）、5 种主流大模型的样本，样本数量 $\geq 10000$  条，包含详细的质量等级标注与特征信息。

4. 文档成果：课题研究报告 1 份（含背景、目标、方法、实验结果、结论等完整内容）；系统使用说明书 1 份（含系统架构、安装部署流程、操作指南等）；相关技术专利申请 1 项（可选）。

演示成果：1 套系统演示案例，包含多场景下的系统应用演示视频与 PPT，直观展示系统的功能与性能。

<b>课题 2</b>	<p style="text-align: center;">AI 驱动的 CAD 电子图纸文档生成与辅助设计</p>
<b>课题描述</b>	<p><b>背景：</b></p> <p>在制造、建筑等领域有海量的 CAD 电子图纸文档，而 CAD 设计师目前还是使用传统软件通过 GUI 完成设计，即使是简单的零部件设计也需要上百次的操作，效率不高；对于有多个零部件需要总装的设备，设计周期长，难以满足快速迭代的需求。与此同时，大量 CAD 电子图纸只是停留阅读、可视化阶段，未能充分发挥用途。</p> <p><b>目标：</b></p> <p>探索人工智能在 AI 辅助 CAD 设计的应用，包括识别 PDF/DWG 等电子图纸的图像、图形，提取产品信息、部件信息、几何造型参数、尺寸、公差等结构化信息，以大模型可以理解的格式进行后续的训练、测试与验证，形成设计师可以使用的知识库与 CAD 参数库，支持自然语言 AI 辅助设计生成建模指令序列，大幅提升零部件建模效率。</p> <p><b>预期工作量：</b></p> <p>用户痛点和需求调研：了解某垂类行业现有的 CAD 电子图纸生成的现状，预计耗时 1-2 个月。</p> <p>现有 AI 生成技术调研：了解目前 AI 生成模型在 CAD 电子图纸生成方面的现状，预计耗时 1-2 个月。</p> <p>系统设计与开发：</p> <p>前期设计：通过对数个 AI 生成模型的测试，确定基础模型，预计耗时 1-2 个月。</p> <p>编码实现：改进基础模型，根据功能复杂度分阶段进行，预计耗时 6-9 个月。</p> <p>测试与优化，预计耗时 2-3 个月。</p> <p>用户反馈与迭代：收集用户反馈，根据反馈进行功能优化和迭代，预计耗时持续进行。</p> <p><b>难点与创新点：</b></p> <p>识别结果的专业度与准确性：垂类 CAD 电子图纸设计制图国标、行标，非常专业；同时图纸部件与标注线数量多，识别过程易相互</p>

干扰，如何准确识别并理解这些内容是一大挑战

结构化数据的定义与生成：图纸的部件从总装到单独零部件，尺寸与复杂度跨度大，定义时既能纵览全局、又需不失细节，同时生成的内容需要大模型可以理解

生成技术的应用：将生成建模指令序列应用于常见的 CAD 软件，需要提较高的有效性，大幅超过设计师传统建模效率。

深度学习模型的应用：利用最新的深度学习模型和合适的数据集，如 DeepCAD，进行的智能识别和指令序列生成。

交互式编辑支持：提供基于 AI 的建议和自动完成功能，使用户在某 CAD 设计时更加高效和准确

### **预期成果：**

软件原型：开发完成的基于某 CAD（开源或者国产）AI 辅助设计界面，提供下载或在线服务。

技术文档：编写详细的技术文档和用户手册，介绍软件的功能、使用方法等。

案例研究：发布使用本工具垂类应用的案例研究，展示其功能和效果。

学术论文/技术博客：在相关的学术会议或技术博客上发表研究成果，分享技术细节和应用经验。

<b>课题 3</b>	<p style="text-align: center;">文本行连续性判定问题</p>
<b>课题描述</b>	<p><b>背景：</b></p> <p>对于从视觉丰富文档（Visually-Rich Document）转录出的纯文本文档，由于转录过程的差异，其结果的文字排布具有多样性。</p> <p>特别是转录结果中的换行符，其来源具有多种可能性，从而具有歧义。其可能表示原始文档中表格单元格的结束；原始文档中段落的结束；代码片段中的行结束；也可能是排版过程在页面边界或者环绕内容边缘所隐式加入的不具有语义含义的排版换行。</p> <p>对于这些隐式加入的排版换行，我们希望在某些文档处理过程中移除，从而恢复其前后文本的连续性。</p> <p>这个问题既可以从计算机视觉角度入手、也可以从自然语言理解角度入手。在这里我们希望得到一个从自然语言理解角度入手的解决方案。</p> <p><b>目标：</b></p> <p>设计一个系统，它接收一些阅读顺序正确的文本行作为输入，已知行与行两两之间有一个带歧义的换行符。对这些换行符是否是隐式加入的排版换行进行预测，输出结论和置信度。</p> <p>该系统至少支持中英文语言书写的文本。其他语言的支持为可选加分项。</p> <p><b>预期工作量：</b></p> <p>根据技术路线、是否自行构建数据集等因素有差异。</p> <p>预计 3~9 个月之间。</p> <p><b>难点与创新点：</b></p> <p>内容理解的准确性</p> <p><b>预期成果：</b></p> <p>软件原型：开发完成的检测软件原型和源码。</p> <p>数据集（如存在）：如果使用了原创或改编采用的数据集，提供用到的数据集。</p> <p>学术论文/技术博客：在相关的学术会议或技术博客上发表研究成果，分享技术细节和应用经验。</p>

<b>课题 4</b>	基于视觉语言模型的适用于 PDF 应用场景的 OCR 系统
<b>课题描述</b>	<p><b>背景：</b></p> <p>OCR (Optical Character Recognition, 光学字符识别) 是将包含文字的图像/视频帧转为可编辑、可检索的机器文本与结构化数据的技术。在 PDF 应用场景中不仅需要识别出文本，还需要识别字符的准确位置和方向，字符所使用的字体相关信息，例如，字体名称、字号，是否加粗和倾斜等。</p> <p>OCR 解决方案经历从传统方法到深度学习再到多模态大模型的技术路线变化。基于深度学习技术的解决方案已经成功工业化，但是基于多模态大模型的解决方案有着比深度学习更大的优势：全局上下文与布局理解强；天然支持多语言/混合语种。代表解决方案有 PaddleOCR-VL, HunyuanOCR 等。</p> <p>但是基于多模态大模型的方案也存在一定的不足或风险：存在“幻觉/不忠实”风险，可能输出图像中不存在的内容；对提示词、解码策略依赖较大。多模态大模型只输出识别到的字符结果，通常不会输出字符结果的置信度，字符的准确位置和方向，也不会识别 PDF 应用场景所需的字体相关信息。</p> <p><b>目标：</b></p> <p>结合多模态大模型（如 PaddleOCR-VL 或者 HunyuanOCR）和其他深度学习模型，实现 PDF 应用场景所需内容的识别。具体包括：字符及字符精确位置、置信度、字符所用字体相关信息（字体名、字号、是否加粗、是否倾斜、是否带衬线、是否是上下标、是否带下划线和删除线）。</p> <p><b>预期工作量：</b></p> <p>技术与调研：分析研究 OCR 在 PDF 下应用场景所需相关技术，预计耗时 1-2 个月。</p> <p>设计与开发：深度学习模型选型，模型训练调优，预计耗时 3-4 个月。</p> <p>用户反馈与迭代：收集用户反馈，根据反馈进行功能优化和迭代，可持续进行。</p>

### 难点与创新点:

充分结合多模态大模型和其他深度学习模型各自的优点，形成整体解决方案。整个系统可以高效、稳定运行，并可持续优化。

### 预期成果:

解决方案: 部署解决方案，使用 API 进行调用。

技术文档: 编写详细的技术文档和用户手册，介绍解决方案的原理、功能、使用方法等。

案例研究: 发布一系列使用本解决方案的研究案例，展示其功能和效果。

学术论文/技术博客: 在相关的学术会议或技术博客上发表研究成果，分享技术细节和应用经验。

课题 5	PDF 文件转 Markdown
课题描述	<p><b>背景:</b></p> <p>随着技术文档、学术资料和产品说明等内容大量以 PDF 形式存在，PDF 已成为信息分发的重要载体。但 PDF 天然以“版式呈现”为主，不利于二次编辑、知识结构化与内容复用。</p> <p>Markdown 作为轻量化、结构化的文本格式，在技术写作、知识库、AI 训练与内容生产等场景中被广泛使用。</p> <p>当前主流 PDF 转换方案多存在以下问题：</p> <p>结构还原能力弱，标题层级、列表、代码块丢失  表格、公式、图片与正文关系难以准确表达  扫描型 PDF 依赖 OCR，识别质量不稳定  转换结果难以直接用于知识管理或 AI 场景</p> <p>因此，有必要研究并实现一套高质量 PDF→Markdown 转换方案，提升文档结构还原能力与内容可用性。</p> <p><b>目标:</b></p> <p>本课题旨在构建一套自动化、可扩展的 PDF 转 Markdown 转换能力，实现从 PDF 文档到结构化 Markdown 内容的高质量转换。</p> <p>具体目标包括：</p> <p>支持文本型与扫描型 PDF 的统一转换  尽可能保留原文档的结构语义（标题、段落、列表、表格等）  输出符合 Markdown 规范、可直接用于编辑和发布的结果  为知识库构建、内容管理及 AI 应用提供标准化输入格式</p> <p><b>预期工作量:</b></p> <p>1、技术调研与方案设计  PDF 文档结构解析方式调研  OCR 与版式分析能力评估  Markdown 表达规范与映射规则设计</p> <p>2、核心转换能力实现  PDF 内容解析与文本抽取  文档结构识别（标题层级、段落、列表）  表格、图片、公式等特殊元素的 Markdown 表达  OCR 识别结果与原始结构的融合处理</p> <p>3、质量评估与优化</p>

转换结果准确率与可读性评估

不同类型 PDF 的适配与调优

异常文档与边界场景处理

#### 4、示例验证与成果整理

选取典型 PDF 样本进行验证

输出示例与对比说明

总结技术方案与应用价值

**难点与创新点：**

**难点**

- PDF 缺乏明确语义结构，内容与版式高度耦合
- 扫描型 PDF 需依赖 OCR，存在识别误差与顺序问题
- 表格、公式等复杂内容难以准确映射为 Markdown
- 不同来源 PDF 差异大，通用性要求高

**创新点**

- 引入结构化解析思路，从“版式还原”转向“语义还原”
  - 将 OCR 结果与版面信息结合，提升扫描文档转换质量
  - 设计 PDF 元素到 Markdown 的映射规则，提高输出可用性
- 面向知识管理与 AI 应用场景，优化 Markdown 结果的结构一致性

**预期成果：**

#### 1、技术成果

- 一套 PDF 转 Markdown 的自动化转换方案
- 支持多类型 PDF 文档的转换能力
- 转换质量与结构还原效果显著优于基础文本抽取

#### 2、成果呈现方式

- 转换工具或服务原型
- PDF 原文与 Markdown 结果的对比示例
- 技术方案说明文档与应用案例说明

#### 3、应用价值

- 支撑技术文档、知识库内容的快速结构化
- 提升文档在 AI 检索、RAG、内容生成等场景下的可用性
- 为后续文档智能处理能力建设奠定基础