

附件 1:

藏语大模型关键技术研发及应用研究 “揭榜挂帅”制项目榜单

一、发榜单位：市委人才办、市科技局、市行审局

二、发榜项目名称：藏语大模型关键技术研发及应用研究

三、项目背景：当前，高校、科研院所、企业等面向国家标准语言以及部分地区方言的语言大模型研究越来越多，但是因资金、技术等原因，面向藏语言的大模型研究及应用还很少。少数企业基于智能语音、扫描翻译和光学字符识别（OCR）等技术，打造出藏汉双语智能产品，但是还没有基于标准化的信息交换用藏文编码的藏语原生大模型，导致藏汉互译能力有待提升。

为全力推动数字西藏建设工作，通过整合藏语语料资源、结合 AI 技术能力以及垂直场景 AI 应用，对藏语原生语音语义大模型进行深度研发，并基于该模型开发场景应用非常必要。该模型可嵌入政务服务工作，实现三大价值提升：一是支持藏语语音指令与文字混合输入，打破面向老龄群体、纯藏语沟通人群的政务服务壁垒；二是赋能跨语言政务协同场景，提升政策传达精准性；三是通过数字技术带动藏语应用生态、赋能文化传承，让藏语言在数字化进程中得到保护，符合了时代的新任务和新要求。

四、功能需求

项目揭榜方配备有专业背景的项目团队，基于丰富的藏语语料，开发高性能的、原生的藏语语音语义大模型，应用在会议、便民服务等场景。

（一）藏语大模型

藏语大模型提供藏汉翻译、藏文及藏语语音识别等能力，适配会议等多个场景应用。

主要功能：

1、藏汉翻译：（1）文本互译：支持藏语与汉语的精准文本互译，覆盖公文、新闻、日常文本对话等场景；（2）语音翻译：支持实时将藏语/汉语语音转换为汉语/藏语文本或语音；

2、藏文识别：OCR 文字识别，支持印刷体、手写体藏文识别，包括现代文档及表格、古籍文献等；

3、藏语语音识别与合成：支持卫藏、安多、康巴三大方言。

（二）会议场景

会议系统集成藏语大模型，覆盖会前、会中、会后全流程，包括实时语音识别、音频文件转写、藏语多方言混合识别、说话人区分、自然语言理解、自然语言生成等能力。

主要功能：

1、支持藏语/汉语语音转文字，藏汉互译；

2、多语言识别，支持卫藏、安多、康巴等藏语方言的实时转写；

- 3、说话人分离与标注，自动区分不同发言人，标注身份；
- 4、实时字幕生成，可选藏/汉翻译并投射至会议屏幕；
- 5、会后生成会议纪要。

（三）便民服务场景

实现藏汉双语医保知识库建设、智能导办系统开发需求。

1、多模态藏汉双语医保知识库建设

整合藏汉双语医保政策、办事指南、材料清单等结构化数据，覆盖城乡居民医保、异地就医等高频事项；建立藏语方言识别模块，支持卫藏等主要方言的语音交互查询。

2、智能导办系统开发

实现“语音/文字双通道”咨询：群众可通过藏语语音查询办理地点（如各级医保局、定点医院窗口）、工作时间及材料清单；动态材料生成：根据参保类型（职工/居民）、业务情形（报销/备案）自动生成个性化材料清单，支持藏汉双语 PDF 导出。

五、项目评价指标

（一）方言适应性指标

支持卫藏、安多、康巴 3 大方言自动识别；

藏语语音识别综合化准确率质量指标（准确率 $\geq 90\%$ ）；

藏汉混合识别准确率质量指标（准确率 $\geq 90\%$ ）；

流式语音识别整句时延性能指标（200ms 以下）；

藏语多方言语音识别语体转换准确率质量指标（准确率 $\geq 90\%$ ）；

（二）实时翻译指标

多领域文本翻译忠实度质量指标（准确率 $\geq 85\%$ ）；

多方言语音翻译忠实度质量指标（准确率 $\geq 80\%$ ）；

多模态翻译忠实度质量指标（准确率 $\geq 80\%$ ）；

翻译结果符合目标语言的语法和表达习惯；、

（三）文本识别指标

标准印刷品藏文识别综合准确率质量指标（准确率 $\geq 99\%$ ）；

藏文文献典藏多风格字体标准印刷品准确率质量指标（准确率 $\geq 99\%$ ）；

手写体藏文识别综合准确率质量指标（准确率 $\geq 89\%$ ）；

藏文文献典藏多风格字体手写体准确率质量指标（准确率 $\geq 89\%$ ）；

复杂场景藏文识别综合准确率质量指标（准确率 $\geq 91\%$ ）；

（四）性能指标

对于简单任务（如短文本翻译、语音识别），响应时间不超过 2 秒；复杂任务（如长文本生成、多语言翻译）响应时间不超过 10 秒；系统支持同时在线用户数不少于 1000 人，7×24 小时稳定运行。

（五）安全性指标

数据安全方面，对用户数据进行加密存储和传输，防止数据泄露；内容安全方面，模型生成内容通过安全审核，无有害、虚假、违法等不良信息。

六、项目目标

（一）项目内容

- （1）藏语原生语音语义大模型软件系统；
- （2）配套技术文档、用户操作手册等；
- （3）提供 1 年的免费技术支持和维护服务。

（二）研发目标

（1）软件系统部署在指定的服务器环境中，能够稳定运行并满足各项功能需求和性能指标；藏语原生语音语义大模型基于标准化的信息交换用藏文编码；

（2）技术文档内容完整、准确，符合行业规范，能够指导系统的开发、维护及使用；

（3）在规定时间内响应技术支持服务相关问题，并提供有效的解决方案。

七、项目揭榜方要求

（一）基本要求

1.项目揭榜方组成

项目揭榜方为区内外有研发能力的高校、科研院所、企业等法人单位。鼓励有社会责任感的企业联合高校、科研院所等，组成一支产学研联合团队，共同致力于藏语大模型的研究与开发应用工作。

2.项目经验要求

项目揭榜方具备人工智能、自然语言处理等相关项目开发经验，至少完成 1 项语言大模型或多语言处理项目，且具备藏语或其他地区方言人工智能处理经验。

3.团队成员要求

项目揭榜方团队成员具有计算机科学、人工智能、自然语言处理及藏语语料领域等专业背景及研发经验，并通过该项目持续培养藏语大模型方面的本地人才。

4.本地研发要求

项目揭榜方配备本地研发与运营团队，且软件系统部署在西藏本地。

（一）技术要求

1.算力基础

拥有自主的高性能计算服务器集群或稳定的云计算资源，算力条件满足项目需求；具备成熟的语音语义处理技术，能够实现藏语语音和语义处理功能开发；具有完善的数据安全管理体系和质量保障体系，确保项目数据和成果的安全性和可靠性。

2.语料基础

基于标准化的信息交换用藏文编码规范，以保障模型训练的有效性及实用性。

3.研发基础

具备自主研发的人工智能算法、模型架构，在国际或国内权威学术期刊/会议上发表过相关领域的研究成果；拥有自然语言处理技术积累，能够为藏语大模型开发提供技术支撑。

八、项目资金来源

政府财政支持经费不超过 500 万元。若企业牵头作为项目揭榜方，配套资金原则上不低于 1:1。

九、项目研发周期

（一）研发期限

自合同签订之日起 6 个月内完成原生藏语语音语义大模型部署；随后 12 个月内，部署会议、便民服务等场景应用，并试运行。

（二）项目进度安排

本项目分期进行建设，一期项目分为建设期和试运行期，具体时间节点如下：

建设期：6 个月，包括语料准备、算力资源部署、语料训练、原生藏语语音语义大模型研发与部署等工作。

试运行期：12 个月，将藏语大模型嵌入会议及便民服务场景，并开展系统运行维护等工作。

序号	项目名称	发榜单位	项目领域	任务要求（研究内容）	成果要求（考核指标）	项目总投资额（万元）	企业出资额（万元）	完成时限	联系人及联系方式	属地
1	政府采购藏语大模型项目	拉萨市科学技术局	数字经济领域	基于丰富的藏语语料，开发高性能的、原生的藏语语音语义大模型，应用在政府会议翻译及智能会议纪要、便民服务场景。 研究内容：1. 研发原生的藏语语音语义大模型，为 AI 多场景应用打下基础；2. 研发藏汉双语会议场景，提供智能翻译、智能会议纪要等功能。3、便民服务场景：1) 多模态藏语医保知识库建设：整合藏汉双语医保政策、办事指南、材料清单等结构化数据，覆盖城乡居民医保、异地就医等高频事项；建立藏语方言识别模块，支持卫藏等主要方言的语音交互查询。2) 智能导办系统开发：实现“语音/文字双通道”咨询：群众可通过藏语语音查询办理地点（如各级医保局、定点医院窗口）、工作时间及材料清单；动态材料生成：根据参保类型（职工/居民）、业务情形（报销/备案）自动生成个性化材料清单，支持藏汉双语 PDF 导出。 主要目标：1. 研发原生的藏语语音语义大	1. 考核指标： （1）方言适应性（支持卫藏、安多、康巴 3 大方言自动识别）指标： 藏语语音识别综合化准确率质量指标（准确率≥90%）； 藏汉混合识别准确率质量指标（准确率≥90%）； 流式语音识别整句时延性能指标(200ms 以下)； 藏语多方言语音识别语体转换准确率质量指标（准确率≥90%）； （2）实时翻译指标： 多领域文本翻译忠实度质量指标（准确率≥85%）； 多方言语音翻译忠实度质量指标（准确率≥80%）； 多模态翻译忠实度质量指标（准确率≥80%）； （3）文本识别指标： 标准印刷品藏文识别综合准确率质量指标（准确率≥99%）； 藏文文献典藏多风格字体标准印刷品准确率质量指标（准确率≥99%）；	1000	500	2 年	——	拉萨市

				模型，为实现多场景应用打下基础，并完成会议场景和便民服务场景。2. 做好供大模型使用的算力资源、硬件及安全配置等。 手写体藏文识别综合准确率质量指标（准确率$\geq 89\%$）； 藏文文献典藏多风格字体手写体准确率质量指标（准确率$\geq 89\%$）； 复杂场景藏文识别综合准确率质量指标（准确率$\geq 91\%$）； （4）性能指标： 对于简单任务（如短文本翻译、语音识别），响应时间不超过 2 秒；复杂任务（如长文本生成、多语言翻译）响应时间不超过 10 秒；系统支持同时在线用户数不少于 1000 人，7×24 小时稳定运行。 （5）安全性指标： 数据安全方面，对用户数据进行加密存储和传输，防止数据泄露；内容安全方面，模型生成内容通过安全审核，无有害、虚假、违法等不良信息。 2. 研发成果： 藏语语音语义大模型软件系统；配套的技术文档、用户操作手册等；提供 1 年的免费技术支持和维护服务。					
--	--	--	--	---	--	--	--	--	--