

项目榜单

榜单名称	面向端侧智能的离线类人自然交互技术研发		
行业领域	人工智能	专业方向	关键技术
(计划)启动时间	2025年1月1日	计划完成时间	2027年12月31日
榜单提出目的	<p>随着人工智能技术的快速发展，大模型在自然语言处理、计算机视觉和语音识别等领域取得了显著的成果。然而，大模型的部署和应用面临着如计算资源消耗大、数据传输延迟、隐私保护等诸多挑战。因此，端侧化成为大模型落地的重要载体。端侧智能主要包括云端大模型和端侧大模型。云端大模型虽然功能强大，但依赖稳定的网络连接，难以满足实时性和隐私保护的需求。相比之下，端侧大模型能够在设备本地进行数据处理和决策，具有低延迟、高隐私保护和低功耗等优势，能够更好地满足即时响应和数据安全的要求。然而端侧智能的应用也面临着计算量大、实时性要求高，受限于运行环境、内存、存储空间等挑战。特别是在语音交互方面，端侧设备需要在低功耗、小内存条件下实现高性能的语音识别、语义理解和语音合成，这对技术提出了更高的要求。</p> <p>本项目紧密围绕端侧大模型应用，重点攻克基于端侧算力的低延迟交互技术、端侧语音大模型自适应技术、端侧语音一文本联合处理技术，离线和在线相结合的多模态语音交互系统等关键技术，解决离线端侧智能人机交互应用的重大技术挑战，赋能多种终端设备，实现更加自然、高效的类人交互体验，保护用户数据隐私和安全。推动“AI+”在各领域的深度融合，为智能终端普及与应用开辟新路径。</p>		
榜单任务内容	<p><b>1.榜单项目拟解决的问题</b></p> <p>(1)基于端侧算力的低延迟交互技术。研究多层次大模型的模型量化和压缩技术，优化计算图，实现多语种低延迟的大模型推理。</p> <p>(2)端侧语音大模型自适应技术。针对用户的发音习惯、口音差异、非母语发音、非规范表达等现象，通过长时差异补偿和混合口音自适应技术，使系统动态适应用户的个性化交互方式。</p> <p>(3)端侧语音一文本联合处理技术。研究端到端热词快速增量学习，语音识别结果纠错，逆文本归一化，及生成式口语摘要技术。</p> <p>(4)离线和在线相结合的多模态语音交互系统。在多种硬件平台，包括国产化平台上实现多模态交互，优化多设备协同交互，实现离线和在线结合交互技术。</p> <p><b>2.技术性能指标及产业化指标</b></p> <p>(1)技术指标</p> <p>1)端侧离线的语音识别、说话人分离、语音合成和文本推理；</p> <p>2)端侧离线的中英文语音识别准确率&gt;97%，支持10种语言；</p> <p>3)端侧离线的说话人分离，总分类错误DER&lt;5%；</p> <p>4)实现1000个热词快速学习，生效时间小于30秒。</p> <p>(2)产业化指标</p> <p>研发成果应用于典型行业用户，实现收入1000万元以上。</p>		

<p>榜单效益目标</p>	<p><b>1.预期经济效益</b></p> <p>智能终端是人工智能的重要载体，端侧AI部署是AI实现规模化扩展及应用落地的关键。2023年我国智能硬件的市场规模达14,031亿元，预计2024年将达到15,033亿元。IDC预计，2024年中国的AI终端占比将达到55%，搭载AI功能的终端设备将超过70%。本项目研发端侧智能离线类人自然交互技术，面向政企、医疗、教育、交通、养老等领域推广应用，形成行业应用示范的标杆。实现直接经济收益1000万元以上，带动人工智能市场规模数十亿元。</p> <p><b>2.对产业发展的突出贡献</b></p> <p>本项目通过突破端侧智能离线类人自然交互领域的“卡脖子”技术，实现规模化产业应用和国产化替代，推动我国人工智能技术的创新发展，增强自主创新能力。项目的应用将赋予智能终端设备类人情感智能和多模态交互能力，提升人机交互的自然流畅度和情感深度。在可穿戴设备、智能家居、智能汽车等领域应用，革新用户与智能设备的交互方式，使之更加人性化、个性化、智能化，同时保障用户数据的安全与隐私。推动智能终端设备中关键组件与技术的国产化，增强产业链的安全性和稳定性，增强中国在全球智能终端市场的竞争力。</p>
---------------	--