



机器人 + 交互环境 (人和场景)



感知器： 机器人 + 环境传感器

- 感知语音： 语音识别、声纹、语种、语音情绪；
- 感知空间： 声事件、场景分割、语义地图、雷达、超声；
- 感知人物： 人脸识别、人脸情绪、唇语识别；
- 感知文字： OCR识别

多模态认知

多模态表达

多模态感知



表达承载： 机器人 + 环境设备

- 语音表达： 语音合成，音色、情绪、歌声、变声；
- 视觉表达： 数字人、肢体动作、表情、微动作；
- 联动表达： 富媒体表达（音乐、图片、视频、网页），IoT多设备驱动；

功能模块	创作平台功能说明	应用及人机交互效果示例	任务类型
文章摘要管理	<ol style="list-style-type: none"> 平台对接客户的公众号，解读新闻； 提供自动文本摘要和人工审核入库功能 	<p>互动解读客户每日公众号内容。 用户：今天我们公司有什么热点消息？ 机器人解读摘要，有屏机器人再跳转到公众号界面。</p>	中文文本摘要
数字化场景管理	<ol style="list-style-type: none"> 平台提供基于文本检索图片、生成图片功能； 图片作为诗词朗诵、讲故事的穿插素材，作为数字人交互时的背景空间素材； 图片和视频生成后人工审核入库，基于交互记录的二次标注； 	<p>人机交互的虚拟空间生成，交互讲解辅助。 用户：冰墩墩长什么样子。 机器人进行图片搜索并展示。 用户：你给我讲讲愚公移山的故事。 机器人讲故事，随着文本进度显示不同的配图。</p>	以文搜图、以文作画 
交互场景管理	<ol style="list-style-type: none"> 平台提供基于视觉场景理解的语义管理功能； 基于交互记录图片和文本标注，意图分解标注； 	<p>基于交互主体，解读交互对象的互动空间。 用户：告诉我你看到了什么？ 用户：我周边有几个人，我右边有几个人，左边的是男生女生？</p>	以图生文、物体检测、VQA 看图提问 
媒体资产管理	<ol style="list-style-type: none"> 平台录入视频和图片时，基于以图生文、视频描述自动构建摘要和索引； 人工审核后二次标注； 	标注和数据支撑	以图生文、视频描述