

# 拟公示算法机制机理内容

算法名称	云果多模态对话生成算法
算法基本原理	<p>本算法识别用户语音提问，根据语义和意图分别生成文本、图片和语音内容。主要包含如下模块：</p> <ol style="list-style-type: none"><li>1. 语音识别：通过语音检测识别声音信号结束，由声学模型进行音频识别，将声音转化文字并进行通用语义识别。</li><li>2. 对话管理：结合通用和专门语义识别，区分私人助理、公共知识、图片、音乐等请求类型，组织不同模态内容生成、多轮对话。</li><li>3. 多模态内容生成：基于大语言模型生成公共类知识内容、图形模型生成通用类自然风景图片、自然语言模型生成私人助理回复，以及回复音乐内容。</li><li>4. 语音合成：通过发音与合成引擎，将文本内容合成为音频。</li></ol>
算法运行机制	<p>本算法接收用户通过智能终端提出语音问题，将音频转化成文本，进行关键字、意图提取。对于通过敏感词审核的用户提问，分别叠加 <b>Prompt</b> 之后由大语言模型进行问题答复、由图像模型按用户意图的生成图像、由音乐平台回复歌曲请求。最后通过语音合成向用户播报答复内容。</p>
算法应用场景	云果 AI 助手系列的语音对话系统。
算法目的意图	通过多模态 AI 模型融合、云端侧 AI 融合，提供高效工作和学习 AI 工具，帮助用户获取个人化 AI。
算法公示情况 (选填)	