

36氪获悉，人机语音交互服务商「一知智能」已完成超亿元人民币B轮融资，由凯泰资本、亿联凯泰以及中信证券投资共同投资。据悉，本轮融资资金将用于算法开发、产品升级、团队建设与新业务开拓。

一知智能是36氪曾报道过的一家企业，于2017年成立，创始团队来自浙大人工智能研究所，是一家专注于人机语音交互技术的SaaS服务商，通过电话、微信、短信等载体为企业触

达用户，并实现精细化智能运营

。公司凭借自研的语音识别、语义理解和语音合成

三大人机交互核心算法，为零售、生活服务等泛消费行业，以及公共服务提供AI场景化智能服务。据一知智能介绍，公司曾先后获得浙江省高新技术企业研究开发中心、杭州市领军型创新团队，中国人工智能学会语音对话与听觉专业委员会单位。

近年来，在AI产业落地的浪潮中，智能语音交互因其交叉学科属性，包含ASR(语音识别)、NLP(自然语言处理)、TTS(文本语音转换)等多项高难度AI技术，至今依然有较高的应用落地门槛。而从场景需求上来看，在商业活动与政务工作中，以微信和电话为主要交流载体的人员排查、信息采集、定向通知、政务咨询等环节，往往是最费时费力的部分，因此大到各地政府的防疫提示来电，小到消费品牌的大促上新通知，投入智能语音交互技术也已成为政企服务实现降本增效的趋势之一。因此，如何迅速适应通话场景，提升AI语音的对话体验，成为业内人机交互服务商进行技术优化的首要考验。

一知智能创始人、CEO陈哲乾在介绍团队的赛道选择时，把创业故事比喻为“拿着锤子找钉子”的科研式创业之路。彼时，求学于浙江大学人工智能研究所的公司创始团队，便在人机智能交互领域有了丰富的研究成果，曾代表浙江大学多次获得国际性NLP比赛冠军。与AI赛道内其他带有科研基因的团队相类似，一知智能的创业方式也是带着较强的技术寻找拥有商业前景的落地场景。经过近3年时间，在调研机器可以取代人类工作的诸多行业领域后，一知团队最终将产品定位在高频次的人机互动载体——电话，再通过已积累的AI语音技术为智能外呼找到应用场景，由此确定了公司核心技术的商业化方向。也因为拥有了核心算法、算力与数据的支持，后续公司业务的发展锚定双向互动需求较多的消费广告与公共服务领域，便进入了“逢山开路，遇水迭桥”的发展阶段。

对比业内同样聚焦于AI语音机器人与智能语音赛道的百应科技、网易七鱼、智齿客服等企业，陈哲乾认为一知智能的主要差异点在于将公司定位于技术研发型服务商，通过人机对话技术的不断迭代来做深消费品牌的会员智能化服务。从这一定位出发，近年来一知智能在核心技术上有了以下升级：

·ASR：

在语音识别提取的过程中，一知智能在webRTCNS(降噪处理代码)的框架基础上，对环境音降噪与人声增强进行双向优化，实现双模并行。同时，公司使用MFCC+resCNN提取算法方案，细化机器人在性别、年龄与情绪识别中的感知颗粒度。

在通用场景识别的E2E（端到端）模型中，一知智能在传统语音识别架构AED（Attention-Encoder-Decoder基于注意力机制的编解码模型）的基础上做了场景化升级，自研了基于上下文文本定制增强模块的Context-Aware Encoder算法，在模型训练过程中，使用额外补充的场景文本作为强化信息，让模型建立起针对特定输入的强化架构，提高特定垂直场景中的语音识别率。

体现在场景应用上，该项技术能过滤沟通中的环境噪音干扰，并使机器人能快速识别用户的年龄、性别与情绪，根据当前情景选择最合适的对话内容。

·NLP：

一知智能已于日前升级了新一代对话架构NLP2.0。在实现意图识别、情感识别、问答识别、任务对话、智能纠错，知识图谱等NLP常见任务之外，新的对话架构还提出了在泛消费领域的专用大规模预训练语言模型“EAZI”——在Transformer架构基础上，基于语言学知识与大量垂直数据，以NLP算法驱动语义理解，改进模型架构与预训练策略，并且针对大量的消费领域信息与公司已积累的消费场景对话进行专项训练。

体现在场景应用上，此项创新在问卷调研、用户满意度调研等外呼需求中，解决了时间、地址、组织名称等实体信息提取的难点，机器人在应对消费场景客户调研中的复杂问题时，也能快速找到对应答案。

·Fastspeech 系列语音合成算法：浙大 一知人工智能联合研究中心，联合微软共同推出FastSpeech1和FastSpeech2两代算法系列，集合了语音合成、情感合成和语音克隆算法，以及相应的语气词回复生成算法。

据一知智能介绍，该算法系列相较于谷歌的tacotron两代算法，在同等实现效果上，速度分别提升了38倍和260倍。体现在场景应用上，机器人在沟通中能够逼真地模拟人类情绪，并能在识别对话方情绪后及时给出回应。