

AI语音关键技术与业务实践研究

Research on Key Technologies and Service Practices of AI Speech

陈浩然,孙记明,刘牧寅(中讯邮电咨询设计院有限公司,北京 100048)

Chen Haoran, Sun Jiming, Liu Muyin (China Information Technology Designing & Consulting Institute Co., Ltd., Beijing 100048, China)

摘要:

研究了AI语音技术在业务领域的实践应用,梳理归纳了AI语音三大关键技术:语音识别、语音合成、自然语言处理的技术原理与技术演进路线;描述了构建在三大技术之上的AI语音产品的架构、功能特色和应用场景;详细说明了AI坐席产品、AI质检产品与云呼叫中心对接方案,并通过对比的方法给出各方案的优势劣势;最后对AI语音技术与产品进行了展望。

关键词:

语音识别;语音合成;自然语言处理;AI坐席
doi:10.12045/j.issn.1007-3043.2021.02.013
文章编号:1007-3043(2021)02-0067-07
中图分类号:TN914
文献标识码:A
开放科学(资源服务)标识码(OSID):



Abstract:

It studies the practical application of AI speech technology in the service field, and summarizes three key technologies of AI speech: speech recognition, speech synthesis, and natural language processing. Then it describes the architecture, functional features and application scenarios of AI speech products based on three technologies above. Furthermore, the connecting schemes of AI agent products, AI quality testing products and cloud call center are described in detail, and the advantages and disadvantages of each scheme are given through comparison. Finally, it prospects development trend of AI voice technology and products.

Keywords:

Automatic speech recognition; Text to speech; Natural language processing; AI agent

引用格式:陈浩然,孙记明,刘牧寅. AI语音关键技术与业务实践研究[J]. 邮电设计技术,2021(2):67-73.

0 前言

当前,在云计算、大数据、AI技术的赋能下,原有以电话为主要手段进行客户联络与营销的传统呼叫中心,正在向云化、数据化、智能化为特征的全渠道多媒体客户服务中心及营销中心转变。尤其是以AI语音技术为特色的“云+网+呼叫中心”的智能呼叫中心产品,解决了行业内部现存各类问题,提升客户服务效率,优化客户体验。而以AI质检为代表的运营管理工具,在辅助客服人员服务过程的同时,对坐席通话

行为与内容进行分析抽查,进而营造出良性发展、更有秩序、更加规范的语音业务生态。综上所述,AI语音技术破除了政企语音市场规模化发展所遇到的障碍。

1 AI语音关键技术

2017年9月6日,微软小冰给人类打出第1个全双工电话,2018年5月10日,Google在I/O大会首次做了AI电话 Duplex。此后AI语音技术在商业价值的驱动下开始了飞速发展。一个完整的利用电话进行人机对话的流程如图1所示,通过语音识别(ASR)、语音合成(TTS)、自然语言处理(NLP)、声纹识别(VPR)、文本分析(TA)等技术手段,依托语音信号处理、深度学

收稿日期:2020-12-02

习、模式识别、数据挖掘等算法,实现人机沟通、信息传递、意图表达、情感交流等行为。

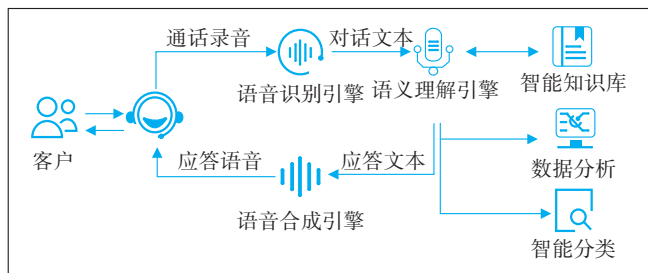


图1 基于电话的人机对话流程

1.1 ASR 技术演进

语音识别即机器理解和识别人类语音信号并转成文本的技术,基本流程有3个环节:语音预处理数字化、特征提取、训练/识别,通用的语音识别算法是通过模式匹配与统计模型相结合的方式来实现,先对已有的音素数据构建参考模板;再构建参考模板和未知数据提取的模板之间测度函数,最终用最佳准则来识别决策。

第1代的通用算法是HMM-GMM混合高斯隐马尔科夫模型,该算法可以将声学 and 句法集成在单一框架中,并使得连续语音识别成为可能,在业内最早实现商用化,并持续商用了近30年。随着深度学习理论下神经网络(DNN)的成熟,受限玻尔兹曼机(RBM)所堆叠出的深度置信网络(DBN)替代了混合高斯模型(GMM),基于DNN-HMM模型成为第2代语音识别的主流算法,显著降低了误识别率,并激发了基于深度神经网络的语音识别研究热潮。新的神经网络层出不穷,循环神经网络(RNN)与传统的神经网络相比,网络结构更有优势更能成功地应用到语音识别任务中,而长短时记忆网络(LSTM)作为RNN的一种典型网络,将识别准确率较之前提高了6%~8%,成为第3代语音识别技术主体。近2年来,研究者们发现RNN的另一种典型网络连接时序分类器(CTC)网络不必再针对每一帧对语音进行对齐的预处理操作,省去了耗时的反复多次迭代,彻底摆脱HMM的框架,预测时的速度更快,识别速度更快,且更换识别语言体系时可以利用相同的框架结构训练,实现了端到端的语音输入到文字识别。该算法模型为语音识别第四代算法模型,称为E2E模型。目前该技术识别率处于国内领先地位,16K语音识别率超97%,8K语音识别率超85%;识别速度实时率可以达到0.2~0.3;且支持英语等外语种、维藏蒙彝朝等少数民族语、粤语四川话闽

南话等方言的识别。

1.2 TTS 技术

语音合成即将文本转成语音的过程,基本流程有4个环节:语义表示、语言编码、发声编码、语音产生,实现了从意向到概念到声码器控制信号再到语音波形的转化。

早期的语音合成是最基本的语音单元,如音素、双音素、半音节、音节作为合成单元,建立语音库,通过合成单元拼接而达到无限词汇的合成,并通过声码器(如串并联混合型共振峰合成器)来模拟发音过程中的声道共振(串联通道产生元音和浊辅音,并联通道产生清辅音)。从20世纪90年代末,主流技术开始使用存放大量真人语音样本的语音库,通过选择合适的拼接语音片段来实现高质量的合成语音。2000年以来语音合成是基于统计建模和机器学习的方法,通过建模加训练快速构建合成系统,构成了第1代语音合成算法的基石,即HMM参数合成方法。和语音识别技术类似,深度神经网络的统计声学建模方法也迅速成为语音合成的技术核心,第2代基于DNN的方法由于对高维声学特征中各个维度相关性和输入输出复杂映射关系能够更为精细化的建模,大大改进了合成语音的音质。而最新的第3代基于端到端的模型预测,将神经网络的训练和声码器的训练过程合二为一,不再需要人工录音、标注,通过机器自主训练具备与真人高相似度的音色,发音连接顺畅,更贴近自然人发音。第三代的语音合成技术可基于应用场景的不同,调整合成的语速、音色、音量等特征,匹配各场景需求,同时在语音素材库中,拥有男声、女声、童声等多种音色,每种音色均有不同发音人,无需重新进行训练,使用同一个模型即可全面满足各种应用场景多样需求。

1.3 NLP 技术

自然语言处理是指在人机交互中机器理解自然语言,并做出响应的技术,解决人工智能完备性上的认知智能问题,具体技术手段仍为马尔可夫统计学建模与机器学习、深度学习相结合的方法。自然语言处理在分析对象上分为词汇级、句子级和段落篇章级,包含的功能要点有语法分析、语义表示、语义关系、语句生成以及单文本/多文本分析;在分析内容上分为词法分析、语法分析、语义分析、语用分析等。近10年来,自然语言处理都是基于神经网络来实现的。第1代自然语言处理利用浅层神经网络,基于词向量技术,将用户输入问题进行分词,利用关键词的词频及词权重进行

关键词赋值,实现了词汇级的理解;第2代自然语言处理基于深度学习和强化学习技术,从词向量扩展到句向量,进而支持基于上下文的语义理解和多轮对话,目前是AI机器人主流应用技术;第3代自然语言处理基于更先进的ELMo算法,支持篇章级的语义抽取,对上下文的语义理解和多轮对话有了更进一步的提升。随着网络深度的不断增加和机器学习自主化的深入,自然语言处理系统一方面能够应用历史积累的数据,只需少量人工给出头部结论,便可自主学习进行知识归类及认知;另一方面能够应用历史积累处理失败数据,无需人工参与,自主进行无监督聚类,挖掘出新的知识点,并对旧的知识点进行优化。

通过ASR、TTS、NLP的技术演进来看,技术的更新迭代周期是指数型的而非线性的,且最终都是通过AI领域最为核心的深度学习理论来实现。

2 AI 语音产品

整合前文所述业界领先的语音识别、语音合成与自然语言理解算法,结合运营商的语音能力开放,再运用数据挖掘、知识图谱、数据可视化等相关辅助技术,可打造一系列新型的AI语音产品,提供智能化语音服务与人工坐席的智能化支撑服务,实现高效率、低成本的营销、客服、质检工作。作为运营商传统语音业务的增值功能,可突破呼叫中心人工坐席各类瓶颈,提升企业的沟通效率,扩大企业的销售成果、品牌影响及市场地位。

2.1 架构设计

AI语音产品的系统软件架构如图2所示。

从软件架构来看,业务管理层面面向不同应用场景,对其业务进行统一运营管理,满足客户预测式外呼、人

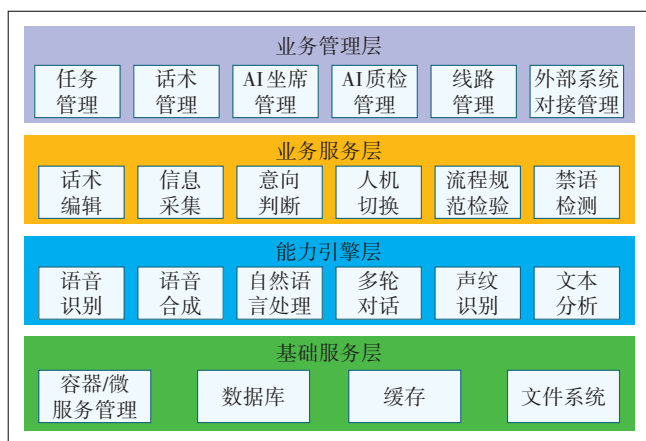


图2 AI语音产品系统架构图

机切换、纯AI机器人3类需求场景;业务服务层整合引擎基础能力,执行具体业务逻辑处理,满足客户整体AI语音能力调用需求场景;能力引擎层提供AI语音的原子能力,当前ASR准确率对于电话信道可达85%,ASR+话术训练后的NLP整体交互准确率可达92%,满足客户语音识别高准确性需求场景;基础服务层对计算资源、存储资源、线路资源进行管理、监控与灾备,满足客户电信级保障的高可用需求场景。

从产品功能来看,AI语音产品可以分为AI坐席产品与AI质检产品。AI坐席产品提供智能电话营销、智能问卷调查、信息通知、活动推荐、客服接待等服务。AI质检产品基础功能包括语音转写、话者分离、关键词检出;核心功能包括话术规范分析、静默分析、语速语调分析、意图识别分析、禁语检测、全文检索、自动定位问题录音、人工复检、录音回溯;管理功能包括结果报表生成、配置管理、词库管理、规则管理、监控管理、计划管理以及系统内的权限管理、资源监控和日志管理。下面分别介绍2个产品的功能特色。

2.2 功能特色

2.2.1 AI坐席特色

AI坐席产品的外呼功能主要定位于营销或回访的应用场景,有3个优势特色:一是赋能电话呼叫业务,实现高效率低成本转化,完成客户精准筛选;二是支持预测式外呼(多并发),外呼事半功倍效率更高;三是通过客户无感知的人机切换,能够在发现重要商机的第一时间人工跟进。

AI坐席产品的呼入功能主要定位于投诉接待、业务咨询等呼入场景,有3个优势特色:一是帮助客户自助解决问题,让客户享受菜单导航式的咨询体验,二是通过情感分析来提升人工客服质量,通过分析负面情绪对应的消息内容,精准定位服务存在的问题,预先或有效及时解决诉求共性问题、突发性事件或苗头性问题;三是结合大数据的舆情智能分析,可以统计出用户的热点问题,准确掌握用户需求,及时了解社会热点及用户的状况,尤其在政府机关、事业单位对热门事件分析、地区舆情分析、市民咨询与投诉热点分析的场景中,优势最为明显。

在外呼场景下,与人工坐席相比AI坐席可提升3倍的客户意向筛选效率,平均节省50%以上人力成本。同时AI坐席产品的呼入功能可大幅提升客服工作效率与客户满意度,并为客户节省45%以上的人力成本,如表1所示。

表1 AI坐席与人工坐席对比

项目	人工坐席	AI坐席
每天拨打量/通	200~300	800~1 200通
每天工作时间/h	8	外呼11; 客服24
年工作日/天	250	360
年拨打电话数量/通	75 000	288 000
工作状态	不可控, 受情绪影响	稳定
数据统计	存在内容缺失、记录耗时、个人主观问题	信息完善、系统高效、客观记录
事件跟进	存在记录较乱、撞单问题	自动分类、及时跟进
客户体验	两极分化, 投诉风险	话术规范、内容全面
成本	人员费用+场地费用+平台使用费用	AI服务年租赁费用

2.2.2 AI质检特色

AI质检主要定位于企业自用和运营商监管场景。

针对企业自用场景, AI质检即可完成通话内容的质检以及语音特征的质检, 其中通话内容的质检是根据质检规则加工和NLU算法识别结果进行比较, 对通话内容中的话术、关键词、意向进行分析与检测, 并从大量语音中提取语调、语速、能量等多种特征, 经过多重神经网络训练, 能够判断对话者的情绪, 并做出简单的推理。进而检验人工坐席的客服人员话术是否规范统一, 是否存在滥用; 同时甄别客户情绪, 当识别到客户投诉意图时, 可进行后续的补偿措施。

针对运营商监管场景, AI质检用于对双方通话的行为、话术等进行统一抽查分析, 查验是否有违规业务的开展。这一过程需要人机协作来进行质检, 即根据不同角色的工作职责, 自动质检后台系统为业务人员提供相对应的人机协作质检功能。

由于AI质检+人工复检覆盖率与效率远超过人工抽检质检, 解决普通质检覆盖面小、对精准服务能力提高支撑不足的问题, 规避了坐席主观记录导致的工单

的真实性和完善性存疑问题, 快速提升精准服务能力, 降低人工话务压力。结合后端对垂直行业的深度数据挖掘分析, 更能充分利用呼入电话中蕴藏的海量用户行为数据, 为企业业务经营提供强大的支撑手段, 对垂直行业领域的销售/服务模式产生深远影响。

2.3 典型场景

2.3.1 预测式外呼

预测式外呼即呼叫中心系统通过自动拨号功能, 对当前空号、忙线与其他无效通话状态进行过滤, 接通后再转给坐席, 减少用户等待或坐席等待的情况发生。而结合AI外呼的预测式外呼, 将通话接通后转给AI坐席, 通过AI的初步意向筛选再转接给人工。过滤+初筛可以大幅减少人工坐席的工作量, 提升人力资源的投入产出比, 让人工坐席解答更为专业的问题。同时根据呼叫接通率和意向初筛比例情况, 可以动态调整AI坐席数量和人工坐席数量, 合理配比优化资源。如图3举例, 如每天外呼并发量是 X , 且接通率仅50%, 那么AI坐席数量可设置为 $50\%X$, 如果意向筛选大概50%才有意向, 那么人工坐席数量可设置为 $25\%X$, 图4为一个信用卡营销的实际案例数据。

2.3.2 AI外呼智能回访

智能回访信息采集是AI外呼另外一个重要应用

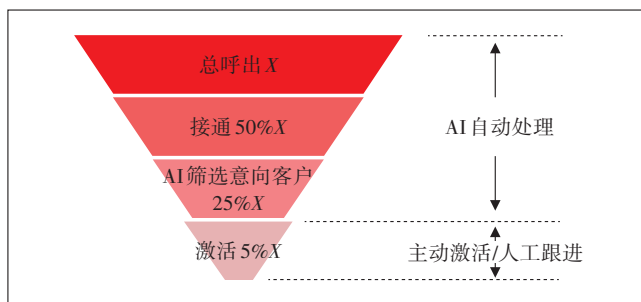


图3 预测式外呼示意图

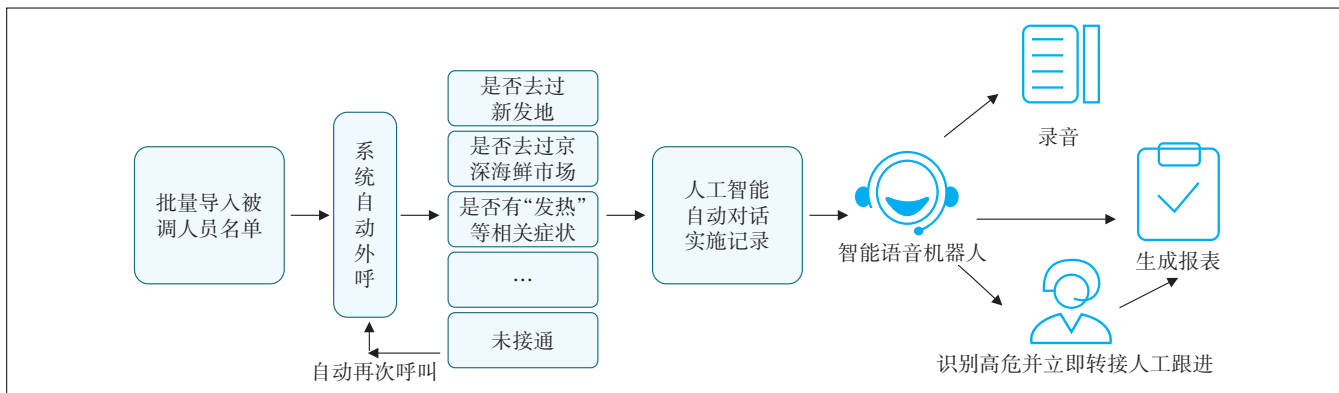


图4 智能回访示意图

场景,设定回访话术并将回访人员名单导入系统后,系统按照需要调查的问题批量发起对被调查者的通话,并通过拟人对话、对被调查者的语音识别、采集被调者的按键反馈来实现回访的数据收集。

该应用场景为2020年的疫情防控提供了较大的帮助,社区通过AI智能回访与辖区人员电话沟通,并自动生成调查结果,当识别到高危人士或调查发现高危情况,系统立刻实时在线将电话自动转接人工,保证高危信息第一时间获得处理和跟进。所有通话中产生的调查结果自动形成标签分类统计,迅速形成报表,以便上报上级单位。

2.3.3 AI客服政务大厅

AI客服所赋能的智慧政务是AI语音技术的一个重要应用场景。为了全面提升政务服务规范化、便利化水平,为群众提供高效、便捷的政务服务,国家指导各政务部门通过“12345”等政务服务热线集中接受社会公众的咨询、求助、意见、建议和投诉,进而整合现有政民互动渠道,提供政策法规、办事程序、生活指南及查询有关部门职能范围等咨询服务。与此同时,国家标准委制定了量化标准,即每周7天,每天24h工作制,电话受理时应在15s之内接听,连续24h内呼叫接通率应大于等于95%。

AI客服恰恰契合智慧政务的一系列要求,依托电话、微信、app、网络在线等众多服务渠道,通过智能语音、文字、图片、视频等多种方式与公众进行远程多媒体交互,并且对外提供智能服务接待、业务导办与办理、智能外呼调研等服务,支撑政府“不见面”服务的顺利开展;对内基于大数据和AI等技术,提供舆情分析、决策辅助等服务,利用智能化手段助力政务服务的高效开展,提高政府服务力、公信力,提升群众满意度。

2.3.4 AI质检通信监管

通话监管是AI质检产品一个较新的应用场景,与面向客服水平提升的质检场景不同,此场景下的目标是检验通话过程是否存在违规的红线业务,如互联网金融推销、诈骗等情况,以及号码在使用过程中是否存在不符合登记用途的通话行为,并不关心业务逻辑和话术的匹配性以及语调语速的规范性。

该场景的难点在于大多通话无标准话术,故在质检流程上引入了2个环节,分类器初步分类与关键词比对。分类器初步分类是指采用深度学习的方式,使用标记好分类的大规模数据集对分类器进行训练,将分类器训练成为能够理解篇章级对话并能通过对话识

别其所属行业的分类器,再将此分类器对实际录音进行推断。推断分类后的录音拥有了分类的标签,再根据分类标签对应的关键词库进行关键词比对,对录音进行最终的违规性判别确认。AI质检和传统质检的全量文本比对相比,引入了分类环节,省去了关键词库每个词汇要全文搜索的计算步骤,计算效率更高,误报率与漏检率更低,更适合于大话务量的语音平台质检。

3 AI语音产品与云呼叫中心对接方案

在技术浪潮推动下,面向呼叫中心、客服中心的智能化升级是大势所趋,但AI语音产品如何更好地内嵌到已有的呼叫中心系统中,AI语音技术如何更好地赋能呼叫中心业务场景是当下智能化升级面临的主要问题。

一个数字化、云化的呼叫中心系统结构通常分为接入管理、通信管理、基础内核、运营监控管理、开放能力5部分,AI语音产品与呼叫中心系统在基础内核、运营监控管理、开放能力3个部分具有耦合关系(见图5)。

3.1 AI坐席与呼叫中心产品对接

早期的AI语音技术是通过媒体资源控制协议(MRCP)以语音识别(ASR)和语音合成(TTS)这种AI语音原子能力的方式,面向语音通信产品提供服务。该协议依赖会话发起协议(SIP)来协调和管理MRCP客户端和服务端之间的会话,依赖会话描述协议(SDP)来描述、发现和进行媒体交互,并依赖SIP和SDP来与媒体服务端之间建立媒体会话交互。当媒体通道建立后,客户端可以通过SIP会话控制语音资源服务端上的媒体处理资源。该协议最后一个版本定稿于2012年,它的缺点在于未考虑到NLP能力,也未考虑人机对话中常见的打断、静音等拟人化场景,故呼叫中心平台要支持静音检测、打断识别、语义分析、情绪识别等功能,并要维系多轮对话、处理话术逻辑、把握人机切换时机,对呼叫中心要求过高,AI整体拟人化效果很低,并非主流方案。

当前AI坐席是作为整体能力提供服务的,即ASR、NLP、TTS集成在一起,其优势在于支持静默提醒、打断功能,拟人化程度非常高;可根据需求编辑跳转人工坐席的话术节点,减少了对呼叫中心平台的适配要求;话术可基于深度学习技术针对不同行业背景进行训练与调优,话术制作可以通过录音师录音替代TTS合成,进一步提升了通话过程中的客户体验。整

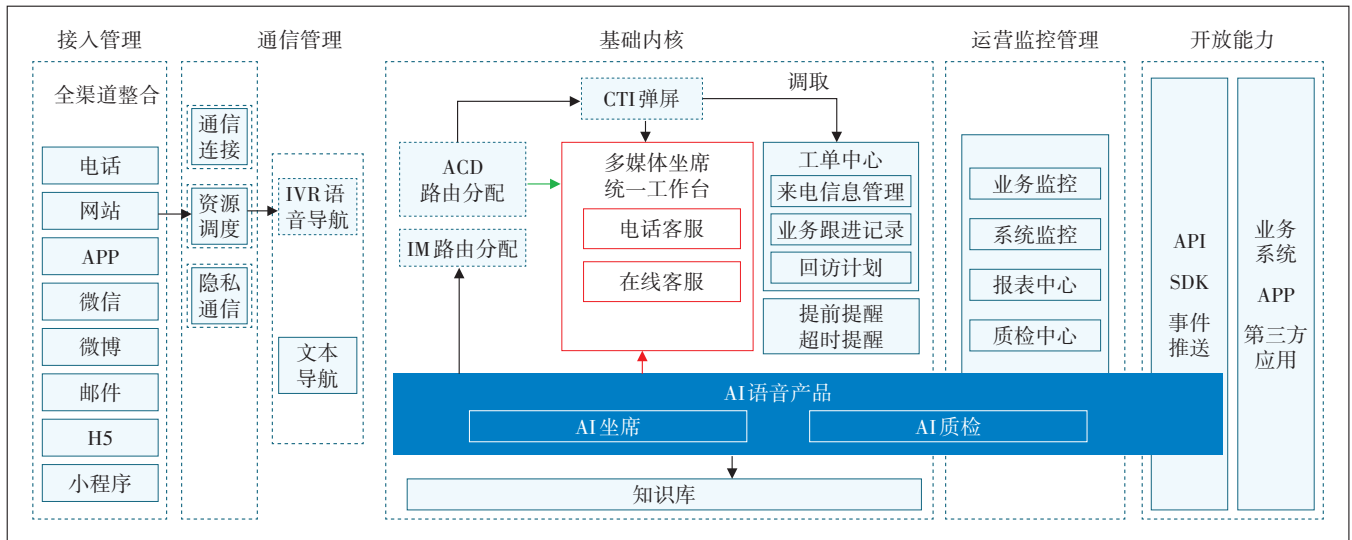


图5 AI语音产品与呼叫中心产品耦合关系图

能力调用有2种具体实现方案,每种方案皆支持机器人意向筛选(AI预测式外呼)、在线转人工坐席(人机耦合)、坐席辅助(监听对话主动介入)的应用场景。

3.1.1 AI坐席API对接

该对接方案是指呼叫中心使用AI语音产品提供的Restful API对其能力进行整体调用,步骤如下:

a) 呼叫中心平台获取 APP_KEY 和 APP_SECRET,配置信息回调地址,用于调用API并接收AI回传的数据。

b) 呼叫中心平台调用创建任务接口,选择话术、设置AI坐席数量,并将AI坐席设置成呼入坐席状态;呼叫中心平台调用客户导入接口,将CRM中的客户名称和电话号码批量导入到任务中。

c) 呼叫中心平台通过双向回拨的方式(双呼)发起呼叫,其中AI语音平台需要将AI坐席的状态(如未呼通、对方忙线、对方挂断)实时通知呼叫中心平台。

d) 当需要AI坐席与人工坐席切换时,AI语音平台话术节点自动调用接口告知呼叫中心平台,呼叫中心平台将话务从AI坐席和被叫的通话切换至人工坐席和被叫的通话。

3.1.2 AI坐席SIP Trunk方式对接

该对接方案是指呼叫中心向AI语音产品提供SIP中继功能,由AI语音平台借助呼叫中心的话务通道发起呼叫,步骤如下:

a) 呼叫中心提供人工坐席队列呼入地址(SIP URI)供AI坐席转接呼叫。

b) 在AI话术中配置转人工节点和转接地址(SIP

URI)。

c) AI语音平台以SIP Trunk形式与呼叫中心建立互联互通,根据任务发起呼叫。

d) AI坐席与客户通话过程中触发了转人工节点, AI语音平台通过SIP协议将呼叫送到呼叫中心人工坐席队列,同时将AI与客户的通话数据传入呼叫中心,由人工坐席与客户继续通话。

3.1.3 不同对接方式对比分析

方式1的主要调用方是呼叫中心平台,优势在于整个呼叫流程由呼叫中心平台进行把控,呼叫频次与业务处理并发能力配置比较灵活;劣势在于呼叫中心平台需要做一些适配开发,且占用了2个大网号码与语音通道。

方式2的主要调用方是AI语音平台,其优势在于仅使用一个号码与语音通道;劣势在于并不是所有呼叫中心都支持SIP Trunk方式对接,或呼叫中心的业务逻辑处理层无法感知通信层的请求,人工切换的业务回调无法实现。

3.2 AI质检与呼叫中心产品对接

AI质检产品根据实时性的需求可分为离线质检与实时质检,离线质检的系统对接方式较为简单,通过http或ftp的方式在呼叫中心平台拉取或由呼叫中心平台推送录音文件与随路数据即可,下文主要介绍实时质检系统对接方案。

3.2.1 AI质检协议方式对接

该方案使用MRCP协议,将ASR原子能力提供给呼叫中心平台,但涉及到主被叫信息以及随路数据的

传输,需要对协议进行修改,步骤如下:

a) 呼叫中心平台呼出或呼入时,通过MRCP协议调用AI语音平台的AI实时转写能力。

b) 呼叫中心平台将通话媒体数据与信令中的主被叫信息传给AI语音平台的质检系统进行实时分析。

c) 将质检后信息以文本方式发至人工坐席所在PC,实现坐席弹屏的话务辅助。

d) 当质检过程中识别到敏感词时,AI语音平台通知呼叫中心平台进行话路强制拆除。

3.2.2 AI质检数据抓取方式对接

该方案通过交换机镜像口获取呼叫中心平台的所有数据,并通过协议分析的方式提取主被叫信息与媒体流,进行解码、转写与质检,步骤如下:

a) AI语音平台的抓包服务器与呼叫中心平台接到同一个交换机上,其中交换机需做镜像口配置。

b) AI语音平台质检系统通过抓包服务器将抓取通话媒体数据与信令中的主被叫信息,传给AI语音平台的质检系统进行实时分析。

c) 将质检后信息以文本方式发至人工坐席所在PC,实现坐席弹屏的话务辅助。

d) 当质检过程中识别到敏感词时,AI语音平台通知坐席,告警提示人工坐席手动拆除话路。

3.2.3 不同对接方式对比分析

方式1主要通过协议方式实现系统对接,其优势在于通话数据在呼叫中心平台进行了复制,一路传送至坐席,一路传送至质检系统进行分析,实时性较好,且可以通过告知呼叫中心平台来进行话路强拆;其劣势在于协议标准没有坐席、技能组等业务侧的关键字段,呼叫中心平台对于非标准字段需要做适配开发。

方式2主要通过抓包方式实现系统对接,其优势在于呼叫中心平台无需做任何更改;劣势在于需要网络设备的配合,对抓包服务器的网络吞吐性能要求较高;抓包与协议解析增加了处理时延;无法对话路进行强拆仅能以文本形式告警给坐席端。

4 思考与展望

以AI坐席产品与AI质检产品为代表的AI语音产品实现了主流AI技术的快速产品化落地,短短几年内产品成熟度不断提高,并几乎在各个垂直行业都创造出极具价值的商业案例,突破了语音业务传统技术带来的瓶颈,产生了巨大的经济利益与社会效益。

然而,AI产品对人工的完全替代还是遥不可及

的,在呼叫中心或客服中心,AI坐席作为人工坐席入口的第一防线,这种辅助性地位仍将长期保持不变。随着AI语音技术的进一步演进,AI语音产品的技术指标也会不断提升,但如自然语言理解、机器学习中的自学习技术需有颠覆性的突破,AI的理解力、回答力、执行力才会有质的飞跃。

在可预期的未来,AI语音能力将会与AI视频能力、文本信息的数据挖掘能力进行进一步整合,输出一个感知智能与认知智能相结合的完整AI能力,并支持渠道融合、数据融合的各类扩展,在对复杂语义、复杂情感的判断上,对工单流转、业务流程的理解上,对垂直行业细分领域的积累上,都会明显优于人工。届时,围绕AI语音技术的智能交互结合各类渠道,将触及行业、企业、个人所有沟通关系,智能交互产品将以独立自主营销拓客、独立自主解决用户问题为根本导向,进行主动联络、意图判断乃至管理决策,营造出崭新的政企智能服务生态。

参考文献:

- [1] 朱频频. 智能客户服务技术与应用[M]. 北京:中国铁道出版社有限公司,2019:120-130.
- [2] 汤晓鸥,陈玉琨. 人工智能基础[M]. 上海:华东师范大学出版社,2018.
- [3] 焦李成,赵进,杨淑媛,等. 深度学习、优化与识别[M]. 北京:清华大学出版社,2017.
- [4] 汤晓鸥,陈玉琨. 人工智能基础[M]. 上海:华东师范大学出版社,2018.
- [5] 韩纪庆,张磊,郑铁然. 语音信号处理[M]. 3版. 北京:清华大学出版社,2019.
- [6] 郑捷. NLP汉语自然语言处理原理与实践[M]. 北京:电子工业出版社,2017.
- [7] 约阿夫·戈尔德贝格. 基于深度学习的自然语言处理[M]. 车万翔,郭江,张伟男,等,译. 北京:机械工业出版社,2018.
- [8] 任华,杨迪,汪少敏. 基于外呼的智能演进网络建设[J]. 电信科学,2019(4):131-138.
- [9] 李兵. 银行业智能客服应用研究与展望[J]. 中国金融电脑,2020,(3):51-54.
- [10] 贺文. 智能客服“倒逼”改变[J]. IT经理世界,2018(10):25-27.

作者简介:

陈浩然,毕业于电信科学技术研究院,高级工程师,硕士,主要从事AI语音业务、视频通信业务、能力开放产品研发工作;孙记明,毕业于北京邮电大学,高级工程师,硕士,主要从事基础通信语音产品、AI语音产品运营工作;刘牧寅,毕业于北京交通大学,高级工程师,硕士,主要从事音视频、网络策略、消息类产品研发及管理工作。