头豹 LeadLeo

400-072-5588

头豹研究院 | 人工智能系列深度研究

2019 年
中国语音助手行业研究报告

行业走势图



TMT 团队

卢佩珊 分析师

邮箱：cs@leadleo.com

相关热点报告

- 人工智能系列深度研究——科创板上市，人工智能行业借势迎来发展
- 人工智能系列深度研究——沙利文助力 2019 全球人工智能健康峰会成功举办，共同推动全球人工智能技术和产业发展
- 人工智能系列深度研究——谷歌上线“猜画小歌”，智能技术与人们生活越来越近
- 人工智能系列深度研究——产品智能化浪潮开启，智能家居风口已至

报告摘要

各企业开发和掌握的语音助手技术的本质区别逐渐缩小，技术壁垒效应逐渐被削弱，技术已不再是企业构建商业模式的最主要基础。全产业链布局及语音助手生态体系构建将成为语音助手服务商持续发展的主要途径，应用场景的开发将是语音助手厂商未来的主要关注点。目前智能手机、智能家居以及智能车载三大场景发展基础良好，需求较大，成为各类型厂商激烈争夺的领域。

热点一： 语音助手市场前景广阔

语音交互成为人工智能的必然趋势，语音助手逐渐与各类型人工智能产品深度融合。随着产品智能化的深入，搭载语音助手的智能设备数将逐步增长。搭载语音助手的智能设备的消费量亦会有所增长，推动中国语音助手市场规模进一步扩大。

热点二： 语音助手应用场景逐渐丰富

语音助手的场景化解决方案将逐渐丰富，语音交互方式作为人机交互的重要演进方向，正逐步渗透到人们的日常生活中，与各类型应用场景相结合。

热点三： 语音助手未来将可听、可视、可触摸

随着人工智能的智能程度提升，语音、图像、视频等模态将趋于融合，语音助手可视化将成为可能。类比人类的语言感知过程，智能人机交互也将向语音、图像、视频等多模态融合方向发展，语音助手通过收集声音、语言、视觉、触觉等感知信息，以自然语言处理为主，辅以唇语识别、表情认定、姿势识别、情绪感知等技术，将数据多维度融合以实现人工智能从感知转变为认知。

目录

1 方法论 5

1.1 方法论 5

1.2 名词解释 6

2 中国语音助手行业市场综述 7

2.1 语音助手定义 7

2.2 语音助手相关技术 7

2.3 语音助手应用领域 8

2.4 中国语音助手产业链分析 10

2.4.1 产业链上游 12

2.4.2 产业链中游 13

2.4.3 产业链下游 13

2.5 中国语音助手市场规模 14

3 中国语音助手行业驱动因素 16

3.1 基础硬件设施进步为行业发展提供基础 16

3.2 智能语音技术的进步提高助手的智能程度 16

4 中国语音助手行业制约因素 18

4.1 语音识别技术未解决实际场景识别问题 18

4.2 机器的自然语言理解能力仍显不足.....18
4.3 数据孤岛导致数据资源得不到全面开发.....19
5 中国语音助手行业相关政策法规.....20
6 中国语音助手行业发展趋势.....22
6.1 语音助手应用场景逐渐丰富.....22
6.2 语音与视觉等模态相互融合.....22
7 中国语音助手行业竞争格局.....24
7.1 中国语音助手行业竞争现状分析.....24
7.2 中国语音助手行业典型企业分析.....25
7.2.1 苏州思必驰信息科技有限公司.....25
7.2.2 北京云知声信息技术有限公司.....26
7.2.3 北京智齿博创科技有限公司.....27

图表目录

图 2-1 中国语音助手产业链.....10

图 2-2 中国用户对语音助手的关注面，2018 年.....14

图 2-3 中国语音助手市场规模（按收入计），2014-2023 年预测.....15

图 5-1 中国语音助手行业相关政策.....21

1 方法论

1.1 方法论

头豹研究院布局中国市场，深入研究 10 大行业，54 个垂直行业的市场变化，已经积累了近 50 万行业研究样本，完成近 10,000 多个独立的研究咨询项目。

- ✓ 研究院依托中国活跃的经济环境，从人工智能、大数据、智能语音等领域着手，研究内容覆盖整个行业的发展周期，伴随着行业中企业的创立，发展，扩张，到企业走向上市及上市后的成熟期，研究院的各行业研究员探索和评估行业中多变的产业模式，企业的商业模式和运营模式，以专业的视野解读行业的沿革。

- ✓ 研究院融合传统与新型的研究方法，采用自主研发的算法，结合行业交叉的大数据，以多元化的调研方法，挖掘定量数据背后的逻辑，分析定性内容背后的观点，客观和真实地阐述行业的现状，前瞻性地预测行业未来的发展趋势，在研究院的每一份研究报告中，完整地呈现行业的过去，现在和未来。

- ✓ 研究院密切关注行业发展最新动向，报告内容及数据会随着行业发展、技术革新、竞争格局变化、政策法规颁布、市场调研深入，保持不断更新与优化。

- ✓ 研究院秉承匠心研究，砥砺前行的宗旨，从战略的角度分析行业，从执行的层面阅读行业，为每一个行业的报告阅读者提供值得品鉴的研究报告。

- ✓ 头豹研究院本次研究于 2019 年 8 月完成。

1.2 名词解释

- **LibriSpeech**：语音数据，内容包括 1,000 小时的英文发音和对应文字。
- **麦克风阵列**：设置两组以上麦克风，对音讯进行侦测，所得到的资料交由数字信号处理器进行比对，用以还原声音的原貌，并消除背景杂音。
- **鲁棒性（Robust)**：系统的健壮性，指系统在不同应用条件下的性能稳定性。
- **语料**：语言材料，是构成语料库的基本单元。

2 中国语音助手行业市场综述

2.1 语音助手定义

语音助手是以语音识别技术和语音合成技术为基础，通过语音信息实现人机交互，给用户反馈信息或替代用户及企业执行任务的虚拟助手。

2.2 语音助手相关技术

语音助手是机器感知与处理人类语音信息的应用，需要人工智能语音相关技术，如语音识别技术、语音合成技术以及自然语言处理技术的辅助，赋予其语音感知与处理能力。

（1） 语音识别技术

语音识别技术以语音为研究对象，通过识别与处理语音信号使计算机自动识别和理解人类口述的语言。计算机通过语音识别和理解的过程将人类口述的语音信号转变为机器可处理的文本。计算机的语音识别过程可概括为：系统对语音信号进行分析和处理，除去冗余信息，提取语音信号中的关键信息和表达语言含义的特征信息，并按照不同语言的语法规则给关键信息划分段落，取出已识别出的字词连接成句，结合语义理解适当调整句子构成，修正已生成的句子内容。

（2） 语音合成技术

语音合成技术是计算机将文本信息转换成语音的技术，赋予机器“讲话”能力。语音合成技术的实现过程主要包括文本分析和语音合成两个步骤。文本分析是计算机基于语言学原理，将文本标准化的过程。计算机将自然语言文本中的数字、缩略语等转换为系统标记词后进行语义理解，并为文字序列配上语音脚本，通过语音脚本将文字序列转换成音韵序列。语音合成是计算机将音韵序列合成为语音波形，输出语音流的过程。

（3） 声纹识别技术

声纹识别技术是计算机通过语音信号提取发声者独有的声门开合频率、口腔大小形状及声道长度等声学特征识别发声者身份的技术。声纹是一种承载语音频谱（语音、语调等）的音频信息，不同生物个体的发音器官具有独特性，发出的声纹信息各不相同，声纹信息可用于身份鉴别。声纹识别技术可用于辨认发声者以及确认发声，即从若干发声者中寻找指定发声者以及确认声音是否由指定发声者发出。

（4） 自然语言处理技术

自然语言处理是计算机理解和生成自然语言的过程，自然语言处理技术使计算机具有识别、分析、理解和生成自然语言文本（包括字、词、句和篇章）的能力。自然语言处理机制涉及自然语言理解和自然语言生成两个流程：①自然语言理解：计算机理解自然语言文本的思想和意图；②自然语言生成：计算机用自然语言文本表述思想和意图。

2.3 语音助手应用领域

语音助手的应用场景广阔，众多智能设备通过接入语音助手提升智能化水平。语音助手具体的应用场景可分为消费级应用场景和行业应用场景。

（1） 消费级应用场景

①语音助手+车载

按键或触控操作对驾驶员而言安全系数较低，在驾驶过程中使用容易引发安全事故。语音控制在车载场景中是替代按键或触控的最佳方案，既能避免驾驶员分心，又能提升驾驶员的驾驶体验。各大汽车厂商开始在前装控制系统中搭载语音控制方案，通过语音助手处理驾驶员对导航、音乐搜索与播放、信息听写等任务的指令。众多语音助手行业巨头纷纷推出语音车载产品，如科大讯飞推出汽车语点系统，百度推出 Carlife 智能车载系统支持语音声控，并与部分汽车厂商达成合作，实现其语音助手在车载场景的投放。

②语音助手+家居

语音助手是智能家居的入口端，智能家居通过语音助手与用户交互。语音助手可与空调、电视、音响、灯具、玩具等家用设备和智能家居控制中枢系统相结合，控制互联的智能家居产品，如长虹、LG、TCL 等主流的智能电视内嵌了语音助手，用户通过命令语音助手可实现查收影视剧、听歌、发微博、开关电视等功能。

③语音助手+智能手机

为了提升用户体验，智能手机普遍配备语音助手功能，如 iphone 配备 Siri，小米手机配备小爱同学等。智能手机通过语音助手解放用户双手，进一步提升其智能水平。

（2） 专业级行业应用场景

①语音助手+医疗行业

语音助手在医疗行业的应用涉及语音导诊机器人、临床病历的语音录入及转写、智能辅助诊疗等。随着市场需求的增加，语音助手企业纷纷入局医疗行业，例如，科大讯飞的智能语音系统已在全国多家医院落地使用，科大讯飞与北京大学口腔医院口腔数字化医疗技术和材料国家工程实验室共建基于语音的门诊病历采集系统，并将该系统在医院实现试点投放。

②语音助手+教育行业

语音助手在教育行业的应用已深入到教育体系的“学、练、测、评”全部环节，主要产品有智能教育机器人、语音测评、人机口语对话训练、互动教学等。例如，科大讯飞将语音助手应用在口语训练、口语考试以及儿童早教智能硬件产品中。

③语音助手+客服行业

语音助手在客户服务领域的应用较为广泛，覆盖金融、交通、电信、旅游等行业。机器人客服相比传统客服，具有成本低、全天候 24 小时坐班、时刻待命等优点。随着人力成本上升，企业对智能客服的需求将会逐步加大。

2.4 中国语音助手产业链分析

语音助手产业链上游市场主体为基础资源提供商，包括硬件供应商（如芯片供应商、服务器供应商和存储供应商等）和软件供应商（如云服务供应商和数据库供应商等）；中游市场主要由语音助手技术提供商组成，包括语音助手算法供应商、语音助手软件应用供应商以及语音助手行业解决方案提供商，负责将语音助手技术转化为具体应用或解决方案，为下游需求端提供服务。语音助手下游应用场景广阔，覆盖智能移动设备、智能家居、车载、金融以及客服等场景，涉及 B 端和 C 端用户主体，包括各类型机构用户、企业用户和个人用户。

图 2-1 中国语音助手产业链



来源：头豹研究院编辑整理


innovation map 创新地图
推广
前哨 2020 科技特训营
掌握创新武器 抓住科技红利
Insights into Tech and the Future
直播时间
每周四20:00-21:00
全年50次直播课程
+私享群互动
随报随听
王煜全
海银资本创始合伙人
得到《全球创新260讲》主理人
扫码报名
微信咨询：InnovationmapSM
电话咨询：157-1284-6605

2.4.1 产业链上游

语音助手产业链上游市场由基础资源供应商组成，涉及网络设备、服务器、芯片、存储、云服务、数据库等软、硬件供应商，负责为语音助手技术和产品开发商提供必要的资源支持。

（1） 芯片

人工智能芯片作为语音助手产业的核心，芯片的技术成熟度影响语音助手的性能。人工智能芯片行业的技术门槛较高，高端芯片技术均被海外芯片巨头垄断，技术壁垒高。中国人工智能芯片依赖进口，导致语音助手服务商的制造成本居高不下，压缩企业利润空间。

人工智能芯片按不同的应用场景可分为通用类人工智能芯片、云端人工智能芯片、终端人工智能芯片。①在通用类人工智能芯片领域，美国的英伟达占据主导地位，中国的人工智能芯片企业实力较薄弱，缺乏国际竞争力。②在云端人工智能芯片领域，英特尔、亚马逊等海外企业占据较大的市场份额，部分中国企业陆续布局云端推断市场，但竞争力较弱。③在终端人工智能芯片领域，中国企业取得较大进展寒武纪、地平线和深鉴科技等一批优质企业相继诞生。

（2） 云服务

云服务供应商为语音助手研发企业提供基础设施平台，解决语音助手技术研发厂商的数据存储、运算以及调用问题。由于性价比、部署方式等因素，语音助手研发企业较多选用公有云服务。

目前，公有云服务供应商有：①通过云服务产业链资源优势拓展至公有云服务行业的企业，如电信运营商、网络设备制造商、IDC 厂商等，此类企业拥有较强的资金实力，加上本身处在公有云产业链上游，基础设施方面优势明显；②大型互联网企业，如亚马逊、腾讯、阿里巴巴等，此类企业资金实力雄厚，客户认可度高，设施齐备、技术成熟，具备发展公有云业务的有利条件；③传统的软件企业，如 Microsoft、Oracle、金蝶等，此类企业的软件

产品的市场认可度高，技术积累丰厚，客户资源丰富，有利于向公有云市场拓展。除此之外，行业中存在不少新兴的创业公司，如青云、Ucloud、七牛云等。

2.4.2 产业链中游

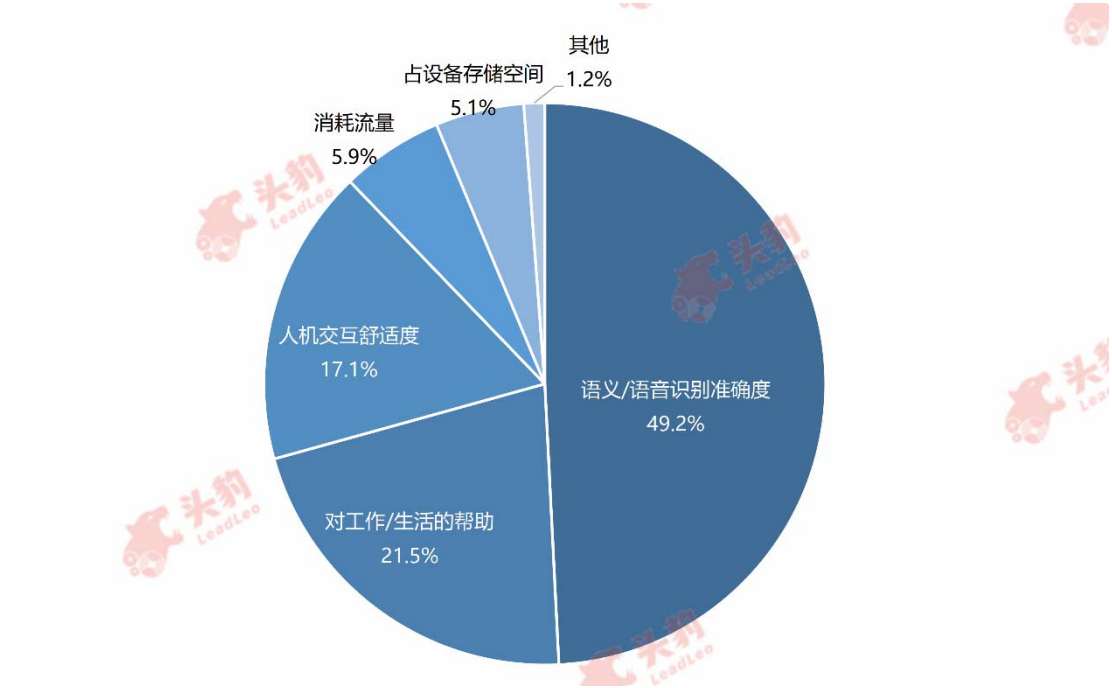
语音助手产业链中游市场主体主要有语音助手算法提供商、行业解决方案提供商以及应用产品开发商。中国的语音助手厂商较多集研发算法、解决方案以及应用产品功能于一身，厂商自主研发语音助手算法，形成一整套语音助手关键技术方案，并将自主研发的语音助手算法以及技术方案内嵌于硬件产品中，典型代表企业有百度、腾讯以及阿里巴巴。例如百度开发了度秘，并将度秘置于百度搜索、百度地图、百度新闻等软件产品以及百度音箱等硬件产品中。部分语音助手技术提供商专注于智能语音技术研发，为下游应用领域提供定制化的语音技术解决方案，如 OPPO 手机的小欧助手和华为手机的小 E 助手均采用科大讯飞的语音技术解决方案。

2.4.3 产业链下游

语音助手下游应用场景广阔，涉及金融、医疗、教育、汽车、消费电子产品等行业，其中智能手机、智能家居以及车载应用是语音助手在 C 端市场的主要应用场景；智能客服、机器人是目前语音助手在 B 端市场能产生营收的主要应用场景。

目前语音助手技术发展并未成熟，语音助手的用户体验尚待提升。头豹研究院调研数据显示，用户最在意语音助手的语义 / 语音识别准确度（占比达 49.2%），68.2%的受访用户认为语音助手的语义/语音识别准确度有待提升。

图 2-2 中国用户对语音助手的关注面，2018 年



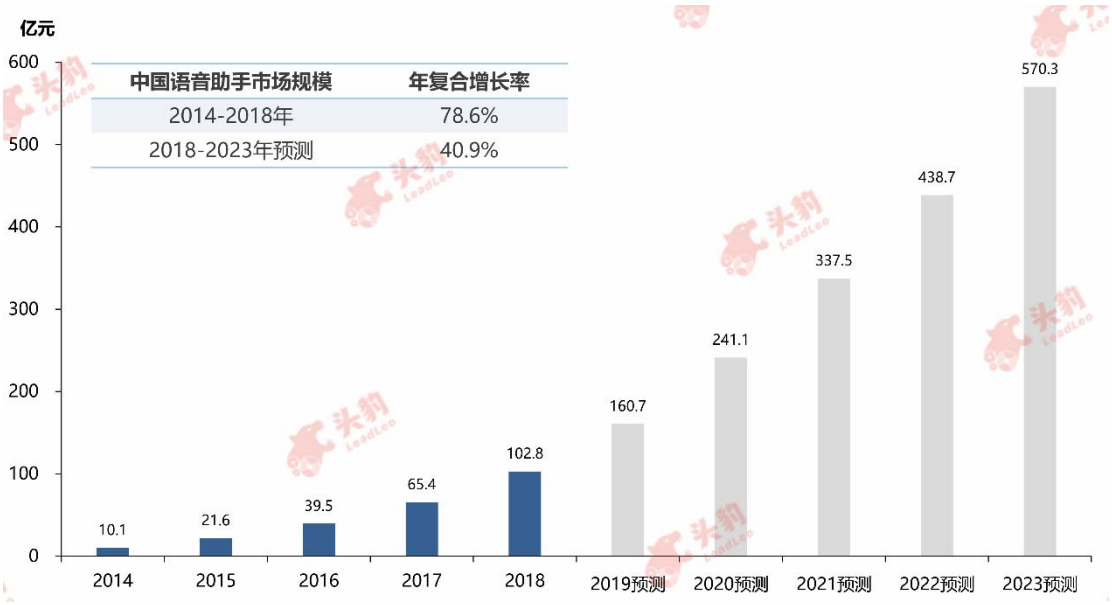
来源：头豹研究院编辑整理

2.5 中国语音助手市场规模

现阶段，搭载语音助手的设备产品主要有智能手机、智能家居产品（如智能冰箱、智能空调、智能洗衣机、智能照明等）以及智能汽车，其中智能手机是语音助手的主要载体；智能家居产品丰富，且产品具有刚性需求，为语音助手提供丰富的落地载体选择；智能汽车正逐步替代传统汽车，其市场潜力带动语音助手需求提升。2014-2018 年间，中国语音助手市场规模实现快速增长，年复合增长率达 78.6%。

语音交互成为人工智能的必然趋势，语音助手逐渐与各类型人工智能产品深度融合。随着产品智能化的深入，搭载语音助手的智能设备数将逐步增长。由于中国居民消费水平提升，消费主力群体代际变迁等因素，搭载语音助手的智能设备的消费量将会有所增长，推动中国语音助手市场规模进一步扩大，预计到 2023 年，中国语音助手市场规模将增长至 570.3 亿元人民币。

图 2-3 中国语音助手市场规模（按收入计），2014-2023 年预测



来源：头豹研究院编辑整理



3 中国语音助手行业驱动因素

3.1 基础硬件设施进步为行业发展提供基础

人工智能芯片的发展提高语音助手行业数据处理能力，增强语音助手的语音交互能力，有利于提升用户体验。人工智能芯片的算力决定了计算机处理语音数据的速度以及语音交互的智能程度，在传统的云端计算方案中，终端设备将数据全部上传至云端，通过云端计算分析后将结果传送回终端，庞大的数据和运算量导致设备运行成本和功耗居高不下，且容易导致终端设备人机交互延时以及卡顿，影响用户体验。

终端人工智能芯片的发展，解决了数据在终端的运算问题，终端深度集成语音芯片将芯片和语音处理算法组合成模块，置于终端设备中，赋予终端设备语音数据处理能力，当用户向设备发送指令，麦克风接受语音信息并将其转换成电信号，芯片模块处理分析电信号，实现终端的实时数据处理，减少因数据传输造成的反应延时。此外，终端深度集成语音芯片拥有较大的数据存储空间，可将语料数据保存于终端，保证隐私数据的安全性。终端深度集成语音芯片的发展，既能减少向云端传输的数据量，降低终端设备运行成本，又能提高语音助手的反应速度，提升用户体验，为语音助手的推广应用提供坚实的基础。

3.2 智能语音技术的进步提高助手的智能程度

智能语音技术进步提升了语音助手的智能化程度，助力语音助手产品的推广。现阶段的语音合成技术已接近人类的发音水平，语音助手开始拥有语调、情绪的变化，能更好地吸引用户与其交互。语音识别技术随着深度神经网络的广泛使用，语音识别准确率取得飞跃式的突破，2018 年 10 月云从科技在开源语音识别数据集 LibriSpeech 上刷新世界纪录，其语音识别准确率可达 97.03%。新一代语音识别算法框架、口语化语音识别、个性化语音识别、智能对话、音视频融合以及语音合成等基础理论研究和技术研究的资源投入进一步加大，部

分场景和领域的技术研发有重大突破，在远距离和复杂场景下的语音采集与识别技术有所提升，如思必驰推出的“环形 6+1 远场”麦克风阵列方案，利用 6 个麦克风呈环形 360°排列，环形中间放置 1 个麦克风辅助拾音，有效地对远场语音进行抗噪处理，其远场 5 米范围内的语音识别准确率达 92%，该方案目前已在智能家居场景中被广泛运用。针对医疗、教育、商贸等专用词汇较多的专业领域，语音助手技术亦有针对性的优化，使语音助手能更好地适应行业发展需求，辅助行业用户智能地执行任务。

4 中国语音助手行业制约因素

4.1 语音识别技术未解决实际场景识别问题

目前的语音识别技术仅在特定领域、特定环境下达到实用化程度，业内企业实现的 97%的语音识别准确率较多是基于安静室内环境产生的人工测评结果，在实际使用场景中，尤其是在自由发音、强噪声、多人同时发声、远端声场等环境下，机器的语音识别性能并不理想。

在语音助手的真实使用场景中，方言、噪音、远场、断句等情况不可避免，语音采集地点变更导致的环境及背景噪音差异、语音信号传输载体方式多样导致的传输信道差异，都将造成语音信号的频谱畸变，导致语音系统难以准确识别用户话语，对用户需求的理解产生偏差。语音识别技术的成熟度将影响语音助手应用产品生态的构建与发展，语音识别技术仍需提升抗干扰能力，其“鲁棒性”问题亟待解决。

4.2 机器的自然语言理解能力仍显不足

现阶段，语音助手的自然语言理解能力与人类助手的自然语言理解能力相差甚远，语音助手为用户带来的人机交互体验不佳，与用户的生活结合程度依旧不深。语音助手与用户日常生活的结合深度较大程度上取决于语音助手的人机交互能力，拥有良好交互能力的语音助手能方便用户支配，但机器的人机交互能力受自然语言理解能力的影响，目前机器对自然语言的理解准确度不高，导致用户对语音助手的人机交互体验不佳。机器由于难以拥有生活常识、文化背景和人类风俗习惯知识储备，对于作为文化载体的自然语言，无法结合具体的语言环境、风俗习惯等因素分析语言内容，只能通过系统设定的规则机械地分析源语的语法结构，语言逻辑无法处理，对于复杂句子结构或一词多义的情况，机器的理解能力有限。机器不具备上下文功能，无法将聊天中的信息与上下文关联，遇到一词多义的情况，机器无法像人类般基于已有知识储备和上下文环境确定多义词的意思，导致语音助手无法准确判定用户

的意图，人机交互不顺畅。

此外，现阶段的自然语言处理模型未能解决用户表达中的常识问题和个性化问题，导致用户对语音助手指派命令时需要使用特定用语，严重影响用户的人机交互体验，如用户对手机语音助手发号“查找附近的餐馆”指令时，语音助手会在地图上执行查找附近餐馆的任务，并显示附近餐馆信息，但若用户说“我饿了”，手机语音助手则不会有任何反应，因为语音助手缺乏“饿了需要进食”的常识，无法为用户自动执行查找餐馆的任务，而人类助手听了“我饿了”的话语，就知道雇主“需要进食”，需要为雇主执行与进食相关的任务。语音助手无法全面理解用户的表达，其不智能程度大幅降低用户与其交互的意愿。

4.3 数据孤岛导致数据资源得不到全面开发

数据是人工智能发展的基石，海量数据为训练人工智能提供原材料。中国人口基数庞大，可所产生海量的数据资源为智能语音的发展提供充足的训练素材，但目前中国的智能语音行业仍缺乏一个标准的跨平台数据共享系统，巨头企业掌握大量数据资源，不对外开放，不利于小企业或初创企业的数据积累。与此同时，各企业利用自有数据资源独立研发，各类型语音软件系统兼容度不高，进一步阻碍数据资源的自由流通。缺少数据的智能语音企业难以提高模型的识别准确度，影响产品的投放以及市场份额的拓展，企业的发展受限不利于行业的快速扩张。缺少统一的数据共享系统在一定程度上会制约中国智能语音行业的快速发展，未来，相关部门需要对数据共享制定规范与准则，帮助数据有效分享，促进数据资源的合理运作。

5 中国语音助手行业相关政策法规

2015 年以来，中央及各地方政府陆续出台相关政策，积极支持人工智能产业的发展，并推进人工智能在各个细分领域的渗透和壮大。2015 年 7 月，国务院发布《关于积极推进“互联网+”行动的指导意见》，将“互联网+人工智能”列为其中 11 项重点行动之一。2017 年 7 月，国务院印发《新一代人工智能发展规划》，将人工智能上升至国家战略，并确立了人工智能发展三步走的目标，为中国人工智能产业发力奠定制度基础。2017 年 10 月，十九大将“人工智能”写进党代会报告，并提出加快建设制造强国，发展先进制造业，推动互联网、大数据、人工智能和实体经济深度融合等意见。截至 2018 年，中央及各地方政府在短短 3 年内出台了 60 余项与人工智能产业发展相关的政策，足以看出国家对人工智能产业化的重视。智能语音助手是人工智能的应用体现，同样受到政策的鼓励发展。

2016 年 5 月，发改委、科技部、工信部及中央网信办联合发布《“互联网”+人工智能三年行动实施方案》要求加快建设文献、语音、图像、视频、地图等多种类数据的海量训练资源库和基础资源服务公共平台，建设支撑超大规模深度学习的新型计算集群，建立完善产业公共服务平台。进一步推进计算机视觉、智能语音处理、生物特征识别、自然语言理解、智能决策控制以及新型人机交互等关键技术的研发和产业化，为产业智能化升级夯实基础。

2017 年 7 月，国务院发布《新一代人工智能发展规划》，要求加快培育人工智能产业领军企业。在无人机、语音识别、图像识别等优势领域加快打造人工智能全球领军企业和品牌，推动国内优势企业、行业组织、科研机构、高校等联合组建中国人工智能产业技术创新联盟，支持龙头骨干企业构建开源硬件工厂、开源软件平台，形成集聚各类资源的创新生态，促进人工智能中小微企业发展，支持各类机构和平台面向人工智能企业提供专业化服务。

2017 年 12 月，工信部印发《促进新一代人工智能产业发展三年行动计划（2018-2020

年)》，鼓励支持新一代语音识别框架、口语化语音识别、个性化语音识别、智能对话、音视频融合、语音合成等技术的创新应用，在智能制造、智能家居等重点领域开展推广应用。到2020年，实现多场景下中文语音识别平均准确率达到96%，5米远场识别率超92%，用户对话意图识别准确率超90%。

2018年11月，工信部发布《新一代人工智能产业创新重点任务揭榜工作方案》，提出面向语音识别、视觉识别、自然语言处理等基础领域及工业、医疗、金融、交通等行业领域建设高质量人工智能训练资源库、标准测试数据集。到2020年，基础语音、视频图像、文本对话等公共训练数据量大幅提升，在工业、医疗、金融、交通等领域汇集一定规模的行业应用数据，用于支持创业创新。

图 5-1 中国语音助手行业相关政策

政策名称	颁布日期	颁布主体	主要内容及影响
《新一代人工智能产业创新重点任务揭榜工作方案》	2018-11	工信部	提出面向语音识别、视觉识别、自然语言处理等基础领域及工业、医疗、金融、交通等行业领域建设高质量人工智能训练资源库、标准测试数据集。到2020年，基础语音、视频图像、文本对话等公共训练数据量大幅提升，在工业、医疗、金融、交通等领域汇集一定规模的行业应用数据，用于支持创业创新
《促进新一代人工智能产业发展三年行动计划（2018-2020年）》	2017-12	工信部	鼓励支持新一代语音识别框架、口语化语音识别、个性化语音识别、智能对话、音视频融合、语音合成等技术的创新应用，在智能制造、智能家居等重点领域开展推广应用。到2020年，实现多场景下中文语音识别平均准确率达到96%，5米远场识别率超92%，用户对话意图识别准确率超90%
《新一代人工智能发展规划》	2017-07	国务院	要求加快培育人工智能产业领军企业，在无人机、语音识别、图像识别等优势领域加快打造人工智能全球领军企业和品牌，推动国内优势企业、行业组织、科研机构、高校等联合组建中国人工智能产业技术创新联盟
《“互联网”+人工智能三年行动实施方案》	2016-05	发改委 科技部 工信部 中央网信办	要求加快建设文献、语音、图像、视频、地图等多种类数据的海量训练资源库和基础资源服务公共平台。进一步推进计算机视觉、智能语音处理、生物特征识别、自然语言理解、智能决策控制以及新型人机交互等关键技术的研发和产业化，为产业智能化升级夯实基础

来源：头豹研究院编辑整理

6 中国语音助手行业发展趋势

6.1 语音助手应用场景逐渐丰富

语音助手的场景化解决方案将逐渐丰富。语音交互方式作为人机交互的重要演进方向，正逐步渗透到人们的日常生活中，与各类型应用场景相结合。与场景贴合是技术实现商业化的主要途径，为适应不同场景的语音交互需求，语音助手必须贴合不同应用场景，深入剖析场景真实环境，结合各行业专业知识体系，提供针对性的行业或应用解决方案。例如，在智能车载领域，语音助手企业开始从车内真实环境出发，为车厂量身定制具有语音控制、语音导航以及语音呼叫等功能的语音助手方案。在医疗领域，适用于医疗行业的语音助手被广泛用于临床文档改良、临床语音识别、辅助医疗质量把控等工作中，充分协助医护人员完成大量程序化的工作，推进智慧医疗的发展进程。语音助手技术巨头科大讯飞已成功地为全国 100 多家医院提供量身定制的智能医疗语音解决方案。随着传统行业业务处理智能化需求上涨，语音交互方式与传统行业的结合将逐步加深，将为语音助手创造更多可投放场景。

6.2 语音与视觉等模态相互融合

随着人工智能的智能程度提升，语音、图像、视频等模态将趋于融合，语音助手可视化将成为可能。深度学习神经网络的引进使得语言模态、文字模态、图像模态、视频模态的编码和解码可在同一个深度学习框架下统一运行，不同模态的对象可被同一模式编码与解码，同一模式的编码与解码可使不同模态对象随意融合，语音分析结果可与图像分析结果结合应用。语音助手将可模拟人类的语言认知过程，将语音和视觉同时作为语音理解源，不仅通过声音来获取信息，同时还用眼睛观察说话者口型、表情的变化，提升语言感知能力。

类比人类的语言感知过程，智能人机交互也将向语音、图像、视频等多模态融合方向发展，语音助手通过收集声音、语言、视觉、触觉等感知信息，以自然语言处理为主，辅以唇

语识别、表情认定、姿势识别、情绪感知等技术，将数据多维度融合以实现人工智能从感知转变为认知。

7 中国语音助手行业竞争格局

7.1 中国语音助手行业竞争现状分析

中国语音助手服务商可分为以百度为代表的**互联网巨头企业**，以科大讯飞为代表的**智能语音技术巨头企业**以及以云知声为代表的**深耕于垂直领域和细分场景的初创企业。**

互联网巨头企业拥有丰富的客户资源以及完善的产品生态，其可将自主研发的语音助手投放于自有产品生态中，同时可利用丰富的 C 端产品经验及用户数据为语音助手的功能化设计提供基础，基于用户需求驱动产品的发展与落地。与此同时，互联网巨头企业通过开放语音平台，以产业内合作的方式，将语音助手植入合作伙伴的产品或相关业务场景中，实现其语音助手应用场景的快速构建，抢先占领广阔的人工智能设备入口端。

智能语音技术巨头企业拥有较强的技术实力，在语音助手行业拥有先发优势，占据较大的市场份额。但随着深度学习算法的成熟，语音数据的增长，语音模型开发难度以及成本有所下降，智能语音行业的技术壁垒逐渐被打破，智能语音科技企业的先发优势逐渐被削弱，较多企业有能力提供语音助手的算法模型，与语音技术巨头企业形成激烈的竞争。智能语音技术巨头企业原有的语音技术授权、提供语音算法模型、技术维护等业务模式的收入逐渐受限，部分智能语音巨头企业开始从单一的智能语音技术提供商向全方位人工智能技术服务商转型，将语音技术转化为产品与服务，为下游各应用领域提供定制化语音助手产品与服务，进一步巩固并增强语音技术优势。

语音助手领域的初创企业较多以垂直领域和细分场景为突破口，研发单一场景或单一行业的语音助手产品或解决方案，避开与互联网巨头企业以及智能语音技术巨头企业的直接竞争。相较于互联网巨头企业和智能语音技术巨头企业，初创企业的布局范围较小，涉及的下游应用场景较狭窄，对于部分综合实力欠缺的初创企业，专注于语音助手应用的细分场景有

助于提升企业在垂直领域的专业性。

根据在行业前 10 语音助手企业拥有超过 2 年产品战略规划经验的专家分析，由于各企业开发和掌握的语音助手技术的本质区别逐渐缩小，技术壁垒效应逐渐被削弱，基于技术销售的商业模式难以为继，技术已不再是企业构建商业模式的最主要基础，全产业链布局及语音助手生态体系构建将成为语音助手服务商持续发展的主要途径，应用场景的开发将是语音助手厂商未来的主要关注点。目前智能手机、智能家居以及智能车载三大场景发展基础良好，需求较大，成为各类型厂商激烈争夺的领域。

7.2 中国语音助手行业典型企业分析

7.2.1 苏州思必驰信息科技有限公司

7.2.1.1 企业简介

思必驰成立于 2007 年，是研发智能语音交互技术的智能语音服务商，为企业和开发者提供语音交互解决方案，包括 DUI 开放平台、企业级智能服务、人机对话操作系统、人工智能芯片模组等。思必驰已于 2018 年 5 月完成由元禾控股、中民投领投，深创投、富士康、联发科跟投的 5 亿元人民币 D 轮融资。

7.2.1.2 产品研发

（1）语音交互

思必驰为用户提供移动端的语音交互技术支持，适用于 Android、iOS 等系统，可应用于语音助手、智能客服、地图导航等领域，具备语音唤醒、识别、合成以及语义理解等多种能力。在智能车载场景中，思必驰为用户提供“云+端”解决方案，利用智能语音操作系统控制智能后视镜、智能车机、便携式导航仪等设备，实现用户双手解放。

（2）智能家居

思必驰拥有软硬一体化的智能家居解决方案，为智能家居产品提供高性能、低成本的声

源定位、个性唤醒、语音识别、语义理解、对话交互等功能，使家居产品可适应近场、远场的语音交互，方便用户控制，提高家居产品的智能化水平。

7.2.1.3 竞争优势

思必驰拥有自主知识产权的人机对话、语音识别、语义理解、语音合成、声纹识别等综合语音技术，其智能人机对话技术突破了传统语音技术不能很好支持复杂语音交互功能的难题，使语音输入不再局限于呆板简单的句式，在复杂环境和自然口语交流的情况下，能保证优异的语音分析精度和稳健的人机对话性能。思必驰的智能人机对话技术能够提供基于不完整或不准确的语音识别结果进行智能语义推理，通过针对特定领域特定应用需求、对话目标、对话行为、对话状态和对话上下文的统计建模，解决传统系统设计中规则无法覆盖实际对话状态的问题，同时增强系统对于错误识别结果和错误推理结果的自适应性，大幅度提升用户的语音交互体验。

思必驰语音识别技术解决了大量词典和语言模型的识别问题，采用了鲁棒特征提取的方法，能够有效降低噪声环境下的错误率，并用动态噪声自适应来使统一的模型适用于各种噪声环境，支持全国各地方不同口音普通话识别，在用户实际使用环境下，引擎的识别率能保持较高水平。

7.2.2 北京云知声信息技术有限公司

7.2.2.1 企业简介

云知声成立于 2012 年，是智能语音识别技术企业，核心技术有语音技术、语言技术、知识计算、大数据分析等，为用户提供智能家居、智能车载、智慧医疗、智能教育等语音技术解决方案、云知声输入法、语控精灵等产品服务。云知声于 2018 年 7 月完成 6 亿元人民币的 C+轮融资，实现 C 轮系列融资总额达 13 亿元人民币，创下语音技术领域单轮融资的最高记录。

7.2.2.2 产品研发

云知声核心技术依靠自主开发的人工智能芯片、AIUI 和 AI Service 三大解决方案实现落地，在家居、车载、医疗和教育等领域有广泛应用，形成了完整的“云端芯”生态闭环。

云知声推出智能语音一体化解决方案，包括集成人工智能芯片、软件以及集成方案，可协助客户从云端和终端两侧同时部署物联网设备。云知声为用户提供芯片设计开源服务，让更多人工智能产品能采用雨燕 SoC。除了雨燕 SoC 和降噪引擎外，其余部分均根据客户的需求分拆产品，轻量化到全套配备方案齐全，定制化产品出货时间快，可快达 1 个月。

7.2.2.3 竞争优势

云知声专注于研发多款定制型芯片，致力于打造可面向智慧城市场景，并提供对语音和图像等多模态计算支持的多模态 AI 芯片。云知声与吉利集团旗下的生态链企业——亿咖通科技共同打造的面向智慧出行场景的多模态车规级 AI 芯片雪豹（Leopard）于 2019 年实现启动量产。云知声依托 AI 芯片的研发能力，将业务拓展至包括智能家居、智能汽车、智能儿童机器人、智慧酒店、智慧交通等场景。

7.2.3 北京智齿博创科技有限公司

7.2.3.1 企业简介

智齿科技成立于 2014 年 5 月，是一家提供智能客户服务的互联网创业企业。智齿科技将自然语言理解、机器学习及大数据技术进行有效整合，构建了包括云呼叫中心、智能机器人客服、人工在线客服、工单系统、智齿 360、智能外呼机器人在内的智能全客服平台和智能外呼平台，为企业用户提供多样化的智能客服服务，为用户解决移动时代、体验经济环境下传统客服所不能解决的企业客服痛点。智齿科技于 2018 年 4 月获得由云启资本领投，耀途资本、上海原龙投资和博雅盛景等机构跟投的 1.5 亿元人民币 B+轮融资。

7.2.3.2 产品研发

智齿科技推出云呼叫中心、客服机器人、在线客服、工单系统、外呼机器人等 SaaS 客服产品，为金融、教育、电商、企服、生活消费、互娱、政企等领域的客户提供一站式客服解决方案，同时构建了多元丰富的统计分析报表，帮助企业用户分析顾客关注焦点、营销活动结果验收、客服人工工作量及效率，为客服主管提供数据支撑，辅助企业科学经营决策。

智齿客服产品支持桌面网站、移动网站、微信、微博、APP 等接入方式，用户只需把代码复制到网站上即可将多平台顾客汇集于同一平台管理，操作简便。

7.2.3.3 竞争优势

智齿科技的智能客服业务行业服务经验丰富。智齿科技在 26 个行业构建知识图谱，其中 16 个行业构建了 40 个细分场景知识图谱，提供行业知识库及行业客服机器人问答，问答语料库内容覆盖面广，能为众多领域行业用户提供智能客服问答服务，目前智齿科技服务过的企业数量超 5 万家，覆盖电商、互联网金融、生活服务、企业服务、在线教育、互动娱乐等多个细分行业，包括用友、富士康、滴滴出行、趣分期、新东方、搜狐、销售易等知名企业用户。

智齿客服率先将自然语言处理技术及机器学习引入客服产品，实现精准理解用户问题并匹配最佳答案，反馈答案准确率高达 98%。智齿客服支持的机器人客服与人工客服自由切换模式，可解决 80%人工客服的同质化回复问题，同时机器人客服可自动组织标准答案，帮助客服人员大幅提升客服回复质量和效率。

头豹研究院简介

- 头豹研究院是中国大陆地区首家 B2B 模式人工智能技术的互联网商业咨询平台，已形成集行业研究、政企咨询、产业规划、会展会议行业服务等业务为一体的一站式行业服务体系，整合多方资源，致力于为用户提供最专业、最完整、最省时的行业和企业数据库服务，帮助用户实现知识共建，产权共享
- 公司致力于以优质商业资源共享为基础，利用大数据、区块链和人工智能等技术，围绕产业焦点、热点问题，基于丰富案例和海量数据，通过开放合作的研究平台，汇集各界智慧，推动产业健康、有序、可持续发展



四大核心服务：

企业服务

为企业提供**定制化报告**服务、**管理咨询**、**战略**调整等服务

云研究院服务

提供行业分析师**外派驻场**服务，平台数据库、报告库及内部研究团队提供技术支持服务

行业排名、展会宣传

行业峰会策划、**奖项**评选、行业**白皮书**等服务

园区规划、产业规划

地方**产业规划**，**园区**企业孵化服务

头豹 LeadLeo

报告阅读渠道

头豹科技创新网 —— www.leadleo.com PC端阅读全行业、千本研报



头豹小程序 —— 微信小程序搜索“头豹”、手机扫上方二维码阅读研报

添加右侧头豹研究院分析师微信，邀您进入行研报告分享交流微信群



图说　表说　专家说　数说

详情请咨询

客服电话
400-072-5588

上海
王先生： 13611634866
李女士： 13061967127

南京
杨先生： 13120628075
唐先生： 18014813521

深圳
郭先生： 15121067239
李先生： 18916233114