



语音理解 头豹词条报告系列

陈文广
2023-04-14 未经平台授权，禁止转载

版权有问题？[点此投诉](#)

行业：

信息传输、软件和信息技术服务业/软件和信息技术服务业/其他信息技术服务业

信息科技/软件服务

关键词：

语音交互技术

智能语音

语音识别

语音交互

自然语言处理

词条目录

行业定义 语音理解是一种利用人工智能技术对语句进行自动识别和... AI访谈	行业分类 针对应用场景和任务实现的不同，语音理解技术可以分为... AI访谈	行业特征 中国语音理解行业的商业模式主要包括产品订阅服务、A... AI访谈	发展历程 语音理解行业目前已达到 4个 阶段 AI访谈
产业链分析 <div>上游分析中游分析下游分析</div> AI访谈	行业规模 2017至2022年中国语音理解技术所带动的市场规模保持... AI访谈 数据图表	政策梳理 语音理解行业相关政策 7篇 AI访谈	竞争格局 通过语音分析相关专利以及市场力两个维度对中国部分与... AI访谈 数据图表

摘要 在本篇报告中，我们将对语音理解技术进行深入的研究和探讨。报告将结合宏观市场角度，从政策、市场规模、发展趋势与竞争格局切入分析，并围绕语音合成技术的产业链条，深入分析中国语音理解行业的发展现状与趋势

语音理解行业定义

语音理解是一种利用人工智能技术对语句进行自动识别和语义理解的过程。语音理解是在语音识别的基础上，进一步对语音的语法和语义知识进行分析和利用，以达到更高层次的人机交互的目的。整个语音理解过程包括语音信号处理、语音识别、语法分析和语义理解等环节，涉及多个交叉学科，包括实验语音学、汉语语法、自然语言理解和知识搜索等，旨在为用户提供更丰富、更准确的人机交互体验^[1]

[1]

1: <https://www.leixue...>

2: <https://www.leixue...>

3: 泪雪网

语音理解分类

针对应用场景和任务实现的不同，语音理解技术可以分为通用型语音理解和特定型语音理解。通用型语音理解具有较强的泛用性，常用于搜索引擎，语言翻译等场景；特定型语音理解具有较强的专业性，适用于医疗、法律等专业性的领域^[2]



[2]

1: [http://baike.cntr...](#)

2: [https://www.woshi...](#)

3: 电子元件技术

[3]

1: [http://baike.cntr...](#)

2: [https://www.woshi...](#)

3: 电子元件技术

语音理解行业特征

中国语音理解行业的商业模式主要包括产品订阅服务、API接口以及定制化解决方案等；在行业门槛方面，目前行业进入门槛较高，主要包括技术门槛、数据门槛、人才门槛和生态门槛；在技术发展方面，语音理解技术将与其他人工智能技术融合发展，赋能下游产业智慧升级^[4]

1 商业模式

语音理解行业的商业模式主要包括产品订阅服务、API接口以及定制化解决方案等

在产品订阅服务方面，语音理解服务商根据市场需求提供多种服务套餐，套餐内可包含不同的功能，如发音人选择、语音处理速度、支持的语言种类等。根据套餐内容的不同，服务商按照使用次数或年度订阅服务的方式进行收费，订阅价格通常在数千元至十万元不等。在API接口调用方面，语音理解服务商提供API接口供客户调用。客户可以将这些接口集成到自己的产品或服务中，从而为终端用户提供语音识别、i理解等功能。API接口的计费方式通常包括按次计费、按使用量计费等。在定制化解决方案，针对特定行业

或企业需求，语音理解技术厂商可以提供定制化的解决方案。这些解决方案通常包括特定领域的语音识别、语义理解和自然语言生成等功能。定制化解决方案的收费模式包括一次性费用、持续维护费用或按项目进度收费等

2 行业门槛

目前语音理解行业的进入门槛处于较高的水平，主要包括技术门槛、数据门槛、人才门槛和生态门槛

在技术门槛方面，语音理解涉及语音识别、自然语言处理、知识图谱、机器学习等多个技术领域，需要具备较强的研发能力和创新能力以及长时间的经验积累，同时需要不断跟进技术的发展和变化，保持技术的领先优势。在数据门槛方面，语音理解需要大量的语音数据和文本数据来训练模型和优化算法，数据的质量和数量直接影响语音理解的效果和性能。同时，数据的收集、标注、存储、管理等也需要投入较多的人力和物力资源。在人才门槛方面，语音理解需要具备多学科背景和专业知识的人才，包括语音信号处理、自然语言处理、机器学习、计算机科学等领域。**目前，中国人工智能相关人才的市场缺口超500万**，对于企业而言，招聘适合的人工智能专家需要花费高昂的用工成本，同时人才的培养和留存对于中小企业而言也是一个较大的挑战。在生态门槛方面，语音理解需要与各种应用场景和行业需求相结合，形成完整的产业链和生态圈；这就需要与各方合作伙伴建立良好的关系，提供定制化的解决方案，满足不同客户的需求。然而对于中小企业而言，在生态资源方面存在明显劣势。总体而言，当前语音理解赛道的进入门槛相对较高。中小企业可与已有成熟技术的企业进行合作或购买技术授权，将这些技术应用于自己的产品和服务中，以降低自身的研发投入，缩短产品上市时间

3 多技术融合发展

多技术融合发展，语音交互赋能产业智慧升级

语音理解作为人工智能感知智能的一部分，其关键作用在于将物理世界中的信息转化为计算机可处理的数据，后续的认知智能奠定基础。因此，除了实现其核心功能外，语音理解还可以与其他AI技术深度融合，应用于更多的生活场景。通过将前端的语音交互与后端的互联网服务相结合，多种技术协同发展，不仅赋能单一技术，还能推动AI语音相关产业创新，助力新兴产业崛起。例如，在AI语音理解技术的推动下，语音机器人、智能客服等新兴产业正快速发展；在公检法领域，通过融合计算机视觉、自然语言处理和语音合成等技术，能够实现智能庭审、电信网络反欺诈、虚拟法官、声纹研判、智能接警和警务智能语音服务等功能，为公检法参与者提供全面而高效的服务

[4] 1: <https://www.xfyun...> 2: <http://www.199it.c...> 3: <https://www.sohu...> 4: 沙利文、科大讯飞、中...

[5] 1: <http://www.199it.c...> 2: <https://www.sohu...> 3: <https://www.xfyun...> 4: 中国语音产业联盟，新

语音理解发展历程

自20世纪50年代以来，语音理解技术大致经历了四个发展阶段。在模板匹配阶段，贝尔实验室开发了一个有限的识别系统，**仅能识别10个英文数字发音**。在模式和特征分析阶段，线性预测编码技术（LPC）和动态时间规整引入语音识别领域，**解决了特征提取和不等长语音匹配问题**。在概率统计建模阶段，隐马尔科夫模型（HMM）和高斯混合模型（DMM）开始被广泛应用，**神经网络技术也逐步发展**。当前，**语音理解技术已进入规模商业化落地阶段**，深度神经网络（DNN）开始占据主导地位，**提高了语音识别和理解的准确性**，近年来端到端语音合成技术逐步成为主流，促使准确率进一步提升。目前，许多消费级和专业级语音理解产品已在市场上广泛应用，例如智能音箱和语音助手等^[6]

萌芽期 · 1952~1960

模板匹配阶段

1952年，贝尔实验室研制出了世界上第一个能识别10个英文数字发音的系统，但只能理解有限的词汇以及内存中的数字，无法将语音转化为完整的句式或者词汇，同时对于不同声音模式，机器能识别的数量极为有限

启动期 · 1970~1980

模式和特征分析阶段

线性预测编码技术（LPC）被成功应用于语音识别，动态规划的思想也被应用到语音识别并提出动态时间系统，有效解决语音信号的特征提取和不等长语音匹配问题，可以通过对声音的模式和特征设置参数，并基于大量词汇进行连续的语音识别

高速发展期 · 1980~2010

概率统计建模阶段

主流算法开始转为概率统计建模，隐马尔科夫模型（HMM）和高斯混合模型（DMM）开始应用，同时神经网络在语音理解领域稳步发展，深度神经网络开始频繁出现在语音理解的主流市场，逐渐有商业化产品落地

成熟期 · 2010~2023

规模商业化落地阶段



算法由传统概率统计转向深度神经网络为主，DNN的出现开始大幅提升语音识别和理解的准确率，2016年首次达到人类水平（95%）；近年来端到端的语音合成开始成为潮流，准确率进一步提升至98%。作为人机交互的入口，消费级产品和专业级产品开始大规模落地，例如智能音箱、语音助手等产品

- [6]

1: <https://mp.weixin...>

2: <https://mp.weixin...>

3: 西南证券、CSDN、科大...
- [7]

1: <https://mp.weixin...>

2: <https://mp.weixin...>

3: 西南证券、CSDN、科大...

中国语音理解产业链分析

中国语音理解行业产业链上游主要分为数据、算力和算法三部分，在数据方面，是以海天瑞声、标贝科技、数据堂为主的数据标注服务商，为语音理解模型提供经过标注的训练数据；在算力方面，中高端的语音芯片仍然被Intel、英伟达等美国厂商所垄断，中国厂商逐渐ASIC语音芯片赛道进行发力，目前已完成技术上的突破，云知声、科大讯飞等厂商已实现的规模化量产；在算法方面，目前中国的语音理解模型的语音识别准确率已达到全球领先水平，例如科大讯飞的语音识别的准确率已高达98%。产业链中游主要为语音理解技术的服务提供商，分为互联网企业和智能语音科技企业两种类型；互联网企业在数据资源、产品生态和技术融合方面具有明显优势，代表企业包括腾讯云、阿里云、百度云等；智能语音科技企业的优势在于专注度和定制化方面，代表企业包括科大讯飞、思必驰、云知声等厂商。产业链下游主要为应用厂商，典型的应用场景包括智能车载（代表厂商包括比亚迪、特斯拉、长城汽车等）、智能音箱（代表厂商包括百度、天猫精灵、小米）等

在上游的数据标注领域，数据标注的质量和数量直接影响语音合成模型的性能和效果，包括准确性、自然度、流畅度、鲁棒性、表现力等方面。因此，投入足够的精力和资源进行高质量的数据标注是提升语音合成技术的关键。分析师认为，随着应用市场对AI模型的智能化水平要求的提高，数据标注质量将成为未来AI模型行业竞争的重要因素。在语音芯片领域，随着人工智能算法基本稳定，且规模化人工智能时代即将到来，ASIC芯片将迎来爆发式增长，具备ASIC芯片自主量产能力的企业将获得先发优势。在中游的语音理解技术服务环节，受语音理解技术的复杂性、研发投入、数据需求及市场竞争等因素制约，分析师认为，语音理解赛道的进入门槛已相对较高。中小企业可与已有成熟技术的企业进行合作或购买技术授权，将这些技术应用于自己的产品和服务中，以降低自身的研发投入，缩短产品上市时间。在下游应用场景方面，智能车载为语音理解技术的典型应用场景，2021年中国新发布乘用车当中，语音交互功能的渗透率达到了84%；预计未来5年将提升至90%以上。智能音箱结合了语音识别、语义理解和语音合成技术，可作为智能家居的“中枢神经”，人类可向智能音箱发出语音命令，从而控制其他智能家居设备。2021年中国智能音箱的出货量达到2,631万台，其中百度、小米、天猫精灵、华为位居前四，合计份额达到96.5%。面对当前行业集中度水平极高的形势，中小创企业的智能家居产品将以接入智能音箱头部厂商的方式为主^[8]

生产制造端

数据标注服务商

上游厂商

北京海天瑞声科技股份有限公司 >

数据堂（北京）科技股份有限公司 >

标贝（北京）科技有限公司 >

查看全部

产业链上游说明

在语音理解算法模型的训练数据方面，训练数据集的生成包括设计、采集、标注和质检等生产环节，或者基于客户提供的原料音频文件执行标注、质检工作，最终形成客户所需的语音合成训练数据集。数据标注的质量和数量直接影响语音合成模型的性能和效果，包括准确性、自然度、流畅度、鲁棒性、表现力等方面。因此，投入足够的精力和资源进行高质量的数据标注是提升语音合成技术的关键。现阶段，一般数据标记的准确率维持在60%-70%区间，高质量的数据标注准确率约为90%。随着应用市场对AI模型的智能化水平要求的提高，数据标注质量将成为未来AI模型行业竞争的重要因素

生产制造端

AI语音芯片提供商

上游厂商

科大讯飞股份有限公司 >

百度（中国）有限公司 >

北京宇音天下科技有限公司 >

查看全部

产业链上游说明

目前中国在AI芯片领域对外依存度仍然较高，中高端的芯片仍然依赖进口，例如GPU近乎被美国企业英伟达、英特尔所垄断，合计占据全球市场约90%的市场份额。在国产替代方面，经过中国多年在ASIC语音芯片领域的加速研发，部分本土企业已实现量产，例如云知声的AI芯片及模组出货量已达千万级水平，积累近800家合作客户。由于云端训练算法十分复杂，短期内GPU等灵活性较高的芯片依然将占据市场主流地位。从长期来看，中国AI芯片产业的发展潜力巨大。随着人工智能算法基本稳定，且规模化人工智能时代即将到来，ASIC芯片将迎来爆发式增长，具备ASIC芯片自主量产能力的企业将获得先发优势

生产制造端

开源模型提供商

上游厂商

微软（中国）有限公司 >

谷歌信息技术（中国）有限公司 >

阿里巴巴达摩院（杭州）科技有限公司 >

查看全部 ▾

产业链上游说明

从算法框架来看，目前主流语音识别模型已经以深度神经网络为主导，神经网络的出现及普及为语音识别准确率的提升起到了重要作用。目前中国语音理解模型的研发已处于全球领先水平，其中语音识别准确率已接近100%，以中国头部智能语音科技厂商——科大讯飞为例，科大讯飞依靠专业的语音识别技术，在近场安静的情况下，语音识别的准确率已经达到了98%；在多人会议场景下语音识别的准确率提升至95%

中 产业链中游

品牌端

互联网企业

中游厂商

腾讯云计算（北京）有限责任公司 >

阿里云计算有限公司 >

百度云计算技术（北京）有限公司 >

查看全部 ▾

产业链中游说明

在数据资源方面，互联网企业通过其庞大的用户群体和业务范围，积累了大量的语音数据和多样化的应用场景。这为语音识别和理解模型的训练提供了丰富的数据资源，有助于提高语音理解技术的准确性和适用性。在产品生态方面，互联网企业具备良好的平台生态，能够将语音理解技术产品和服务与自家的应用场景相结合，形成一个完整的技术和服务体系。在技术融合方面，互联网企业具有多样化且跨领域的业务布局，能够将语音理解技术与其他AI技术（如自然语言处理、计算机视觉等）相结合，赋能更多“智慧化”的产品和服务

品牌端

智能语音科技企业



中游厂商

科大讯飞股份有限公司 >

思必驰科技股份有限公司 >

云知声智能科技股份有限公司 >

查看全部

产业链中游说明

相比互联网厂商，智能语音科技企业在专注度和定制化方面存在明显优势。在专注度方面，智能语音科技厂商更为专注于语音理解相关技术与垂直赛道的研发与布局，对于市场需求及技术趋势更为了解，有助于智能语音科技厂商及时地把握市场机遇，抢占先发优势。在定制化方面，智能语音科技厂商能够根据客户的具体需求和应用场景，提供高度定制化的语音理解解决方案，满足不同行业和领域用户的个性化需求，提升客户满意度

产业链下游

渠道端及终端客户

智能汽车厂商

渠道端

比亚迪股份有限公司 >

广州小鹏汽车科技有限公司 >

特斯拉（上海）有限公司 >

查看全部

产业链下游说明

在智能车载领域，智能语音的应用涉及到语音识别、语义理解、自然语言生成和语音合成等环节，驾驶员通过语音助手实现与汽车的高效交互，完成多媒体娱乐、车辆控制、智能导航、车况监控等任务的指令。在车载场景中，语音交互系统是替代按键或触控的最佳方案，可最大程度地保障驾驶员在驾驶过程中的功能使用的安全与效率，因此语音交互系统逐渐成为智能汽车的标配产品。在渗透率方面，2021年中国新发布乘用车当中，语音交互功能的渗透率达到了84%，显然，语音交互功能基本上已成为智能化座舱的一项标配功能。未来在日趋智能化的汽车座舱领域，语音交互将成为不可或缺的部分，智能语音在智能汽车领域的渗透率将持续上升，预计未来5年超过90%

渠道端及终端客户

智能音箱

渠道端

小米科技有限责任公司 >

浙江天猫技术有限公司 >

北京百度网讯科技有限公司 >

查看全部

产业链下游说明

智能音箱结合了语音识别、语义理解和语音合成技术，一方面，智能音箱能够产生自然流畅的语音反馈，与用户进行语音交互，例如播报天气、新闻资讯等；另一方面，智能音箱可作为智能家居的“中枢神经”，人类可向智能音箱发出语音命令，从而控制其他智能家居设备，例如智能空调、智能电视、扫地机器人等在应用价值方面，语音合成高效提升用户体验。**2021年中国智能音箱的出货量达到2,631万台，其中百度、小米、天猫精灵、华为位居前四，合计份额达到96.5%。从发展趋势上来看，由于智能语音技术行业进入壁垒较高，且智能音箱对产品生态的要求较高，预计未来市场份额将向头部企业进一步聚拢在发展趋势方面，随着语音理解技术的不断进步，智能家居设备的语音交互将变成更加富有逻辑、自然和流畅，用户体验将进一步提升。同时，语音理解技术将与图像识别、触摸等其他模态相融合，实现多模态交互**

- [8]

1: <https://36kr.com/p...>

2: <https://36kr.com/p...>

3: <https://36kr.com/p...>

4: <https://www.leiph...>

5: <https://m.jrj.com.c...>

6: <https://smarthom...>

7: 云知声、九鼎投资、科...
- [9]

1: <https://www.dong...>

2: <https://m.jrj.com.c...>

3: <https://smarthom...>

4: <https://36kr.com/p...>

5: <https://36kr.com/p...>

6: <https://www.leiph...>

7: <https://www.leixue...>

8: 科大讯飞、云知声

中国语音理解行业市场规模

2017至2022年中国语音理解技术所带动的市场规模保持高速增长，由1,256.64亿元逐步增长至6,957.72亿元，CAGR为40.8%；预计未来5年将增长至26,189.24亿元。

2017至2022年，全球智能手机出货量维持在10亿台以上的规模（例如2017的出货量为14.72亿台，2022年为12.1亿台），以及智能音箱的出货量由2017年的3,200万台快速上升至2022年的1.2亿台，叠加疫情催化线上智能语音的市场应用与需求，推动中国语音理解所带动的相关市场规模保持高速增长，于2022年达到6,957.72亿元

分析师认为，在政策端，2022年末中国中央经济工作会议明确提出要大力发展数字经济，人工智能作为数字经济发展的**重要引擎，未来将得到政策的重点扶持，预计人工智能技术在数字经济产业中的渗透率将由2022年的9%逐步上升至2027年的22%。**在技术端，随着数据、算力和算法等基础设施的不断完善与发展，**语音识别模型的智能化水平随之提高，语音识别准确率及其产品与服务性能不断上升，从而推动市场规模增长。在应用**

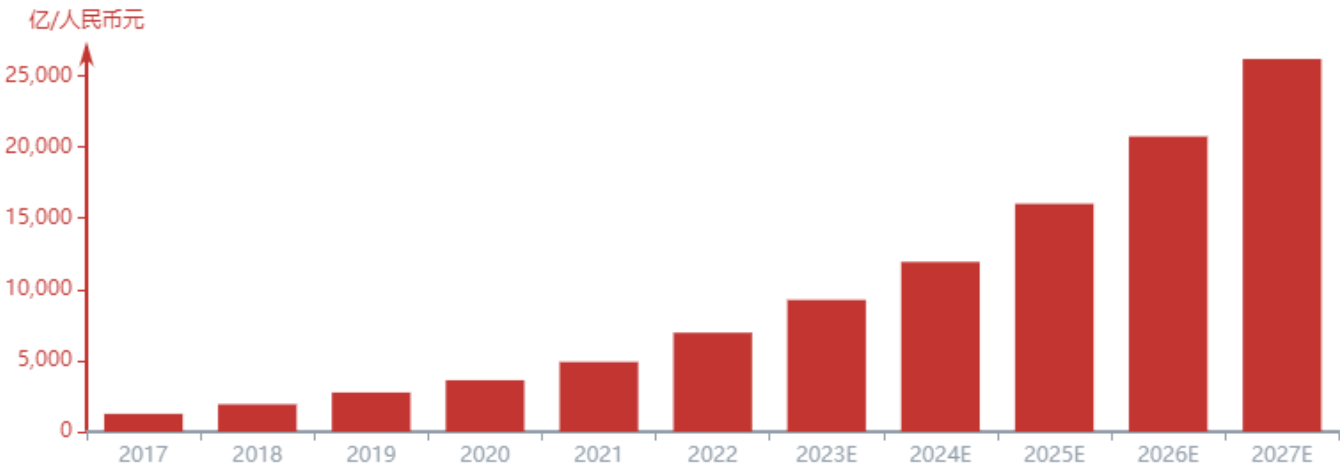
端，随着下游应用市场对语音理解技术的需求增加，例如智能车载的语音交互的渗透率将由2021年84%提升至2027年的90%以上，语音理解技术将向下游市场横向与纵向加速渗透，带动市场规模快速增长。综上所述，未来预计在政策扶持、技术进步、产业融合等驱动因素的共同作用下，预计中国语音理解所带动的市场规模将于2027年达到26,189.24亿元，CAGR为30.4%^[10]

中国语音理解所带动的相关市场规模

中国互联网协会、华为、中国通信院

下载原始数据

中国语音理解所带动的相关市场规模



中国语音理解所带动的相关市场规模=中国智能语音所带动的相关市场规模*语音理解在智能语音中的价值量占比

[10] 1: [http://www.199it.c...](http://www.199it.com) 2: [https://finance.sin...](https://finance.sina.com.cn) 3: [https://www.esmc...](https://www.esmc.com.cn) 4: [https://new.qq.co...](https://new.qq.com)
5: [http://finance.peo...](http://finance.people.com.cn) 6: [https://www.isc.or...](https://www.isc.org.cn) 7: [https://www.afenjie...](https://www.afenjie.com) 8: [https://baijiahao.b...](https://baijiahao.baidu.com)
9: [https://t.cj.sina.co...](https://t.cj.sina.com.cn) 10: 中国通信院、华为、中...

中国语音理解行业相关政策梳理

[11]	政策名称	颁布主体	生效日期	影响
	《关于推动12345政务服务便民热线与110报警服务台高效对接联动的意见》	国务院办公厅	2020-05-15	8
政策内容	强化科技赋能，进一步加强12345平台和网上12345能力建设，开发智能推荐、语音自动转写、智能服、自助派单功能			

政策解读	《意见》提倡利用语音理解技术将语音信息自动转换成文字信息，降低人工转写成本，提高受理效率，同时减少因人工操作产生的错误。此外，政策鼓励采用智能客服系统，通过语音理解技术识别用户的需求，自动为用户提供服务或引导用户进行自助服务，减轻客服人员的工作负担，提高服务效率和满意度			
政策性质	鼓励性政策			

[11]	政策名称	颁布主体	生效日期	影响
	《关于进一步优化地方政务服务便民热线的指导意见》	国务院	2021-01-06	8
政策内容	加强自助下单、智能文本客服、智能语音等智能化应用，方便企业和群众反映诉求建议			
政策解读	《指导意见》鼓励在政务服务便民热线中引入智能语音技术，利用语音理解技术开发自助下单功能，使得企业和群众可以通过语音指令进行服务请求，由此提高服务质量和效率，更好地满足企业和群众的需求			
政策性质	指导性政策			

[11]	政策名称	颁布主体	生效日期	影响
	《天津市人民政府办公厅关于印发天津市新型基础设施建设三年行动方案（2021—2023年）的通知》	天津市人民政府	2021-03-03	6
政策内容	强化“津治通”平台支撑能力，引入智能视频识别和智能语音服务，提升数据支撑决策能力和平台指挥调度能力			
政策解读	该政策提倡在“津治通”平台中引入智能语音服务，通过运用先进的语音识别和语义理解技术，使“津治通”平台在处理公共事务、应急响应等方面具备更强大的指挥调度能力，并将语音数据转化为可供分析的文本数据，从而提高政府部门的数据支撑决策能力，由此提升政府部门的政务服务效率和质量			
政策性质	规范类政策			

[11]	政策名称	颁布主体	生效日期	影响
	《省政府办公厅关于印发江苏省“十四五”现代服务业发展规划的通知》	江苏省人民政府	2021-07-09	

政策内容	加强计算机视觉、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
政策解读	《通知》旨在加强现代服务业发展，尤其重视人工智能产业的发展。该规划提出要推进人工智能开发框架、算法库、工具集等研发突破，这意味着政府将支持和鼓励相关研究机构和企业加大在语音理解领域的研发力度，以实现技术突破。同时规划明确提出要加强智能语音处理、自然语言理解等关键技术的研发和产业化，这表明政府对语音理解技术的发展给予高度关注，将促进该领域技术水平的提升
政策性质	指导性政策

[11]	政策名称	颁布主体	生效日期	影响
	《黑龙江省人民政府关于印发黑龙江省“十四五”数字政府建设规划的通知》	黑龙江省人民政府	2022-01-02	7
政策内容	培育人工智能语音企业，提供智能语音产品，加强智能语音技术在政务服务、呼叫中心、跨境贸易、智慧养老等场景的应用			
政策解读	政策鼓励企业研发和生产具有市场潜力和创新价值的智能语音产品，以满足政府、企业以及市场在语音技术方面的需求；此外，政策强调将智能语音技术应用于政务服务领域，例如在线咨询、办事指南等，以提高政务服务的效率和智能化水平			
政策性质	鼓励性政策			

[11]	政策名称	颁布主体	生效日期	影响
	《上海打造未来产业创新高地发展壮大未来产业集群行动方案》	上海市人民政府	2022-11-28	7
政策内容	加快超大模型向机器视觉、智能语音语义、自然语言处理、人机交互等领域应用，推动AI普惠化			
政策解读	该《方案》鼓励将超大规模的人工智能模型应用于语音理解领域，以提高语音识别和理解的准确性和效率，并推动AI技术在更多领域和行业的广泛应用，实现技术融合和产业创新，让更多人受益于AI技术带来的便利和效率提升			
政策性质	指导性政策			



[11]	政策名称	颁布主体	生效日期	影响
	《鼓励外商投资产业目录(2020年版)》	商务部	2022-12-27	6
政策内容	鼓励智能语音、量子通信等设备研发、制造			
政策解读	政策明确鼓励外国投资者投资于智能语音领域，包括研发、制造等各个环节，促进中国在语音识别、语音合成、语音理解等关键技术方面的创新，提升中国在全球智能语音产业的竞争力，同时有助于引进国际资金和先进技术，加速中国智能语音产业的发展			
政策性质	鼓励性政策			

[11]

1: <https://public.zzs...>

2: <https://news.cctv.c...>

3: <http://www.beijing...>

4: <https://www.hlj.go...>

5: <https://fzggw.ah.g...>

6: <http://www.jiangs...>

7: 国务院办公厅、商务部...

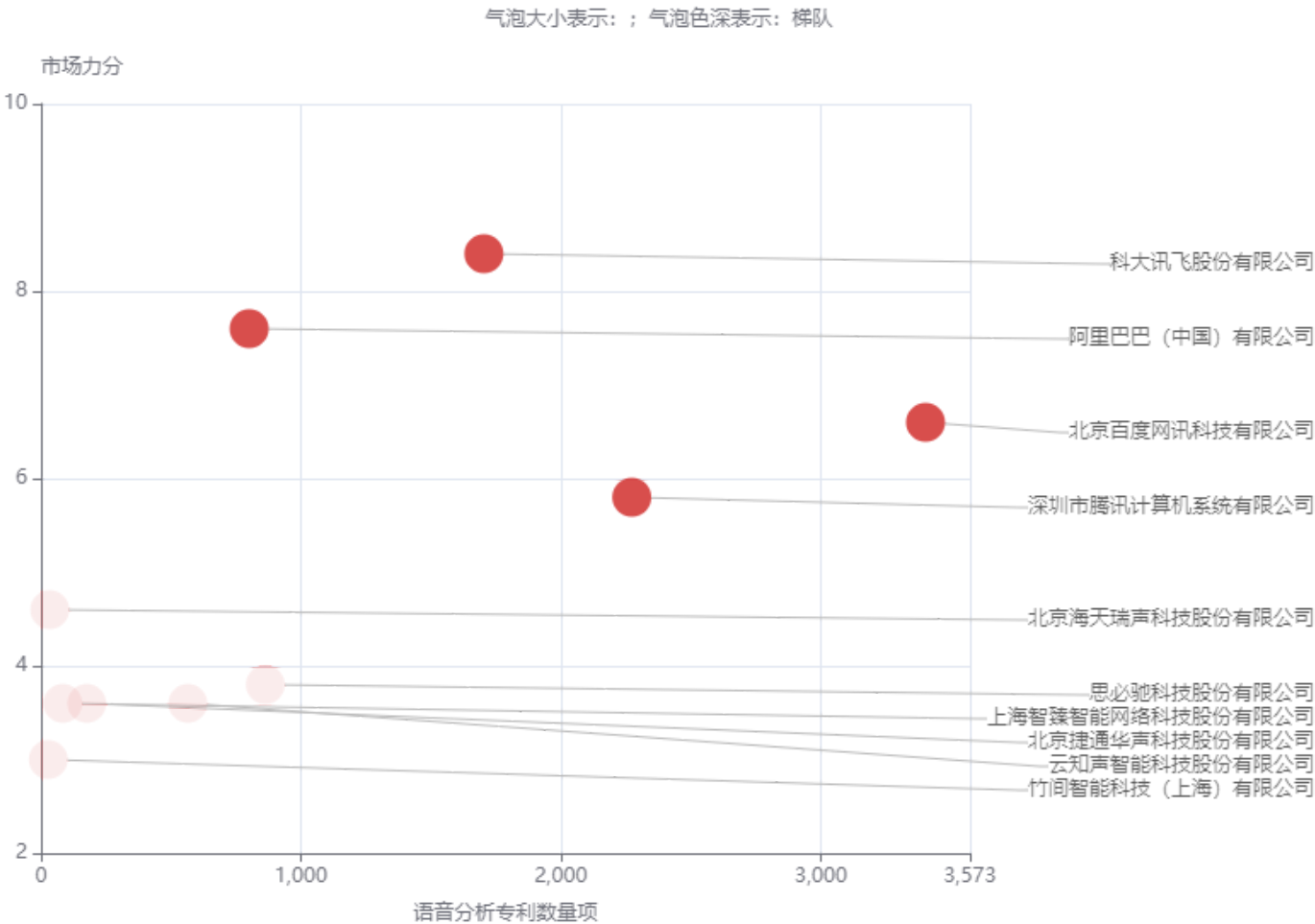
中国语音理解行业竞争格局

通过语音分析相关专利以及市场力两个维度对中国部分与智能语音相关的企业进行比较分析，位于领导者象限的企业主要为互联网头部厂商（BAT）与智能语音厂商科大讯飞，具有专利数量多、研发能力强、市场占有率高和产品生态能力强等特征；位于挑战者象限的企业在技术能力以及市场力方面具备一定的竞争力，在某些细分领域具备行业领先的实力，但与领导者象限内的企业仍存在一定的差距，代表厂商包括思必驰、云知声、竹间智能等人工智能科技厂商

互联网大厂凭借在数据规模、研发能力、资金规模和应用场景方面的优势，占据中国语音理解市场的领先地位。在数据方面，头部互联网厂商拥有海量的语音数据和用户场景，并且可利用大数据和云计算提升语音理解的效果和性能。例如腾讯云凭借庞大的社交数据平台，积累了数十万小时的语音标注数据，这些数据可用于语音理解模型的训练。同时，头部互联网厂商拥有强大的技术研发能力和资金优势（BAT三大厂商的研发投入达到数百亿/年的规模，是中小企业的数十倍），能够掌握和引领语音识别与合成的前沿技术；此外，互联网厂商拥有丰富的应用场景和商业模式，能够将语音识别与合成技术应用于社交娱乐、影视制作、教育、广告营销等领域。位于挑战者象限的企业主要为人工智能技术与服务提供商，在市场力及技术专利方面具备一定的竞争力，在某些细分领域具备行业领先的实力，但与领导者象限内的企业存在一定的差距，例如位于挑战者象限的思必驰在语音分析领域的专利数量为862项，高于位于领导者象限的阿里云的800项，但在市场占有率、产品生态以及客户资源方面仍存在明显差距，思必驰的市场力评分为3.8，低于阿里云的7.6

随着人工智能技术的不断发展，应用市场对于语音理解的要求不断抬升，行业进入门槛亦随着提升。师认为，位于领导者象限的头部厂商的市占率缓慢增长，中国语音理解市场的行业集中度将进一步上升，CR4将达

到75%以上；位于挑战者象限的企业竞争将持续加剧^[12]



[14]

上市公司速览

[12]

1: [http://plus.soopat...](#)

2: [https://data.eastm...](#)

3: [https://cloud.tenc...](#)

4: 德勤、Soopat、沙利文...

[13]

1: [http://plus.soopat...](#)

2: Soopat

[14]

1: [https://data.eastm...](#)

2: [https://www.speech...](#)

3: [https://www.sinov...](#)

4: [https://ai.unisoun...](#)

5: [https://www.aispe...](#)

6: [https://cloud.tenc...](#)

7: [https://www.emoti...](#)

8: [https://ai.aliyun.co...](#)

9: [https://www.xfyun...](#)

10: [https://www.xiaoi...](#)

11: 德勤、东方财富、科大...

语音理解代表企业分析

1 腾讯云计算（北京）有限责任公司

公司信息

企业状态	存续	注册资本	104250万人民币
------	----	------	------------

企业总部	市辖区	行业	互联网和相关服务
法人	谢兰芳	统一社会信用代码	911101085636549482
企业类型	有限责任公司(法人独资)	成立时间	2010-10-21
品牌名称	腾讯云计算（北京）有限责任公司		
经营范围	利用互联网经营游戏产品运营、网络游戏虚拟货币发行、从事互联网文化产品的展览、比赛... 查看更多		

▪ 竞争优势

1.语音资源丰富，借助庞大的用户资源，积累了千亿级互联网语料数据。2.技术积累扎实，腾讯在智能语音领域超过拥有超2,000项专利。3.服务与落地经验领先。为数十个腾讯亿级用户产品提供智能语音的服务支撑。

2 百度云计算技术（北京）有限公司

▪ 公司信息

企业状态	存续	注册资本	10000万美元
企业总部	市辖区	行业	软件和信息技术服务业
法人	崔珊珊	统一社会信用代码	91110302589136401T
企业类型	有限责任公司(台港澳法人独资)	成立时间	2012-01-05
品牌名称	百度云计算技术（北京）有限公司		
经营范围	计算机软件开发、生产；提供技术服务、技术咨询；计算机网络系统集成；销售自产产品； ... 查看更多		

▪ 竞争优势

1.“自增强循环”模式。拥有强大用户规模与产品生态，庞大的数据量反哺数据算法不断完善和进步。2.自然语言理解定制开发成本低。文心大模型降低了90%以上的数据标注量，90%的算力投入和85%以上的开发时长。


3 科大讯飞股份有限公司【002230】

▪ 公司信息

企业状态	存续	注册资本	232375.2783万人民币
企业总部	合肥市	行业	研究和试验发展
法人	刘庆峰	统一社会信用代码	91340000711771143J
企业类型	其他股份有限公司(上市)	成立时间	1999-12-30
股票类型	A股	品牌名称	科大讯飞股份有限公司

经营范围

增值电信业务；专业技术人员培训；计算机软、硬件开发、生产和销售及技术服务；系统工... 查看更多

▪ 财务数据分析									
财务指标	2013	2014	2015	2016	2017	2018	2019	2020	2021
销售现金流/营业收入	0.88	0.85	0.96	0.99	0.92	0.96	0.92	0.97	-
资产负债率(%)	20.0224	24.5358	22.2538	30.6817	40.3875	46.3375	41.6202	47.7703	44.775
营业总收入同比增长(%)	59.9238	41.5968	40.8734	32.7766	63.9731	45.4119	27.3008	29.2297	40.607
归属净利润同比增长(%)	52.9489	36.0031	12.0878	13.9047	-10.2707	24.7058	51.1211	66.4831	-
应收账款周转天数(天)	155.2862	189.3242	188.7089	174.9526	143.8159	135.0895	151.3877	145.873	127
流动比率	3.7561	2.3836	2.9774	2.1947	1.6112	1.3354	1.6648	1.4413	1.63
每股经营现金流(元)	0.632	0.5231	0.4007	0.2275	0.2612	0.5487	0.6966	1.0207	0.384
毛利率(%)	53.0057	55.6343	48.8951	50.5196	51.3794	50.0261	46.0202	45.1162	-
流动负债/总负债(%)	82.4078	84.8198	85.7468	78.8999	83.4317	81.9754	82.0683	87.5915	82.941
速动比率	3.5613	2.1896	2.6381	1.8456	1.2479	1.0898	1.3901	1.1547	1.421
摊薄总资产收益率(%)	9.0096	8.1809	6.4392	5.2837	4.0345	4.315	5.3276	6.4169	5.536
营业总收入滚动环比增长(%)	46.6687	46.3676	30.1759	73.622	60.2335	27.0359	49.4961	95.599	-
扣非净利润滚动环比增长(%)	201.2757	191.2552	207.3742	8790.7143	455.0873	5358.7679	997.3604	793.0686	-
加权净资产收益率(%)	10.63	10.79	9.17	7.5	5.74	6.94	8.22	10.97	-
基本每股收益(元)	0.64	0.47	0.34	0.37	0.33	0.27	0.4	0.64	0.7
净利率(%)	22.2092	21.884	17.4578	14.9611	8.8009	7.8054	9.3571	11.0696	8.7952
总资产周转率(次)	0.4057	0.3738	0.3688	0.3532	0.4584	0.5528	0.5694	0.5797	0.651
归属净利润滚动环比增长(%)	140.1001	123.0864	150.3192	820.3134	333.7751	264.9265	141.9939	173.4243	

每股公积金(元)	4.6141	2.4136	2.8846	3.1329	3.7834	2.1924	3.1699	3.2988	4.4466
存货周转天数(天)	64.082	69.0979	69.4806	100.4324	101.6145	87.8027	61.7369	80.7121	80
营业总收入(元)	12.54亿	17.75亿	25.01亿	33.20亿	54.45亿	79.17亿	100.79亿	130.25亿	183.14亿
每股未分配利润(元)	1.3161	1.1009	0.91	1.1507	1.3041	1.0581	1.3038	1.6917	2.0884
稀释每股收益(元)	0.64	0.47	0.34	0.37	0.32	0.26	0.38	0.61	0.68
归属净利润(元)	2.79亿	3.79亿	4.25亿	4.84亿	4.35亿	5.42亿	8.19亿	13.64亿	15.56亿
扣非每股收益(元)	0.5	0.36	0.26	0.2	0.27	0.13	0.24	0.36	0.44
经营现金流/营业收入	0.632	0.5231	0.4007	0.2275	0.2612	0.5487	0.6966	1.0207	0.384

▪ 竞争优势

1. 深耕语音领域的顶尖技术。科大讯飞在语音领域深耕多年，拥有ASR、TTS等核心语音技术，并位于行业顶尖水平。2. 拥有海量语音数据资源，有助于提高语音识别和合成的准确性。

法律声明

权利归属：头豹上关于页面内容的补充说明、描述，以及其中包含的头豹标识、版面设计、排版方式、文本、图片、图形等，相关知识产权归头豹所有，均受著作权法、商标法及其它法律保护。

尊重原创：头豹上发布的内容（包括但不限于页面中呈现的数据、文字、图表、图像等），著作权均归发布者所有。头豹有权但无义务对用户发布的内容进行审核，有权根据相关证据结合法律法规对侵权信息进行处理。头豹不对发布者发布内容的知识产权权属进行保证，并且尊重权利人的知识产权及其他合法权益。如果权利人认为头豹平台上发布者发布的内容侵犯自身的知识产权及其他合法权益，可依法向头豹（联系邮箱：support@leadleo.com）发出书面说明，并提供具有证明效力的证据材料。头豹在书面审核相关材料后，有权根据《中华人民共和国侵权责任法》等法律法规删除相关内容，并依法保留相关数据。

内容使用：未经发布方及头豹事先书面许可，任何人不得以任何方式直接或间接地复制、再造、传播、出版、引用、改编、汇编上述内容，或用于任何商业目的。任何第三方如需转载、引用或基于任何商业目的使用本页面上的任何内容（包括但不限于数据、文字、图表、图像等），可根据页面相关的指引进行授权操作；或联系头豹取得相应授权，联系邮箱：support@leadleo.com。

合作维权：头豹已获得发布方的授权，如果任何第三方侵犯了发布方相关的权利，发布方或将授权头豹或其指定的代理人代表头豹自身或发布方对该第三方提出警告、投诉、发起诉讼、进行上诉，或谈判和解，或在认为必要的情况下参与共同维权。



完整性：以上声明和本页内容以及本平台所有内容（包括但不限于文字、图片、图表、视频、数据）构成不可分割的部分，在未详细阅读并认可本声明所有条款的前提下，请勿对本页面以及头豹所有内容做任何形式的浏览、点击、引用或下载。

