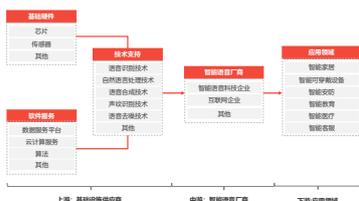


2019 年 中国智能语音行业研究报告

行业走势图



人工智能团队

王则焜 分析师
邮箱: agnes.Wang@frostchina.com

相关热点报告

- 人工智能系列深度研究——2019 年中国人工智能行业研究报告
- 人工智能系列深度研究——2019 年中国智能家居行业研究报告
- 人工智能系列深度研究——2020 年中国医疗机器人行业研究报告
- 人工智能系列深度研究——2020 年中国工业机器人行业研究报告

报告摘要

智能语音技术是人工智能领域的重要分支。智能语音技术涉及多类型学科,其核心技术包括语音合成、语音识别、声纹识别、自然语言理解、语音去噪等关键技术。伴随智能语音技术的发展,智能语音的应用覆盖多个场景,如智能家居、智能车载、智能医疗、智能客服、智能教育等。互联网企业、智能语音技术企业以及智能语音初创企业纷纷布局中国智能语音市场,推动智能语音行业市场规模持续扩容,2023 年中国智能语音行业市场规模将达到 655.1 亿元。

热点一：智能语音核心技术的发展

智能语音行业的核心技术包括语音合成技术、语音识别技术、声纹识别技术、自然语言理解技术以及语音去噪技术。除自然语言理解技术和语音去噪技术外,其余四项核心技术在智能语音领域的商业化应用较为广泛,核心技术的发展推动各应用领域智能语音产品的落地。

热点二：VR 游戏将改变玩家的社交方式

在智能语音行业中,深度学习、大数据、云计算等辅助技术的发展提高了语音识别的准确率、降低了数据存储所需的硬件成本,为智能语音行业的发展提供契机,助力智能语音行业发展。

热点三：语音交互的便利性优势突出

语音交互技术的发展帮助用户实现了本能表达,解放消费者感官的占用,优化用户的消费体验。语言交互是利用语言信息进行交互的技术,在空间上打破了用户与智能设备的距离限制。在智能语音领域,远场识别技术的发展进一步缩小了用户与智能设备的距离限制,当用户与智能设备保持较远距离时,用户仍然可以通过语音指令与智能语音设备进行交互。语音交互的便利性将惠及各个用户群体的生活,如老年人群体、盲人群体。语音交互技术为更广泛的消费者群体带来“伴随式”体验。语音交互的便利性将推动下游应用场景对于智能语音设备的需求,助力智能语音行业发展。

目录

1	方法论.....	5
1.1	方法论.....	5
1.2	名词解释.....	5
2	中国智能语音行业市场综述.....	8
2.1	中国智能语音行业定义.....	8
2.2	中国智能语音行业主要应用技术分类.....	8
2.2.1	语音合成技术.....	8
2.2.2	语音识别技术.....	8
2.2.3	声纹识别技术.....	9
2.2.4	自然语言理解技术.....	9
2.2.5	语音去噪技术.....	9
2.3	中国智能语音行业市场规模.....	10
2.4	中国智能语音行业产业链分析.....	11
2.4.1	中国智能语音行业产业链上游.....	12
2.4.2	中国智能语音行业产业链中游.....	13
2.4.3	中国智能语音行业产业链下游.....	15
3	中国智能语音行业驱动因素.....	17
3.1	智能语音核心技术的发展.....	17

3.1.1	语音合成技术.....	17
3.1.2	语音识别技术.....	17
3.1.3	声纹识别技术.....	17
3.2	辅助技术的发展为行业发展创造机遇.....	18
3.2.1	深度学习.....	18
3.2.2	云计算.....	19
3.2.3	大数据.....	19
3.3	语音交互的便利性优势.....	19
4	中国智能语音行业制约因素.....	21
4.1	行业基础层技术存在短板.....	21
4.2	技术瓶颈限制行业发展.....	21
4.2.1	语音识别技术“鲁棒性”问题凸显.....	21
4.2.2	自然语言处理技术尚不成熟.....	22
5	中国智能语音行业政策.....	23
5.1	国家性政策红利助力行业发展.....	23
5.2	地方性政策红利助力行业发展.....	23
6	中国智能语音行业发展趋势.....	25
6.1	行业将大力发展深度集成语音 AI 芯片.....	25
6.2	行业将构建智能语音专利池.....	25

7	中国智能语音行业竞争格局	27
7.1	智能语音科技企业	27
7.2	互联网巨头	27
7.3	初创企业	28
7.4	典型代表企业分析——蓦然认知	28
7.4.1	企业简介	28
7.4.2	产品分类	28
7.4.3	竞争优势	29
7.5	典型代表企业分析——思必驰	29
7.5.1	企业简介	29
7.5.2	产品分类	30
7.5.3	竞争优势	30
7.6	典型代表企业分析——云知声	31
7.6.1	企业简介	31
7.6.2	产品分类	31
7.6.3	竞争优势	32

图表目录

图 2-1 中国智能语音行业市场规模, 2014 年-2023 年预测	11
图 2-2 中国智能语音行业产业链.....	12
图 4-1 中国智能语音行业国家性政策	23

1 方法论

1.1 方法论

沙利文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、语音识别、大数据等领域着手，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。
- ✓ 弗若斯特沙利文本次研究于 2019 年 05 月完成。

1.2 名词解释

- VUI——Voice User Interface，语音交互界面。
- GUI——Graphical User Interface，图形用户界面，采用图形方式显示的计算机操作用户界面。

-
- 声门——两瓣声带之间的开口，肺部压出的空气通过声带的出口，不同的声门大小导致不同的语音音色。
 - 算法——基于特定的计算模型，旨在解决某一信息处理问题而设计的一个指令序列。
 - 算力——计算机计算哈希函数输出的速度。
 - 定制化芯片——通过定制化半导体解决方案，面向特定领域提供的定制化芯片。
 - 类脑类芯片——仿照人类大脑的信息处理方式所制造的芯片，该类芯片可实现数据并行传送，分布式处理，并能够以低功耗实时处理海量数据。
 - HMM 模型——Hidden Markov Model，隐性马尔可夫统计模型，用来描述一个含有隐含未知参数的马尔可夫过程。
 - API 接口——Application Programming Interface，应用程序编程接口。
 - 鲁棒性——Robust，即系统的健壮性，是指系统在不同应用条件下的性能稳定性。
 - 语音芯片——具备语音交互功能的专用集成电路。
 - 嵌入式语音软件——满足资源和运算受限环境下的各种嵌入式应用需求的语音软件（平台），如手机、GPS 导航等。
 - 语音合成技术——将计算机自身产生的、或外部输入的文字信息转变为可以听得懂的、流利的语言输出的技术。
 - 语音识别技术——将人类语音中的词汇内容转换成计算机可以处理的输入语料，最终实现词句词义识别的技术。
 - 声纹识别技术——通过语音信号提取说话人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，进而识别出说话人身份的技术。
 - 自然语言理解技术——利用处理语言的技术，使计算机理解人类语言的含义，并通过对话的方式回答用户提出的问题。

-
- 语音去噪技术——控制语音通信过程中的语音质量，提高系统对于语音理解的准确性，缓解噪音污染对于语音收录影响的技术。
 - 蓦然认知——一家以认知计算、自然语言理解技术为核心的人工智能公司。
 - 思必驰——一家专注于人性化智能语音交互技术的智能语音服务商。
 - 云知声——一家专注于物联网人工智能服务的智能语音识别服务商。
 - 寒武纪——中国智能芯片生产服务商，企业专注于打造各类智能云服务器、智能终端以及智能机器人的核心处理器芯片。
 - 地平线——中国人工智能芯片和算法软件服务商。

2 中国智能语音行业市场综述

2.1 中国智能语音行业定义

智能语音技术是人工智能领域的重要分支。智能语音技术涉及多类型学科，其核心技术包括语音合成、语音识别、声纹识别、自然语言理解、语音去噪等关键技术。智能语音行业是以语音为研究对象，对语音语义进行识别、理解以及生成，使机器具备自然语言处理能力，并且利用其核心技术赋予机器“听觉”、“理解能力”以及“语言能力”。伴随智能语音技术的发展，智能语音的应用覆盖多个场景，如智能家居、智能车载、智能医疗、智能客服、智能教育等。

2.2 中国智能语音行业主要应用技术分类

2.2.1 语音合成技术

语音合成技术是通过计算机将外部输入的文字信息转变成自然流畅的语言，赋予机器“讲话”的能力。语音合成技术覆盖声学、语言学、数字信号处理、计算机科学等多个学科。语音合成技术的实现过程主要分为文本分析和语音合成两个步骤。文本分析是基于语言学原理，将文本标准化，将原始文本中的数字、缩略语等转换为对应的标准词，然后进行语言处理。在文本分析的过程中，系统将为每一个字词赋予单独的语音脚本，并依据规则对文本进行分割标记，将文字序列转换成音韵序列。语音合成技术通过不同的算法，将音韵序列生成语音波形，合成高质量的语音流输出。

2.2.2 语音识别技术

语音识别技术是将人类语音中的词汇内容转换成计算机可以处理的输入内容的技术。语音识别技术通过将用户输入的指令进行特征提取，形成特征数据流，然后与系统中已有的语音模型进行比对，寻找系统中最为接近的语音内容。实现语音识别的过程主要分为四步：(1)

选择识别单元，即确定选择识别的对象，然后根据识别对象的语音特点、词汇量大小等条件确定识别对象为单词、音节或音素；(2) 提取特征参数，从语音波形中提取出重要的反应语音特征的相关信息；(3) 建立声学模型和语言模型，进行训练和识别；(4) 是后期的处理，包括音字转换、词法、句法和文法的处理等。

2.2.3 声纹识别技术

声纹识别技术是基于声纹信息识别人类身份的生物特征识别技术。声纹识别技术通过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，进而识别出发声者的身份。声纹识别技术的作用主要包括两方面：(1) 发声者辨认，主要用于在从某一语音材料的若干发声者中寻找指定发声者；(2) 发声者确认，主要用于确认某一语料是否由指定发声者发出。声纹识别技术实现原理和语音识别技术原理类似，但声纹技术识别主要是对其发声者身份的进行判断，因此实现过程相较于语音识别更简单。未来，声纹识别技术的主要发展方向为降低发声者身体状况、说话的方式、录音信道及环境噪音对声纹信息的干扰，提高声纹信息技术在干扰因素下的识别准确度。

2.2.4 自然语言理解技术

自然语言理解技术是通过利用处理语言技术，使计算机理解人类语言的含义，并通过对话的方式回答用户提出的问题。自然语言理解技术指将表达语音的一种方式映射为计算机能理解的表达方式，其使用原理是根据上下文辨识一个多义词在指定句子中的确切意义，并根据句子的结构和词义推导该句子的句义。未来，自然语言理解技术将引入部分规则机制，利用规则和统计结合的方式弥补计算机对系统语言理解的不足。同时自然语言理解技术将开放学习机制，修正统计数据，弥补语料统计数据的局限性。

2.2.5 语音去噪技术

语音去噪技术通过控制语音通信过程中的语音质量，提高系统对于语音理解的准确性，

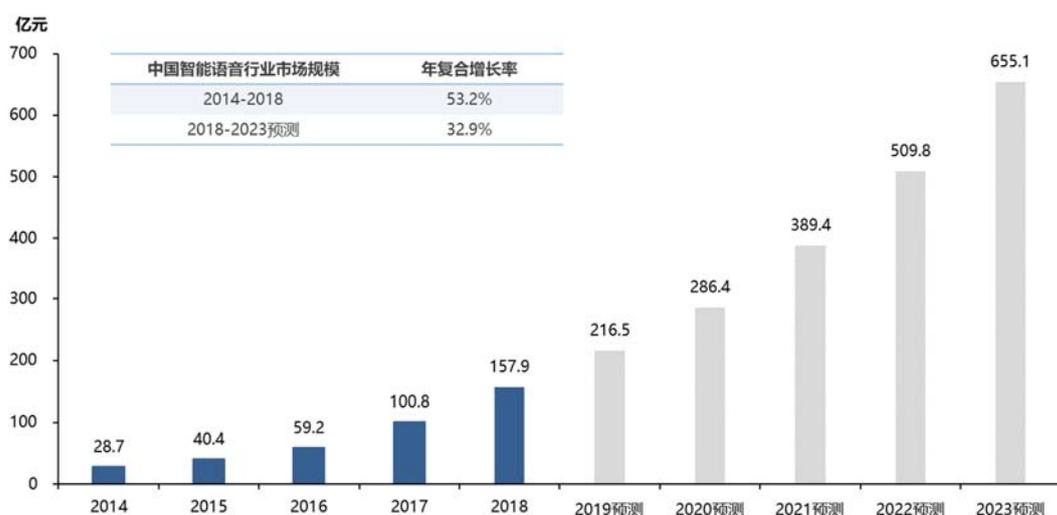
缓解噪音污染对于语音收录影响。在语音通信的过程中，实现波束形成、回声消除和噪声抑制都需要复杂的算法和大量的信号处理，回声消除算法是目前语音去噪技术较常用的算法类型。有效的回声消除算法需要持续的在一颗 DSP 芯片上运行，但有限的 DSP 芯片资源有限将影响数据传输的高效性和实时性，影响语音处理算法的算力以及语音处理系统的性能。语音去噪技术的成熟度是提高语音识别系统性能的保障。未来，语音去噪技术将不断提高与实际环境使用的结合度，减少噪音干扰对语音语义识别的影响。

2.3 中国智能语音行业市场规模

据沙利文数据统计，2014 年中国智能语音行业市场规模仅有 28.7 亿元（见图 2-1），得益于政策环境和技术实力的支持，中国智能语音行业得以快速发展。2018 年中国智能语音行业的市场规模增长至 157.9 亿元，2014 年至 2018 年中国智能语音行业的年复合增长率达到 53.2%。

具有智能语音多年从业背景的行业专家表示，中国政府在智能语音技术研发及产业化方面的利好政策频出，为智能语音产业的发展创造了良好的政策环境。在“中国制造 2025”、《新一代人工智能发展规划》以及《促进新一代人工智能产业发展三年行动计划（2018-2020 年）》中，政府均将推动智能制造行业上升为国家战略，积极推动智能产品在各场景的集成应用。尤其在《行动规划》中，政府的发展重点覆盖智能语音行业的基础层至应用层，在助力智能语音底层软硬件建设的同时，加速各领域智能产品的落地，推动智能语音行业的发展。其次，语音识别、自然语言识别等核心技术的发展推动智能语音产品的商业化落地，进一步深化智能语音在产业下游各场景的应用。同时，大数据、云计算等辅助技术性能的提升为智能语音行业发展提供助力。

图 2-1 中国智能语音行业市场规模，2014-2023 年预测



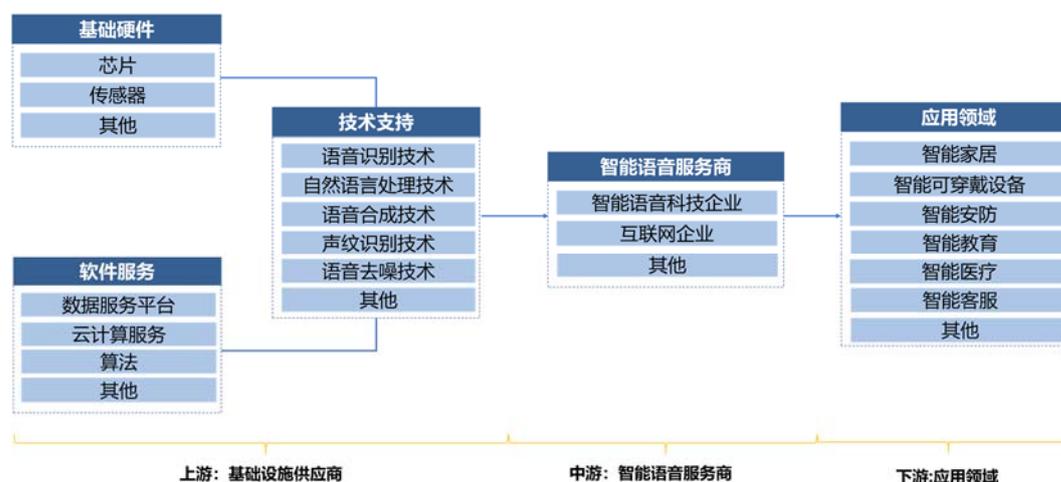
来源: fsTEAM 软件采编, 沙利文数据中心编制

目前, 中国智能语音行业处于快速发展阶段, 各场景智能语音产品的商业化应用逐渐走向成熟。人工智能、语音识别技术等核心技术的快速发展推动下游各领域对智能语音的应用需求不断扩大, 吸引政府、资本的持续关注。互联网企业、智能语音技术企业以及智能语音初创企业纷纷布局中国智能语音市场, 推动智能语音行业市场规模持续扩容。沙利文预测, 2023 年中国智能语音行业市场规模将达到 655.1 亿元。

2.4 中国智能语音行业产业链分析

中国智能语音产业链的上游参与者分为基础硬件供应商和软件服务商两类。基础硬件供应商主要为智能语音行业的上游提供人工智能芯片、传感器等智能硬件。软件服务商主要包括数据服务平台服务商、云计算服务商等参与主体。行业中游的主要参与者主要包括智能语音科技企业、互联网企业等参与主体。中国智能语音产业下游为智能语音产品及服务所覆盖的应用领域, 主要包括家居、医疗、教育等场景。

图 2-2 中国智能语音行业产业链



来源: 沙利文研究院绘制

2.4.1 中国智能语音行业产业链上游

中国智能语音产业链上游基础层主要为智能语音设备的运行提供算力。上游的基础设施供应商分为基础硬件供应商和软件服务商两类。基础硬件供应商为行业上游提供人工智能芯片、传感器等智能硬件。软件服务商主要包括数据服务平台商、云计算服务商等参与主体。

➤ 基础硬件——芯片

人工智能芯片作为智能语音产业的核心,芯片的技术成熟度将影响智能语音设备的性能。人工智能芯片行业的技术门槛较高,海外厂商技术领先,中国人工智能芯片较依赖进口,导致智能语音服务商的制造成本居高不下,压缩企业利润空间。人工智能芯片按照不同应用场景可分为通用类 AI 芯片、云端 AI 芯片、终端 AI 芯片。在通用类 AI 芯片领域,美国英伟达的 GPU 芯片占据主导地位,中国 AI 芯片企业人工智能基础层建设实力较薄弱,缺乏国际竞争力。在云端 AI 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,但竞争实力较弱。在终端 AI 芯片领域,中国企业取得了较大的突破,代表企业包括寒武纪、地平线和深鉴科技等优质厂商。伴随定制化芯片和类脑芯片的发展,中国人工智能芯片将逐渐打破海外芯片企业对中国市场的垄断。

➤ 软件服务

(1) 算法

智能语音算法主要应用于数据的计算、分析和语音识别，中国在智能语音算法领域优势较弱。算法及算法框架的技术门槛较高，以 Facebook、谷歌、IBM、微软为主的海外科技巨头占据算法行业较大的市场份额，中国仅有少数几家科技公司拥有针对算法的开放平台，如百度。核心环节技术实力薄弱削弱中游智能语音服务商的国际竞争实力。智能语音在各领域应用程度的加深导致行业对于核心算法的需求将逐渐提升，算法制约亟待解决。

(2) 云计算服务

中国云计算服务商众多，可以分为以阿里、腾讯为代表的互联网企业、以华为为代表的传统 IT 企业、以中国移动、中国电信、中国联通为代表的运营商、以及自主研发初创企业。云计算服务为智能语音设备提供强大的运算能力和资源整合能力，提高了智能语音后台技术的智能化水平，进一步提高语音识别能力，推动智能语音技术应用的落地。

2.4.2 中国智能语音行业产业链中游

中国智能语音行业中游主要包括智能语音科技企业、互联网企业等参与主体。

➤ 智能语音科技企业

中国国内专注于智能语音技术研发的代表性企业为科大讯飞和捷通华声。在智能语音行业发展前期，智能语音科技企业以研发单一的智能语音技术为主，利用其技术优势获取企业收益。近年来，伴随智能语音行业的发展，深度学习技术的应用、海量语音语料数据的积累、以及计算机算力的提升，使智能语音技术的调用得以简化，技术门槛逐渐降低，导致智能语音科技企业的技术优势逐渐被削弱，智能语音科技企业逐渐从单一的智能语音技术商转型为全方位人工智能技术服务商。智能语音科技企业依托其技术优势，积极布局下游各应用场景，通过为下游应用领域提供高端定制化的智能语音服务解决方案，加速企业产品和服务对下游的渗透。

前哨 | 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

➤ 互联网企业

基于智能语音领域较好的发展前景，以 BAT 为代表的互联网巨头开始布局中国智能语音领域。互联网企业通过开放语音生态系统，以产业内合作的方式，将语音技术植入合作方的产品中，并利用其 C 端优势推动产品在下游各应用场景的落地。互联网巨头用户流量优势明显，企业对于市场产品喜好的把控优于其他行业参与者。互联网企业可以基于对用户偏好的分析，对下游消费者进行产品的精准投放。

2.4.3 中国智能语音行业产业链下游

➤ 智能家居

在智能家居领域，智能语音通过与智能电视、智能音箱、智能照明等智能终端，以及智能家居控制中枢系统相结合。并利用语音交互技术实现对所有智能家用设备的控制，从而打破单一家用产品的智能化，构建智能家居生态。伴随语音交互、对话式交互技术的发展，用户只需向智能家居中枢系统发出指令，再由智能家居中枢系统通过语音语义识别，将自然指令转化为机器语言，向各智能终端设备发出服务信号。未来，智能语音技术的发展将推动智能语音在智能家居领域的应用将逐渐加深。

➤ 智能医疗

智能语音在医疗行业的应用主要体现在两个方面：(1) 利用智能语音技术实现病人电子病历与临床报告语音录入与转写，建立整合语音电子病历。语音电子病历的应用帮助医生在诊疗过程中实时完成病例编写，在提高医生的工作效率与工作质量同时，患者可以通过语音电子病例系统下载完整的诊疗过程和病历；(2) 伴随语音病历的积累，医院可以利用大数据技术和深度学习技术挖掘医学案例和语音资料的价值，利用智能语音技术实现辅助治疗。

➤ 车载语音

智能语音在汽车领域的应用较普遍，当用户在驾车行驶过程中活动受限时，语音交互将

成为车载场景中最适合的交互方式。在车载场景中，智能车载产品主要包括导航路线规划、语音接听电话、音乐搜索与播放、信息听写等功能。智能语音技术的发展将推动智能车载场景开发更多服务类型，如社交、娱乐、餐饮等。目前，较多智能语音服务商专注于智能车载场景的产品开发，如蓦然认知自主研发的对话应用、对话式车机 OS、智能语音座舱等产品。智能语音车载产品的落地可以在保障用户安全的前提下提升驾乘体验，打造成成熟的智能车联网系统。

➤ 智能教育

智能语音在教育领域的应用主要围绕教育体系中“学、练、测、评”等核心需求，搭建“平台+内容+终端+应用”的完整教育教学生态体系，推动教育信息化产业的快速发展。近年来，政府将教育信息化作为促进教育公平和推进教育现代化的有效手段，并且将教育信息化上升为国家战略，相关支持政策陆续出台。基于政策的支持和智能语音技术的发展，智能语音在教育领域的应用逐渐深入，结合智能语音的智能教育产品逐渐落地，其中包括智能课堂、互动教学工具、教学质量测评与分析工具、资源平台等产品。

➤ 智能客服

智能语音在客户服务领域的应用日渐深入，主要形式包括语音问答、语音质检、语料挖掘、隐私保护等。相较于传统客服，智能客服的引入和应用将有效降低企业成本，智能问答、语音质检等服务减少人工客服坐席数量及员工培训成本。同时，智能客服可以确保服务的标准化输出，且满足 24 小时全天候在线服务。此外，智能客服的应用将最大程度上保障客户隐私，隐藏客户的真实身份。因此，企业基于成本及服务标准化等方面的要求，对于智能客服的需求将逐渐提高，智能语音在客户服务领域的应用程度有望持续加深。

3 中国智能语音行业驱动因素

3.1 智能语音核心技术的发展

智能语音行业的核心技术包括语音合成技术、语音识别技术、声纹识别技术、自然语言理解技术以及语音去噪技术。除自然语言理解技术和语音去噪技术外，其余四项核心技术在智能语音领域的商业化应用较为广泛，核心技术的发展推动各应用领域智能语音产品的落地。

3.1.1 语音合成技术

在语音合成技术中，最具代表性的语音合成方法为 HMM 模型，可以在不受人工干预的情况下，构建较小尺寸的合成系统，适合应用于嵌入式设备。近年来，语音合成技术的发展已接近自然人的发音水平，并从单一广播风格发展至拥有语调、情绪的程度。语音合成技术的发展为智能语音在车载导航、移动终端等设备的应用奠定基础。

3.1.2 语音识别技术

中国语音识别技术发展较快，其发展水平已实现与国外同步。语音识别技术的发展推动了语音听写技术在医疗、报社、银行、电信等领域的应用，同时推动嵌入式识别系统进入智能终端领域。伴随以语言学为基础的文法模型的转变，语音识别技术可以保证计算机更好的理解自然语言。并且人工神经网络技术在语音识别领域的应用，可以有效缓解噪声环境下的语音识别准确度不高的问题，助力智能语音设备为用户带来较好的消费体验。

3.1.3 声纹识别技术

声纹识别技术通过探测语音信号提取用户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，进而识别发声者身份。声纹识别技术的应用可以有效解决远程身份认证、网络交易、生存认证等问题，并且保障信息传输的安全性、完整性和可靠性。因此，声纹识别

技术的发展推动智能语音在国家安全、电话银行、智能门禁等领域的广泛应用。

3.2 辅助技术的发展为行业发展创造机遇

在智能语音行业中，深度学习、大数据、云计算等辅助技术的发展提高了语音识别的准确率、降低了数据存储所需的硬件成本，为智能语音行业的发展提供契机，助力智能语音行业发展。



推广

deansel

改变营销增长格局 布局品牌私域生态

鼎栈—专注品牌私域增长落地的顾问公司

扫码咨询

网站: www.deansel.com
邮箱: info@deansel.com
电话: 156-0190-7109

杨永康
鼎栈创始人&CEO

3.2.1 深度学习

深度学习是推动智能语音行业发展的关键性技术，基于深度神经网络的语音识别技术和机器翻译技术可以大幅提高自然语言处理的准确率，为用户带来较好的消费体验。深度学习技术可以伴随数据规模的增加不断提升其性能，当海量数据持续输入存储，设备的分辨性能将不会被影响，同时保证智能语音系统使用的流畅性。深度学习技术的应用使机器可从数据中直接提取特征，产品设计者无需对每一个问题设计特征提取器，设计环节的简化将有效节省产品设计及生产成本，降低技术难度。深度学习技术的应用推动智能语音行业技术门槛的降低，吸引更多行业参与者布局智能语音市场，推动行业发展。

3.2.2 云计算

云计算服务为智能语音设备提供强大的运算能力和资源整合能力,提高智能语音后台技术的智能化水平,使终端设备获得云端服务器级别的运算能力,进一步提高语音识别能力,推动智能语音技术应用逐步落地。各智能语音服务商可以在传统云计算结构上建立新的服务层,提高智能语音设备在语音识别、自然语言处理等方面的核心能力。

3.2.3 大数据

大数据技术的发展为智能语音后端数据存储提供支持。大数据技术特有的分布式存储技术能够将单一体量巨大的语音数据文件切分成不同模块,使其能够存放于多台计算机所建构的集群中。多台计算机之间通过相互通信,将集群内的存储空间进行资源整合、数据虚拟化,并对外提供文件访问服务的文件系统。大数据技术在智能语音领域的应用极大的降低了数据存储所需的硬件成本。

大数据技术可以对单一体量的数据进行切分、存储、再整合,语音语料的数据存储只需多个性能基本达标的存储硬件,不再受限于计算机硬件的性能。

大数据技术的应用可以保证所存储数据的安全,由于大数据技术可以将单条语音数据重复分配至集群中的多个计算机中。当数据不慎丢失或破坏时,系统可以在集群中提取到完整数据。

大数据技术在智能语音领域降低了海量数据存储的难度,同时替代了高性能存储硬件的使用,为日益增长的进量数据以及原始数据的存储提供技术支持。

3.3 语音交互的便利性优势

语音交互技术的发展帮助用户实现了本能表达,解放消费者感官的占用,优化用户的消费体验。

语言交互是利用语言信息进行交互的技术,在空间上打破了用户与智能设备的距离限制。

在智能语音领域，远场识别技术的发展进一步缩小了用户与智能设备的距离限制，当用户与智能设备保持较远距离时，用户仍然可以通过语音口令与智能语音设备进行交互。

除改善用户体验的优势外，语音交互还具有明显的速度优势。在用户与智能语音进行语音交互时，语音输入效率明显高于文字输入效率。自然语言处理技术的发展缩短了语音识别的处理时间，具有明显的效率优势。

此外，语音交互的便利性将惠及各个用户群体的生活，如老年人群体、盲人群体。语音交互技术为更广泛的消费者群体带来“伴随式”体验。语音交互的便利性将推动下游应用场景对于智能语音设备的需求，助力智能语音行业发展。

4 中国智能语音行业制约因素

4.1 行业基础层技术存在短板

目前，中国智能语音行业基础层的核心关键技术积累较薄弱，算法、芯片及基础元器件的原创科技实力与国外企业相差较大。智能语音行业基础层核心技术实力欠缺，如芯片、传感器等核心元器件，不仅影响行业原创科技成果的输出，并且核心技术或生产环节长期依赖国外，不利于中国智能语音企业建立国际竞争优势。

智能语音芯片等核心基础元器件从投入到规模化生产的周期较长，在产品研发至生产期间所需资金投入较大且回报周期长，致力于基础层核心元器件研发的企业需面临较大的资金压力。近年来，智能语音行业较好的发展前景吸引资本涌入，中国国内资本持续加码智能语音领域，但资本多集中在智能语音行业的应用层，在回报周期较长的基础层领域缺乏布局。

智能语音基础层核心元器件的技术瓶颈导致企业的国际竞争力不足。中国近 90%的芯片需依靠进口，行业对于进口芯片等基础元器件的进口依赖，消耗了大量的外汇储备，导致智能语音产品的生产成本居高不下，压缩行业的获利空间，制约智能语音行业的发展。

4.2 技术瓶颈限制行业发展

4.2.1 语音识别技术“鲁棒性”问题凸显

语音识别技术在智能语音下游的应用逐渐广泛，由于语音信号的频谱受到复杂的环境因素干扰，如信道差异、测试环境差异等，语音识别系统的“鲁棒性”受影响，智能语音硬件对于用户需求识别的准确度降低，影响消费者体验。语音识别的完整过程包含语音信号识别、静音切除、声学特征提取等多个环节。基于语音信号的多样性和复杂性，语音识别系统在保证环境安静、语料的清晰、标准的情况下才能实现用户需求的准确识别。

在智能语音硬件的真实使用场景中，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,导致智能语音系统对于用户需求的准确识别产生偏差。语音识别技术的成熟度将影响各应用领域智能生态的构建及发展,语音识别技术的“鲁棒性”问题亟待解决。

4.2.2 自然语言处理技术尚不成熟

目前,自然语言处理(NLP)技术在人机交互领域仅停留在浅度处理层面,智能语音设备无法保障智能语音系统对于用户需求识别的准确率。自然语言处理(NLP)技术主要包含三个层面,分别为词法分析、句法分析、语义分析。在语义分析层面,机器对句子的理解只能做到语义角色标注层面,即标出句中的句子成分和主被动关系等,词义消歧是制约自然语言处理(NLP)技术的最大阻碍。语音识别系统在完成切词、标注词性以及词语识别后,需要对每个词语进行理解。但由于一词多义的情况无法避免,机器无法比拟人力基于语言环境和知识储备消除各种歧义的能力,因此基于自然语言处理技术的语音识别系统在实际操作情况并不理想。

5 中国智能语音行业政策

5.1 国家性政策红利助力行业发展

近年来，中国人工智能产业发展势头较好，人工智能在教育、医疗、环境保护等领域的应用程度不断加深。政府的利好政策频出，旨在推进人工智能技术与各应用领域的深度融合。智能语音作为人工智能领域较成熟的技术分支，其技术的成熟将加速各应用领域智能化产品的落地，因此，政府对于人工智能行业的利好政策将带动智能语音行业的发展。

图 5-1 中国智能语音行业国家性政策

政策名称	颁布年份	颁布主体	主要内容及影响
《中国制造2025》	2015年	国务院	加大科技创新力度，推动移动互联网、云计算、大数据等技术在智能制造领域的突破
《新一代人工智能发展规划》	2017年	国务院	以提升新一代人工智能科技创新能力为主攻方向，发展智能经济，建设智能社会，维护国家安全，构筑知识群、技术群、产业群互动融合和人才、制度、文化相互支撑的生态系统
《促进新一代人工智能产业发展三年行动计划（2018-2020年）》	2017年	工业和信息化部	《行动规划》按照“系统布局、重点突破、协同创新、开放有序”的原则，为人工智能行业的发展提出了多方面的任务

来源：沙利文研究院绘制

2017年，政府为推动人工智能与实体经济的深度融合，积极落实“中国制造2025”和《新一代人工智能发展规划》部署。由工业和信息化部印发的《促进新一代人工智能产业发展三年行动计划（2018-2020年）》按照“系统布局、重点突破、协同创新、开放有序”的原则，为人工智能行业的发展提出了多方面的任务：（1）重点培育和发展车联网汽车、智能服务机器人、智能语音交互系统、智能翻译系统、智能医疗等智能化产品，推动智能产品的经济社会的集成应用；（2）重点发展智能传感器、神经网络芯片等关键环节，夯实人工智能行业基础层软硬件的建设。《行动规划》的发展重点覆盖智能语音行业的基础层至应用层，在助力智能语音底层软硬件建设的同时，加速各领域智能产品的落地，推动智能语音行业的发展。

5.2 地方性政策红利助力行业发展

中央政府在人工智能领域利好政策的不断出台，各地方政府对于人工智能产业的重视程

度日益增强。在《促进新一代人工智能产业发展三年行动计划（2018-2020年）》中，政府表明将充分利用现有资源和手段，加强部省联动，依托国家新型工业化产业示范基地建设等工作。政府将支持有条件的地区发挥自身资源优势，培育人工智能领军企业，探索建设人工智能产业集聚区。政府鼓励各省市建设相关领域的制造业创新中心，设立重点实验室，鼓励行业合理开放数据，支持人工智能在重点行业和关键领域的应用力度，促进人工智能产业发展。截至目前，中国已有逾 20 个省市出台了人工智能产业鼓励政策，其中江苏、上海、广东、北京、浙江是国家大力发展人工智能的热点区域。在国家政策以及地方配套政策的推动下，人工智能、智能语音等领域产业链初具规模，与产业下游应用领域的结合不断深入，推动经济社会各领域从数字化、网络化向智能化加速提升。

6 中国智能语音行业发展趋势

6.1 行业将大力发展深度集成语音 AI 芯片

基于语音交互为用户带来的便利性，智能语音技术在各行业应用程度日益加深，各领域智能终端产品逐渐普及。在传统智能语音解决方案中，语音识别技术需在云端进行计算，庞大的数据和运算量导致设备运行成本和功耗居高不下，且易出现语音交互“时延”的问题，影响用户的消费体验。AI 芯片的算力决定了智能语音处理的速度以及语音交互的便利性，因此行业将大力发展深度集成语音 AI 芯片。

深度集成语音 AI 芯片利用主要算法和芯片架构深度集成，将芯片和语音识别算法技术组合成语音模块，放置在设备中。用户只需向设备发送指令，设备中的语音模块通过麦克风将语音导入设备变成电信号，语音模块再将电信号提取出来进行运算，从而实现“人机交互”。深度集成语音 AI 芯片拥有较大的数据存储空间，无需将语音语料上传至云端进行计算，保护信息安全性。深度集成语音 AI 芯片降低了设备的数据运算量及功耗，在节约设备运行成本的同时，提高智能语音处理速度，提升消费者的用户体验。

6.2 行业将构建智能语音专利池

目前，中国智能语音行业进入快速成长期，语音技术在各下游场景的商业化应用日益广泛，推动智能语音产业分工的细化以及科学技术的密集化。落地的智能语音产品会涉及众多专利技术，且分属于不同所有者，众多的专利请求、复杂的专利授权、以及高额的专利侵权赔偿等问题将阻碍智能语音新技术的实践运用。因此，中国智能语音行业有望建立智能语音专利池，引导和促进智能语音技术跨行业的实施和应用，推动语音相关产业的发展。

智能语音专利池的建立可以有效降低交易成本，消除专利实施中的授权障碍，减少企业间的专利纠纷、有利于专利技术的推广应用。专利权人可以向智能语音专利池贡献自己的核心专利，同时享有免费使用其他专利权人专利的权利，免去了专利权相互交叉许可的交易成本。专利的有效管理消除了专利实施中重复授权的障碍，有效降低了专利许可中的交易成本，满足行业需求的同时，提高社会经济效益。

构建智能语音专利池可以促进专利权人的相互合作，形成专利权人的利益联盟，使其共同开发研究智能语音产业的新型技术，保持智能语音专利池的先进性，实现利益最大化。此外，智能语音专利池的建立将提高智能语音企业的国际竞争力，实现优势互补，以对抗海外智能语音企业。

7 中国智能语音行业竞争格局

目前，中国国内智能语音服务商共分为三类：(1) 以百度为代表的互联网巨头；(2) 以科大讯飞为代表的智能语音服务商；(3) 以云知声、图灵机器为代表的深耕于垂直领域和细分场景的初创企业。部分海外智能语音服务商尝试布局中国智能语音市场，但受限于服务器不在本地，中文语音语料资源匮乏等因素，海外智能语音产品在中国智能语音市场的占有率并未形成一定规模。

7.1 智能语音科技企业

以科大讯飞为代表的中国国内智能语音科技企业拥有较强的技术优势，该类型企业凭借其较强的研发优势吸引资本涌入，在智能语音市场占据较大的市场份额。深度学习技术的发展、产业基础设施的改善，中国智能语音行业的技术壁垒逐渐被打破，其语音模型创建环节被省略，技术开发难度、成本逐渐降低。

智能语音行业技术门槛的降低，导致企业原有的商业模式被改变，原有的技术授权、提供技术解决方案、技术维护等高业务模式逐渐向技术免费、增值服务付费的模式转变，中国国内智能语音科技企业的先发优势逐渐被削弱。部分智能语音科技企业逐渐从单一的智能语音技术商转型为全方位人工智能技术服务商，企业利用其核心技术优势为下游各应用领域提供定制化产品和服务，巩固并增强其竞争优势。

7.2 互联网巨头

中国国内互联网巨头企业凭借其庞大的用户资源、各细分场景的数据积累在语音交互技术领域持续发力。以百度为代表的互联网巨头企业拥有强大的搜索引擎和完善的知识库资源存储，有助于提高语音识别在各应用场景下的识别准确率。在智能语音领域，互联网企业巨头为市场驱动型企业，拥有较强的 C 端优势。互联网企业利用其 C 端产品经验及用户数据

为产品的功能化设计提供基础，基于客户需求驱动产品的发展和落地。相较于以技术为主要驱动力的智能语音科技企业，互联网巨头的智能语音产品将更加符合用户需要，有利于互联网巨头抢占更多的市场份额，强化其竞争优势。

7.3 初创企业

智能语音初创企业以垂直领域和细分场景为突破口，切入智能语音产业，实现企业的快速增长。相较于互联网企业和智能语音科技企业，初创企业的布局范围较小，涉猎的下游应用场景较狭窄。但对于部分综合实力欠缺的初创企业，专注于智能语音细分场景有助于提升企业在该垂直领域的专业性，为未来横向切入其他智能语音场景打好基础。

7.4 典型代表企业分析——蓦然认知

7.4.1 企业简介

北京蓦然认知科技有限公司成立于 2016 年 5 月，是以认知计算、自然语言理解技术为核心的的人工智能公司，主要对外提供信号处理、语音、语义、云端服务等自动对接的一站式智能交互解决方案。蓦然认知专注于认知计算、自然语言理解等技术的研发，且拥有自主的语音交互技术。企业的对话应用、对话式车机 OS、智能语音座舱等核心产品主要应用于智能车载、智能家居、智能客服三大领域。

7.4.2 产品分类

➤ 对话应用

对话应用是企业打造车联网应用生态的入口，企业致力于打造去 APP 化的语音代理，融合“VUI+GUI”的综合性功能，以语音对流为轴心，实现跨场景融合，跨设备调用，连接各场景实现服务串联。对话应用在云端上表现为一个完整的服务分发平台，平台一端通过连接各种数字化服务，以弥补知识图谱中关于“动态知识”的不足。另一端通过运行在云端的对话算法来连接各种设备端的语音代理，为用户提供服务。

➤ 对话式车机 OS

对话车机 OS 是以对话应用为基础，整合了降噪模组、语音识别、语义理解、多轮对话、图像、大数据技术打造的车联网场景一体化。对话式车机 OS 以多轮对话的方式操作应用，通过获取内容和服务，实现多场景融合的交互能力，适配多种底层系统，打造车内交互中枢以提升用户体验。对话式车机 OS 针对车载场景，支持导航、音乐、停车、电话等出行相关的日常生活服务。企业在对话式车机 OS 应用中构建出行相关的知识图谱，建立后台数据库，建造全新的智能车联网应用生态。

➤ 智能语音座舱

蓦然认知的智能语音座舱以对话车机 OS 为核心，以语音交互为轴线，通过连通车机大屏及车内前后各功能屏，给予车内人员实时互动、内容共享的全新娱乐与车载生态服务体验。产品以统一账号连接手机端、家居端和座舱端，将移动座舱空间与家庭空间连接起来，完成任务与信息的跨场景交互，打造跨设备应用体验。

7.4.3 竞争优势

优蓦然认知的技术团队深耕自然语言处理行业多年，企业以语义理解、对话系统、学习系统、自动服务对接、语音识别、语音合成、声纹识别、唤醒+离线命令词、远场降噪、语言生成为核心，设计出能够解放用户双手的交互系统。企业的知识系统覆盖从最开始的关键词搜索、结构化搜索、构建知识图谱、上下文理解、连续对话等。知识系统包括由易到难的功能模块，为后台建立丰富严密的知识图谱。企业完整的知识体系为智能硬件的使用提供技术支持，基于用户的使用习惯提供针对性服务，以提升人与机器的交互效率。

7.5 典型代表企业分析——思必驰

7.5.1 企业简介

思必驰是专注与人性化智能语音交互技术的智能语音服务商，公司为企业和开发者提供

自然语言交互解决方案，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 5 亿元第四轮融资。

7.5.2 产品分类

➤ 智能车载

在智能车载场景中，企业为用户提供“云+端”决绝方案，利用智能语音操作系统控制智能后视镜、智能车机、便携式导航仪等设施，实现双手解放。

➤ 智能家居

在智能家居领域，企业为用户提供软硬一体化的解决方案，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。企业利用其近场远场语音交互优势控制用户的家居场景，为用户提供智能化服务。

7.5.3 竞争优势

➤ 技术优势

企业在语音识别、语音合成、对话系统、软硬件结合等方面拥有较强的技术优势。在国际抗噪语音识别测试中，企业曾两次获得单系统识别的最佳结果。并且在语音合成方面，思必驰采用参数化统一模型合成语音，用一个公式去描述声音的生成，自由定制多种合成音，让语音语料更加清晰的同时，使得语音更加个性和独特的表现力，提高系统的试错能力，降低设备的错误率。

➤ 资本优势

思必驰曾获联想之星、启迪、阿里巴巴、元禾控股、中民投资本、深创投、富士康、MTK 等多方投资。企业先后获得剑桥大学商业计划奖、剑桥大学 Downing 基金、英国政府研发基金。企业作为政府重点培育企业，曾获国家电子发展基金、工信部电子发展基金、科技部创新基金、苏州科技计划领军企业基金等支持，为企业的发展提供资本支持。

7.6 典型代表企业分析——云知声

7.6.1 企业简介

云知声是一家专注于物联网人工智能服务的智能语音识别服务商,企业的核心技术体系为语音技术、语言技术、知识计算、大数据分析等。截至 2019 年,云知声共计完成 6 轮融资,总融资金额过 10 亿元,投资方包含启明创投、高通、京东数科等。企业的合作伙伴数量超过 2 万家,覆盖用户达 2 亿,其中开放语音云覆盖的城市超过 470 个,覆盖设备超过 9,000 万台。

7.6.2 产品分类

➤ 智能家居方案—UniHome

企业通过在家居硬件中植入 AI 芯片,组成不同的硬件平台、安装不同系统下的操控系统,并将系统主控集成。主控系统利用语音识别等技术分辨用户的需求,提供语音交互、互联内容的功能,满足不同智能家居场景的硬件交互需求。

➤ 智能车载方案—UniCar

智能车载方案是集降噪、语音交互、云端计算和内容服务为一体的解决方案。在车载垂直场景,智能车载方案可满足用户车载导航、电话、娱乐、社交、社交五大场景的功能诉求。

➤ 智能医疗方案—UniHealth

企业为用户提供智能医疗方案,基于医疗人工智能技术和大数据分析进行持续探索,实现智能语音交互的知识问答和病历查询,进行健康风险预测和患者分群分析。

➤ 智能教育方案—UniEdu

在智能教育领域,企业提供的语音评测服务采用先进的语音评测技术和自然语言处理技术为客户提供全方位智能化的语言学习产品后台服务。企业利用云计算技术将自动口语评测服务放在云端,并开放 API 接口供客户远程使用。

7.6.3 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。企业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。云知声依托其 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、智慧交通等诸多场景。

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



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”、手机扫右侧二维码阅读研报



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

李先生：18916233114

李女士：18049912451