

释放数字生产力，留存探索语音内容

智能语音转写行业研究报告

©2022.12 iResearch Inc.



前言

研究背景：

在工具不发达的年代，会议记录主要依靠人力完成，以多人合作的分工形式提升记录效率。后随着记录工具不断升级和专业培训，人工转写的效率也在不断提升，专业速录师可依靠速录机完成会议等场景的转写需求，但成本较高。后随着互联网及人工智能技术的不断发展，智能语音转写产品应运而生。尤其在 2011 年，大量研究人员转向深度学习在智能语音领域的研究，利用大数据、机器学习和大算力“三驾马车”，让语音识别的识别准确度再一次得到明显提升，智能语音技术迎来落地应用的发展期。”工欲善其事，必先利其器“，智能化的语音转写服务以价优、质高、便捷的优势满足了转写记录这一交流场景的需求痛点，并在远程办公、新媒体、国际化交流的需求背景下，未来保持强劲市场增长力。

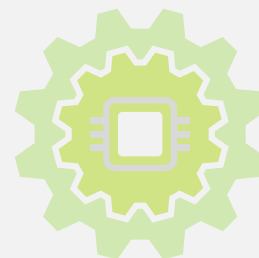
对此，艾瑞发布《中国智能语音转写行业研究报告》，从语音识别-智能转写产品角度出发，确立智能语音转写服务的范围定义，描绘智能语音转写服务的产业图谱与需求市场，梳理智能语音转写服务在 SaaS 软件服务及本地解决方案的不同产品形式、商业模式及厂商格局，并为中国智能语音转写行业的趋势发展提供分析判断，**希望通过本报告，为读者呈现中国智能语音转写的产业发展背景、行业厂商动态、产品发展机遇的多维视角，欢迎各界探讨指正。**

研究对象：

作为语音识别技术的产品应用，智能语音转写产品是可以支持长音频识别的语音转文字服务，分为实时语音转写与非实时语音转写，可为信息处理和数据挖掘提供基础。

研究方法：

本报告通过业内资深的专家访谈、桌面研究、产品对比研究、行业数据统计与行业规模推算输出相应研究成果。



报告撰写

艾瑞咨询
产业数字化研究部
人工智能研究组



智能语音 产业发展

近年来，智能语音技术与互联网、企业服务、消费硬件、传媒、医疗健康等各行业的深度融合带来了新的用户需求增长和商业模式创新，创造产业经济价值、繁荣产业生态，算法模型、优质数据集与多样化应用场景助力产业规模走高。**部分智能语音产品如语音助手、语音转写、智能客服等取得产品价值突破或商业上的显著成就，语音识别相关产品多已进入稳步上升期。**但在细分产品的交互体验、使用效果、场景优化等方面仍面临长期求索。人们面对“AI”时希望得到的自然、类人、甚至高信息密度的交互体验，仍然是一个宏伟的开放性课题。



智能语音 转写产品

语音识别产品早期主要是语音听写，即用户说一句、机器识别一句；后来发展成语音转写，更聚焦于人人交流场景。**智能语音转写是可以支持长音视频的语音转文字服务，附加产品服务、多语种翻译、内容分析等智能化功能，满足用户在会议、庭审、采访、直播、视频制作、客服质检等场景中的实时与非实时语音转写需求。**随着语音识别准确性及效率的提升、多语种与方言转写服务丰富，以及上下文纠正、标点过滤、自定义热词配置、声纹角色分离、语音内容分析提取等功能的逐步优化，**智能语音转写服务的商业化落地与多场景复用持续推进，成为语音识别产品的“排头兵”。**



智能语音 转写市场

在人力成本、协同办公、传媒音视频、会展交流、跨国沟通等多重因素驱动下，中国智能转写市场不断注入需求活力，**2021年中国智能语音转写市场规模已约为10亿元。**未来，随着智能转写的技术突破、功能丰富及场景泛化，智能转写市场规模将加速上扬，**预计2026年市场规模将达到38亿。**从产品形态来看，智能转写产品主要包括SaaS类产品与本地化部署解决方案两大类。其中，**SaaS市场头部聚集效应显著，讯飞听见与搜狗听写位列第一梯队，讯飞听见在转写准确率尤其是小语种和方言等、产品丰富度、品牌影响力和发展潜力维度拔得头筹。**未来，SaaS形式API调用与垂类解决方案将形成合力，构成智能语音转写产业既快且稳的增长飞轮，高生态活性加硬解决方案实力的企业将更能突出重围，抢占更多增量市场。



智能语音转写 趋势洞察

从技术趋势来看，语音识别技术的精度和速度仍取决于实际应用环境，面对“混合语种”“嘈杂环境”下的“多人”“交互”“重叠”等多重因素交织的复杂语音场景，语音转写技术应用仍有待突破；**从场景价值来看**，如今智能转写应用领域大多仅服务于从语音到文字转写内容的实现，未来转写应用可结合自然语言理解、机器学习、知识图谱等AI技术，拓展转写产品的场景边界，深入挖掘转写内容价值，以更高阶、智能的辅助替代角色，为客户提供问题预警、策略总结、决策分析等功能服务；**从厂商策略来看**，各家将以构建自身产品生态，加强外部场景合作为策略核心，基于自身企业特点选择差异化侧重，共同推进转写技术的应用渗透与市场发展。

智能语音转写行业——发展背景篇

1

智能语音转写行业——市场分析篇

2

智能语音转写行业典型企业案例

3

智能语音转写行业——发展趋势篇

4

智能语音产业的宏观背景

数字信息输入输出的重要载体，人工智能产业落地“先锋军”

智能语音技术指通过声音信号的前端处理、语音识别（ASR）、自然语言处理（NLP）、语音合成（TTS）等技术形成完整的人机语音交互流程，是实现人与机器交流的纽带，也是数字信息输入与输出的重要载体。近年来，智能语音技术与互联网、企业服务、消费硬件、传媒、医疗健康等行业的深度融合带来了新的用户需求增长和商业模式创新，创造产业经济价值、繁荣产业生态。智能语音产业的迅速发展促进了我国数字经济发展、提高了社会治理的智能化水平、推动了我国人工智能技术创新的战略突破。作为人工智能产业落地的“先锋军”，智能语音产业得到了国家和地方政策的有力支持，且随着参与者不断进入智能语音赛道，“百舸争流，千帆竞发”，产业技术水平和产品竞争力不断提高。

中国智能语音产业典型应用场景及政策汇总（部分）

发布日期	相关机构	重点内容
2022-05	国务院办公厅	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、自助派单功能
2021-11	工信部	工业和信息化部批复组建国家智能语音创新中心，将围绕 多语种语音识别、语音合成语义理解和专用人工智能语音芯片 等研发方向，构建集共性技术研发、测试验证、中试孵化和成果转化于一体的创新平台
2021-01	国务院办公厅	提出加强 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
2020-10	工信部	鼓励 智能家居产品普及语音控制功能 ，推动基于智能语音识别技术的智能音箱、智能可穿戴设备及其他智能家电产品开发，老年人可通过语音方式实现便捷化操作
2019-02	最高人民法院	全面提升语音识别技术在 庭审语音同步转录 中的应用效能，建成全国法院智能语音云平台，实现全国法院语音识别的模型共享和统一管理
2018-04	国务院办公厅	开展智能医学影像识别、病理分型和多学科会诊以及多种 医疗健康场景下的智能语音技术应用 ，提高医疗服务效率
2017-07	科技部	公布了首批国家新一代人工智能开放创新平台，包括自动驾驶、城市大脑医疗影像和 智能语音
2017-07	司法部	大力发展电子公证、法律服务智能保障等业务模式，推进 人工智能语音热线和社交网络法律服务机器人 技术研发，促进公共法律服务提档



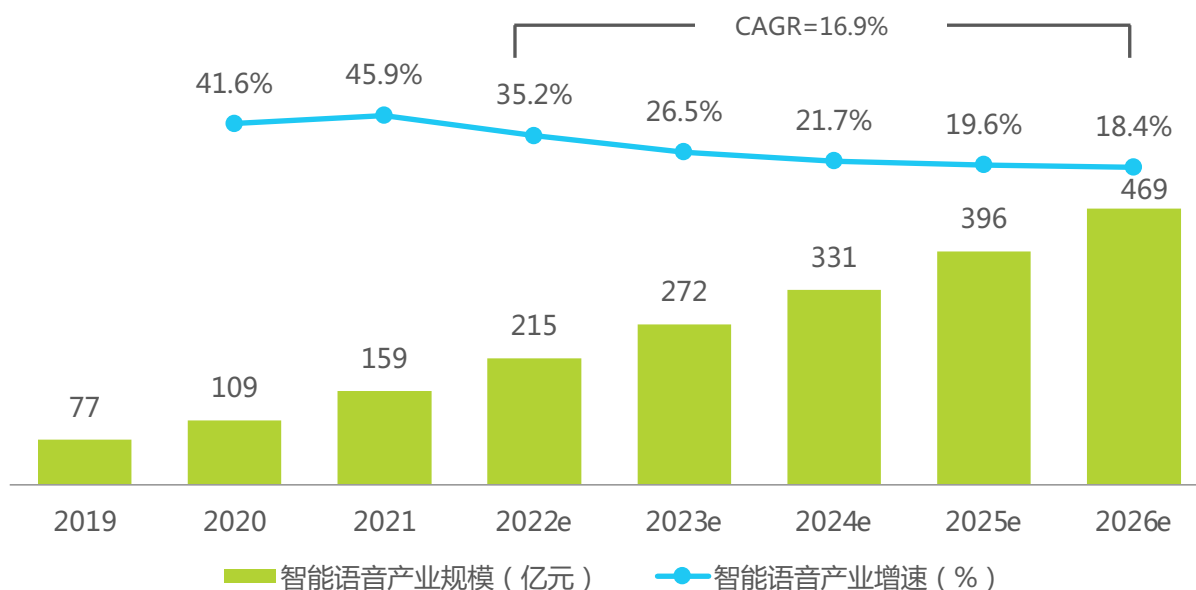
来源：艾瑞研究院根据公开资料自主研究绘制。

智能语音产业的市场规模

2022年智能语音市场规模达215亿元，产业规模持续走高

近年来，我国人工智能产业维持稳步增长态势；其中，智能语音产业基于语音识别等算法模型突破、优质数据集积累和丰富的下游应用场景创新，已进入规模化深耕阶段。我国头部智能语音企业、大型互联网企业等纷纷以“开放平台+垂直赛道”的发展模式，一方面通过语音开放平台为各行业开发者提供智能语音技术支撑，协作场景与产品创新，助力产业规模增长；另一方面凭借各自在消费硬件、协同办公、视频直播等领域的行业理解与用户生态，持续拓展智能车载、娱乐传媒、协同办公、智慧医疗、在线教育、智能家居等垂直行业赛道，以语音为信息的出入口，构建泛语音产业生态集群。2022年中国智能语音产业规模可达215亿元且维持较高增速，预计到2026年产业规模可达469亿元。

2019-2026年中国智能语音产业规模



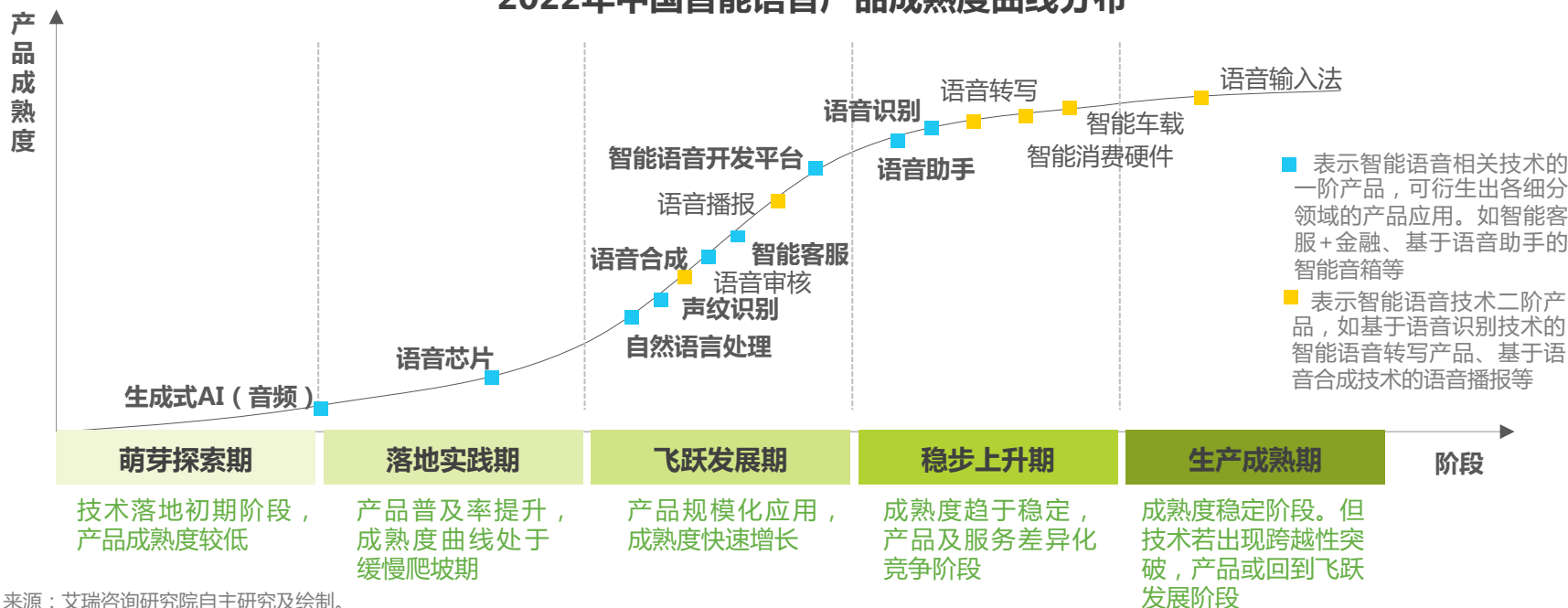
注释：智能语音典型产品包括对话机器人、智能硬件中的AI语音助手以及教育、医疗、司法、公安、互联网等垂直行业中的智能语音产品及应用。
来源：艾瑞咨询研究院根据专家访谈，结合艾瑞统计模型自主研究绘制。

智能语音产业的产品成熟度

语音识别相关产品多已进入稳步上升期

人类对机器语音识别的探索始于20世纪50年代，迄今已逾70年。2016年，在深度神经网络的帮助下，机器语音识别准确率第一次达到人类水平，意味着智能语音技术落地期到来。后随着近场语音识别准确率提升、远场语音识别和唤醒发展、全双工语音交互出现、基于NLP的对话和问答能力逐渐成熟、知识图谱技术助力对话引擎以及针对实际应用中的算法优化，智能语音技术的落地可用性不断突破。但其背后涉及的声学研究、模式识别研究、通用NLP研究及垂直场景的深度语义理解等还未成熟到拼成一个没有明显短板的“木桶”。因此尽管部分智能语音产品如语音助手、语音转写、智能客服等已取得了产品价值突破或商业上的显著成就，但在细分产品的交互体验、使用效果、场景优化等方面仍面临长期求索。人们面对“AI”时希望得到的自然、类人、甚至高信息密度的交互体验，仍然是一个宏伟的开放性课题。

2022年中国智能语音产品成熟度曲线分布



来源：艾瑞咨询研究院自主研究及绘制。

智能语音转写的定义与分类

语音识别产品的重要输出形态，分为实时与非实时语音转写

语音识别产品早期主要是语音听写，即用户说一句、机器识别一句；后来发展成语音转写，更聚焦于人人交流场景。智能语音转写是可以支持长音视频的语音转文字服务，分为实时语音转写与非实时语音转写，可为信息处理和数据挖掘提供基础。适用于线上线下会议记录转写、影视字幕制作、媒体新闻工作、会议翻译等多个应用情境。作为数字化劳动力，解决刚需问题，有效提高办公效率。随着语音识别准确性及效率的提升、多语种与方言转写服务丰富，以及上下文纠正、标点过滤、语气词过滤、自定义热词配置、声纹角色分离、语音内容分析提取等智能化服务功能的逐步优化，智能语音转写服务的商业化落地与多场景复用持续推进，成为语音识别产品的“排头兵”。

智能语音转写产品定义与分类



预览已结束，完整报告链接和二维码如下：

https://www.yunbaogao.cn/report/index/report?reportId=1_50174

