

中国智能语音助手企业案例研究 报告

2018年





近年来，人工智能发展迅速，也得到越来越多的关注。其中，语音类人工智能中的重要行业应用“智能语音助手”也成为了人工智能的一个主要应用领域。目前行业里的智能语音助手通常应用于消费级产品和专业级行业应用两大领域。



智能语音助手用户多分布在经济发达地区，开始尝试使用智能语音助手的最主要原因是想通过语音操控从而解放双手，并且智能语音助手对语音语义识别的准确性也成为了用户选择及使用最重要的考量因素。



该份案例报告通过对智能语音行业的市场动态、趋势，智能语音助手用户的行为习惯和需求以及典型企业案例进行分析，希望为行业发展提供参考。

概念定义

人工智能，英文全称为“Artificial Intelligence”（AI），是计算机科学领域的一个分支，包涵研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统等。人工智能企图通过模拟人类的思考反应方式，生产出具有与人类相似思考反应能力的机器，核心的三大基础技术是计算机视觉、智能语音和机器学习。

其中，智能语音技术目的在于实现人机语言的通信，使人与机器之间能够通过纯语音进行交互，包括语音识别（ASR）、语音合成（TTS）和自然语言处理（NLP）三项主要技术。



中国智能语音助手行业发展背景 1

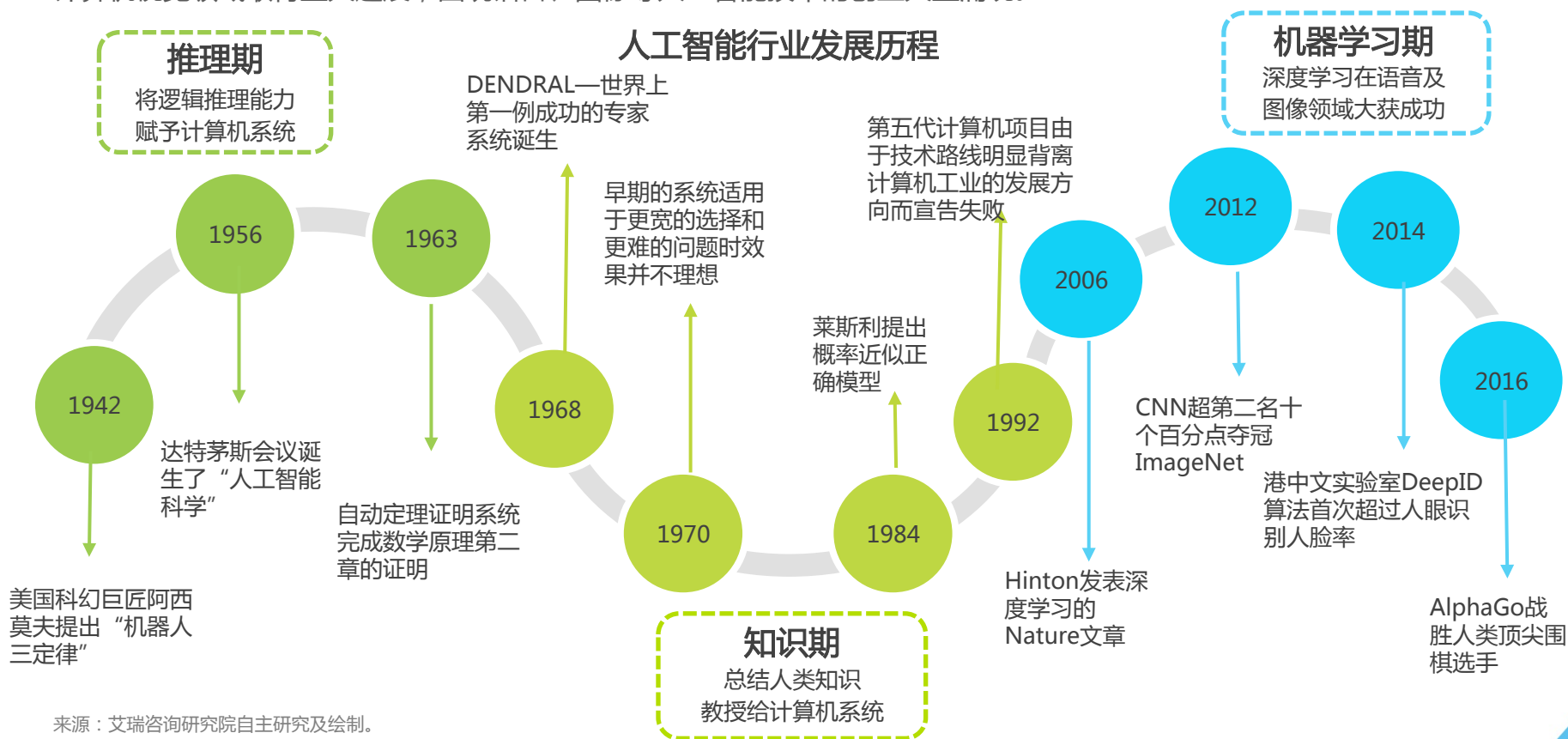
中国智能语音助手用户研究 2

咪咕灵犀案例分析 3

人工智能行业发展现状

语音识别、计算机视觉领域取得重大进展

在20世纪50年代到70年代初，人工智能的研究尚处于“推理期”，人们认为如果赋予机器逻辑推理能力，机器就能具有智能。到了20世纪70年代，人们意识到人类之所以能够判断、决策，除了推理能力之外，还需要具备一定的知识。发展到20世纪80年代，机器学习真正成为一个独立的学科领域，相关技术层出不穷。2010年后，“人工智能”相继在语音识别、计算机视觉领域取得重大进展，围绕语音、图像等人工智能技术的创业大量涌现。



来源：艾瑞咨询研究院自主研究及绘制。

人工智能分类

虚拟语音助手是人工智能的重要应用领域

计算机视觉、智能语音和机器学习是人工智能的三大核心基础技术，目前研发出的人工智能应用大多是这三种技术综合运用的结果，只是其中的主次之分不同。依据核心基础技术类目，可以将人工智能分为两大类，即感官智能和决策智能，其中视觉智能、语音智能和深度学习智能是感官智能和决策智能下的三大子赛道。目前，语音智能的一个重要行业应用就是虚拟助手，即“智能语音助手”。它的核心在于人类通过纯语音信息实现与机器的交互，让智能机器“助手”帮忙完成指派的任务。

人工智能行业分类

感官智能

决策智能



视觉类
计算机视觉

目前主要应用于辅助驾驶、图像/视频编辑、图片识别、人脸识别、文字识别、视频监控、工业视觉检测、三维视觉、医疗影像诊断等。



语音类
智能语音

目前主要应用于智能家居、虚拟助手、可穿戴设备、智能车载、智能医疗、陪伴机器人等。



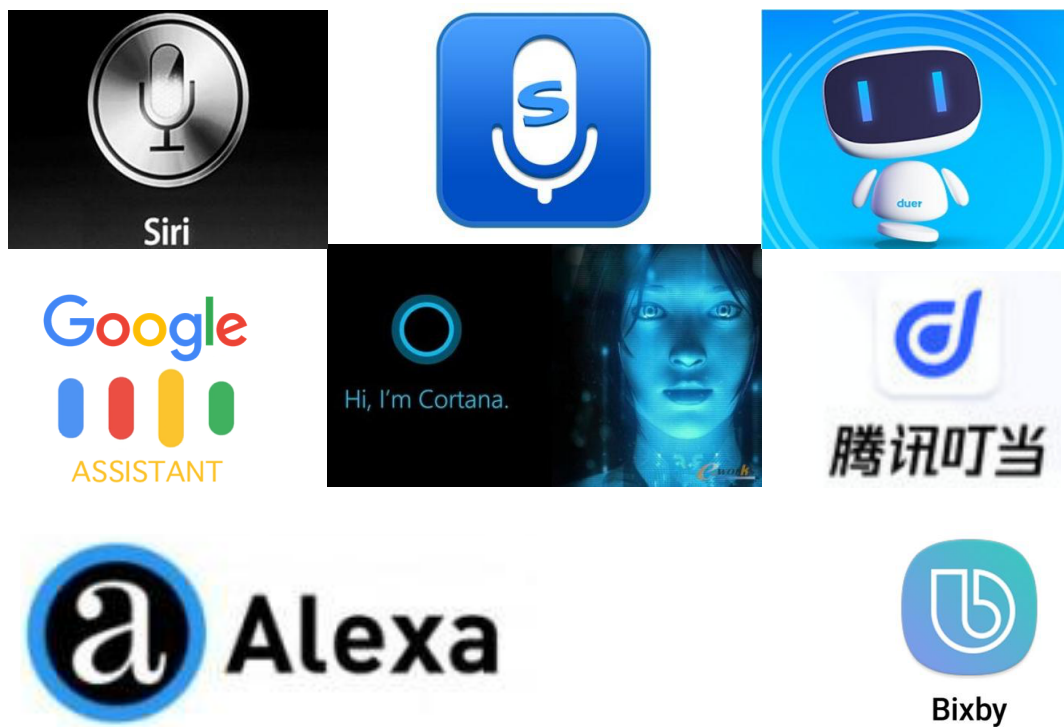
深度学习类
机器学习

目前主要应用于精准营销、风险预估、健康监测、优化运营、AlphaGo等。

智能语音助手行业现状

互联网巨头纷纷进入布局

从2010年开始，互联网巨头们纷纷通过自主研发或并购/参股的方式开始探索智能语音产业，其中，智能语音虚拟助手成为重点布局对象，此外，为占据一定的市场先机，苹果、谷歌、微软、亚马逊、百度、腾讯、搜狗等巨头们也陆续开始在智能车载、智能家居、智能医疗、可穿戴设备等诸多细分市场寻求突破。



智能语音技术及其发展

语音识别、声纹识别、语音合成等

人类因为具有语言的能力而区别于其他物种，智能语音技术即研究人与计算机直接以自然语言的方式进行有效的沟通的各种理论和方法，涉及机器翻译、阅读理解、对话问答等，因为语言在词法、句法、语义等不同层面的不确定性及数据资源的有限性、背景知识的复杂性等各方面限制，智能语音技术仍有非常大的提升空间，仅在特定领域可取得较好的应用，鲁棒性存在大量挑战。在自然语言处理之前，声纹识别可根据说话人的声纹特征识别出说话人，语音识别技术可赋予机器感知能力（在深度学习的驱动下，目前近场语音识别准确率可达98%，远场、抗噪、多人等非限定或配合条件下的识别有待改进），将声音转为文字供机器处理，在机器生成语言之后，语音合成技术可将语言转化为声音，形成完整的自然人机语音交互，这样的语音交互系统可看作一个虚拟对话机器人，具体流程如下。

智能语音交互系统的技术流程



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_21259

