

# 智能语音助手用户研究： 理论进展与实践启示\*

赵一鸣<sup>1,2,3</sup> 陈宇<sup>2,4</sup> 刘齐平<sup>5</sup>

(1. 武汉大学信息资源研究中心, 武汉 430072; 2. 武汉大学信息管理学院, 武汉 430072;  
3. 武汉大学大数据研究院, 武汉 430072; 4. 武汉大学图书情报国家级实验教学示范中心, 武汉 430072;  
5. 湖北经济学院信息管理学院, 武汉 430205)

**摘要:** 智能语音助手已经成为国际研究热点, 而该主题的中文研究成果还非常匮乏。使用“检索-评估-综合分析”的系统性综述方法, 研究智能语音助手用户研究领域的英文文献, 通过对文献基本信息、研究设计、研究发现和实践启示等进行编码, 揭示该领域在用户认知、用户行为、用户评价3个方面的研究进展, 并为相关研究设计和产品设计提供建议。

**关键词:** 智能语音助手; 用户研究; 系统性综述方法

**中图分类号:** G353.11      **DOI:** 10.3772/j.issn.1673-2286.2023.05.004

**引文格式:** 赵一鸣, 陈宇, 刘齐平. 智能语音助手用户研究: 理论进展与实践启示[J]. 数字图书馆论坛, 2023(5): 26-34.

随着人工智能对信息领域的影响日益显著, 智能语音助手作为AI时代的智能化人机交互界面之一, 已经得到学界和业界越来越广泛的关注和应用, 如医疗健康、智能家居、自动驾驶、个人与集体的知识管理等<sup>[1]</sup>。当前, 国际主流的智能语音助手包括Apple Siri、Google Assistant (原Google Now)、Microsoft Cortana、Amazon Alexa等, 国内比较成功的智能语音助手有百度小度、小米小爱、咪咕灵犀、天猫精灵等。

自新冠疫情暴发以来, 越来越多的人在家办公, 使用智能语音助手的群体占比从2020年1月的46%上升到52%, 甚至更高<sup>[2]</sup>。可以预见, 在未来会有越来越多的用户使用智能语音助手, 而且智能语音助手对老年人、残疾人等信息弱势群体比较友好, 语音交互的便捷模式能够拉近老年人等信息弱势群体与智能产品的距

离, 有助于解决国务院提出的切实解决老年人运用智能技术困难的问题。因此, 开展智能语音助手用户研究有非常重要的理论和实践意义。

在中文社会科学引文索引(CSSCI)的题名、关键词字段中使用“智能语音助手”作为关键词检索, 发现CSSCI中与智能语音助手相关的论文非常少<sup>[1,3-4]</sup>。与此同时, 国际上存在较多的该主题英文论文。相比于传统的文献综述, 系统性综述具有更透明化、更结构化的分析过程, 能够最大限度地避免文献选择偏差, 从分散的文献中总结重要的研究主题, 已广泛应用于社会学、管理学等领域<sup>[5]</sup>。在知识较为分散的语音助手研究领域, 系统性综述法是合理且有效的。因此, 本文采用系统性综述的方法, 围绕国际上的语音助手用户研究展开分析和综述, 从分散的研究中揭示相关的研究热点,

收稿日期: 2023-04-01

\* 本研究得到国家自然科学基金面上项目“‘搜索即学习’视角下的用户学习状态感知与智能响应研究”(编号: 72274146)、国家自然科学基金面上项目“探寻式搜索过程中的路径识别与评价研究”(编号: 71874130)、国家自然科学基金创新研究群体项目“信息资源管理”(编号: 71921002)资助。

弥补当前研究的不足。

本文主要回答两个研究问题: ①国际上智能语音助手用户研究的研究热点包括哪些方面; ②从智能语音助手用户研究中得到哪些实践启示。其研究价值在于: ①可以弥补国内智能语音助手用户研究方面的不足, 进一步推动智能语音助手场景下的用户研究, 把握现有研究的重点、热点、不足以及研究的趋势, 有助于中国学者对未来研究内容的把握, 丰富该领域的研究成果; ②总结学者提出的设计建议, 有助于智能语音助手的设计优化, 从而更好地为智能语音助手产品的设计开发提供指导性的建议, 促进人工智能生态圈的和谐发展。

## 1 研究方法过程

遵循系统性综述研究方法的SALSA (Search-Appraisal-Synthesis-Analysis) 框架<sup>[6]</sup>, 分为文献数据获取、数据评估、数据编码和数据分析4个步骤, 对现有的文献进行搜索、评估、综合和分析。

### 1.1 数据获取

(1) 选择数据库。由于智能语音助手研究分散在不同领域, 选择了6个覆盖智能语音助手相关文献的主要数据库, 它们分别是Web of Science、ACM (Association for Computing Machinery) 数据库、IEEE Xplore、ScienceDirect、EBSCO数据库和Scopus。这6个数据库具有可信度和权威性, 并且覆盖了智能语音助手用户研究的主要出版源。

(2) 确定检索式。通过探索式搜索, 逐步确定检索策略。以Feine等<sup>[7]</sup>的研究为基础, 先使用“智能语音助手”及其相关术语在标题范围内检索, 再在检索结果中使用“用户研究”及其相关术语在摘要或全文中进行检索。在时限方面, 智能语音助手在2010年之后才逐步得到应用, 因此, 检索文献的时间范围为2010年1月—2022年10月。

(3) 补充数据。为了保证检索结果的全面性, 以代表性智能语音助手的名字为检索词进行补充检索。补充检索的检索式为Abstract=(Alexa OR Siri OR Bixby OR “Google Now” OR Cortana) AND ALL=(participant\* OR subject\*), 并对检索结果进行人工

筛选。

### 1.2 数据评估

本步骤的目标是确定研究文献的纳入/排除标准, 以便保留最具代表性的作品。本文关注的智能语音助手是一种基于软件的系统, 旨在使用自然语言与人类进行语音交互。这不仅排除了基于文本的对话系统, 也排除了使用受限的命令短语或预定义的关键字集的对话系统<sup>[8]</sup>。被纳入的研究文献必须与用户研究有关, 纯技术类、系统设计与开发类文献被排除在外。此外, 文献必须为全英文撰写, 收录文献的出版物必须是会议论文集或期刊, 短文章、展板、说明、书评和简讯被排除在外。

### 1.3 数据编码

通过搜索和评估筛选出来的文献构成了本研究的数据集。对于数据集中的每一篇文献, 依据表1中的字段提取相应特征。需要说明的是, 通过对数据集内所有文献的内容进行归纳, 得到了三大研究主题: 用户认知、用户行为和用户评价, 并以此作为研究发现的主要编码。

表1 文献数据编码表

类别	项目
文献基本信息	出版年
	出版物
	作者和单位
研究设计	研究方法
	用户实验任务
	使用设备
	实验人群
研究发现	研究场景
	用户认知
	用户行为
实践启示	用户评价
	面向用户与环境
	面向功能设计
	面向隐私与安全
	面向个性化设计

## 1.4 数据分析

在6个数据库中共检索到2287篇文献,删除重复的129篇文献后,由研究人员阅读每篇文献的标题和摘要以对剩余文献进行预选。根据纳入/排除标准,研究人员在剔除1793篇不符合标准的文献之后,对剩下的365篇文献进行全文阅读。经过层层筛选,最终选定107篇文献作为分析的数据集。

## 2 智能语音助手情境下的用户认知、行为与评价

通过对数据集中每篇文献的“结果”“讨论”和“结论”部分的内容进行编码,并进行综合分析,归纳出3个主题,分别为智能语音助手情境下的用户认知、用户行为与用户评价。

### 2.1 用户认知

对智能语音助手用户认知的研究主要集中在感知有用性、感知易用性、感知价值、隐私风险感知、学习和期望等方面。

(1) 感知有用性。感知有用性是指用户感知到的产品功能的有用程度。感知有用性是常见的用户接受语音助手的影响因素<sup>[9]</sup>。绝大多数用户对智能语音助手表现出积极的态度,认为它们能够使困难的任务变得容易,并简化用户的日常工作。但是,也有一些因素影响了用户的感知有用性,比如智能语音助手收听时间过短、固定位置限制了互动、不能持续地帮助用户实现他们的目标<sup>[10]</sup>等,尤其是在处理复杂问题上,语音识别质量限制了以英语为第二语言使用者和有口音使用者<sup>[11]</sup>的感知可用性。虽然智能语音助手被认为是盲人最友好的现代智能产品,取代了许多不同的设备,但是其视觉反馈设计(如移动设备屏幕显示自动语音识别内容、智能音箱表面灯光提示)对盲人用户来说是不友好的<sup>[12]</sup>。

(2) 感知易用性。语音交互的易用性为用户接受并使用智能语音助手的常见原因之一<sup>[13]</sup>。与文本交互相比,语音交互具有更高的感知效率、更低的认知成本、更多的乐趣和更高的服务满意度<sup>[14]</sup>。用户更喜欢语音输入的方式,因为这是一种比打字更快的输入形式,

用户不必担心单词的拼写问题,并且可以从智能语音助手获得音频回复<sup>[15]</sup>。用户可以很容易地通过唤醒词或设备上的物理按钮来激活智能语音助手<sup>[12]</sup>,智能语音助手使用门槛较低。但是,语音识别技术的痛点使得在某些情况下,用户需要仔细考虑他们的措辞,并且用一种非常清晰且特定的方式组织语言以便智能语音助手正常理解<sup>[11]</sup>,这在一定程度上影响了用户的正常使用。

(3) 感知价值。感知价值是指用户感知到的使用价值,无论是普通用户群体还是信息弱势群体都对智能语音助手使用价值表示了肯定。用户分享他们与对话式AI互动的动机包括功能价值、享乐价值和社交价值<sup>[16]</sup>。智能语音助手带来的象征性好处、社会存在感和社会吸引力亦对用户使用该技术产生了积极影响<sup>[17]</sup>。用户普遍觉得语音交互比触控交互更有趣、更有吸引力,也更快乐和活跃。同时,智能语音助手在唤起有视觉和/或运动障碍的用户的独立感和授权感方面亦发挥了重要作用<sup>[12]</sup>。低收入的老年弱势群体希望使用智能音箱作为日常助手和健康服务工具<sup>[18]</sup>。在个人生活上,智能语音助手也能为独立生活的老年人远程监控个性化的家庭锻炼活动<sup>[19]</sup>。智能语音助手在支持用户心理健康方面具有较大的潜力,比如在专业指导下通过音乐疗法更好地帮助用户进行心理治疗<sup>[20]</sup>,也能在新冠疫情期间提升学生自主学习效果<sup>[21]</sup>。

(4) 隐私风险感知。研究发现,隐私风险与用户对语音助手的态度呈负相关关系<sup>[22]</sup>。用户普遍都有对个人信息隐私的担忧,比如黑客攻击、记录私人对话<sup>[23]</sup>等,尤其是在发现某些家用智能语音助手可能会在没有唤醒词的情况下记录语音时,用户的担忧会增加。用户隐私风险感知的主要影响因素是用户数据去向的不确定性和系统数据使用的整体透明度。

(5) 学习。学习是指作用于认知层面并激发用户学习的心理与行为。智能语音助手对用户某一行为的干预能够提升用户的认知并让用户产生相应的学习行为<sup>[24]</sup>。智能语音助手不仅可以作为一种信息来源满足儿童的好奇心,在儿童的自主学习中发挥作用,还能通过有声书提高有阅读障碍的用户的学习能力。

(6) 期望。期望是指用户对智能语音助手的期待,比如在多轮对话、多意图理解、自动关联上下文和拟人性等方面的期待。对于日常使用的智能语音助手,一方面,用户期望对话的自然性,这表现在使用的连续性和连贯性上:用户期望在同一时间进行一系列任务时,智

能语音助手能始终保持收听状态, 智能语音助手可以对数据进行简要分析, 与用户实现互动<sup>[25]</sup>, 从而实现多轮对话, 而不是每次都需要用唤醒词或点击物理按钮来重新激活对话; 同时, 大多数用户希望智能语音助手能够自动关联上下文从而推断出当前任务的语境<sup>[11]</sup>。另一方面, 用户期望智能语音助手具有更多拟人的特性, 如在交互中出现幽默或讽刺的互动<sup>[11]</sup>, 成为用户的管家并使用户的生活更加智能化等。除此之外, 还有用户表达了对功能的更多期待: 用户希望智能语音助手能够实现多任务处理和跨设备集成, 希望能够使用Siri控制所有的苹果产品, 而不仅局限于智能手机, 这也在一定程度上显示了构建智能语音助手生态链的必要性。智能语音助手可轻易提供简单信息, 但用户更期望获取更有深度的信息<sup>[26]</sup>。

## 2.2 用户行为

用户行为方面的关注点可以归纳为3个方面: 用户选择的设备放置位置、用户在使用功能上的偏好、人机交互行为。

(1) 用户选择的设备放置位置。智能语音助手放置的位置一般与用户特定的使用目的相契合。对于家用智能音箱, 卧室是最常见的位置, 但Amazon Alexa用户更倾向于将其放在客厅或厨房<sup>[27]</sup>。对于移动设备上的智能语音助手, 用户更倾向于在私人场所使用其传递信息<sup>[28]</sup>; 在公共场所, 用户因担心引起他人注意或打扰他人, 造成社会性尴尬或不快感, 避免使用移动设备上的智能语音助手<sup>[12]</sup>。

(2) 用户在使用功能上的偏好。相比于聊天机器人, 用户更喜欢使用智能语音助手来达成任务和社交目的。在获取信息方面, 用户对聊天机器人和智能语音助手的态度没有区别<sup>[14]</sup>。道德感强的参与者往往会拒绝对抗性的智能语音助手<sup>[29]</sup>。智能语音助手的载体类型决定了用户在使用功能上的偏好。移动设备上的智能语音助手常用于访问应用程序(如打电话、发短信、查地图等)和信息查询(如天气查询、网页搜索等)<sup>[28]</sup>。以智能音箱为载体的智能语音助手常被用于搜索、娱乐(播放音乐、讲笑话和玩游戏等)以及物联网控制(开关灯和调节室温等)。用户需要智能语音助手提供交互指导时, 更偏向智能语音助手提供基于示例的指导, 这也为用户偏好研究提供了方向。

随着使用时间的延长, 基于智能语音助手用户对

功能的偏好, 信息资源管理领域信息搜索停止行为和停止规则研究有了新的情境。例如, 当用户使用智能语音助手进行多次信息搜索, 产生搜索困难时, 用户可能更倾向于使用智能语音助手处理简单的任务(例如天气查询、音乐快速查询等), 并避免将其用于更复杂的信息搜索<sup>[27]</sup>。

(3) 人机交互行为。用户与智能语音助手的交互范式可以归纳为“触发词、问题、答案”。在交互中, 用户与智能语音助手的交互行为呈现出协同化、多主体的特征<sup>[30]</sup>, 并且正在逐渐形成新的“社交礼仪”。用户在交替或协同使用智能语音助手时, 其他用户会保持沉默以等待设备响应<sup>[15]</sup>, 智能语音助手在轮流讨论、回答提问的情境中被更多地使用。智能语音助手常常被多用户使用, 能够支持用户使用语音查询式进行协同信息检索。此外, 不同年龄和性别的用户与智能语音助手的交互行为存在差异<sup>[31]</sup>。

同时, 智能语音助手用户研究扩展了信息资源管理学科关于查询拓展、查询重构的研究边界, 丰富了相关的理论。针对请求失败的情形, 用户一般会采取以下行为策略: 重复查询、重新措辞、简化查询语句、添加语境信息、改变句子结构、回忆以往成功的交互经历以重新制定对话策略、重启系统。对于特定用户, 比如非母语用户, 有时很难通过一次交互得到满意的答案, 需要多次修改和重复语音查询式。儿童和老年用户查询重构的策略则显得比较单一, 尤其是儿童在与智能语音助手进行交互时, 往往会对看护者的会话修复策略形成依赖<sup>[25]</sup>。对于没有技术知识的用户而言, 他们的交互模式倾向于人与人之间的对话, 具体表现为以更完整、自然的句子进行互动, 然后对其进行修正<sup>[11]</sup>。

## 2.3 用户评价

已有研究主要集中在拟人性评价和满意度评价两个方面。

(1) 拟人性。拟人性是指智能语音助手人性化的特征, 包括可接近性、情感、专业性、智能性以及个性等<sup>[32]</sup>。可接近性表现为用户与智能语音助手在情感/社会上的联系程度, 较缺乏社会联系的人更倾向于将智能语音助手拟人化, 有时还会通过它来寻求社会联系。特别是在新冠疫情期间, 用户更加重视智能语音助手的拟人性<sup>[33]</sup>。智能语音助手的情感梯度声学实现也会影响用户拟人性评价<sup>[34]</sup>。从智能语音助手的主要功

能来看,用户对其专业性和智能性的评价是较为积极的,认为它们能够扮演好助理的角色。但是用户普遍认为这种人性化特征是有限的,评价集中体现在缺乏感情、不够亲切等方面。但必须承认的是,智能语音助手的拟人性特征在一定程度上改善了用户体验。当智能语音助手与使用者的个性相符时,用户会对它产生更多的信任与好感,具有共情能力的语音助手更善于调节用户情绪和行为<sup>[35]</sup>。

(2) 满意度。满意度描述了用户对智能语音助手的积极情感评价,主要来源于智能语音助手的性能和个性化方面。首先,语音识别质量对用户满意度有显著影响,以英语为母语的人比非英语母语的人对英语智能语音助手有更高的满意度<sup>[36]</sup>。由于语速较慢,音间沉默时间较长,老年用户的言语可懂度略低,这可能导致老年用户满意度较低<sup>[15]</sup>,但老年用户对智能语音助手的整体评价还是非常正面的。系统的沟通风格、对情境的适应性也会影响用户的感知质量,进而影响用户的满意度。其次,用户的满意度还取决于投入的时间以及交互程度。经常使用智能语音助手的人有着更高的满意度<sup>[11]</sup>,这是因为他们越来越习惯于智能语音助手的响应,而成功的交互更有可能激励他们不断探索智能语音助手的功能边界。此外,由于智能语音助手与传统交互工具不同,用户评价研究更需借助面部表情测量及特征提取<sup>[37]</sup>。

### 3 智能语音助手产品设计的实践启示

#### 3.1 面向用户与环境的智能语音助手产品设计

随着家用智能语音助手得到广泛应用,智能语音助手产品设计应该面向长期使用、面向用户、面向环境、面向特殊人群需求。例如,建议智能语音助手的界面设计支持用户更好理解,并将多模态语音用户交互(Voice-User Interface, VUI)应用于不表现出试错行为的用户<sup>[38]</sup>。手机等移动场景下的产品设计要考虑到环境因素并作出适当的响应,例如在公共场所提供“谨慎”模式以保护用户隐私<sup>[12]</sup>。智能语音助手交互模式便捷,拉近了信息弱势群体与智能产品的距离,设计者应更多地考虑信息弱势群体,比如:考虑为盲人提供替代视觉提示的功能,使盲人用户也能平等地使用智能产品;还可以考虑提供一个无障碍选项,让残疾人在使用

智能语音助手时有更长的输入时间;考虑设计具有面部表情的情感智能语音助手,对阿尔茨海默病患者进行护理<sup>[39]</sup>。

#### 3.2 面向功能的智能语音助手产品设计

设计者应该提供能对用户产生功利价值和效率价值的实用功能,而不是仅以娱乐体验吸引用户,比如:可以通过语音命令创建快捷方式来帮助用户完成重复性任务<sup>[40]</sup>;可以主动问候用户,开始一个非正式的聊天,或在任务中进行更频繁的交互,使任务更有趣。效率价值也体现在设备集成<sup>[12]</sup>和第三方服务整合<sup>[41]</sup>上:一方面,智能语音助手能够操作更多设备,如智能手机、电视、电脑等,提高用户的跨平台协同使用效率;另一方面,它能扩展更多应用程序的使用场景,有助于智能服务融入用户的生活。

在功能上,建议强化智能语音助手在客厅、卧室等生活使用场景中对特定命令的响应能力,并通过混合语言环境训练模型,实现更好的语音识别能力和语言表达能力。为了改善用户体验,设计者可以向用户解释为什么等待时间较长、为什么语音识别失败<sup>[12]</sup>、如何改进查询结果等问题。此外,现有的研究还探索了智能语音助手其他可能的使用场景,比如智能语音助手作为主持人,可以提升群体讨论的质量<sup>[42]</sup>

#### 3.3 面向隐私与安全的智能语音助手产品设计

设计者的首要任务是确保用户与智能语音助手交互的安全性与隐私性<sup>[15]</sup>。现有研究建议增强数据透明度、功能可控性和设备安全性,使用户能够掌握语音数据的去向、监听功能的开闭,并防止黑客的入侵。此外,还应当优化智能语音助手对于儿童数据的保护功能,防止错误识别语音导致的儿童个人信息泄露。

#### 3.4 面向个性化服务的智能语音助手产品设计

智能语音助手的用户界面设计需要考虑不同用户差异化的初始期望。针对不同用户的偏好,智能语音助手的人格特征应有不同<sup>[43]</sup>。智能语音助手系统应实现个性化<sup>[12]</sup>,通过挖掘用户重复的使用模式,理解那些语

义更为模糊的命令。智能语音助手应允许用户自定义语速、音色和音量等<sup>[12]</sup>。此外,还可以通过询问用户的偏好和期望的行为<sup>[44]</sup>、根据不同的用户类型配置不同的系统文件、提高语音助手的情感表达能力等方式来实现个性化的设计。例如在军事环境中,用户更渴望智能语音助手有更具象的手势或情感。

此外,设计者应该加强智能语音助手的类人对话能力,考虑人机交互关系对用户心理的影响。可以采用更加人性化的语音反馈技术,比如使用简洁、幽默、充满情感的表达方式等,并根据用户的反馈调整对话风格,及时依据理解错误的信息进行自我修复,以提高用户满意度。

## 4 讨论

通过对智能语音助手用户研究文献进行系统性综述分析,回答了上文提出的两个研究问题。

### 4.1 国际上智能语音助手用户研究的热点内容

已有研究成果主要集中在用户认知、行为与评价3个方面。这3个方面的研究并不是孤立存在的,而是相互联系并相互影响的,形成了完整的研究闭环,如图1所示。

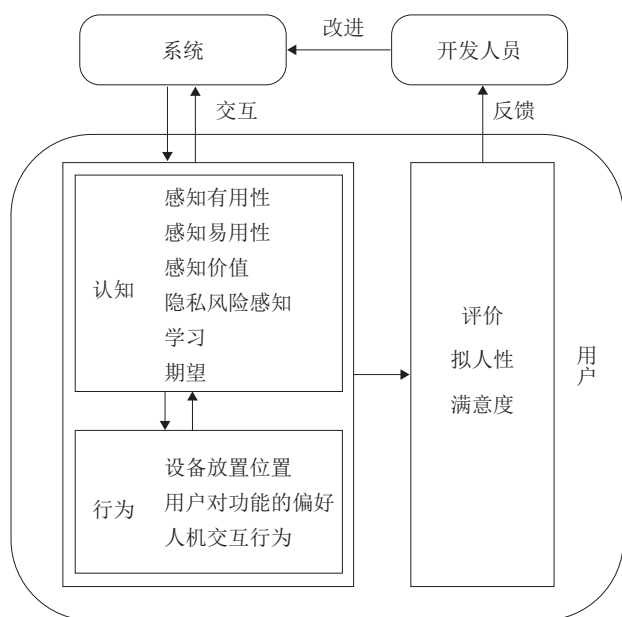


图1 智能语音助手用户研究的研究热点

用户与智能语音助手的交互最先作用于用户认知层面,比如易用性、有用性、价值和隐私风险感知等,这些感知可能来自用户与智能语音助手的交互过程,并影响用户后续的使用行为,比如用户在初次使用时感知到某个智能语音助手的易用性和有用性,这种感知可能会影响用户后续使用智能语音助手的频率,促进用户对智能语音助手功能的探索等。随着使用频率的增加,用户会产生思维与思想上的变化,如对智能语音助手的功能表现出更高的期望等。

用户对智能语音助手的认知以及与其的交互会影响用户评价,设计者可以从用户评价中收集用户意见,改进产品的功能与设计,这一行为又会影响用户的认知与行为,这对智能语音助手而言则是一个迭代和进步的过程。

就目前的文献数据来看,由于研究的设备多为Siri和Amazon Alexa,研究的使用场景集中于日常生活,用户研究涉及的智能语音助手功能比较宽泛。少部分文献就具体场景中的用户行为与交互展开研究,如多人闲聊、汽车驾驶、演讲辅导、阅读陪伴、工作反思和智能语音助手被辱骂等场景。学者多从具体场景或功能出发研究用户的某一类特定行为。

### 4.2 从智能语音助手用户研究中得到的实践启示

开展用户研究有助于了解用户需求与行为方式,帮助设计者更好地设计出有价值的技术产品。可以从性能提升、隐私保护、个性化服务3个角度对本文提出的4个方面的实践启示作进一步的解读。

在性能提升角度,智能语音助手改进优化的重点应聚焦于语音识别能力、响应能力和主动性水平:减轻环境噪声、用户口音等对语音识别的影响;通过提供适当的反馈(如为用户提供查询失败的依据或者改进查询效果的建议),给用户正确的指引以提高智能语音助手的响应能力;尝试打破被动依赖用户的固有模式,主动与用户建立联系,提供增值性服务,打造语音交互生态圈。

在隐私保护角度,语音是在人们日常生活中产生的很普遍且重要的一类数据,设计者应提高用户数据存储和被使用的透明度,告知用户查看、删除个人语音数据等方式,提供隐私保障条款等。此外,考虑到在公共场合使用智能语音助手容易暴露隐私的风险,可以为

用户提供一种“谨慎”模式<sup>[12]</sup>,在公共场所下保护用户的个人信息,从而提升用户对智能语音助手的信任度。

在个性化服务角度,设计者可以考虑根据不同的用户类型配置不同的系统文件,使智能语音助手实时学习与用户交互的内容,挖掘重复的使用模式以适应用户的习惯,从而逐渐具备主动学习和自主学习用户习惯的能力。允许用户自定义语音输出设置,如语速、音色和强度等,从而提升用户体验。此外,设计者应专注于提升智能语音助手的拟人化水平,采用更人性化的表达方式,从而获得更高水平的用户满意度。

## 5 结语

本文使用系统性综述的方法,对国际上的智能语音助手用户研究进行了深度分析,对已有研究中的智能语音助手用户研究设计进行了剖析,揭示了国际上智能语音助手用户研究的研究热点,为智能语音助手的设计和 optimization 提出了建议,可以为国内未来的智能语音助手用户研究提供参考和借鉴。

研究结果表明,智能语音助手用户研究为图书情报领域提供了全新的研究场景和研究机会,比如语义搜索、人机对话以及用户与智能语音助手交互时发生的协同信息检索、查询重构、查询优化和信息搜索停止行为等。同时,语音查询技术拓展了传统信息搜索研究的边界:不仅丰富了信息搜索的内涵与外延,还作为人工智能时代的新交互模式,衍生出很多新的方向。本文的局限在于只采集了6个代表性数据库中的英文文献,可能遗漏了个别相关文献或者非英文文献。

## 参考文献

- [1] 赵一鸣,朱奕蓉,吴林容. 智能语音助手的知识服务能力评价研究[J]. 图书与情报, 2019(4): 132-140.
- [2] 千家智能家居网. 调查显示:新冠疫情提升人们使用智能语音助手频率[EB/OL]. [2022-11-12]. [http://smarthome.qianjia.com/html/2020-05/09\\_365045.html](http://smarthome.qianjia.com/html/2020-05/09_365045.html).
- [3] 喻国明,杨名宜. 平台型智能媒介的机制构建与评估方法:以智能音箱为例[J]. 新疆师范大学学报(哲学社会科学版), 2019, 40(2): 120-126.
- [4] 凯特·克劳福德,弗拉丹·卓勒. 人工智能系统解剖:作为人工劳动力、数据和全球资源剖析图的亚马逊智能语音助手[J]. 新美术, 2020, 41(2): 97-111.
- [5] 孟凡坤,吴湘玲. 重新审视“智慧城市”:三个基本研究问题:基于英文文献系统性综述[J]. 公共管理与政策评论, 2022, 11(2): 148-168.
- [6] GRANT M J, BOOTH A. A typology of reviews: an analysis of 14 review types and associated methodologies [J]. Health Information and Libraries Journal, 2009, 26(2): 91-108.
- [7] FEINE J, GNEWUCH U, MORANA S, et al. A taxonomy of social cues for conversational agents [J]. International Journal of Human-Computer Studies, 2019, 132: 138-161.
- [8] MCTEAR M F. The rise of the conversational interface: a new kid on the block? [C] //International Workshop on Future and Emerging Trends in Language Technology. Lecture Notes in Computer Science, 2017, 10314: 38-49.
- [9] ASLAM W, SIDDIQUI D, ARIF I, et al. Chatbots in the frontline: drivers of acceptance [J/OL]. Kybernetes, 2022, ahead-of-print. <https://doi.org/10.1108/K-11-2021-1119>.
- [10] CORBETT E, WEBER A. What can I say? Addressing user experience challenges of a mobile voice user interface for accessibility [C] //Proceedings of the 18<sup>th</sup> International Conference on Human-Computer Interaction with Mobile Devices and Services, 2016: 72-82.
- [11] LUGER E, SELLEN A. “like having a really bad PA”: the gulf between user expectation and experience of conversational agents [C] //Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 2016: 5286-5297.
- [12] ABDOLRAHMANI A, KUBER R, BRANHAM S M. “Siri talks at you”: an empirical investigation of voice-activated personal assistant (VAPA) usage by individuals who are blind [C] // Proceedings of the 20<sup>th</sup> International ACM SIGACCESS Conference on Computers and Accessibility, 2018: 249-258.
- [13] KADYLA K T, BLOCKER K A, KOVAC C E, et al. Understanding the potential of digital home assistant devices for older adults through their initial perceptions and attitudes [J]. Gerontechnology, 2022, 21(1): 1-10.
- [14] PARNELL S I, KLEIN S H, GAISER F. Do we know and do we care? Algorithms and attitude towards conversational user interfaces: comparing chatbots and voice assistants [C] // Proceedings of the 4<sup>th</sup> Conference on Conversational User Interfaces, 2022: 1-6.
- [15] BALASURIYA S S, SITBON L, BAYOR A A, et al. Use of voice activated interfaces by people with intellectual disability [C] //Proceedings of the 30<sup>th</sup> Australian Conference

- on Computer-Human Interaction, 2018: 102-112.
- [16] NG Y L, LIN Z H. Exploring conversation topics in conversational artificial intelligence-based social mediated communities of practice [J]. *Computers in Human Behavior*, 2022, 134: 107326.
- [17] LEE S, RYU H, PARK B, et al. Using physiological recordings for studying user experience: case of conversational agent-equipped TV [J]. *International Journal of Human-Computer Interaction*, 2020, 36 (9) : 815-827.
- [18] CHUNG J, BLEICH M, WHEELER D C, et al. Attitudes and perceptions toward voice-operated smart speakers among low-income senior housing residents: comparison of pre- and post-installation surveys [J/OL]. *Gerontology and Geriatric Medicine*, 2021, 7. <https://doi.org/10.1177/23337214211005869>.
- [19] JANSON S P, DALLA VIA J, DALY R M, et al. Delivery of home-based exercise interventions in older adults facilitated by Amazon Alexa: a 12-week feasibility trial [J]. *The Journal of Nutrition, Health & Aging*, 2022, 26 (1) : 96-102.
- [20] SIEGERT I, BUSCH M, METZNER S, et al. Music-guided imagination and digital voice assistant-study design and first results on the application of voice assistants for music-guided stress reduction [C] // *International Conference on Human-Computer Interaction. Lecture Notes in Computer Science*, 2022, 13337: 347-362.
- [21] ALDAYEL S S. The effectiveness of using mobile interactive voice assistant applications in developing academic self-efficacy of Saudi university students during the COVID-19 pandemic [EB/OL]. [2022-11-12]. <https://digitalcommons.aaru.edu.jo/isl/vol11/iss6/36/>.
- [22] BUTEAU E, LEE J. Hey Alexa, why do we use voice assistants? the driving factors of voice assistant technology use [J]. *Communication Research Reports*, 2021, 38 (5) : 336-345.
- [23] MCCARTHY A, GASTER B R, LEGG P. Shouting Through Letterboxes: a study on attack susceptibility of voice assistants [C] // *2020 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)*, 2020: 1-8.
- [24] XU Y, WARSCHAUER M. Exploring young children's engagement in joint reading with a conversational agent [C] // *Proceedings of the Interaction Design and Children Conference*, 2020: 216-228.
- [25] DU Y, ZHANG K, RAMABADRAN S, et al. "Alexa, what is that sound?" a video analysis of child-agent communication from two Amazon alexa games [C] // *IDC' 21: Interaction Design and Children*, 2021: 513-520.
- [26] COGHLAN S, WAYCOTT J, NUI L, et al. Swipe a screen or say the word: older adults' preferences for information-seeking with touchscreen and voice-user interfaces [C] // *Proceedings of the 33<sup>rd</sup> Australian Conference on Human-Computer Interaction*, 2021: 130-143.
- [27] LOPATOVSKA I, RINK K, KNIGHT I, et al. Talk to me: exploring user interactions with the Amazon Alexa [J]. *Journal of Librarianship and Information Science*, 2019, 51 (4) : 984-997.
- [28] EASWARA M A, VU K P L. Privacy concerns for use of voice activated personal assistant in the public space [J]. *International Journal of Human-Computer Interaction*, 2015, 31 (4) : 307-335.
- [29] VÖLKE S T, MEINDL S, HUSSMANN H. Manipulating and evaluating levels of personality perceptions of voice assistants through enactment-based dialogue design [C] // *Proceedings of the 3<sup>rd</sup> Conference on Conversational User Interfaces*, 2021: 1-12.
- [30] BENETEAU E, RICHARDS O K, ZHANG M R, et al. Communication breakdowns between families and Alexa [C] // *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019: 1-13.
- [31] ZELLOU G, COHN M, FERENC S B. Age- and gender-related differences in speech alignment toward humans and voice-AI [J]. *Frontiers in Communication*, 2021, 5: 600361.
- [32] KUZMINYKH A, SUN J, GOVINDARAJU N, et al. Genie in the bottle: anthropomorphized perceptions of conversational agents [C] // *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020: 1-13.
- [33] DINH C M, PARK S S. Humanizing chatbots: the effect of fear from the COVID-19 pandemic [C] // *Advances in Digital Marketing and eCommerce. Springer Proceedings in Business and Economics*, 2022: 27-35.
- [34] COHN M, RAVEH E, PREDECK K, et al. Differences in gradient emotion perception: human vs. Alexa voices [C] // *Interspeech 2020, ISCA: ISCA*, 2020: 1818-1822.
- [35] SAMROSE S, ANBARASU K, JOSHI A, et al. Mitigating boredom using an empathetic conversational agent [C] // *Proceedings of the 20<sup>th</sup> ACM International Conference on Intelligent Virtual Agents*, 2020: 1-8.



- [36] PAL D, ARPNIKANONDT C, FUNILKUL S, et al. User experience with smart voice assistants: the accent perspective [C] // 2019 10<sup>th</sup> International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2019: 1-6.
- [37] SWOBODA D, BOASEN J, LÉGER P M, et al. Comparing the effectiveness of speech and physiological features in explaining emotional responses during voice user interface interactions [J]. Applied Sciences, 2022, 12 (3): 1269.
- [38] MYERS C M, PARDO L F L, ACOSTA-RUIZ A, et al. “Try, try, try again”: sequence analysis of user interaction data with a voice user interface [C] // Proceedings of the 3<sup>rd</sup> Conference on Conversational User Interfaces, 2021: 1-8.
- [39] WAIRAGKAR M, DE LIMA M R, HARRISON M, et al. Conversational artificial intelligence and affective social robot for monitoring health and well-being of people with dementia [J]. Alzheimer's & Dementia, 2021, 17: e053276.
- [40] BENTLEY F, LUVOGT C, SILVERMAN M, et al. Understanding the long-term use of smart speaker assistants [J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2018, (3): 1-24.
- [41] ESAU M, KRAUB V, LAWO D, et al. Losing its touch: understanding user perception of multimodal interaction and smart assistance [C] // DIS'22: Designing Interactive Systems Conference, 2022: 1288-1299.
- [42] KIM S, EUN J, SEERING J, et al. Moderator chatbot for deliberative discussion: effects of discussion structure and discussant facilitation [J]. Proceedings of the ACM on Human-Computer Interaction, 5 (CSCW1): 1-26.
- [43] VÖLKE S T, SCHOEDEL R, KAYA L L, et al. User perceptions of extraversion in chatbots after repeated use [C] // Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, 2022: 1-18.
- [44] AZMANDIAN M, ARROYO-PALACIOS J, OSMAN S. Guiding the behavior design of virtual assistants [C] // Proceedings of the 19<sup>th</sup> ACM International Conference on Intelligent Virtual Agents, 2019: 16-18.

## 作者简介

赵一鸣, 男, 博士, 教授, 研究方向: 信息组织与检索、用户信息行为。

陈宇, 女, 硕士研究生, 研究方向: 用户信息行为。

刘齐平, 女, 博士, 讲师, 通信作者, 研究方向: 用户信息行为、平台经济, E-mail: 1043531305@qq.com。

### User Study on Intelligent Voice Assistant: Theoretical Progress and Practical Implications

ZHAO YiMing<sup>1,2,3</sup> CHEN Yu<sup>2,4</sup> LIU QiPing<sup>5</sup>

( 1. Center for Studies of Information Resources, Wuhan University, Wuhan 430072, P. R. China; 2. School of Information Management, Wuhan University, Wuhan 430072, P. R. China; 3. Big Data Institute, Wuhan University, Wuhan 430072, P. R. China; 4. National Demonstration Center for Experimental Library and Information Science Education at Wuhan University, Wuhan 430072, P. R. China; 5. School of Information Management, Hubei University of Economics, Wuhan 430205, P. R. China )

Abstract: Intelligent voice assistant has become a research highlight in the world, but the findings on this topic in domestic journals are still very scarce, so it is urgent to systematically sort out the latest international research progress. In this paper, a systematic review method of “search-appraisal-synthesis-analysis” is used to conduct in-depth analysis of the English papers on user study in the context of intelligent voice assistant. By coding and analyzing the basic information, research design, research findings, and practical implications of the papers, the research progress in user cognition, user behavior, and user evaluation is revealed. This paper puts forward some ideas for the user study and product design of intelligent voice assistant.

Keywords: Intelligent Voice Assistant; User Study; Systematic Literature Review

(责任编辑: 雷雪)