

我国智能语音领域技术标准化研究

谢子萦* 王佳胜 黄继雄

(广东省标准化研究院)

摘要: 依据现行新一代人工智能标准体系建设指南规定,本文分别以国家、行业、团体发布的有效标准3个维度系统梳理了人工智能涉及智能语音领域技术标准化研究现状。从统计数据可以得出,我国智能语音领域的标准化研究已趋向成熟。但是仍存在标准更新滞后、标准覆盖范围有限、缺乏协同合作等突出问题,最后对智能语音领域相关标准化工作提出建议。

关键词: 人工智能, 标准化, 智能语音, 现状

DOI编码: 10.3969/j.issn.1674-5698.2024.12.017

Research on Standardization of Technologies in China's Intelligent Voice Industry

XIE Zi-ying* WANG Jia-sheng HUANG Ji-xiong

(Guangdong Institute of Standardization)

Abstract: According to the current guidelines for the construction of standards system of the next-generation artificial intelligence, the standardization research in the intelligent voice field related to artificial intelligence is systematically sorted out in the paper from three dimensions of valid national, sectoral, and association standards. From the statistical data, it can be concluded that the standardization research in the intelligent voice field of China has become mature. However, there are still prominent problems such as delayed standard updates, limited standard coverage, and lack of coordination and cooperation. Finally, suggestions are put forward for relevant standardization work in the intelligent voice field.

Keywords: artificial intelligence, standardization, intelligent voice, status

0 引言

智能语音技术起源于20世纪50年代至60年代的早期语音识别研究^[1],当时科学家们试图构建能够理解和响应人类语音的系统。早期工作主要集中在模板匹配和基于规则的方法上,但由于计算能力有限和语音数据的复杂性,进展较为缓

慢。进入21世纪,特别是在2000年后,随着深度学习和神经网络技术的兴起,智能语音技术实现了质的飞跃。深度神经网络模型被应用于语音特征提取、声学建模和语言模型构建等方面,显著提高了语音识别和语音合成的准确率和自然度^[2]。

随着移动互联网、物联网、云计算和大数据时代的到来,人机交互方式的需求发生了深刻变化。

作者简介: 谢子萦,通信作者,本科,助理工程师,研究方向为人工智能标准化。

王佳胜,博士,正高级工程师,研究方向为人工智能标准化。

黄继雄,硕士,高级工程师,研究方向为标准化。

用户对于便捷、高效、自然的交互体验有了更高的追求,这催生了智能语音技术在智能手机、智能家居、可穿戴设备、车载信息系统等各种场景中的广泛应用^[3]。随着智能语音技术在智能终端的不断进步和广泛应用,相关行业领域正在迎来显著增长,预计如:智能手机、智能家居、可穿戴设备这些领域的市场价值将实现上千亿元的大幅提升^[4]。

智能语音技术的研究背景是由科技进步、社会需求以及多学科交叉融合共同塑造的,因此智能语音领域标准体系中的各项标准都是相互联系不可分割的,需要坚持整体观以制定高效协同的智能语音技术标准^[5]。本文通过对智能语音领域当前已发布的国家、行业及团体标准以及正在研制的国家标准计划进行分析,得出智能语音技术标准化研究进展。基于综合分析结果,指出我国在推进智能语音技术标准化过程中面临的主要问题,并针对这些问题提出具体改进建议。

1 我国智能语音领域技术标准化现状

1.1 整体标准化现状

智能语音行业中国、家、行业、地方和团体各类标准的制定进展,很大程度上展示了该领域在国内的成长状态及政府、行业对其监管的程度。如:国家推荐性系列标准,涵盖了基础技术、服务质量、安全性等多个层面,旨在规范智能语音产品的功能、性能和安全性要求,推动国内市场的规范化和技术创新。我国智能语音行业已经逐步建立了较为完善的国家标准化体系,包含一系列国家标准、行业标准、地方标准和团体标准。智能语音标准不仅局限于家庭消费电子产品,还延伸到了教育、医疗、政务、电信、电商、金融等企业级应用领域。针对不同的应用场景,争取做到有对应的标准化要求,以确保智能语音技术在各行业健康有序发展,这不仅有效规范了市场行为,也为我国智能语音技术的自主创新和产业发展提供了强有力的支持。

1.2 国家标准

智能语音行业国家标准制定现状在很大程度上反映了该行业在国内的发展水平和政府监管力度。智能语音方向已发布的国家标准总计17项(见表1),均为推荐性标准,内容涵盖术语和定义、语义库要求、通用规范、测试方法、终端设备要求和

表1 智能语音现行国家标准文件及国家标准计划列表

序号	标准编号	标准名称	备注
1	GB/T 21024-2007	中文语音合成系统通用技术规范	推荐性标准
2	GB/T 21023-2007	中文语音识别系统通用技术规范	推荐性标准
3	GB/T 34083-2017	中文语音识别互联网服务接口规范	推荐性标准
4	GB/T 34145-2017	中文语音合成互联网服务接口规范	推荐性标准
5	GB/T 35312-2017	中文语音识别终端服务接口规范	推荐性标准
6	GB/T 41813.1-2022	信息技术 智能语音交互测试方法 第1部分:语音识别	推荐性标准
7	GB/T 41813.2-2022	信息技术 智能语音交互测试方法 第2部分:语义理解	推荐性标准
8	GB/T 36464.5-2018	信息技术 智能语音交互系统 第5部分:车载终端	推荐性标准
9	GB/T 36464.4-2018	信息技术 智能语音交互系统 第4部分:移动终端	推荐性标准
10	GB/T 36464.3-2018	信息技术 智能语音交互系统 第3部分:智能客服	推荐性标准
11	GB/T 36464.2-2018	信息技术 智能语音交互系统 第2部分:智能家居	推荐性标准
12	GB/T 36464.1-2020	信息技术 智能语音交互系统 第1部分:通用规范	推荐性标准
13	GB/T 5271.29-2006	信息技术 词汇 第29部分:人工智能 语音识别与合成	推荐性标准
14	GB/T 36339-2018	智能客服语义库技术要求	推荐性标准
15	GB/T 5271.29-2006	信息技术 词汇 第29部分:人工智能 语音识别与合成	推荐性标准
16	GB/T 44089-2024	信息技术 全双工语音交互系统通用技术要求	推荐性标准
17	GB/T 43969-2024	智能语音控制器通用安全技术要求	推荐性标准
18	20221472-T-607	智能家用电器的语音交互技术 第1部分:通用要求	
19	20242205-T-607	智能家用电器的语音交互技术 第2部分:测试方法	
20	20213581-T-339	道路车辆 免提通话和语音交互性能要求及试验方法	

具体场景应用要求等^[6]。在研的国家标准计划有3项(见表1),均为智能语音产品在应用场景中的规范要求,这表明国家对智能语音行业的监管力度逐步增大,对智能语音交互系统的要求日益严格,显示出国家对这一新兴技术领域的重视愈加提高。国家旨在通过制定和实施一系列标准,完善建设标准体系,确保行业中的产品和服务能够达到一定的质量要求,即通过有效标准化手段规范市场秩序,推动行业良性竞争。同时,随着相关国家标准的发布和修订,有助于用权威手段解决市场上存在的例如:测试方法具有差异性、服务接口不统一、应用场景不同等问题,从而增加公众信任度。此外,规范性文件也更加明确且具有针对性,有利于引导整个产业链向可持续的方向发展。总之,加强对智能语音技术的规范化、标准化,是利用国家标准推动该领域长期稳定发展的关键所在。

1.3 行业标准

行业标准方面,智能语音方向发布并现行标准9项(见表2),均为推荐性行业标准,规范了移动通信、物联网、智能家居、车载信息服务等多个重要领域,并且对这些领域的智能语音技术要求及测试方法进行了详细的规范。通过针对行业内特定应用场景设定统一的技术指标和评价体系,有利于提高整个行业的技术水平和服务质量,促进健康有序的发展。特别值得一提的是,在某些特定的应用场景中都是结合自身领域特点分别制定了更加细致化的行业标准来指导智能语音技术的应用。公安、金融领域考虑到信息安全和隐私保护的重要性,在这些领域发布的标准会更加注重数据加密传输、用户身份验证等方面的要求,以保证过程中

收集或存储的信息能够得到妥善处理。而在智能电视领域针对家庭多媒体这一场景,则更加强调用户体验优化、识别准确性、应答效率等功能实现。通过为不同的应用场景量身定制标准,不仅有利于解决行业内的共性问题,还能激发更多智能语音技术创新应用的出现。

1.4 团体标准

团体标准通常由行业内具有影响力的企业或技术领导者牵头制定,团体标准的制定和发布主要是为了应对快速发展的技术趋势以及市场中不断涌现的新需求,例如:针对新兴的语音识别算法、语音合成技术、多模态交互方式等制定相应的技术标准。相比国家和行业标准,团体标准的一个显著优势是制定流程相对灵活,由于不需要经过漫长的制定过程,因此能迅速地根据科技成果的最新发展调整内容,并及时反映到标准当中。目前智能语音领域发布并现行团体标准13项(见表3),主要以智能语音技术与产品评估方法为方向进行标准化制定。通过推动团体标准的发展和完善,有利于企业间通过一起制定标准交流最新的科技成果,形成良好的市场竞争环境。

2 我国智能语音领域技术标准化存在问题

2.1 标准更新滞后

通过对当前智能语音领域标准制定现状的研究,我们可以观察到,这一产业正逐步实现与不同行业的深度融合,尤其是在金融、医疗和安全监控等行业中。然而,在智能语音技术迅速发展的背景下,标准的制定往往难以跟上技术创新的速度。

表2 智能语音现行行业标准文件列表

序号	标准编号	标准名称	备注
1	YD/T 4157-2022	车载通信终端语音识别功能技术要求和测试方法	推荐性标准
2	SJ/T 11380-2008	自动声纹识别(说话人识别)技术规范	推荐性标准
3	GA/T 1590-2019	警务智能语音服务平台技术要求	推荐性标准
4	GA/T 1591-2019	警务智能语音服务平台接口规范	推荐性标准
5	SJ/T 11712-2018	智能电视语音识别 测试方法	推荐性标准
6	SJ/T 11713-2018	智能电视语音识别 通用技术要求	推荐性标准
7	JR/T 0164—2018	移动金融基于声纹识别的安全应用技术规范	推荐性标准
8	FYB/T 52045-2020	法院语音云平台建设及接入规范	推荐性标准
9	YD/T 4394.5-2023	自然语言处理技术及产品评估方法 第5部分:智能客服系统	推荐性标准

表3 智能语音现行团体标准文件列表

序号	标准编号	标准名称	团体名称
1	T/CCSA 375.1-2022	智能语音技术与产品评估方法 第1部分：中文语音合成基础服务	中国通信标准化协会
2	T/CCSA 375.2-2022	智能语音技术与产品评估方法 第2部分：中文语音识别基础服务	中国通信标准化协会
3	T/CCSA 375.3-2022	智能语音技术与产品评估方法 第3部分：声纹识别（说话人识别）基础服务	中国通信标准化协会
4	T/CCSA 375.4-2022	智能语音技术与产品评估方法 第4部分：全双工语音交互系统	中国通信标准化协会
5	T/GDAQI 132-2023	粤语智能语音服务系统技术要求	广东省质量检验协会
6	T/QGCML 1908-2023	智能语音识别转文字系统技术规范	全国城市工业品贸易中心联合会
7	T/QGCML 433-2022	AI智能语音客服系统通用技术要求	全国城市工业品贸易中心联合会
8	T/HEBQIA 050-2021	智能语音遥控器嵌入式软件技术要求	河北省质量信息协会
9	T/ZZB 2300-2021	蓝牙智能语音遥控器	浙江省质量协会
10	T/SIOT 606-2021	智能语音与视觉交互软件接口要求	上海市物联网行业协会
11	T/ZSA 16-2020	电梯智能语音交互系统技术规范	中关村标准化协会
12	T/CESA 1127-2020	基于人工智能语音交互的养老服务平台要求	中国电子工业标准化技术协会
13	T/SIOT 605-2020	智能语音核心组件接口规范	上海市物联网行业协会

这种滞后性不仅体现在新标准的发布时间上，还包括现行标准对新兴技术和应用场景的适应性不足。标准制定过程通常需要经过广泛的协商、测试和验证，以确保其科学性和实用性。这一过程涉及多个利益相关方，包括政府机构、行业组织、企业和研究机构等，各方之间达成共识往往需要较长时间，这进一步延长了标准从立项到正式发布的时间周期。

2.2 标准覆盖范围有限

由于智能语音技术快速发展且应用场景广泛，智能语音技术是一个跨学科领域，它的研究和发展集合了众多学科的知识^[7]，包括但不限于：数字信号处理（DSP），用于预处理和解析语音信号；人工智能（AI）与机器学习（ML），用于模式识别和自适应学习；自然语言处理（NLP），负责理解语音背后的含义；声学，研究声音产生和传播机制，优化语音识别性能；心理学与情感学，探索语音表达的情感维度，提升交互体验的真实性；统计学和数学，提供量化分析手段和模型构建的基础^[8]。现有标准往往难以全面覆盖所有新兴技术和特定使用场景，例如：虽然已有不少针对普通话等主要语言的语音识别与合成标准，但对于地方方言以及小语种的支持仍然不足。现有的智能

语音技术在处理方言时，由于缺乏统一的发音、词汇、语法和表达习惯等标准^[9]，导致语音识别引擎难以适应所有方言变体，识别准确率受限。

2.3 缺乏协同合作

智能语音领域团体标准文本公开程度较低，体现出企业间缺乏协同合作的问题。由于各企业间存在竞争关系，导致关键技术和研究成果难以共享，形成了技术壁垒；同时，不同企业、科研机构收集的语音数据集往往互不兼容，限制了模型训练的效果和通用性。此外，学术界与企业界缺乏有效沟通机制也阻碍了理论创新向实际应用转化。

3 我国人工智能关键领域技术标准化发展建议

3.1 提高标准制定时效性

为了加快智能语音领域技术标准的制定速度，可以构建一个多方参与、快速响应的标准工作组，该工作组应包括来自政府、企业、科研机构以及用户群体等多方面的代表。可以在制定标准时采用模块化和灵活的标准框架，以便于及时更新和适应新技术的发展，确保标准既能跟上创新步伐又能满足市场需求。

3.2 拓宽标准覆盖广度

为了应对智能语音领域标准覆盖范围有限的问题,可以加强行业内以及跨行业合作^[10],建立统一的智能语音技术标准框架,确保技术兼容与数据互通;加快新兴技术和应用场景的标准制定速度,紧跟技术发展步伐;在标准研究制定时对不同语言地区不同使用场景等因素进行全维度考量,吸纳语言学等领域的专业知识,提升标准的适用性和实用性,逐步完善智能语音技术标准体系,促进行业规范发展。

3.3 促进标准化工作协同

为解决智能语音领域标准化工作中缺乏协同合作的问题,应构建多方参与的协作平台,鼓励企业、研究机构及政府之间建立常态化的沟通机制,通过成立联合工作组或行业协会来协调各方利益,

共同制定和推广统一的技术标准与数据格式;同时,推动标准共享项目的发展,共享标准中的关键技术指标和测试方法等,加速创新成果的应用转化。此外,定期举办会议和技术论坛,加强同行间的交流与合作,以形成合力推进智能语音技术领域的标准化进程。

4 结 语

从本文可以看出,在制定相关智能语音交互技术标准的过程中,应结合最新研究成果和实践经验,选取科学、合理、可行的技术要素,提高标准制定时效性,拓宽标准覆盖广度,促进标准化工作协同,引导和推动智能语音应用行业的健康发展。

参考文献

- [1] 张筱兰,王保论.智能语音技术在教学中的应用研究[J].现代教育技术,2011,21(11):91-94+90.
- [2] 唐永军.基于深度学习的智能语音助手研究[J].现代信息技术,2021,5(12):75-79.
- [3] 许为.九论以用户为中心的设计:智能时代的“用户体验3.0”范式[J].应用心理学,2024,30(02):99-117.
- [4] 胡郁,严峻.智能语音交互技术及其标准化[J].信息技术与标准化,2015(04):14-17.
- [5] 雷静,王佳胜.基于关键要素的人工智能标准化研究[J].标准科学,2018(11):68-72.
- [6] 朱肖曼,申志伟,时文丰,等.我国人工智能关键领域技术标准研究[J].网络安全与数据治理,2023,42(09):65-71.
- [7] 陈鑫源.智能语音交互技术及其标准化[J].电声技术,2018,42(05):78-80.
- [8] 张宣,杨易臻.智能语音技术,为“声音”带来更多可能[N].新华日报,2023-11-15(011).
- [9] 郑晔,欧智坚,杨艇.闽南语智能交互引擎开发和应用研究[J].广播与电视技术,2022,49(11):124-127.
- [10] 许小春.关于智能语音识别技术的应用与发展研究[J].科技风,2022(25):1-3.