



中华人民共和国国家标准

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

Information technology—Intelligent speech interaction system—
Part 1: General specifications

2020-04-28 发布

2020-11-01 实施

国家市场监督管理总局 发布
国家标准化管理委员会

目 次

前言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 系统通用功能框架	4
5 语音交互界面要求	5
5.1 语音采集	5
5.2 语音播报	5
5.3 输入输出	5
5.4 环境噪声适应能力	5
6 数据资源要求	5
6.1 音频数据	5
6.2 文本数据	5
7 前端处理要求	6
7.1 语音唤醒	6
7.2 声源定位	6
7.3 声纹识别	6
7.4 语音增强	7
7.5 格式转换	7
7.6 重采样	7
8 语音处理要求	7
8.1 语音识别	7
8.2 语义理解	8
8.3 语音合成	8
8.4 端点检测	8
8.5 语音编解码	9
8.6 全双工交互	9
8.7 情感计算	9
9 服务接口要求	9
10 应用业务处理要求	9
附录 A(资料性附录) 部分参数及其计算方法	10
A.1 概述	10
A.2 拾音距离	10
A.3 语音交互成功率	10
A.4 语音唤醒	10

A.5	语音识别	11
A.6	语义理解	11
A.7	语音合成	12
A.8	语音质量	12
A.9	声纹识别率	13
A.10	语音编解压缩率	13
A.11	语音增强	13
A.12	声源定位	13
A.13	语音打断成功率	13
参考文献	15

前 言

GB/T 36464《信息技术 智能语音交互系统》拟分为以下几个部分：

- 第 1 部分：通用规范；
- 第 2 部分：智能家居；
- 第 3 部分：智能客服；
- 第 4 部分：移动终端；
- 第 5 部分：车载终端；

……

本部分为 GB/T 36464 的第 1 部分。

本部分按照 GB/T 1.1—2009 给出的规则起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本部分由全国信息技术标准化技术委员会(SAC/TC 28)提出并归口。

本部分起草单位：中国电子技术标准化研究院、科大讯飞股份有限公司、沈阳新松机器人自动化股份有限公司、深圳市优必选科技股份有限公司、中国科学院自动化研究所、上海智臻智能网络科技股份有限公司、中国盲人协会、中国电信集团有限公司、百度在线网络技术(北京)有限公司、中科讯飞互联(北京)信息科技有限公司、上海计算机软件技术开发中心、北京小米移动软件有限公司、西宁市大数据服务管理局、中汽研(天津)汽车工程研究院有限公司、西宁大数据有限公司、极限元(杭州)智能科技股份有限公司、中国移动通信有限公司研究院、厦门盈趣科技股份有限公司、深圳市腾讯计算机系统有限公司、重庆中科云从科技有限公司、中国医学科学院生物医学工程研究所、广东省标准化研究院、苏州思必驰信息科技有限公司、出门问问信息科技有限公司、哈尔滨亿时代数码科技开发有限公司、四川赛闯检测股份有限公司。

本部分主要起草人：董建、马万钟、徐洋、张锋、袁杰、陶建华、王海坤、叶雷鸣、李洪亮、李庆忠、杨震、吕彦锋、宋文林、穆亚敏、潘榕、蔡立志、朱亚军、周伟、孟宪明、孙巍、温正棋、鲍薇、杨本植、陈建成、戴小兰、翟尤、李军、蒲江波、王佳胜、龙梦竹、陈威钢、赵婕、冯丽。

信息技术 智能语音交互系统

第1部分:通用规范

1 范围

GB/T 36464 的本部分给出了智能语音交互系统通用功能框架,规定了语音交互界面、数据资源、前端处理、语音处理、服务接口、应用业务处理等功能单元要求。

本部分适用于智能语音交互系统的通用设计、开发、应用和维护。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件,仅注日期的版本适用于本文件。凡是不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

GB/T 11460 信息技术 汉字字型要求和检测方法

GB 18030 信息技术 中文编码字符集

GB/T 21024—2007 中文语音合成系统通用技术规范

GB/T 34083—2017 中文语音识别互联网服务接口规范

GB/T 34145—2017 中文语音合成互联网服务接口规范

SJ/T 11380—2008 自动声纹识别(说话人识别)技术规范

3 术语和定义

下列术语和定义适用于本文件。

3.1

语音交互 speech interaction

人类和功能单元之间通过语音进行的信息传递和交流活动。

[GB/T 36464.2—2018,定义 3.1]

3.2

语音交互系统 speech interaction system

由功能单元(或其组合)、数据资源等组成的能够实现与人类之间进行语音交互的系统。

[GB/T 36464.2—2018,定义 3.2]

3.3

智能语音交互系统 smart speech interaction system

以语音识别、语义理解、语音合成等全部或部分人工智能技术为基础,由智能软硬件组成,具备智能人机交互能力的语音交互系统。

3.4

人机交互 human machine interaction

人类和功能单元之间,为完成确定任务,以一定的交互方式进行的信息传递和交流活动。