



中华人民共和国国家标准

GB/T 46483—2025

信息技术 客服型虚拟数字人通用 技术要求

Information technology—General technical requirements for customer
service virtual digital human

2025-10-05 发布

2025-10-05 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 缩略语 2

5 参考框架 3

6 功能要求 4

 6.1 形象生成 4

 6.2 视觉交互 5

 6.3 语音交互 6

 6.4 情感交互 7

 6.5 形象驱动 8

 6.6 运营维护 9

7 性能要求 9

 7.1 数字人形象 9

 7.2 视觉交互 10

 7.3 语音交互 10

 7.4 情感交互 11

参考文献 12

前 言

本文件按照 GB/T 1.1—2020《标准化工作导则 第 1 部分：标准化文件的结构和起草规则》的规定起草。

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专利的责任。

本文件由全国信息技术标准化技术委员会(SAC/TC 28)提出并归口。

本文件起草单位：北京市商汤科技开发有限公司、中国电子技术标准化研究院、深圳赛西信息技术有限公司、厦门赛西科技发展有限责任公司、北京理工大学、北京津发科技股份有限公司、北京赢动实名认证服务有限公司、中移(杭州)信息技术有限公司、广州趣丸网络科技有限公司、科大讯飞股份有限公司、鹏城实验室、北京海纳数聚科技有限公司、咪咕文化科技有限公司、南方电网电力科技股份有限公司、中仪英斯泰克科技有限公司、深圳众投互联信息技术有限公司、北京国际云转播科技有限公司、成都市凯瑞医疗科技有限公司、浙江云澎科技有限公司、三人行数据(广东)股份公司、世优(北京)科技股份有限公司、杭州一知智能科技有限公司、雷鸟创新技术(深圳)有限公司、睿魔智能科技(深圳)有限公司、广东天网智城科技有限公司、联通在线信息科技有限公司、道有道科技集团股份公司、深圳媲美科技有限公司、广东粤电信息科技有限公司、广州开发区粤电新能源有限公司、北京北纬三十度网络科技有限公司。

本文件主要起草人：徐柏琦、潘榕、李亚健、李斌、许亲亲、吴庚、王和俊、董桂官、孙齐锋、耿一丹、史培宁、曾洁琪、梁继允、冯南飞、柯绍棠、翁冬冬、赵起超、吕克勤、陈大庆、任少峰、王清菊、何山、郭洁、张旭、毕蕾、石磊、张顺四、郑培文、贺钦、唐传广、张美静、王秋霖、刘浙东、阮锐师、纪智辉、陈哲乾、欧阳琼林、张明、何社超、高德扬、周建修、张纾翔、陈禹明、陈科、邓先才、张君杰、牛龙飞、梁良、邵臻霖、张云澎。

信息技术 客服型虚拟数字人通用 技术要求

1 范围

本文件规定了客服型虚拟数字人系统参考框架、功能和性能的要求。
本文件适用于客服型虚拟数字人系统的设计、开发、测试、应用和维护等。
注：在本文件中，为方便表述，“虚拟数字人”简称为“数字人”，该简称仅用于文件内部，不作为标准术语。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；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 36464.1—2020 信息技术 智能语音交互系统 第1部分：通用规范
GB/T 34083—2017 中文语音识别互联网服务接口规范

3 术语和定义

下列术语和定义适用于本文件。

3.1

虚拟数字人 **virtual digital human**

通过计算机图形学、计算机视觉和语音交互及人工智能生成内容（AIGC）等技术，进行形象、声音、动作等模型训练后，借助真人或计算驱动、在多模态输出设备呈现的虚拟人物。

3.2

客服型虚拟数字人 **customer service virtual digital human**

面向特定行业应用，具有开展客户服务功能的虚拟数字人。

3.3

2D 虚拟数字人 **2D virtual digital human**

以二维平面形式呈现的数字人。

注：2D 虚拟数字人包括 2D 拟真人、2D 卡通等不同风格。2D 拟真人数字人形象通常基于真人形象采集或 AIGC 生成的人物图片、视频，借助人工智能技术训练生成；2D 卡通数字人形象通常基于手绘或图像处理软件等方式绘制而成。

3.4

3D 虚拟数字人 **3D virtual digital human**

以三维立体形式呈现的数字人。

注：3D 虚拟数字人包括 3D 超写实、3D 拟真人、3D 卡通等不同风格。3D 虚拟数字人形象通常基于 3D 建模软件或 AIGC 算法创建，通过复杂的建模、材质贴图、绑定等操作生成，或通过神经辐射场（NeRF, Neural Radiance Field）、3D 高斯泼溅（3DGS, 3D Gaussian Splatting）等算法生成。

3.5

语音唤醒 **speech wakeup; voice trigger**

处于音频流监听状态的语音交互系统，在检测到特定的特征或事件出现后，切换到命令字识别、连续语音识别等其他处理状态的过程。