

团 体 标 准

T/ISC XXX—XXXX

移动终端辅助通话应用能力要求 与评价指南

Mobile Terminal Call Assistance Application Capability Requirements and
Evaluation Guide

在提交反馈意见时，请将您知道的相关专利与支持性文件一并附上。

（征求意见稿）

XXXX—XX—XX 发布

XXXX—XX—XX 实施

中 国 互 联 网 协 会 发 布

目 次

前言	3
1 范围	5
2 规范性引用文件	5
3 术语和定义	5
3.1 辅助通话应用 auxiliary call application	5
3.2 通话转写 call transcription	5
3.3 通话摘要 call summary	5
3.4 智能代接 intelligent call answering	5
3.5 实时通话翻译 real-time call translation	6
3.6 语音增强 speech enhancement	6
4 符号和缩略语	6
5 辅助通话应用能力框架	6
5.1 概述	6
5.2 具体要求	6
5.3 能力建设	7
5.4 用户控制	7
5.5 技术呈现	7
6 基础能力	7
6.1 通话信息采集	7
6.2 语音识别	7
6.3 辅助处理与输出	8
6.4 噪声抑制与人声增强	8
7 核心能力	8
7.1 通话转写能力	8
7.2 通话摘要能力	9
7.3 智能代接能力	9
7.4 实时通话翻译能力	10
7.5 语音增强能力	11
8 支撑保障能力	11
8.1 通话体验与终端适配能力	11
8.1.1 识别与转写体验	11
8.1.2 响应与运行体验	11
8.1.3 终端与网络适配	11
8.2 无障碍与适老化能力	12
8.2.1 视障辅助	12
8.2.2 操作无障碍与适老化	12
8.3 安全与隐私保护能力	12

8.3.1 数据保护	12
8.3.2 用户授权与隐私透明	12
8.3.3 内容与 AI 使用安全	13
9 应用能力评价建议	13
9.1 评价原则	13
9.2 评价参考	13
9.3 典型应用场景	14
9.4 实践反馈与持续改进	14

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。

本文件由中国互联网协会提出并归口。

本文件起草单位：

本文件主要起草人：

本文件及其所代替文件的历次版本发布情况为：

——

T/ISC XXX—XXXX

移动终端辅助通话应用具体要求与评价指南

1 范围

本文件规定了面向移动终端辅助通话应用的发展与应用实践，提出移动终端辅助通话应用的能力框架、典型具体要求、应用场景及评价参考，为相关企业和机构开展产品设计、技术开发、应用推广及用户选择提供参考。

本文件适用于：

- a) 从事移动终端辅助通话应用研发、设计和运营的企业或机构，用于产品能力规划与优化；
- b) 移动终端、通信服务、人工智能等相关企业和机构，用于辅助通话应用能力建设和应用实践参考；
- c) 用户单位及相关组织，用于了解和比较移动终端辅助通话应用的能力覆盖情况。

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 41867-2022 信息技术 人工智能术语

YD/T 3329-2018 移动通信终端无障碍技术要求

3 术语和定义

GB/T 41867-2022 及 YD/T 3329-2018 界定的和下列术语和定义适用于本文件。

3.1 辅助通话应用 auxiliary call application

运行于智能手机、平板等移动终端，依托语音识别、自然语言处理、语音合成等技术，提供通话转写、通话摘要、智能代接、实时翻译、语音增强等通话辅助功能的应用程序（含系统内置功能及第三方应用）。

3.2 通话转写 call transcription

将通话过程中的语音内容实时转换为文字，区分说话人并呈现的功能。

3.3 通话摘要 call summary

基于通话转写等内容，自动提取时间、地点、任务、联系人等关键信息，形成结构化或摘要化内容，帮助用户快速理解通话要点的功能。

3.4 智能代接 intelligent call answering

用户无法或不便直接接听来电时，应用按照用户预设规则自动响应来电，并可通过语音或文字与来电方进行辅助交互、记录通话内容的功能。

3.5 实时通话翻译 real-time call translation

通话过程中呈现多语言语音互译，支持跨语言沟通的功能。

3.6 语音增强 speech enhancement

通过 AI 算法降噪、放大目标人声，提升嘈杂环境下通话清晰度的功能。

4 符号和缩略语

下列符号和缩略语适用于本文件。

AI 人工智能 (Artificial Intelligence)

5 辅助通话应用能力框架

5.1 概述

移动终端辅助通话应用宜围绕“基础能力层、核心能力层、支撑保障层”形成能力体系，其中，基础能力层为辅助通话应用提供通话信息获取、识别、呈现和增强等基础支撑；核心能力层面面向通话内容理解和智能交互形成主要应用能力；支撑保障层从通话体验、终端适配、无障碍与适老化、安全与隐私等方面保障能力可靠、便捷、安全地使用。不同产品可结合目标用户、终端形态、技术条件和应用场景选择相应能力，不要求所有产品具备全部能力。

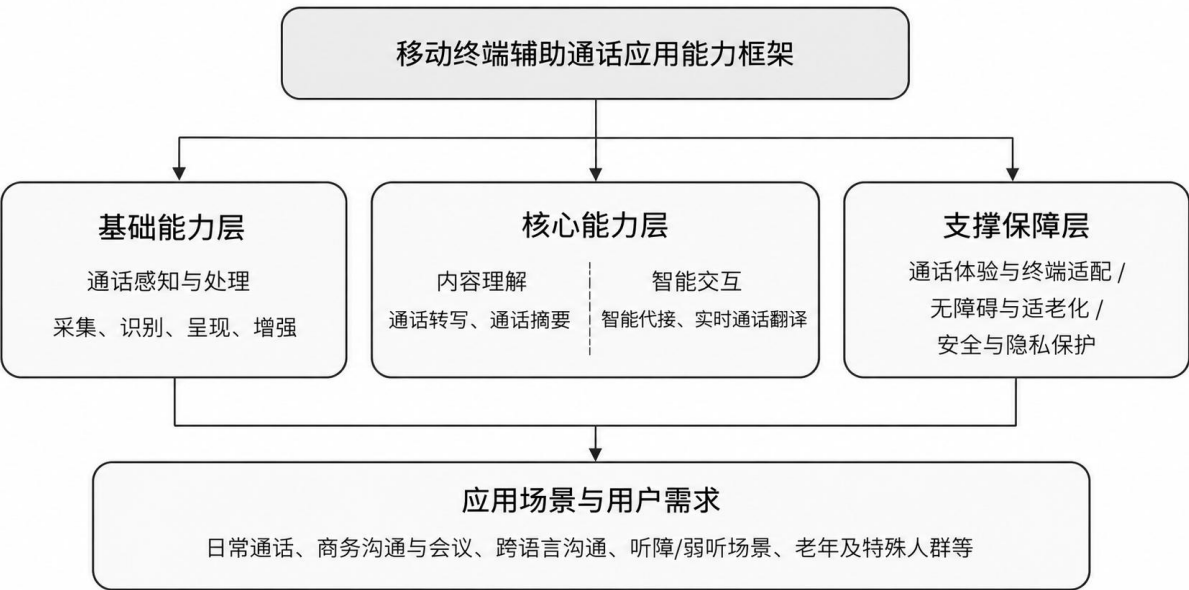


图 1 辅助通话应用能力框架

5.2 具体要求

- a) 基础能力：对通话语音进行采集、识别、呈现和增强的通话感知与处理能力，用于获取通话信息、识别通话内容、完成辅助处理和输出，并提升通话清晰度，为核心能力提供基础支撑；
- b) 核心能力：面向用户对通话信息理解和通话过程辅助的主要需求，分为内容理解和智能交互。其中，内容理解能力用于将通话信息转化为可读取、可理解和可利用的信息，主要包括

通话转写能力和通话摘要能力；智能交互能力用于根据通话上下文和用户授权参与通话过程，主要包括智能代接能力和实时通话翻译能力；

- c) 服务保障能力：包括通话体验与终端适配能力、无障碍与适老化能力、安全与隐私保护能力，用于保障核心功能在不同终端、网络、用户和使用条件下的可用性、可理解性和安全性。

5.3 能力建设

- a) 辅助通话应用能力建设宜以实际用户需求和典型通话场景为基础，优先保障核心功能的可理解性、稳定性和可控性，再逐步增加智能化功能；
- b) 产品设计可依据目标用户和典型应用场景，对基础能力、核心能力和支撑保障能力进行组合。能力组合宜与用户实际需求相匹配，避免功能堆叠，并明确主要能力范围、使用方式、自动化状态及用户控制方式。

5.4 用户控制

- a) 对录音、自动接听、自动回复等可能影响用户权益的功能，宜提供清晰提示、明确状态以及便捷的开启、关闭和设置入口。
- b) 对转写、摘要、翻译及 AI 生成回复等智能化结果，宜提供必要的核验、修改和纠错机制；涉及重要信息时，不宜将模型生成结果直接作为唯一依据。
- c) 产品说明、界面提示或帮助信息宜清晰说明主要辅助能力的功能范围、使用方式及可能存在的局限。

5.5 技术呈现

- a) 根据终端算力、网络条件和应用场景合理选择端侧、云侧或端云协同的技术路线，并关注弱网、噪声、口音、多方通话等实际环境下的连续使用体验；
- b) 对于转写、摘要、翻译及 AI 生成回复等内容，提供必要的核验、修改和纠错机制，涉及重要信息时，不宜将模型生成结果直接作为唯一依据。

6 基础能力

通话感知与处理能力是辅助通话应用开展内容理解和智能交互的基础。产品宜根据具体应用需求形成与通话场景相适应的采集、识别、呈现和增强能力。

6.1 通话信息采集

表 1 通话信息采集能力参考

能力维度	具体要求	评价参考
通话信息采集	能够根据辅助通话功能需要获取通话语音等相关信息，并保持采集过程连续。	关注采集是否与功能需求相匹配，以及通话过程中的连续性。

6.2 语音识别

表 2 语音识别能力参考

能力维度	具体要求	评价参考
语音识别	能够将通话语音转换为可供后续处理和呈现的信息，并支持必要的说话人区分。	关注识别结果的连续性、可理解性和说话人区分情况。

6.3 辅助处理与输出

表 3 辅助处理与输出能力参考

能力维度	具体要求	评价参考
辅助处理与输出	能够依据用户需求和应用场景对通话信息进行处理、组织和输出，为转写、摘要、代接、翻译等能力提供支撑。	关注辅助处理结果是否清晰、连续，并与目标功能相匹配。

6.4 噪声抑制与人声增强

表 4 噪声抑制与人声增强能力参考

能力维度	具体要求	评价参考
噪声抑制与人声增强	能够降低环境噪声及非目标声音对通话的干扰，并增强目标语音的清晰度和可辨识度。	关注噪声抑制效果、目标语音清晰度以及声音自然度。

7 核心能力

7.1 通话转写能力

7.1.1 实时转写

表 5 实时转写能力参考

能力维度	具体要求	评价参考
环境适应	能够适应不同通话环境，保持转写连续性和基本可理解性。	关注不同使用环境下转写的连续性和可理解性。
实时同步	转写内容与通话过程保持合理同步，并及时呈现。	关注转写结果与通话过程的同步性和延迟体验。
说话人区分	能够对通话参与者进行清晰、稳定的说话人区分。	关注说话人标识是否清晰、稳定。

7.1.2 转写内容管理

表 6 转写内容管理实践参考

能力维度	具体要求	评价参考
内容检索	支持按时间、联系人、关键词等维度检索，可结合语义理解提升内容查找效率。	关注检索方式是否便捷、结果是否具有相关性和可理解性。
内容查看与整理	支持对已保存的转写内容进行查看、分类、标记或必要的编辑整理。	关注内容呈现是否清晰，整理操作是否便捷。
内容删除	支持用户删除不再需要的通话记录及相关转写内容，并明确删除范围和结果。	关注用户是否能够清晰、便捷地管理和删除相关内容。
内容导出与分享	支持以适当方式导出或分享通话转写内容。	关注导出、分享方式是否符合使用需求，并注意相关数据保护。

7.2 通话摘要能力

7.2.1 摘要生成

表 7 摘要生成能力参考

能力维度	具体要求	评价参考
摘要生成	在通话结束后及时提供摘要，并根据通话长度和复杂程度优化生成体验。	关注摘要及时性、完整性、可读性。
关键信息提取	能够根据通话内容提取时间、地点、任务、联系人等关键信息，并以便于理解的方式呈现。	关注关键信息的完整性、准确性和可理解性。

7.2.2 摘要应用

表 8 摘要应用实践参考

能力维度	具体要求	评价参考
摘要分享与同步	摘要可通过适当方式分享或同步至日历、待办等相关功能。	关注分享和同步是否便捷，关键信息传递是否完整。
摘要编辑与修正	支持用户对摘要结果进行必要的查看、修改和纠错。	关注修改和纠错是否方便，结果是否易于理解。

7.3 智能代接能力

7.3.1 自动应答

表 9 自动应答能力参考

能力维度	具体要求	评价参考
自动响应	用户无法或不便直接接听来电时，可按照用户预设规则进行自动响应。	关注自动响应是否与用户设置一致，功能状态是否清晰。

7.3.2 多轮交互

表 10 自动应答能力参考

能力维度	具体要求	评价参考
多轮智能交互	可结合上下文进行多轮交互，设置明确的功能边界，对无法可靠处理的请求及时转交用户或结束交互。	关注交互是否连贯、是否能够理解基本意图，以及异常情况下是否有退出或人工接管机制。

7.3.3 代接记录管理

表 11 代接记录管理实践参考

能力维度	具体要求	评价参考
代接内容留存	完成代接后，可支持文字或录音等形式的记录留存，并提供相应查看方式。	关注代接记录是否清晰可查，以及用户对记录的管理是否便捷。

7.4 实时通话翻译能力

7.4.1 多语言支持

表 12 多语言支持实践参考

能力维度	具体要求	评价参考
语种覆盖	根据目标用户和应用场景支持相应语种之间的双向翻译，并逐步扩展语种覆盖范围。	关注目标场景下语种覆盖的适用性及语言切换的便利性。

7.4.2 翻译质量与交互体验

表 13 翻译质量与交互体验参考

能力维度	具体要求	评价参考
翻译内容呈现	保持语音、文字与译文之间的合理对应关系，并减少明显等待和重复操作。	关注翻译过程是否自然、连续，用户是否容易理解当前交互状态。
结果核验与修正	支持用户对翻译结果进行必要的查看、核验和修正。	关注用户对翻译结果进行核验和修正是否方便。

7.5 语音增强能力

表 14 语音增强实践参考

能力维度	具体要求	评价参考
噪声抑制与人声增强	能够降低环境噪声及非目标声音对通话的干扰，并增强目标语音的清晰度和可辨识度。	关注噪声抑制效果、目标人声清晰度以及声音自然度。

8 支撑保障能力

8.1 通话体验与终端适配能力

8.1.1 识别与转写体验

表 15 识别与转写实践参考

能力维度	具体要求	评价参考
方言及混合语言	可根据目标用户逐步支持常见方言、口语化表达以及中外语言混合使用。	关注产品定位与目标用户需求下的覆盖程度。

8.1.2 响应与运行体验

表 16 响应与运行实践参考

能力维度	具体要求	评价参考
长时间运行	关注长时间通话中的稳定性、资源占用和异常恢复能力。	关注连续使用过程中是否出现明显中断、闪退或功能失效。

8.1.3 终端与网络适配

表 17 终端与网络适配实践参考

能力维度	具体要求	评价参考
网络适配	针对 4G、5G、Wi-Fi 及弱网等环境优化，对于关键能力可探索端侧处理或端云协同。	关注不同网络条件下核心功能的连续性，并向用户明确适用条件。

8.2 无障碍与适老化能力

8.2.1 视障辅助

表 18 视障辅助实践参考

能力维度	具体要求	评价参考
界面显示与辅助呈现	支持字体、界面等适当调整，并通过清晰的文字、图标或文字信息呈现通话辅助结果。	关注不同显示设置下的信息可读性和操作便利性。

8.2.2 操作无障碍与适老化

表 19 操作无障碍与适老化实践参考

能力维度	具体要求	评价参考
适老化交互	采用清晰的文字、图标和提示，减少复杂层级，并对自动化和 AI 功能提供易理解的说明。	关注老年用户能否理解功能状态并自主完成关键操作。

8.3 安全与隐私保护能力

8.3.1 数据保护

表 20 数据保护实践参考

能力维度	具体要求	评价参考
数据查看与控制	提供通话记录、转写内容和摘要等数据的查看、导出、删除等管理能力，并支持用户对相关功能进行控制。	关注用户是否能够清晰、便捷地管理和控制相关数据。

8.3.2 用户授权与隐私透明

表 21 用户授权与隐私透明实践参考

能力维度	具体要求	评价参考
授权与隐私透明	对录音、自动接听、自动回复等涉及用户权益的功能，以清晰方式说明用途、数据处理方式及用户控制方式，并在必要时取得用户授权。	关注授权提示是否清晰易懂、隐私信息是否易于获取，以及用户能否自主控制相关功能。

8.3.3 内容与 AI 使用安全

表 22 内容与 AI 使用安全实践参考

能力维度	具体要求	评价参考
AI 生成内容提示	通话摘要、翻译、自动回复等由 AI 生成的内容宜通过适当方式提示用户，并提供必要的核验和修改途径。	关注用户能否识别 AI 生成内容，以及对结果进行核验、修改和纠错是否方便。

9 应用能力评价建议

9.1 评价原则

移动终端辅助通话应用能力评价宜采用场景化、综合性和参考性方式开展，重点关注产品能力覆盖、功能表现、用户体验与可控性、适配与可用性、无障碍与适老化、安全与隐私等维度，不以本文件作为产品准入、强制检测或合格判定的依据。

评价可结合产品资料、功能体验、典型场景观察及用户反馈等信息进行综合判断，以形成对能力适用性和用户体验的整体认识。

表 23 应用能力评价维度

评价维度	关注内容	参考方式
能力覆盖	核心通话辅助能力是否覆盖目标用户和典型场景。	结合产品说明、功能体验和应用场景观察。
功能表现	转写、摘要、代接、翻译、语音增强等能力的及时性、连续性和可理解性。	结合功能体验和典型场景进行综合判断。
用户体验	操作便捷性、状态可理解性、用户可控性及异常情况下的退出或人工接管能力。	关注实际使用过程中的交互表现。
适配与可用性	不同终端、网络及语言环境下的适用性。	结合产品定位和典型使用条件观察。
无障碍与适老化	辅助信息呈现、操作便利性以及重点功能的可理解性。	结合目标用户需求和应用场景观察。
安全与隐私	授权、数据管理、隐私透明和 AI 生成内容提示等保障能力。	结合产品说明、界面提示和功能体验综合判断。

9.2 评价参考

评价参考主要用于说明评价时关注的内容和信息来源，不规定统一的测试方法、测试环境、测试步骤或合格判定规则。评价时宜根据产品定位和目标用户选择与实际应用相关的能力维度进行综合观察。

表 24 应用能力评价参考

评价维度	关注内容	参考方式
能力覆盖	核心通话辅助能力是否覆盖目标用户和典型场景。	结合产品说明、功能体验和应用场景观察。
功能表现	转写、摘要、代接、翻译、语音增强等能力的及时性、连续性和可理解性。	结合功能体验和典型场景进行综合判断。
用户体验	操作便捷性、状态可理解性、用户可控性及异常情况下的退出或人工接管能力。	关注实际使用过程中的交互表现。

	的退出或人工接管能力。	
适配与可用性	不同终端、网络及语言环境下的适用性。	结合产品定位和典型使用条件观察。
无障碍与适老化	辅助信息呈现、操作便利性以及重点功能的可理解性。	结合目标用户需求和应用场景观察。
安全与隐私	授权、数据管理、隐私透明和 AI 生成内容提示等保障能力。	结合产品说明、界面提示和功能体验综合判断。

9.3 典型应用场景

表 25 典型应用场景

典型场景	可重点关注的能力
日常通话	通话转写、通话摘要、信息管理、语音增强等。
商务及会议通话	通话转写、摘要、关键信息提取、内容检索与分享等。
跨语言沟通	实时通话翻译、翻译结果呈现、核验与修正等。
听障及弱听用户	实时转写、字幕及文字辅助、清晰提示等。
老年用户及其他特殊用户	清晰提示、简化操作、辅助交互、自动化功能可控性等。

9.4 实践反馈与持续改进

企业和相关机构可结合产品实际应用情况开展用户反馈、场景观察和版本复盘，持续完善能力。对于快速演进的新技术和新场景，可通过实践案例不断积累经验，并适时更新本指南相关内容。