

豆包式实时通话 App 设计规范

ASR · LLM · TTS 全链路流式交互与界面标准

1. 产品背景与目标

本规范旨在定义“豆包式”实时通话功能的完整交互与视觉标准。核心技术路径遵循 **ASR（语音转文字）→ LLM（大模型推理）→ TTS（文字转语音）** 的全流式处理架构。设计目标是在移动端（iOS/Android）提供低延迟、高容错、自然流畅的人机对话体验。

关键成功指标 (KPI)

<500ms

首包响应延迟

<5%

词错误率 (WER)

99.9%

会话稳定性

1.1 设计范围与核心流程

本设计覆盖从应用启动到通话结束的全生命周期，重点解决网络抖动、模型幻觉及语音打断（Barge-in）场景下的交互一致性。