

即时通信 IM

AI 智能语音



腾讯云

【 版权声明 】

©2013–2026 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

【 商标声明 】



及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【 服务声明 】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【 联系我们 】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或 95716。

文档目录

- AI 智能语音
 - 功能介绍
 - 开通服务

AI 智能语音 功能介绍

最近更新时间：2026-08-11 16:55:35

方案概述

即时通信 IM AI 智能语音服务，基于腾讯领先的 AI 语音识别引擎，将**语音转文字（ASR）**、**文本转语音（TTS）**及**翻译**的全链路处理能力嵌入 IM SDK，支持多种场景灵活使用。帮助开发者在自己的应用内快速构建微信、元宝同款的“AI 语音输入”交互体验。其核心优势体现在以下三个维度：

- 通过 ASR 语音识别引擎，可高效将语音转化为文本，广泛应用于聊天语音输入、语音消息识别等场景。
- 在语音转文字的基础上，翻译能力可实现跨语言内容的转写与沟通，让用户用母语说话，对方直接看到翻译后的文字。
- TTS 文字转语音技术，可将文本转化为自然流畅的语音输出，轻松构建语音播报、AI 变声、个性化语音消息等多样化应用。

功能优势

AI 语音识别精准高效

依托腾讯在 AI 能力上的深厚积累，引入大模型语音识别引擎，在含噪声、口音混杂等复杂音频环境下，识别准确率显著提升。支持中文普通话、英语、粤语、多地方言及多种小语种的精准识别，覆盖主流语言沟通需求。

毫秒级实时转写

语音输入的实时识别延迟可控制在毫秒级，实现“边说边出文字”的流畅体验。超高转写速率确保用户在口述过程中，文字内容同步输出，不打断表达节奏。

智能语义理解

AI 语音交互不仅能“听见”，更能“听懂”。系统智能理解上下文语义，自动断句、智能添加标点符号，无需用户二次编辑修改。

应用场景

微信同款的“语音输入”交互体验

从元宝、豆包等 AI 应用带火语音交互，到微信、企微强化“AI 语音输入”能力，语音正在成为移动端输入的主流方式之一。用户只需按住说话，语音即可被实时转写为文字并填入输入框，点击即可发送消息。相比传统的文字输入方式，AI 语音输入具备更高的输入效率、更低的操作成本和更自然的交互体验，是 IM 应用输入交互的全新升级方向。

腾讯云 IM AI 语音交互提供多种主流 UI 交互模式，旨在帮助开发者将这一国民级交互体验快速融入自己的 IM 应用。

交互模式	说明
长按输入框语音输入	长按输入框后进入语音录入状态，松手后自动将识别文字填入输入框。
点击语音转文字按钮	点击输入框右侧语音转文字图标，进入语音输入模式。
按住说话 + 滑动转文字	按住说话后，滑动手指至“转文字”按钮区域松手，实现语音转文字发送。

以上交互模式延续用户在微信及企微中的使用习惯，降低功能学习成本，提升用户体验。

其他场景推荐

应用场景	使用方式与价值
跨语言沟通	语音输入结合翻译能力，用户用自己的语言说话，应用实时转写后直接翻译成对方熟悉的语言，以原文和译文对照形式呈现在消息气泡中。无需在翻译工具和聊天窗口间反复切换，让跨语言聊天像母语沟通一样自然。
企业办公	口语化表达经 ASR 后，进一步通过 AI 修饰，自动转为更规范的书面文本。系统在保留原意的基础上，去除重复、停顿、口头禅，改写为更清晰正式、适合发送给同事或客户的文字。
AI 机器人对话	用户通过语音与 AI 机器人自然交互，语音实时转写为文字发送给机器人；机器人回复的文本通过 TTS 合成为语音播放，实现全链路语音交互体验，让用户像打电话一样与 AI 自由对话。
有声内容与语音播报	通过 TTS 将文本内容转换为自然语音，适用于消息朗读、通知播报、无障碍阅读等场景。支持多音色切换和声音克隆，满足个性化语音输出需求。

能力全览

语音转文本（ASR）

语音消息转文字

特性及支持语种见下表。

说明	语种
该场景将已发送出去的语音消息转文字	可支持中文普通话、中英混、方言和小语种，包括：zh（中文）、en（英语）、yue（粤语）、ar（阿拉伯语）、de（德语）、fr（法语）、es（西班牙语）、pt（葡萄牙语）、id（印尼语）、it（意大利语）、ko（韩语）、ru（俄语）、th（泰语）、vi（越南语）、ja（日语）、tr（土耳其语）、hi（印地语）、ms（马来语）、nl（荷兰语）、sv（瑞典语）、da（丹麦语）、fi（芬兰语）、pl（波兰语）、cs（捷克语）、fil（菲律宾语）、fa（波斯语）、el（希腊语）、ro（罗马尼亚语）、hu（匈牙利语）、mk（马其顿语）。

说明：

- 通过含 UI 的场景方案接入，详见 [语音转文字](#)，或者通过 SDK 客户端接入 ASR 功能，详见 [语音消息转文字](#)。
- 上文表格中列出的功能配合 IM 使用时，亦可使用 REST API 方式调用，完整的参数说明请参阅 [语音转文本（SpeechToText）](#)。
- 如果您有更多其他语言需求，请 [联系我们](#)。

录音转文字

特性及支持语种见下表。上传 60 秒以内的语音，同步返回识别结果

说明	语种
该场景可支持将还未发出的音频识别为文字	可支持中文普通话、中英混、方言和小语种，包括：zh（中文）、en（英语）、yue（粤语）、ar（阿拉伯语）、de（德语）、fr（法语）、es（西班牙语）、pt（葡萄牙语）、id（印尼语）、it（意大利语）、ko（韩语）、ru（俄语）、th（泰语）、vi（越南语）、ja（日语）、tr（土耳其语）、hi（印地语）、ms（马来语）、nl（荷兰语）、sv（瑞典语）、da（丹麦语）、fi（芬兰语）、pl（波兰语）、cs（捷克语）、fil（菲律宾语）、fa（波斯语）、el（希腊语）、ro（罗马尼亚语）、hu（匈牙利语）、mk（马其顿语）。

说明：

- 通过含 UI 的场景方案接入，详见 [语音转文字](#)，或者通过 SDK 客户端接入 ASR 功能，详见 [录音转文字](#)。
- 如果您有更多其他语言需求，请 [联系我们](#)。

语音合成（TTS）

将文本内容转化为自然流畅的语音输出，可用于语音播报、AI 变声、个性化语音消息生成等场景。

实时语音合成

说明	语种
将文本内容实时转换为语音音频文件。适用于 AI 机器人的流式输出、AI 对话等即时场景。支持系统预置音色或用户克隆的自定义音色。	中文、英语、日语、粤语

一句话语音合成

说明	语种
----	----

针对短文本场景的轻量级语音合成，适用于通知播报、消息朗读、语音提醒等场景。快速返回合成结果，满足低延迟需求。

中文、英语、日语、粤语

声音克隆：基于少量音频样本生成专属音色

基于一段参考音频，克隆出专属自定义音色。克隆成功后返回音色 ID，可在后续 TTS 调用中使用，实现个性化声音定制。

说明：
文字转语音及音色克隆功能在 IM SDK 8.9 及以上版本支持，详见 [文字转语音](#)。

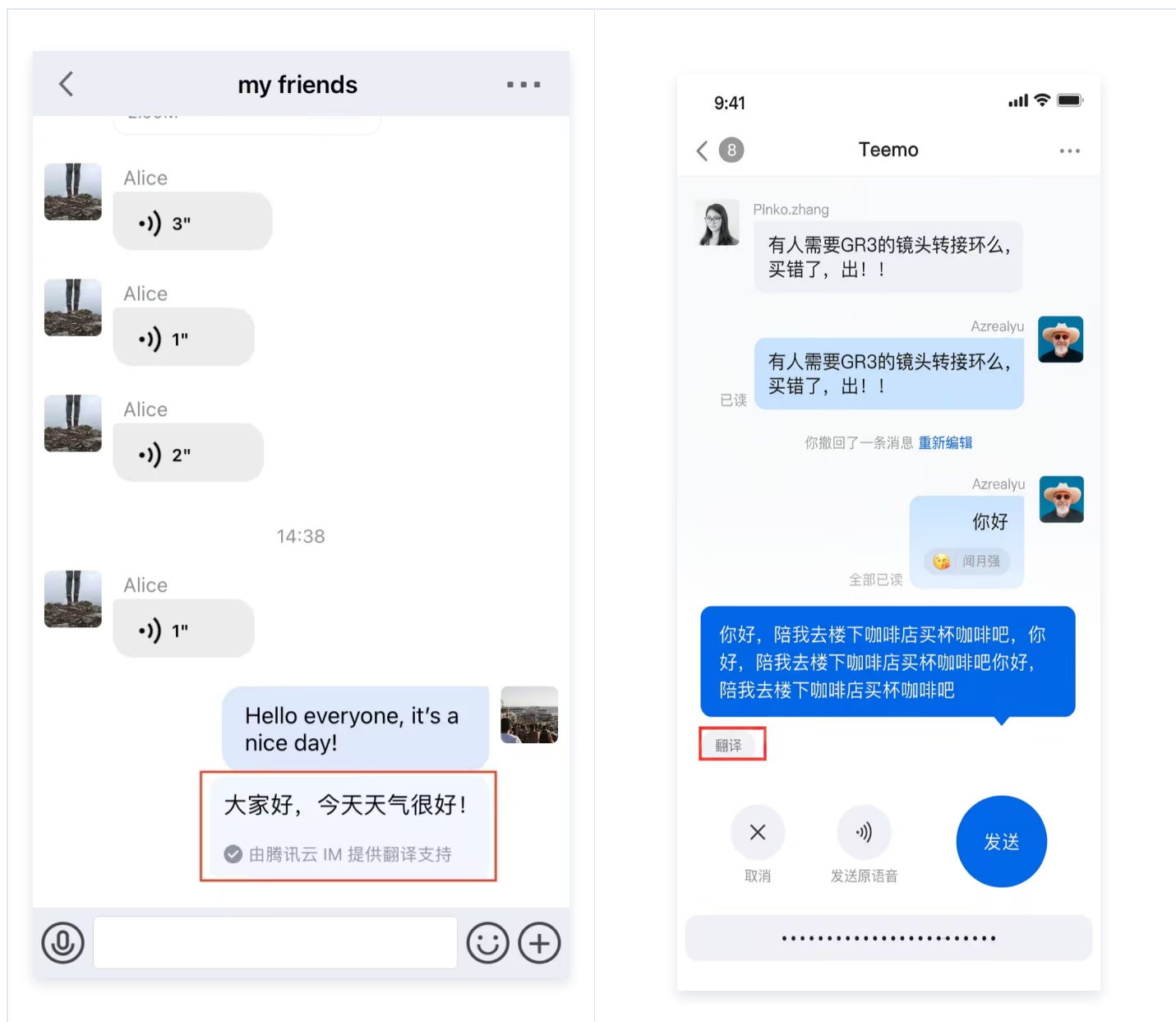
翻译

基于大语言模型技术的深度语义理解能力，系统能够自主适应目标语言表达习惯，实现自然流畅翻译效果。相较于传统翻译，翻译结果的机械感显著降低，在对话交流、泛娱互动等复杂场景下的效果提升明显。
特性及支持语种见下表。

说明	语种
文本翻译	● 国内：重点支持 33 个语种互译（中文、英语、法语、葡萄牙语、西班牙语、日语、土耳其语、俄语、阿拉伯语、韩语、泰语、意大利语、德语、越南语、马来语、印尼语、菲律宾语、印地语、繁体中文、波兰语、捷克语、荷兰语、高棉语、缅甸语、波斯语、古吉拉特语、乌尔都语、泰卢固语、马拉地语、希伯来语、孟加拉语、泰米尔语、乌克兰语），支持 5 种少数民族语言/方言（藏语、哈萨克语、蒙古语、维吾尔语、粤语）。 ● 境外：目前已经支持的翻译语言请参见 文本翻译语言。

效果展示

文本消息翻译效果	翻译效果
----------	------



接入方式

即时通信 IM AI 语音交互能力支持以下接入方式：

- 无 UI 接入：支持通过 IM SDK [客户端接入](#)；部分功能同时支持通过服务端 REST API 接入。具体能力及对应接入方式，可参考各功能章节中的链接。
- 含 UI 接入：支持通过 IM UIKit [客户端接入](#)。

⚠ AI 能力合规使用提示：

- 本服务基于大语言模型/语音大模型技术，生成结果可能因场景复杂性、数据质量、环境因素等原因而存在偏差，腾讯云不对生成内容的准确性、完整性、合法性、不侵权性提供担保。
- 您在使用本服务时应遵守《互联网信息服务深度合成管理规定》、《生成式人工智能服务管理暂行办法》《个人信息保护法》等相关法律法规，确保使用本服务处理的数据已获得相关主体的合法授权，不

得使用本服务生成虚假信息、诈骗内容、色情低俗内容、损害他人合法权益的内容及其他违反法律法规要求的内容，不得用于伪造他人声音或从事任何违法活动。

计费方式

具体费用详情请参见 [增值服务计费说明](#)。

开通服务

最近更新时间：2026-08-05 09:49:40

本文介绍如何开通即时通信 IM AI 智能语音服务。您可以根据 [AI 智能语音功能介绍](#) 了解所包含服务，并参考下述指引免费体验开通和购买开通正式 AI 智能语音服务。

前提条件

- 已 [注册腾讯云账号](#)。
- 已创建即时通信 IM 应用，并获取 SDKAppID。

开通方式

该服务需额外开通后方可调用相关能力，未开通时调用将返回错误。以下三种方式任选其一，配置完成后即可使用所有 IM AI 智能语音服务。

方式一：免费试用

在 [IM 控制台](#) 选择[语音转文字插件](#)或[文本消息翻译插件](#)，免费试用该插件，获取插件试用资格后，即可同步解锁 AI 智能语音服务。

获取语音转文字插件或翻译插件的试用资格详细请参见 [免费试用说明](#)。

方式二：购买插件包

购买 [语音转文字插件包](#) 或 [文本消息翻译插件](#) 任意一个插件包即可使用所有 IM AI 智能语音服务。

❗ 说明：

已购买语音转文字或文本消息翻译插件包的用户，可开通所有 AI 智能语音服务。

需注意：插件包赠送的用量仅限抵扣对应服务。例如，购买翻译插件赠送的用量只能抵扣翻译服务消耗，无法用于其他 AI 智能语音能力的抵扣。

方式三：开通企业版

开通 IM 企业版可免费开通使用所有 IM AI 智能语音服务的能力，具体用量计费规则请参考 [增值服务计费说明](#)，开通方式请参见 [概述及开通指引](#)。

功能接入

开通完成后，您可以参考以下功能文档进行接入：

- [AI 智能语音功能介绍](#)