

实时音视频

AI 实时转录与翻译



腾讯云

【 版权声明 】

©2013–2025 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 实时转录与翻译

功能介绍

开通服务

快速接入

AI 实时转录与翻译功能介绍

最近更新时间：2025-12-11 17:40:21

方案概述

基于腾讯RTC全球传输网络，AI 实时转录解决方案提供实时语音转文本和翻译能力。通过语音识别引擎（ASR）可对指定房间的音频流进行实时识别并将语音转换为文本内容（STT），快速构建直播实时字幕、会议内容实时记录、视频通话与语聊内容转写等多种场景下的语音识别应用。语音识别转文本后支持开启多语言翻译能力，实现跨语言内容实时翻译转写。

基本架构

AI 实时转录解决方案核心功能为实时语音转文字能力。翻译功能为可选项，未开启时不会产生翻译用量费用。

- 实时语音转文本：如下方架构图所示，音频流经过 TRTC 全球网络节点传输到服务端进行音频处理后，由语音识别引擎（ASR）将语音转写为文本。系统支持多语种、热词权重配置、VAD 检测以及实时流式识别。对于需生成字幕或实时内容记录的场景，在此流程即可直接输出最终转写结果回调。
- 实时翻译链路（可选择开启）：针对跨语言理解的场景，可在语音转写的基础上开启翻译功能。系统会将转写后的文本通过 LLM 翻译引擎进行翻译处理，并同时输出原文转写内容以及多语种翻译结果。



版本与能力支持

AI 实时转录服务包含语音转文本和实时翻译两大核心能力。其中，语音转文本能力分为基础、标准和高级三个版本引擎，可参考下文表格并根据使用场景需求选择适合的版本接入。

- 接入流程：具体的接入方式与流程参考 [快速接入](#)。
- 参数说明：下文表格中列出的“模型参数值”对应转录 API 接口 RecognizeConfig 中的 Language 字段，完整的参数定义说明请参阅 [转录 API 说明](#)。

	版本类型	特性说明	语种与模型型号
语音转文本	基础版语言引擎	基础通用语音识别模型。在近场、非复杂噪声的声学环境下具备良好的响应速度与识别准确率。	<p>通用场景：</p> <ul style="list-style-type: none">● "zh": 默认的8k中文（简体）模型，对电话场景支持较好。● "zh-TW": 16k繁体中文模型。● "en": 16k英语通用模型。 <p>垂类场景：</p> <ul style="list-style-type: none">● "16k_zh_edu": 中文教育。● "16k_zh_medical": 中文医疗。● "16k_zh_court": 中文法庭。
	标准版语言引擎	基于大模型引擎，语音识别性能大幅增强，在噪声回音大、人声远小等复杂音频环境的识别准确率显著提升。常见应用场景如会议、直播、语聊、游戏等实时字幕以及实时转写记录等。高度契合 RTC 实时互动相关场景。	<ul style="list-style-type: none">● "8k_zh_large": 大模型引擎，针对电话音频的识别支持较好。● "16k_zh_large": 大模型引擎，同时支持中文、英文、多种中文方言等语言的识别。● "16k_multi_lang": 多语种大模型引擎，当前模型同时支持英语、日语、韩语、阿拉伯语、菲律宾语、法语、印地语、印尼语、马来语、葡萄牙语、西班牙语、泰语、土耳其语、越南语、德语的识别，可实现15个语种的自动识别(句子/段落级别)。● "16k_zh_en": 最新大模型引擎，模型同时支持中文、英语，对多种中文方言和中英混合场景识别更优。
	高级版语言引擎	覆盖更多小语种、方言的精准识别需求，并支持配置3种语言自动识别。	<p>越南语、日语、韩语、印度尼西亚语、泰语、葡萄牙语、土耳其语、阿拉伯语、西班牙语、印地语、法语、马来语、菲律宾语、德语、意大利语、俄语、瑞典语、丹麦语、挪威语、中国粤语</p> <p>具体的语种配置参数枚举列表，请参考 转录 API 说明。</p> <p>如果您有更多其他语言需求，请 联系我们。</p>

实时翻译	实时翻译引擎	基于最新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，在非正式标准的口语对话交流、泛娱互动等复杂场景下的效果提升明显。	支持15种语言翻译：中文、英语、西班牙语、葡萄牙语、法语、德语、俄语、阿拉伯语、日语、韩语、越南语、马来语、印度尼西亚语、意大利语、泰语。 支持术语干预功能，提升模型在复杂应用场景下的效果。
------	--------	--------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------

计费方式

AI 实时转录费用主要由两个独立部分组成：RTC 音频通话费用、AI 语音转文本费用；若需要翻译功能则在前两者基础上额外增加实时翻译费用。以上能力都需要通过购买 AI 智能识别 或 TRTC 包月套餐包解锁，超量支持后付费，具体能力解锁所需套餐版本和费用详情请参见 [AI 智能识别计费说明](#)。

应用场景

应用场景	使用方式与价值
在线教育	通过 AI 转录对讲师与学员发言实时转写，课上提供实时字幕辅助理解，同步实时沉淀课程笔记与关键发言，可用于实时总结，课后复习、回放、存档。开启翻译功能后，可支持多语课堂或跨国教学场景。
社交娱乐	在社交娱乐场景中，通过AI转录能够对在线直播、游戏语聊、互动问答等场景，对发言内容进行实时转写，为观众提供准确字幕增强理解。同时支持输出多语言翻译内容，让不同地区用户能够更自然地参与互动。
办公协作	在线会议、线上面试、商务沟通、研讨会等场景下，通过 AI 转录生成实时字幕与内容记录，会中提升理解效率，会后可用于内容纪要整理。开启翻译功能后，可同时输出多语言转写，支持跨国团队无障碍协作。
客服中心	在线客服、智能外呼、呼叫中心等场景中，通过实时转写帮助客服准确理解用户内容。通话结束后沉淀完整文本记录，用于质检、合规与服务分析。开启翻译功能后，支持客服快速处理跨语言咨询，提高响应效率。

开通服务

最近更新时间：2025-12-11 17:40:21

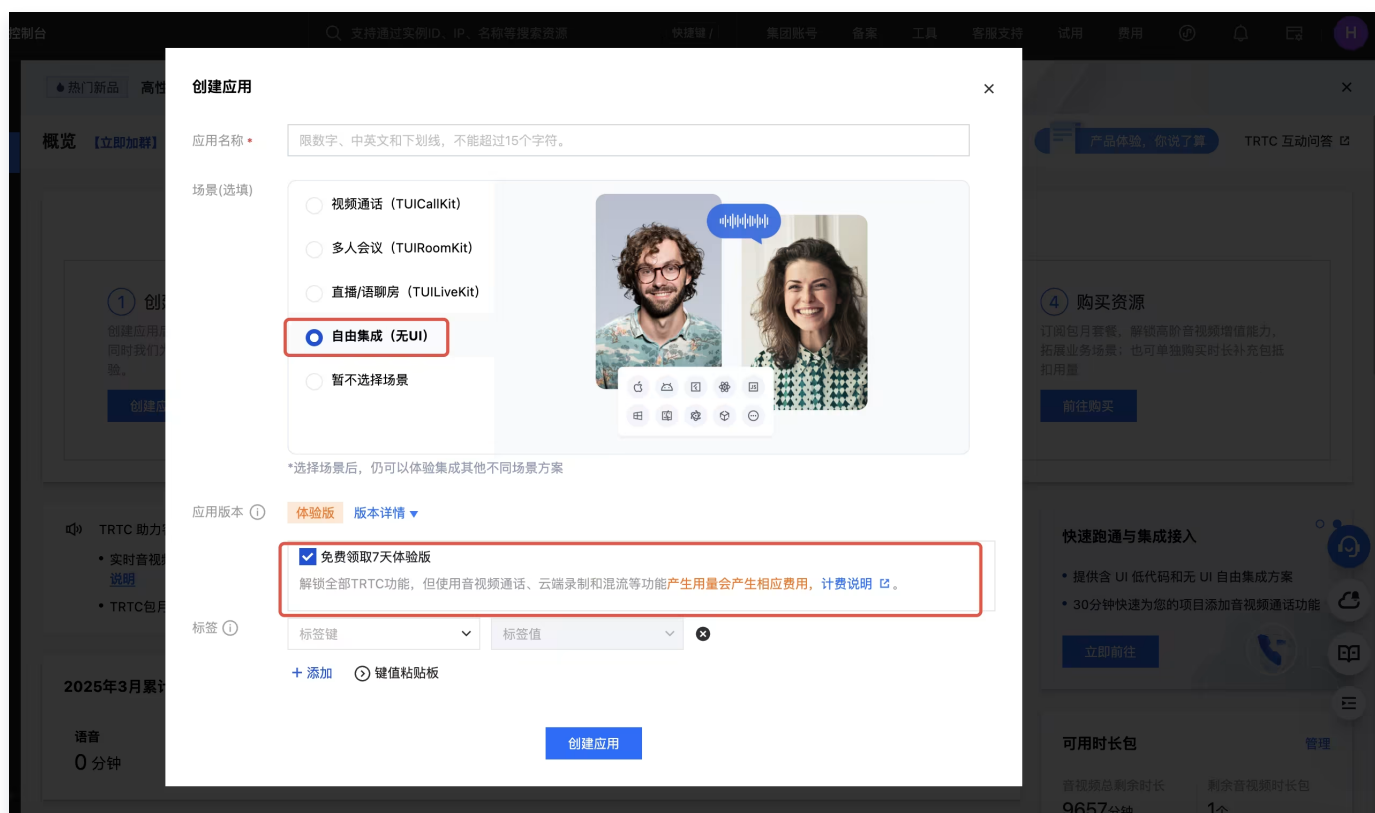
本文介绍如何 [开通体验版](#)、[开通正式版](#) 和 [续费正式版](#) 服务，您可以根据 [AI 智能识别计费说明](#) 了解所需要的服务及费用，并参考下述指引免费体验开通和购买开通正式版 AI 实时转录服务。

开通体验版

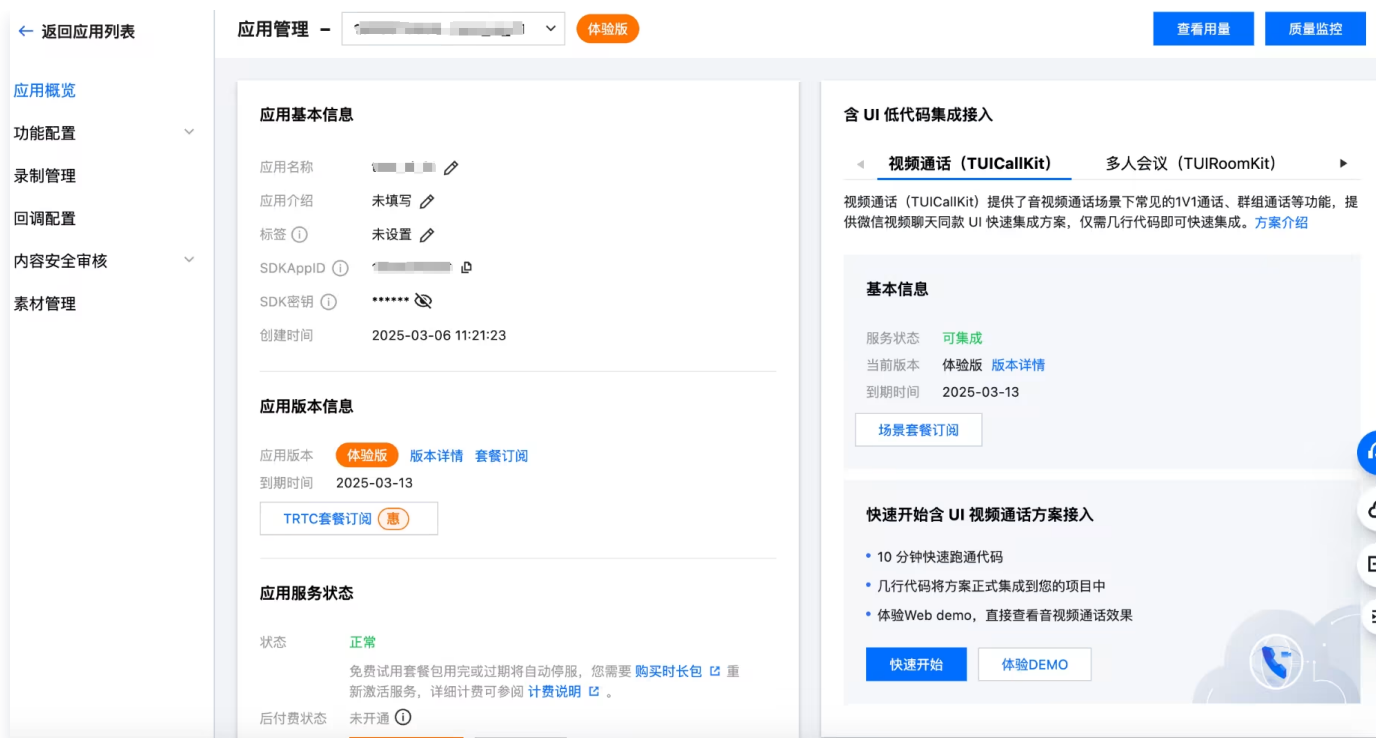
1. 登录 [实时音视频 TRTC 控制台](#)，单击**创建应用**。



2. 在弹窗中输入您的应用名称，选择**自由集成（无UI）**场景，勾选**免费领取7天体验版**，单击**创建应用**。



3. 进入应用管理页面，成功开通**自由集成（无UI）**的7天试用服务。领取完成后，可以看到体验版的基本信息，包括服务状态、版本信息和功能详情、到期时间。这里的 SDKAppID 以及 SDK 密钥（SDKSecretKey）会在后续步骤中使用到。



4. 体验版开通完成后，您可以根据 [快速接入](#) 文档接入实时转录和翻译功能进行体验。

⚠ 注意：

- 领取体验版后仅开通自由集成（无UI）7天的体验资格，测试过程中所产生的音视频时长等资源消耗（例如音频通话费用实时语音转文本服务本、实时翻译服务仍然会产生后付费产品费用，具体可参见 [AI 智能识别费用详情](#)），会按照实时音视频 TRTC 标准计费规则计费。
- 每个 SDKAppID 可免费体验2次，每次有效期均为7天；同时一个账号下所有 SDKAppID 的体验总次数为10次。
- 新账号首次可前往 [试用中心](#) 免费领取10000分钟音视频时长。

5. 若您当前 SDKAppID 领取的体验版已到期，您可单击[版本详情](#)，在版本详情弹窗中单击[体验续期](#)直接使用第二次体验机会，此时体验版有效期将延长7天。

⚠ 注意：

通过[体验续期](#)延长体验版有效期，将消耗1次免费体验机会。

开通正式版

[AI 转录的费用](#) 主要包含两个方面：音频通话费用、以及语音转文字费用；若开启翻译功能则会增加实时翻译费用。

- 其中，语音转文字功能需购买 AI 智能识别套餐包（轻量版/基础版/尊享版）或包月套餐尊享版 Plus /旗舰版 Plus 任一版本才可使用。
- 实时翻译功能需要购买 AI 智能识别套餐包（轻量版/尊享版）或包月套餐旗舰版 Plus 才可使用。
- 音频通话费用在使用过程中将默认开启后付费。

下文将详细说明如何开通正式版的 AI 智能识别套餐包并进行续费的步骤。

1. 登录 [实时音视频购买页](#)。
2. 根据实际业务需求选择配置。接入方案选择无 UI 自定义开发，以及选择需要购买 AI 智能识别套餐的 SDKAppID，购买后无法修改。

选择配置

接入方案

无 UI 自定义开发

直接集成 TRTC SDK，灵活开发多样化的需求

✓ 订阅解锁更多平台和场景（小程序通话等）

✓ 单独购买时长包享受更多优惠

含 UI 低代码快速开发

基于 TRTC 与 IM 的场景组件，1 天即可完成接入

✓ 音视频通话场景 —— TUICallKit

✓ 多人音视频会议场景 —— TUIRoomKit

资源选择

订阅包月套餐

单独购买时长包

购买 AI 智能识别套餐

购买行业版

使用须知

✓ 赠送时长包有效期为 1 个月，用尽后，默认按后付费以刊例价计费，到期未消耗的分钟数将自动清零

ⓘ 套餐包到期后，AI 智能识别相关功能停用，建议开启自动续费防止相关业务受影响

选择应用

创建应用

请确认应用选择正确，购买后无法修改

3. 套餐类型选择 AI 智能识别套餐，并选择您需要开通的套餐版本，AI 智能识别套餐详情可参见 [AI 智能识别详细说明](#)。

AI 智能识别套餐选择

AI 智能识别-基础版

支持多语种语音转文字功能，按月订阅并支持自动续期，每个周期赠送 10,000 分钟抵扣量，超量日结后付费

元/月

AI 智能识别-尊享版

支持多语种语音转文字功能，按月订阅并支持自动续期，每个周期赠送 20,000 分钟抵扣量，超量日结后付费

元/月

AI 智能识别-轻量版（限时 元）

支持 AI 智能识别-尊享版所有功能，有效期一个月，赠送 1,000 分钟抵扣量，超量日结后付费，每个账号限购 1 次

进入活动页购买

时长

1个月

3个月

6个月

12个月

2年

3年

自动续费

☐ 账户余额足够时，到期后自动按月续费

4. 购买完成后，您可以回到 [实时音视频 TRTC 控制台](#)，在 [时长包管理](#) 页面查看具体的 AI 智能识别套餐的使用情况。开通完成后，您可以根据 [快速接入](#) 文档进行转录功能的接入。

音视频时长补充包

录制&混流时长补充包

AI 智能识别时长包

购买时长包

抵扣明细

所有套餐包剩余时长：1000分钟

有效

无效

时长包ID	绑定sdkappid	套餐版本	时长包名称	时长包时长(分)	剩余时长(分钟)/(百分比)	生效时间	过期时间	状态	操作
1		AI 智能识别-轻量版	AI 智能识别赠送时长包-按月	1000	1000 / 100.00%			待使用 剩余31天	告警设置 抵扣明细

10

条 / 页

1

/ 1 页

续费正式版

您可对 AI 智能识别套餐进行续费，具体操作如下：

1. 登录 [续费管理控制台](#)，选择您想要续费的 AI 智能识别套餐，单击续费。

手动续费项 (1)

自动续费项 (0)

到期不续费 (0)

输入资源ID/资源别名搜索

Q

🔔

🔄

⌵

批量续费

设为自动续费

设为到期不续

统一到期日

<input type="checkbox"/>	资源ID/资源名	子产品名称	产品描述	地域 / 可用区	所属项目	到期时间	资源状态	续费周期	单价	操作
<input type="checkbox"/>	<div><div></div><div></div><div></div></div>	TRTC AI 智能识...	-	华南地区 (广州... 广州一区	默认项目	2025-04-10	运行中	1个月	<div>00元/月</div>	<div>续费</div> <div>设为自动续费</div> <div>更多</div>

2. 按照指示完成续费即可。

续费

已选中1个资源

资源ID/资源名	产品描述	当前到期时间	续费时长	续费后到期时间	单价	续费金额
	-	2025-04-10	1个月	2025-05-10	00元/月	00元

自动续费

☐ 账户余额足够时，资源到期后按续费周期自动续费

优惠信息

当前暂无可用优惠券。

价格

00元

最终实付金额以支付订单时为准

下一步

取消

快速接入

最近更新时间：2025-12-11 17:40:21

本文将介绍如何快速配置接入 TRTC AI 转录的实时语音转文本和实时翻译功能。

前提条件

- 登录 [TRTC 控制台](#)，开通 TRTC 服务并 [创建应用](#)。
- 前往 [控制台 > 功能配置 > 增值功能](#) 开启 AI 智能识别的语音转文字、实时翻译功能。



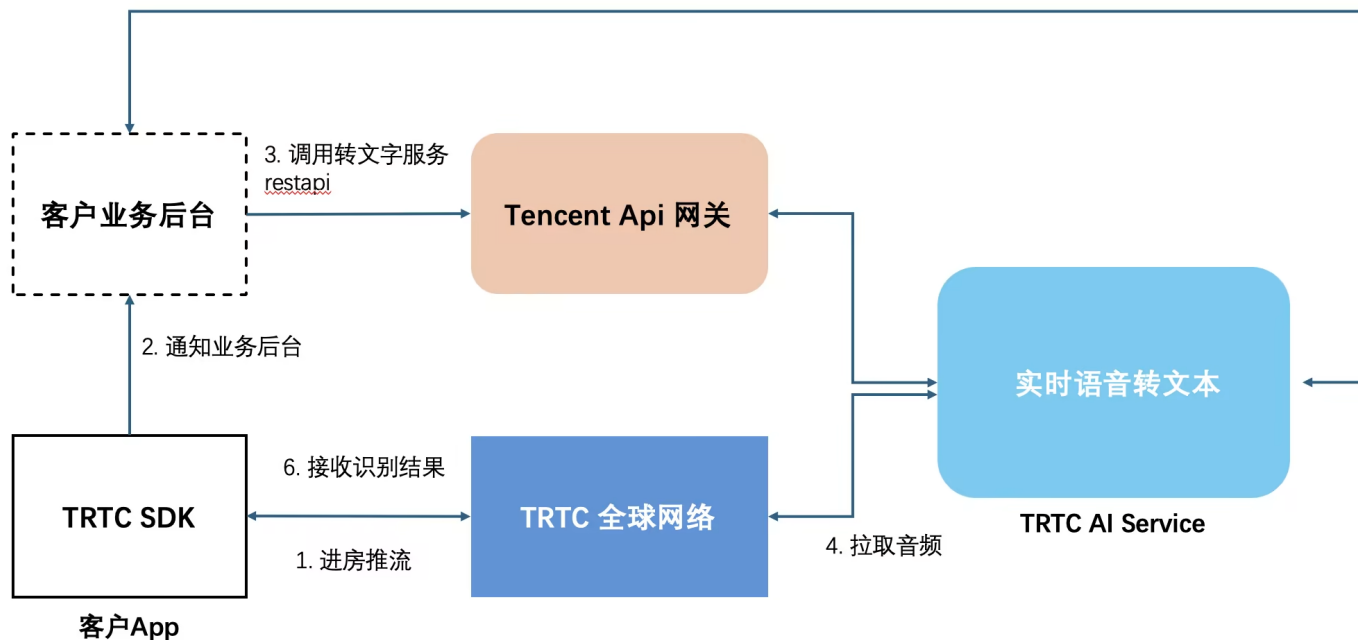
说明：

- 开启前需领取[包月套餐体验版](#)、购买 [AI 智能识别套餐包](#)、或订阅 TRTC 包月套餐部分版本才可使用，具体版本支持的功能详情可参见 [TRTC 包月套餐](#)，您可根据您的需要进行订阅。
- 语音转文字和实时翻译功能会根据调用量产生费用，详情请参见 [费用详情](#)。

功能说明

任务发起后，TRTC AI Service 通过识别机器人进入 TRTC 房间拉流指定用户或所有用户的流，进行语音转文字识别，将识别结果实时回调给客户端和服务端。

5. 回调识别结果



接入说明

步骤1: 集成 TRTC SDK

导入 TRTC SDK 到项目中，进入 TRTC 房间并开启本地麦克风音频采集与发布，相关指引可参考：

- [iOS 无 UI 集成指引](#)
- [Android 无 UI 集成指引](#)
- [Web&H5 无 UI 集成指引](#)
- [小程序 无 UI 集成指引](#)
- [Flutter 无 UI 集成指引](#)

步骤2: 后台发起语音转文字、实时翻译任务

TRTC 提供以下云 API 用于发起和管理语音转文字和实时翻译，具体如下：

- 开始语音转文字任务：[StartAITranscription](#)
- 查询语音转文字任务：[DescribeAITranscription](#)
- 停止语音转文字任务：[StopAITranscription](#)
- 开始实时翻译任务：在 [StartAITranscription](#) 的输入参数中配置 [TranslationConfig](#) 参数，在 TargetLanguages 中指定目标翻译语种，例如 [“en” , “ja”]。目前支持配置以下语种翻译：

语言代码	对应语种
"zh"	中文

"en"	英语
"es"	西班牙语
"pt"	葡萄牙语
"fr"	法语
"de"	德语
"ru"	俄语
"ar"	阿拉伯语
"ja"	日语
"ko"	韩语
"vi"	越南语
"ms"	马来语
"id"	印度尼西亚语
"it"	意大利语
"th"	泰语

⚠ 注意：

- 实时翻译功能目前支持15种语言：中文、英语、西班牙语、葡萄牙语、法语、德语、俄语、阿拉伯语、日语、韩语、越南语、马来语、印度尼西亚语、意大利语、泰语，如需其他语种支持可以 [联系我们](#)。
- 由于受不同语境或语种差异的影响，AI 翻译的译文内容适用于辅助参考，不应作为唯一专业意见或结论。

步骤3：接收语音识别结果

方式一：通过客户端 SDK 接收文字消息

通过 TRTC SDK [接收自定义消息功能](#)，在客户端上监听回调来接收实时的语音转文字、实时翻译的结果数据。客户端回调消息格式如下，以 Web 端为例：

```
trtc.on(TRTC.EVENT.CUSTOM_MESSAGE, event => { // receive custom message
  // event.userId: 语音识别机器人的userId
  // event.cmdId: 消息Id，转录和字幕固定为1
  // event.seq: 消息的序号
```

```
// event.data: ArrayBuffer 类型， 转录或字幕的内容，见下方data字段说明
const data = new TextDecoder().decode(event.data)

// data 字段说明如下

console.log(`received custom msg from ${event.userId}, message: ${data}`)
})
```

data 字段说明（实时的字幕消息）

字段名	类型	含义
type	Integer	10000：实时字幕与完整的一句话下发消息类型。
sender	String	说话人的 UserID。
receiver	Array	接收者 userid 列表，该消息实际是在房间内广播。
payload.text	String	识别出的文本。Unicode 编码。
payload.start_time	String	消息产生的时间 任务启动后的绝对时间。
payload.end_time	String	消息结束的时间 任务启动后的绝对时间。
payload.end	Boolean	如果为 true，代表这是一句完整的话。

```
{
  "type": 10000,
  "sender": "user_a",
  "payload": {
    "text": "",
    "start_time": "00:00:02",
    "end_time": "00:00:05",
    "end": true
  }
}
```

实时的翻译消息

```
{
  "type": 10000,
  "sender": "ai_951073",
  "payload": {
    "start_time_ms": 1760,
    "end_time_ms": 5530,
    "end": false,
    "roundid": "e6330a3c-eed7-40bb-8229-9bbe733a313f", // 一轮对话的 ID
    "translation_text": "simultaneous interpretation of the meeting", //
    翻译文本
    "translation_language": "en", // 翻译的语言代码
    "taskid": "x-dPLCz" // 转录任务唯一标识
  }
}
```

❗ 说明：

回调示例说明：

转录：会将完整的一句话转录并推送

“今天天气怎么样？”

字幕 & 翻译：将字幕句子或者翻译句子分段推送，后一段会包含前一段，确保实时性。

“今天”

“今天天气”

“今天天气怎么样”

顺序说明：字幕消息 > 字幕消息 > > 字幕消息(end = true)

方式二：通过服务端回调接收

语音转文字服务同时提供了服务端事件回调，便于您的服务接收实时对话的消息，查看 [详细回调事件](#)。

⚠ 注意：

语音转文字和实时翻译功能单个 SDKAppId 任务并发数限制**100路**，如需提升，请联系您的商务经理或[提交工单](#)评估处理。