

AudioX 语音克隆与语音生成

产品使用指南

2025 年 3 月

©北京深度逻辑智能科技有限公司版权所有，保留所有版权，文中出现的其他内容信息版权均为各自所有者所有。

禁止任何未经逻辑智能允许或授权的整体或部分抄袭的行为。

本文档的作者与出版人不对文中的错误或疏漏承担责任，也不对因本文描述在使用产品或服务等方面所带来的直接或间接损失负责。

逻辑智能对本文档的版本更新与内容变更不负有告知义务。

在使用本产品对音视频等内容进行操作时，请确保您拥有该内容的合法版权或已获得版权所有者的明确授权。任何侵犯他人版权的行为，将由用户自行承担法律责任。

平台禁止翻译、传播任何违法、色情、暴力、诽谤、侮辱、侵犯他人隐私等不良内容。用户应确保提交的音视频内容符合所在国家或地区的法律法规及社会道德规范。

一、 基本信息

基于音频大模型实现了自然、高度拟人、高感情的语音克隆和语音合成能力。

其中语音克隆可以充分保留原有说话人/配音的音色、情感和语速等特征，而语音合成除了支持使用克隆生成的 Speaker 外，也支持直接使用多种覆盖了大多数应用场景的预制 Speaker，解决了传统 AI 语音机械感强、语音质量差的问题，实现在影视配音、文化传媒、电商出海、语音客服、教育出海、数字人、IP 形象、有声出版、智能硬件等多个领域中 AI 语音内容的批量输出。

二、 使用方法

本产品使用 API 密钥进行身份验证。在完成购买后，请联系人工客服获取 API-KEY，并得到最新版本的接口地址和接口文档。

在语音克隆接口中，您可以通过上传一段语音文件的 base64 编码，来完成语音克隆。被克隆的语音文件时长通常在 5 秒左右，所包含语音的品质会影响克隆的质量，所以请提供干净、清晰的语音。发送指令后，系统会返回一个新的 Speaker ID，这一过程将消耗您已购买的克隆次数。

在语音合成接口中，您可以控制 text（待合成的文本内容）、speaker（说话人）、lang（待合成的语言）等参数，来完成语音合成。其中，既可以使用您通过克隆生成的 Speaker（如有），也可以直接使用接口文档中提供的预制 Speaker。发送指令后，系统会返回合成好的 AI 语音，这一过程将消耗您已购买的字符数。

三、 更多内容

您可以通过以下地址：

<https://apifox.com/apidoc/shared-e4da1232-46a9-4fd7-b521-72f6049e4c00/268721090e0>

访问本产品的接口文档以获取更多内容。