



产品数据表 · 捕获 · 语音转文字 STT

会议内容要逐字打出来？ 从现在起，每段录音都会以文字返回。

语音转文字 STT 可将会议、通话和口述录音转换为文字。它自动识别语言，Ultra 等级可区分发言人。专为办公、销售和项目团队打造，可在应用中使用，也可通过 API 调用。

一览

最大 50 MB 的音频和视频文件，自动语言识别，三个质量等级，Ultra 中的发言人区分，以及通过 API 提供的词级时间戳、SRT 和 VTT。

01 · 内容概述

先开会，再打字？ 现在文字会自动送达。

项目会议纪要今天必须发出，而录音时长为一小时。使用语音转文字 STT，您只需上传文件，即可取回文字。在 Ultra 中，文字按参与者分开，可直接复制或翻译。

输入

音频（MP3、WAV、OGG、FLAC、M4A）和视频（MP4、WebM、AVI、MOV），最大 50 MB

语言

自动识别；通过 API 可使用 100 个语言代码，应用中可选 15 种固定语言

输出

纯文本转写稿，Ultra 中按发言人以对话形式呈现；可下载为 TXT

访问方式

应用中位于

Media AI › 语音工作室
下、REST API 和 MCP



自动识别语言

语音转文字 STT 会自行识别录音的语言。如果您已知语言，可以设定，通过 API 可从 100 个语言代码中选择。

三个质量等级

Basic 快速返回纯文本，Premium 增加带词级时间戳的分段，Ultra 还会区分发言人。等级位于 [转录质量](#) 下。

区分发言人

在 Ultra 中，语音转文字 STT 会识别参与者，并以对话形式显示转写稿。点击一下即可重命名发言人。

词汇表

借助 [专业词汇](#)，语音转文字 STT 能更准确地识别您的知识库和价目表中的专业术语。通过 API，您可从 Premium 起添加自己的热词。

通过 API 生成字幕

通过 API，Premium 及以上可返回 SRT 和 VTT 以及词级时间戳；Ultra 另可返回用于发言人分析的 RTTM。

交给 Translate Studio

[在翻译工作室中使用](#) 可将转写稿直接交给 Translate Studio，无需复制粘贴。

拖入文件，选择等级。 直接使用转写稿。

在 Voice Studio 中只需四个步骤。通过 REST API 和 MCP，您可以从自己的流程中直接发送音频文件。

1

打开 Voice Studio

在 Media AI › 语音工作室 下，选择 语音转文本 (STT) 选项卡。从起始页也可通过 快捷工具 › Media › 语音转文本 进入。

2

拖入文件

将音频或视频文件拖入输入区域。语言保持自动，或手动选择。

3

选择等级与选项

在 转录质量 下选择 Basic、Premium 或 Ultra。在 Ultra 中，您可以在 对话者识别 下设定发言人数量。

4

使用转写稿

复制它，使用 下载为 TXT 保存为文本文件，或一键交给 Translate Studio。

技术数据

功能	详情
输入格式	MP3、WAV、OGG、FLAC、M4A、MP4、WebM、AVI、MOV；应用中最大 50 MB，每次运行一个文件
语言	自动识别；通过 API 可使用 100 个语言代码，应用中有 15 种固定选项
质量等级	Basic（API 默认）、Premium、Ultra（应用默认）
API 输出	output_format 可选 json、srt 和 vtt（自 Premium 起），rttm（Ultra）
发言人区分	Ultra 中可选 min_speakers 和 max_speakers，范围 1 至 20
REST API	POST /job/add/paperoffice_voice___stt，参数 audio_file
MCP	/mcp-full 上的 po_stt_transcribe

03 · 实际应用

会议、口述、培训视频。 每一项都是完成的文字。

项目会议纪要

项目负责人在 Ultra 中上传录音，将发言人 1 和 2 重命名为对应姓名，然后将对话复制到会议纪要中。

外勤口述

测量员在路上用手机口述，回到办公室后上传 M4A 文件。Basic 返回报告所需的纯文本，词汇表则可识别产品名称。

培训视频的字幕

开发人员通过 API 发送视频，参数为 `quality=premium` 和 `output_format=srt`，并取回带时间戳的字幕文件。

常见问题

语音转文字 STT 能识别哪些语言？

语言会自动识别。通过 API 可从 100 个语言代码中设定语言，应用中则提供 15 种语言作为固定选项。

语音转文字 STT 能区分发言人吗？

可以，在 Ultra 中。转写稿以对话形式呈现，发言人可以重命名。通过 API 您还可获得发言人数量、重叠语段和 RTTM。

我能获得 SRT 或 VTT 格式的字幕吗？

可以，通过 API 并使用 `output_format=srt` 或 `vtt`，从 Premium 等级起可用。在应用中，您可将转写稿保存为 TXT 文件。

如何计费？

按用量计费，取决于质量等级。当前价格见定价页面。

开始使用

让您的下一段录音自动整理成文字。

登录您的 PaperOffice 账户，打开 [Media AI](#) > 语音工作室 并选择 语音转文本 (STT)。拖入文件，选择等级，开始转写。

paperoffice.ai/zh/ai-voice-stt-yuyin-zhuanwen-shibie-stt



了解更多模块

paperoffice.ai/zh/ai-voice-stt-yuyin-zhuanwen-shibie-stt

价格

paperoffice.ai/en/pricing

联系我们

paperoffice.ai/zh/lianxi