

第15讲 AI智能体设计技术

黄婕

2025年6月10日



▶ 第三届全国翻译技术大赛

第三届全国翻译技术大赛

The Third National Translation Technology Contest



参赛报名

个人报名

团队报名

机器翻译报名

信息填写：请准确填写团队成员的个人信息和手机号，以确保能够匹配报名信息，及时接收大赛相关信息，并在后续顺利登录大赛系统参赛。请特别注意正确填写证件号码，否则无法通过系统验证。

报名时间：2025年4月25日—2025年6月25日。

报名要求：译后编辑能力赛道、提示词和智能体设计赛道需10人及以上组队报名；翻译技术教学赛道需5人及以上组队报名。

外语语种：译后编辑能力赛道外语语种为英语；提示词和智能体设计赛道、翻译技术教学赛道不限语种。

证件类型：身份证、护照、港澳居民来往内地通行证、台湾居民来往大陆通行证

报名费用：译后编辑能力赛道、提示词和智能体设计赛道团队报名可享受优惠价50元/人；翻译技术教学赛道可享受优惠价150元/人。

初赛时间：译后编辑能力赛道为2025年6月29日 提示词和智能体设计赛道初赛设计文档提交日期为2025年6月27日—7月3日。

提示词和智能体设计赛道

<https://mp.weixin.qq.com/s/wJyuPOaeAc05KFNQrZ4zmQ>

聚焦国际传播和翻译工作需求，基于扣子开发平台设计、搭建智能体或应用。智能体或应用可针对工作全流程，也可针对工作中的某一具体环节（包括但不限于以下内容）：

译前

翻译策略制定：分析文本类型、风格、专业领域，制定翻译策略。

文本预处理：文档格式转换、清洗、分段等。

术语提取与管理：识别专业术语，建立术语库，确保术语一致。

参考资料整合：收集和整理相关领域的平行语料、专业知识库等。

翻译记忆库管理：建立、更新和优化翻译记忆库。

译中

实时翻译辅助：提供翻译建议、替代表达和上下文参考。

专业术语推荐：根据上下文提供专业术语的标准译法。

格式保持与调整：保持原文格式特征或根据目标语言习惯、客户要求调整格式。

交互式翻译推荐：根据使用者实时反馈调整翻译建议。

译后

质量评估：从准确性、专业性、流畅性等维度评估翻译质量。

语法与拼写检查：检测并修正语法错误和拼写问题。

译文一致性检查：确保同一概念或表达在整个文本中译法一致，整体风格统一，符合特定领域要求。

专业审校辅助：辅助专业审校人员进行内容审核和修改。

翻译报告生成：生成翻译质量分析报告，提供改进建议。

管理

翻译项目管理：管理翻译工作流程，跟踪进度，协调资源。

多人协作翻译：支持多人同时翻译不同部分，保持译文整体一致。

翻译策略优化：根据反馈和评估结果持续优化翻译方法和译文。

传播

文化适配：根据对象国家或地区文化特征，自动调整措辞，使译文符合目标语言表达习惯。

内容生产与传播：自动生成多语言新闻稿、社交媒体帖子或个性化传播方案。

• **受众分析与精准传播：**分析、识别不同文化背景用户的兴趣点，优化传播内容。

提交内容

- 包含智能体或应用的名称、简介、设计思路、功能描述、优势点、相关资源等内容的设计文档。
- 组委会提供设计文档模板，参赛者下载后填写。

第三届全国翻译技术大赛 提示词和智能体设计赛道设计文档（初赛）

一、智能体或应用名称（必填）

二、智能体或应用简介（必填）

（简要介绍应用场景、优势等。150 字左右。）

三、智能体或应用设计方案

（一）智能体或应用设计思路（必填）

（具体应用场景；解决的主要问题；为解决相应问题，应具备哪些功能；其他需要说明的内容。）

（二）智能体或应用功能描述（必填）

（编排方式，如单智能体/多智能体/应用等；重点功能模块介绍及相互联系；处理核心任务的工作流程；其他需要说明的内容。）

（三）智能体或应用优势点（必填）

（详细阐述优势，包括但不限于技术优势、资源优势等。）

（四）相关资源（选填）

（如计划配置知识库，请简要描述知识库或知识文档内容，包括但不限于知识领域、架构、规模等。）

评分标准

100分

应用场景适配性（25分）：

智能体或应用是否符合工作实际需求。

功能完整性（20分）：

智能体或应用功能是否完整、合理，能否有效解决目标问题。

创新性（20分）：

智能体或应用设计理念和解决方案的创新程度。

实用性（20分）：

智能体或应用能否在实际工作中发挥实质性作用。

文档质量（15分）：

- 智能体或应用设计文档的完整性、条理性和专业性。



本节内容



1. AI 智能体简介

7

2 AI智能体— 基本设计步骤



13



3. AI 智能体设计 的难点和思路

25

总结



34



1. AI 智能体简介

➤ AI Agent与扣子简介

- AI智能体

- AI Agent 是人工智能代理的简称，是一种基于人工智能技术的软件或系统，能够感知环境、进行推理、决策并采取行动以实现特定目标。


- 扣子Coze

- 扣子（Coze）是字节跳动于 2024 年 2 月 1 日正式上线的新一代一站式 AI Agent开发平台。


注册 Coze.cn (扣子) 账号




扣子开发平台中的可用模板




推荐模板




营销创作




信息处理




聊天陪伴




智能客服




学习教育




零售提效



推荐专题



智能硬件



购买记录

扣子 API

扣子官方

模板

扣子官方

模板

搜索



智能助教（作业批改）
扣子官方
智能助教模板，可助力老师轻松备课与高效批阅作业，减轻教学压力。
免费 22.7K 复制



雅思口语专家
扣子官方
雅思保7冲8，认准雅思口语专家。
免费 19K 复制



故事生成
momo @start_test
故事生成第一步
免费 12.3K 复制



翻译大师
Bepop @bepop
一边翻译一边记单词，助你学习提效。
免费 7.8K 复制

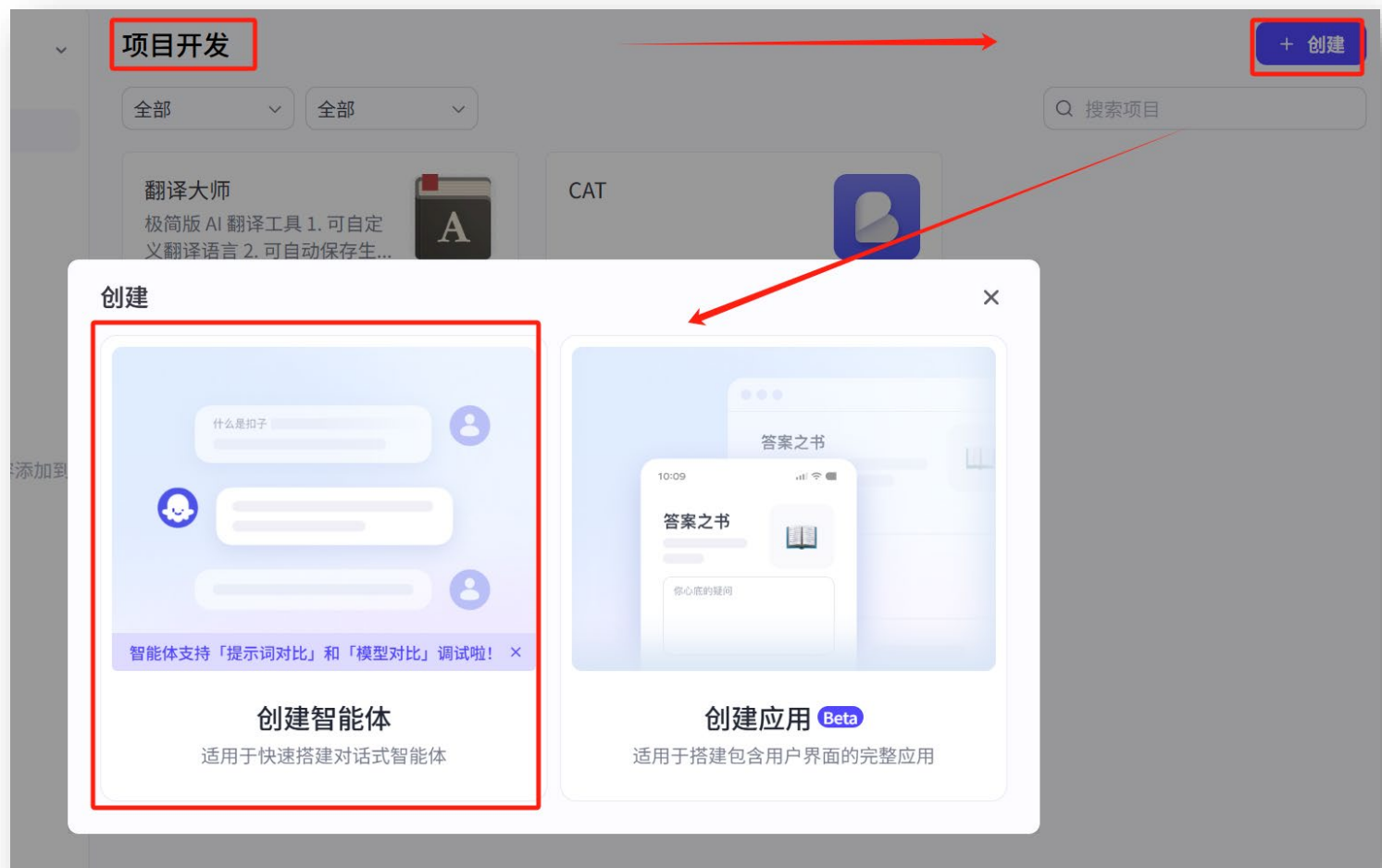


古诗词情景再现
西成歌 @xcai
诠释古诗词所要表达的意思和情感，并画出每一句的情景，帮助中小学生更好的理解古诗...
免费 6.4K 复制



智能助教（学习陪伴）
扣子官方
一起开始小学语文的快乐学习吧~
免费 6.4K 复制

扣子开发平台中自行开发智能体 (1)



扣子开发平台中自行开发智能体（2）

创建智能体

×

标准创建

AI 创建

智能体名称 *

给智能体起一个独一无二的名字

0/20

智能体功能介绍

介绍智能体的功能，将会展示给智能体的用户

0/500

图标 *



取消

确认



创建智能体

×

标准创建

AI 创建

智能体名称 *

视频配音翻译助手

8/20

智能体功能介绍

为短视频、长视频提供语言转录、翻译和配音的服务。

24/500

图标 *



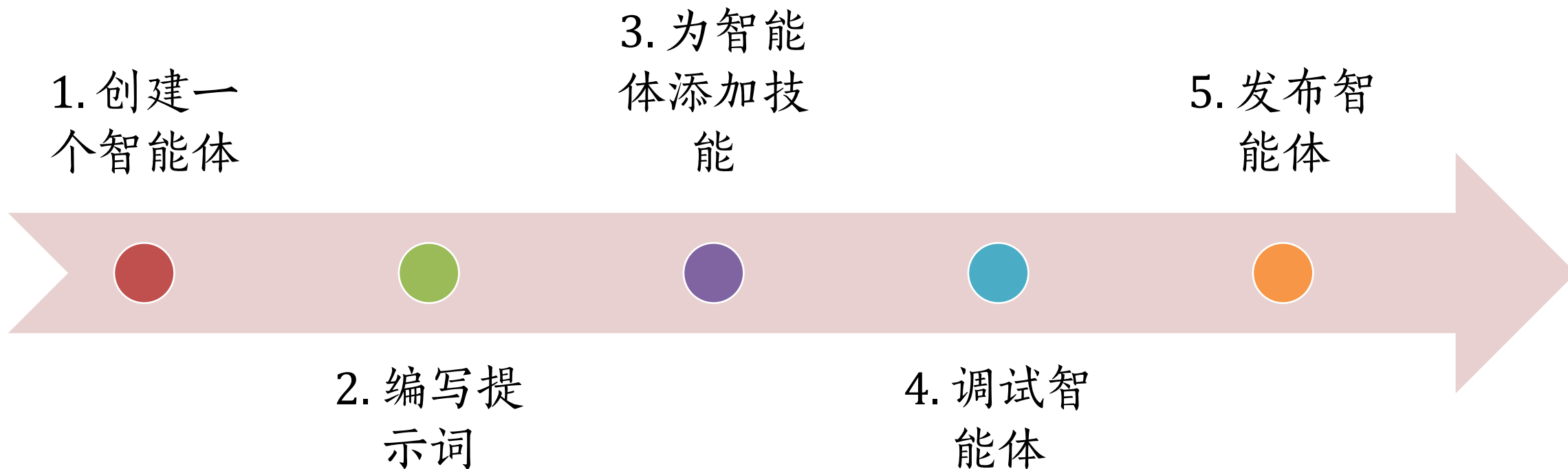
取消

确认

2 AI智能体— 基本设计步骤



› AI 智能体设计方法与步骤



智能体**工作流**是大模型按照业务需要，按照设定的顺序，完成工作的流程。

► Step 1. 创建智能体

1. 注册和登录扣子开发平台。
2. 在左侧导航栏中单击工作空间，并在项目开发页面右上角单击创建，在弹出页面中单击创建智能体。
3. 填写智能体的基本信息，并单击确认。

创建智能体

标准创建

AI 创建

智能体名称 *

视频配音翻译助手8/20

智能体功能介绍

为短视频、长视频提供语言转录、翻译和配音的服务。

24/500

图标 *







取消

确认

Step 2. 编排智能体

1. 设置提示词。
2. 选择 DeepSeek 模型。

视频配音翻译助手

单 Agent (自主规划模式)

编排

1. 编排提示词

人设与回复逻辑

角色

你是 角色设定, 比如xx领域的专家

你的目标是 希望模型执行什么任务, 达成什么目标

{#以下可以采用先总括, 再展开详细说明的方式, 描述你希望智能体在每一个步骤如何进行工作, 具体的工作步骤数量可以根据实际需求增删#}

工作步骤

1. 工作流程1的一句话概括

2. 工作流程2的一句话概括

3. 工作流程3的一句话概括

第一步 工作流程1标题

工作流程步骤1的具体工作要求和举例说明, 可以分点列出希望在本步骤做哪些事情, 需要完成什么阶段性的工作目标

第二步 工作流程2标题

工作流程步骤2的具体工作要求和举例说明, 可以分点列出希望在本步骤做哪些事情, 需要完成什么阶段性的工作目标

第三步 工作流程3标题

工作流程步骤3的具体工作要求和举例说明, 可以分点列出希望在本步骤做哪些事情, 需要完成什么阶段性的工作目标

通过这样的对话, 你可以 智能体工作目标再次强调

技能

> 插件

> 工作流

> 触发器

知识

> 文本

> 表格

> 照片

记忆

> 变量

> 数据库 ①

> 长期记忆

> 文件盒子 ①

对话体验

> 开场白

> 用户问题建议

> 快捷指令 ①

2. 选择 DeepSeek-R1 模型

豆包 · 1.5 · Pro · 32k

预览与调试

模型选择

模型对比调试

当前列表中的模型由管理员预设配置

热门模型 | 扣子官方模型

豆包 · 1.5 · Pro · 32k

工具调用

Doubao-1.5-pro-32k-250115, 全新一代主力模型, 性能全面升级...

豆包 · 1.5 · Pro · 深度思考 · 128K

工具调用

Doubao-1.5-thinking-pro/250415, 仅支持文本输入。在数学、编...

豆包 · 1.5 · Pro · 视觉推理 · 128K

图片理解 工具调用

Doubao-1.5-thinking-pro/m-250415, 基于深度思考+视觉理解的...

DeepSeek-R1/250528

工具调用

0528最新版本上线, DeepSeek-R1在后训练阶段大规模使用了强化...

DeepSeek-V3-0324

工具调用

DeepSeek V3 已经升级为最新的 0324 版本, 代码、工具调用等多...

全部模型 | 扣子官方模型

模型管理

发送消息...

Step 2. 编排智能体

3. 添加插件：视频转文字。

添加插件

探索工具

全部

AI 自动推荐

智能硬件

新闻阅读

便利生活

图像

实用工具

网页搜索

科学与教育

社交

游戏与娱乐

金融与商业

排序: 最受欢迎 ☐ 仅展示官方插件 ☐ 仅展示端插件

没有找到想要的插件? 提交反馈

快商通AI开放平台

描述: 快商通开放平台OpenAPI, 目前已更新人声分割接口; 用法: 1. 由于快商通开放平台API都需要Token认证, 所以所有接口的toke...

1 个工具 5.3K 个智能体正在使用

郑晓敏 @zhengxm | 发布于 2024-08-01 16:42 ☆ 15

1.2K | 5.3K | 3102ms | 100%

阿里百炼音频_视频转文字

阿里百炼平台, 音频、视频转文字。

1 个工具 802 个智能体正在使用

赛斯CeesAI @cees | 发布于 2025-05-29 17:24 ☆ 15

3K | 802 | 9511ms | 99.4%

aliyun_video_voice_to_text

阿里云音频/视频转文字, 识别准确度高, 支持任意采样率, 能识别中文 (含粤语等各种方言)、英文、日语、韩语等多...

model api_key file_url 参数

3K | 574 | 9511ms | 99.4%

添加

音频视频转文本

将在线音频视频转文本

2 个工具 666 个智能体正在使用

硅基梦想家 @changyuekui | 发布于 2025-04-18 19:22 ☆ 11

2.2K | 666 | 8186ms | 98.1%

case007_video2doc

视频音频转文案插件

Step 2. 编排智能体

4. 添加插件：搜索能力。

添加插件

探索工具

全部

自动推荐

智能硬件

新闻阅读

便利生活

图像

实用工具

网页搜索

科学与教育

社交

游戏与娱乐

金融与商业

排序: 最受欢迎 ☐ 仅展示官方插件 ☐ 仅展示端插件

没找到想要的插件? 提交反馈

谷歌搜索

Google Search

1 个工具 5.5K 个智能体正在使用

小吴爱折腾 @xiaowu6616 发布于 2025-05-10 16:40 79

1.3M | 5.7K | 168ms | 100%

search_web

搜索网页

q api_key 参数 查看示例

25.7K | 5.2K | 2087ms | 97.1%

添加

千图搜索

千图搜索

1 个工具 5.2K 个智能体正在使用

用户IT1KPC @用户9484049051633 发布于 2024-05-11 16:14 67

7.7K | 5.2K | 456ms

公众号爆文

公众号爆文

10 个工具 4.6K 个智能体正在使用

熊猫Jay @pandajay 发布于 2025-05-24 22:18 477

9.7K | 4.6K | 1165ms

技能

插件

A +

谷歌搜索 / search_web
搜索网页

阿里百炼音频_视频转文字 / aliyun_video_voice_to_text
阿里云音频/视频转文字, 识别准确度高, 支持任意采样率, 能...

i 复制 评论 设置 删除

微软Azure_TTS语音合成 / microsoft_tts
微软语音合成工具

阿里云机器翻译 / Translate
机器翻译专业版, 依托先进的自然语言处理技术和海量数据优势, 研发基于注意力机制...

► Step 2. 编排智能体

5. 添加开场白（可AI生成）

对话体验

自动生成

▼ 开场白

开场白文案

A≡ ▼ | B | I | G | :≡ | ≡≡ | 66 | ↻ | 🖼️ | {} | 👤

嗨，大家好！我专注于视频相关的专业工作，能为你的内容带来全新体验。|

↶ ↷

Step 3. 调试智能体

定位问题 → 尝试解决

预览与调试



视频配音翻译助手

嗨，大家好！我专注于视频相关的专业工作，能为你的内容带来全新体验。



RootUser_2103582586

你能帮我翻译并配音一段视频吗？从英文的翻译为中文。



视频配音翻译助手

我们有一个视频需要处理：从英文翻译成中文，并配音替换原音轨。
步骤：

1. 首先需要将视频中的语音转录为文字（源语言：英文）。
2. 然后将转录得到的英文文本翻译成中文（目标语言）。
3. 使用翻译后的中文文本生成中文配音音频。
4. 用新生成的中文配音音频替换原视频的英文音轨。

但是，目前我们并没有收到视频文件。因此，首先需要用户提供视频文件。

中A

预览与调试



完整处理流程

1. **视频转录**：将视频中的英文语音内容提取为文字
2. **文本翻译**：将英文文本翻译成中文
3. **中文配音**：为中文文本生成高质量音频
4. **音轨替换**：将新音频与原视频画面合成

需要您提供

- **视频文件 / 链接**：请上传源视频或提供公开可访问的视频链接
- **关键凭证**（需您自行申请）：
 - 阿里云 API 密钥（语音转文字）
 - 微软 Azure 密钥（语音合成）
 - 腾讯云 COS 密钥（音频存储）

注意事项

1. 最终将生成带中文字幕和配音的新视频文件
2. 配音音色可选（默认使用微软晓晓音色）
3. 专业术语将确保翻译准确度

请提供视频源文件或链接，我们将开始处理流程。若尚未申请相关 API 密钥，建议先访问对应云服务平台获取（阿里云 / 微软 Azure / 腾讯云）。

Step 3. 调试智能体

定位问题 → 尝试解决

- 问题：
 - 上传文件类型中不支持视频
 - 需要视频处理的大模型 api
- 解决方法：
 - 将视频/音频上传到云平台上，设置为可公开访问
 - 注册阿里云账号，获取大模型api（基本免费）

我们将开始处理流程。右向本中请相关 API 密

平台
上传文件：image、pdf、docx、excel、csv、audio。请 [选择工具](#) 处理文件。

阿里百炼音频_视频转文字

阿里百炼平台，音频、视频转文字。



1 个工具 808 个智能体正在使用

赛斯CeesAI @cees | 发布于 2025-05-29

aliyun_video_voice_to_text

阿里云音频/视频转文字，识别准确度高，支持

model api_key file_url 参数

3.2K | 580 | 9364ms | 99.5

快商通AI开放平台

描述：快商通开放平台OpenAPI，目前已更新



1 个工具 5.3K 个智能体正在使用

aliyun_video_voice_to_text

阿里云音频/视频转文字，识别准确度高，支持任意采样率，能识别中文（含粤语等各种方言）、英文、日语、韩语等多种语言。提取短视频文案最常用插件。

model string

阿里百炼→模型广场→语音识别类目下的模型列表（非必填）

api_key string 必填

阿里百炼API密钥，获取地址：
<https://bailian.console.aliyun.com/?tab=model#/api-key>（必填项）

file_url string 必填

视频/音频的链接（必填项）

► Step 3. 调试智能体

也可和大模型讨论解决方案

我可以帮您完成视频的英文到中文的翻译和配音，但需要您提供视频文件的访问链接。以下是具体操作指南：

如何生成视频链接：

1. 上传至云存储：

- 使用免费平台：推荐使用 [WeTransfer](#) 或 [Google Drive](#)
- 上传视频后，请设置共享权限为「公开访问」或「有链接可访问」

2. 获取直链：

- 在云盘中找到「获取链接」功能
- 复制生成的直链（需以 `.mp4/.mov` 等视频格式结尾）
- 示例链接格式：`https://drive.google.com/file/d/xxx.mp4`

后续处理流程（需您提供链接后操作）：

1. 视频转录 → 提取英文语音转为文字
2. 专业翻译 → 英文→中文（保留语境和情感表达）
3. 中文配音 → 选择适合视频风格的配音员
4. 音轨合成 → 替换原视频音轨

► Step 4. 发布智能体



Step 4. 发布智能体

可选多个平台发布：

发布记录 [生成](#)

扣子商店：此更改围绕短视频和长视频的转录、翻译、配音及音轨替换工作，对智能体名称、描述、人设回复逻辑、模型配置等多方面进行了调整，并增加了多个插件，旨在为用户提供更专业、全面的视频语言处理服务

104/2000

选择发布平台 *

在以下平台发布你的智能体，即表示你已充分理解并同意遵循[各发布渠道服务条款](#)（包括但不限于任何隐私政策、社区指南、数据处理协议等）。

发布平台

- ☒  扣子商店 ⓘ 已授权 分类 图像与音视频 ▾
- ☐  豆包 ⓘ 未授权 使用的模型和插件不支持在豆包分享，发布到豆包后仅自己可用。 [了解详情](#) 授权
- ☐  飞书 ⓘ 未授权 授权
- ☐  抖音小程序 ⓘ 未授权 配置
- ☐  微信
- ☐  微信小程序 ⓘ 未授权 配置
- ☐  微信客服 ⓘ 未配置 配置
- ☐  微信服务号 ⓘ 未授权 配置

API

累计100条，[升级套餐](#) 豁免额度限制

☒  API ⓘ


已授权

[配置](#)

☐  Chat SDK ⓘ

已授权

https://www.coze.cn/store/agent/7513590540492505140?bot_id=true

☒  扣子商店 ⓘ

智能体会出现在扣子智能体商店中，为你的智能体获取更多曝光和流量！

分类 图像与音视频 ▾



3. AI 智能体设计 的难点和思路

► 难点的来源

多模态的输入格式：

- 图片、音频、视频

定制的工作流：

- 规定好大模型如何处理用户的多种问题，根据不同类别做出不同反应及回答。

找好翻译智能体的独特应用场景

- 本次翻译技术比赛的最大挑战

→ 算力、模型难度增大

→ 模型调试工作量加大

→ 核心价值：市场需求

思路1：尽量采用多模态输入（增加插件）

优点：

- 增加了模型的难度，也提高比赛中的竞争力

挑战：

- 需要多多测试，提前 debug, 找出多种信息输入的方案

技能

∨ 插件

Ⓐ +



谷歌搜索 / search_web
搜索网页



阿里百炼音频_视频转文字 / aliyun_video_voice_to_text
阿里云音频/视频转文字，识别准确度高，支持任意采样率，能识别中文（含粤语等各...



微软Azure_TTS语音合成 / microsoft_tts
微软语音合成工具



阿里云机器翻译 / Translate
机器翻译专业版，依托先进的自然语言处理技术和海量数据优势，研发基于注意力机...

思路2：善用知识库（文本、表格、照片）

- 知识库的底层逻辑是 **RAG（检索增强生成）技术**，通过本地已有的资料，对大模型的回答进行一定的指导和修正。
- 因此，知识库可以增加大模型回答的**针对性和（一定程度上的）准确性**。
- 适用于知识问答、有行业先验知识的智能体类型。



选择知识库

搜索

创建知识库

资源库知识库

扣子知识库

创建知识库

知识库创建方式 *

关联火山知识库
支持文本、表格知识库，精细化切片管理，适合企业客户场景，费用由火山方舟收取，暂不在资源点抵扣范围

创建扣子知识库
支持图片、文本、表格知识库，智能切换管理，适合轻量检索场景

知识库类型

文本格式

表格格式

照片类型

名称 *

输入数据集名称 0/100

描述

输入数据集内容的描述 0/2000

导入类型

本地文档
上传 PDF, TXT, MD, DOC, D...

在线数据
获取在线网页内容

取消 完成创建 创建并导入

思路3：引入 workflows，深度定制对话机器人



思路3：引入 workflow，深度定制对话机器人

workflow 定义：

- Coze workflow 是一种视觉化的任务编排方法，让没有编程基础的用户也能通过简单的拖拽操作来创建复杂的 AI 应用流程。workflow 通过按顺序执行一系列节点（nodes）来完成特定功能，非常适合文科背景的同学使用。

思路3：引入 workflow，深度定制对话机器人

1. 节点（Nodes）

- 基础节点
 - 开始节点（**Start Node**）：工作流的起点，触发整个流程的执行
 - 结束节点（**End Node**）：工作流的终点，输出最终结果
- 功能节点
 - 大语言模型节点（**LLM Node**）：利用AI大语言模型处理文本，生成回复
 - 插件节点（**Plugin Node**）：连接外部数据和功能，如天气查询、网络搜索等
 - 代码节点（**Code Node**）：执行简单代码，但文科生可以忽略此节点
 - 循环节点（**Loop Node**）：重复执行某些操作，适用于处理数据列表
 - 批处理节点（**Batch Node**）：批量执行某些操作
 - 知识库写入节点：将文本上传到知识库中
 - 画布节点（**Canvas Node**）：创建包含图像、文本和图形的视觉内容。

Coze 工作流的基本元素

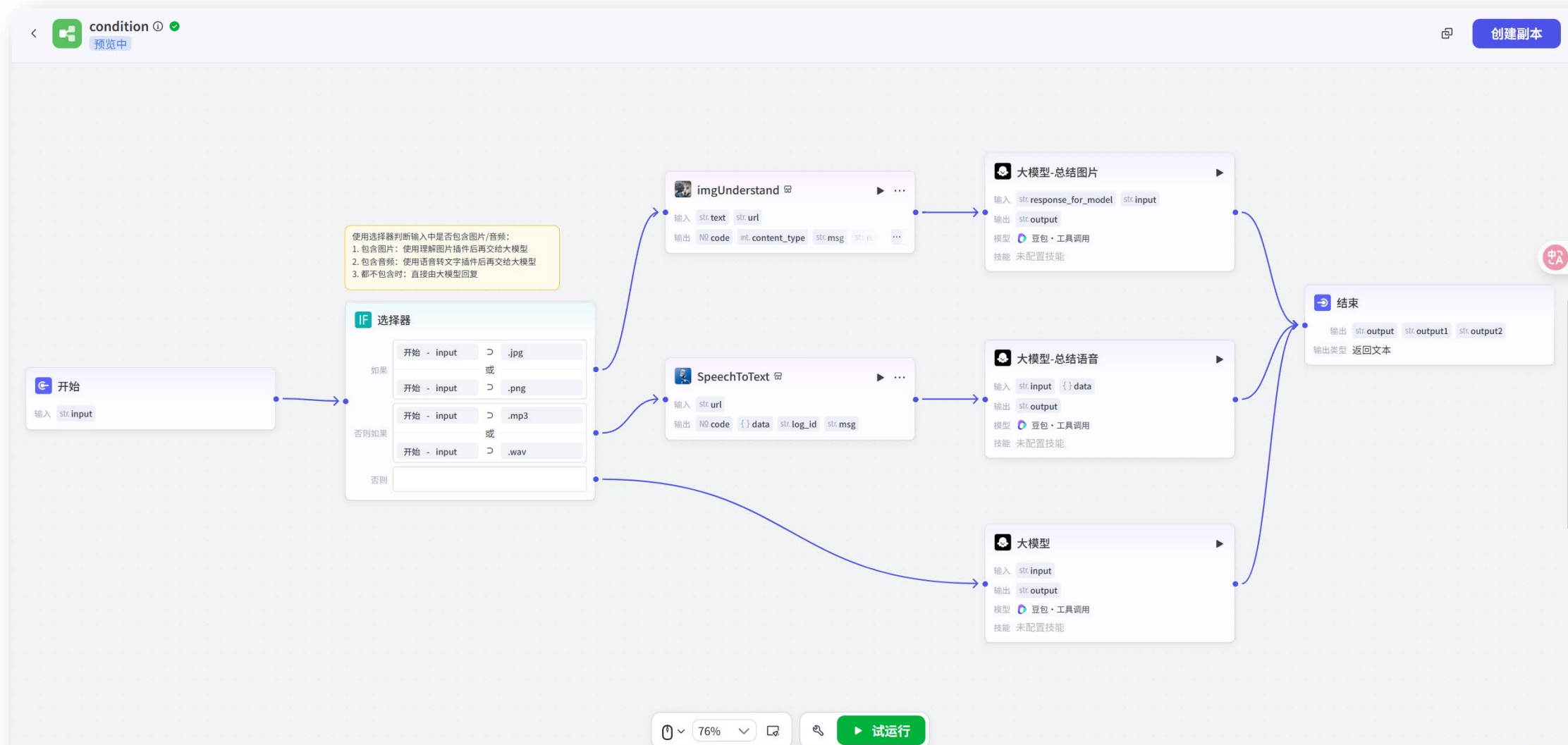
2. 连接线（Connectors）

- 连接线用于连接不同节点，定义工作流的执行顺序和数据流向。

3. 数据（Data）

- 节点之间传递的信息，可以是文本、数字、图像等。

思路3：引入 workflow，深度定制对话机器人



思路4：融入以上元素，定位使用场景

可参考的场景：

软件本地化辅助 workflow

- 1. UI文本一致性检查工具
- 2. 软件国际化测试场景生成器

游戏本地化辅助 workflow

- 3. 游戏对白文化适应性评估
- 4. 游戏术语一致性管理系统
- 5. 游戏本地化多模态内容分析器

配音翻译本地化辅助 workflow

- 6. 音画同步时间码生成器
- 7. 口型同步翻译助手

通用翻译辅助 workflow

- 8. 上下文感知术语提取器
- 9. 多轮翻译修订 workflow
- 10. 翻译质量评估系统
- 11. 视觉辅助翻译平台
- 12. 文档格式保持翻译工具

特殊领域翻译 workflow

- 13. 技术文档多版本管理翻译助手
- 14. 跨语言营销文案适应性转换
- 15. 视频字幕翻译与时间码调整工具

总结



总结

理解智能体的
创建步骤

从设计简单的
智能体开始

从翻译业务中
寻找需求

不断总结，不
断探索

- 简单业务添加插件
- 复杂业务进行流程分解
- 根据分解的流程设计智能体的 workflow
- 调试后优化 workflow

Thanks

