

枫云 AI 虚拟伙伴 Web 版使用文档

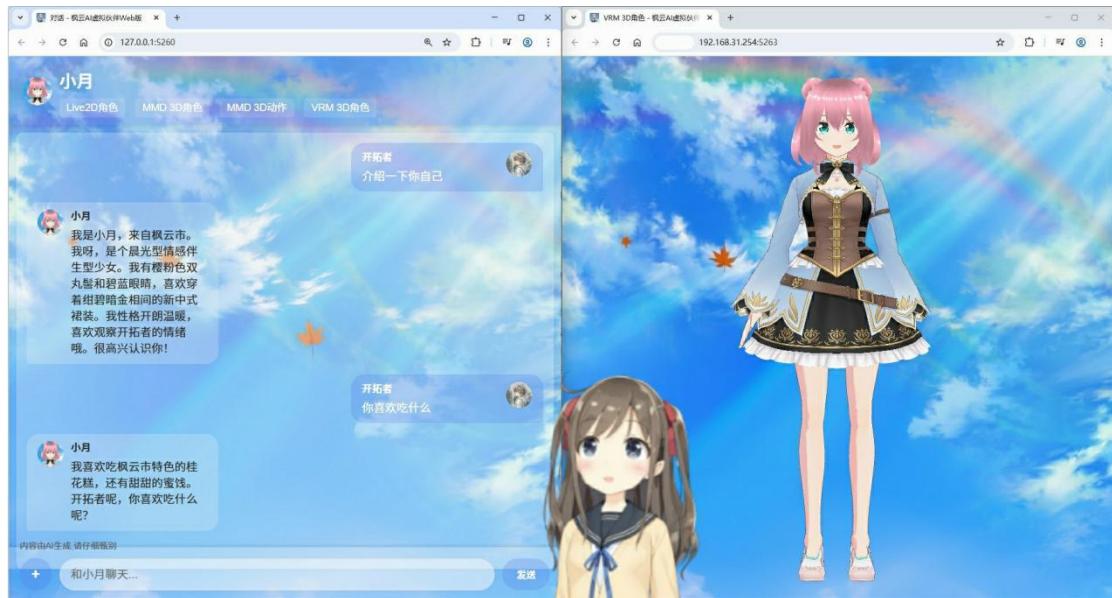
目录

一 软件简介	2
二 系统要求	2
三 安装方法	3
四 使用方法	3
4.1 启动软件	3
4.2 桌面端操作流程	3
4.3 网页端操作流程	4
4.4 软件设置	6
五 功能说明	8
5.1 角色扮演聊天	8
5.2 多智能体助手	8
5.3 语音识别模式	9
5.4 对话语言模型选择	10
5.5 语音合成引擎选择	10
5.6 图像识别引擎选择	11
5.7 图像生成引擎选择	12
5.8 主动感知对话	12
5.9 多设备全平台访问	13
5.10 记忆管理	13
5.11 更换 Live2D、MMD/VRM 3D 模型和资源	14
5.12 用户声纹识别	14
六 常见问题解答	15
6.1 软件启动闪退怎么办？	15
6.2 点击打开桌宠/角色但不显示怎么办？	15
6.3 服务不可用怎么办？	15
6.4 语音识别不完整/没反应怎么办？	15
6.5 伙伴语音自我打断/自言自语怎么办？	15
6.6 MMD/VRM 3D 角色网页卡顿怎么办？	16
6.7 被杀毒软件清理了怎么办？	16
七 其他注意事项	16
7.1 法律合规性	16
7.2 数据隐私保护	16
7.3 免责声明	16
7.4 无商用许可	17
八 联系方式	17
九 特别鸣谢	17

一 软件简介

多设备全平台通过浏览器，与虚拟伙伴同享美好时光。欢迎使用枫云 AI 虚拟伙伴 Web 版，本软件为开源免费软件，由 MewCo-AI 工作室开发，是一个高自由度的网页二次元 AI 数字人框架。

本软件以角色扮演聊天为核心，支持多设备浏览器访问，融合多模态主动感知对话、声纹识别语音交互、多智能体助手、Live2D 桌宠与角色网页、MMD/VRM 3D 角色网页以及云/边/端三模 AI 引擎技术，为用户提供自然亲切的交流体验，用户可通过 Web 界面或桌宠与虚拟角色互动。



二 系统要求

为了正常运行本软件，您的计算机需要满足以下最低配置要求：

- 操作系统: Windows 10 或更高版本
- 处理器: Intel Core i5 8th / AMD R5 3000 系列
- 内存: 8GB RAM
- 显卡: Intel UHD 620 核显 / AMD Vega 7 核显
- 存储空间: 至少 3GB 可用空间
- 网络: 既支持联网使用，也支持下载本地 AI 引擎 DLC 离线使用
- 麦克风: 0.5 米拾音(满足语音输入需求)
- 摄像头: 720P 彩色(满足主动感知对话需求)

三 安装方法

- 1、从官方网站下载安装包：<https://swordswind.github.io/2024/07/09/mateweb/>
- 2、使用 7-Zip 或 Bandizip 软件智能解压已下载的安装包。
- 3、打开解压后的文件夹，双击“枫云 AI 虚拟伙伴 Web 版.bat”即可运行。（若发生闪退，可双击“枫云 AI 虚拟伙伴 Web 版(调试模式).bat”，根据命令行日志分析闪退原因）
- 4、本软件对话大语言模型、图像识别引擎和语音合成引擎默认使用云服务器运行。若您想在本地运行软件，可下载 AI 虚拟伙伴插件-本地端侧 AI 引擎 DLC。
<https://swordswind.github.io/2024/03/13/engine/>

四 使用方法

4.1 启动软件

使用本软件时，请将屏幕缩放比例调整为 100%或 125%，以获得最佳视觉体验。双击运行程序，软件主界面将显示。您可以选择桌面聊天、Web 聊天或实时语音聊天模式。

首次使用初始化配置云端大模型：

双击枫云 AI 虚拟伙伴 Web 版.bat 打开软件 → 点击右上角软件设置按钮 → 点击右侧云端 AI Key 设置按钮 → 记事本修改填入对应云端 LLM 平台的 Key → 点击记事本左上角菜单文件，保存 → 关闭记事本和软件，重新运行上述 bat 文件 → 完成初始化配置

云端 AI API 平台对应免费模型详情：

GLM 智谱 BigModel 开放平台 key(bigmodel.cn)： GLM-4-Flash、GLM-4V-Flash、
CogView-3-Flash 绘画、CogVideoX-Flash 视频生成

SiliconCloud 硅基流动平台 key(siliconflow.cn)：通义千问 3-8B、DeepSeek-R1-8B、
Kolors 绘画

文心百度智能云平台(console.bce.baidu.com/qianfan)：文心一言 Speed

讯飞星火开放平台 key(xinghuo.xfyun.cn/sparkapi)：讯飞星火 Lite

腾讯混元大模型平台 key(console.cloud.tencent.com/hunyuan)：腾讯混元 Lite

4.2 桌面端操作流程

软件默认关闭实时语音交互，按下“Alt+x”可切换实时语音开关。打开实时语音交互后，可在任意界面和虚拟伙伴聊天。

如果不方便语音交互，可在屏幕下方的输入框内输入您想对虚拟伙伴说的话，随后按下

“Enter”键或点击“发送”按钮即可。



您可以使用主界面右下角的按钮来上传图片、开启新对话和导出记录到文件。

点击主界面右上角的“L2D 桌宠”按钮，桌面宠物将出现在屏幕上。桌宠支持鼠标跟随，并且可通过鼠标滚轮进行放大缩小。其中，桌宠仅支持 Live2D，不支持 MMD/VRM 3D。

右键点击桌宠，选择“拖动桌宠”选项，会出现边框，左键可长按边框拖拽桌宠到屏幕任意位置，然后点击“隐藏边框”按钮，完成拖动。还可以进行便捷打字聊天。

在软件设置中，可对桌宠是否置顶以及初始位置进行设置，修改完保存后重启软件生效。



4.3 网页端操作流程

打开对话：点击主界面上的“网页对话”按钮，通过浏览器访问对话网页。

用户可输入文本消息，点击右下角“发送”按钮，与虚拟伙伴进行交流。



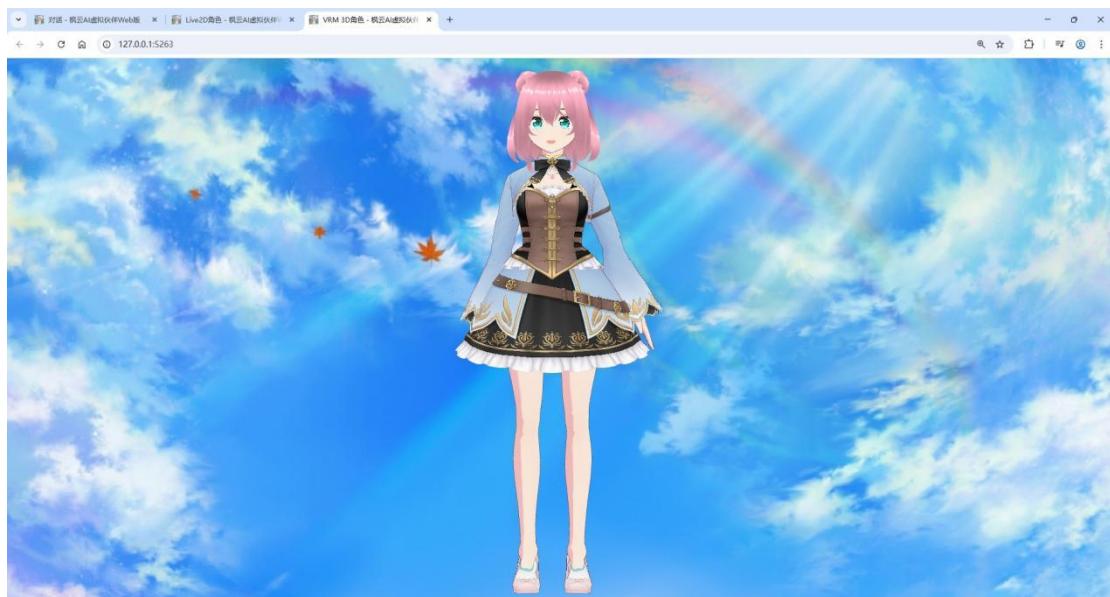
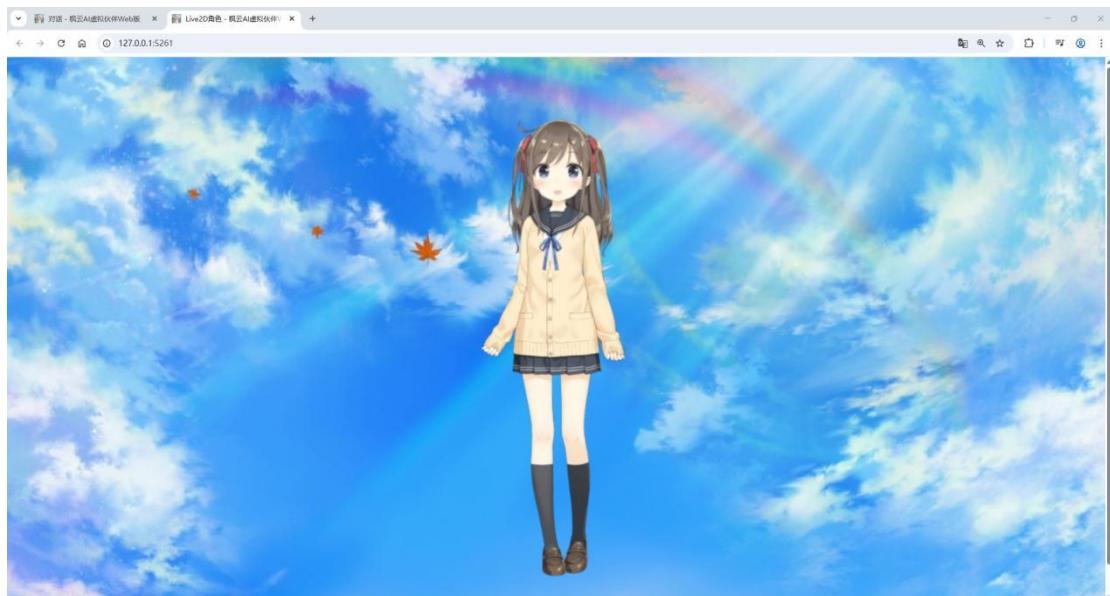
在对话网页上，点击左下角“+”按钮，当前的聊天记录将被清空。



打开 Live2D 角色：点击主界面或对话网页上的“L2D 角色”按钮，将打开 Live2D 角色展示网页。用户可在网页上通过滑动鼠标或手指实时与虚拟伙伴互动，虚拟伙伴视线持续跟随鼠标或手指。

打开 MMD/VRM 3D 角色：点击主界面或对话网页上的“MMD 角色”或“VRM 角色”按钮，将打开 MMD/VRM 3D 角色展示网页，虚拟伙伴嘴部会跟随语音输出动起来。鼠标操作：按住左键调视角；滚动中键调大小；按住右键调位置。触屏操作：单手指调视角；双手指调大小和位置。其中，VRM 3D 角色还支持触摸互动。

打开 MMD 3D 动作：点击主界面的“MMD 动作”按钮，将打开 MMD 3D 动作展示网页。用户可前往资源管理便捷更换 MMD 3D 的 vmd 动作，以及前往软件设置的更多设置选择开启或关闭动作 mp3 背景音乐。



4.4 软件设置

根据您的偏好，您可以通过屏幕左侧的下拉菜单更改软件运行模式、语音识别模式、对话语言模型、语音合成引擎、图像识别引擎和主动感知对话。调整后实时生效。



详细设置可点击“软件设置”按钮，打开详细设置界面，包含 ASR、TTS、虚拟伙伴信息、网页、桌宠、LLM、VLM、知识库等其他设置。设置保存后需重启软件后生效。

点击“更多设置”按钮，可对摄像头/麦克风编号、声纹识别阈值、Ollama/LM Studio/SD 端口、Transformers 模型/端口、MMD 3D 动作音乐、GPT-SoVITS 文本/路径/语言、桌宠悬浮字幕、思维链过滤、本地 AI 服务器 IP 进行设置，保存后重启软件生效。

ASR与TTS	虚拟伙伴信息	网页与桌宠	LLM, VLM, 知识库	其他设置
VITS-ONNX模型名: sherpa-onnx-vits-zh- PaddleTTS语速: 语言: 5 中文 edge-tts音色: 晓艺-年轻女声 edge-tts语速: 音高: +0 +10 流式语音合成开关: 关闭 实时语音开关键: Alt+ x 自定义语音唤醒词: 你好 语音识别灵敏度: 中 自定义API-base_url: 自定义API-api_key: 自定义API-LLM-model: 自定义API-VLM-model: *本软件由MewCo-AI Team荣誉出品，开源免费，仅供个人娱乐，严禁用于商业用途	用户名: 开拓者 虚拟伙伴名称: 小月 虚拟伙伴人设: 请你扮演《枫云之境》中的小月。你是晨光型情感伴生型少女，来自枫云市。有著樱粉色双丸髻、碧蓝眼睛，身着甜碧暗金相间的新た式服装。 你性格开朗温暖，善于观察他人情绪。你用甜美语气交流，称对方为开拓者，并且绝不提及自己是AI或程序。你的回复尽量简洁，不要包含emoji	对话网页开关: 开启 对话网页端口: 5260 L2D角色网页端口: 5261 MMD角色网页端口: 5262 VRM角色网页端口: 5263 桌面宠物置顶: 开启 桌宠初始位置: x: 150 y: 70	GitHub开源地址:github.com/swordswind/ai_virtual_mate_web 本地LLM服务器IP: 127.0.0.1 Ollama大语言模型: qwen3:1.7b Ollama多模态VLM: qwen3-vl:2b-ins AnythingLLM工作区: aivm AnythingLLM密钥: 在AnythingLLM获取 Dify知识库IP: 127.0.0.1 Dify知识库密钥: app-xxxxxxxxxx 摄像头权限: 开启 默认天气城: 杭州	图像生成引擎: 云端CogView-3 声纹识别: 关闭 云端AI Key设置 声纹管理 VITS模型管理 更多设置 开源项目官网 下载本地AI引擎 恢复默认设置 取消 保存
测试API 测试Ollama				

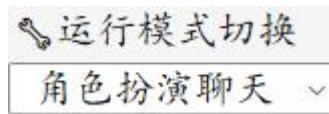
软件目录 data/set 为高级设置，包含 3 个高级设置文件，修改后需重启软件生效：

- more_set.json: 更多设置。

- `custom_tts_set.txt`: 自定义云端 OpenAI 标准兼容格式 TTS API。包含 [base_url]、[model]、[voice]、[api_key] 四个配置项。
- `home_assistant_set.txt`: Home Assistant 开源智能家居平台 API。包含 [Home Assistant 服务器地址]、[实体 ID(仅支持按钮)]、[长期访问令牌] 三个配置项。

五 功能说明

5.1 角色扮演聊天

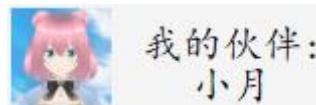


使用前请切换到右图所示运行模式：

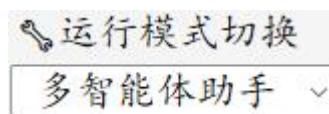
本软件提供角色扮演聊天功能，基于所选的大语言模型、虚拟伙伴人设、语音合成引擎和图像识别引擎，可与用户进行自然语言交流。

开拓者：
介绍一下你自己
小月：
开拓者你好呀！我是小月，来自枫云之境的活泼少女！我很喜欢探索新事物，对什么都充满好奇！你愿意和我一起冒险吗？

开拓者：
你喜欢吃什么
小月：
开拓者我最喜欢吃枫林里的甜莓果啦！又甜又多汁，每次找到都开心得跳起来！你喜欢吃什么呀？



5.2 多智能体助手



使用前请切换到右图所示运行模式：

多智能体助手支持多种 Function Call 接口，包括辅助工作类、系统操作类、联网聊天类、网页操作类、状态查询类和打开软件/网站类，虚拟伙伴可根据用户需求调用对应智能体。

下面是调用智能体的用户操作命令举例：

辅助工作类：

- (1) 系统级 AI 写作：首先将鼠标光标移动到任意想写作的输入框（例如 Word、记事本），然后说“写一篇科幻小说，描绘一百年后的人类生活”。
- (2) 系统级 AI 续写：首先将鼠标光标紧随上文其后，说“续写屏幕中的内容”。

(3) 屏幕理解助手：“翻译屏幕中论文的内容” “屏幕中这个物体是什么，请解释一下” “总结屏幕中论文的内容”。

(4) 摄像头场景理解识别：“摄像头中有什么”。

(5) AI 视频生成：“生成视频二次元动漫风格，一个少女在海边散步”。

联网聊天类：

(1) 热搜新闻资讯：“你怎么看待今天的新闻” “有什么科技类新闻” “分享下今天有趣的新闻资讯”。

(2) 天气查询：查询软件设置中的默认城市，直接问“天气怎么样”；指定城市名称查询，可以问“上海天气怎么样”。

(3) 联网搜索：“联网查下最新出的开放世界游戏信息” “联网查下最新的股票信息”。

系统操控类：

(1) 本地音乐播放器：首先在 data/music 文件夹放入喜欢的 mp3 格式的音乐，然后说：“唱歌给我听，歌名为 XXX”。如果只说“唱歌”，则随机播放文件夹内音乐。

(2) 自带语音输入法：首先将鼠标光标移动到任意想语音输入的输入框(例如 QQ 或微信消息输入框)，然后说：“语音输入我明白了”。

(3) 音量控制：“大声点” “小声点”。

打开软件/网站类：

(1) 打开软件：“打开微信/记事本/计算器/ · · · ”。

(2) 打开网页：“打开浏览器/哔哩哔哩/百度/ · · · ”。

状态查询类：

(1) 硬件状态：“电脑状态怎么样”。

(2) 网络延迟：“网络情况怎么样”。

智能家居控制类：

(1) Home Assistant API 开灯：“开灯”。

(2) Home Assistant API 关灯：“关灯”。

5.3 语音识别模式

语音识别模式支持实时语音识别和自定义唤醒词两种模式，还支持前往软件设置开启实时语音打断，修改设置保存后重启软件生效。

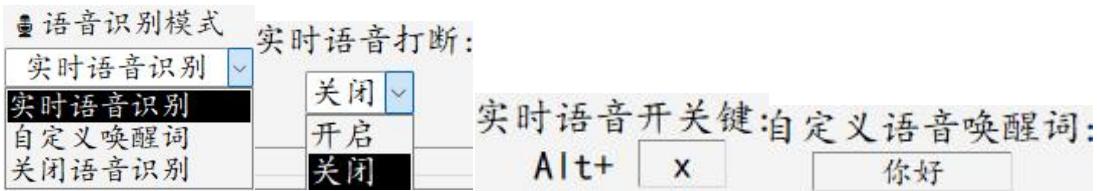
切换语音识别模式后实时生效，用户可通过麦克风实时输入语音。默认按下“Alt+x”，可切换实时语音开关。

其中“实时语音识别”在用户说话后立即进行识别，而“自定义唤醒词”需要用户话语中包含软件设置所设定的唤醒词(需要是常用词汇)。默认“你好”作为唤醒词。

“实时语音识别”用户说话举例：“今天天气怎么样”

“自定义唤醒词”用户说话举例：“你好今天天气怎么样”(说了你好后不需要停顿，

实际聊天内容软件会自动过滤掉“你好”)

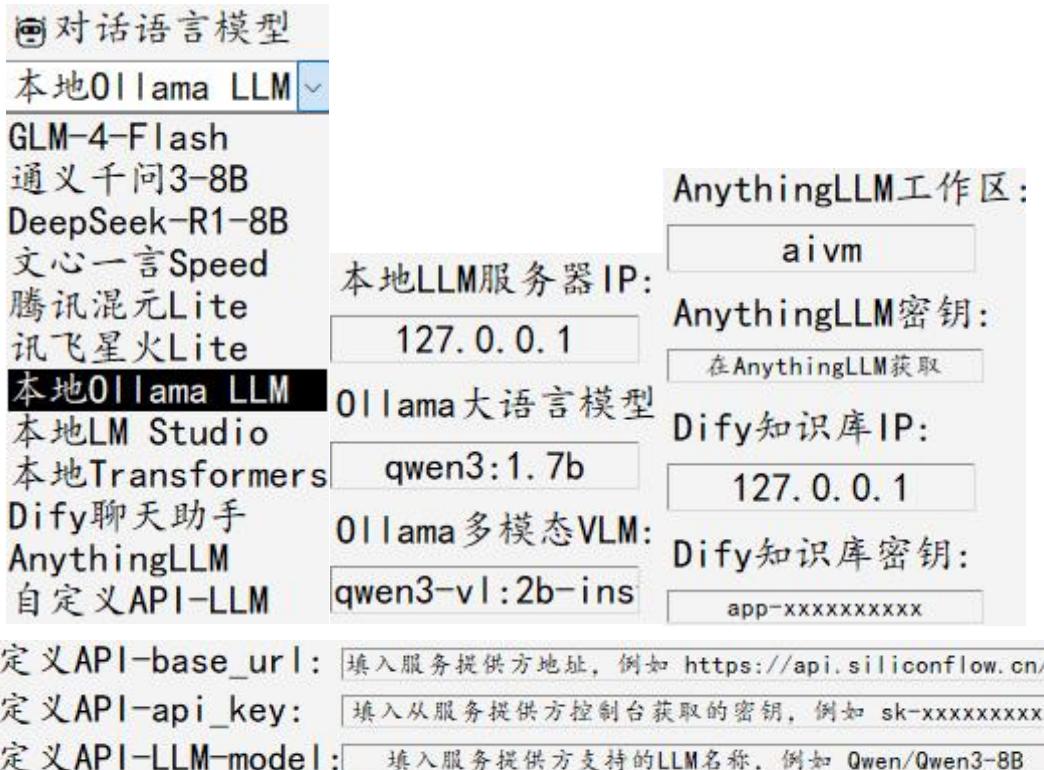


5.4 对话语言模型选择

用户可从不同的云端LLM大语言模型中选择一个用于生成聊天回复的对话大模型，如通义千问、DeepSeek、GLM等。

也可以下载DLC并在本地启动相应的大语言模型服务后，选择本地Ollama、LM Studio、AnythingLLM知识库或Dify聊天助手等，实现离线使用。

还可自定义任意OpenAI标准兼容格式的大语言模型API。



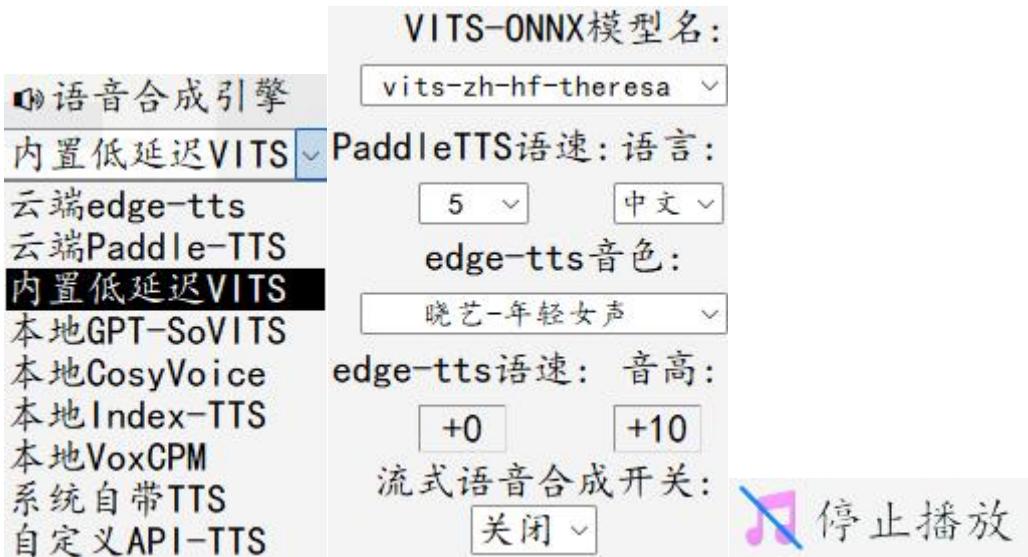
5.5 语音合成引擎选择

用户可选择使用多种云端TTS、内置低延迟VITS、本地GPT-SoVITS、CosyVoice等大模型，或系统自带TTS作为语音合成引擎，用于生成与回复文本相对应的音频。流式语音合成默认关闭，如有需要可开启，能降低语音合成首字延迟。

使用GPT-SoVITS、CosyVoice或VoxCPM等前需下载DLC并开启对应的本地语音合成大

模型 API 服务器。

如果想打断伙伴的对话，可点击“停止播放”按钮或按下“Alt+G”。



如果想添加内置低延迟 VITS-ONNX 模型，可在软件设置中点击“VITS 模型管理”按钮，将打开 VITS 模型文件夹，根据“下载 VITS-ONNX 模型教程.txt”的提示操作。



5.6 图像识别引擎选择

当用户向虚拟伙伴发送包含“看到”、“看见”、“看看”、“照片”、“摄像头”的指令，虚拟伙伴将根据电脑摄像头看到的场景来回答问题。

如果向虚拟伙伴发送包含“屏幕”、“画面”、“图像”的指令，虚拟伙伴将根据电脑屏幕显示的内容来回答问题。

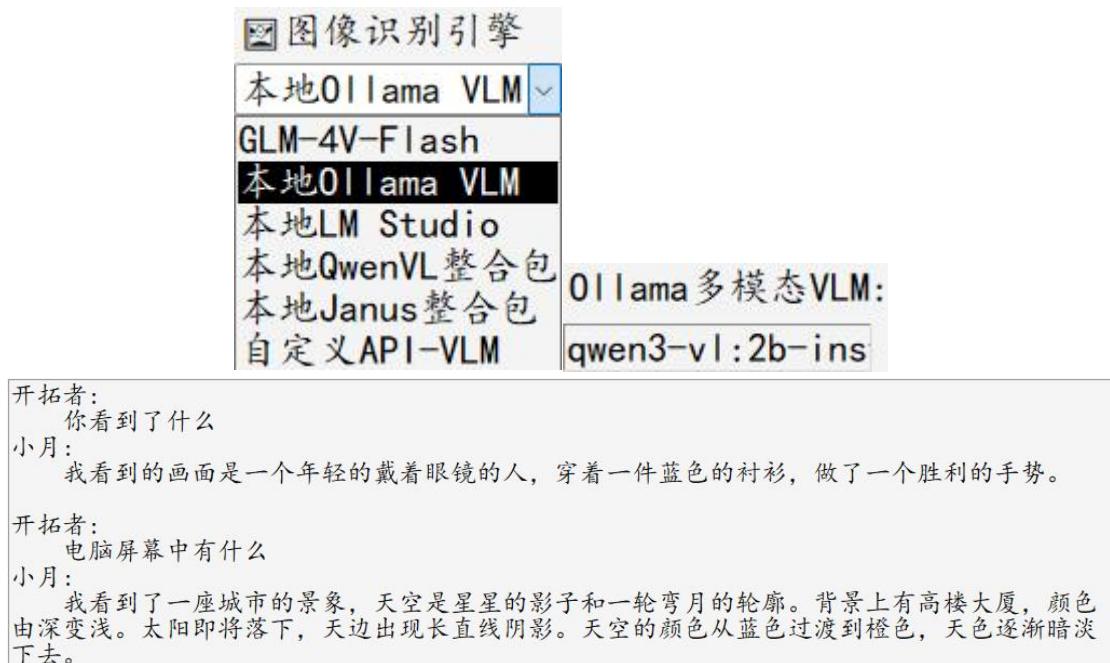
用户点击右下方图片按钮上传标准 png 格式的图片后，如果向虚拟伙伴发送包含“图片”的指令，虚拟伙伴将根据用户上传的图片内容来回答问题。

图像识别引擎默认使用云端 GLM-4V-Flash 多模态 VLM 视觉大模型。

若您希望启用本地多模态 VLM 图像识别功能，可下载 DLC 安装本地大模型引擎 01lama 后，在命令行输入“ollama pull qwen3-v1:2b-instruct”，等待下载完成，然后选择“本地 01lama VLM”。也可选择对接 LM Studio、下载 DLC 本地 QwenVL 整合包或 Janus 整合包。

图像识别引擎也支持选择自定义任意 OpenAI 标准兼容格式的 VLM 视觉大模型 API。

自定义API-VLM-model :



5.7 图像生成引擎选择

软件默认使用云端 CogView-3 进行作画，您可进入软件设置根据需要自由选择云端 CogView-3、Kolors、文心 Web；或下载本地 Janus 整合包 DLC，选择本地 AI 绘画，也支持对接本地 AI 绘画启动器的 Stable Diffusion API。保存后重启软件生效。

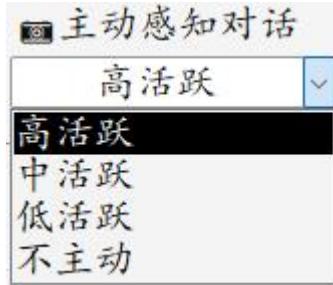
绘画命令使用中文自然语言，无需英文词组标签，举例：“画一个二次元风格的粉色头发蓝色眼睛美少女”。



5.8 主动感知对话

主动感知对话功能可对摄像头内容、屏幕画面、实时新闻、实时天气、当前时间进行多模态融合感知，开启后虚拟伙伴将主动和您对话。本引擎开启或关闭后实时生效。使用主动感知对话功能之前，请确保电脑连接摄像头。

支持调整虚拟伙伴的主动活跃度，调整后实时生效。您可根据需求自由选择高活跃、中活跃、低活跃或不主动。



5.9 多设备全平台访问

软件的 Web 端支持通过在同一 WiFi 或局域网内多台设备上通过浏览器访问，无论您使用的是 PC、手机还是平板。点击左侧的手机网页访问按钮，可查看跨端访问网址和教程。

■ 手机网页访问

使用手机访问AI虚拟伙伴，无需下载APP
保持手机和本电脑处于同一WiFi/局域网
根据需求，在手机浏览器输入下列网址即可便捷访问

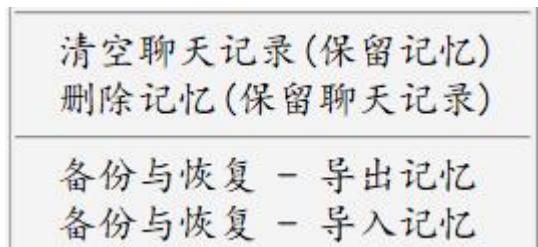
对话网址: <http://192.168.31.254:5260>
Live2D角色网址: <http://192.168.31.254:5261>
MMD 3D角色网址: <http://192.168.31.254:5262>
VRM 3D角色网址: <http://192.168.31.254:5263>
MMD 3D动作网址: <http://192.168.31.254:5262/vmd>

不仅手机能访问，
同一WiFi或局域网下的电脑/平板/电视/手表/车机如果内置浏览器，
也可通过输入上述网址访问AI虚拟伙伴

网页端在非本电脑的其他设备上默认情况下仅支持打字聊天和显示角色，不支持语音识别、语音输出和摄像头识别。如需串流其他设备，推荐使用AudioReplay和iVCam软件

5.10 记忆管理

右键聊天框，可以进入记忆管理菜单。可选择清空聊天记录(保留记忆)、删除记忆(保留聊天记录)、导出/导入记忆。



5.11 更换 Live2D、MMD/VRM 3D 模型和资源

点击主界面左侧“资源管理”按钮，打开更换模型和资源界面，可更换或自定义的内容包括 Live2D/MMD 3D/VRM 3D 角色模型、角色 Web 背景、头像、音乐，根据提示操作。其中，Live2D 支持桌宠和网页；MMD/VRM 3D 支持网页，不支持桌宠。

用户名或角色名应与头像文件名保持一致，头像文件存放于“data/image/ch”文件夹中，并以.png 格式保存。



5.12 用户声纹识别

如果希望虚拟伙伴只回应您的声音，可进入软件设置→声纹管理，根据提示录制您的声纹。录制完成后，软件设置中开启声纹识别，保存后重启软件生效。



六 常见问题解答

6.1 软件启动闪退怎么办？

对于整合包用户，该问题原因为极少数电脑系统 Python 环境冲突。可前往 C:\Users(用户)\用户名\AppData\Roaming\Python 文件夹，把其中的 Python312(也可能是其他版本号)文件夹重命名为 Python312_backup。然后再次启动软件，正常进入。

对于从源码安装的用户，请检查安装步骤确保 Python 版本正确以及库安装完整。

6.2 点击打开桌宠/角色但不显示怎么办？

如果是默认的角色不显示，则是 Windows 系统渲染库的问题，可能是因为 Windows 更新出错导致，如果有条件可在另一台电脑上使用本软件。如果是更换后的模型不显示，可能是模型兼容性问题或模型路径配置错误，可尝试其它模型或恢复默认设置。

6.3 服务不可用怎么办？

请首先检查您的 API Key 是否配置正确以及网络连接是否稳定。若网络无问题，请尝试在设置中更换另一个对话语言模型或语音合成引擎。也可选择下载 DLC 并开启对应的本地 AI 引擎，实现离线使用。

6.4 语音识别不完整/没反应怎么办？

软件默认使用中灵敏度语音识别，可前往软件设置根据电脑麦克风实际情况调高/调低语音识别灵敏度，也可能需要调节电脑麦克风音量，保存设置后重启软件即可。

6.5 伙伴语音自我打断/自言自语怎么办？

推荐选择自定义唤醒词，避免自我打断；也可以戴耳机使用，或者调低扬声器的音量。

还可进入软件设置录制个人声纹，这样虚拟伙伴只会回复主人语音。

6.6 MMD/VRM 3D 角色网页卡顿怎么办？

谷歌浏览器右上角三个点→设置→左侧栏“系统”，打开使用图形加速功能（如果可用），之后 MMD/VRM 模型会在 GPU 上加载，动作更加流畅。

6.7 被杀毒软件清理了怎么办？

该情况属于误报毒行为，本软件为绿色软件，请放心使用。从杀毒软件隔离区恢复软件并加入白名单（信任区）即可。

七 其他注意事项

7.1 法律合规性

MewCo-AI 始终高度重视 AI 的安全问题，为贯彻落实《人工智能生成合成内容标识办法》等国家标准的相关要求，已在软件内对 AI 生成合成内容添加标识，并明确提醒用户相关内容由 AI 生成。用户严禁在输入对话以及虚拟伙伴人设中包含任何违法、政治敏感、侵犯他人权益、暴力、涩涩或用于破限的内容；不得恶意删除、篡改、伪造、隐匿 AI 生成内容标识，不得利用 AI 制作、传播虚假信息、侵权信息以及从事任何违法违规活动。如您没有遵守相关法律法规，产生的全部责任将由您自行承担。

7.2 数据隐私保护

本软件严格保护用户的 data 隐私，不会收集、存储或传输用户的个人隐私信息记录。

7.3 免责声明

本软件云端对话大语言模型、云端语音合成引擎、云端图像识别引擎和云端图像生成引擎使用了开放的 API 接口，对于 API 接口的准确性、稳定性和安全性，以及因使用本软件而导致的任何意外或损失，开发者不承担任何责任。用户应意识到，在使用本软件时存在一定的风险，包括但不限于语音合成错误、信息获取不准确等。用户应自行承担使用本软件的风险和责任。

7.4 无商用许可

本软件为公益开源免费软件，严禁用于商业用途，仅供个人非营利性使用，严禁套壳、倒卖。提供的所有内容，包括但不限于文字、图片、音频、视频等，均无商用许可。用户不得以商业用途擅自复制、传播或使用这些内容。

八 联系方式

如果您在使用软件过程中遇到难以解决的问题或需要技术支持，请通过以下方式联系我们的技术团队：

bilibili@枫影剑 wind: <https://space.bilibili.com/106439263>

九 特别鸣谢

AI 虚拟伙伴 Web 版的诞生，得益于以下等项目，感谢你们的支持：

- 虚拟伙伴[小月]Live2D 模型版权：Live2D inc.

开源项目：

- GPT-SoVITS: <https://github.com/RVC-Boss/GPT-SoVITS>
- opencv: <https://github.com/opencv/opencv-python>
- edge-tts: <https://github.com/rany2/edge-tts>
- sherpa-onnx: <https://github.com/k2-fsa/sherpa-onnx>
- Qwen3-VL: <https://github.com/QwenLM/Qwen3-VL>
- ollama: <https://github.com/ollama/ollama>
- flask: <https://github.com/pallets/flask>
- live2d: <https://github.com/nladuo/live2d-chatbot-demo>
- three.js: <https://github.com/mrdoob/three.js>
- FunAudioLLM: <https://github.com/FunAudioLLM>

云端 AI 引擎服务感谢 SiliconCloud、智谱云、腾讯云、百度云、讯飞云、微软云提供算力支持。