

GLM-PC 用户手册

GLM-PC（牛牛）是什么？

GLM-PC（牛牛）是一款革新性的智能桌面助手，基于先进的 CogAgent 大模型，不仅具备视觉理解能力，更拥有类人的深度思考和任务规划能力。它能够像人类一样理解、分析并执行复杂的计算机操作任务，包括但不限于发送微信消息、预订机票、网页浏览获取信息等。

GLM-PC 的核心优势在于其强大的认知决策能力：

- 深度场景理解：能够准确理解屏幕界面的上下文信息，洞察任务执行的关键要素
- 智能任务规划：可以制定最优执行路径，并在遇到障碍时动态调整策略
- 自适应学习：具备自我纠错和优化能力，能根据实时反馈调整执行方案
- 上下文推理：在执行过程中持续评估环境变化，做出符合用户意图的智能决策

通过简单的自然语言指令，GLM-PC 能够将用户的需求转化为一系列精确的操作步骤，在执行过程中展现出类人的思考模式和决策能力，为用户提供真正智能化的计算机操作体验。

注意事项

GLM-PC 是一个测试中的产品，在使用计算机和互联网交互时，这些风险会加剧。在使用过程中为了最大程度降低风险，请考虑采取以下预防措施：

1. 使用具有最小权限的专用虚拟机或容器，以防止直接系统攻击或事故。
2. 避免让模型访问敏感数据（例如帐户登录信息），以防止信息被盗。
3. 将互联网访问限制在允许的域列表中，以减少对有害内容的暴露风险。
4. 要求人类对可能导致有意义的现实后果的决定以及任何需要肯定同意的任务进行确认操作，例如接受 cookie、执行金融交易或同意服务条款。

安装指南

- 现阶段仅支持 M 系列 Mac 电脑，推荐 M1 和 M3 芯片的系列
- Mac 支持的操作系统 14 及以上，芯片支持 M1 M2 M3，推荐 M1 和 M3，浏览器建议使用 Chrome、Safari、Edge 和 Firefox，推荐显示器分辨率 1512×982
- Windows 支持的操作系统 10 及以上，浏览器建议使用 Chrome、QQ、Edge 和 Firefox，推荐显示器分辨率 2560×1600，缩放 175%，或显示器分辨率 2880×1800，缩放 200%

下载链接：<https://cogagent.aminer.cn/static/pkg/GLM-PC-arm64.dmg>

安装引导：[安装指南](#)

- 远程遥控（仅支持极速模式指令）：现阶段支持一台手机设备和一台电脑设备来使用应用；在手机设备上只能通过浏览器中输入 <https://cogagent.aminer.cn/m> 并登陆才能给应用发消息，主要的功能是远程操作和定时任务；在发消息前需要把应用设置为远程遥控模式，目前仅极速模式可支持远程遥控功能。

使用指南

1. 在 GLM-PC 对话框中输入您想进行的操作
2. GLM-PC 会评估您计算机上的工具并决定操作计划
3. GLM-PC 在新建对话的时候可以选择极速模式和深度思考模式，极速模式注重高效性能，深度思考模式支持将复杂任务，可根据任务指令的复杂程度选择模式；



极速模式



深度思考模式

4. GLM-PC 在计算机中使用工具操作，直至完成任务

- 在 GLM-PC 操作过程中，计算机屏幕上会显示控制条，用户可以随时点击停止键、暂停键；停止键会结束当前任务；暂停键可以暂停当前的任务操作，可以点击继续键从当前界面继续任务
- 当遇到敏感操作时，包括但不限于在登陆界面、信息发送界面、支付界面、文件删除界面，GLM-PC 会自动暂停，等待用户操作或者进行确认，用户需要点击控制条上的继续按钮继续任务
- 如需关闭敏感操作提示，可在个人中心里点击敏感操作白名单的“确认”按钮



5. GLM-PC 执行任务有失败概率，GLM-PC 在遇到错误提示后会自己进行重试，如果依然不成功，您可以修改指令再试一次

极速模式功能推荐

CogAgent 作为通用视觉语言模型，理论上可以跟用户一样进行任意计算机操作，为了贴近用户的使用需求，我们针对一些特定场景进行了优化，包括但不限于：

即时通讯

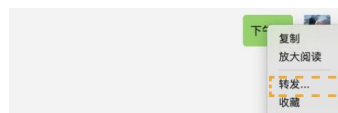
- 微信、飞书、钉钉、邮件发送和转发消息，包括群聊和联系人



使用飞书发送消息



使用微信发送消息

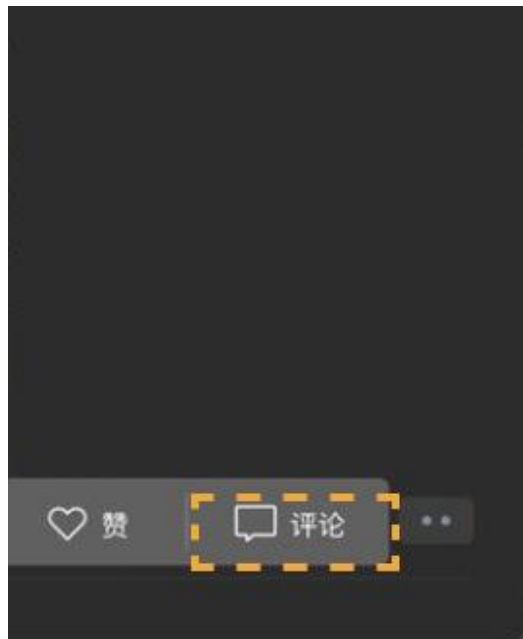


使用微信转发消息

2. 微信朋友圈点赞、评论

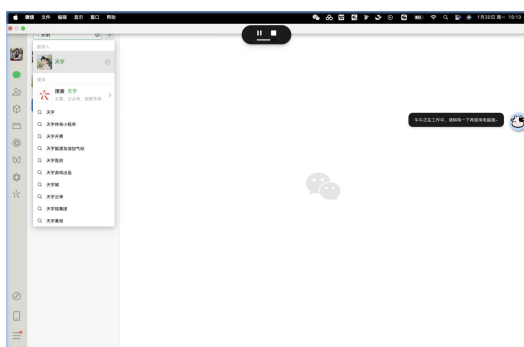


微信给朋友圈 XXX 点赞

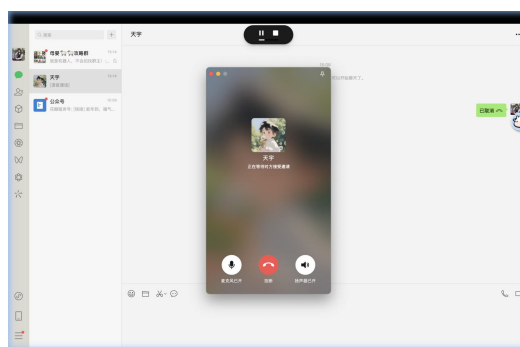


微信给朋友圈 XXX 评论

3. 微信、飞书、钉钉发表情包，视频和语音通话



微信搜索联系人



用微信打语音通话

会议

1. 预定腾讯、飞书和钉钉会议，并把会议邀请发送给参会人

开始 2024/11/18 周一 12:00

时长 30分钟

最多支持100人参会，会议人数超过2人时，启动60分钟限时。[升级能力](#)

☐ 会中时长不足时自动使用加时卡 ①

时区 (GMT+08:00) 中国标准时间 - 北京

重复 不重复

日历 ☒ 添加到日历

☐ Outlook ☐ iCal ☒ 其他日历

邀请 成员

请输入成员名称

安全 ☐ 入会密码

☐ 开启等候室

☒ 允许成员在主持人开会前加入会议

预定

预定腾讯会议



并把会议邀请发送给飞书 xxx

2. 参加腾讯、飞书和钉钉会议，并开启录制和会议纪要功能

☐ 当前设备始终使用此名称入会

会议设置

☒ 自动连接音频

☐ 入会开启摄像头

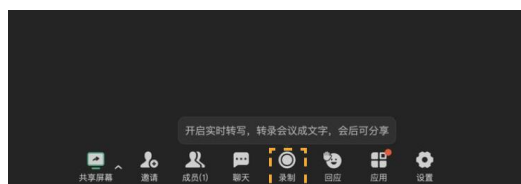
☒ 入会开启麦克风

☒ 入会开启美颜

[点此查看和设置美颜效果](#)

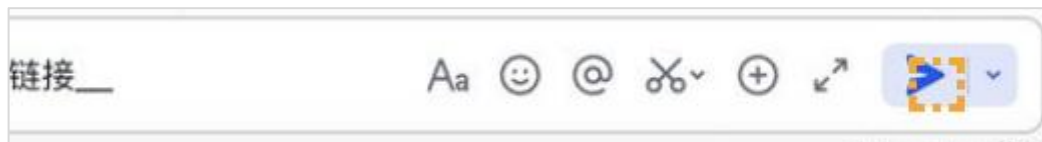
加入会议

参加腾讯会议



并开启录制

3. 把会议纪要发送给参会人



文档上传下载和保存

1. 下载互联网上的文档，保存在本地



下载互联网上的文档，保存在本地

2. 上传文档到微信、飞书、钉钉聊天框



上传文档 XXX 到 XXX 聊天框

文档处理

1. 打开文档



打开微信和 XXX 聊天框中的文档

2. 文档快速阅读、图片理解、文本引用、内容总结、内容解释、内容翻译



快速阅读 xxx 文档

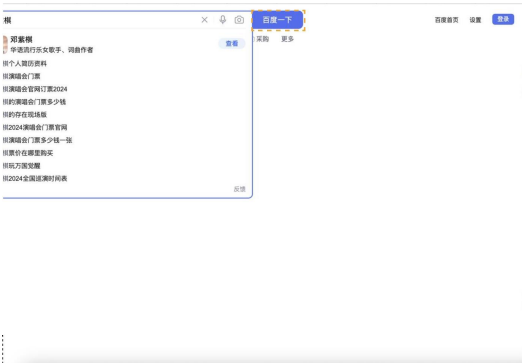
3. 支持 Office(word, ppt, excel) 、WPS、PDF、腾讯文档、飞书文档、钉钉文档



打开飞书 xxx 聊天框的文档

浏览器操作

1. 打开浏览器，跳转到指定 Url



用百度搜索邓紫棋



用百度搜索薛之谦的歌曲

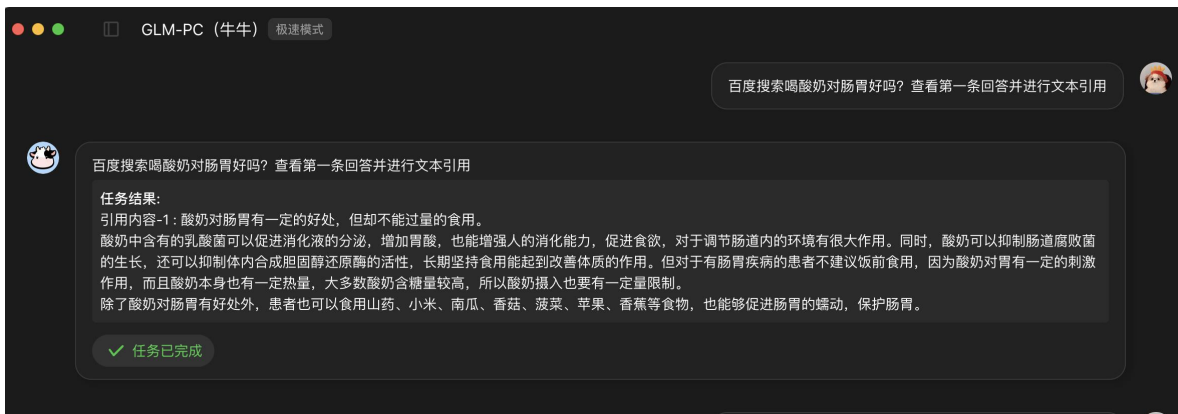


百度搜索喝酸奶对肠胃好吗



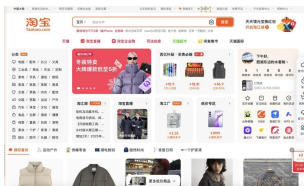
打开指定网址搜索

2. 快速阅读网页，阅读后进行图片理解、文本引用、内容总结、内容解释、内容翻译

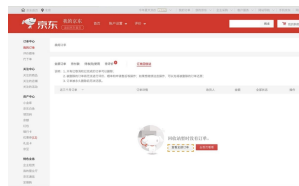


3. 功能类网页操作

- a. 购物类网页的操作：包括淘宝、京东的商品选购、店铺查找、评价撰写、评价总结、购物车操作、收藏操作等



用淘宝查看收藏夹



用京东查看订单历史



用淘宝查看收货地址

- b. 行程规划类网页的操作：携程、飞猪、美团的查车次、查酒店、查门票、查机票、订车票、订酒店、订门票、订机票、行程规划、查询周边、评价生成等



用携程查北京到广州的机票



携程查北京的酒店

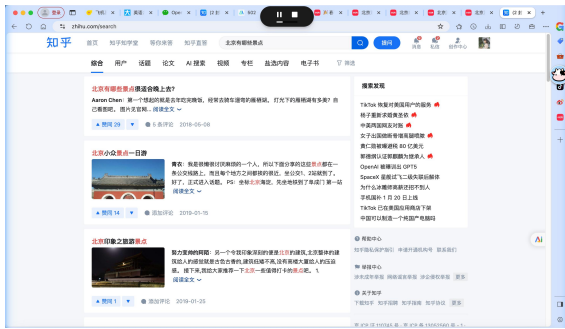


用飞猪查北京到上海的机票

4. 信息查询类网页操作（更多见 agent 搜索）

Agent 搜索（支持文字和图片引用）

1. 打开搜索引擎（支持百度、微信公众号、知乎、小红书、多个 AI 搜索平台等）



用知乎搜索北京有哪些景点



微信搜索公众号天安门广场

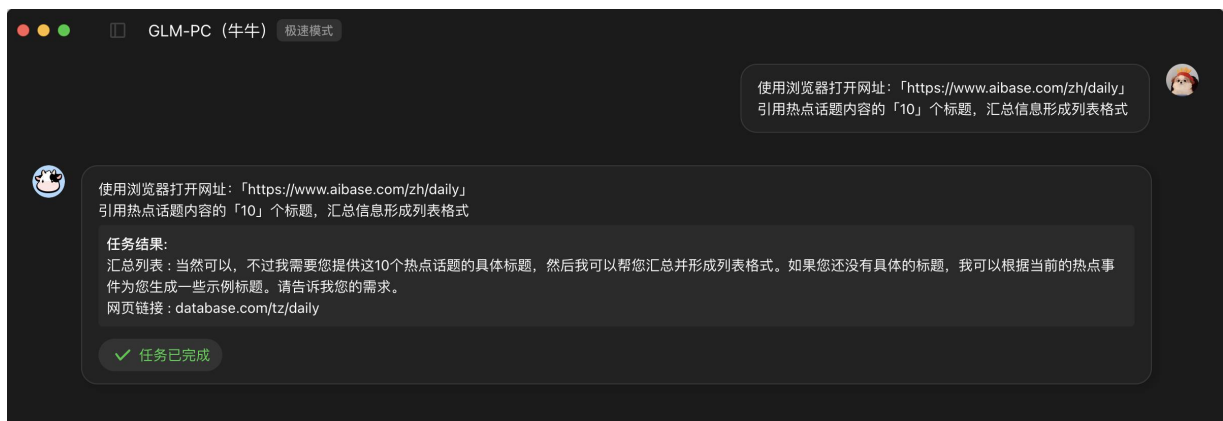


用百度搜索 2024 年全国 GDP 排名



小红书搜北京热门滑雪场

2. 快速阅读链接页面的内容，阅读后进行图片理解、文本引用、内容总结、内容解释、内容翻译



3. 搜索结果的来源引用展示



电脑设置

1. 调整显示模式（夜间模式/白天模式）



收藏任务

GLM-PC 任务执行成功后，可以点击“收藏任务”把任务指令加入收藏，一键点击即可快速触发



远程遥控

第一步：在设置中打开 GLM-PC 的远程遥控模式



第二步：手机上登陆 <https://cogagent.aminer.cn/m>，可以远程给 GLM-PC 发指令消息，让 GLM-PC 进行电脑操作。GLM-PC 每执行一步会返回操作时的屏幕截图。如果有敏感操作，会等用户进行确认后再操作



远程给 GLM-PC 发指令消息



GLM-PC 每执行一步会返回操作时的截图

定时任务（仅限极速模式）

手机上登陆 <https://cogagent.aminer.cn/m>，可以给 GLM-PC 发指令消息，让 GLM-PC 在指定的时间进行电脑操作。GLM-PC 每执行一步会返回操作时的屏幕截图。如果有敏感操作，会等用户进行确认后再操作



让 GLM-PC(牛牛)遥控器在指定的时间进行电脑操作



敏感操作，等用户进行确认后再操作

深度思考模式功能推荐

深度思考模式支持极速模式下的所有能力，但对于任务，深度思考模式下，GLM-PC 会花更多时间进行任务分解、制定计划 and 自我反思，整个任务所需时间更长。以下功能是深度思考特有的能力：

思考和问答

深度思考模式支持跟用户的多轮对话，用户可以直接提问、可以对牛牛的回复进行追问。



主动提问

如果用户的指令缺少执行任务所需要的信息，牛牛会发起主动提问



自我纠错和尝试

如果牛牛最初的执行计划遇到了阻碍，会尝试自我纠错并用新的计划进行再次尝试

使用牛顿法解方程： $x^3-x^2-x-1=0$



思考

1. 用户需要使用牛顿法求解方程 $x^3-x^2-x-1=0$ 。
2. 牛顿法是一种迭代求解方程的方法，需要方程的导数。
3. 计算方程 $f(x) = x^3 - x^2 - x - 1$ 的导数 $f'(x)$ 。

结果

无

思考

1. 根据observation信息可知，目前缺少async function main函数定义或调用，缺少await LOG日志输出。
2. 重新生成包含async function main函数定义和调用的代码，并确保有await LOG日志输出。

结果

方程的根是：
NaN

思考

1. 根据observation信息可知，方程的根是NaN，表示计算过程中出现了问题。
2. 检查代码，可能是初始猜测值x0或迭代次数maxIterations不合适，导致未找到收敛解。
3. 尝试调整初始猜测值和迭代次数，重新进行计算。

结果

方程的根是：
1.8392868100680193

思考

1. 根据observation信息可知，已经成功计算出方程的根。
2. 用户需求已完成，可以输出结束标记。

✓ 任务已完成

循环、多步推理和长链条任务

GLM-PC（牛牛）具备出色的复杂任务处理能力，特别擅长执行以下类型的操作：

(1) 循环任务处理

- 自动处理重复性操作，如批量文件处理、数据录入等
- 智能识别循环模式，无需手动设定重复次数
- 在循环过程中保持高精度和稳定性

(2) 多步推理执行

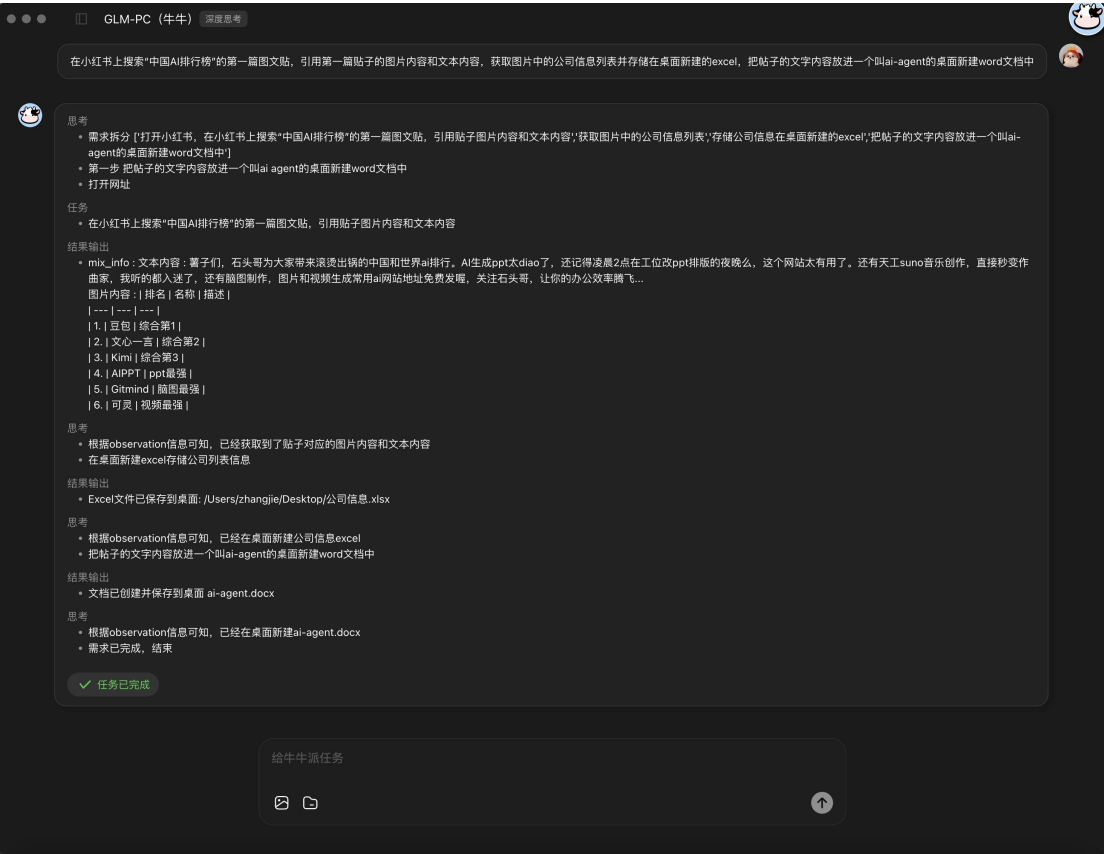
- 将复杂指令分解为清晰的执行步骤
- 根据上下文动态调整执行策略

- 在任务执行过程中进行实时决策优化

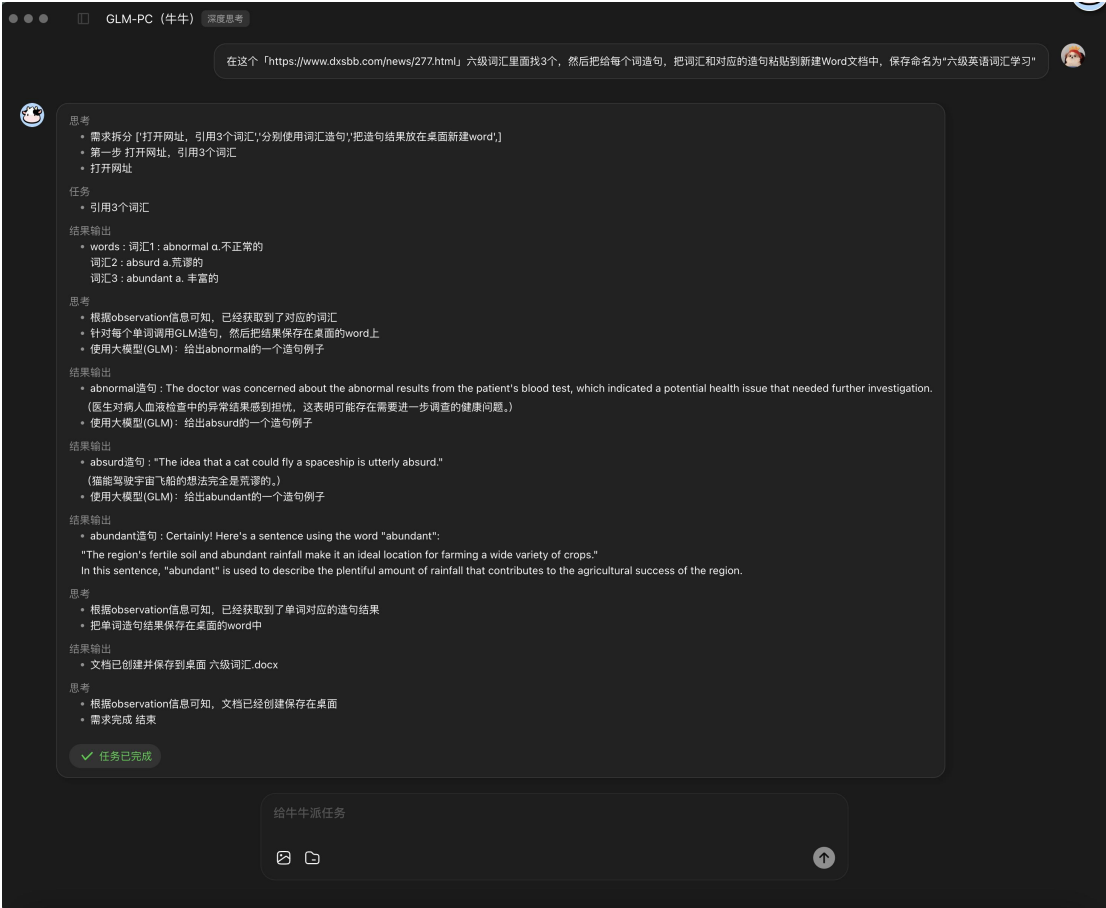
(3) 长链条任务管理

- 可靠处理涉及多个应用程序的复杂 workflow
- 准确追踪和管理长序列操作步骤
- 即使在长时间任务中也能保持注意力和准确性

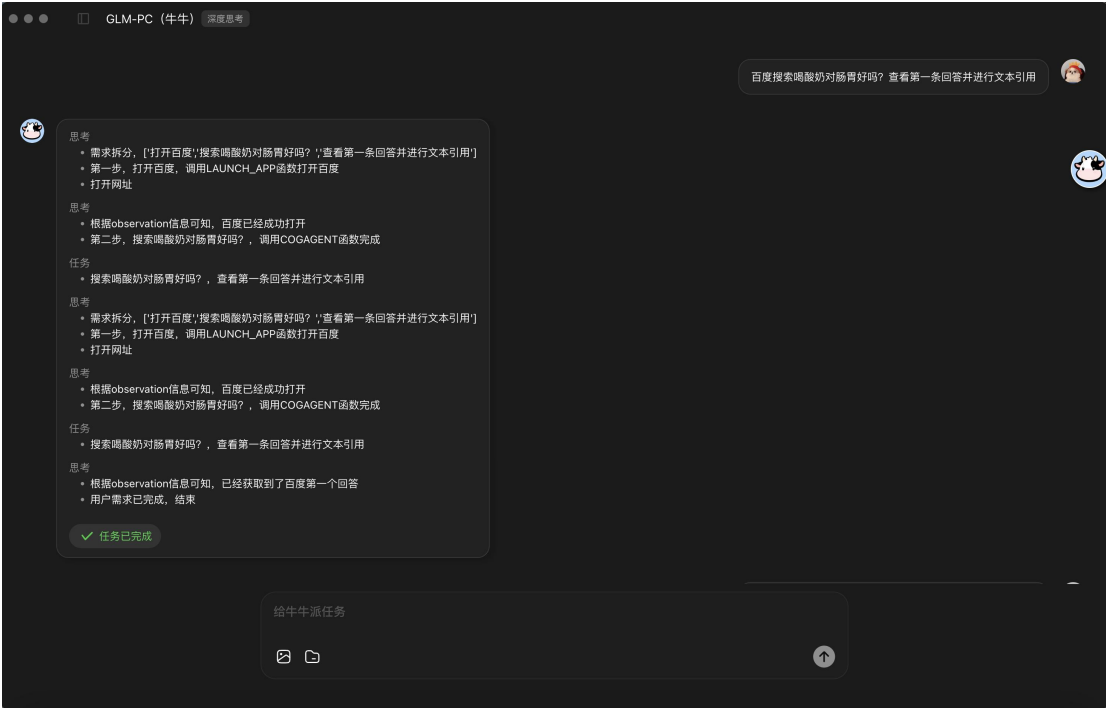
1. 示例 1: 快速阅读网页，阅读后进行图片理解、文本引用、内容总结、内容解释、内容翻译



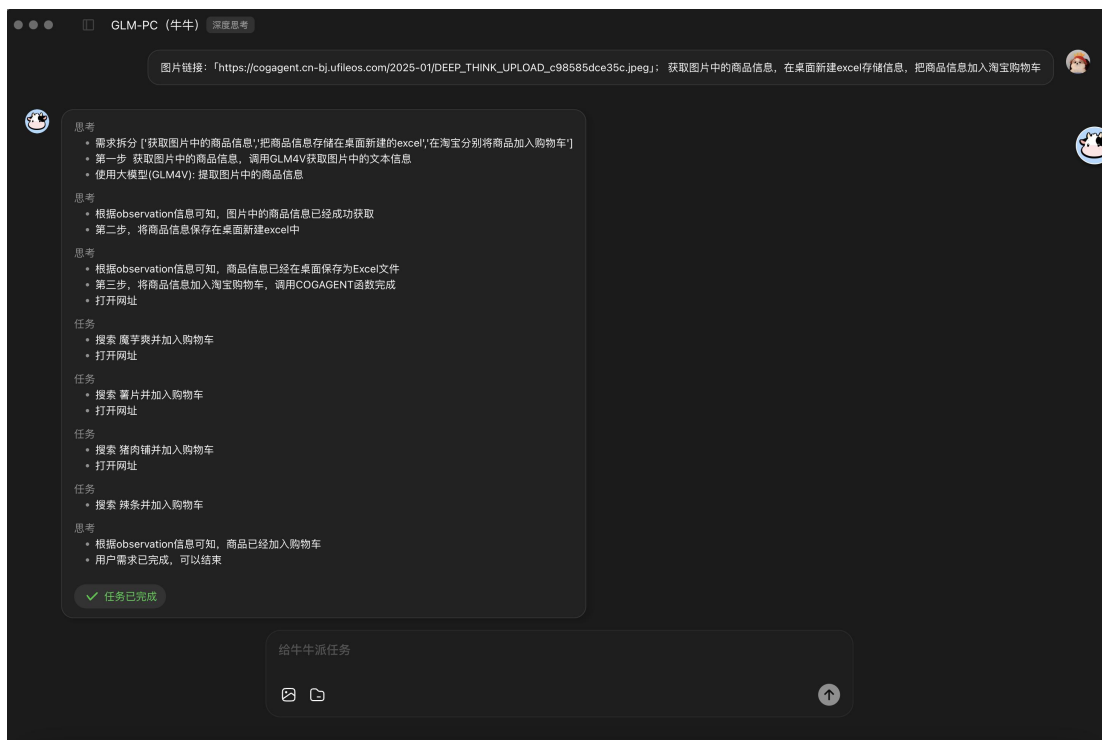
2. 示例 2: 快速阅读链接页面的内容，阅读后进行图片理解、文本引用、内容总结、内容解释、内容翻译



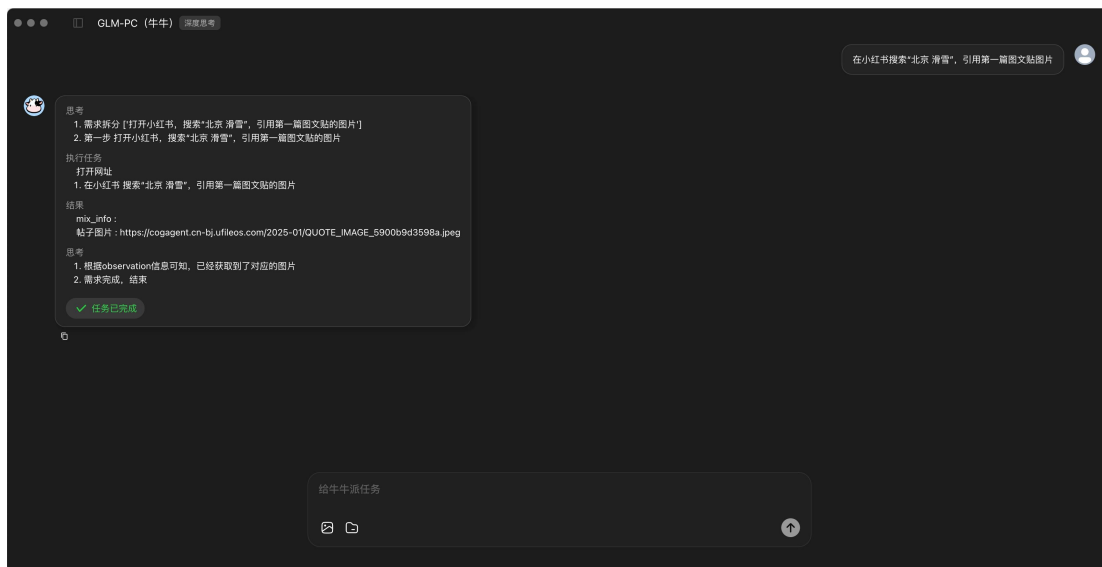
3. 示例 3: 搜索结果的来源引用展示



4. 示例 4: 快速识别链接页面的图片, 理解后跳转到淘宝进行商品加购



5. 示例 5: 引用图片的结果展示（点击 url 查看图片）



上传图片、根据图片做任务

GLM-PC（牛牛）具备强大的图片识别和处理能力，让您的工作事半功倍：

(1) 支持上传图片：PNG、JPG、JPEG 等格式，单次仅支持上传 1 张图片

(2) 智能图片理解

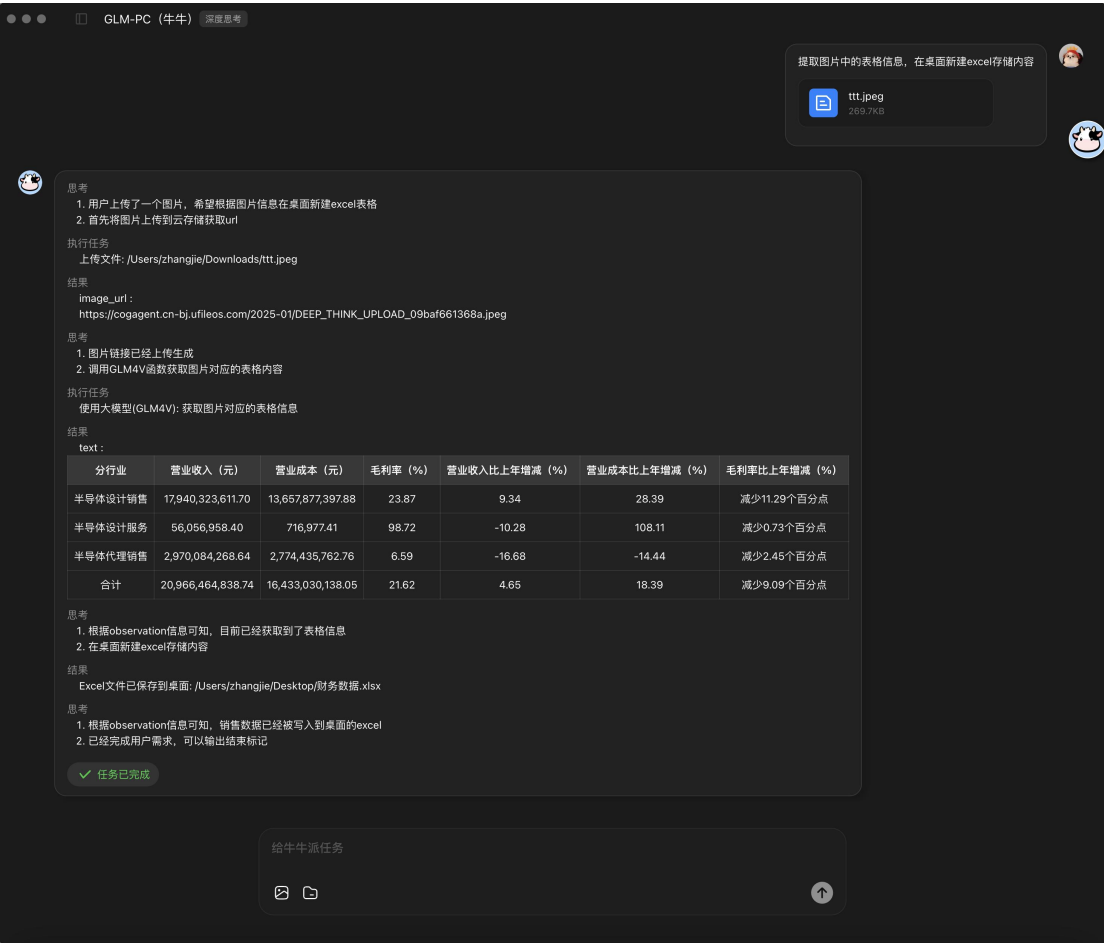
- 准确识别图片中的文字、界面元素和视觉信息
- 理解图片内容与上下文的关联
- 自动提取关键信息，辅助任务执行

(3) 基于图片的任务执行

- 根据图片内容智能规划任务步骤
- 可执行图片相关的复杂操作指令
- 支持"看图办事"式的自然交互方式

使用示例：

- "帮我把这张截图变成桌面上的 Excel 表格"
- "请将这个界面上的信息填写到表单中"



上传文件、根据文件做任务

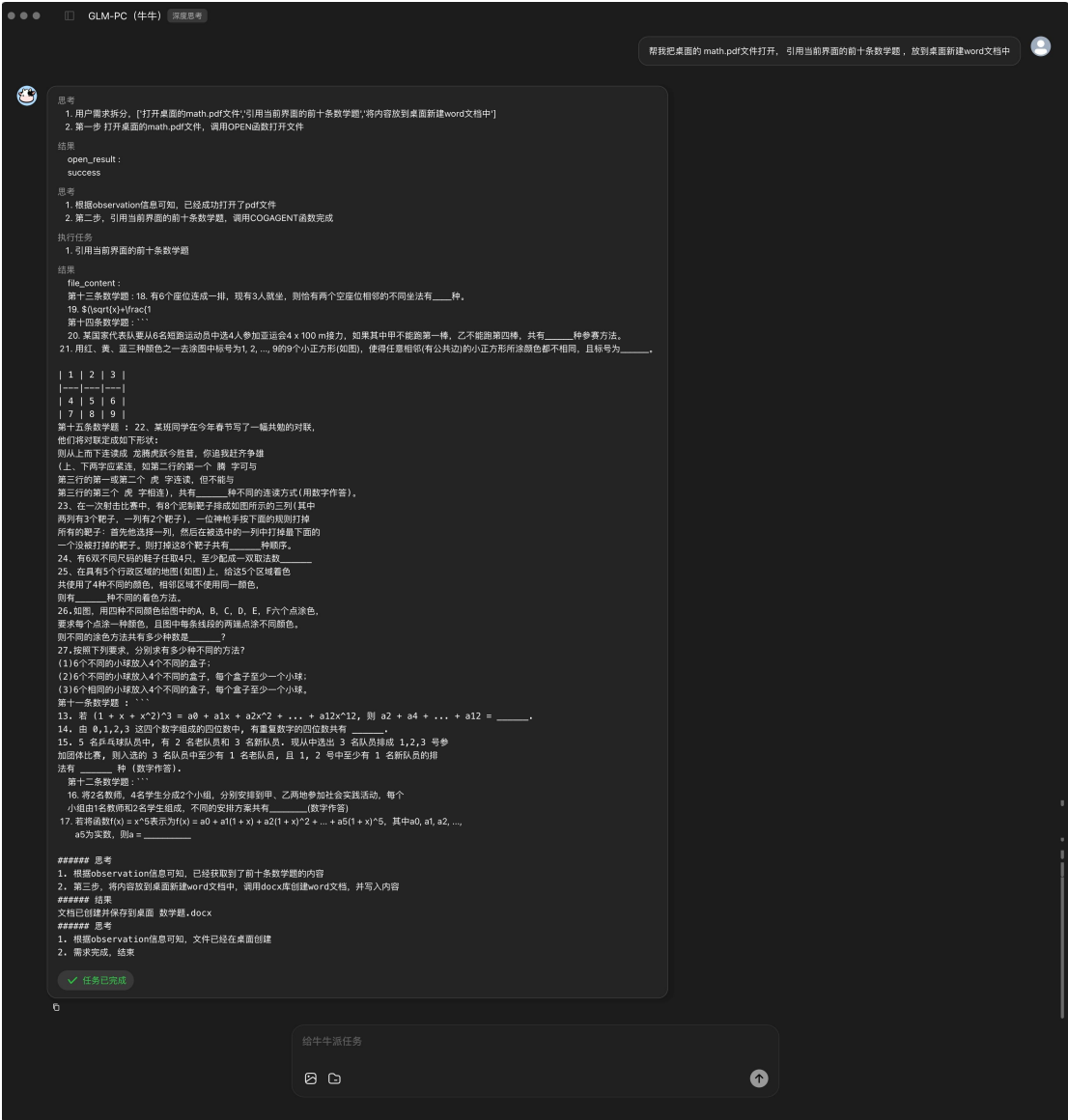
GLM-PC (牛牛) 能够灵活处理各类文件，为您提供全方位的文件智能处理服务：

(1) 多样化文件上传

- 支持常见办公文档（Word、Excel、txt）等多种格式
- 支持单个文件处理需求

(2) 智能文件理解

- 深度解析文件内容和结构
- 自动识别文件格式和编码
- 智能提取关键信息和数据



示例指令合集

更多应用牛牛的案例可以见如下文档中的指令和录屏：

[测试指令演示视频](#)

意见反馈

如果您有想对我们说的建议或者反馈，可以用文字、图片、视频进行反馈，可以在应用的用户个人中心（点击左下角的头像可以进入用户个人中心）进行反馈，或者用邮件发送到 cogagent@zhipuai.cn

相关文档链接：

[安装指南](#)

[教学视频集](#)