




在线平台快速上手指南

使用须知

本平台是用于处理复杂任务和专业任务的智能体平台，不是聊天平台，通常需要较多轮次迭代，执行时间大约 5-10 分钟，如果您希望快速聊天，请使用更合适的软件（例如 Deepseek/豆包）。


1. 请先选择平台，目前有主平台（用于通用任务、编程、写作等），专业图文写作平台（ColorDoc，用于国家项目申请书撰写、专业图文报告撰写等），专利助手（用于撰写专利交底书、专利说明书等）。各个平台的数据不互通，但账号互通。
2. 如果您想参考已有案例，可以在主页面填写空的 API Key，并点击连接按钮，工作目录就会列出参考案例，点击下箭头可展开某个工作目录具体查看文件，其中 workspace 中的文件为交付给用户的文件。
3. 如果您希望开始新的任务，请先注册，注册后会获得 API key 并自动填入。如果您丢失了 API key，可以采用同一个用户名及手机再次注册。


工作流程及核心功能使用说明


- (1) 工作目录中，选择一个目录（变为蓝色）或创建一个工作目录（点击工作目录标题旁边的  按钮）
- (2) 选择技能，如果没有合适的技能也可以不选择，但建议填写，填写技能后，可以不填写需求直接运行。**强烈建议选择一个技能（效果更好）。**
- (3) 上传您的背景文件（如果有），例如参考文档等，支持多文件上传（建议不要传输大于 10MB 的文件），可以从您的操作系统桌面或者文件浏览器直接拖入工作目录。
- (4) 在提示词对话框中填写您的任务需求，如果需要强调是对某个文件的修改，可以@文件名，或者从工作目录中直接拖拽文件到提示词对话框。
- (5) 点击任务发送按钮发送任务，任务开始执行。
- (6) 执行过程中可以从工作目录中找到生成的文件，其中 md 结尾的文件是大模型直接生成的，如果你想二次修改，请选择此文件，word/pdf 文件为导出的，大模型不能直接编辑。md 文件支持编辑，请在预览窗口中点击右上方的源码模式按钮即可编辑。相关程序（html/python 等）是大模型直接输出的，可以让大模型直接继续修改。
- (7) 如果您发现大模型走偏了，可以点击停止按钮，并输入新的需求后继续运行。
- (8) 如果任务已停止，新启动的任务默认只会知道历史任务的需求，而不会知道执行过程详细，如果你需要强调阅读历史详细，可以@manager.out 这个文件包括了历史执行轨迹。
- (9) 在 markdown 预览中，网络搜索的图片可以通过上一张/下一张切换（因为一个搜索关键词会下载多张图片），生成的 SVG 图可以编辑，点击编辑按钮后会进入 SVG 编辑模式，对 SVG 中的文字重叠、位置摆放问题进行简单修正，如果您需要更强大的修正，可以下载 SVG 图后，通过 powerpoint 软件加载此图片，并转换为形状，此时您就可以自由编辑此图了。您也可以通过 AI 润色按钮生成一张经过了润色的图（采用 nano banana pro），润色后的图不再支持编辑。
- (10) 生成的 word 默认采用 png 图（非矢量），pdf/md 文档默认采用矢量图。
- (11) 大模型生成的图片通常位于 images 文件夹。其中 mermaid 文件（.mmd）支持修改源码并重新编译为图片。
- (12) 网络搜索结果位于 web_search_results 文件夹，网络搜图位于 web_search_results/images 文件夹。
- (13) 如果您需要和智能体更详细的对齐需求，可以使用 plan 模式，此模式下大模型会向您提问，对齐一些需求模糊点后，撰写 plan.md 之后再切换到 agent 模式执行。

(14) 如果您需要多智能体协同工作, 需要先打开多智能体开关, 并在提示词中写清让主智能体何时启动其他智能体 (可以明确要求其使用 `spawn_agent` 发起子智能体), 否则主智能体可能会自己执行完所有的任务。如果需要智能体之间通信, 可以明确要求使用内建的智能体广播通信或 P2P 通信工具进行消息收发。收发的消息位于 `mailbox` 文件夹中。

主要功能按键介绍

工具栏中的主要按钮如下:  分别为: 多智能体开关 (默认不需要打开), 网络搜索开关, 无限执行模式开关, 上传文件按钮, 深度思考开关, 任务执行按钮。

聊天窗口右上角的四个按钮  的作用依次是: 加载聊天记录 (选择了某个工作目录后, 可以加载历史工作记录到聊天窗口), 多智能体可视化视图按钮, 虚拟终端按钮 (需要在 `config.txt` 中打开这个功能才会出现), 历史记录清理按钮。

工作目录的某个文件的功能按钮  为执行按钮 (仅当为可运行程序才会有), 回到历史版本按钮, 下载按钮, 删除按钮。鼠标悬停在某个按钮上可以看到按钮的解释。

1. 软件概述

1.1 产品介绍

AGI Agent 是一个基于大语言模型 (LLM) 驱动的 L3 级别自主智能体系统。该软件采用多轮迭代的工作机制，能够根据用户需求自主调用各种工具并处理反馈结果，实现对工作区文件的更新和外部环境的改变。作为目前唯一支持 5 类图像格式自动插入的富图像文档撰写智能体，AGI Agent 具有多智能体协同、多层次长期记忆、具身智能感知等特色功能，强调智能体的通用性及自主决策能力。该系统不仅能够生成图文并茂的专业文档，更能在复杂的场景中提供全方位的智能化支持。

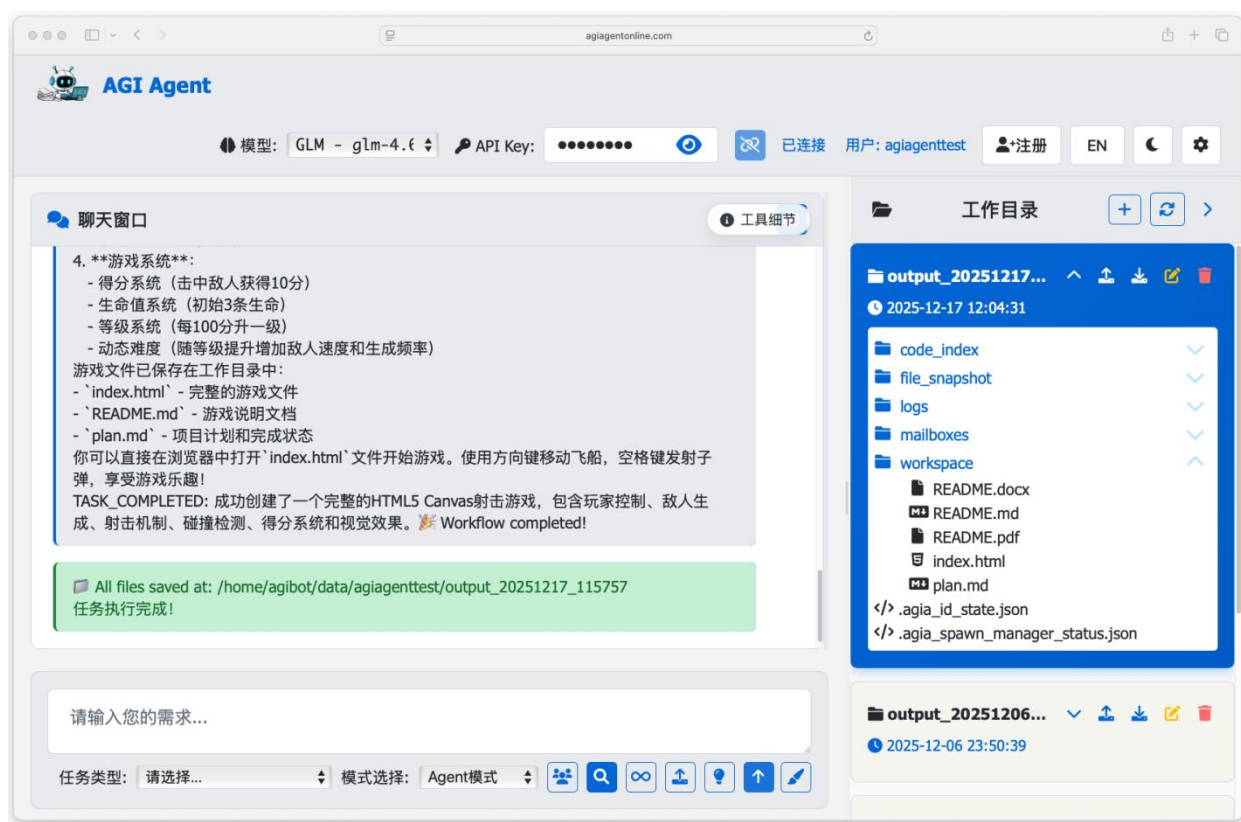


图 1. AGI Agent Web GUI 界面布局设计图，展示完整的用户界面和功能模块

1.2 核心优势

在富图像文档撰写方面，AGI Agent 提供了独特的多类型图像自动生成和插入功能。该系统支持 5 类图像格式的自动插入，包括 Mermaid 流程图、SVG 矢量图、网络搜索图像、AI 生成图像和代码分析图表，使得生成的文档图文并茂，彻底告别传统文档的枯燥无味。从文档质量角度来看，系统生成的文字具有学术论文和专业报告的水准，不像普通大模型那样只列观点不写细节，而是注重深度分析和详细描述，文字优雅专业，适合高质量的学术写作应用。

在编程方面，作为氛围编程（Vibe Coding）平台，可以用于编写典型的 python/c++ 程序，用户只需要撰写提示词（需求），并选择合适的工作模式，即可得到对应的完整的代码仓库，实现基于自然语言的应用开发。

在通用计算机任务执行方面，智能体可以根据用户要求执行典型的多步骤计算机任务，包括目录整理、安装软件、执行操作系统及 shell 命令，网络搜索等。

另外，系统具备强大的多智能体协同能力，可以自主创建专业化的子智能体，每个智能体都配备独特的提示词、模型类型和专用工具库，在共享工作空间下高效协同工作。

1.3 应用场景

AGI Agent 广泛适用于复杂报告调研与生成、项目级代码编写、计算机自动运行、多智能体研究等应用场景。在学术研究领域，该软件特别擅长生成完整的研究报告、项目申请书、学术论文、专利文档等，单次可生成 30-50 页的长篇文档。此外，系统内建了 10 余个专业写作模板，涵盖报告撰写、国家项目申请书、标书技术投标稿、博客文章、专利交底书、丰富图片文档等多种文档类型，风格多样，满足不同的写作需求。从技术应用角度来看，该系统还支持数据分析、实验设计、代码开发等专业写作全流程工作，通过多模态感知能力处理图像、音频、视频等多种类型的输入数据提供全方位的智能化支持。

2. 系统要求

2.1 硬件要求

AGI Agent 对硬件配置要求较为灵活，能够在多种硬件环境下稳定运行。首先，在处理器方面，系统支持 x86_64 架构的 Intel 或 AMD 处理器，推荐使用四核心以上的 CPU 以获得更好的性能表现，特别是在多智能体协同工作时，多核心处理器能够显著提升并行处理能力。其次，内存需求方面，最低要求 4GB RAM，但推荐配置 8GB 或更高内存，特别是在处理大型文档生成、多模态数据处理或运行多智能体协同任务时，充足的内存能够确保系统稳定运行和快速响应。再次，存储空间需求相对较小，基础安装仅需 500MB 磁盘空间，但建议预留 2GB 以上空间用于工作文件、图像缓存、长期记忆数据和生成文档的存储，确保系统具备足够的扩展空间。

2.2 操作系统兼容性

AGI Agent 具有优秀的跨平台兼容性，全面支持主流操作系统，确保不同用户都能享受到完整的功能体验。在 Linux 系统方面，支持 Ubuntu 18.04+、CentOS 7+、Debian 9+ 等主流发行版，提供完整的功能支持和最佳的性能表现，特别是在服务器环境和自动化部署场景中表现优异。另外，macOS 系统同样得到完全支持，兼容 macOS 10.15+ 版本，确保苹果用户也能享受到包括富图像文档生成、多智能体协同等在内的完整功能体验。

2.3 软件依赖与模型支持

系统的核心依赖是 Python 3.6 或更高版本，建议使用 Python 3.8+ 以获得最佳的兼容性和性能表现。此外，稳定的网络连接是必需的，用于 API 调用、网络搜索功能、图像搜索和模型服务的访问。对于网页抓取功能，系统需要安装 Playwright 浏览器驱动，这是可选的依赖项，仅在需要网页内容抓取时才需要安装。从大语言模型支持角度来看，系统支持多种主流 AI 模型，包括 Anthropic Claude 系列（Claude Sonnet 4 提供最佳智能程度）、OpenAI GPT 系列（GPT-4 Turbo 平衡性能与速度）、Google Gemini 系列以及国产模型如 Kimi K2（优秀的 Agent 性能）、DeepSeek V3（极佳性价比）、火山大模型、Qwen 系列等，用户可以根据需求和预算选择合适的模型配置。

3. 安装指南

3.1 环境准备

在开始安装 AGI Agent 之前，需要确保系统环境满足基本要求并进行必要的准备工作。首先，验证 Python 版本是否符合要求，可以通过在终端或命令提示符中执行 `python --version` 命令来查看当前 Python 版本，如果系统中尚未安装 Python 或版本过低，需要前往 Python 官网下载并安装最新版本。其次，确保系统具有稳定的网络连接，因为安装过程中需要下载依赖包，运行时也需要访问大语言模型的 API 服务、进行网络搜索和图像搜索等功能。最后，强烈建议在安装前创建一个专用的 Python 虚拟环境，通过执行 `python -m venv agia_env` 和 `source agia_env/bin/activate`（Linux/macOS）来避免与其他 Python 项目产生依赖冲突，确保系统的稳定性和可维护性。

3.2 源码安装（推荐）

从源码安装是推荐的安装方式，能够获得最新的功能和更新，特别是富图像文档生成和多智能体协同等核心特性。

首先，从github下载 `git clone https://github.com/agi-hub/AGIAgent`

然后，执行一键安装脚本

```
./install.sh
```

整个安装过程通常在几分钟内完成。

对于高级用户，可以手工执行如下安装过程：

1. 安装 Python 依赖包：通过执行 `pip install -r requirements.txt` 命令来安装所有必需的依赖项，如果网络不稳定可以使用国内镜像源 `pip install -r requirements.txt -i https://pypi.tuna.tsinghua.edu.cn/simple/`

2. 安装 Playwright 浏览器驱动

执行 `playwright install-deps` 和 `playwright install chromium` 命令

3. 安装 pandoc 软件

执行 `sudo apt install pandoc`

4. 安装 latex 软件系统（如果需要 pdf 输出）

执行 `sudo apt-get install -y texlive-xetex texlive-fonts-recommended texlive-fonts-extra fonts-noto-cjk`

4. 配置设置

4.1 基础配置

AGI Agent 的配置主要通过 `config/config.txt` 文件进行管理，在基础配置方面，最重要的是 API 相关设置，包括 `api_key`、`api_base` 和 `model` 三个核心参数，用户需要根据所选择的大语言模型服务提供商，填写相应的 API 密钥和服务地址。

请至少配置模型到 `# GUI API configuration` 字段，对于 GLM-4.7 模型，可以使用 anthropic 接口，如下：

```
# GUI API configuration
api_key=your_api_key(可从bigmodel.cn申请)
api_base=https://open.bigmodel.cn/api/anthropic
model=glm-4.7
max_tokens=16384
```

其中 `max_tokens` 指最大输出 token 数量，默认设置为 16K tokens。

另外，语言设置也是重要的基础配置项，通过 `LANG=zh` 可以将界面和输出设置为中文，确保生成的富图像文档具有良好的中文支持和本土化体验。

AGI Agent 支持多种主流大语言模型，用户可以根据不同的应用需求和预算选择最适合的模型配置。从文档质量角度来看，Claude Sonnet 4 提供最佳的智能程度和富图像文档生成质量，特别适合需要高质量学术写作和专业报告的场景。从平衡性能考虑，GLM-4.6 在速度与质量之间达到了很好的平衡，适合大多数日常应用场景，包括多智能体协同任务。从成本效益角度分析，DeepSeek V3.2 提供了极佳的性价比，国产模型成本可控，几十页、几万字文档开销小于 1 元人民币，而 Kimi K2 则在 Agent 和编码性能方面表现优秀，特别适合多智能体协同和代码生成任务。每种模型都有其特定的配置要求和 API 格式，用户需要根据选择的模型进行相应的参数调整。

4.3

5. 核心功能特性

5.1 富图像文档生成系统

AGI Agent 在富图像文档生成方面具有全球领先的技术优势，是唯一支持 5 类图像格式自动插入的智能文档撰写系统。该系统能够自动生成包含多种类型图像的高质量文档，包括 Mermaid 流程图、SVG 矢量图、网络搜索图像、AI 生成图像和代码分析图表等五大类图像格式，实现真正的图文并茂效果。从技术实现角度来看，系统内置了 SVG 图编辑器、Mermaid 源码编辑器、Markdown 编辑器及预览器，搜索到的图像可以一键更换，使得图文文档编辑更加便捷高效。另外，为了满足高质量出版物的需求，生

成的图像同时提供 SVG 矢量格式和 PNG 位图格式，确保在不同应用场景下的兼容性和显示效果。系统支持输出 Markdown、Word、PDF、LaTeX 等多种文档格式，特别是 LaTeX 格式对学术论文撰写具有重要价值，能够满足各种专业文档发布和印刷的需求。

5.2 代码开发与数据分析

AGI Agent 在代码开发与数据分析方面提供了全面的支持，能够处理从简单脚本到复杂项目的各种编程和分析任务。系统具备智能代码搜索功能，结合语义搜索和关键词搜索技术，能够快速定位和分析现有代码库中的相关内容，支持增量索引和动态更新。从代码质量角度来看，系统生成的代码不仅功能完整，而且遵循最佳实践和编码规范，包含详细的注释和文档说明，支持多种编程语言和框架。再次，系统还支持代码测试和调试，能够自动生成测试用例并执行验证，确保代码的可靠性和稳定性。对于数据分析任务，系统能够自动处理各种数据格式，生成专业的数据处理脚本、创建多样化的可视化图表、编写详细的分析报告，所有结果都会整合到富图像文档中。系统还支持实验设计和结果分析，能够协助用户完成从数据收集、处理、分析到报告生成的完整流程。

5.3 多智能体协同架构

在多智能体协同方面，AGI Agent 实现了智能体的自主创建和专业化分工，构建了完全自主的协同工作机制。系统能够根据任务复杂度和专业领域需求，自主决定创建相应的专业智能体，每个智能体都配备独特的提示词、模型类型和专用工具库，形成术业专攻的专业化分工体系。从协作机制来看，不同角色的子智能体能够在共享工作空间下高效协同工作，通过点对点及广播消息通信机制实现无缝互联互通，支持复杂任务的并行处理和协同优化。此外，系统还集成了邮箱信件查看机制，智能体之间可以进行更加复杂和结构化的信息交换，确保协同工作的高效性和准确性。这种多智能体架构特别适合处理需要多个专业领域知识的复杂项目，如大型报告撰写、跨学科研究、多维度数据分析等，能够显著提高工作效率和输出质量。

5.4 专业任务模板

AGI Agent 内建了 10 余个专业任务模板，涵盖了主要的文档类型，每个模板都经过精心设计和优化，确保生成的文档具有专业的质量和风格。这些模板包括研究报告撰写、国家项目申请书、标书技术投标稿、博客文章、专利交底书、丰富图片文档、编写程序，风格多样，满足不同场景的写作需求。

6. 使用说明

AGI Agent 提供了现代化的 Web 图形界面，为用户提供直观便捷的操作体验和可视化的任务管理功能。

6.1 服务器端程序的配置和启动

运行 GUI/app.py 启动后台服务:

```
python GUI/app.py
```

终端具有如下打印则代表启动成功:

```
Starting AGIAgent GUI Server on port 5008
```

```
* Serving Flask app 'app'
```

```
* Debug mode: off
```

如果不希望在 5008 端口启动, 可以使用 `--port <端口号>` 指定。

如果希望驻留程序运行, 请使用驻留方式启动:

```
cd GUI/daemon
```

```
./start_monitor_daemon.sh
```

此时关闭终端也不会影响驻留程序运行。

最后, 查看机器 ip 地址, 用于用户登陆, 例如使用 `ipconfig` 或 `ifconfig` 获得。

6.2 Web GUI 用户界面前端的启动

Web GUI 可对整个执行过程中实时显示执行进度和详细日志信息, 用户可以观察到图像生成过程、智能体创建情况、工具调用状态等详细信息。从监控角度来看, 用户可以随时查看任务的执行状态, 包括当前执行轮数、多智能体协同进度、富图像文档生成状态、生成的文件列表等信息。另外, 系统还支持任务的暂停和继续功能, 用户可以在任务执行过程中进行干预和调整, 特别是在多智能体协同任务中可以实时调整协作策略。任务完成后, 用户可以通过下载按钮一键下载整个工作区的内容, 包括生成的富图像文档、代码文件、数据分析结果等, 方便后续的使用和分享。

6.2.1 用户登陆与注册

使用浏览器登陆主页, 地址为: <http://server-ip:5008>。成功访问时, 浏览器会显示如下 Web GUI 页面。

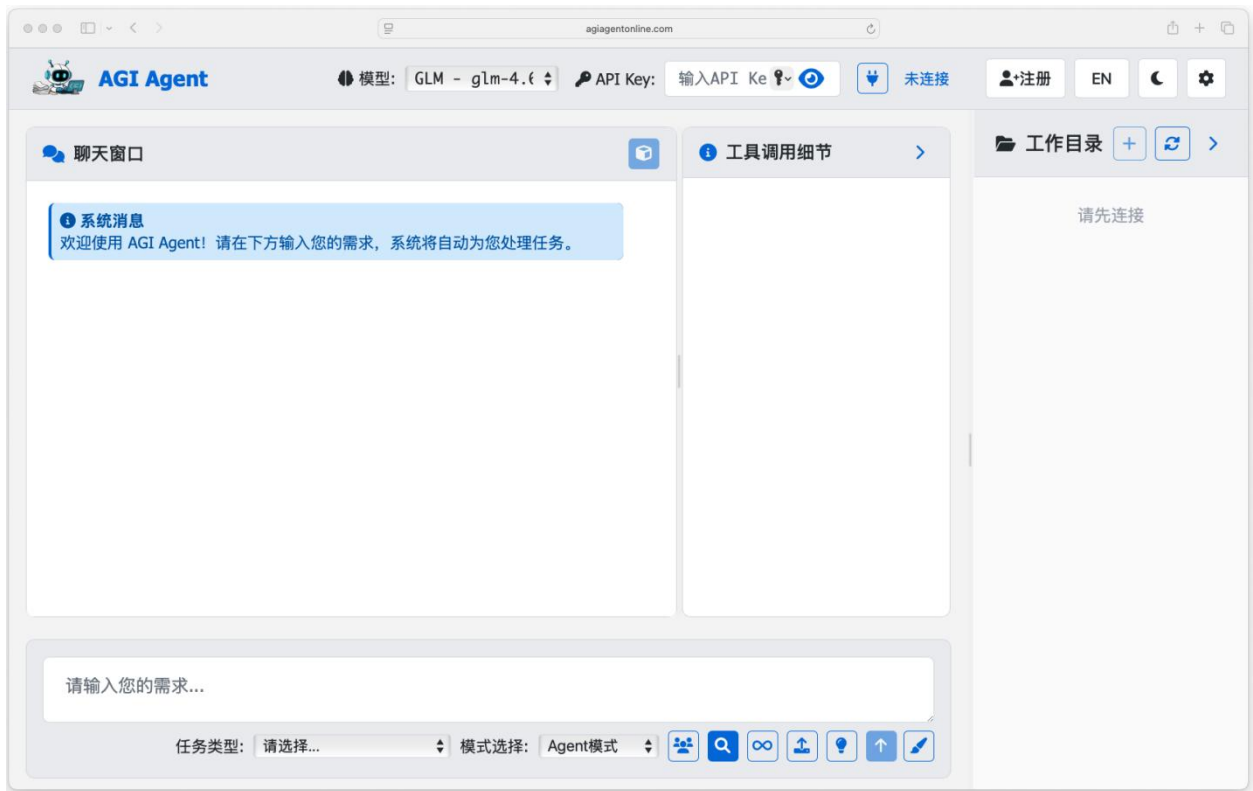


图 2. 登陆主页

6.2.1.1 注册与找回

系统采用单密钥免登录方式，主页面只需要填写 API Key，如果以前登录过，可以不重复输出。如果没有 API Key，可以如下方式获取。

首先点击主页的注册按钮进行用户注册，获得 API Key，如下图所示。如果忘记了 API Key，可以使用同样的用户名和手机进行注册，可以拿回 API Key。

注：系统后台不会记录你的 API Key，而是记录的 APIKey 的 sha256 Hash 编码，因此不要试图从系统后台找回 APIKey。系统只会记录手机号的后 4 位，不会记录完整手机号。系统后台会使用你的用户名作为用户子目录的路径名。你的用户数据位置位于

数据根目录/你的用户名/工作目录

其中数据根目录在 config.txt 中通过变量 `gui_default_data_directory` 配置

工作目录新建任务时会动态生成的时间戳名称，类似于 'output_20251021_145421'

此外，系统包括了一个测试账号 Username=agiagenttest，APIkey=agiatest

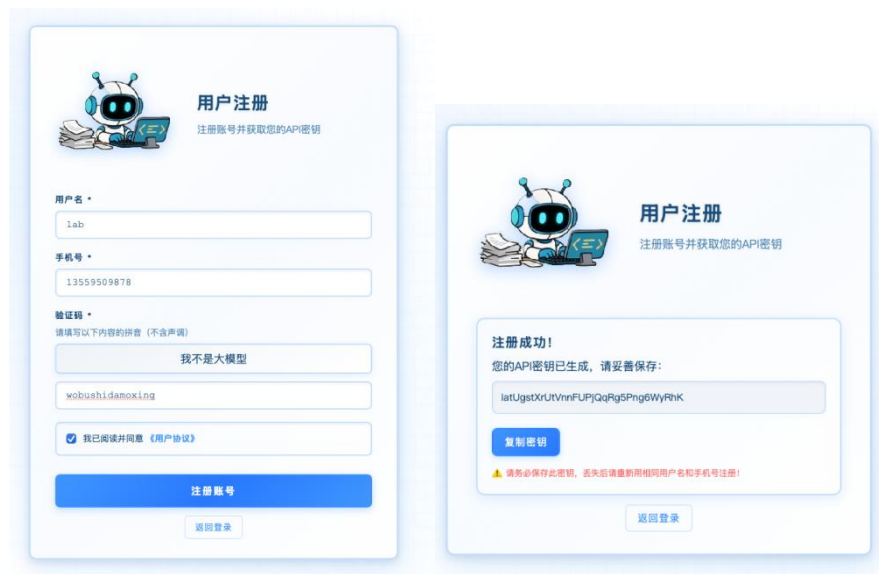


图 3. 用户登陆

6.2.2.2 登陆

进入主页后，API Key 会自动填充，请点击 API Key 右侧的插头按钮，系统会切换到已连接状态。
注意：在未连接状态下，系统无法发起新的任务，无法创建文件夹。请先连接。连接后，工作目录侧会显示以前所有工作使用的工作目录。

6.2.3 选择工作目录

本软件的所有输出物会放置在一个工作目录中，点击工作目录的+按钮，可以新建工作目录。左键点击某个工作目录，变为蓝色时，代表选中了该工作目录。
工作目录可以折叠/展开，用于查看详细文件。
详细工作目录描述见工作目录一节。

6.2.4 启动任务

(1) 请选择任务类型

对于普通编程任务，无需选择（处于请选择状态即可）
对于图文编写任务，可以选择对应的任务类型。

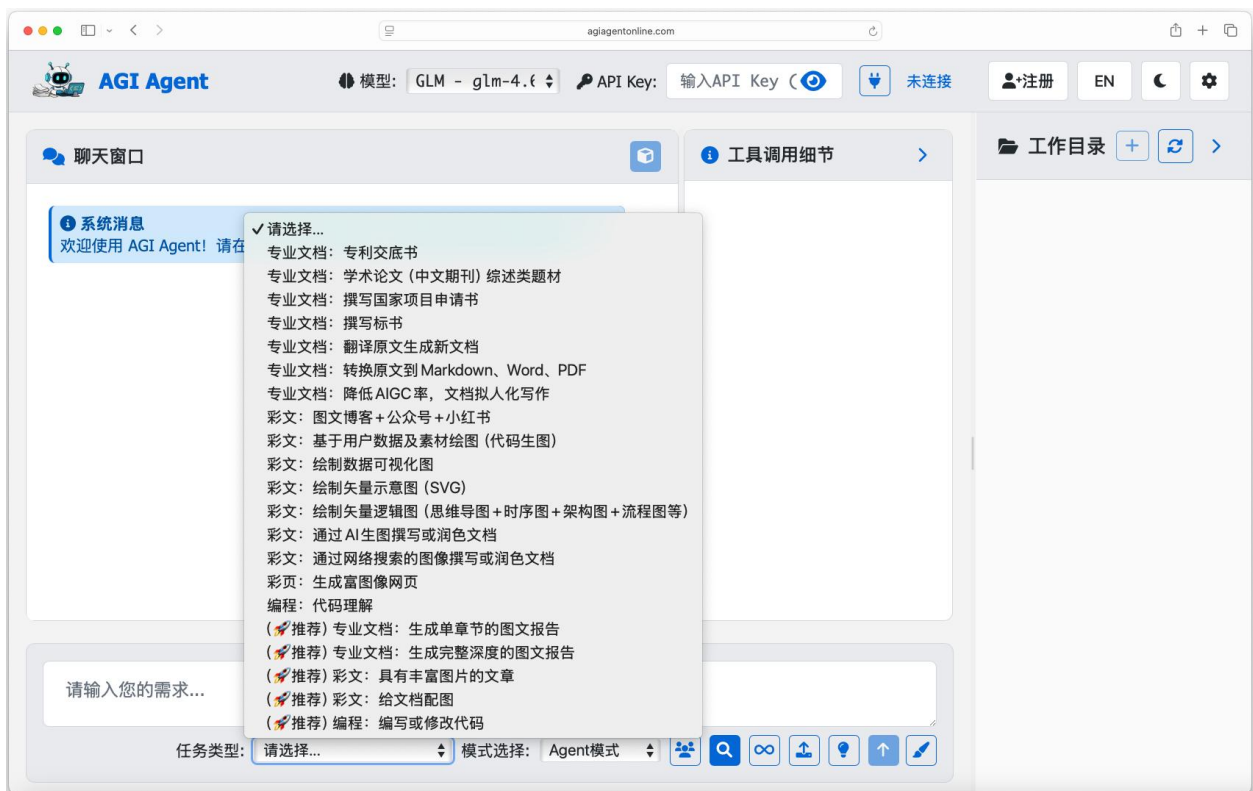


图 4. 任务类型选择

(2) 选择模式，Agent 模式即直接智能体执行模型。Plan 模式会和用户进行深入的需求对齐，形成 plan 文件，之后才会切换到 Agent 模式运行任务。

(3) 点击上箭头  按钮，即可启动任务。

6.2.5 工具栏介绍

右侧的一系列按钮在悬停时可以得到对应的解释。其中：

第一个按钮为多智能体开关，打开（底色变蓝色）代表打开了多智能体相关功能，此时主智能体 (manager) 会自动创建子智能体，完成部分任务。

第二个按钮为网络搜索开关，打开则允许进行网络搜索。

第三个按钮为无限运行按钮，在未开启时，智能体在运行一定轮数后就会停止，轮数可在设置界面配置，默认 100 轮。开启此按钮后，智能体不会自己结束任务，且没有轮数限制。任务结束后会进入 IDLE 状态并等待用户的追加任务。但如果系统网络中断或页面退出则会退出任务。

第四个按钮为上传文件按钮，可以上传资料到工作目录。

第五个按钮为任务执行按钮，请先在需求框中输入需求，再点这个按钮执行任务。如果你选择了某个任务类型，也可以不输入任何需求（因为任务类型已明确了需求）。

第六个按钮为清除历史记录按钮。会清空聊天窗口的显示。



图 5. 控制按钮区域

6.2.6 配置页面介绍

其他配置项可以在页面右上角的配置按钮（齿轮形状）后找到，其中执行轮次可以更改智能体在非无限模式下的执行轮数。启动长期记忆选项可以启用长期记忆功能。MCP 工具可以启用 MCP 工具，默认的 MCP 工具在 config/mcp_servers.json 中配置，并不勾选 MCP 工具配置按钮时生效。勾选 MCP 工具配置之后，则会读取 config/mcp_servers_GUI.json 文件，只有在这里配置过的工具才会显示到此界面。勾选后的工具才会生效，且会覆盖默认的 MCP 配置。

启用中文分词只在撰写带有中文的代码时有用，用于搜索代码空间时具有更高精度。

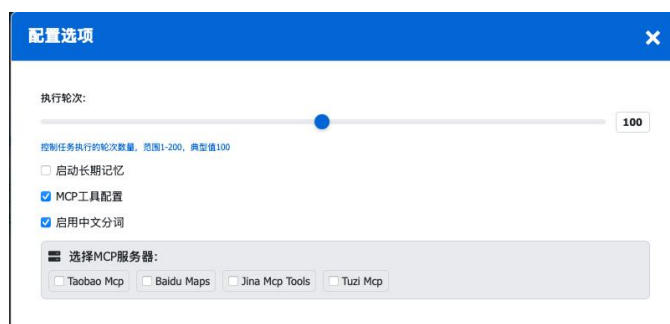


图 6. 额外配置窗口

6.2.7 模型修改界面

在模型选择框可以找到自定义选项，之后进入模型配置页面，用户可以自行配置 API Key，Base URL 和模型名称。支持 Anthropic 协议和 OpenAI 协议。保存后，会自动更新到 config.txt 文件的 # GUI API configuration 字段。



图 7. 模型配置

6.2.8 示例应用

目前任务类型包含了两个示例应用：

“编程：撰写神经元发放模拟算法并运行验证”

“编程：撰写脉冲神经网络训练代码并验证训练过程”

无需填写需求，点击运行即可开始运行该任务。

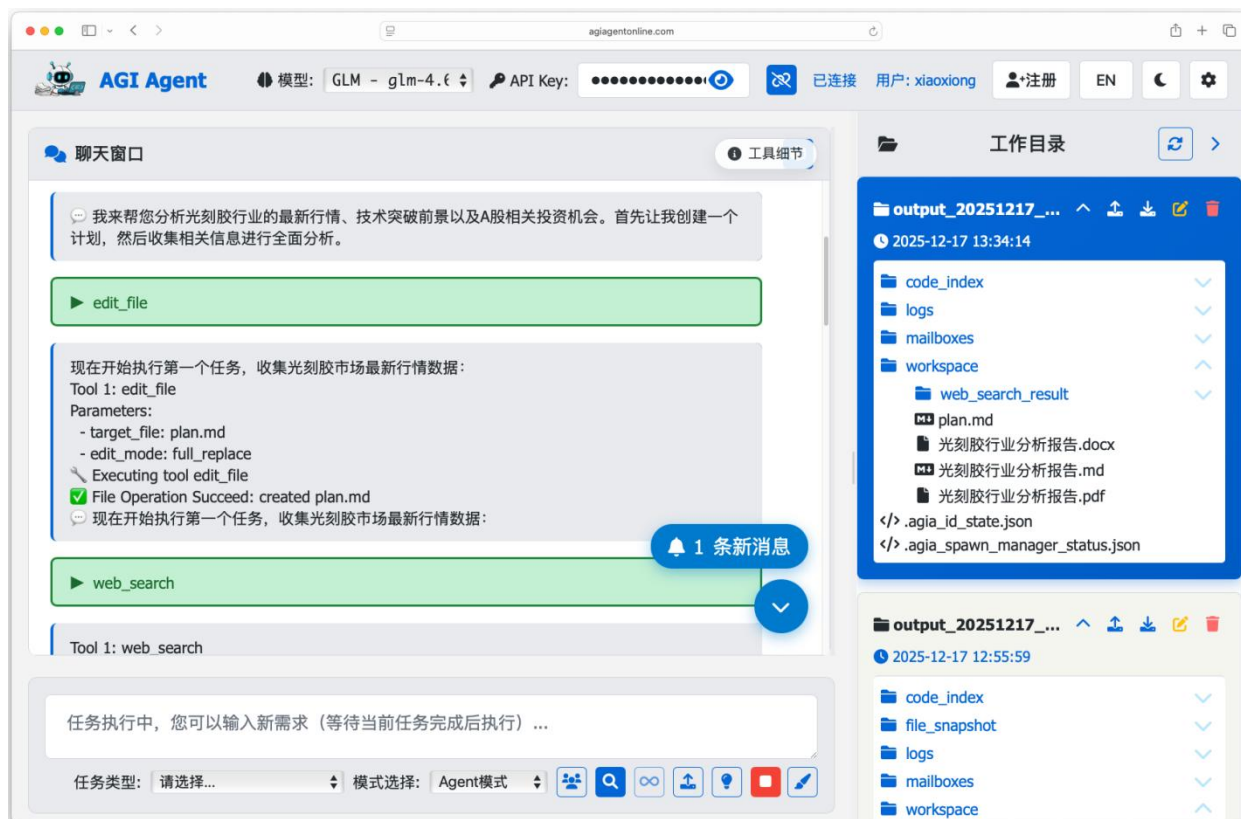


图 8. AGI Agent 工具调用过程


其中左侧聊天窗口中，会显示大模型的聊天消息、工具调用（绿色框）以及工具执行结果（Tool 1: XXXX 等）。

此外工具调用细节对话框会显示工具调用细节，用户也可以点击聊天信息的绿色框看到相同的内容。

右侧的工作目录会定时更新，如果更新不够及时，可以点击工作目录文字右侧的刷新按钮。

注：由于当前大模型编程仍为前沿领域，目前尚无法做到全自动和 100%成功，仍然为辅助编写手段，用户需通过多轮次对话及少量代码完善来完成完整项目。

6.2.9 多智能体视图

点击智能体模式按钮，可以看到实时的多智能体创建、消息交互、每个智能体的工具调用情况，运行轮数等内容（仅在多智能体打开时建议使用）。

6.2 工作目录管理与文件操作

工作目录是智能体工作的输入输出文件位置，大模型默认只会在工作目录中读取、修改、写入文件。在工作目录管理方面，软件提供了完整的文件和目录操作功能，支持富图像文档的创建、编辑和管理。用户可以通过界面创建新的工作目录，或者选择现有的工作目录进行操作，选中的工作目录会以蓝色高亮状态显示，确保操作的准确性。从文件管理角度来看，用户可以直接在浏览器中上传、下载和管理项目文件，支持拖拽上传和批量操作，特别适合处理包含多种图像格式的复杂文档项目。此外，系统还提供了强大的文件预览功能，支持 Markdown、代码、图像、SVG、Mermaid 图表等多种文件格式的在线预览和编辑，用户无需下载文件就能查看和修改内容。所有生成的富图像文档、代码文件、数据分析结果都会自动归类到相应的目录中，便于用户查找和管理，同时支持一键下载整个工作区的功能。

6.2.1 工作目录的文件规则

对于某个工作目录，通常包括如下子目录：

1. 输出的所有内容位于 workspace 目录，用户所需的文件应该都位于这个目录。

其中：

- workspace/images 为大模型通过图文混排（在 markdown 中嵌入撰写的图像 SVG 源码/mermaid 源码）自动被工具解析生成的图像。
 - workspace/web_search_result 为通过 web_search 工具搜索到的网页源文件及对应抽取过滤得到的纯文本文件。
 - workspace/web_search_result/images 为 image_search 工具搜索得到的图像文件。
 - 运行日志位于 logs 目录，包括主智能体的历史对话文件及日志文件（manager.out/manager.log）及其他智能体的对应文件（agent_XXX.out, agent_XXX.log）。
2. 多智能体的邮箱信件位于 mailbox 目录。
3. 代码检索的向量信息存储于 code_index 目录，这部分目录用户不需要使用，且在下载完整工作目录时也不会被下载。

6.2.2 文件预览与编辑

Web GUI 界面集成了强大的富图像编辑和预览功能，为用户提供了完整的图文文档创作环境。系统内置了 SVG 图编辑器，用户可以直接在浏览器中编辑和调整 SVG 矢量图，支持图形元素的添加、修改、样式调整等操作，确保图像的专业性和美观性。从 Mermaid 图表角度来看，系统提供了专门的 Mermaid 源码编辑器，用户可以实时编辑流程图、时序图、类图等各种类型的图表，编辑器支持语法高亮和实时预览功能。此外，系统还集成了 Markdown 编辑器和预览器，支持富文本编辑和实时预览，用户可以在编辑过程中即时查看文档的最终效果。对于网络搜索获得的图像，系统提供了一键更换功能，用户可以轻松替换不合适的图像，优化文档的视觉效果。所有编辑操作都支持实时保存和版本管理，确保用户工作的安全性和连续性，同时支持多种格式的导出，包括 PDF、Word、LaTeX 等格式，满足不同应用场景的需求。

Markdown 预览界面顶部具有源码模式按钮，点击后可以进入源码编辑模型，用户可修改该文件。修改后请及时保存。此外，顶部还具有 Word/PDF/Latex 格式的导出按钮，会实现 markdown 文件到上述三个格式的导出和自动下载，导出后，工作目录中也会出现这些文件。



图 11. 单击 md 文件可以预览生成的文稿，且页面上方具有 word/pdf/latex 文稿生成按钮
源码模式下用户可以手工编辑文稿

Markdown 预览遇到 SVG 图像时，可以按编辑按钮进入 SVG 编辑状态，通过鼠标拖拽对图像的元素（例如文字，形状）进行拖拽微调，实现简单的 SVG 编辑。SVG 编辑并保存后，Markdown 文件会更新。

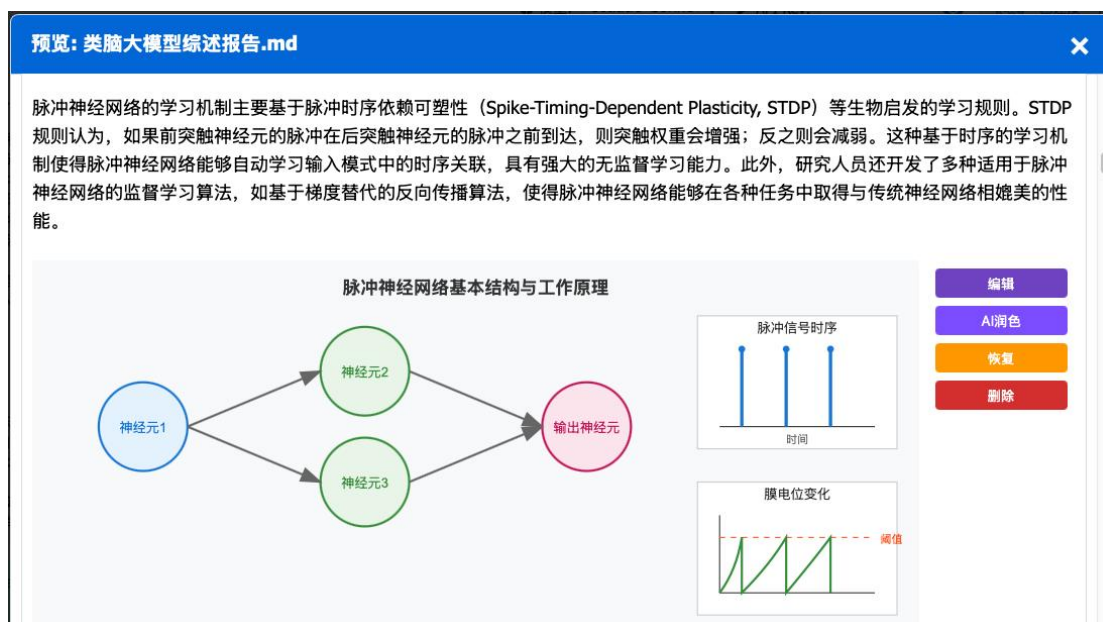


图 12. 图文混排案例 (SVG 图)

Markdown 预览遇到网络搜索的图像时，可以按上一张、下一张实现切换。由于下载图像时，同一个搜索词会下载多张图像，因此可以在上述图像中找出更符合上下文的一张。

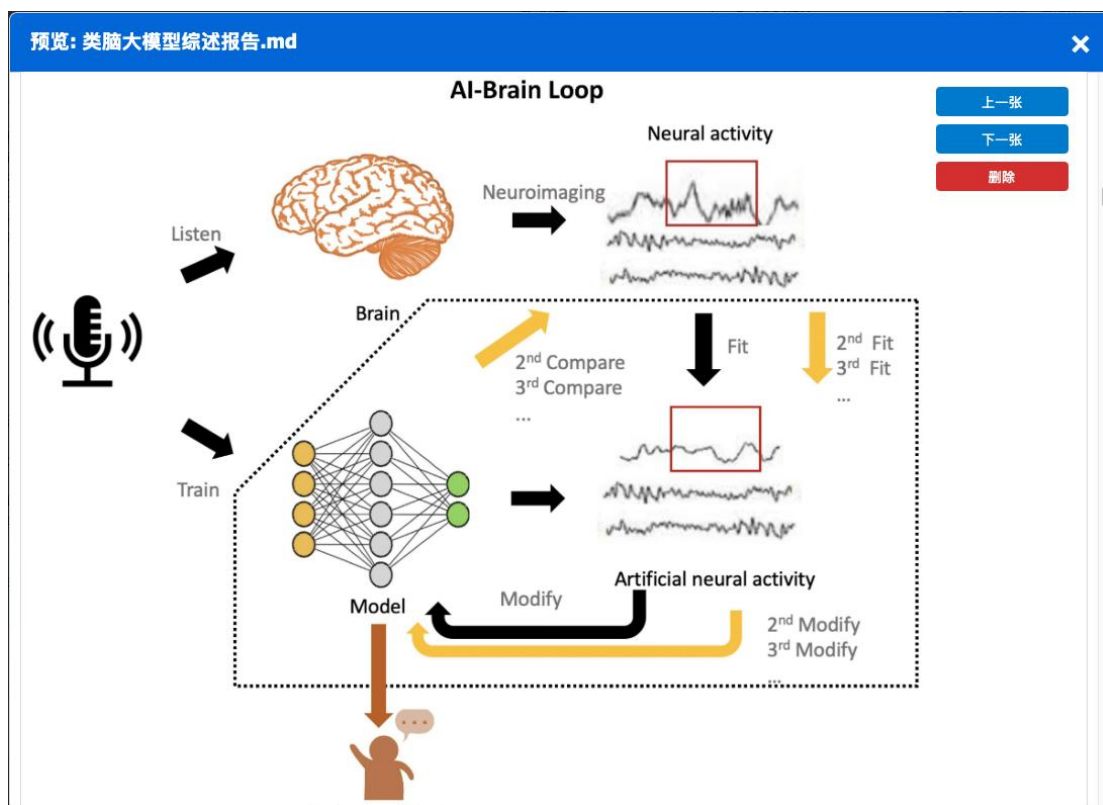


图 13. 图文混排案例（网络搜索图像），可以通过上一张、下一张进行切换

图像（SVG、PNG）预览窗口可以查看图像及放大预览。Mermaid 源码可打开进入编辑页面，编辑后可以点击生成新的图像。

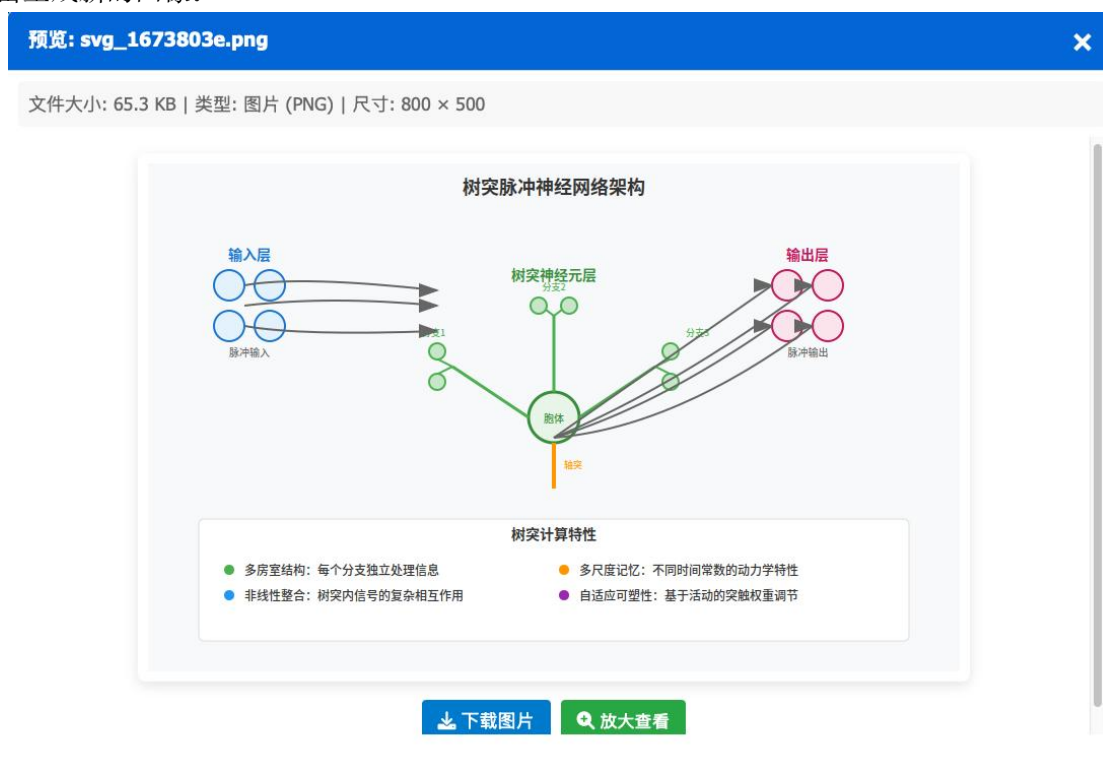


图 14. 界面可以查看生成的所有 PNG/SVG 图

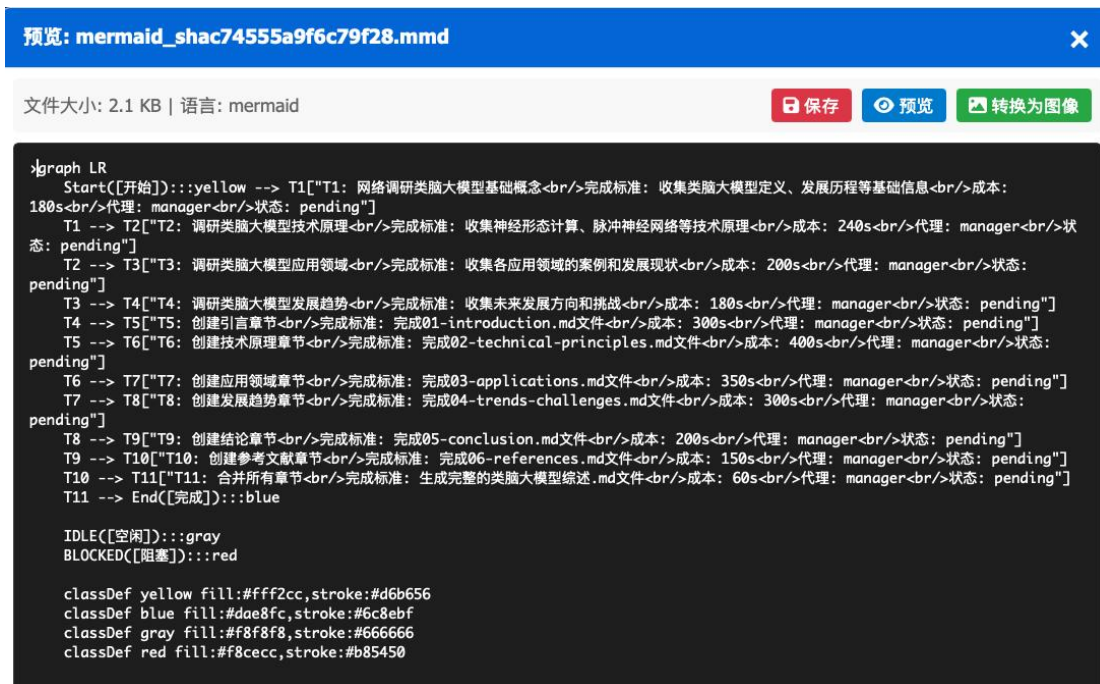


图 15. 界面支持对流程图（Mermaid 图）进行源码预览及用户修改，修改后可以转换为图像

对于聊天历史，会存储于 manager.out 中，更多信息存储与 manager.log 中，可以点击查看。

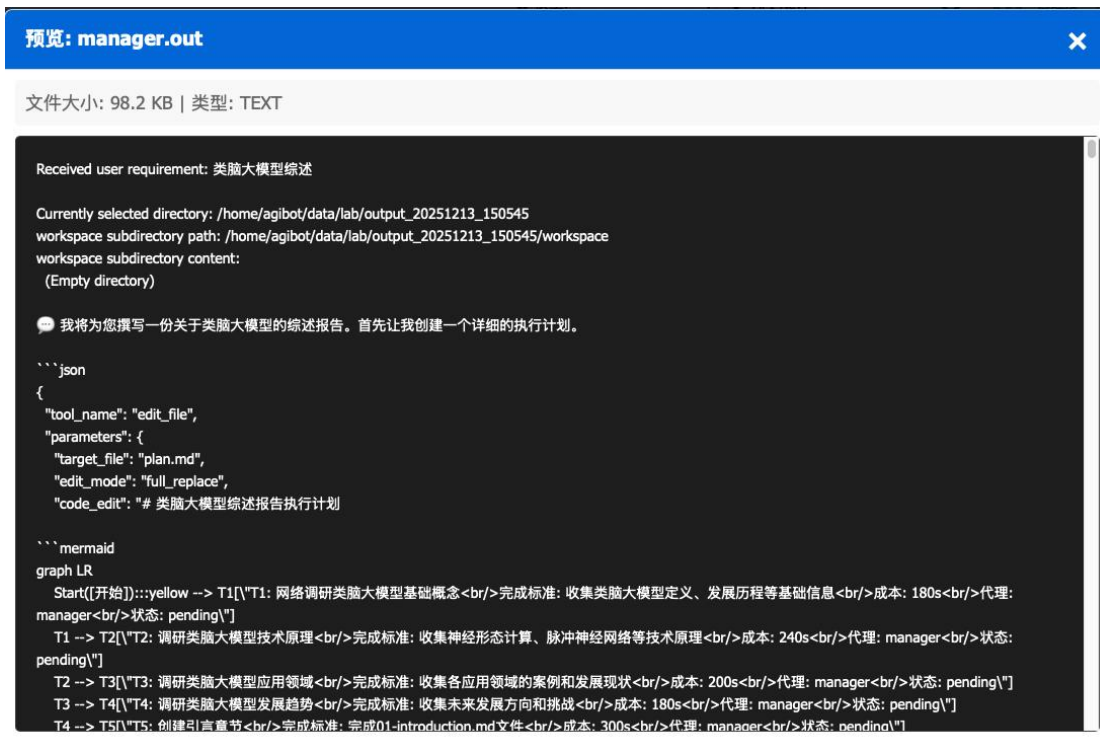


图 16. logs/manager.out 可查看输出日志，logs/manager.log 可查看更多细节输出信息

注：目前并不是所有文件格式都可以预览，目前工作目录支持如下格式预览：各类代码文件，Markdown 文件（md），PDF 文件，json 文件，mermaid 源代码文件（mmd），图像文件（SVG 文件，PNG 文件），使用手册

json 文件, txt 等文本文件, html 网页等。Word/PPT/Excel 仍需要下载后查看。

6.2.4 用户资料的上传

工作区可以上传各类文件, 目前支持 Word (Docx 格式), Excel (xlsx 格式), PDF, 图像及各类代码语言的识别。用户可直接点击上传按钮, 直接上传即可。上传后可通过提示词进行读取或进行可视化展示, 例如上传了某个 excel 文档, 可以写提示词: 将 excel 内容转化为可视化图像。上传的纯文本或代码, 会自动进行文件索引, 并可在撰写代码时使用。其他多模态格式不会被索引。

6.2.5 多轮交互功能

当用户对某个文件的效果不满意, 或希望在提示词中指定某个文件时, 可以用鼠标将某个文件拖动到需求描述栏, 放开左键后, 需求描述栏会出现@文件名形式的提示, 代表需求提示词中提及了某个文件。用户也可以自行在需求框输入文件名或模糊的文件名, 大模型会调用工具进行文件检索。

7 典型应用案例

7.1 专业图文报告

步骤:

1. 点击技能 ->  推荐) 专业文档: 生成完整深度的图文报告 , 或者其他类型的文件, 例如


专业文档: 专利交底书

彩文: 图文博客+公众号+小红书

2. 在需求提示框中输入报告的要求
3. 可选的, 可以上传自己的资料到工作目录
4. 点击回车或者任务发送按钮启动任务
5. 报告会分章节撰写, 最后合并为长文, 并同步给出 pdf/word 格式

7.2 写一个在线小程序

步骤:

1. 点击技能->  编程: 写个小游戏或软件(网页运行)
2. 在需求提示框中输入报告的要求
3. 生成完毕后, 工作目录会出现 html 格式的文件, 双击即可运行

7.3 用户文件分析

步骤:

直接输入需求, 例如“根据@文件 1 @文件 2 写一些数据分析代码, 绘制图表, 之后总结报告”, 其中文件 1 和文件 2 是你上传的文件, 可以从工作目录中拖拽到需求窗口。支持图像、word、pdf、excel 等常用格式。

附 1：高级配置

除了基础配置外，config.txt 文件还提供了丰富的高级配置选项以满足专业用户对富图像文档生成、多智能体协同等功能需求，参考如下表格。

表 1. 高级配置选项

基础配置

参数名	类型	默认值	说明
LANG	字符串	zh	语言设置，en 为英文，zh 为中文

API 配置

参数名	类型	默认值	说明
api_key	字符串	无	API 密钥，用于访问大模型服务
api_base	字符串	无	API 基础地址，大模型服务的端点 URL
model	字符串	无	使用的模型名称（如 glm-4.6, gpt-4.1, claude-sonnet-4-0 等）
max_tokens	整数	无	模型输出的最大 token 数量，根据模型能力设置

注意：配置文件中可以配置多个 API 提供商 (OpenAI、Anthropic、Zhipu、Bailian、Google Gemini、DeepSeek、Volcengine Doubao、Moonshot、SiliconFlow、Ollama 等)，但只有最后一个配置生效。

工具调用

参数名	类型	默认值	说明
Tool_calling_format	布尔值	False	工具调用格式，True 为标准工具调用 API，False 为基于聊天的工具调用

系统配置

参数名	类型	默认值	说明
gui_default_data_directory	字符串	../data	GUI 默认用户数据目录, 用于显示文件列表和管理工作区目录

文件转换配置

参数名	类型	默认值	说明
auto_convert_to_word	布尔值	True	是否自动将 Markdown 文件转换为 Word 格式
auto_convert_to_pdf	布尔值	True	是否自动将 Markdown 文件转换为 PDF 格式
auto_convert_to_latex	布尔值	False	是否自动将 Markdown 文件转换为 LaTeX 格式

如果用户不希望调整，可以直接保持默认配置即可。

系统还具有长期记忆功能的详细配置，包括记忆存储路径、RAG 检索阈值、摘要触发长度等参数，确保历史经验能够有效指导当前任务执行。参考 config_memory.txt 文件。

附 2. 工具配置

智能体默认配置的工具列表位于 prompts/tool_prompt.json 中，每个工具为一个 json 字段，目前具备十余个内建工具，包含文件操作、网络检索、代码索引等功能，一些额外（不常用工具）位于 prompts/additional_tool_prompt.json 中；内建工具列表芥蓝见表 2。

表 2. 工具概览表

工具名称	工具作用	必需参数	可选参数
workspace_search	在代码库中进行语义搜索，查找与查询最相关的代码片段	query	target_directories
read_file	读取文件内容，支持读取整个文件或指定行范围	target_file, should_read_entire_file	start_line_one_indexed , end_line_one_indexed _inclusive
run_terminal_cmd	执行终端命令	command	-
grep_search	基于正则表达式的快速文本搜索，查找精确的模式匹配	query	case_sensitive, exclude_pattern, include_pattern
edit_file	编辑现有文件或创建新文件	target_file, edit_mode, code_edit	old_code
web_search	搜索网络获取实时信息	search_term	-
search_img	根据查询获取相关图片，	query	-

	保存到本地文件，返回图片文件路径列表		
fetch_webpage_content	从网页 URL 获取内容进行分析或信息提取	url	search_term
talk_to_user	向用户显示问题或消息并等待键盘输入（带超时）	query	timeout
get_sensor_data	获取物理世界信息，包括图像、视频、音频和传感器数据	type, source	para
merge_file	将多个文件合并为单个输出文件	file_list, output_file	-
convert_docs_to_markdown	将各种文档格式转换为 Markdown 格式	file_path	-
compress_history	使用 AI 摘要压缩对话历史以减少上下文长度	-	keep_recent_rounds
plan_tools	根据搜索查询规划和动态加载下一轮的工具	query	-
idle	空闲工具，表示本轮不执行任何操作，主要用于多智能体同步	-	reason, sleep