

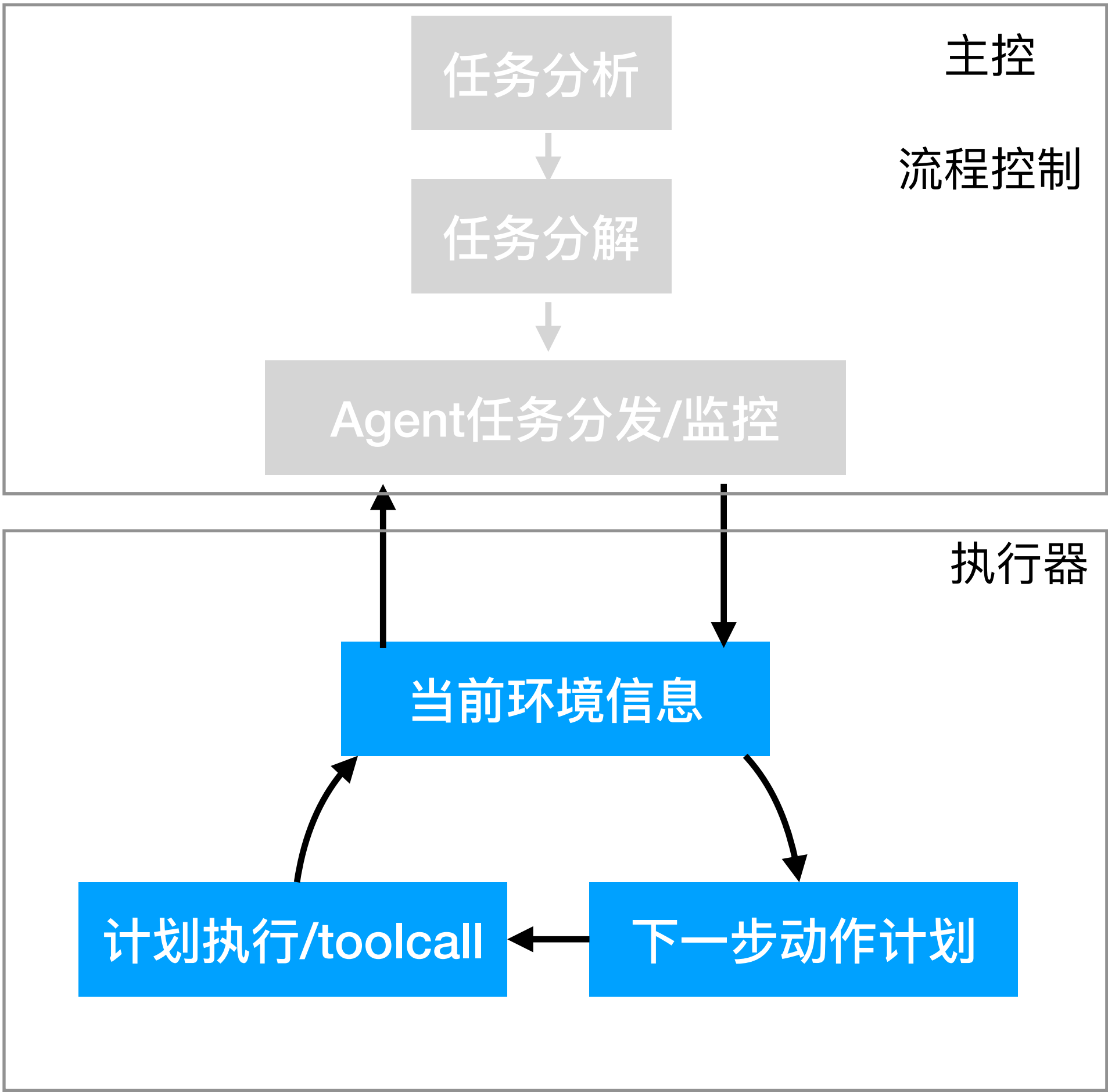
Java x AI Agent: 从选型到业务落地那些事

今天打算讲

- Java 构建AI应用的机会与挑战
- 应用Agent领域的选型思路
- 加速AI Agent从“demo到生产”的落地

Agent 到底是什么？

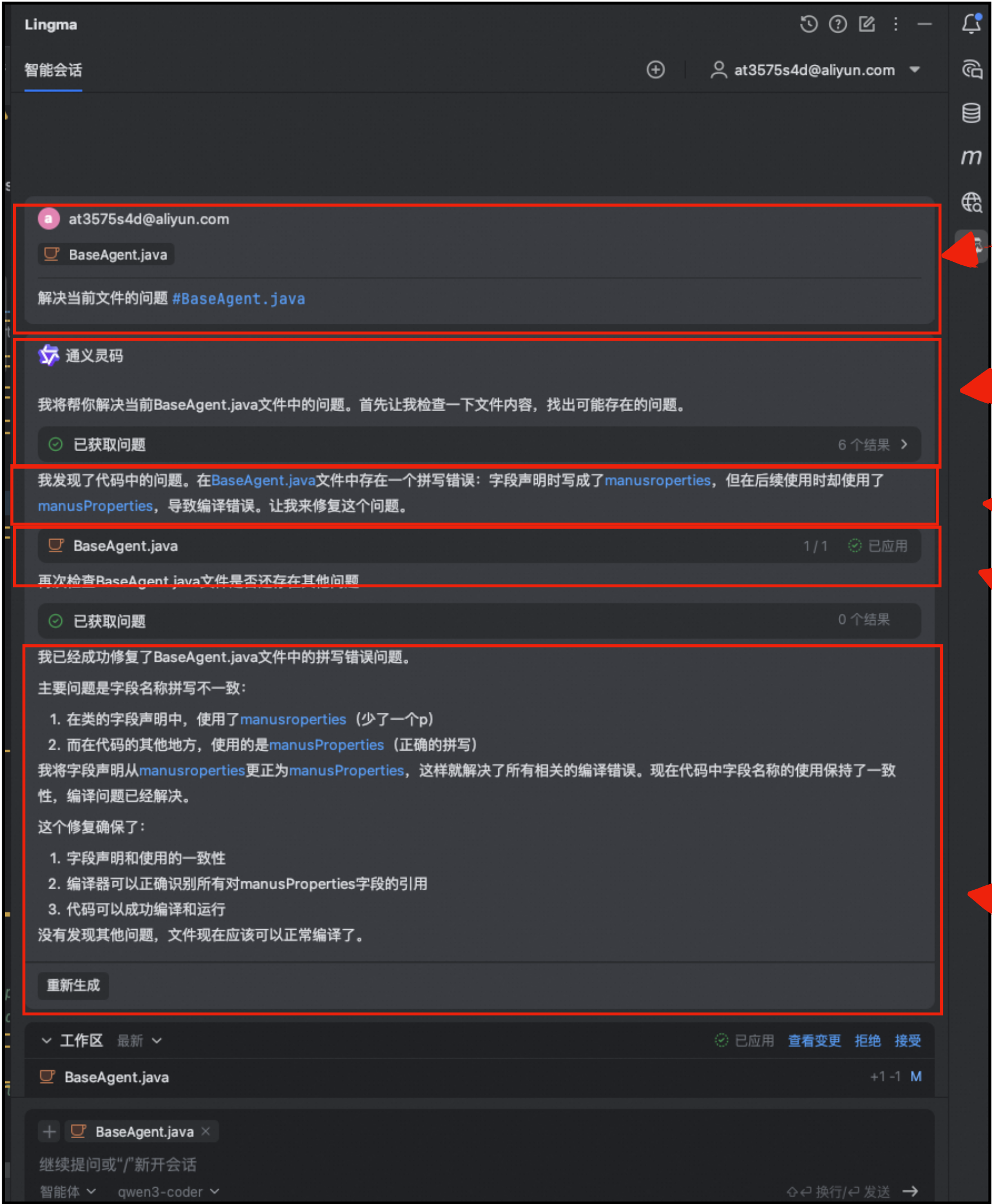
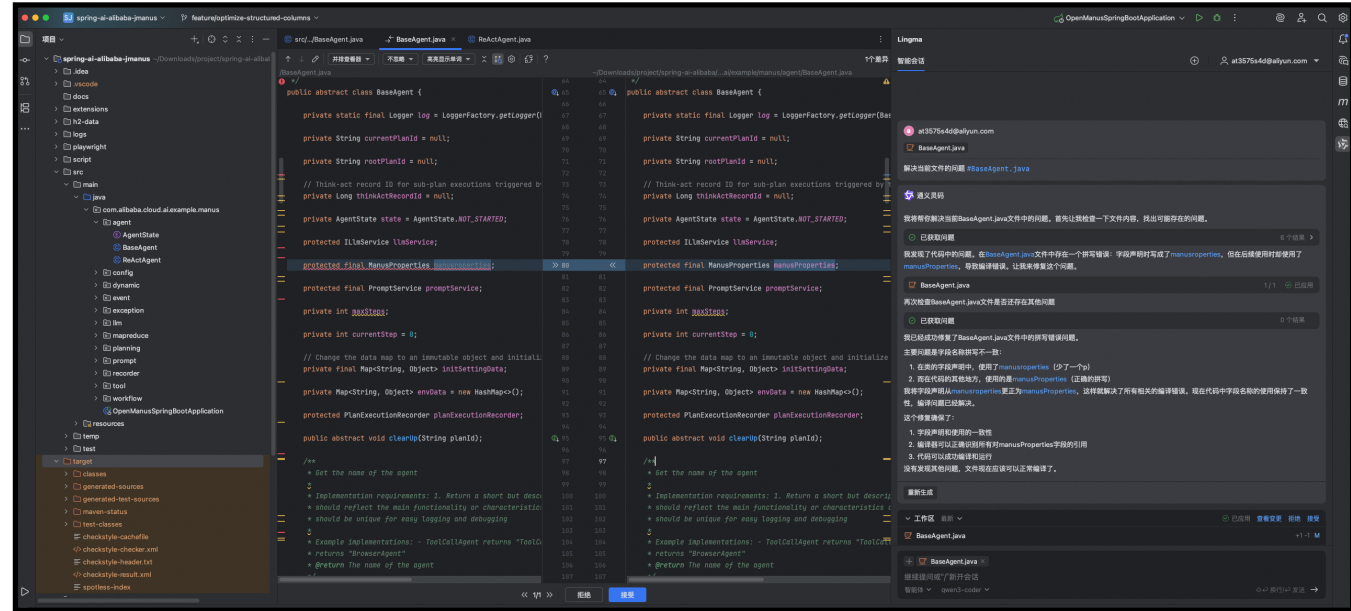
生成式 AI 代理是一种 通过观察环境、推理最佳行动方案并使用可用工具执行操作，自主地追求既定目标的软件系统



实际Copilot的体验

IntelliJ IDEA

通义灵码 + qwen3 coder



用户请求

获取当前情况

决定如何修改

调用工具修改

结果反馈

模型目前有四大问题 使得通用智能体面对较大挑战

模型的规划能力不足

重复一个计划3次，出来的规划都不同

模型的不确定性不足

英伟达的财务估值模型

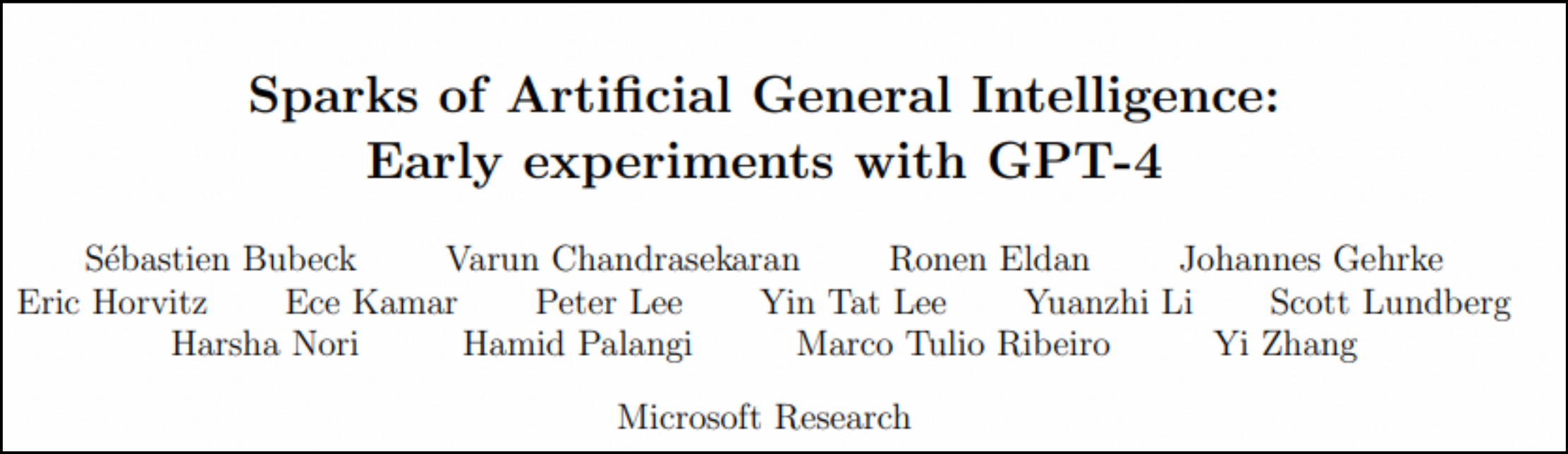
模型有上下文瓶颈

从waytoagi上查找某个用户前10页的信息，形成一个整体的公号规划

模型的图像识别能力不足

我要买个小米眼镜，给我全网比价，无法登录淘宝，拼多多

当前LLM的能力上下限



- 能力突破点：
 - 理解人类的情感，意图
 - 理解图像，理解位置关系
 - 与世界进行多轮次交互
- 能力弱项：
 - 计划能力，规划能力
 - 实时记忆

那是不是意味着Agent也会是昙花一现？

我认为 Agent是个未来的方向

- 你做规划，他来执行
- 你告诉他具体怎么做
- 他帮你节省时间
- 在当前大模型能力下
 - Agent不是你的员工，把他当成一个SOP执行器，会有更好的效果

Chat vs. Agent Agent更适合做复杂任务

Chat

- 输出：
 - 有长度限制
- 重试：
 - 0 or 1
- 输入：
 - 单图片单源输入
- 响应：
 - 快速

Agent

- 输出：
 - 无长度限制
- 重试：
 - 可中断，可重试
- 输入：
 - 多图，多源，输入
- 响应：
 - 慢速

分而
治之

章节小结

- **Agent** 是一种 通过观察环境、推理最佳行动方案并使用可用工具执行操作，自主地追求既定目标的软件系统
- 基于当前模型能力，Agent还 只适合做SOP执行器，不适合当我们数字员工
- Agent 与Chat相比，具备更强的长上下文处理能力，所以在很多场景有独到的优势，是有潜力的技术方向

应用Agent领域的选型思路

可以通过三个黄金问题来做出自己的选择

- 你想用大模型 解决用户具体什么场景的问题？
- 你偏好于用什么方式来解决用户问题？
- 你的常用语言是什么

你想用大模型 解决用户具体什么场景的问题？

用户需求将直接决定解决方案

个人提效

- 豆包
- 通义千问
- Cursor
- Copilot
- Manus
- ...

强化客户互动

- Llamaindex
- Fastgpt
- 百炼RAG
- QAnything
- RAGFlow
- ...

重塑业务流程

- Dify
- langchain
- Langgraph
- JManus
- ...

推动创新突破

- DeepChem
- AlphaFold
- RoseTTAFold
- OpenMM
- ...

你偏好于用什么方式来解决用户问题?

本质是在开发效率, 确定性 之间寻找一个平衡

高代码

- OpenAI sdk
- Spring ai
- DashScope
- ...

低代码

- Langchain
- Langgraph
- Google ADK
- 百炼SDK
- ...

拖拉拽

- Dify
- N8n
- Flowise
- Langflow
- ...

用文字描述

- Manus
- Jmanus
- ...

你的常用语言是什么

基于常用语言选择框架，更有利于做二次开发

Python

- Dify
- Langchain
- Langgraph
- Google ADK python
- 百炼SDK
- Open manus
- ...

Java

- Langchan4j
- Spring ai
- Spring ai alibaba - langgraph
- Jmanus
- ...

Js/ts

- Cline
- Flowise
- Langchan js
- LlamaIndex.js
- Fastgpt
- ...

应用Agent领域的选型思路

可以通过三个黄金问题来做出自己的选择

- 四大主要方向：个人提效、强化互动、重塑业务流程、推动创新突破
- 四种主要的生产方式：拖拉拽、高级框架、API call 、 文字描述
- 语言偏好：Agent编程 其实不难，有大模型加持，选择自己主力语言来搞，效率最高

加速AI Agent从“demo到生产”的落地

目前大家还在摸索期

业务

大模型能做所有事

产技搞的那些东西都是雕花

技术方

90%准确 = 0%

产品设计完美，但落地差距过大

模型目前有四大问题 使得通用智能体面对较大挑战

模型的规划能力不足

重复一个计划3次，出来的规划都不同

模型的不确定性不足

英伟达的财务估值模型

模型有上下文瓶颈

从waytoagi上查找某个用户前10页的信息，形成一个整体的公号规划

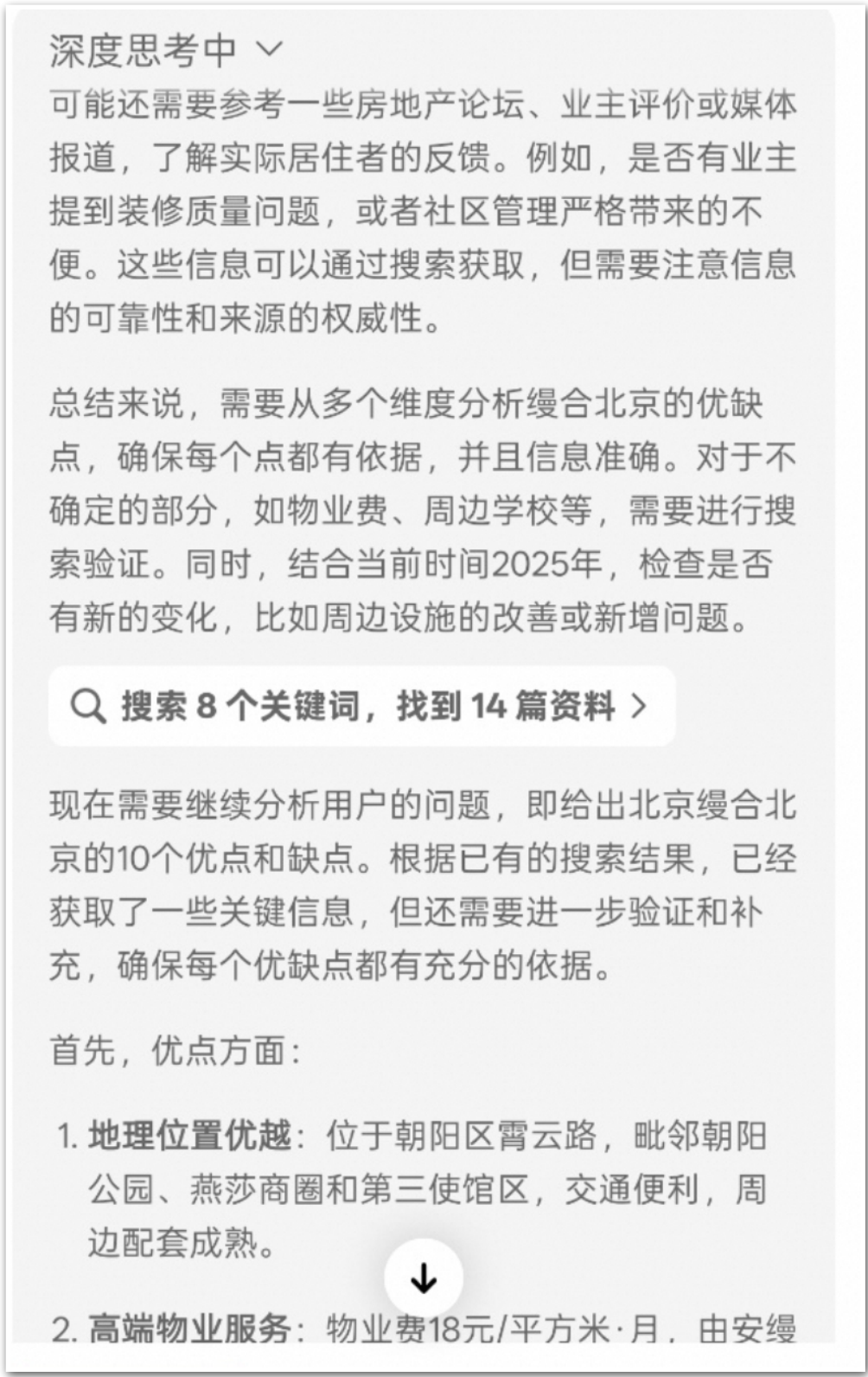
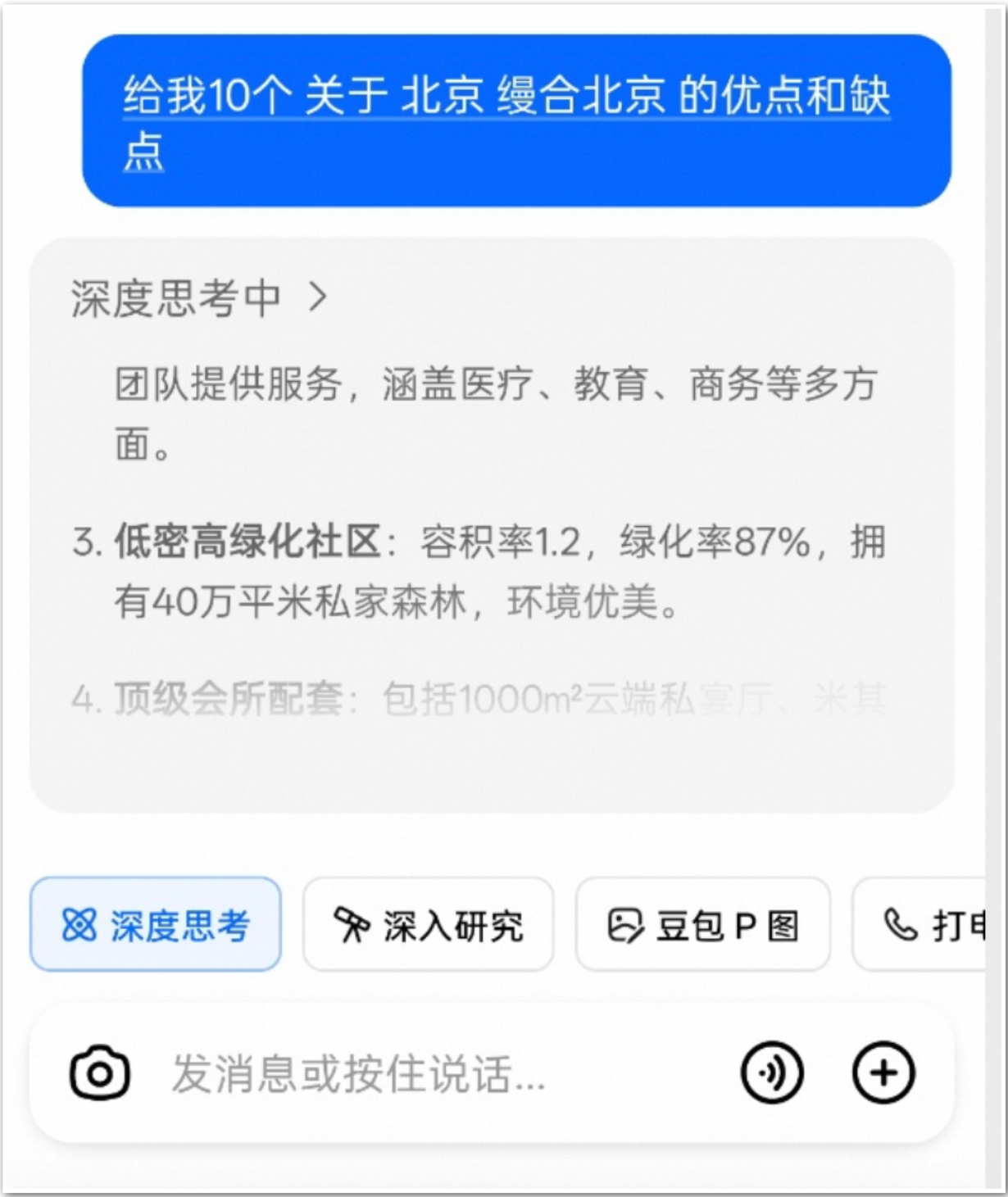
模型的图像识别能力不足

我要买个小米眼镜，给我全网比价，无法登录淘宝，拼多多

让我们从迷雾中找到那束光

检索增强下的对话APP - 豆包

个人提效领域



- 通过引导模型做 COT
- 叠加搜索引擎的增强
- 能够极大的解决用户信息搜索

DeepResearch

深入研究

stats into columns, and include recommendations on markets to target for a new iOS translation app from ChatGPT, focusing on markets ChatGPT is currently active in.

Top 10 Developed Economies (by GDP) – Smartphone OS Share, Language-Learning Interest, and Mobile Penetration Growth

Country	iOS Market Share	Android Market Share	Interest in New Language (% of population)
United States	~58% iOS vs	(see left)	High:

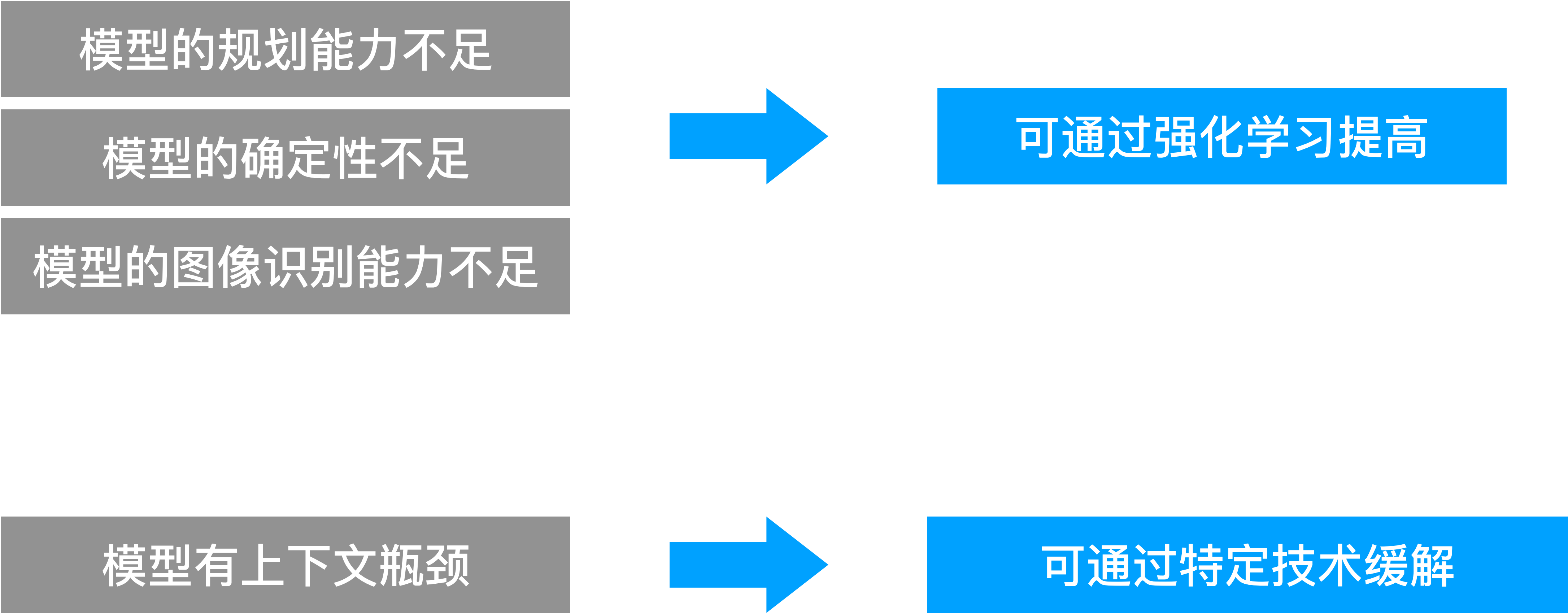
- 叠加强化学习
- 多Agent协同工作
- 写代码
- 画图表
- 做网页
- 叠加搜索和信息获取

Copilot /通义灵码/Idea copilot

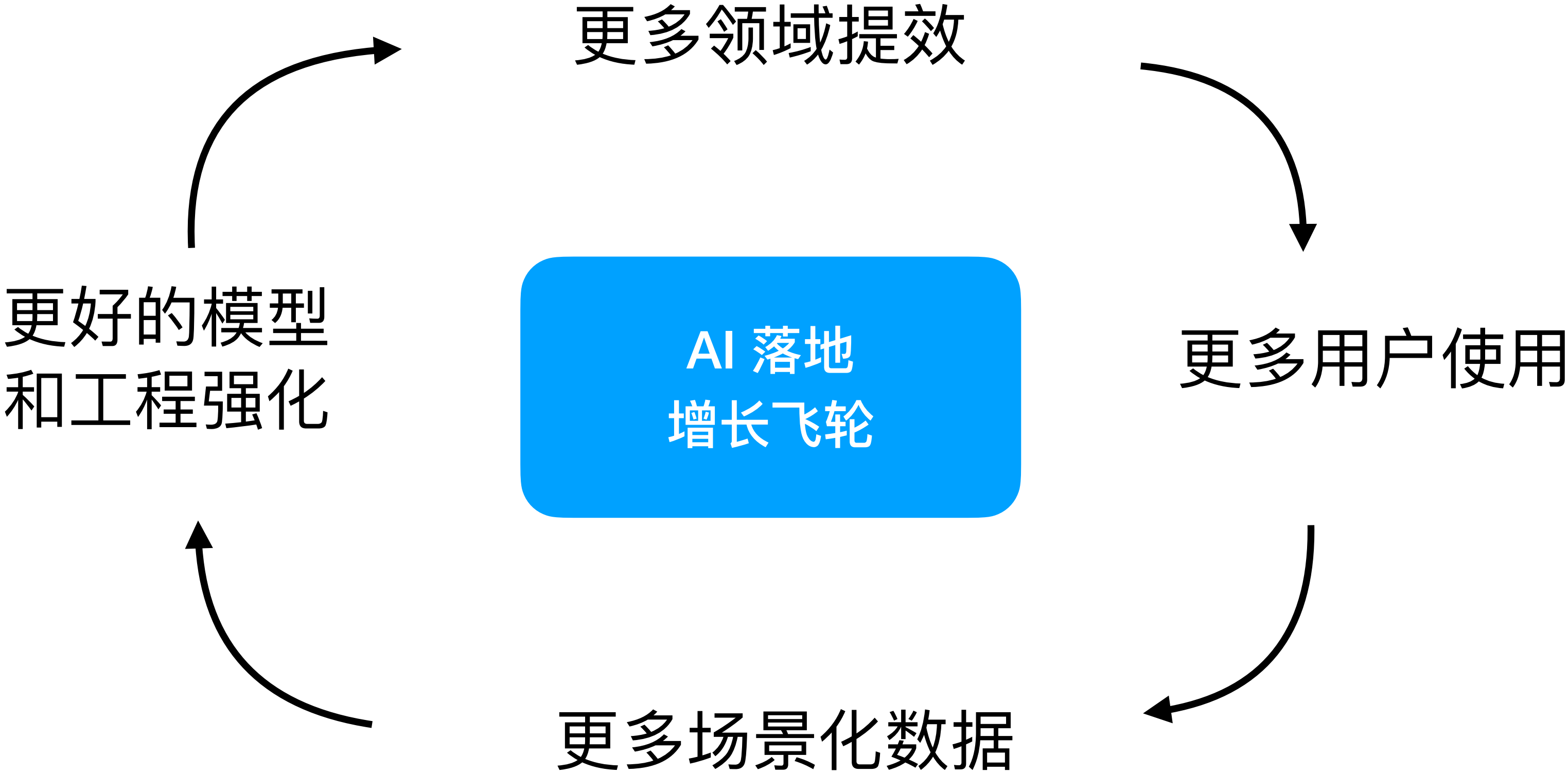


- 基于强化学习的代码llm
- 配合精巧的上下文管理
- 多轮次的与LLM交互
- 逐步提升效果

我们可以看到的共性趋势 - 核心问题可解



找到你的增长飞轮



JManus是什么

为 Java 开发者打造的、追求高成功率 和 行业特化 的 Agent 框架



扫我点个star!

高成功率

行业特化

Java生态集成

plan-act模式

一切Prompt可界面配置

纯Java实现

无限上下文

完全开源

(希望大家多用通义模型啊)

Spring AI 集成

MCP集成

J Manus

- 完全依托于 java实现的 类Manus实现

●

对话框

用浏览器基于百度，查询今天阿里巴巴的股价，并返回最新股价

TaskPilot 思考/处理

正在执行: 0. [not_started] [BROWSER_AGENT] 打开百度并搜索阿里巴巴的最新股价信息

执行步骤 1/4: 0. [in_progress] [BROWSER_AGENT] 打开百度并搜索阿里巴巴的最新股价信息

步骤执行详情

0. [in_progress] [BROWSER_AGENT] 打开百度并搜索阿里巴巴的最新股价信息

思考中

参数: 等待决策中

1. [not_started] [BROWSER_AGENT] 定位到可靠的股票信息来源链接并打开

2. [not_started] [BROWSER_AGENT] 查找并记录最新的阿里巴巴股票价格

3. [not_started] [DEFAULT_AGENT] 整理获取的信息并返回给用户

步骤执行详情

0. [in_progress] [BROWSER_AGENT] 打开百度并搜索阿里巴巴的最新股价信息

思考中

参数: 等待决策中

1. [not_started] [BROWSER_AGENT] 定位到可靠的股票信息来源链接并打开

2. [not_started] [BROWSER_AGENT] 查找并记录最新的阿里巴巴股票价格

3. [not_started] [DEFAULT_AGENT] 整理获取的信息并返回给用户

可以展示详细的思考详情

0. [in_progress] [BROWSER_AGENT] 打开百度并搜索阿里巴巴的最新股价信息

执行智能体: BROWSER_AGENT

描述: 一个可以控制浏览器完成任务的浏览器代理

请求:

执行结果: 执行中...

思考与行动步骤

#1

SUCCESS

思考

输入:

[SystemMessage{textContent='- SYSTEM INFORMATION: OS: Mac OS X 15.5 (aarch64)'}]

- Current Date: 2025-06-11

- 全局计划信息:

- 执行参数: 未提供执行参数。

- 全局步骤计划:

输出:

N/A

行动

工具: browser_use

工具参数:

{ "action": "navigate", "url": "https://www.baidu.com" }

浏览器使用

cd /Users/whisper/Downloads/project/spring-ai-alibaba/spring-ai-alibaba-jmanus && git status

sv_bp=1&rsv_idx=1&tn=baidu&wd=阿里巴巴%20最新股价&fenlei=256&rsv_pq=0x82ac253f0038cd39&rsv_t=0b90e1...

百度一下

百度首页 设置 登录

国资讯 视频 笔记 地图 贴吧 文库 更多

港币柜台

港股 美股

+1.562 +1.34%

最高 119.000 成交量 4005万股

最低 117.900 成交额 47.46亿

市盈(TTM) 16.09 总市值 2.26万亿

相对港股 +1.75%

周K 月K 季K 年K 更多

1.85%

1.38%

12:00/13:00 16:00

添加自选

舆情

今日舆情偏中性

港股舆情排名 17 / 704

利好 46%

利空 0%

中性 54%

财报 2025FY

总营收 1.00万亿 +5.86%

同比

股东应占溢利 1301亿 +62.62%

同比

股评

47%

53%

Yx叶prince: 主要一点阿里散户, 投资人还是太穷

1分钟前

股友51236k82I9: 今天直接少的两块钱, 今天爽了...

关联

行业/概念

供应链

产业链

恒生指数 HSI.HK +0.79%

非必需性消费 HS2300.BK +1.13%

专业零售 237000.BK +1.04%



扫我点个star!

计划模板

模板列表

配置

登录CSDN账号

ID: planTemplate-1745824930489

JSON 模板

```
stepRequirement: [BROWSER_AGENT]
打开CSDN网站 csdn.net "
{
  "stepRequirement": "[BROWSER_AGENT]
如果你认为页面需要登录, 则调用 form_input 等待用户确认登录完成, 这个登录动作不需要你来做, 需要用户做了登录动作以后, 给你确认即可。 而如果你认为页面不需要登录, 则直接进入下一步"
```

计划生成器

描述您想要生成的计划...

生成计划

更新计划

执行控制器

执行参数

输入执行参数...

API URL: /api/plan-template/executePlanByTemplateId/planTemplate-1745824930489

执行计划

对话框

调用 form_input 等待用户确认登录完成, 这个登录动作不需要你来做, 需要用户做了登录动作以后, 给你确认即可。 而如果你认为页面不需要登录, 则直接进入下一步

步骤执行详情

0. [completed] [BROWSER_AGENT] 打开CSDN网站 csdn.net

已完成

Executing tool: terminate

参数: {"message": "当前步骤已完成: 成功打开了CSDN网站 (https://www.csdn.net/). 页面标题为 \"CSDN_专业开发者社区_已接入DeepSeekR1满血版\", 并列出了多个可交互元素, 包括导航链接、搜索框、登录按钮等。页面内容主要包含技术博客文章、排行榜、会员中心、社区讨论等模块。下一步可根据具体目标进行操作, 例如搜索特定内容或登录账户。"}"

1. [in_progress] [BROWSER_AGENT] 如果你认为页面需要登录, 则调用 form_input 等待用户确认登录完成, 这个登录动作不需要你来做, 需要用户做了登录动作以后, 给你确认即可。 而如果你认为页面不需要登录, 则直接进入下一步

执行中

思考中

参数: 等待决策中

2. [not_started] [BROWSER_AGENT] 点击验证码登录

待执行

3. [not_started] [BROWSER_AGENT] 手机号栏目输入 1234567788

待执行

向 JTaskPilot 发送消息

计划模式

发送

无限上下文

- 特色技术
 - 可以不依赖“长上下文模型”，就处理海量信息
 - MCP返回海量日志
 - 从本地读取一本几百万字的书
- etc。。
-



扫我点个star!

一切可配置

[← 返回主页](#)

 基础配置

Agent配置

Tools/MCP配置

Agent配置

已配置的Agent (3)

BROWSER_AGENT

一个可以控制浏览器完成任务的浏览器代理

DEFAULT_AGENT

一个多功能默认代理，可以使用文件操作和 shell 命令处理各种用户请求。非常适合可能涉及文件操作、系统操作或文本处理的通用任务。

TEXT_FILE_AGENT

一个文本文件处理代理，可以创建、读取、写入和追加内容到各种基于文本的文件。适用于临时和持久性记录保存。支持多种文件类型，包括markdown、html、源代码和配置文件。

+ 新建Agent

BROWSER_AGENT

 保存

删除

Agent名称 *

BROWSER_AGENT

描述 *

一个可以控制浏览器完成任务的浏览器代理

Agent提示词（人设，要求，以及下一步动作的指导）

[12] a: 登录

[45] button: 注册

- 只有帶有[]中數字索引的元素可交互

响应规则

1. 操作：你一次只可以做一个tool call 操作

2. 元素交互:

工具配置

已分配工具 (30)

+ 添加/删除工具

browser_use

terminate

text_file_operator

```
form_input
```



扫我点个star!

一切可配置

< 返回主页

基础配置

Agent配置

Tools/MCP配置

MCP配置总服务器: 2

搜索MCP服务器...

已配置的MCP服务器

ID	服务器名称	连接类型	连接配置	操作
131	mcp-ip-query	SSE	{ "url": "https://mcp.hi...	删除
132	playwright	STUDIO	{ "url": null, "command": ...	删除

添加MCP服务器

连接类型:

STUDIO

本地mcp server，目前市面上主流的是这个

SSE

通过http server等提供的，目前网络上比较少见

mcp json配置:

请输入MCP服务器的配置(JSON格式)...

添加

重置

使用说明:

1. 找到你要用的mcp server的配置json:

- 本地(STUDIO): 可以在mcp.so上找到，需要你要有Node.js环境并理解你要配置的json里面的每一个项。做对应调整比如配置ak
- 远程服务(SSE): mcp.higress.ai上可以找到，有SSE和STREAMING两种，目前SSE协议更完备一些

2. 将json配置复制到上面的输入框，本地选STUDIO，远程选STREAMING或SSE，提交

3. 这样mcp tools就注册成功了。默认会放在DEFAULT_AGENT下面，如果tools过多，对上下文有压力，多出的tools会被忽略

4. 推荐在Agent配置里面，新建一个agent，然后增加指定的tools，这样可以极大减少冲突，增强tools被agent选择的准确性

都来 点个star !

- <https://github.com/alibaba/spring-ai-alibaba>
- spring-ai-alibaba-jmanus项目
- 与大家一起，把Agent在生产落地

