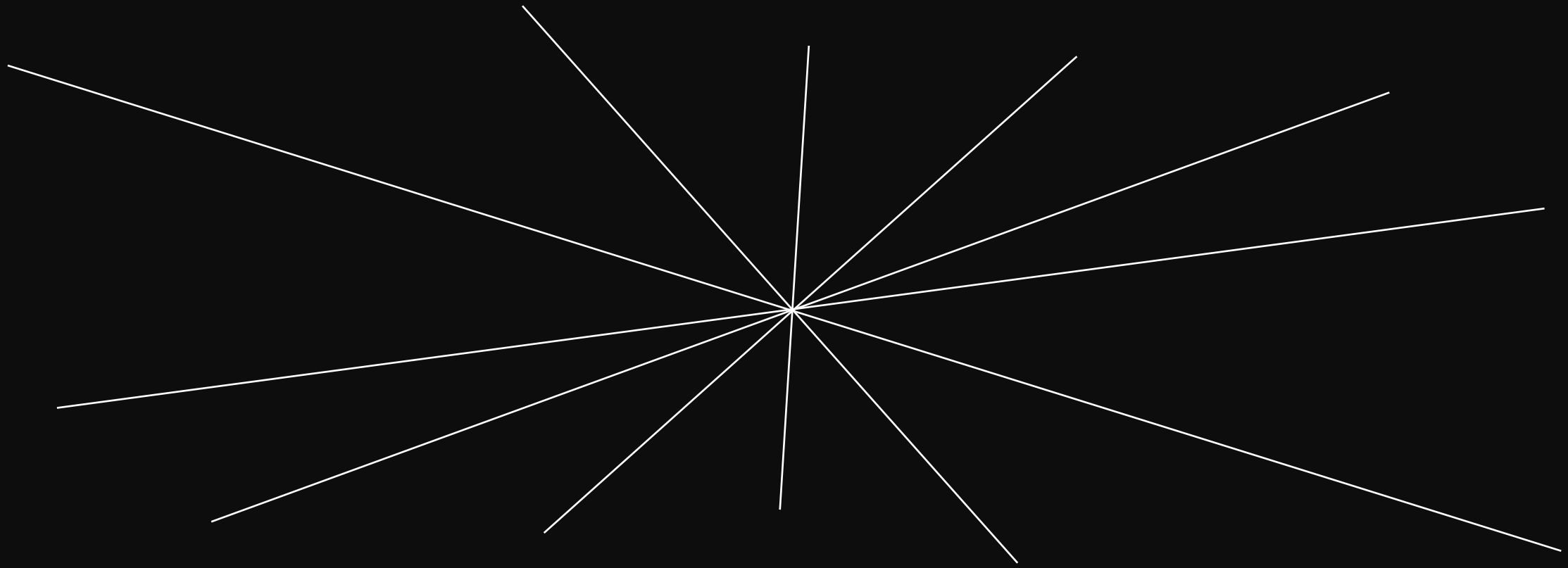
Agentic Workflow



从混沌到秩序：AI重塑了我的工作流

Shared by nimbus, May 2024

nimbus（银海）

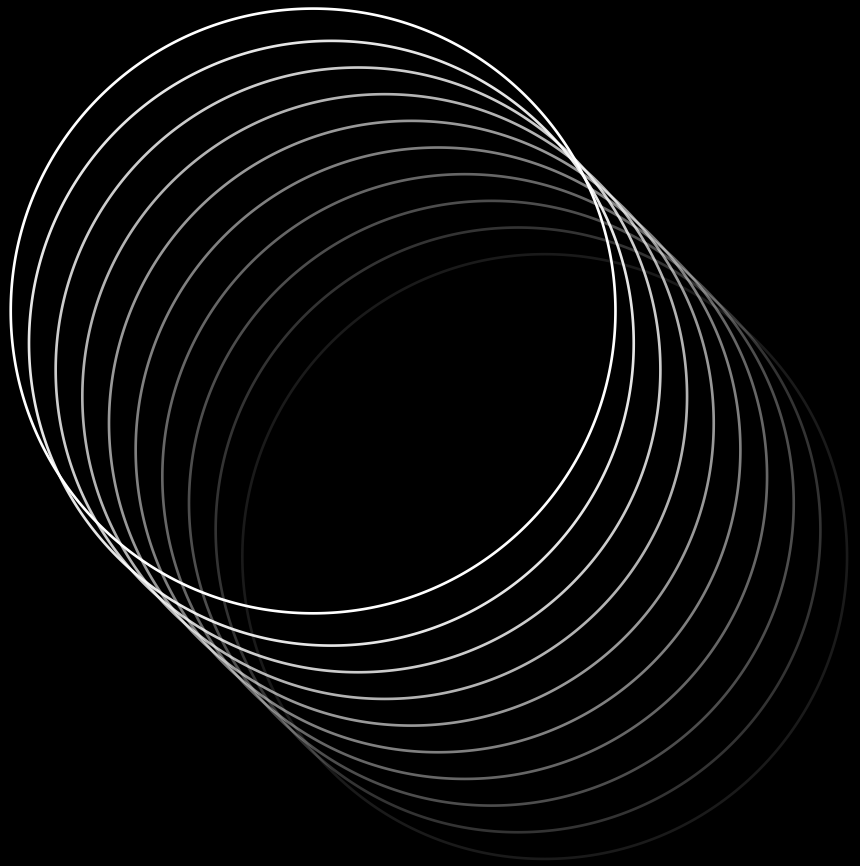
AI 产品一枚，可能来自火星

Datawhale 社区 *AIGC* 板块负责人

WaytoAGI 知识库共创者

百度飞桨、科大讯飞合作讲师

AI- Native 应用：*Pailido* / *AI* 拍立得 作者



Inhai.wiki

初见 *Agentic Workflow*

你好，重新认识一下，工作流!

Agentic Workflow 能玩出花?

- 你将会看到

我对 *Agentic aWorkflow* 也有一些偏见

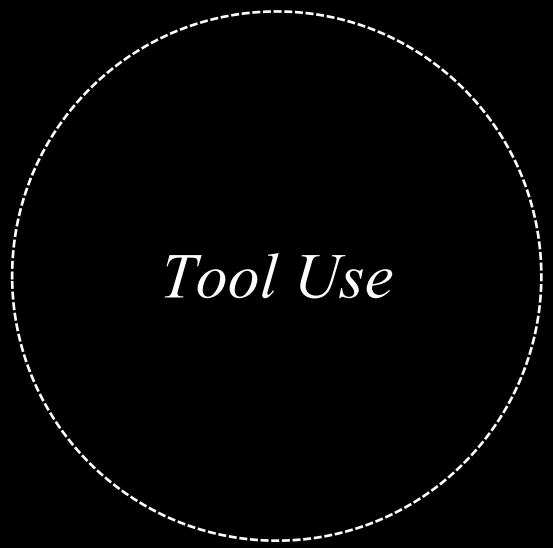
AI 拍立得产品设计逻辑浅谈

Agent 的出现重塑了我的工作流

你好，*AI Agentic Workflow*

在今年的 4 月初，吴恩达老师在美国红杉做了一场演讲，介绍了 4 种主要的 *Agentic Workflow* 设计模式。

让 *Agent* 审视和修正
自己生成的输出



LLM 生成代码、
调用 *API* 等工具进行操作

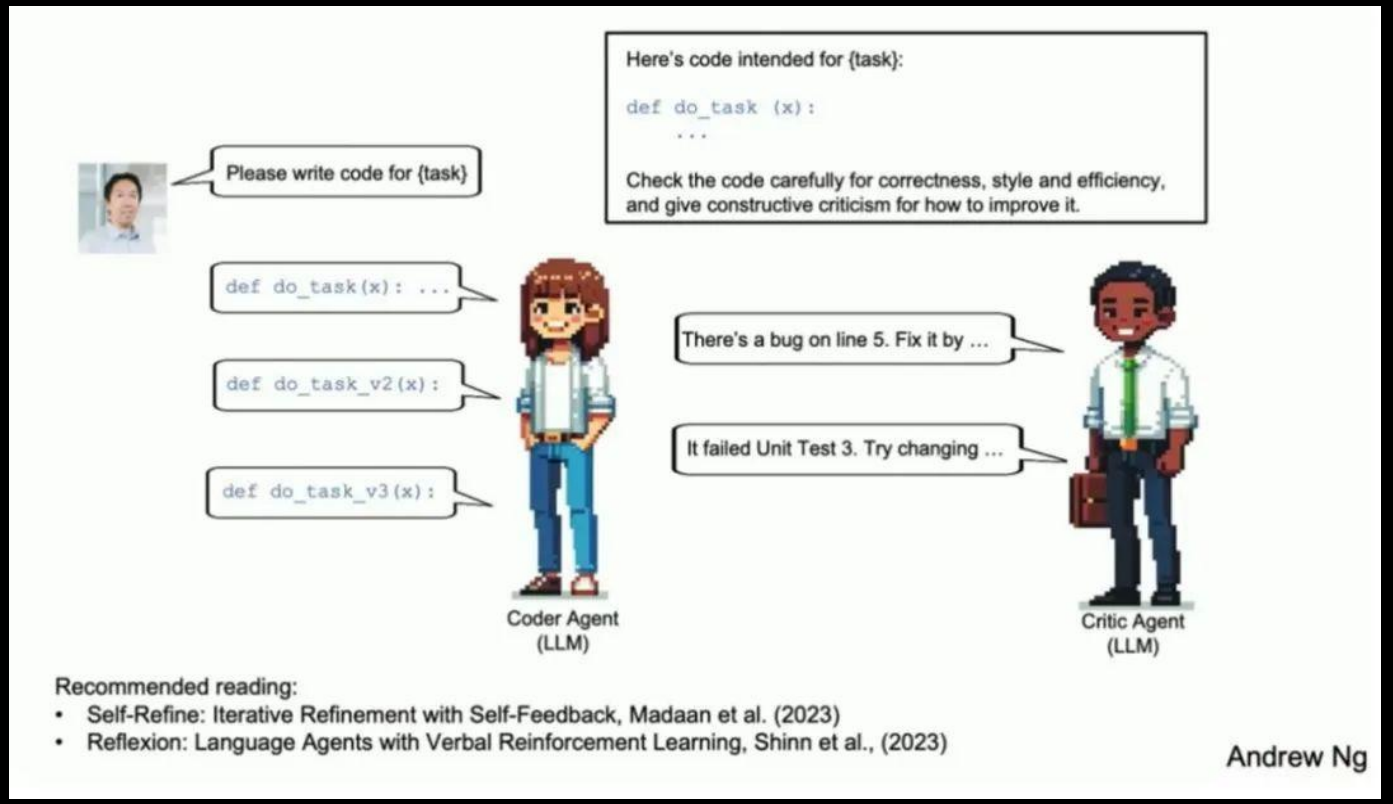
让 *Agent* 分解复杂任务
并按计划执行

多个 *Agent* 扮演不同角
色合作完成任务

Inhai.wiki

Reflection

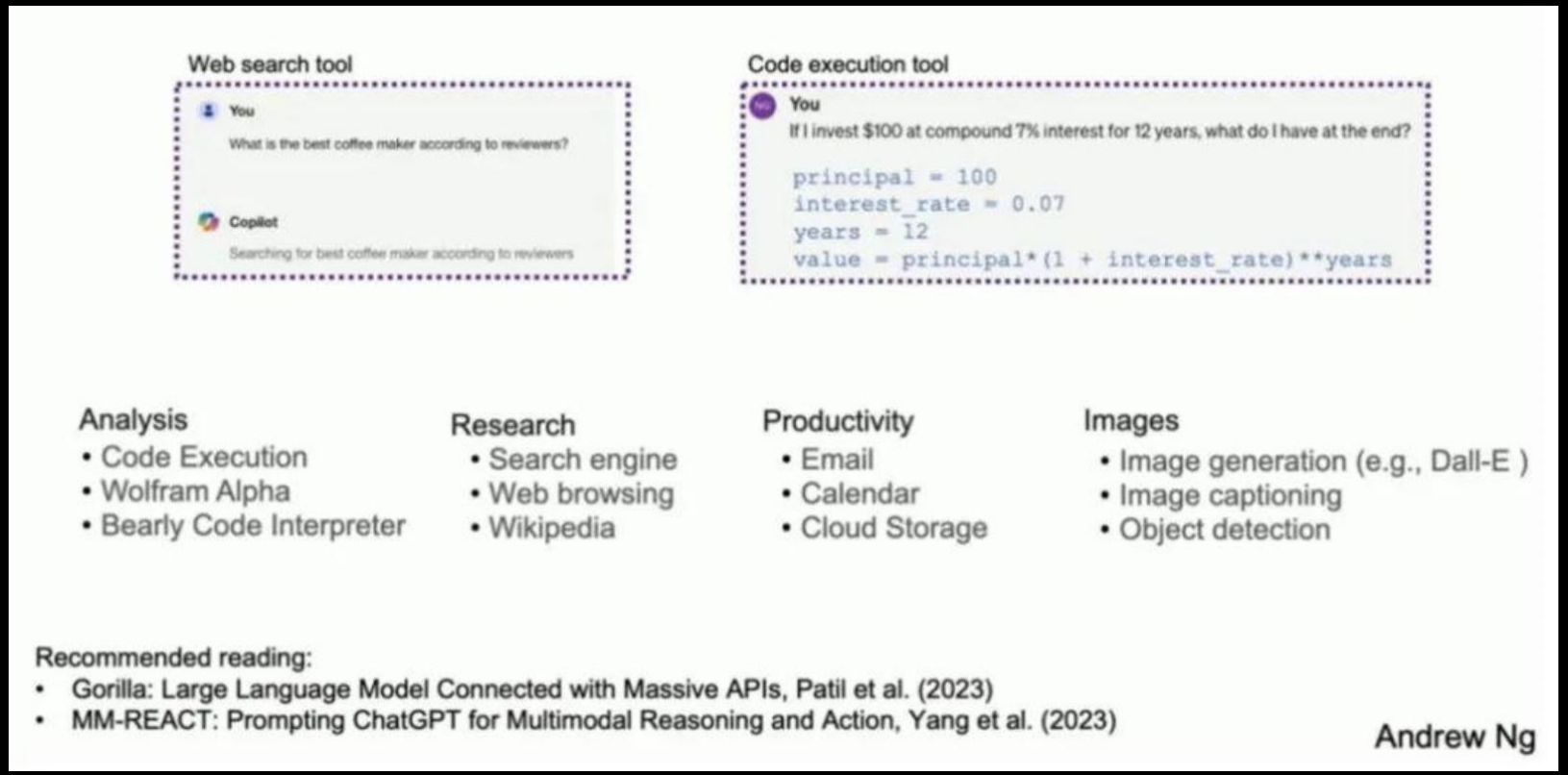
反思其实在根本上是一个博弈的过程，如果你让大模型写一段代码，它会立刻给你反馈。这时你可以将它输出的代码片段再输入回去，让大模型仔细检查代码的准确性和结构规范性，并给出评论。然后，你可以将这些反馈结果再次输入给大模型，它可能会输出一个比第一版更好的代码，如果有两个 Agent：一个负责 Coding，另一个负责 Code Review，效果会更佳。



图为：两个 *Agent* 在博弈写代码的过程

Tool Use

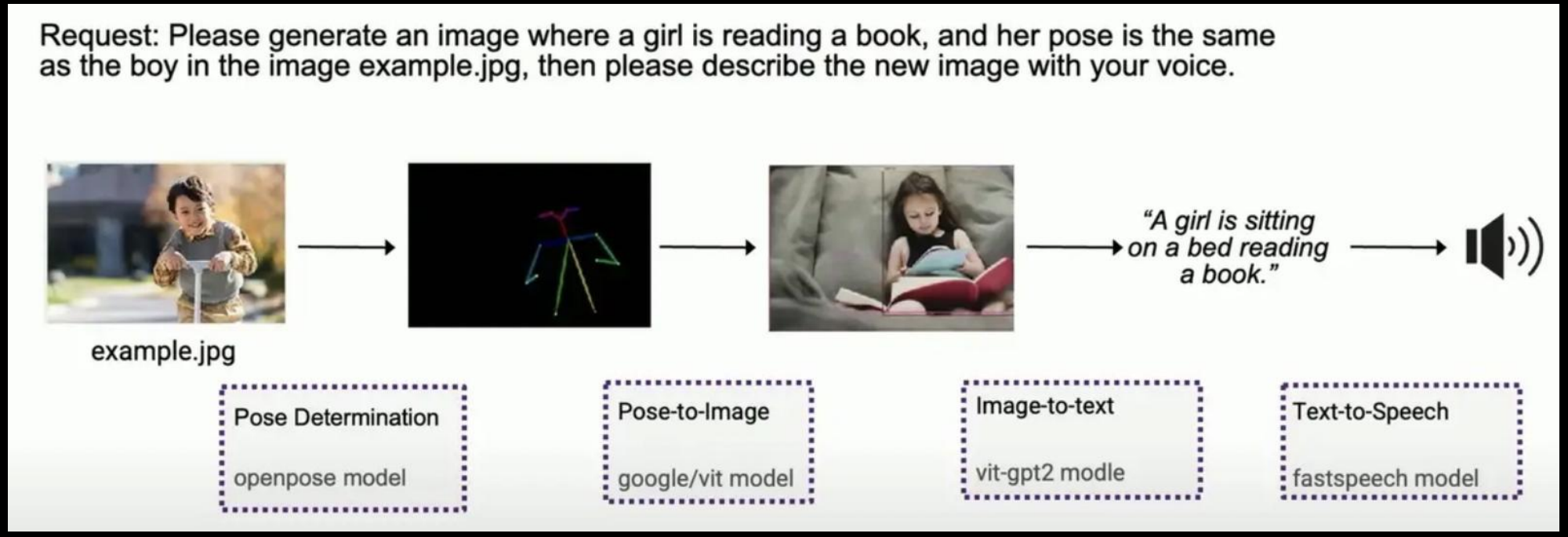
如果大家使用 *Kimi Chat* 来查询某个问题，你会发现它会在互联网上检索相关内容，并基于检索结果进行总结分析，最后给出结论。这其实是大模型利用「网页搜索」工具的一个典型例子。



图为：罗列了多种工具的使用

Planning

Agent 通过自行规划工作流路径，面向于简单的或者一些线性流程的运行。比如下图中：*Agent* 会先识别男孩的姿势，并可能找到一个姿势提取模型来识别姿势，在接下来要找到一个姿势图像模型来合成一个新的女孩图像，然后再使用图像理解文本的模型，并在最后使用语音合成输出，完成这个流程任务。



“请生成一张图片，其中一个女孩正在阅读一本书，她的姿势与图片 *example.jpg* 中的男孩相同，然后用声音描述新的图像。”

Multiagent Collaboration

吴恩达通过开源项目 *ChatDev* 进行举例，你可以让一个大语言模型扮演不同的角色，比如让一个 *Agent* 扮演公司 *CEO*、产品经理、设计师、代码工程师或测试人员，这些 *Agent* 会相互协作，根据需求共同开发一个应用或者复杂程序。

图为：*ChatDev* 工具中多个 *Agent* 协作完成任务

Multiagent Debate

Task	Single agent	Multi-agent
Biographies	66.0%	**73.8%**
MMLU	63.9%	**71.1%**
Chess move	29.3%	**45.2%**

(Du et al., 2023)

图为：单个 *Agent VS* 多个 *Agents*

AI Agent 基本框架

OpenAI 的研究主管 *Lilian Weng* 曾经写过一篇博客叫做《 LLM Powered Autonomous Agents 》，其中就很好的介绍了 *Agent* 的设计框架，她提出了“*Agent = LLM* + 规划 + 记忆 +工具使用”的基础架构，其中大模型 *LLM* 扮演了 *Agent* 的“大脑”。

Planning

规划

主要包括子目标分解、反思与改进。将大型任务分解为较小可管理的子目标处理复杂的任务。而反思和改进指可以对过去的行动进行自我批评和自我反思，从错误中学习并改进未来的步骤，从而提高最终结果的质量。

Memory

记忆

分为短期记忆和长期记忆。其中短期记忆是指的将所有的上下文学习看成是利用模型的短期记忆来学习；而长期记忆是提供了长期存储和召回信息的能力，它们通常通过利用外部的向量存储和快速检索来存储和召回信息。

Tools

工具

通过学会调用外部不同类型API来获取模型权重（通常在预训练后很难修改）中缺少的额外信息，包括当前信息，代码执行能力，访问专有信息源等（例如获取此时此刻的天气、联网网搜索等）

Action

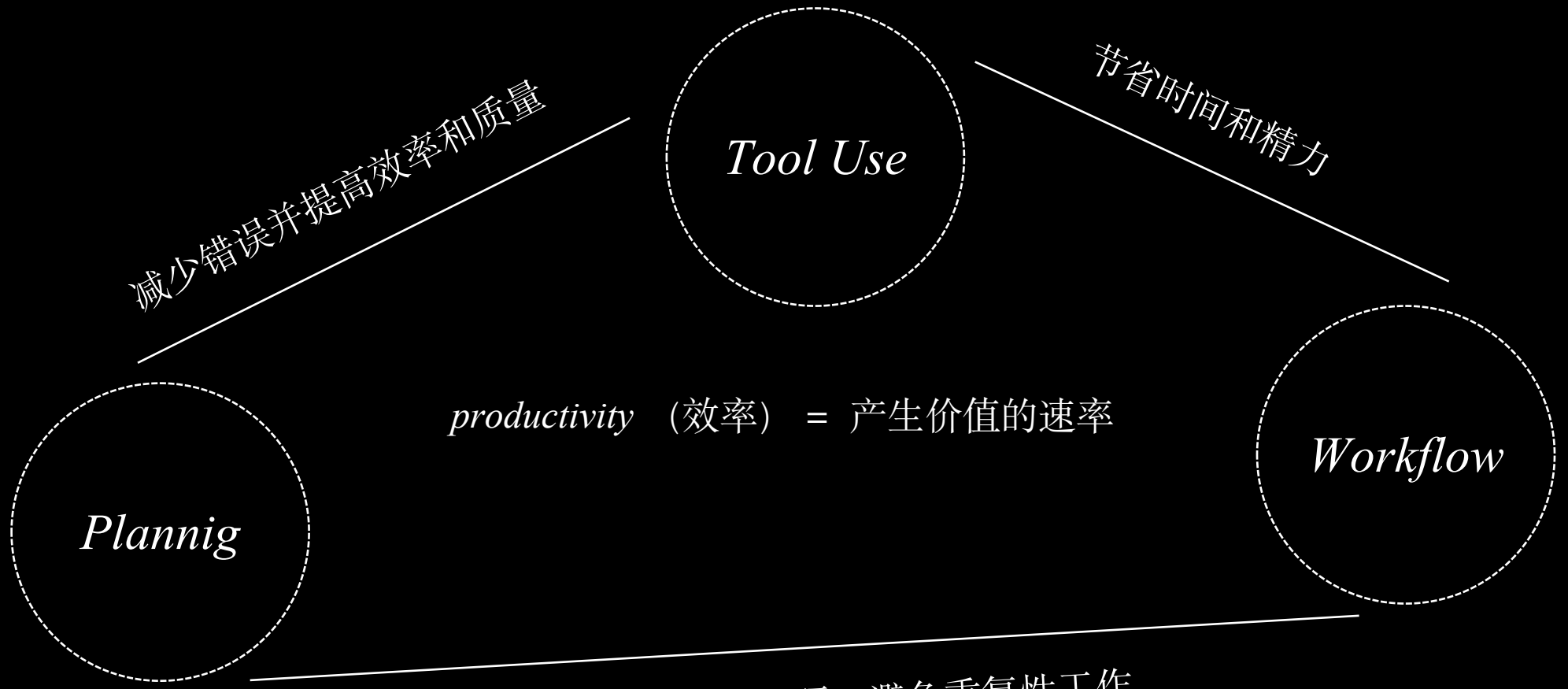
动作

根据上述大模型结合问句 *(Query)* 、上下文的规划 *(Context)* 、各类工具，最终大模型才能决策出最终需要执行的动作是什么。

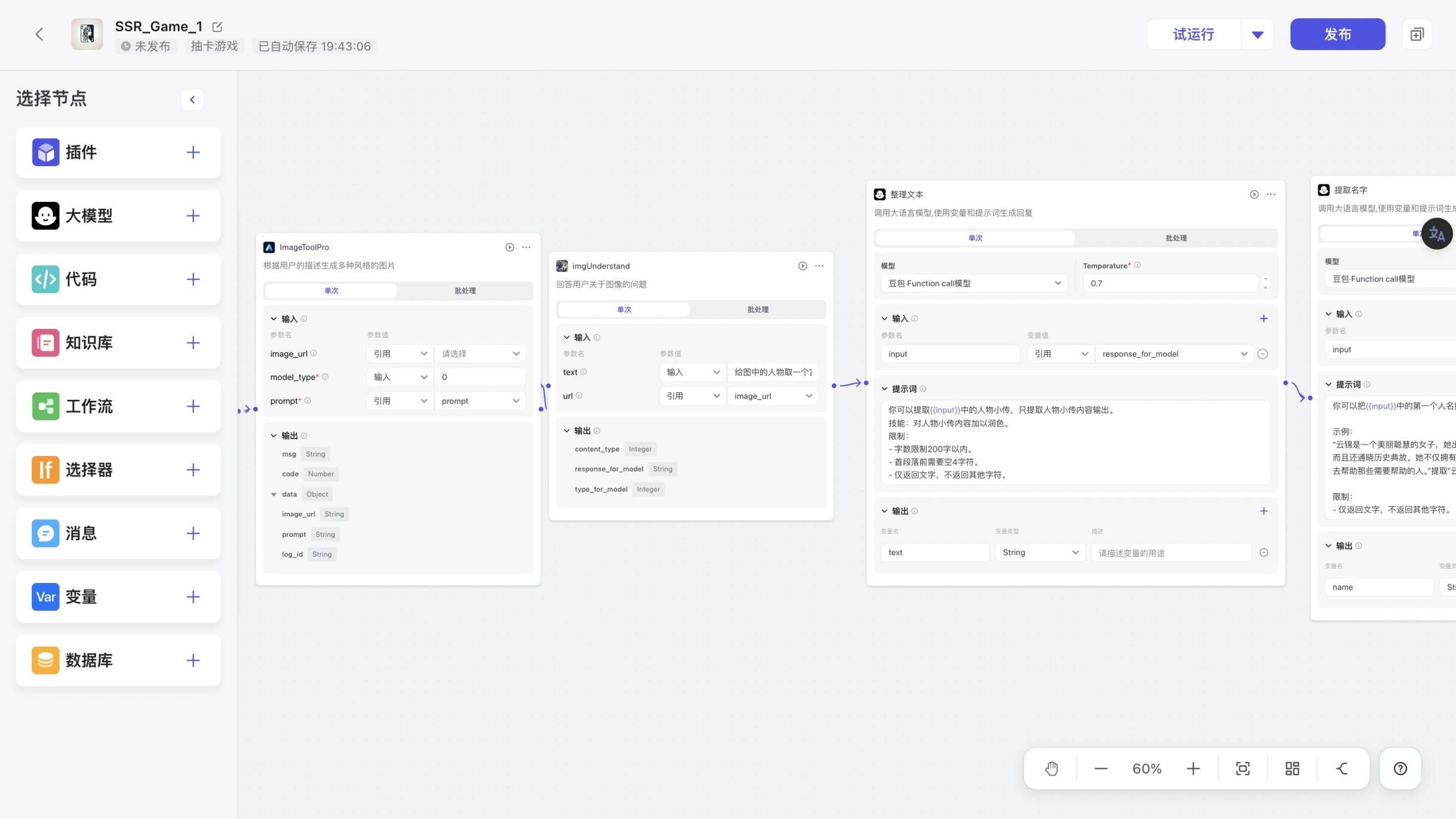
Agentic Workflow 解决什么问题

Agentic Workflow 通过将一个复杂的任务分解成较小的步骤，在整个过程中中融入了更多人类参与到流程中的规划与定义。

它减少了对 *Prompt Engineering* 和模型推理能力的依赖，提高了 *LLM* 应用面向复杂任务的性能，更丰富、更精确。



productivity （效率） = 产生价值的速率


SSR_Game_1
未发布
抽卡游戏
已自动保存 19:43:06
试运行
发布
选择节点
插件
大模型
代码
知识库
工作流
选择器
消息
变量
数据库
ImageToolPro
根据用户的描述生成多种风格的图片
单次
批处理
输入
参数名
参数值
image_url
引用
请选择
model_type*
输入
0
prompt*
引用
prompt
输出
msg String
code Number
data Object
image_url String
prompt String
log_id String
imgUnderstand
回答用户关于图像的问题
单次
批处理
输入
参数名
参数值
text
输入
给图中的人物取一个
url
引用
image_url
输出
content_type Integer
response_for_model String
type_for_model Integer
整理文本
调用大语言模型,使用变量和提示词生成回复
单次
批处理
模型
豆包·Function call模型
Temperature*
0.7
输入
参数名
变量值
input
引用
response_for_model
提示词
你可以提取{{input}}中的人物小传，只提取人物小传内容输出。
技能：对人物小传内容加以润色。
限制：
- 字数限制200字以内。
- 首段落前需要空4字符。
- 仅返回文字，不返回其他字符。
输出
变量名
变量类型
描述
text
String
请描述变量的用途
提取名字
调用大语言模型,使用变量和提示词生成
模型
豆包·Function call模型
输入
参数名
input
提示词
你可以把{{input}}中的第一个人名
示例：
“云锦是一个美丽聪慧的女子，她
而且还通晓历史典故。她不仅拥有
去帮助那些需要帮助的人。”提取“云
限制：
- 仅返回文字，不返回其他字符。
输出
变量名
name
60%

编排

单 Agent 模式 | 豆包·Function call模型

人设与回复逻辑

优化

使用自然语言填写 Bot 的人物设定、功能和工作流程

技能

- 插件
- 工作流
- 触发器

知识

自动调用

- 文本
- 表格

记忆

- 变量
- 数据库
- 长期记忆 关闭

对话体验

- 开场白 Markdown 编辑器
- 用户问题建议 开启
- 快捷指令
- 背景图片

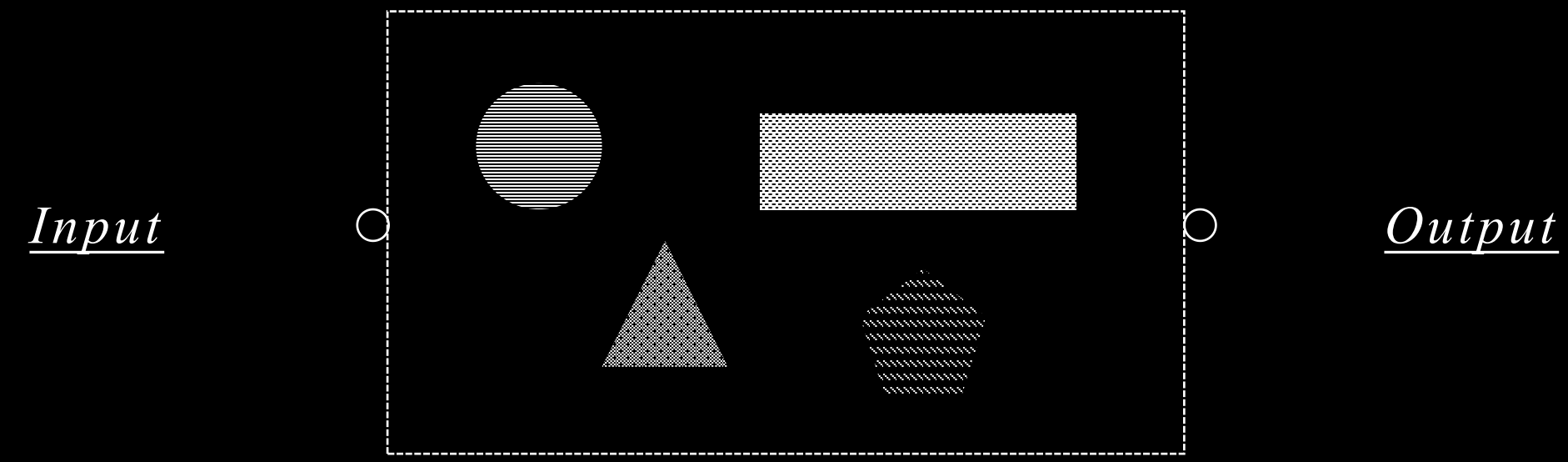
预览与调试

workflow

内容由AI生成，无法确保真实准确，仅供参考。

Agentic Workflow 的“套娃”设计

体验过不同 *Agent* 流程编排开发平台的同学会发现，*workflow* 会成为一个组件被调用，同时 *workflow* 中又能够嵌套新的 *workflow*，实际上不管是基础节点、插件工具、*LLM*、逻辑条件处理等，都实际上是一个以输入、输出的组装的模块，不同的组件之间通过连接构成一个更大的模块。



大模型根源的“不太聪明”
是加上*workflow*也解决不了的

*LangGPT 提示词框架*工作流设计

与传统的 *Prompt* 从输入直接到输出的映射方式相比，*LangGPT* 提示词框架应用了*CoT（Chain of Thought）*完成了从输入到思维链再到输出的映射，即*<input——>reasoning chain——>output>*。

```
## Workflow
1. You will analysis the given the personal information.
2. Create a summary of my diet and exercise plan.
3. Create a detailed workout program for my exercise plan.
4. Create a detailed Meal Plan for my diet.
5. Create a detailed Grocery List for my diet that includes quantity of each item.
6. Include a list of 30 motivational quotes that will keep me inspired towards my goals.
```

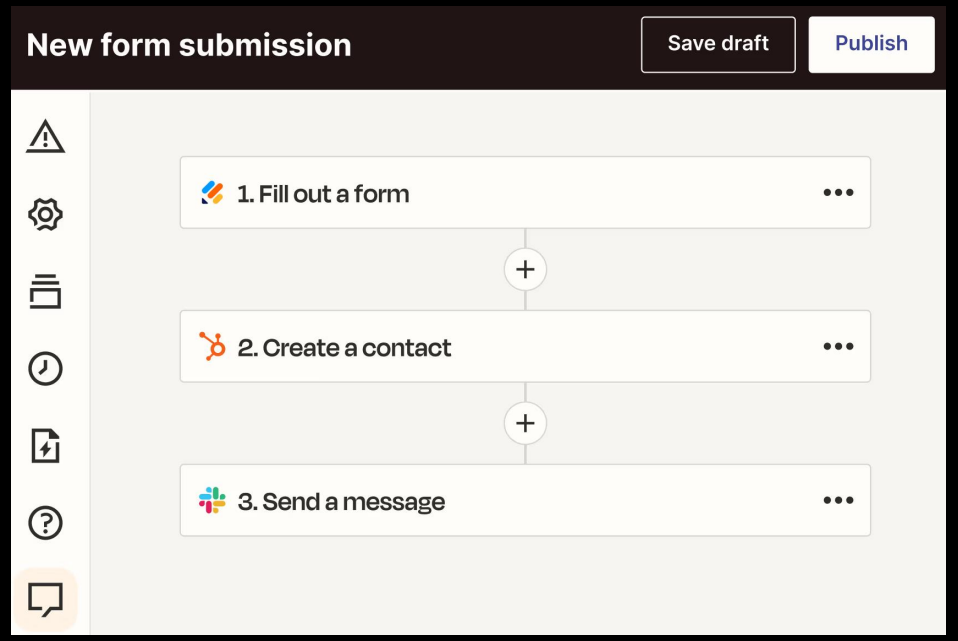
LangGPT 结构化提示词片段

“Let’s think step by step”

RPA 的工作流设计

流程机器人（*RPA*）软件的目标是使符合某些适用性标准的基于桌面的业务流程和工作流程实现自动化，一般来说这些操作在很大程度上是重复的，数量比较多的，并且可以通过严格的规则和结果来定义，现在越来越多的*RPA*软件带上了*LLM*。

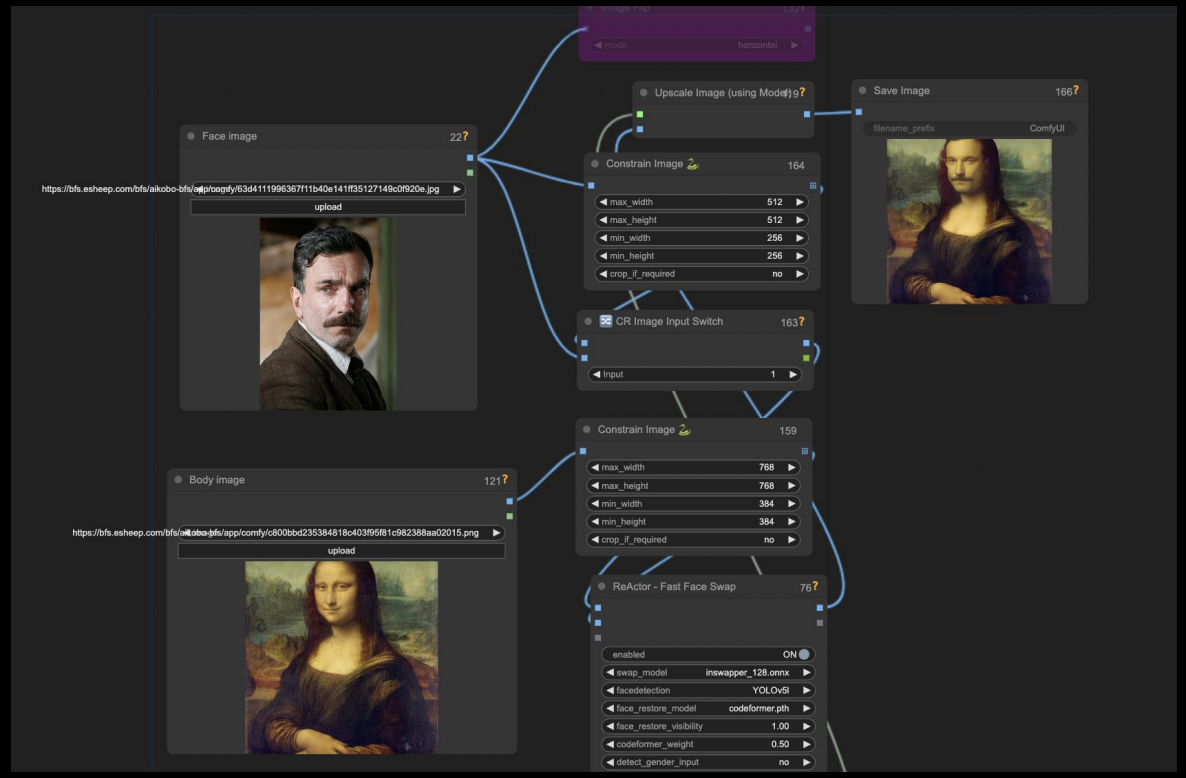
图为：*影刀RPA 工作流*



图为：*Zapier 工作流*

ComfyUI 的工作流设计

ComfyUI 是将开源绘画模型 *Stable Diffusion* 进行工作流化操作模式，用户需要在流程编辑器中配置出每一个的*pipeline*，并通过不同节点和连线来完成模型的操作和图片内容生成，提高了流程的可复用性，同时它的 *DSL* 配置文件还支持导出导入。



ComfyUI 流程编辑器


```
{
  "last_node_id": 104,
  "last_link_id": 158,
  "nodes": [
    {
      "id": 19,
      "type": "Text List to Text",
      "pos": [
        648.4678654583099,
        356.2712236829971
      ],
      "size": {
        "0": 240,
        "1": 60
      },
      "flags": {},
      "order": 19,
      "mode": 0,
      "inputs": [
        {
          "name": "text_list",
          "type": "LIST",
          "link": 24,
          "label": "text_list",
          "slot_index": 0
        }
      ],
      "outputs": [
        {
          "name": "STRING",
          "type": "STRING",
          ...
```


DSL 配置文件

Inhai.wiki

Dify.AI 可被复制的工作流设计

在 *Dify.AI* 中，我很兴奋的看到它的工作流设计语言跟 *ComfyUI* 会有一些相似之处，都是定义了一套*DSL*语言，并且非常方便的可以使用导入导出的功能进行工作流的复用。

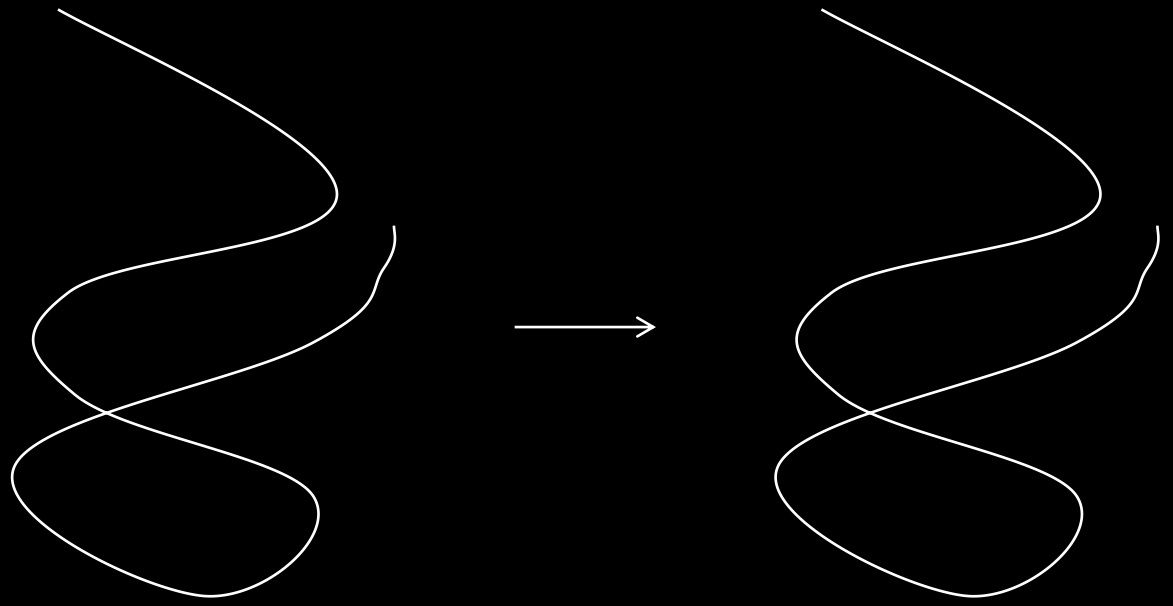
Dify.AI 导出 *DSL* 文件

```
app:
  description: ''
  icon: "\U0001F916"
  icon_background: '#FFEAD5'
  mode: workflow
  name: Agentic Workflow
workflow:
  features:
    file_upload:
      image:
        enabled: true
        number_limits: 3
        transfer_methods:
        - local_file
        - remote_url
    opening_statement: ''
    retriever_resource:
      enabled: false
    sensitive_word_avoidance:
      enabled: false
    speech_to_text:
      enabled: false
    suggested_questions: []
    suggested_questions_after_answer:
      enabled: false
    text_to_speech:
      enabled: false
```

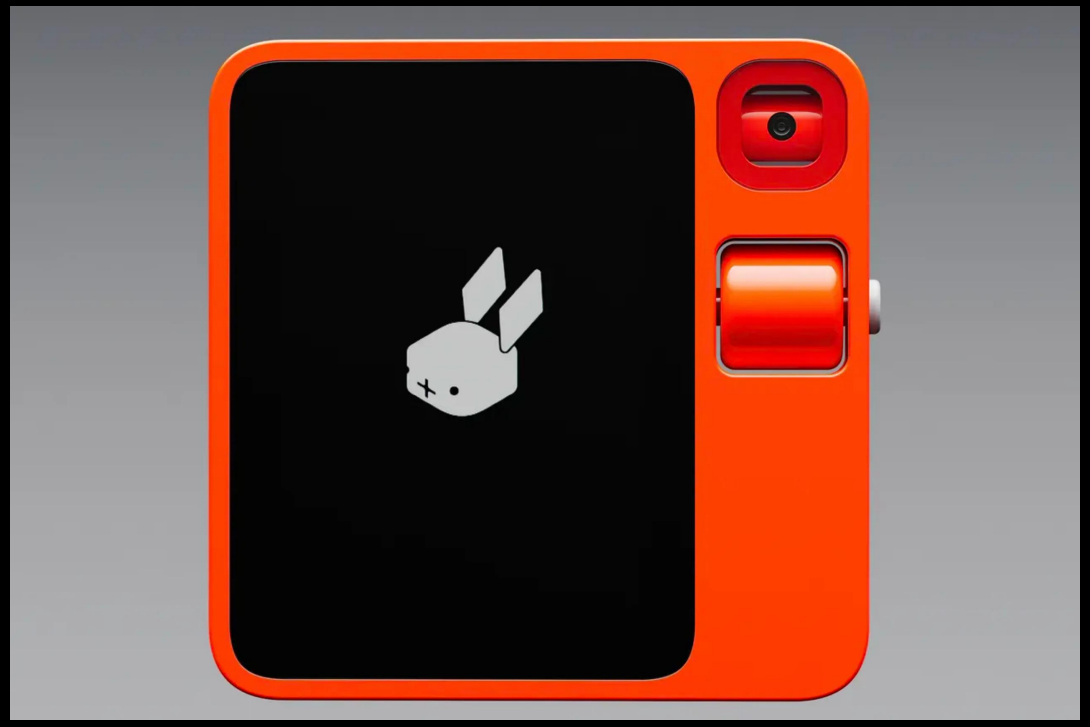
Dify.AI 通过 *DSL* 文件创建工作流

模仿式工作流是最快的学习方法

Large Action Model 采用称为“通过演示进行模仿”的技术。检查人们在单击按钮或输入数据时如何与界面互动，然后准确地模仿这些操作，他们收集知识并从用户提供的示例中学习，使他们更能适应进一步的变化并能够处理不同的任务。



学习复杂行为路径，完成模仿后复现



图为：*Rabbit R1*

别神话了 *AI Agent*
*Agentic Workflow*可能是“伪需求”

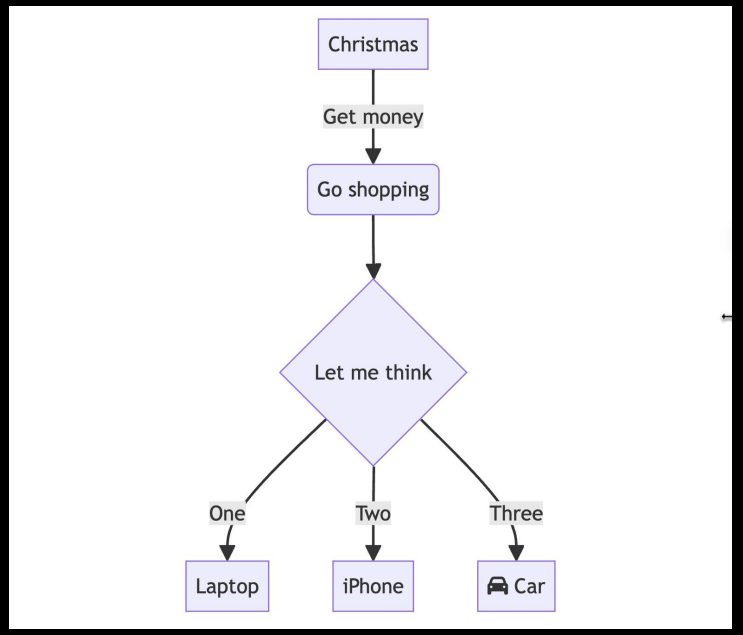
Idea Time：通过自然语言创建工作流

复杂的工作流搭建怎么会如此麻烦...这似乎跟我我理想中的*Agentic Workflow*并不太一样！有没有一种更加方便高效的方式，让我能够在短时间内创作一个符合我预期的*Agentic Workflow*原型？有了，通过自然语言来构建*DSL*并还原工作流。

*Mermaid*语法构建流程图

“生成一个购物的流程图”

```
flowchart TD
    A[Christmas] -->|Get money| B(Go shopping)
    B --> C{Let me think}
    C -->|One| D[Laptop]
    C -->|Two| E[iPhone]
    C -->|Three| F[fa:fa-car Car]
```



一个小思考题
*Agentic Workflow*该给谁用？

Pailido | AI 拍立得

这是一款文案快速生成的 *AI-Naitive* 产品，各个场景由 *AI Agent* 驱动，仅需选中场景后点击拍摄即可快速生成对应文案。
使用多模态模型，理解图片特征和输出场景期待，搞定小红书文案、外卖点评写作、闲鱼商品发布文案...真的太快了!

Img → *AI Agent* → AA

扫码体验

Inhai.wiki

Reshape your workflow with AI ?

or

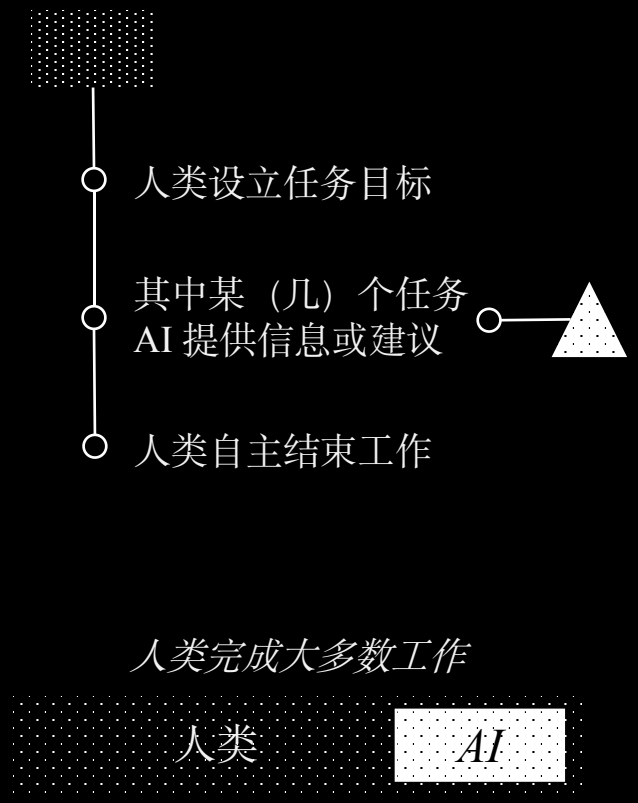
Reshape your AI workflow ?

Inhai.wiki

AI 与人的协同关系

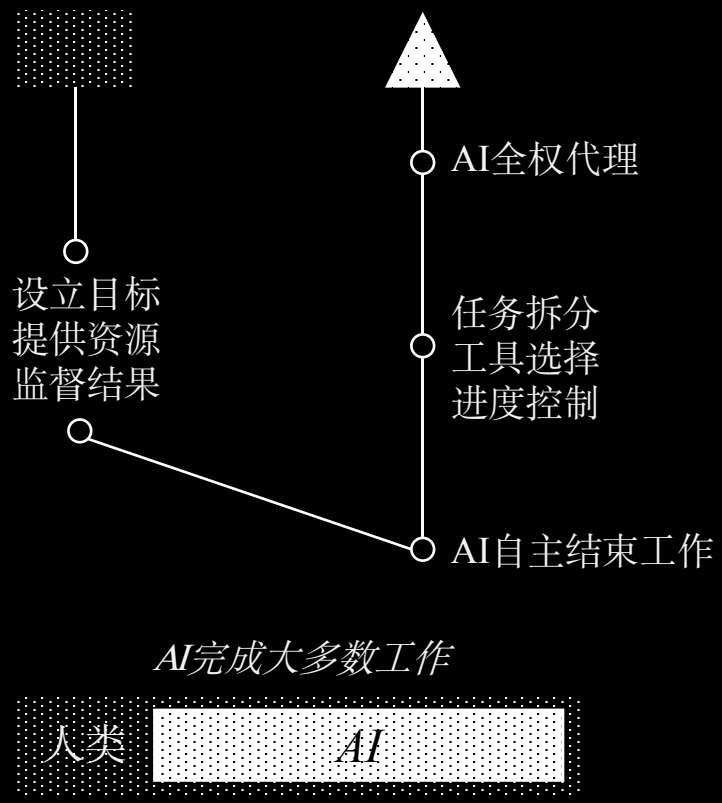
生成式 AI 的人机协同分为 3 种产品设计模式：Embedding（嵌入式）、Copilot（副驾驶）、Agent（智能代理）

在这 3 种模式下，人与 AI 的协作流程也是有所差异。



Embedding 模式

Copilot 模式

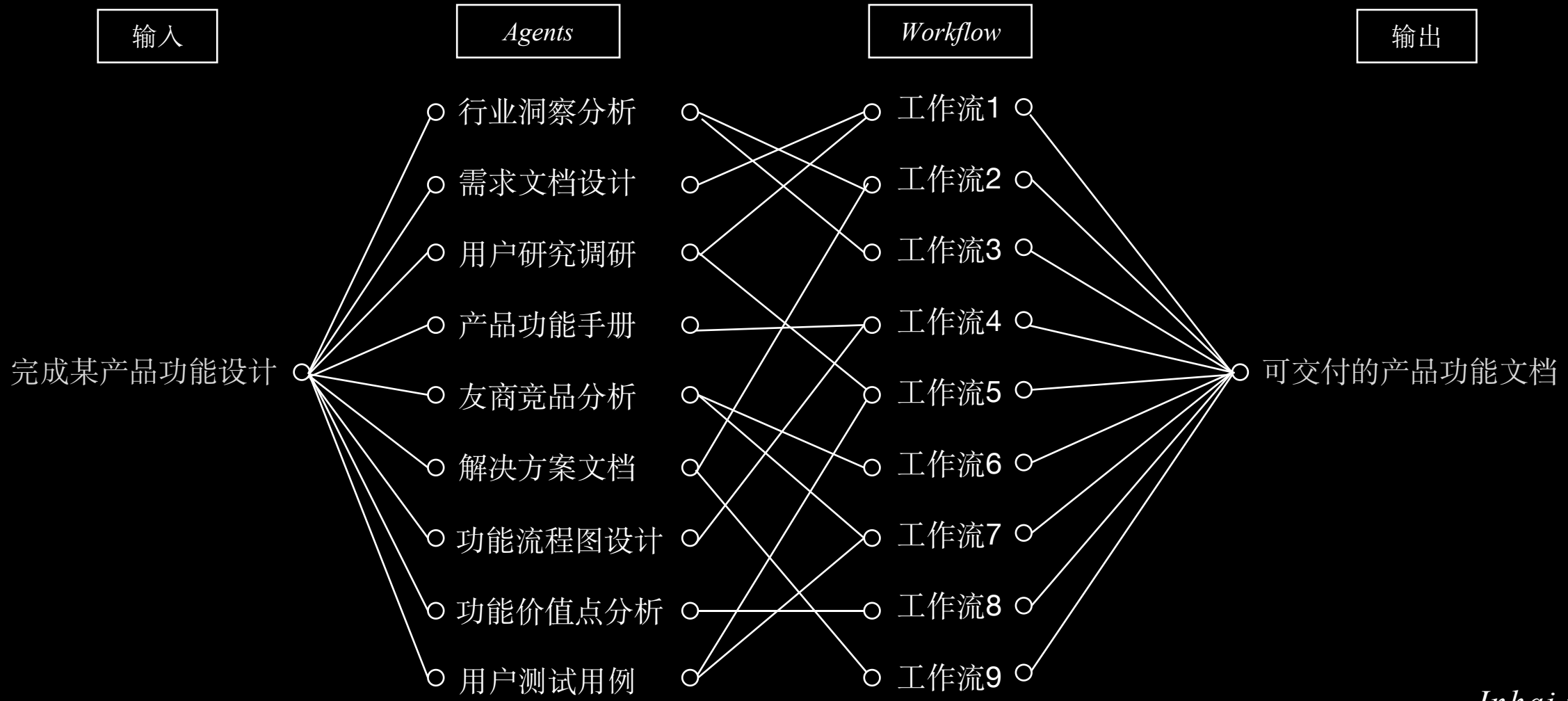


Agents 模式

Inhai.wiki

*Agentic Workflow*驱动角色工作流变革

使用*Multiagent Collaboration*的方法，让不同角色的 *Agent* 按照根据任务要求自主规划选择工具、流程进行协作完成一件任务。



规划

道德

逻辑

文化

格式化

识别

分类

反馈

记忆

交互

解释

输入

翻译

预测

检索

适应

推理

多任务

元认知

分析

输出

情感

优化

可视化

总结

评估

生成

隐私

处理

提取

信息源

抽象

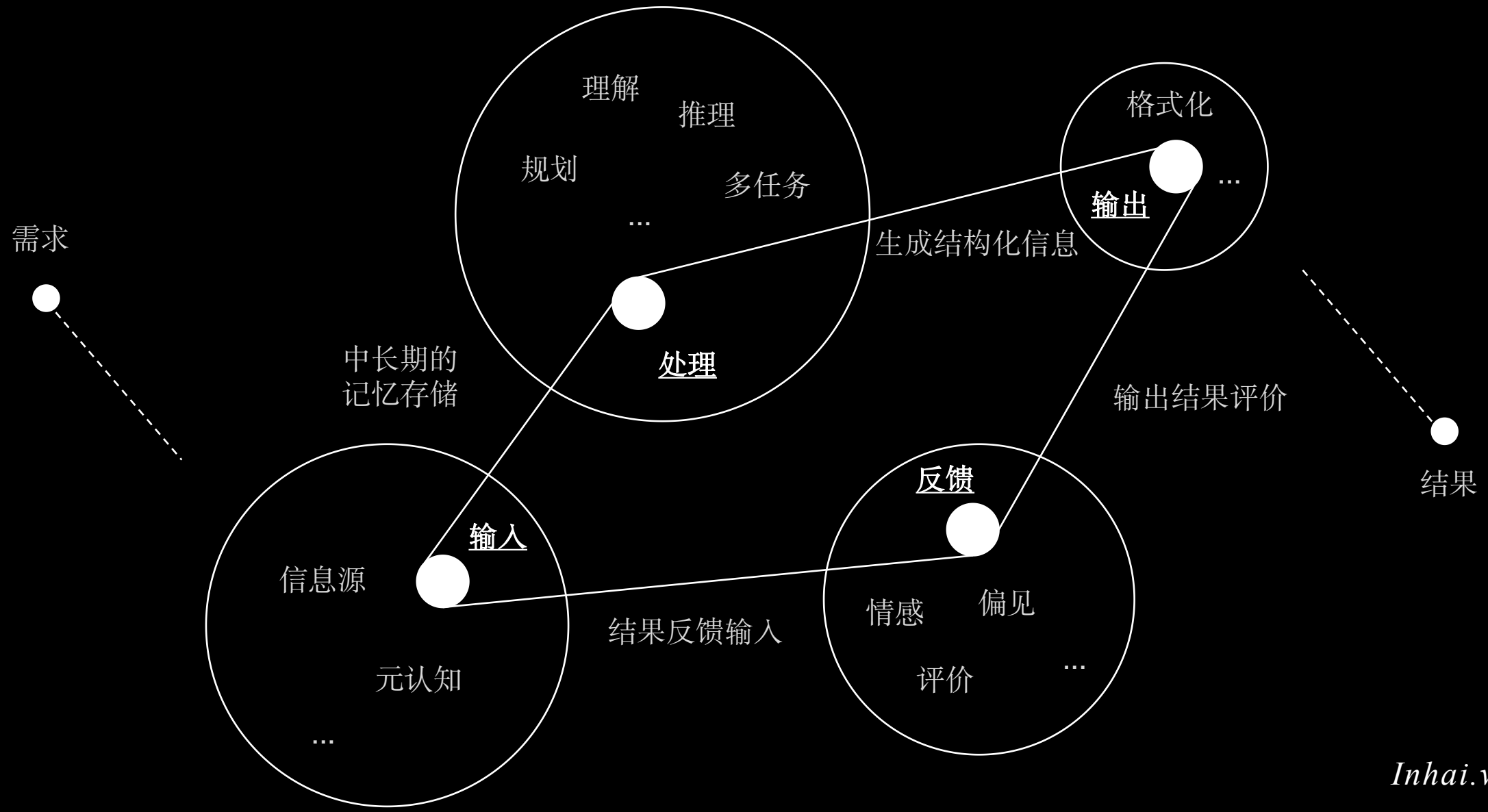
改写

决策

模拟

偏见

从原子能力到协同工作流



Inhai.wiki

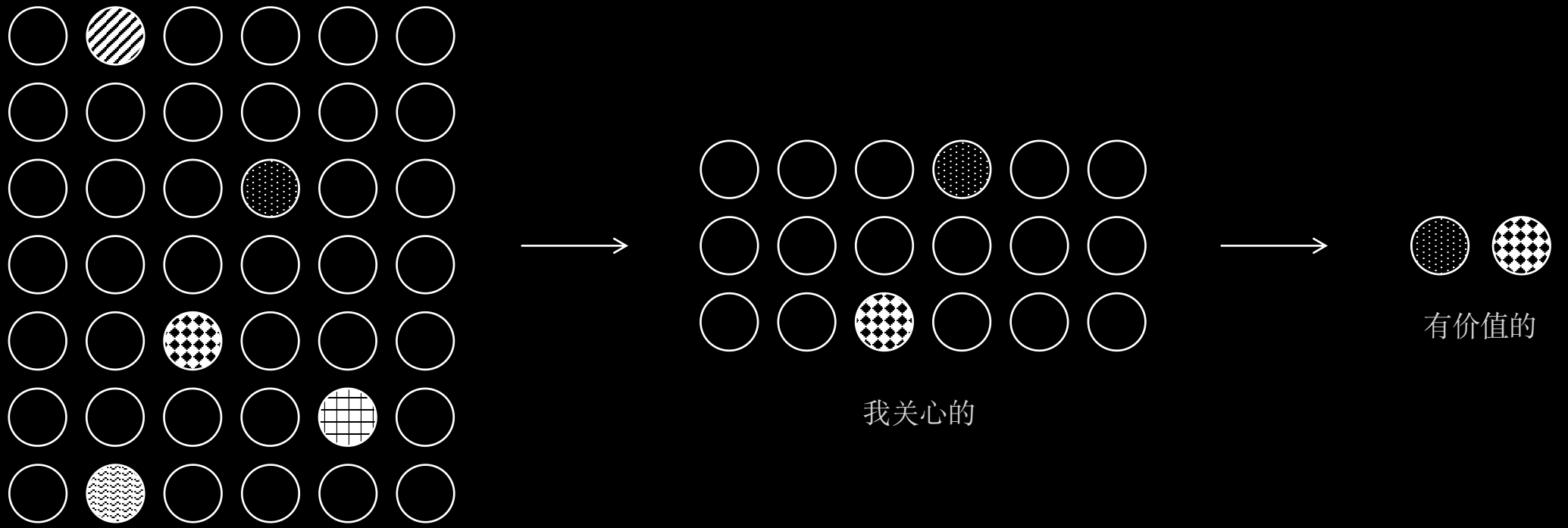
重塑获取信息的方式

搜索引擎作为互联网基础设施，同时也是互联网的入口，对于用户而言，从解决问题出发，搜索引擎和基于大模型的聊天机器人的目标从根本上是一致的。自 2022 年底 ChatGPT 发布，其通过问答形式被认为将对传统搜索引擎带来颠覆。

辅助高效的处理信息

阅读完一份 *10* 万字的 *PDF* 研究报告需要多久？这份报告主要讲了什么内容？有没有我要关注的点？

智能摘要是目前我用的超顺手的一个功能，能够辅助我快速的筛选信息，什么值得看，什么不容错过，实现信息的降噪。



信息表达更简便

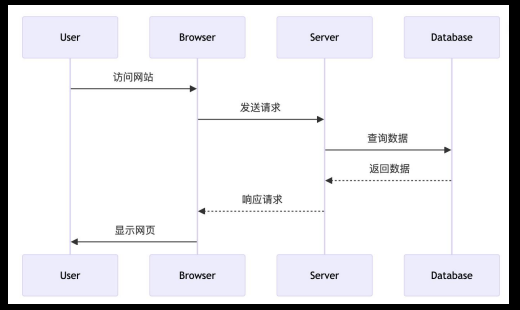
放在以往很难想象，如果要实现下面这俩张图，可能会设计一系列的思考、草稿、理清逻辑等等流程。

现在用自然语言描述一句话就给你生成了这样美观可用的图片，极大的降低了不同角色的创作门槛和周期，是真的简便。

AA ⟶ *Img*

“画一只闪电皮卡丘” ⟶

“画一张用户访问网站的流程图” ⟶



新的产品设计方式

使用 *AI* 帮我完成需求文档

使用 *AI* 进行产品需求分析

使用 *AI* 设计访谈提纲

使用 *AI* 绘制用户旅程图

使用 *AI* 进行竞品调研

使用 *AI* 进行搞定用户画像

使用 *AI* 完成指导我设计用户研究问卷

使用 *AI* 进行一场头脑风暴

AI 设计产品测试用例

使用 *AI* 设计 *PPT* 大纲

使用 *AI* 绘制产品功能流程图

设计产品图标

使用 *AI* 设计产品封面图

Inhai.wiki

我的工作流上使用的产品

输入

即刻	推特	小红书	微信公众号	*NetNewsWire Rss*
Medium	*Chrome*	飞书知识库	*pinterest*	知识星球

处理

ChatGPT	*Coze*	*Cubox*	*Obsidian*	*Readwise*
Kimi	通义系列	*Upscayl*	*Upscayl*	

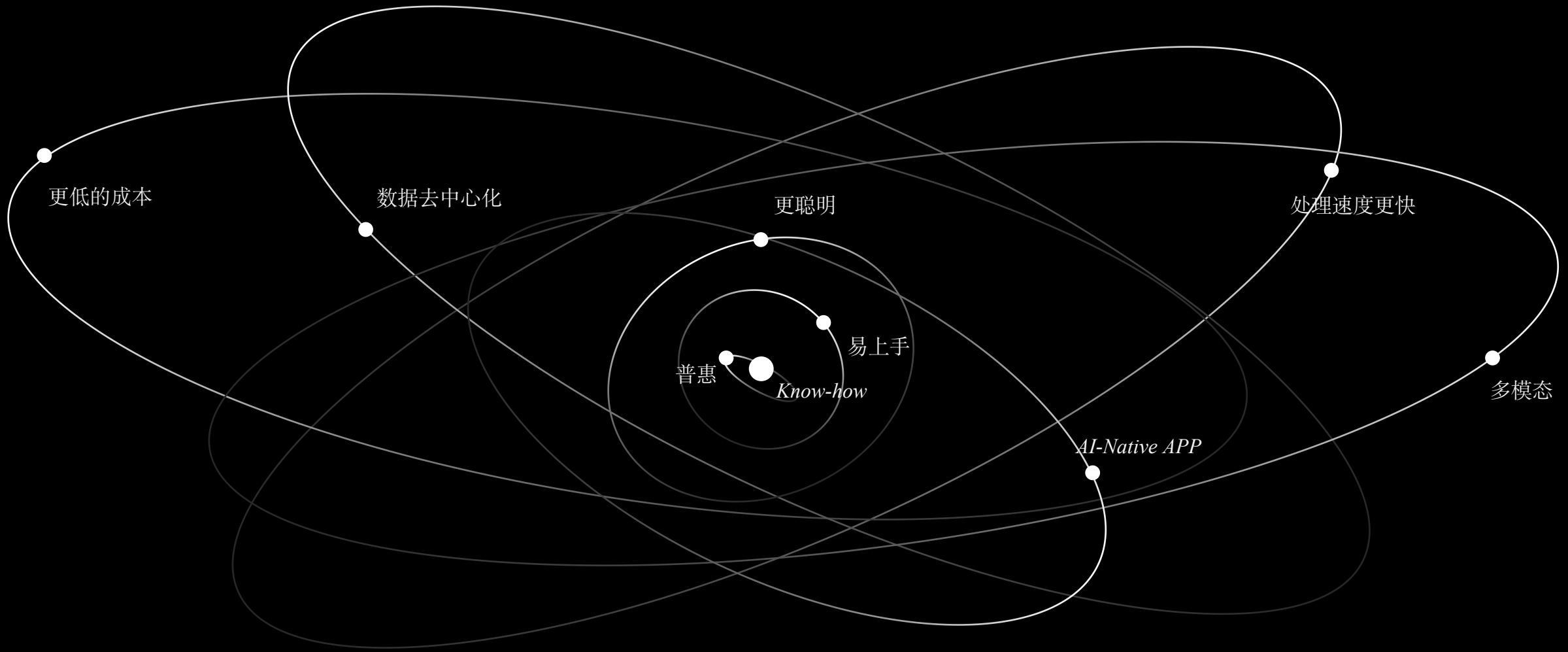
输出

美图设计室	即梦 *Dreamina*	*Figma*	*Anaconda 3*	*excalidraw*
ComfyUI	*esheep*	*Typora*		

Inhai.wiki

曾被认为的异想天开的想法
都可能会是 *AI Agent* 的未来

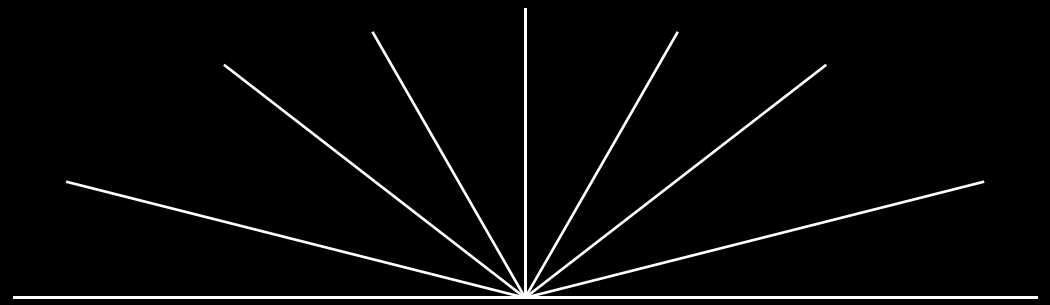
Agents Universe



Inhai.wiki

```
def user_feedback():
    if like_this_share or contact_with_me:
        do_action('点赞')
        do_action('分享')
        add_wechat_friend('168007300')
```

Inhai.wiki



Thanks

Inhai.wiki