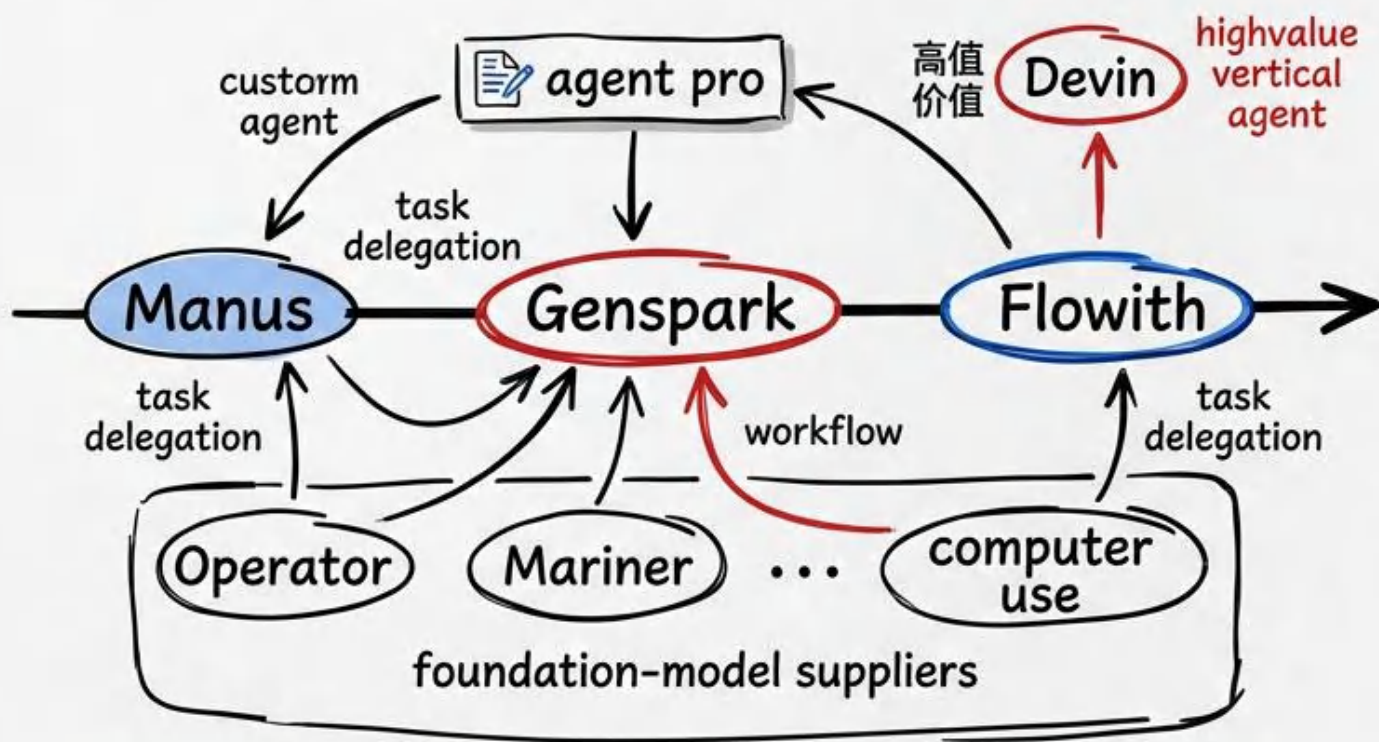


全球通用智能体竞争报告

产品层重构版：真正面向任务交付的通用 agent 竞争

- 这份报告不再以基座模型为中心，而以**真正的通用 agent 产品层**为中心。
- Manus、Genspark、Flowwith 是主轴；Operator、Mariner、computer use 是底座能力层；Devin 是高价值垂直代理。
- 核心问题不是谁更会答题 II，而是谁更像一个能接活、能交付、能持续协作的数字同事。



@ 清新研究团队简介

- 领导学术研究团队近30人。指导大数据、AI、人形机器人等多个产业团队。
- 团队坚持：整体主义、实证主义、社会建构、进步主义。

沈阳：清华大学新闻学院/人工智能学院双聘教授、博导



六大研究方向：

视频号：@清新研究；公众号：@清新研究



1. AI大模型理论与哲学



2. AI文艺



3. AI应用



4. 新媒体与网络舆论



5. 大数据

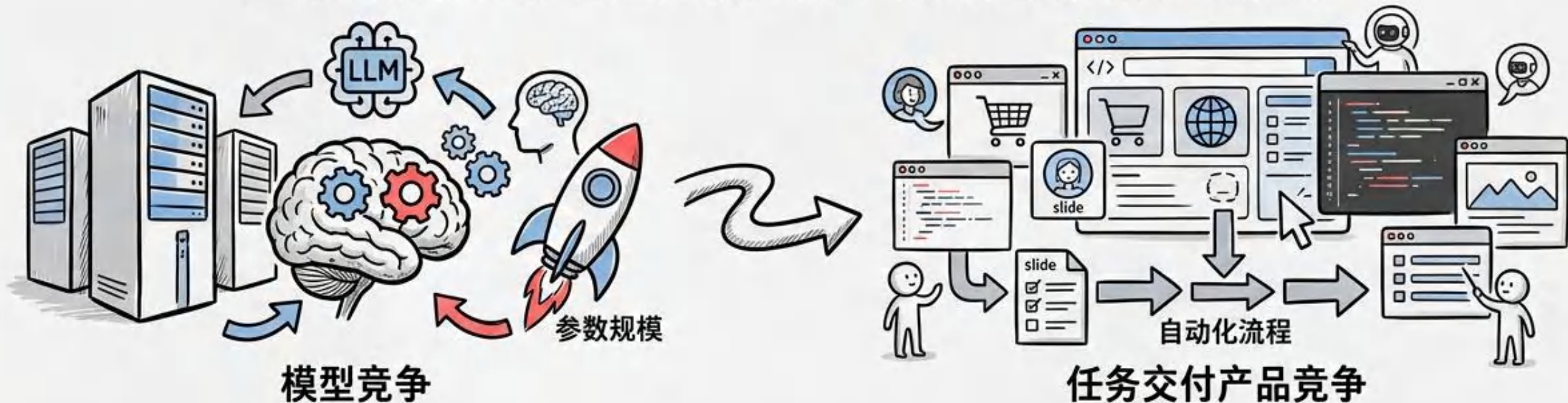


6. XR应用

✉ 邮箱：124739259@qq.com | 🐦 微博：@清新研究 | 公众号：@清新研究

执行摘要

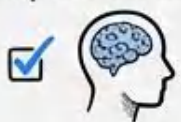
通用智能体竞争已经从模型竞争转向任务交付产品竞争



Manus、Genspark、Flowith 更接近用户心智中的“真正通用智能体”。

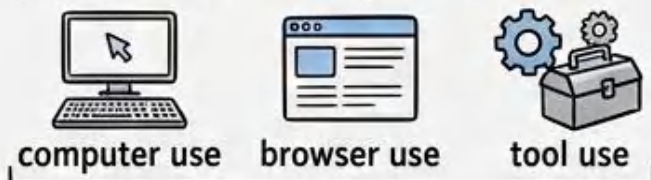


Oocopic



真正的通用

OpenAI、Google、Anthropic 更像是在提供 computer use、browser use 和 tool use 的原语。



原语

未来护城河会更多体现在任务交付、工作台入口、记忆与治理，而不是参数规模。



参数规模

目录

一、什么才算真正的通用智能体



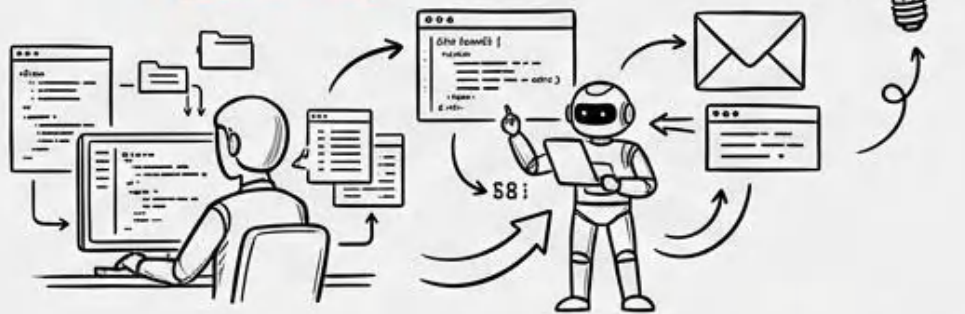
二、产品层主战场： Manus / Genspark / Flowith



三、底座能力层： OpenAI / Google / Anthropic



四、垂直高价值代理：Devin



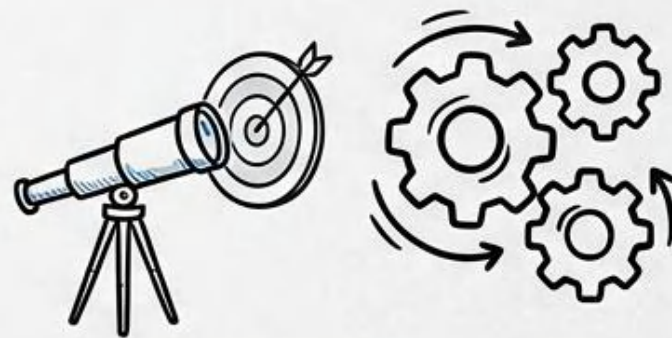
五、竞争维度、原创概念与未来判断





第一部分：重新定义研究对象

先把基座模型、agent 原语和真正的通用 agent 产品层区分开



@ 清新研究团队 | 2026年3月

什么才算真正的通用智能体

接任务、拆任务、执行任务、交付结果



为什么不能再把“通用智能体报告”写成“基座模型报告”

★ 模型层重要，但不是最终竞争单位

★ 基座模型 (Foundation Models)

基座模型层 (能力上限)

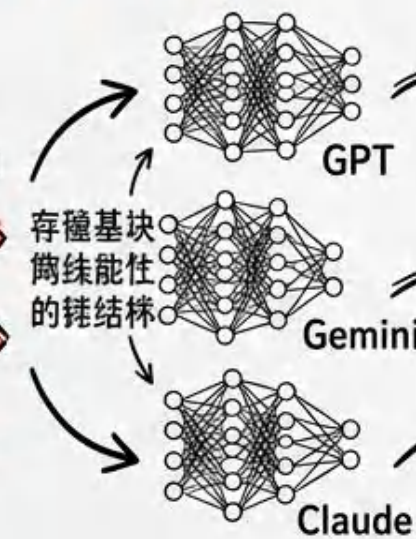
- 决定能力上限
- 强大的基座，但不是产品

只是工具

不自用绞竞果作息

不能直接解决用户问题

- 不自动等于用户实际使用的通用智能体产品



利用原座模型制作
为最终竞争重作息

通用智能体产品 (Agents Products) !

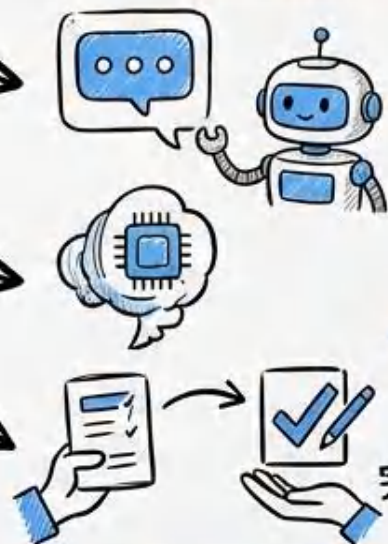
产品层 (用户实际使用)

- 真正的用户竞争单位
- 把任务接过去并做完

因此报告主轴必须
转向产品层

(立外一任求例)

垂直领域智能体
(Vertical Agents)
对季应所用例

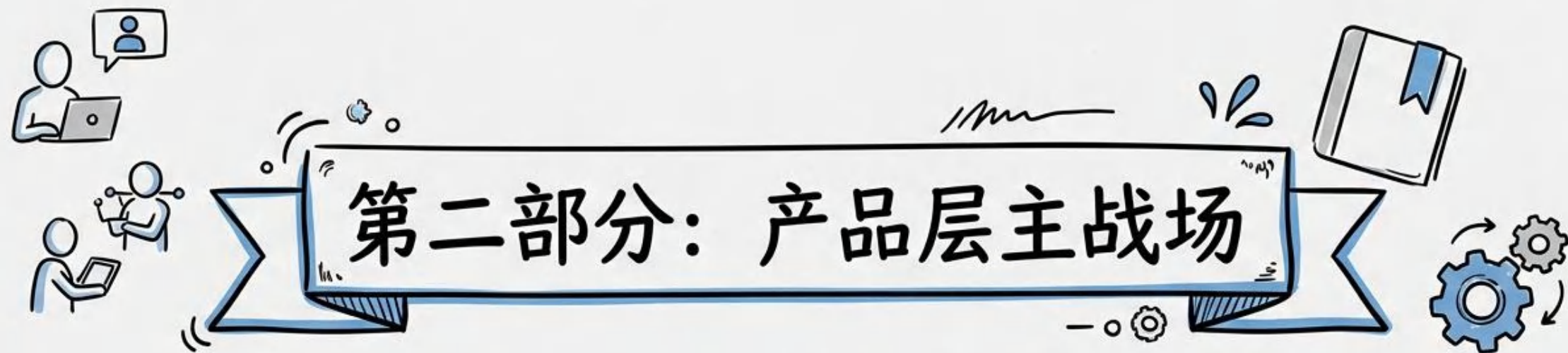


广义智能体报告
(Agent Reports) VS

新的三层框架

产品层、原语层、垂直代理层





第二部分：产品层主战场

➔ 真正的通用 agent 产品竞争，应该从这里开始写

— 这一部分把 **Manus**、**Genspark**、**Flowith** 放到报告中心。

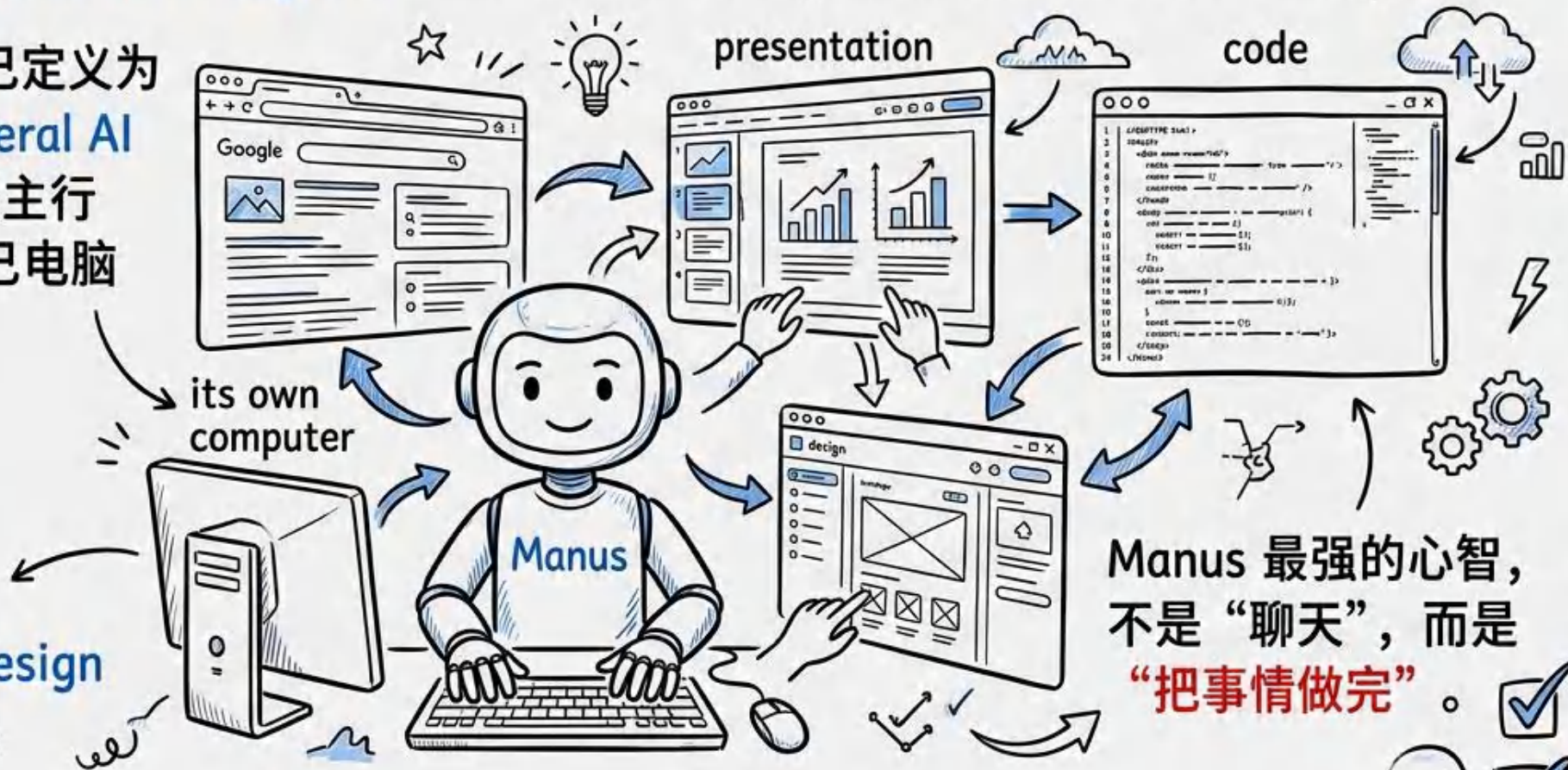


Manus: 最典型的任务交付型通用 agent

★ autonomous general AI agent + virtual colleague with its own computer

★ Manus 官方把自己定义为 autonomous general AI agent，强调像是主行强调像一个有自己电脑的虚拟同事。

★ 它的产品重心是 create slides、build websites、develop apps、design 这类可交付任务。



为什么 Manus 必用能体 必须在报告中心

它代表了通用智能体从“回答器”转向“交付器”

【传统/有限智能体】



- 单纯 Browser Agent (浏览器代理)
- 只能访问和提取网页信息
- 交付的是“信息”而非“成果”
- 局限于浏览器, 无法自主生成文档

✗ “回答器”

Compared: Broad Agents vs. Vertical Agents

Broad Agents → Vertical Agents

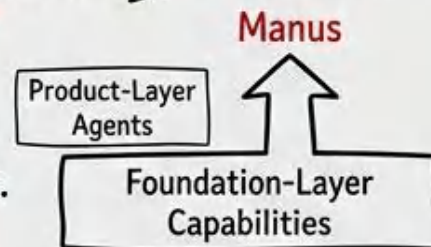
【Manus 通用智能体】



- 争夺默认任务承接方位置
- 这使它比单纯 browser agent 更接近用户对“通用智能体”的直觉
- 交付的是“成果”而非“信息”
- 拥有 Foundation-Layer 能力

✓ “交付器”

Compared: Product-Layer Agents vs. Foundation-Layer Capabilities



Manus 的关键差异点

本地浏览器执行 + 文件交付 + 任务端到端完成

— Browser Operator 让 Manus 直接在用户本地浏览器里使用现有登录态和活动标签页。

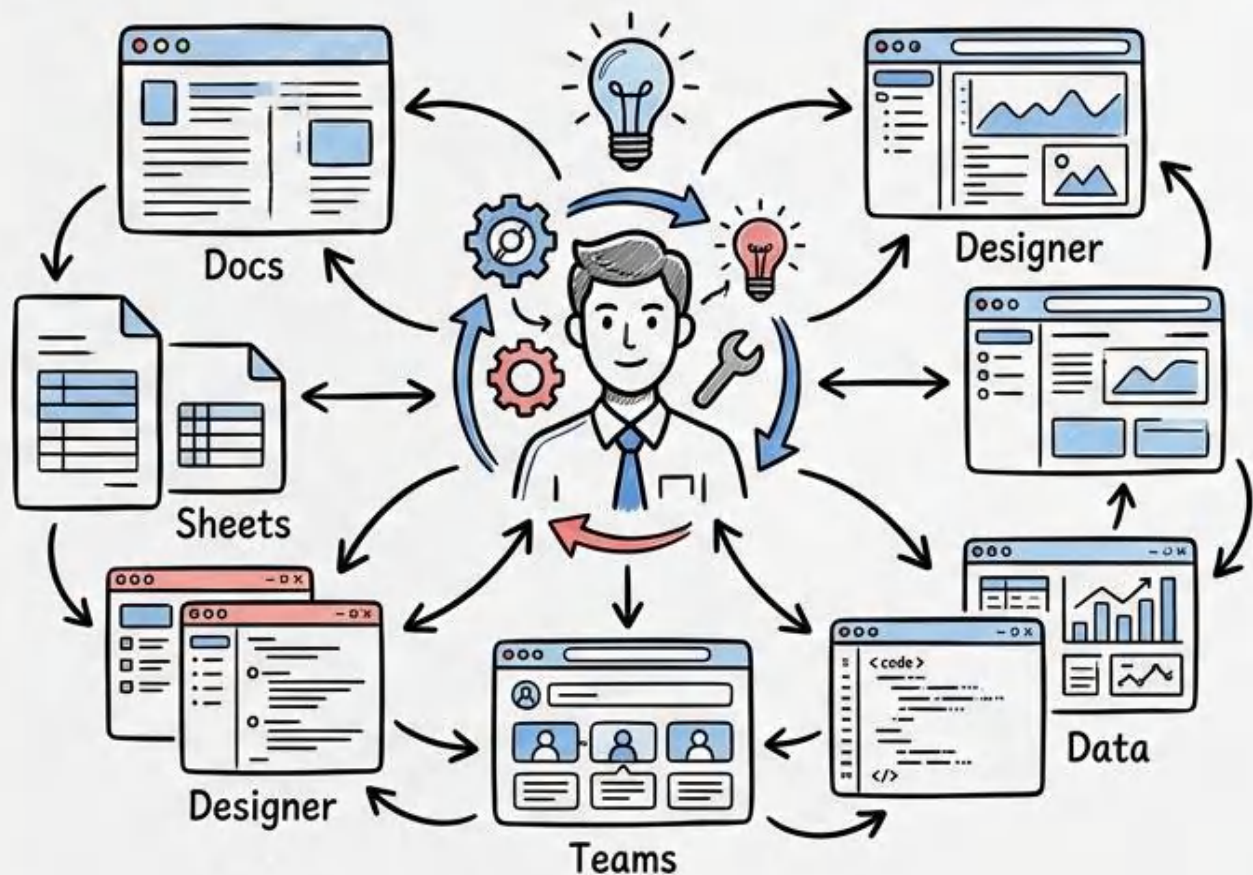
— 同时它能生成 PPT、网站等真实文件成果。



— 这让它比单纯在云端点网页的agent 更接近生产环境。

Genspark: 把通用智能体做成 all-in-one AI workspace

Super Agent + Docs/Sheets/Slides/Designer/Teams



- Genspark 首页把自己定义为 your **all-in-one AI workspace**.
- Help Center 把 Super Agent 定义为 **thinks, plans, and acts** to complete your everyday tasks.
- 它更像一个 **agent-工作套件**, 而不只是单个代理。



Genspark 与 Manus 的路线差异

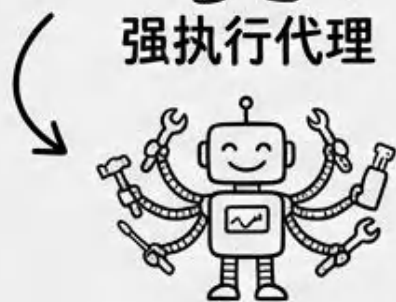
强执行代理 vs agent-native productivity suite

Manus

- Manus 更像“给我任务，我去完成”的强执行代理。
- 争夺交付心智

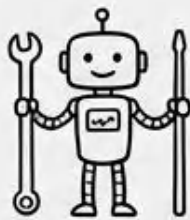


强执行代理



Broad Agent

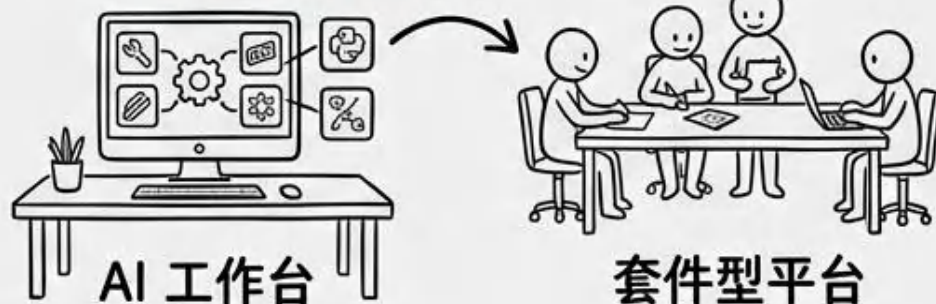
VS



Vertical Agent

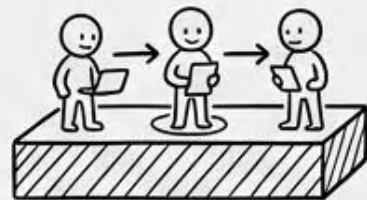
Genspark

- Genspark 更像“把常见工作类型都放进一个 AI 工作台”的套件型平台。
- 争夺工作台入口



Product-layer Agents

VS

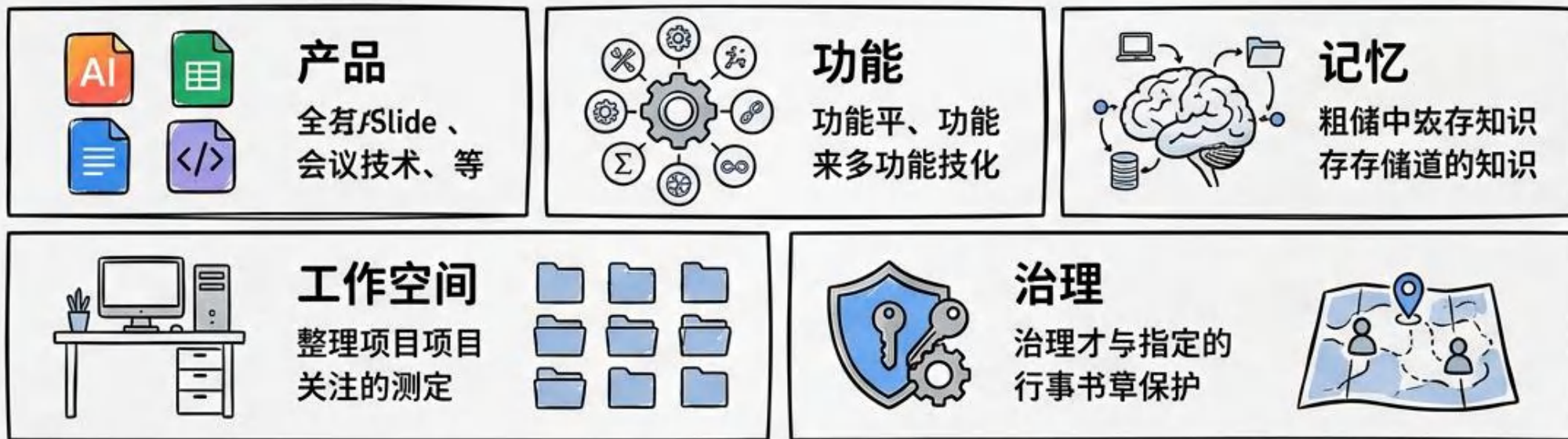


Foundation-layer Capabilities

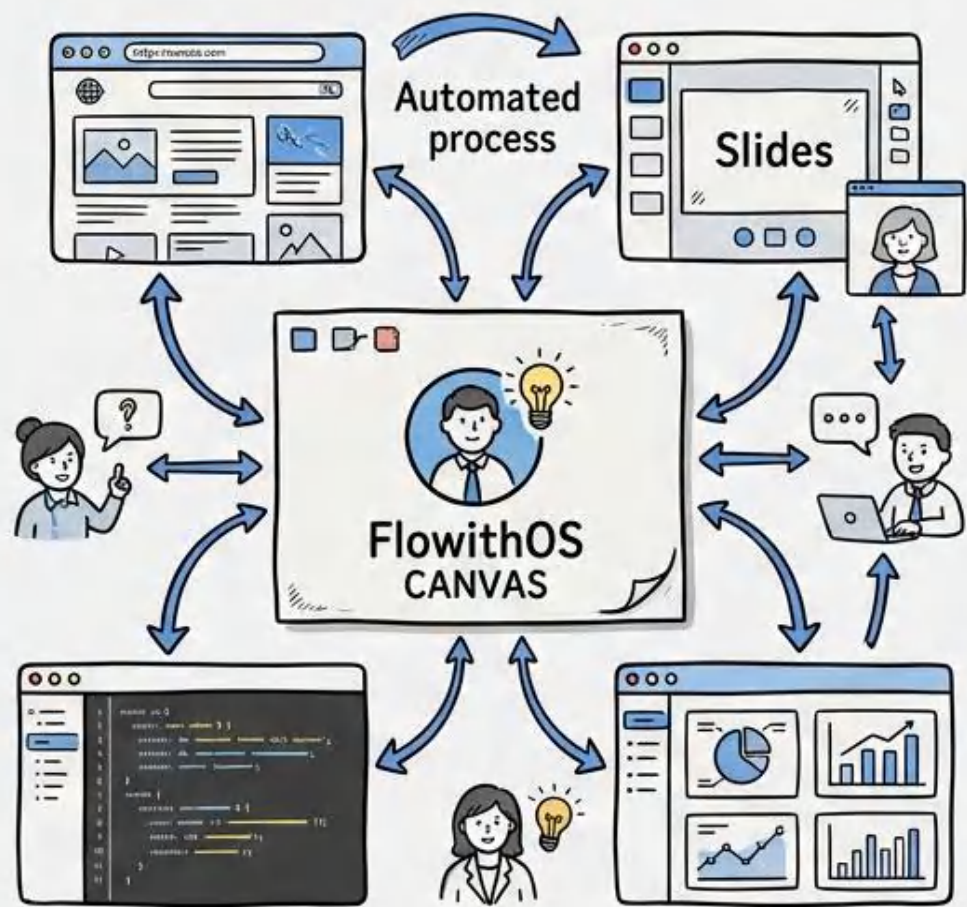
Genspark 的平台化价值

为什么它不只是一个 Super Agent

- 它把 AI Slides、AI Sheets、AI Docs、AI Designer、AI Developer、Teams 等模块拼成统一工作台。
- 这种结构天然更容易形成迁移成本和团队入口。
- 一旦用户把文档、表格、会议和 agent 执行都迁进来，平台黏性会迅速增加。



Flowith: 把通用智能体做成 canvas-first agent workspace



Neo / Oracle / FlowithOS

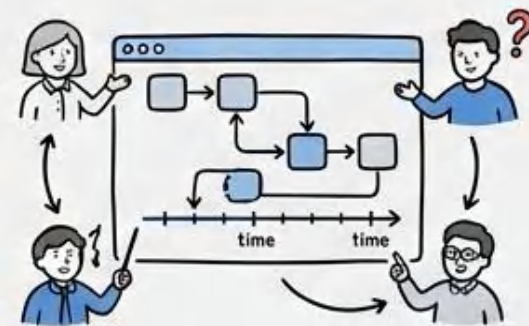
- Flowith 官方把产品定义为 agentic AI workspace, 把知识、创作和执行连接在单一 flow 中。



- Oracle 和 Neo 强调 autonomous planning、task breakdown、tool invocation 和 self-optimization。



- 它代表的是“长期上下文 + 可视化 workflow + agent 协作空间”路线。

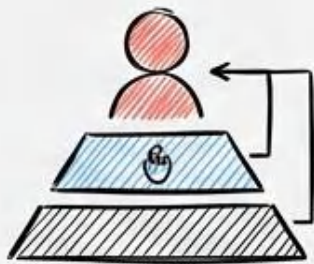


Flowith 与 Manus、Genspark 的差异

● 不是只做执行，也不是只做套件，而是做上下文空间

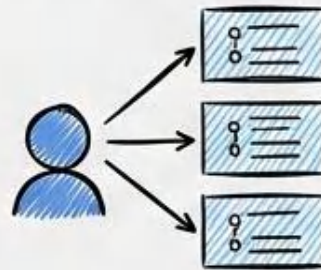
- Flowith 的核心不是单次任务交付，而是长期项目与复杂研究过程中的连续协作。
- 它通过 Canvas、Recipe、Nodes 和 Knowledge Garden 把 agent 行为显式化。
- 这使它在深度研究、长期知识工作和复杂项目场景中更有辨识度。

产品层 Agent vs 基础层能力



产品层 Agent vs 基础层能力

泛型 Agent vs 垂直 Agent

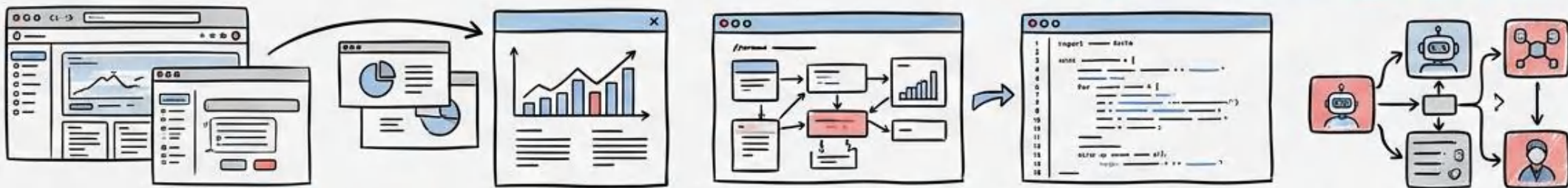


泛型 Agent vs 垂直 Agent

FlowithOS 的意义

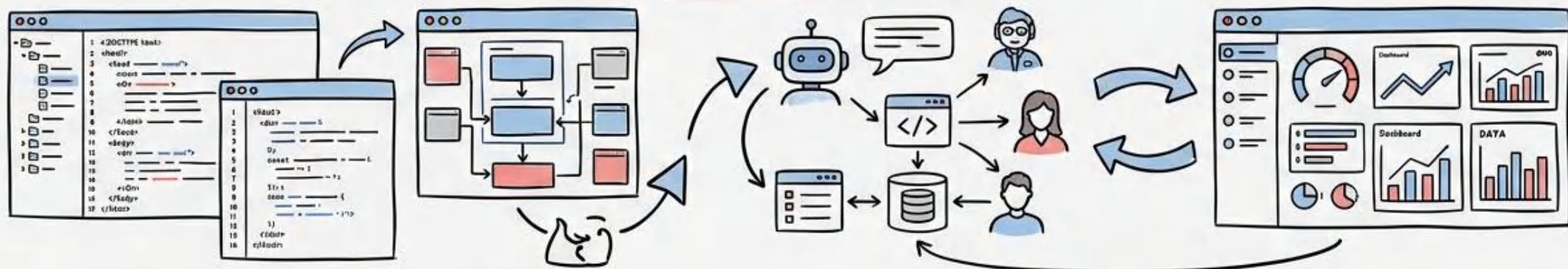
它在争夺 “AI Agent Operating System” 心智

— FlowithOS 官方直接把产品定义为 next-generation **AI Agent Operating System**。



— 关键词是 **self-improvement**、**memory**、**speed**，而不是单纯聊天或搜索。

— 这说明 Flowith 的野心是成为**默认 agent 环境**，而不是单个工具。



产品层三强对比

Manus / Genspark / Flowith 三条不同路线

	产品核心	关键功能	工作模式
Manus	 最像数字同事	 最强调任务交付与结果生成	 自动执行流程
Genspark	 最像 agent-native 工作套件	 最强调统一入口和产品模块	 集成多重 Agent
Flowith	 最像长期项目工作台	 最强调上下文、记忆和可视化任务流	 人机协同项目管理



第三部分：底座能力层

PART 3: FOUNDATION CAPABILITY LAYER

真正的通用 agent 产品，背后都依赖这些动作原语



核心原语



感知



理解



操作



交互

这一部分解释为什么 OpenAI、Google、Anthropic 很重要但含重要，不应替代产品层。



OpenAI: 从 Operator 到 ChatGPT agent

computer use + integrated task execution

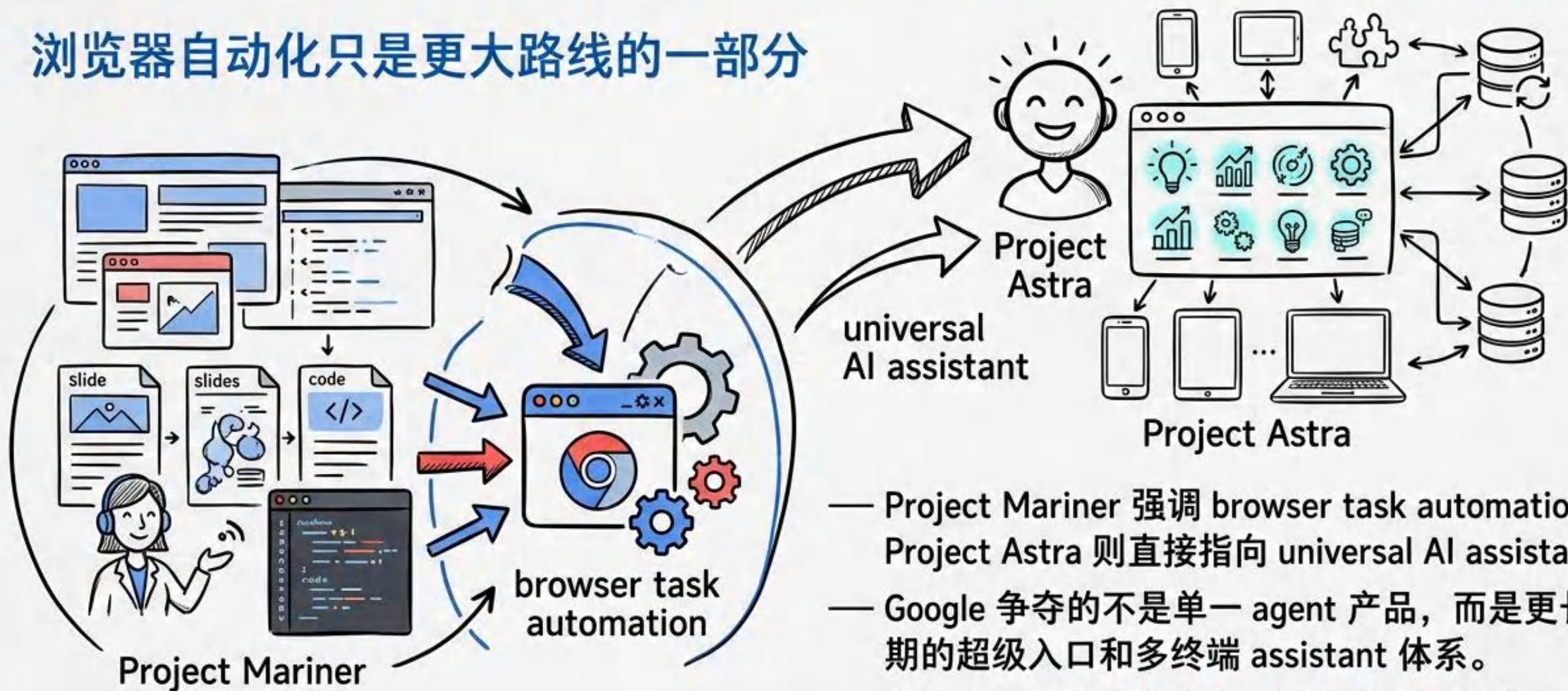
— Operator 是能用自己的浏览器替用户执行任务的研究预览。

— ChatGPT agent 则进一步把研究、表格、预订和 slides 等任务整合到统一产品里。



Google: Project Mariner 与 universal assistant 路线

浏览器自动化只是更大路线的一部分



- Project Mariner 强调 browser task automation, Project Astra 则直接指向 universal AI assistant。
- Google 争夺的不是单一 agent 产品, 而是更长期的超级入口和多终端 assistant 体系。
- 这使它更像“**平台底座 + 默认入口**”的选手。

Anthropic: computer use 作为 agent 动作层能力



重点在工具和 GUI 操作，而不是消费级 agent 品牌

- Anthropic 的 computer use 让开发者可以在沙箱计算环境里操作 GUI。
- 它的重要性在于增强 agent 的 执行能力，而不是直接争夺消费级通用 agent 入口。
- 这是一种更偏底层动作原语的角色。

为什么原语层重要，但不等于最终产品层

能点按钮，不等于就是用户默认的通用智能体

— 原语层决定 agent 能不能看屏幕、点网页、切工具、跑复杂流程。



— 但真正决定用户心智，是这些能力被如何包装成默认工作入口。



— 所以底座厂商重要，但不应替代产品层成为报告主角。



第四部分：高价值垂直代理



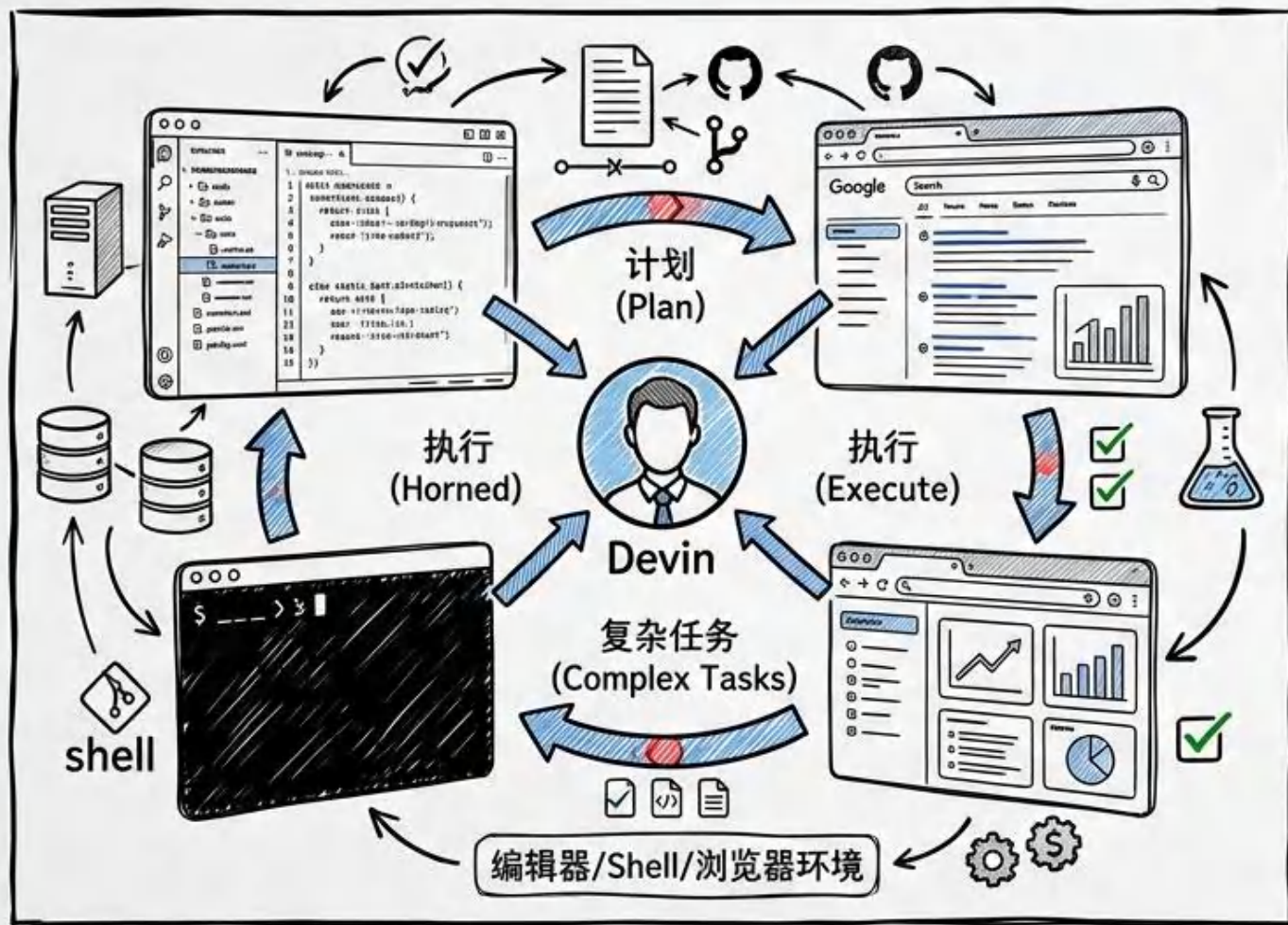
Devin 很强，但它更像**深度代理**而不是广义通用 agent



这一部分解释为什么 Devin **必须写**，但不应和 **Manus** 归为同一层。



Devin: 最典型的高价值垂直 agent



AI software engineer / collaborative AI teammate

- Cognition 把 Devin 定位为 AI software engineer 和 collaborative AI teammate。
- 它可以计划、执行复杂工程任务，并在 shell、编辑器和浏览器环境中工作。
- 后续版本还加入 **self-verify**、**auto-fix** 和 **computer use** 测试能力。

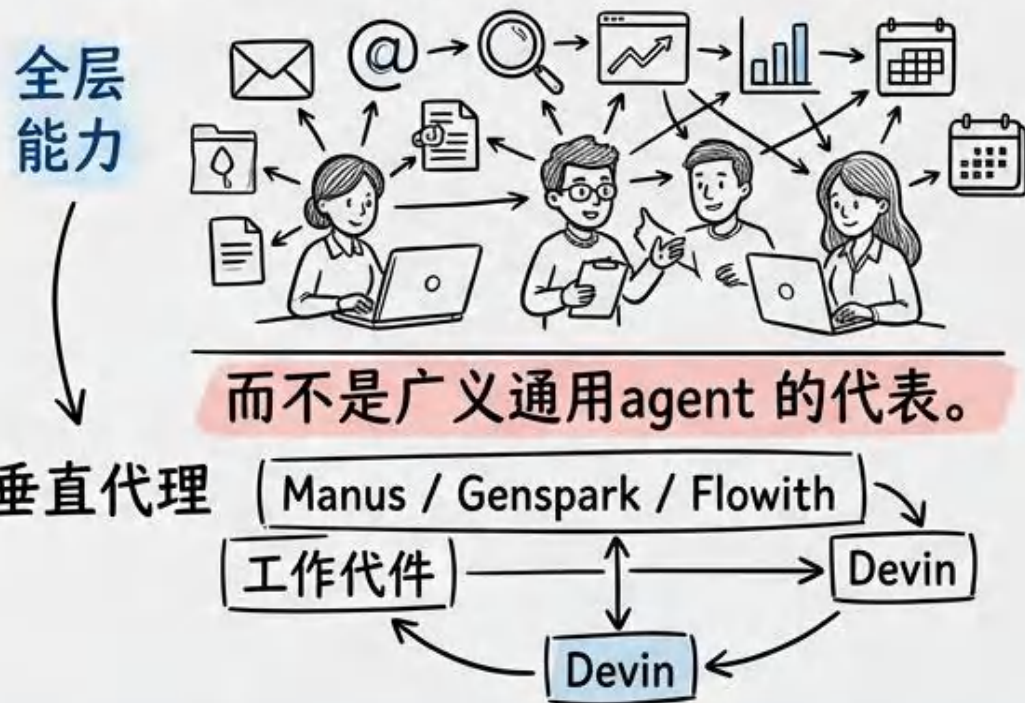
为什么 Devin 不能直接等同于“通用智能体”

深度极强，但广度不等于通用 agent

- Devin 的核心场景仍然是**软件工程**，而不是所有任务类型。



- 而不是**广义通用 agent**的代表。
- 未来更可能是 Manus/Genspark/Flowith 与 Devin 并存，而不是相互替代。





第五部分：真正的竞争维度

比较通用智能体，不能再只看模型强弱

这一部分把竞争框架从**模型分数**，改成**任务交付**和**工作台竞争**。

• 模型分数 ~~98%~~



• 任务交付



• 工作台竞争



五个真正重要的竞争维度

任务交付、环境控制、记忆、入口、治理



维度一：任务交付能力

核心能力：高效完成用户特定任务
(Task Delivery)



维度二：环境控制能力

浏览器/电脑/工具控制。环境掌握
(Environment Control)



维度三：workspace 与记忆

Workspace/项目流。工作记忆存储
(Workspace & Memory)



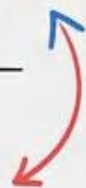
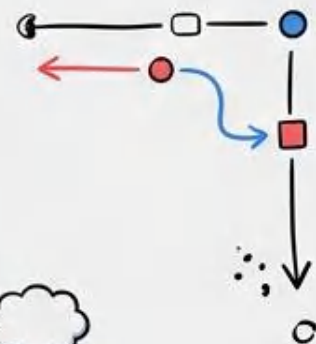
维度四：用户入口与平台黏性

核心入口点。高平台黏性
(User Entrance & Stickiness)



维度五：企业治理与控制面

企业安全。控制台、策略管理
(Governance & Control Plane)



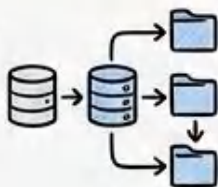
为什么未来比的是“默认任务承接方”

不是谁最聪明，而是谁最先接管你的 workflows

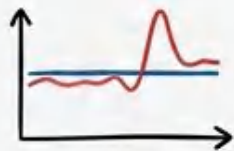
— 用户不会长期围绕抽象模型迁移，而会围绕能把活做完的 **agent** 迁移。



— 谁成为**默认任务承接方**，谁就会拥有更强的数据、记忆和工作流优势。



— 这比一次性的模型领先更更持久。



Product-layer Agents

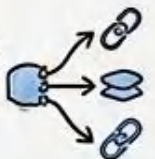
• 专注完成特定任务



• 更强的执行力和整合能力



• 流程和工作流和整合能力



• **更接近用户价值**

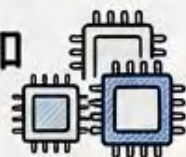


Foundation-layer Capabilities

• 提供通用模型能力



• 更侧重算法和算力优势



• 计能量算力重点

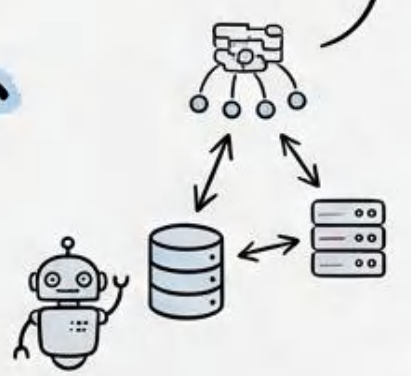
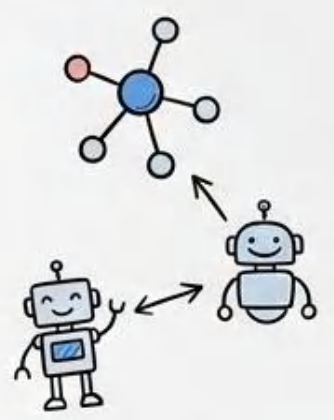


• 基础，但距离直接解决问题有一定距离

第六部分：原创核心概念

★ 用五个概念重新解释通用智能体竞争 ★

— 以下五个概念用于帮助理解产品层、原语层和垂直代理层之间的关系。





任务交付型智能体

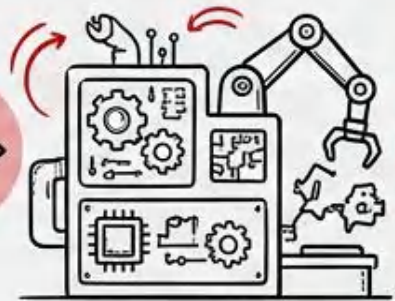
团队首创概念

任务交付型智能体 = 不是回答用户，而是替用户完成任务并交付结果

— 真正值钱的地方不在聊天，而在交付。
[用户输入 (任务)] -> [智能体 (执行器)] -> [任务结果 (交付物)]



用户需求

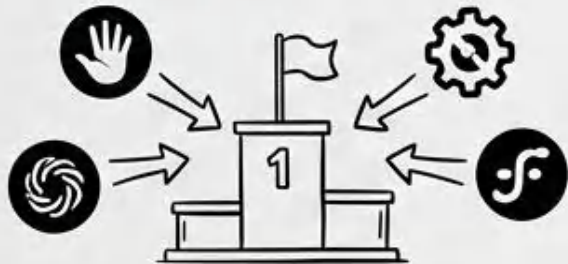


智能体



替用户完成 交付结果

— Manus、Genspark、Flowith 都在争夺这个位置。



— 未来用户选择 agent, 更像选择数同事而不是选择问答模型。

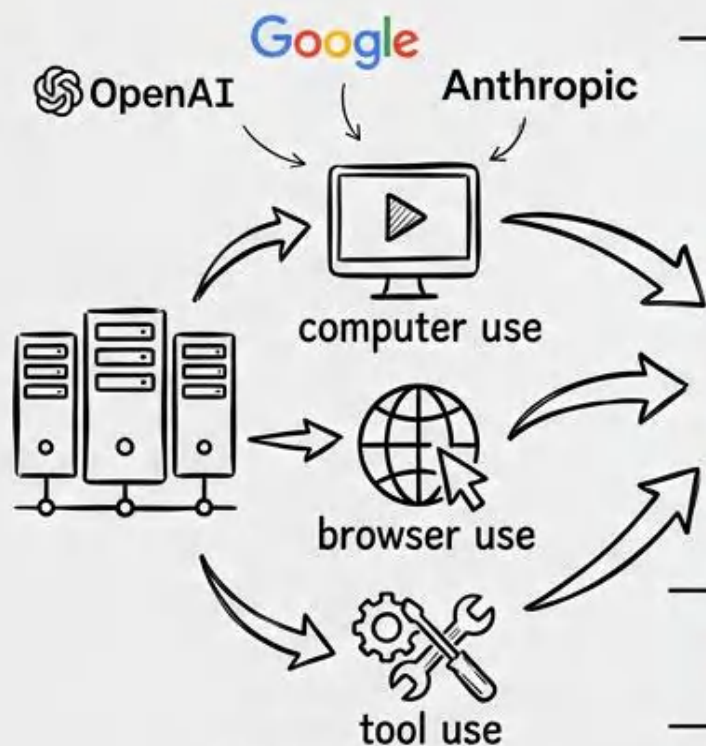


能干 高效

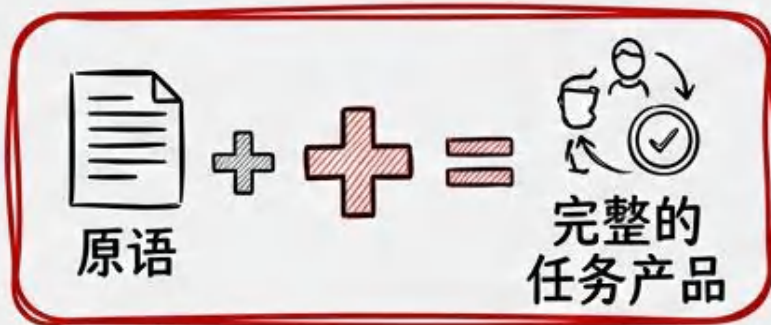
原语层与产品层分离

团队首创概念

- 原语层与产品层分离 = 底座厂商提供**动作能力**，产品层玩家提供**最终体验**



— OpenAI、Google、Anthropic 更像是在提供 **computer use、browser use、tool use** 原语。



— OpenAI、Google、Anthropic 更像是在提供 **computer use、browser use、tool use** 原语。

— 真正占领用户心智的，是把这些原语做成**完整任务产品**的人。

— 这就是为什么产品层报告**不能再被底座层抢走中心**。



产品层
(最终体验)

原语层
(底座厂商)

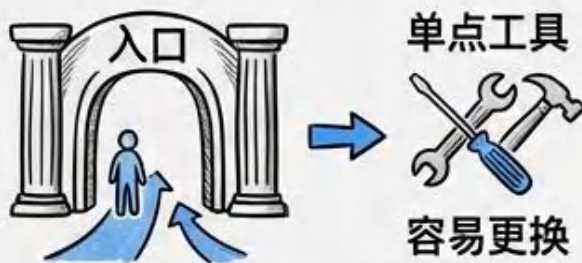
工作台护城河

工作台护城河 = 用户把常见工作都迁入同一 workspace 后形成的高切换成本

— Genspark 和 Flowith 都在争夺这种护城河



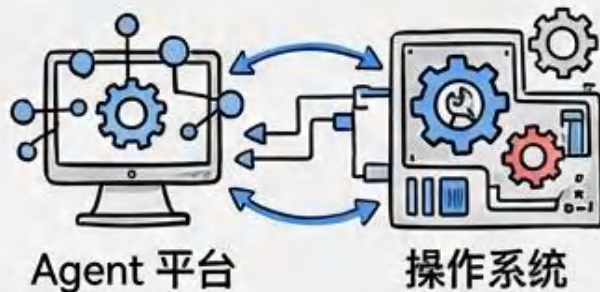
— 入口一旦形成，迁移成本会远高于单点工具



— 入口一旦形成，迁移成本会远高于单点工具



— 未来 agent 平台竞争会越来越像操作系统竞争

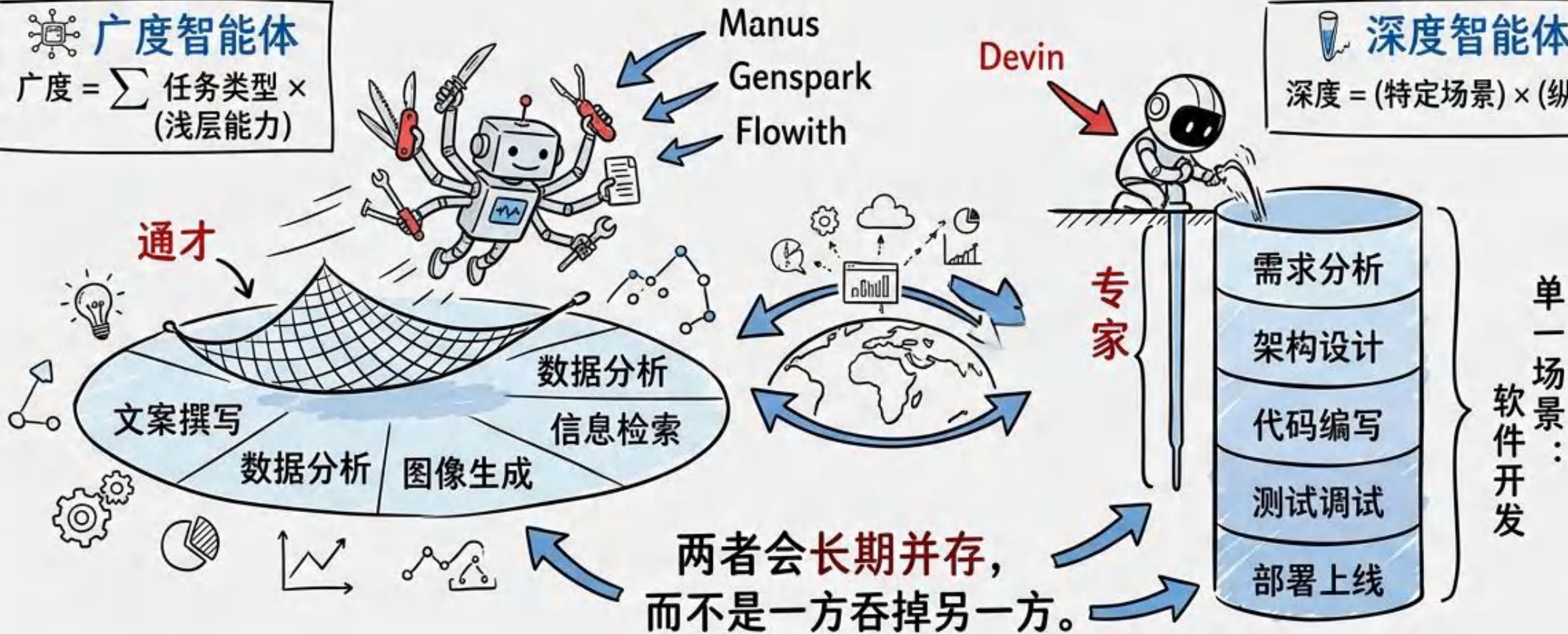


广度智能体与深度智能体

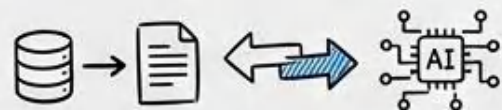
→ 广度智能体覆盖更多任务类型，深度智能体在单一场景做得更深

广度智能体
广度 = \sum 任务类型 × (浅层能力)

深度智能体
深度 = (特定场景) × (纵深能力)



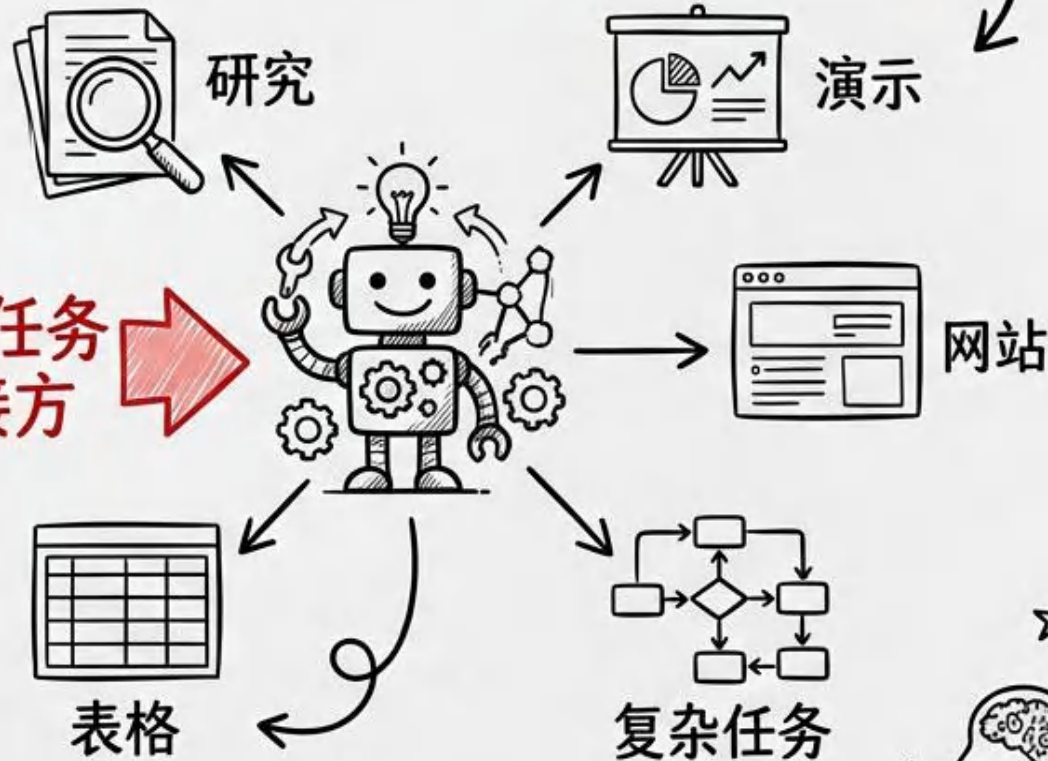
交付替换权



交付替换权 = agent 成为用户默认任务承接方后的持续支配力

- 一旦某个 agent 默认接研究、演示、网站、表格和复杂任务，它就**不再只是工具**。
- 它会变成**新的工作入口**。
- 谁先拿到这层位置，谁就更接近通用智能体时代的**平台红利**。

默认任务承接方





第七部分：未来两年判断



接下来两年，通用智能体竞争会越来越像产品与入口竞争



→ 本章收束为几个最值得执行层关注的判断。



未来两年的三条主判断

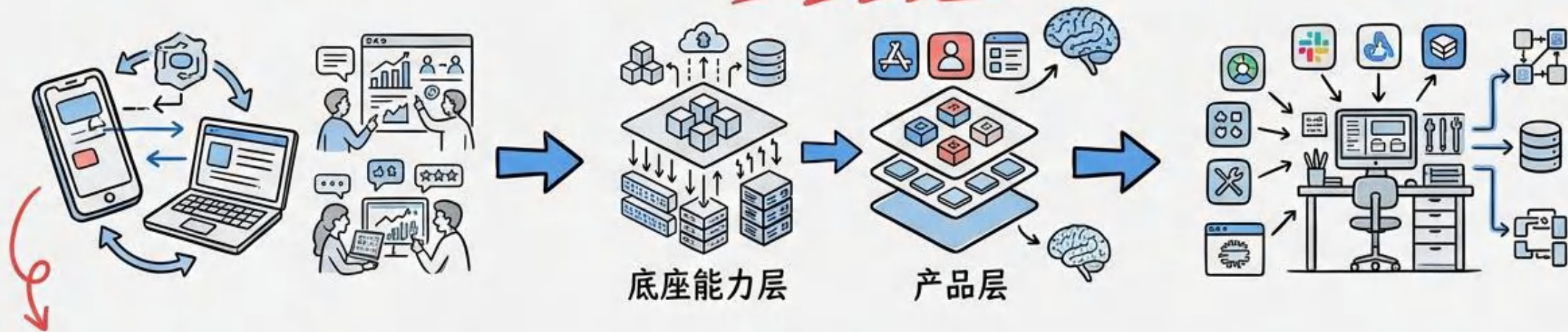
产品竞争强化



原语层继续分化



工作台价值上升



判断一：通用智能体竞争会越来越像**产品竞争**，而不是**模型排名竞争**。

判断二：底座能力层和产品层会**继续分离**。

判断三：工作台、记忆和入口的**价值会持续上升**。

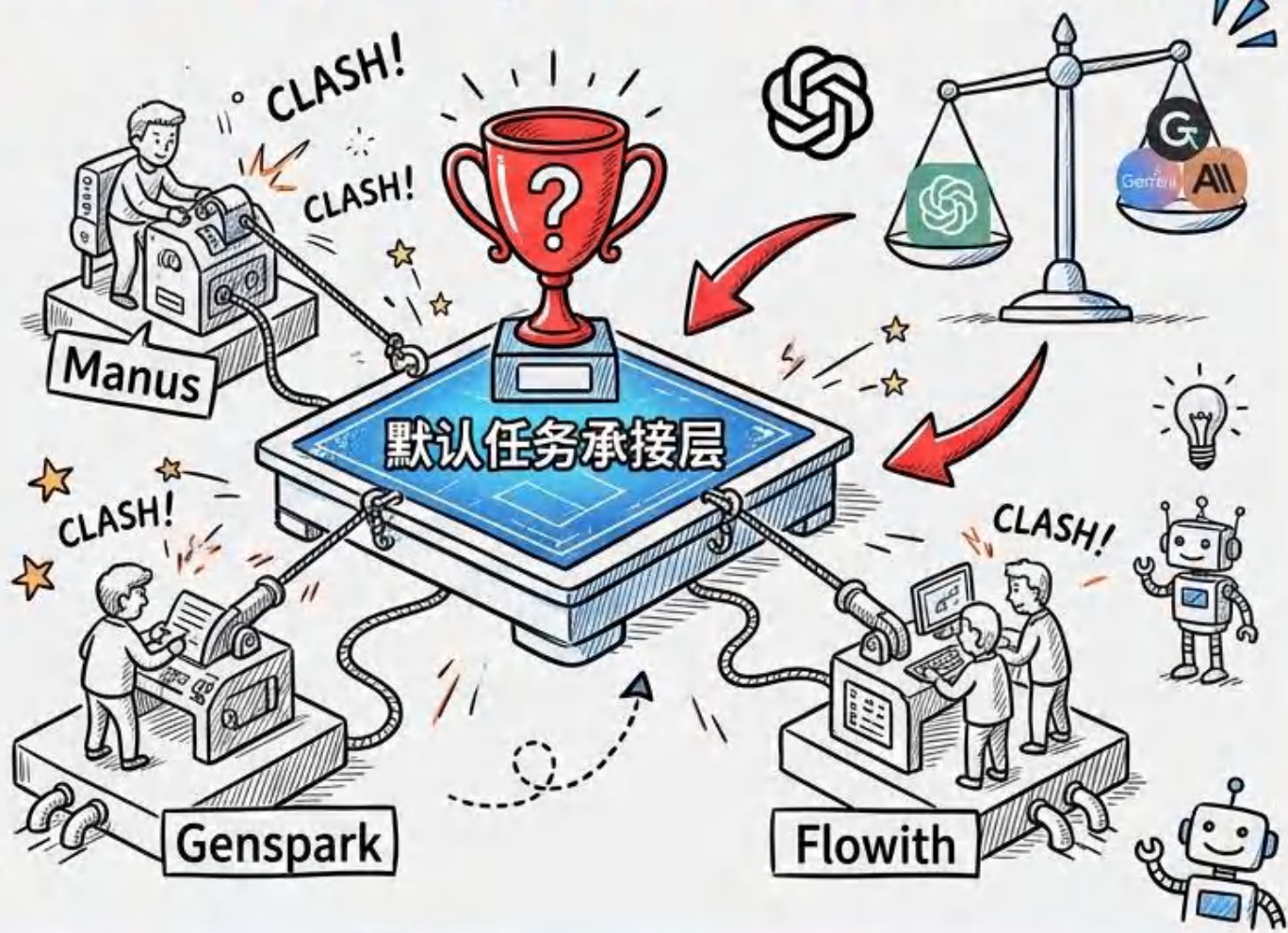
最终判断

真正的通用智能体报告，应该把 Manus / Genspark / Flowith 放在中心

— 如果报告主轴还是 GPT、Gemini、Claude，那就还不是通用智能体报告。

— 真正的主角是那些已经开始接任务、拆任务、做任务、交付任务的产品。

— 未来的胜负将更多取决于谁能成为 默认任务承接方。



智灵动力 Desearch

深度、精准、可验证——这才是 AI 研究报告该有的样子

深度挖掘

不是表面面植到信息，
但是逻辑严格的行业调听

数据可验

每个结论面包括据来源
拒绝 AI 幻幻想，支持从网页 / 行业数库
其让复杂研究表一目可能解

智能可视化

- 不是“交差式摘要”，而是像专业分析师一样交付深度研究报告
- 自动构建研究框架、精准检索高可信信源、生成结构严谨的数据溯源报告
- 一次提问，全场景交付：报告、图表、PPT、播客脚本、音频



AI 研究分分析师

个人研究者：学术综述 / 前沿技术追踪
市场分析师：竞品监控 / 行业趋势研判
企业战略部：新市场进入策略 / 投资可行性分析
投资机构：尽调速报 / 赛道扫描
咨询顾问：项目前期框架搭建

A. 数据可溯源 + 可视化

- 数据自动标注来源
- 关键数据自动生成图表

B. 公私数据融合

- 支持上传内部文档、财报、竞品分析、专利库
- 构建专属知识图谱，输出高度定制化内容

C. 三种研究模式

- 普通模式：快速扫盲
- 深度模式：系统性分析
- 专家模式：战略级报告

D. 本地化定制解决方案

- 私有化部署
- 多人协同
- 审计追踪
- 数据本地隔离

- 10分钟入门一个新行业
- 5分钟追踪社媒平台热门信息
- 3分钟了解企业全景画像
- 一文解读政策影响
- 一键内容智能转化为 Word / PPT / 播客脚本 / 音频

案例部经来玩的使用如例

C端：单篇报告约 1~10 元

B端：年费制 + 私有部署 + 团队账号 + 定制训练

感谢观看

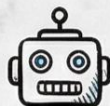
THANK YOU



执行人: 杜靖洋



提议人: 沈阳



具体实施: OpenClaw



资料搜索及整理: ZeeLin Desearch



PPT生成: Gemini 

