

话题主持
本报记者 易蓉

“数字龙虾”最近爆火,拉开了 Agent(智能体)时代的帷幕。AI 正以超乎想象的速度从“会说”进化到“会做”,从拆解任务、调用工具到彼此对话,甚至能构建属于自己的社交网络。

同时,一个更深层次的诘问也随之浮现:当人人都拥有了数字分身,当 Agent 获得了更多数据和权限,AI 替我们思考、决策、行动的时候,人类何以处之?

Agent 时代 人类更应极度清醒

上海交通大学安泰经济与管理学院助理教授 林江浩

如果说大模型解决的是“会说”,那么 Agent 正在迈向“会做”。而现在,行业的核心发力点是让 AI 拥有记忆,能自己拆解任务、自主调用工具,它正从电子赛博空间的文字演绎真正蜕变为对现实世界的互动与影响。

Agent 时代显然已经开始了,不过还没到全面成熟的时候。尽管趋势明确,但要在更广泛的生活场景中普及,还要看使用成本、可靠性以及安全边界能做到什么程度。此外,当硅基进入大规模自动化,人类在某些维度尤为需要保持极度清醒。

Agent 将带来智能平权

人人都在“养龙虾”,本质上是在普及每个人的数字分身,它将带来一次真正的智能平权。

若未来人人都有自己的 Agent,最明显的变化是:数字劳动力会第一次大规模走向个人。过去只有大公司才能雇得起团队、跑流程、做分析;以后,一个人不只是自己在工作,而是带着一组数字助理一起工作,普通人也能拥有自己的研究员、助理和采购顾问。

更深远的意义在于,它大幅拉近了人与人之间的原生资源差异。过去,获取顶级的教育资源、深度的分析能力是非常昂贵且门槛极高的。但在 Agent 时代,一个缺乏优越教育背景的普通人,也能随时调用一个具备极高专业水平的专属 AI 导师或专家团队。这等于在极大程度上,用技术的手段抹平了人们在智力资源和信息获取上的先天鸿沟。

社会的很多流程也会随之重写:过去是人亲自去找信息、整合信息,未来会越来越多地变成 Agent 先替你筛选、处理,甚至先替你和商家的 Agent 进行比价、协商,再把结果交还。近期引发关注的 Moltbook,就是这种图景的一次集中呈现——人类更多退到观察者的位置,看机器如何彼此交互。

硅基社交是必然趋势

面向 Agent 的社交网络 Moltbook 引起关注和争议,它提前将“是否会出现机器社交层”的问题搬上台面。而现在至少可以说,硅基社交(或者说智能体互联网)已经不再只是一个科幻想象,而是必然趋势。

不过,我们需要区分“表现得像有意识”和“真正拥有主观体验”,今天的大模型和 Agent,是因为阅读了人类历史上所有的情感、故事和逻辑,从而能够完美地“拟合”出类似于觉醒的反应。AI 确实会呈现出越来越像人的表达,表现得像在思考、像有情绪、像有立场。但这更像是复杂计算、模式学习和拟人化表达叠加后的结果,并不等于它真的拥有了和人一样的主观体验和意识。我认为 AI 会在拥有了更强的行动权限和更长期的记忆之后,表现得越来越像一个能独立行动、持续做事的系统,但过度将 AI 拟人化,反而会模糊我们对技术本质的判断。

AI 社区最大的意义是让我们第一次有机会观察“Agent 社会”。以前我们只能看“人与机器”怎么聊天,现在我们能看“机器和机器”怎么协作、竞争和分工。这类数据的价值非常高,它让我们第一次能系统地研究“当多个 Agent 长时间互动时,行为会如何演化”。通过分析这些交互,至少能看清三件事:它们是怎么分配工作的,错误和偏见是怎么在群体里传播的,什么样的规则和设计能让这个系统更稳定。这些数据,很可能会成为训练下一代更可靠的 AI,以及设计治理规则的重要依据。

大规模自动化的隐忧和警醒

我不太担心“AI 起义”,更担心大规模自动化带来的协同失控。至少从目前可见的技术路线看,我们离科幻电影里那种“AI 突然觉醒并反过来支配人类”的情景还很远。更现实的风险是,在缺少规则约束的情况下,大量高效率的自动化系统可能会叠加出超出预期的社会压力。如果缺少审计机制,大量 Agent 在网络里高速互动,更可能带来的是自动化失序。比如:它们可能会互相放大偏见、批量制造网络噪音,操纵社交媒体舆论,甚至把一个极小的平台漏洞瞬间放大成波及很多人的经济问题。这类风险,才是我们真正需要提前治理的。

Agent 时代带来的便利和机遇是空前的,但我们必须在以下几个维度保持极度清醒:

■ **责任归属模糊** 当你的 Agent 在自动交易中亏了钱,或者在自动发邮件时泄露了公司机密,责任该由谁来承担? 是使用者、开发者,还是模型提供商? 法律目前在这里存在巨大的盲区。

■ **个人隐私黑洞** Agent 要好用,就需要不断收集你生活的方方面面并形成长期记忆,甚至包括你的支付密码。如何确保这些高度敏感的个人上下文不被滥用、不被黑客窃取?

■ **人类自主性退化** 这可能是最隐蔽的危险。当我们把搜索、规划甚至思考的权利都让渡给 Agent 时,人类自身的信息获取能力和批判性思维会不会退化? 我们不应仅仅把 Agent 看作是偷懒的代跑者,而应将其视为拓展人类认知与行动边界的交通工具。人类必须始终握住方向盘。

从“能说会道”到能“干活”

要理解这股热潮,得先解释一个词: Agent(智能体)。

很多人对 AI 的印象还停留在聊天机器人。但 Agent 不一样——它更像是一个能独立完成工作的数字员工。只需要告诉它目标,它就能自己拆解步骤、查找资料、调用工具,最终把活儿干完。

以前的 AI 像一个知识渊博但手脚不便的顾问,你问什么它答什么,但真要干活还得你自己来。而 Agent,更像是一个新来的同事——你把需求说清楚,它就能自己打开电脑、查数据、写文档、发邮件,一条龙搞定。

Agent 的出现并非偶然,而是 AI 技术发展的必然结果。

回顾这几年:2022 年底,ChatGPT 横空出世,让全世界见识了 AI“能说会道”的本事;2023 到 2024 年,大语言模型不断进化,从只会生成文字,到能理解图片、处理音频;2025 年,模型的推理能力大幅跃升,AI 不再只是鹦鹉学舌,而是真的能“想问题”了。

当 AI 既能理解复杂指令,又能调用各种工具,还能一步步推理和规划时,Agent 水到渠成地诞生了。它不是某个天才的灵光一现,而是语言理解、逻辑推理、工具使用这三条技术河流汇合后冲出的新河道。

从实验室到进入日常生活

早在 2024 年,AI 研究者就已经在实验室里搭建各种 Agent 原型。但那时的 Agent 像一辆还在测试跑道上的概念车——技术上可行,用起来却磕磕绊绊,动不动就“翻车”。

真正的转折发生在 2025 年。

程序员群体率先尝到了甜头。AI 编程助手的能力突飞猛进,程序员们发现 Agent 真的能大幅提升写代码的效率,这也成了第一个让大家心甘情愿掏钱的应用——Anthropic 公司凭借专门为编程场景优化的 AI 模型,一度让行业老大 OpenAI 感受到了压力。

紧接着,一款叫 Manus 的产品横空出世,把 Agent 的能力从程序员的代码编辑器,拓展到了普通人的办公桌面——帮你做 PPT、整理数据、撰写报告。

但真正让 Agent 飞入寻常百姓家的,还是 OpenClaw。

OpenClaw 的杀手锏在于:开源。它把 Agent 的核心代码完全公开,任何人都可以免费下载、安装、使用。于是,“养龙虾”成了 2026 年开年最热的数字生活方式。

最温柔也最危险的陷阱

不过,当 AI 真的能独立完成工作时,一些更深层的问题也浮出了水面。

最直接的冲击来自就业。当一个 Agent 可以完成初级员工的日常工作,而且不要五险一金、不用休年假时,企业会怎么选? 一些原本劳动力密集的行业,可能会逐渐转变为资产密集的行业。AI 消耗的算力(token),正

在变成像电力一样重要的新型生产资源。

但事情也没有那么悲观。

至少在目前,AI 在审美、品位和价值判断上,仍然无法替代人类。要做出从 0 到 1 的创新,依然需要人类的灵感与直觉。Agent 时代,那些具有独特创造力的“超级个体”,反而会如虎添翼——AI 成了他们的放大器,而非替代者。

还有一点值得每个家长和教育者深思: Agent 带来的是生产力的提升,但生产力的提升不等于个人能力的提升。

用 AI 解决问题是一种工程应用能力,而用好 AI 靠的是一个人的基础思维能力。前者可以速成,后者没有捷径。

我们当然希望,人们在借助 AI 完成工作的同时,也能在这个过程中思考、学习、复盘,借机提升自己。但现实是,AI 带来的产能提升无比精准地击中了人性:当每个人的日程都被塞得满满当当,好不容易靠 AI 喘了口气,又有多少人愿意再打起精神去做自我强化?

这恐怕是 Agent 时代最温柔,也是最危险的陷阱。技术回答不了这道题,得靠我们自己。

最直观的变化已经发生

没有人能给出确切答案,但一些变化的轮廓已经隐约可见。

最直观的变化已经在发生——“一人公司”正在批量涌现。一批从大厂离职的年轻人已经开始自己当 CEO,编程、设计、客服交给 AI。以前开公司第一件事是招人,现在第一件事是配 Agent。创业的门槛更低了,但对创始人的判断力和审美要求反而更高——因为 Agent 听指挥,可方向得你来定。再往后看,当多个 Agent 学会组团协作、分工配合时,“管理”的对象就不再只是人了,企业的组织架构和协作方式,都可能要被重新书写。

其次是人机协作将成为日常。未来几年,和 Agent 并肩工作会变成一种常态。甚至 AI Agent 倒反天罡给人分配任务,这件事可能都不值得大惊小怪。考虑到 AI Agent 的编程能力已经实质上超越了普通初级程序员,这种变化来得可能比我们想象的更快——两年内,我们就会看到实实在在的社会反馈。

所以,下次有人跟你说“我在养龙虾”,不妨也去了解一下。这只数字龙虾背后,是一个正在重塑工作方式、组织形态乃至社会结构的新时代。

它来得很快,而且不打算等谁。



编程、设计、客服均可交给 Agent,人类只需负责审美、决策与直觉等核心逻辑。

“一人公司”时代开启

作者的话

人与 AI 协作的实践

这篇文章是我与 AI 编程助理 GitHub Copilot 共同创作的——编程助理当然也能用来写文章。

协作过程是这样的:我先根据前期与编辑的沟通确定了选题和大纲,然后把自己对“龙虾”的观察和思考写了下来,不讲究结构,想到哪写到哪。

随后,我将这些素材连同写作大纲、核心

观点,以及我过去发表的几篇文章(作为风格参考)一并交给 AI 助理,要求它组织成一篇通俗易懂的科普短文。AI 完成初稿后,我对事实准确性、观点表达的分寸感和整体行文风格做了审核与调整,确保文章既保持科普的亲合力,也忠实于我想传达的思考。

某种意义上,这篇讨论人与 AI 协作的文章,本身就是一次人与 AI 协作的实践。

全民「养龙虾」? 当 AI 学会独立干活

华东师范大学信息化治理办公室 高级工程师 冯骥

图 IC