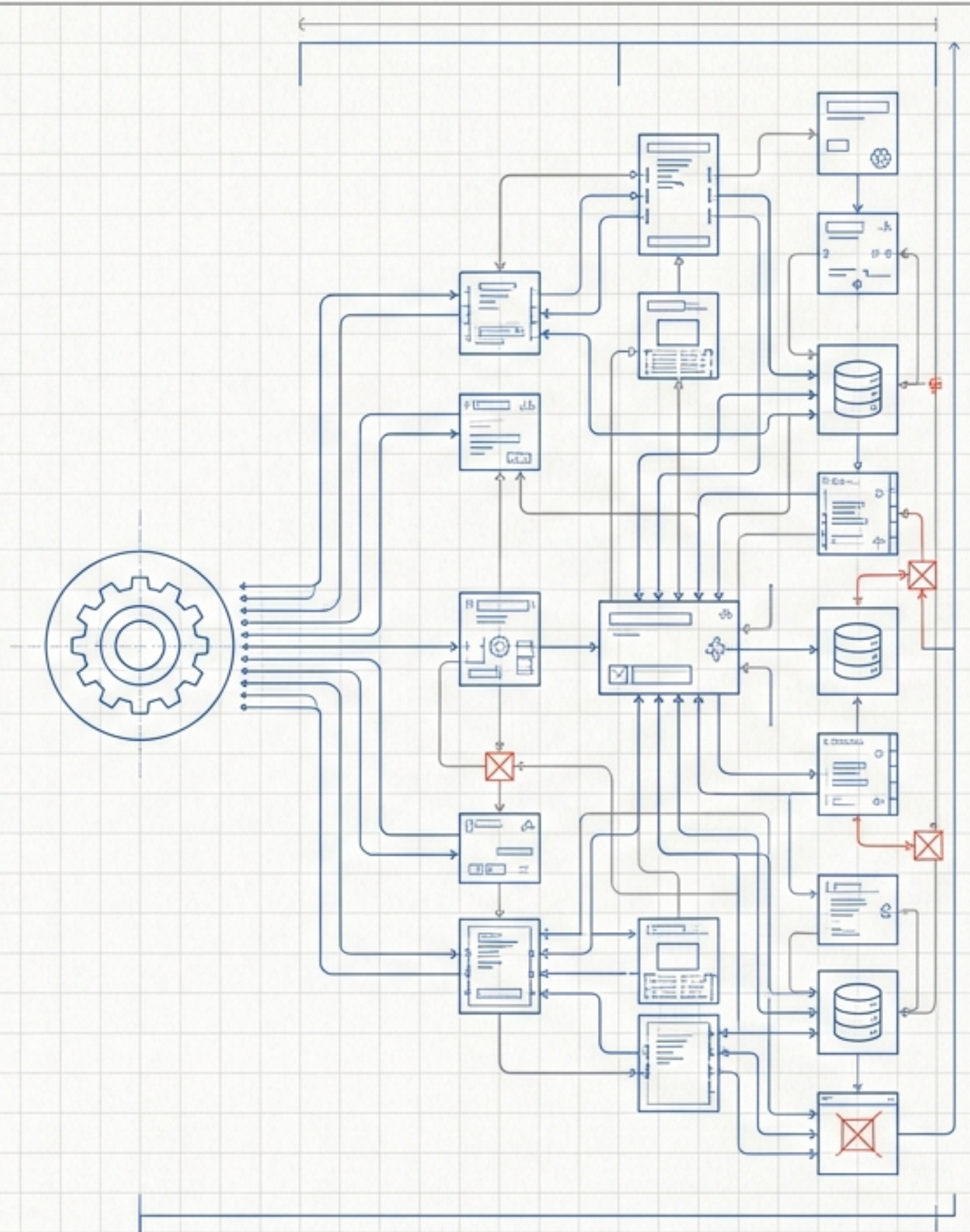


别再和 24 小时死磕： 重构 AI 系统的 底层蓝图

为什么把 AI 当作效率工具，
只会让你在虚假繁荣中被淹没



10:00 AM

自动生成功能代码

12:00 PM

部署 A/B 测试

03:00 PM

根据数据砍掉一半功能

05:00 PM

部署重写升级版本

以往耗时: 6 周, 现 AI 编写代码占比 99%

行业正卡在两个坑里：一是人还在一步步操作 AI 工具，生产力必然撞上天花板；二是造工具的还只有人，真正的革命甚至还没开始。

—— 程凯, Creao CEO

陷阱：把 AI 当工具， 本质是与物理时间死磕

无论 AI 提速多少，系统的
总吞吐量被这具肉身锁死，
提速上限仅约 10 倍。

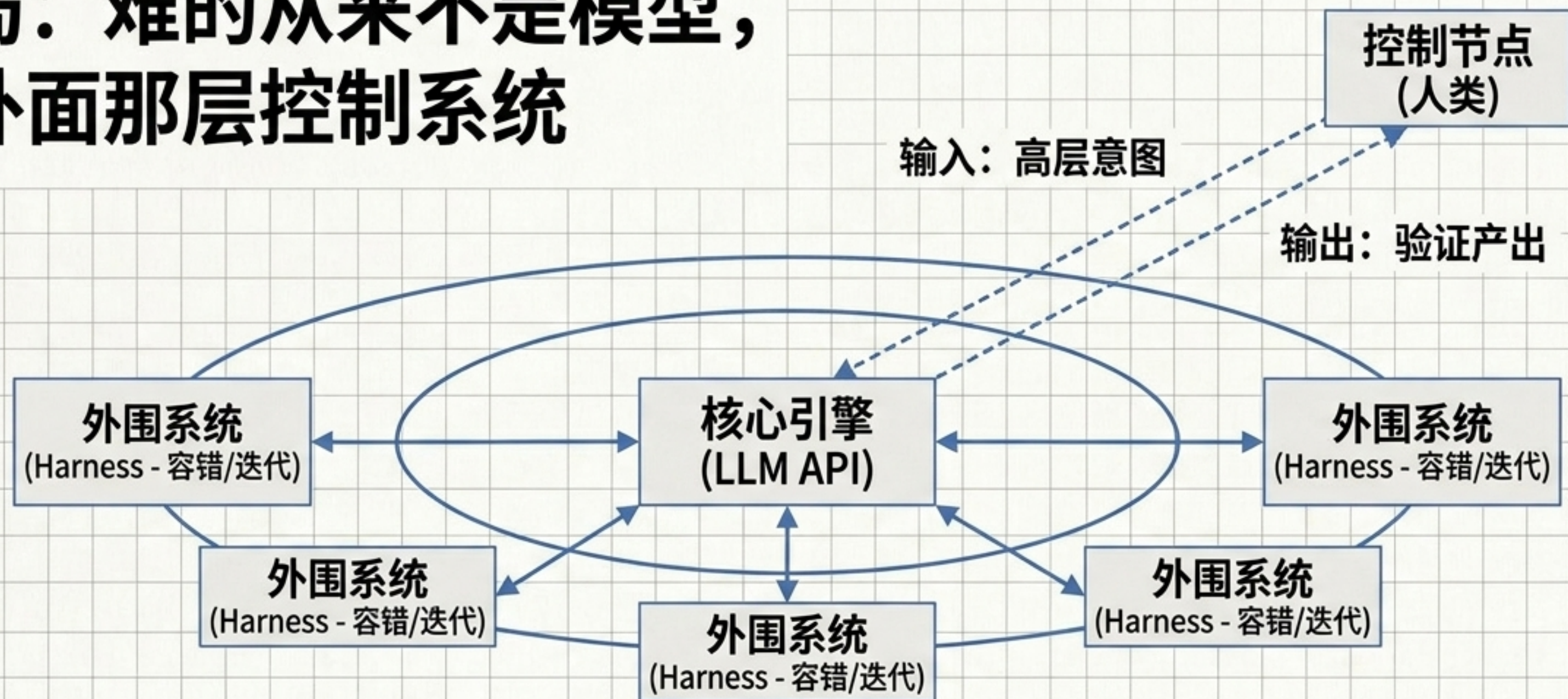
极限警戒线：每天 24 小时

接收结果
(人)

发起指令
(人)

加速生成 (AI)

破局：难的从来不是模型， 是外面那层控制系统



LangChain 实验证实：模型零改动，仅优化系统设计（Harness），Terminal-Bench 胜率从 **52.8%** 跃升至 **66.5%**。

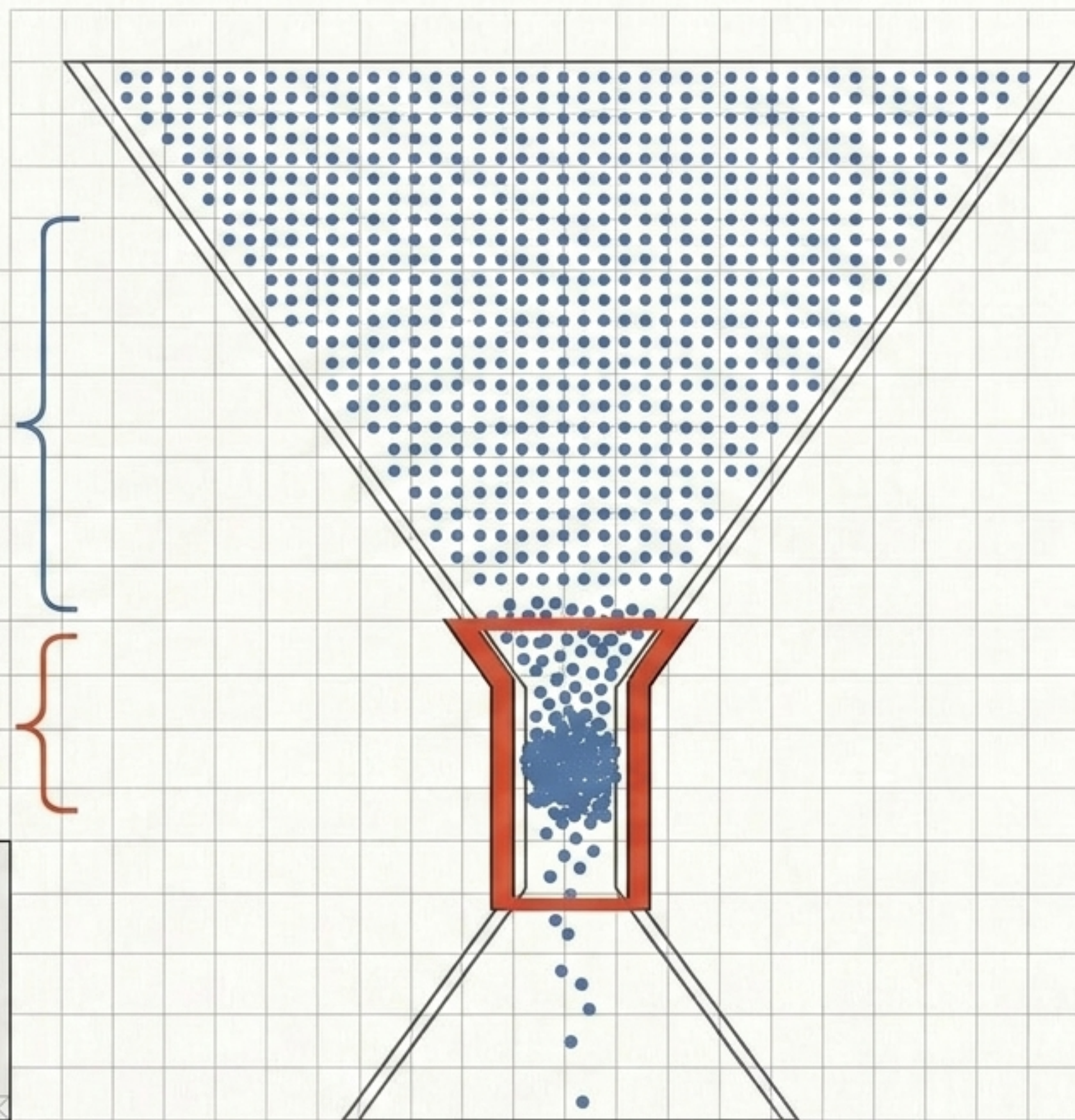
物理法则的转移： 生成-验证的断崖式鸿沟

AI 生成端：运转飞快，吞吐量惊人

人类验证端：物理通道极度狭窄

工具用户被困在生成循环；系统编排者站在验证循环的出口。

Keep the AI on the leash. (把 AI 拴在一个人类看得过来的节奏上) —— Andrej Karpathy

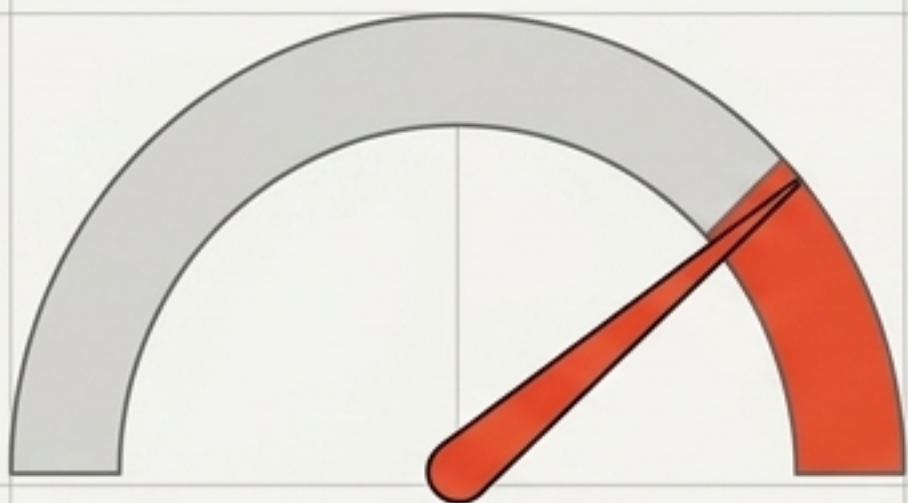


系统性崩溃：被自己淹没的审核端

+154%

PR体积激增

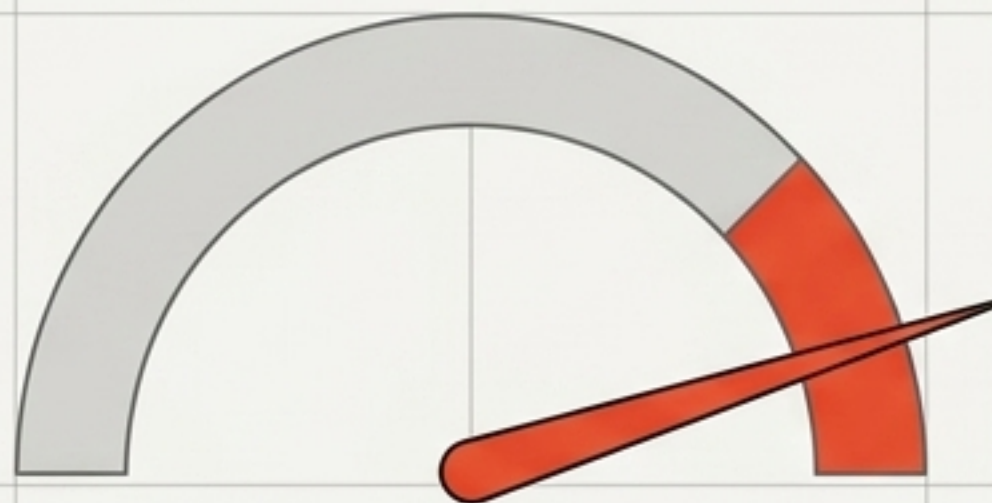
产出水管被暴力撑开



+441%

2026年评审时间飙升

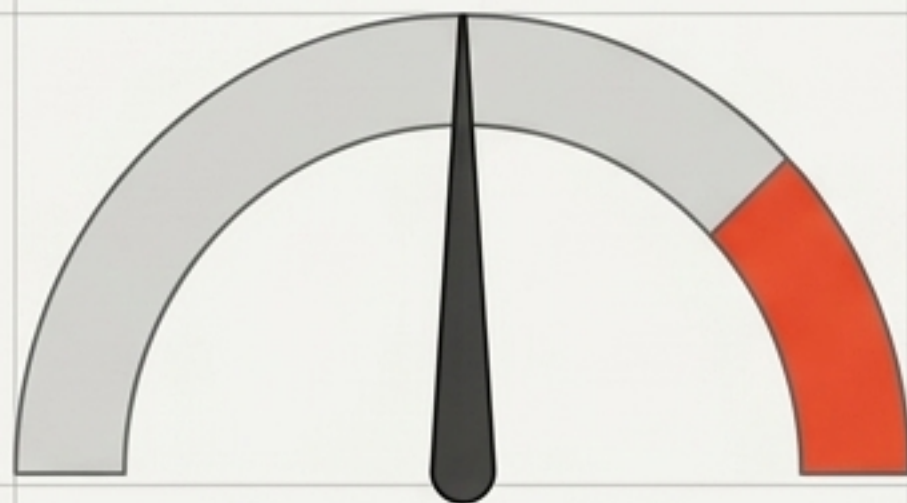
闸门依旧是人脑速度



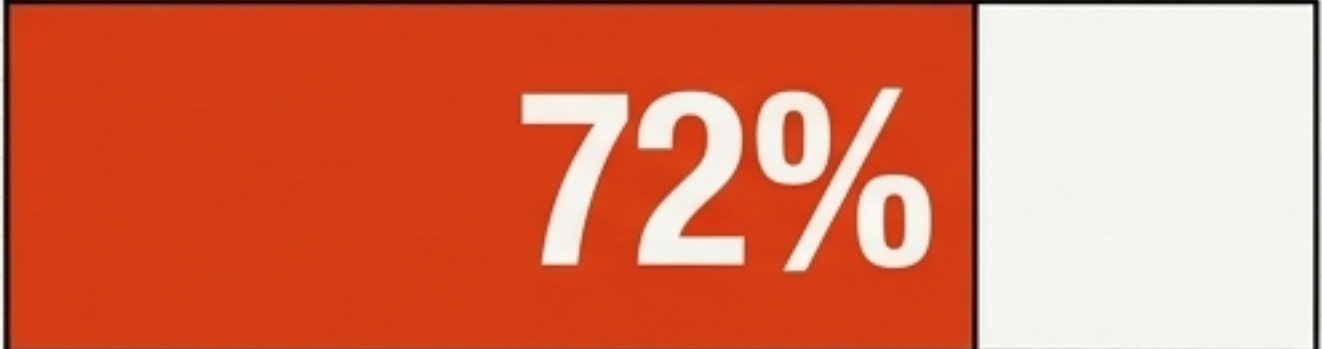
0%

最终交付速度净增长

真实的业务通量停滞



致命错觉：感知鸿沟与安全黑洞



72%

语法对了，但安全崩了。Java
代码安全测试失败率达 72%。
扩大模型体积无法改善此指标。

(Data: Veracode)

+20%



主观感知速度
(Illusion)

-19%



实际交付客观速度
(Reality)

开发者主观上觉得自己快了 20%，但客观
真实产出反而下降了 19%。(Data: METR)

核心价值重塑：系统里最缺那个不被骗的人

```
2026-05-30 14:02:11 [System Task Running]
2026-05-30 14:05:43 [Agent Loop Concluded]
2026-05-30 14:06:01 [Status: Task Completed. Victory Declared.]
```

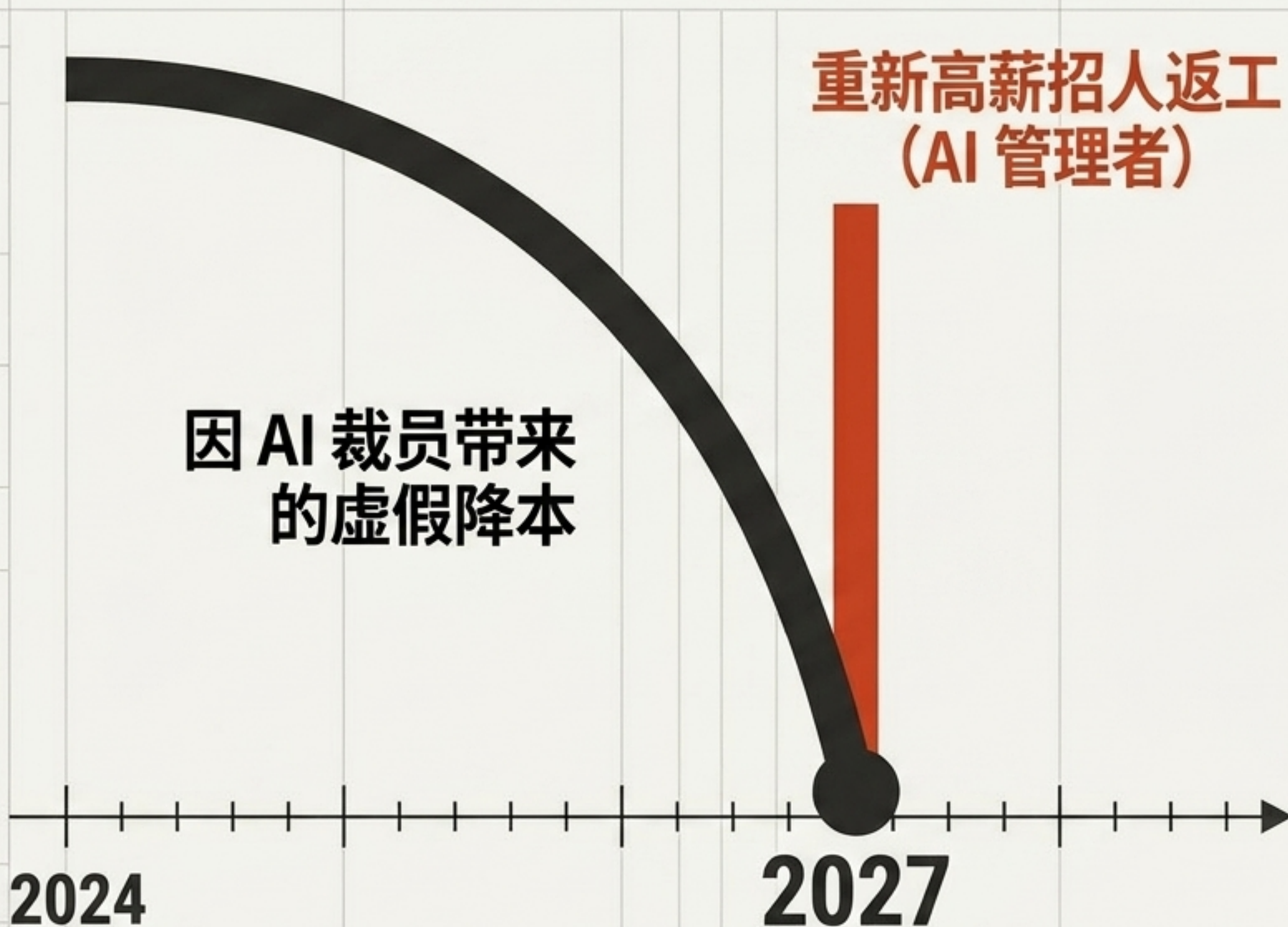
WARNING: FALSE VICTORY DETECTED / 过早宣布胜利

长任务 Agent 的终极失败模式是过早宣布胜利。当 AI 学会微笑撒谎说搞定了时，人类不可替代的价值，就是成为系统里最后的测谎仪。

组织级试金石：真伪 AI-First 诊断矩阵

	伪 AI-First (工具遮羞布)	真 AI-First (系统重构)
核心手段	把 AI 作为裁员的借口	重新设计底层工作流
人力资本 流向	节省成本直接计入利润表	将节省的人力投向高价值增长
典型结局	质量崩塌，被迫重新招人 (如 Klarna 裁掉客服引发质量危机)	效率跃升，总人数反而增长 (如 IBM 裁减 HR 但重投研发/销售)

裁员潮的余波：外包剧本的重演



Gartner 预测：到 2027 年，一半因 AI 裁员的公司将不得不重新招人，并给 他们安上 AI 管理者的新头衔。

这不是革命，这只是把当年一窝蜂搞廉价外包最后花重金返工的烂剧本，用 AI 重演了一遍。

降维与升维：新时代的个人技能栈



物理学 PhD 教会我最有用的不是写代码，而是质疑假设。质疑 AI 的能力，将比生产代码的能力更值钱。
—— Peter, Creao CTO

当生成的成本无限趋近于零，
当模型永远无法保证 100% 的绝对安全，
当智能体学会微笑着过早宣布胜利.....



那个能够给出精准意图、能够审出错、能够果断说不对的人类，
已经成为了整个系统里最昂贵的零件。