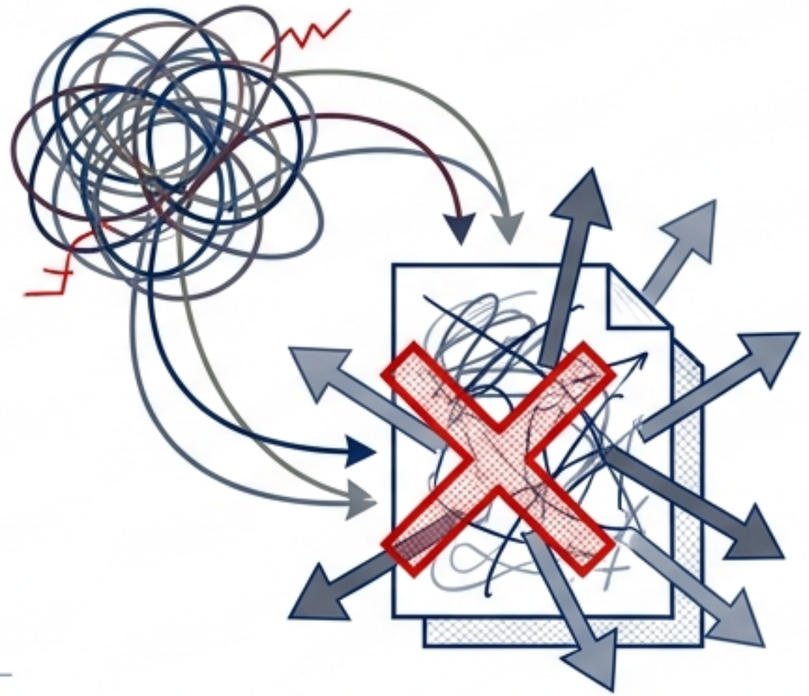
指挥 Agent 舰队：重构投资尽调研究的系统架构

跨越单点提示词的性能天花板，打造具备验证、迭代与纠错能力的自动化研究流

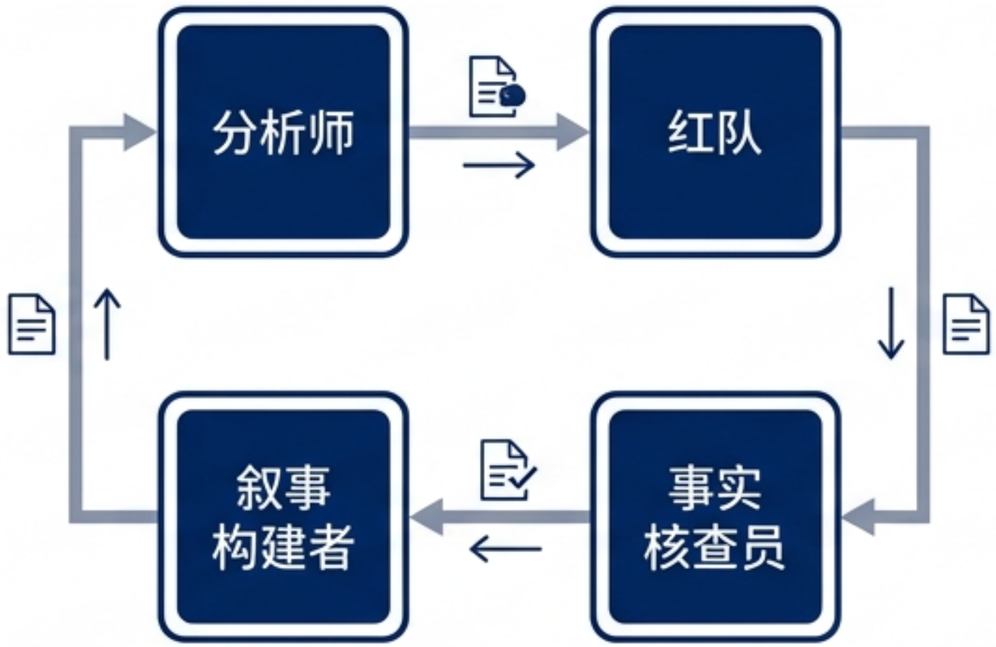
一个人同时扮演四个角色，只会输出内部打架的混乱报告

SINGLE AGENT OPERATIONS



单兵作战：自说自话，缺乏交叉验证

FLEET COLLABORATION

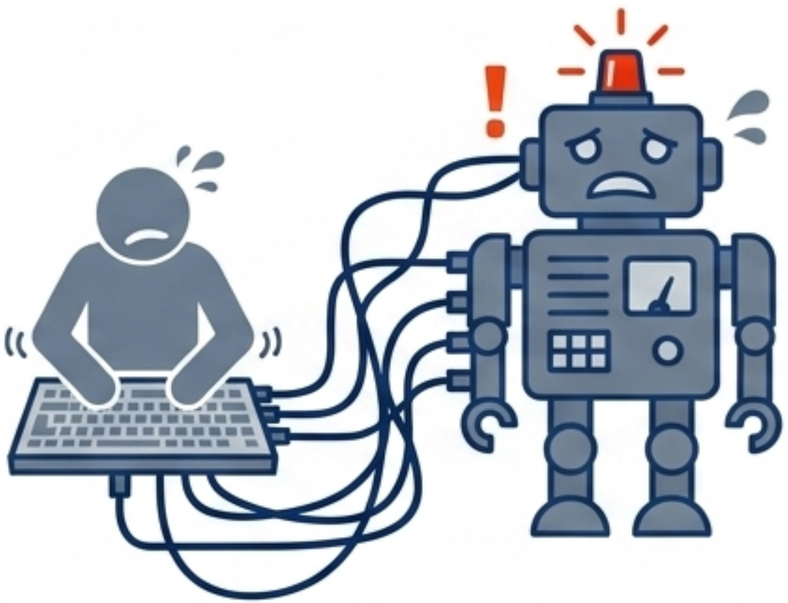


舰队协作：明确角色界限，显式交接

问题出在把本该是一支团队的工作交给了一个人。三个各自能力很强的人丢在一起不等于好团队，缺的是分工、交接和验证规则。

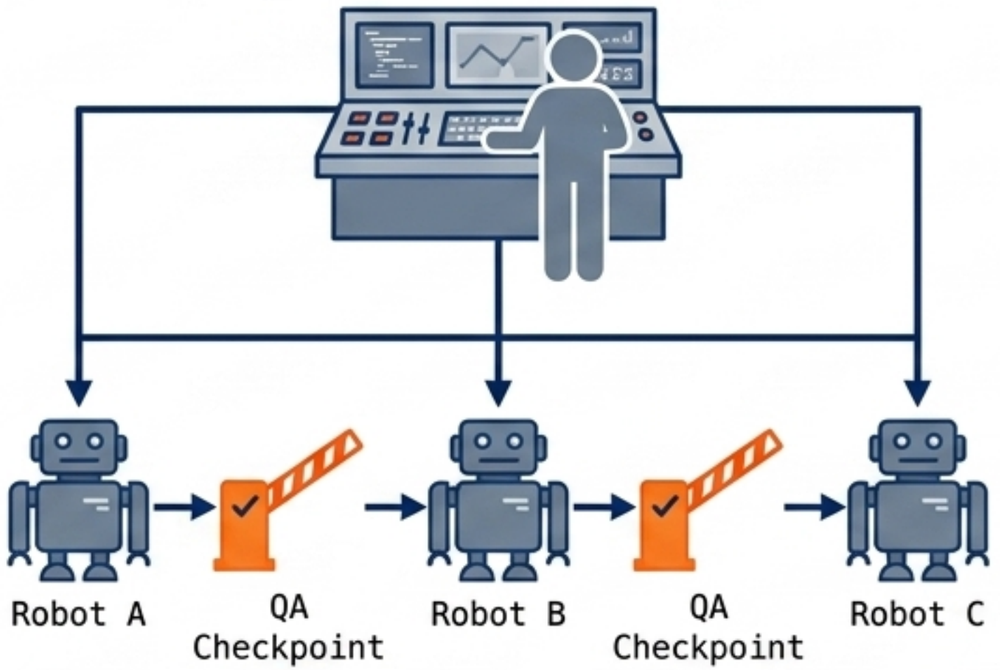
认知转换：从“提示词工程师”到“系统架构师”

旧模式：提示词工程师



试图通过完美的提示词让 AI 一次性产出完美的长篇报告。

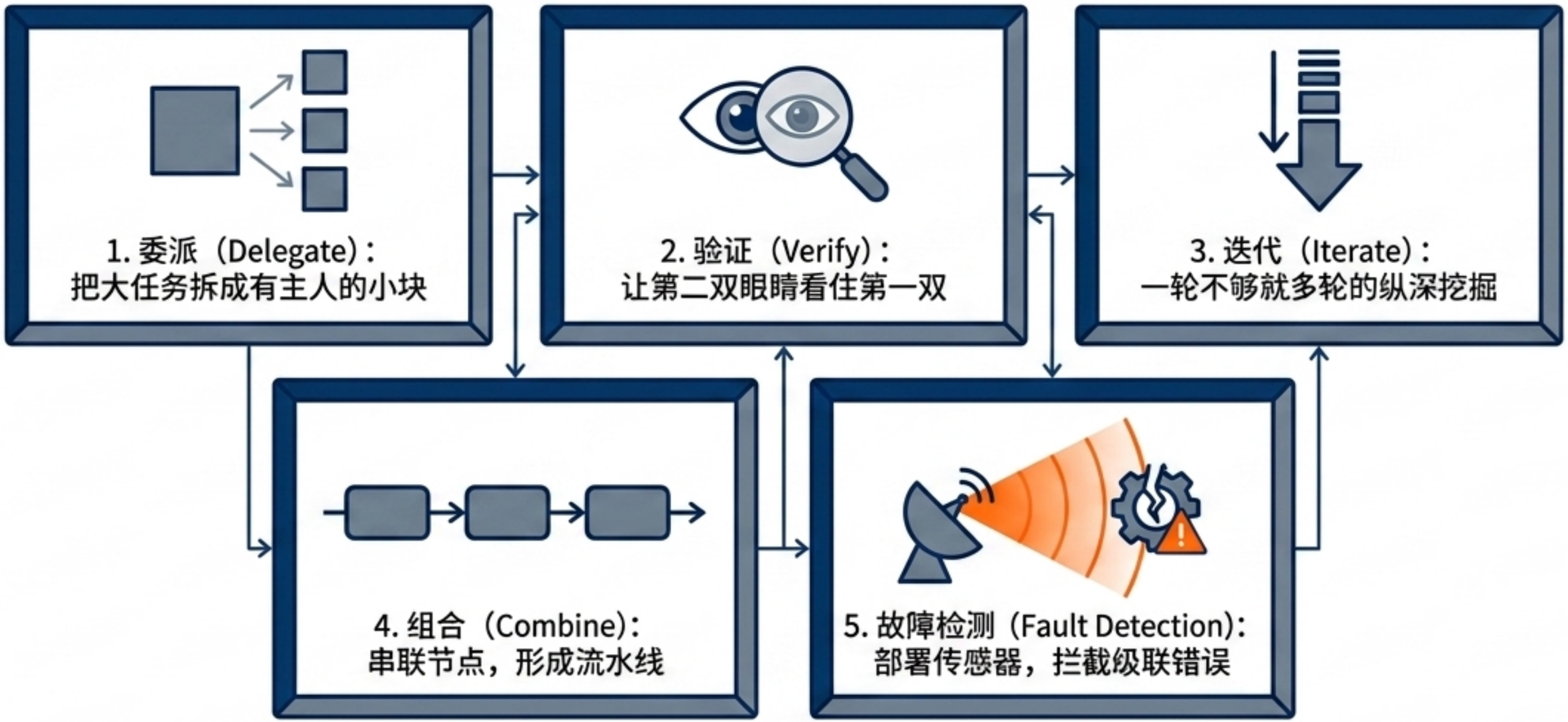
新模式：系统架构师



搭建生产线。定义输入输出，设置质量检查点（QA），处理系统分歧。

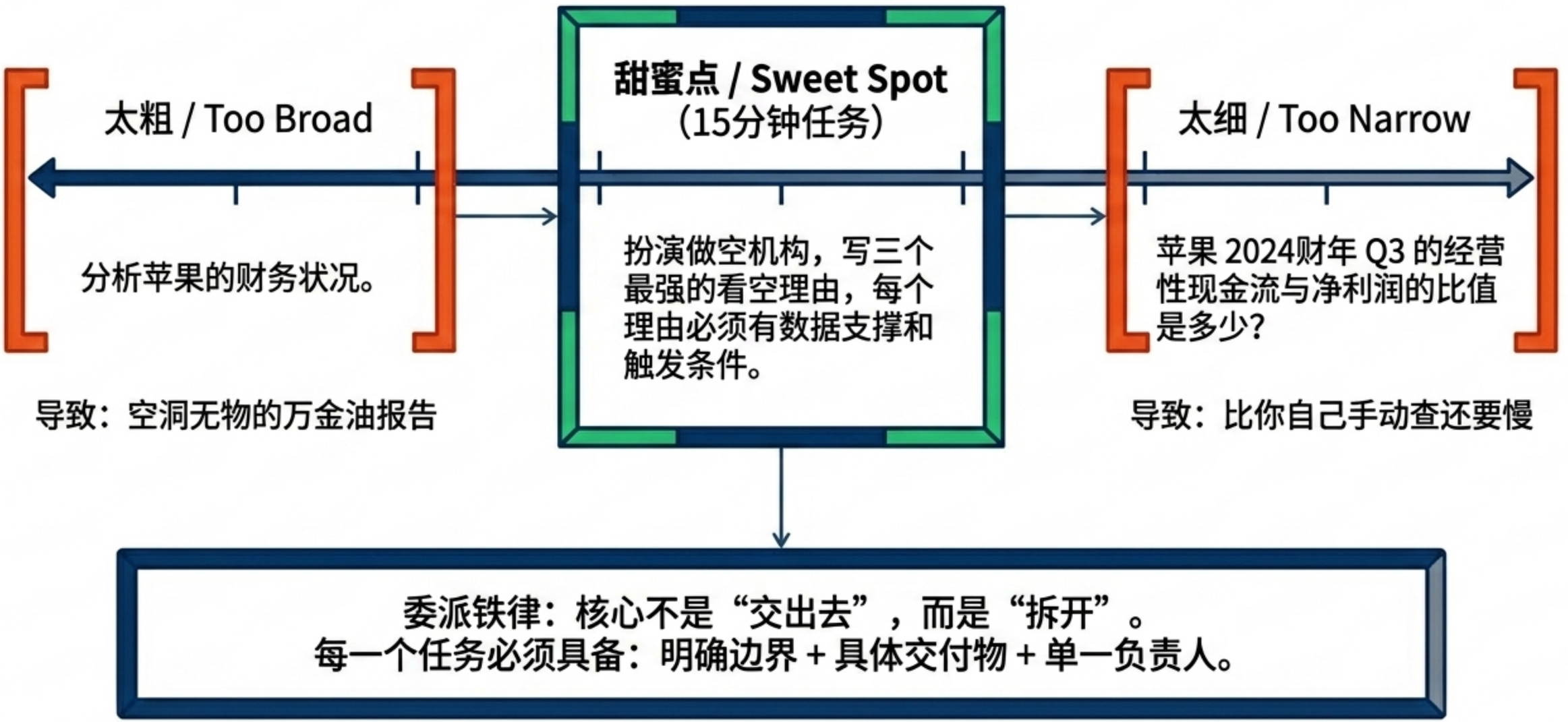
你的核心价值不再是“让 AI 替你思考”，而是“设计让 AI 互相验证的流程”。

覆盖 90% 投资研究的五个核心协作模式

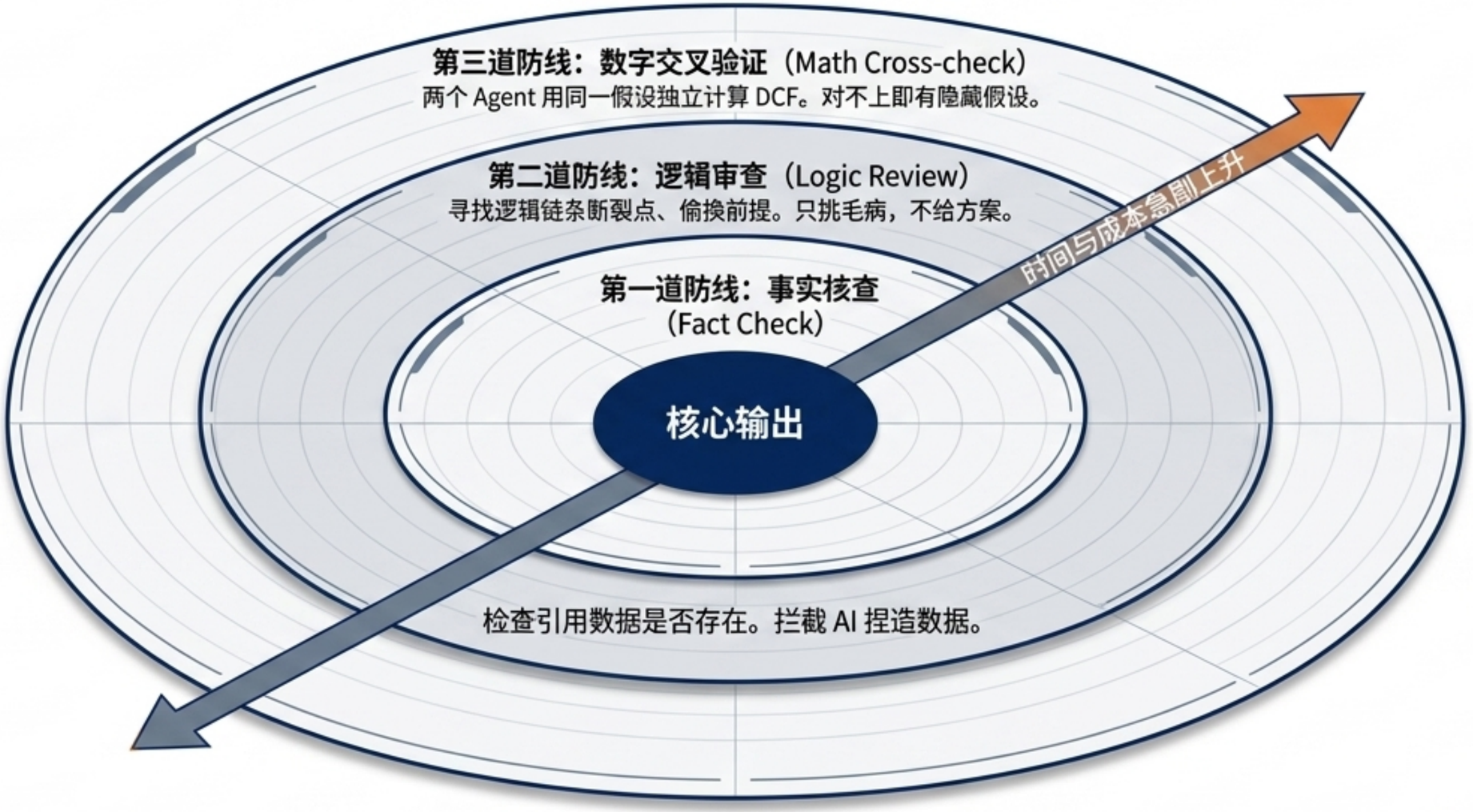


注：这不是一套死板的流程，而是五块可以自由拼装的系统积木。

模式一·委派：寻找 15 分钟的“任务甜蜜点”

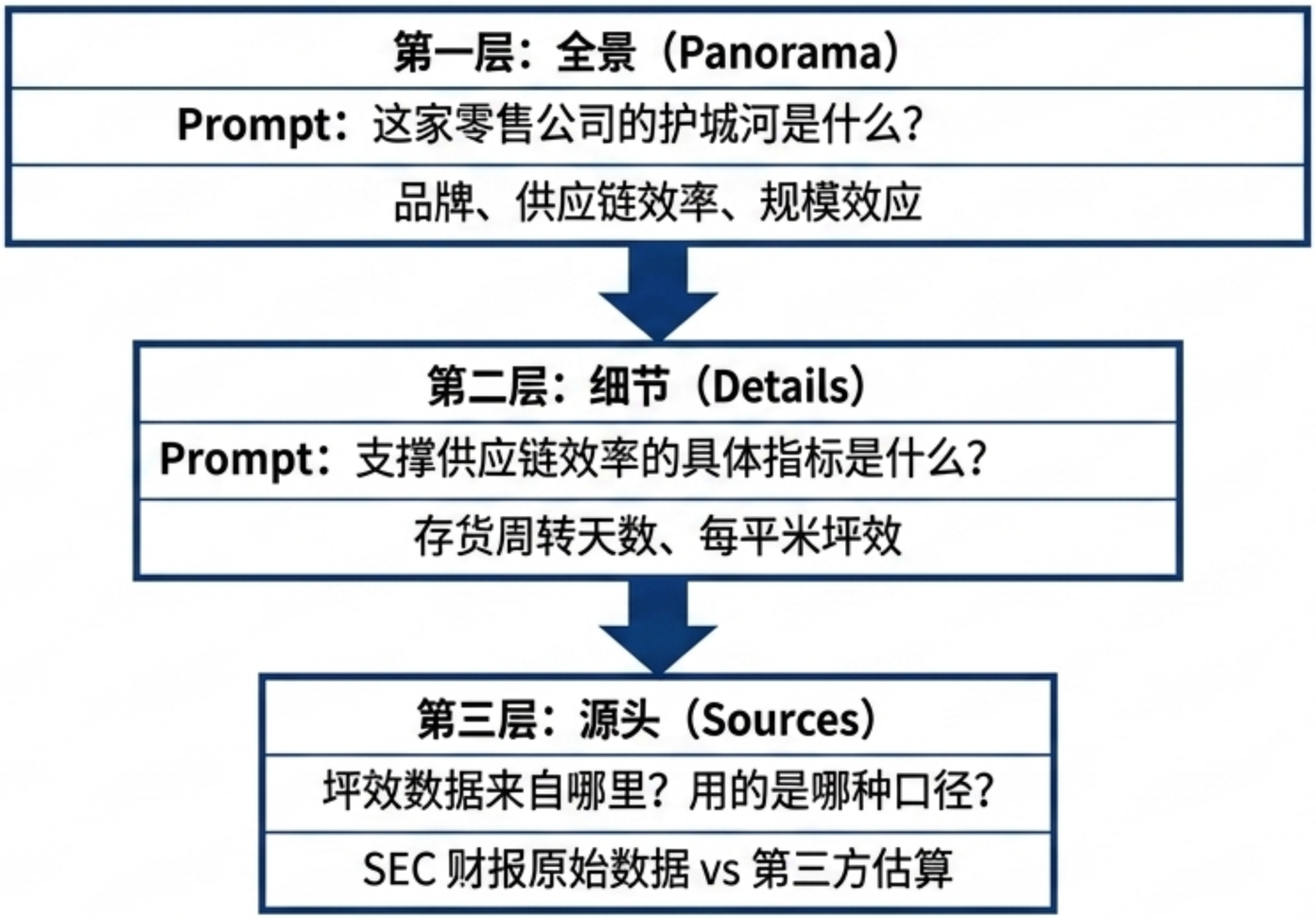


模式二 · 验证：纵深防御体系 （Defense in Depth）



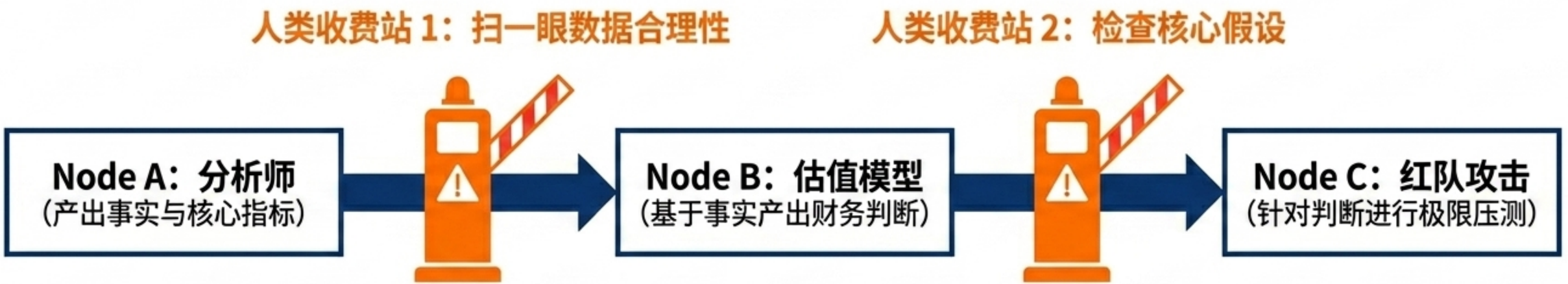
金融现实：三步全跑完，成本是单次查询的 5-8 倍。
把验证当保险：在损失最大的地方买最厚。日常研究仅做事实核查即可。

模式三·迭代：结构化的向下钻取（Drill-Down Funnel）



系统警告：超过三轮还没解决的问题，立刻停止追问！继续迭代只会逼迫 AI 发挥创意编造数据，此时必须切换为人类介入判断。

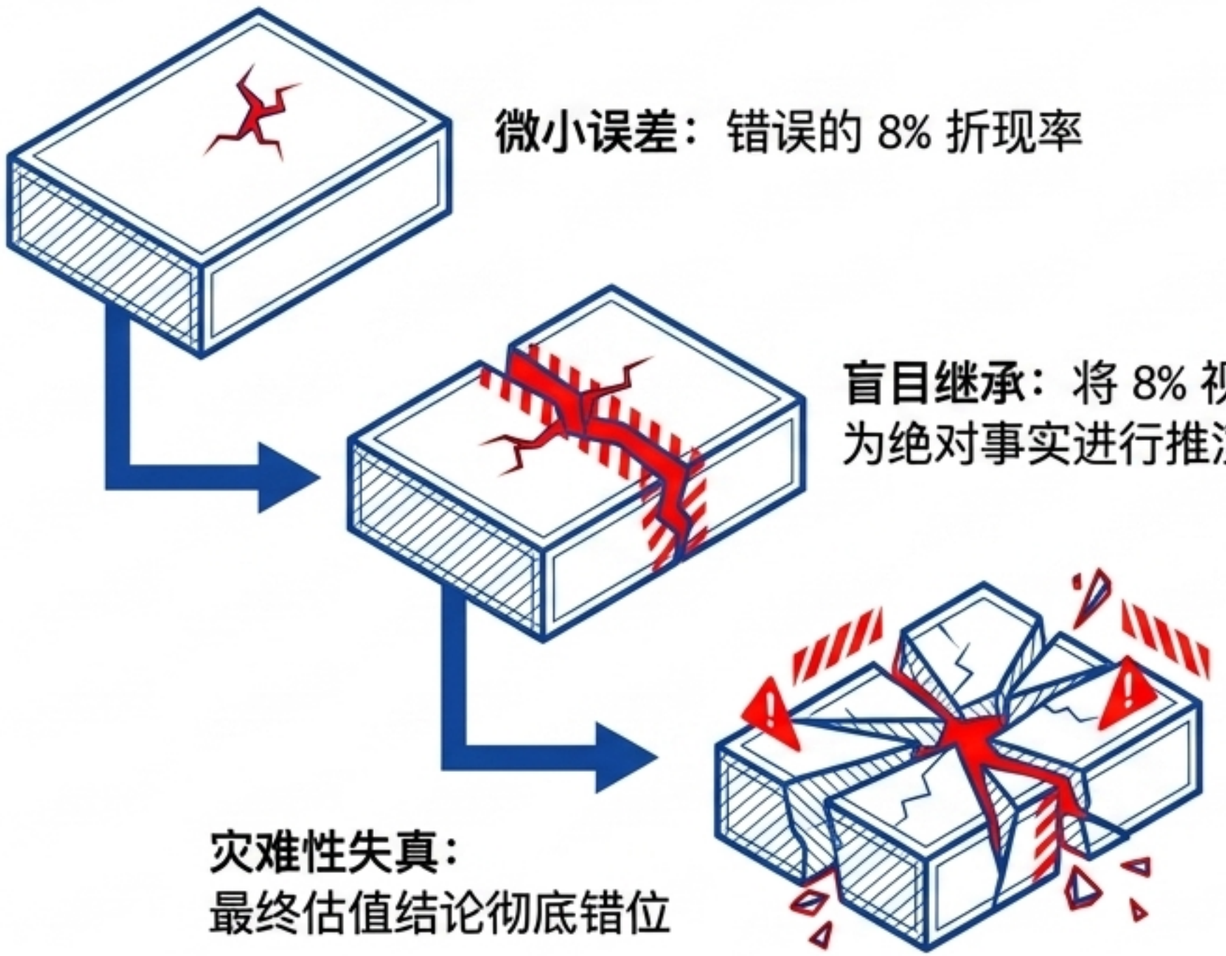
模式四·组合：带有“人类收费站”的流水线



核心认知：全自动流水线是一个“错误放大器”。组合模式的最大风险是太舒服。必须在关键节点插入人工检查，以物理切断级联错误（Cascade Errors）的传播路径。

模式五 · 故障检测：拦截“假设继承”与数据污染

The Assumption Inheritance Cascade



微小误差：错误的 8% 折现率

盲目继承：将 8% 视为绝对事实进行推演

灾难性失真：
最终估值结论彻底错位

Pre-flight Checklist（部署传感器）

Pre-flight Checklist（部署传感器）

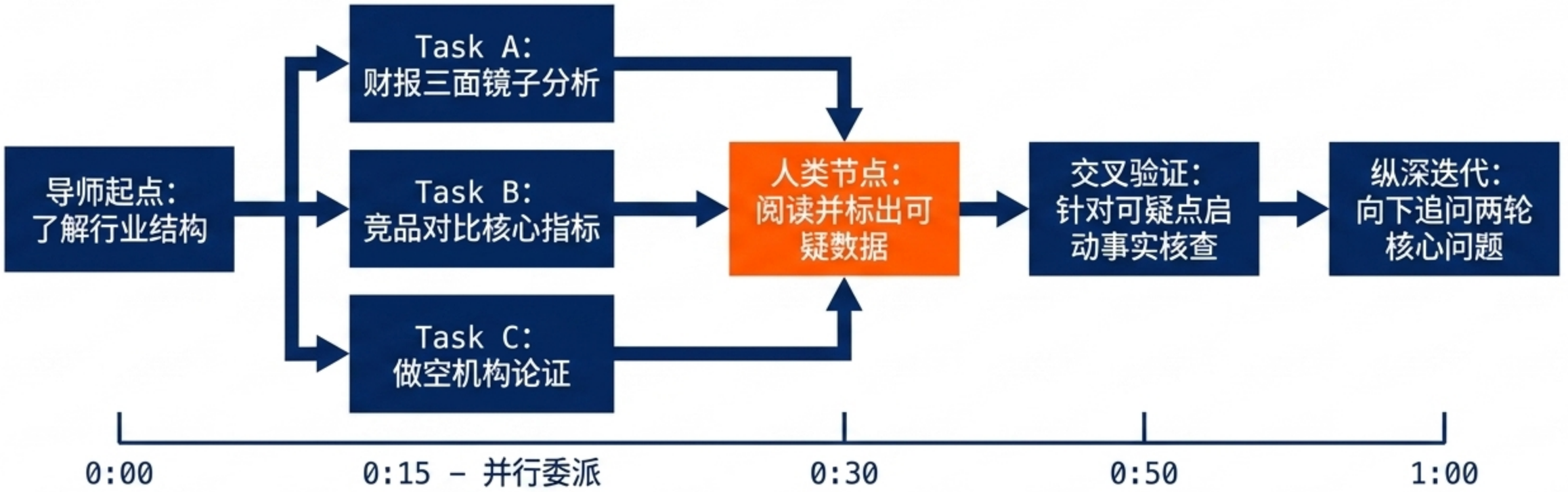
- ☑ “**数据来源故障**”：必须追问关键数字的原始文件出处。
- ☑ “**假设继承故障**”：将最终结论与最初假设对齐压测（“如果这个假设错了，哪个**最致命？**”）。
- ☑ “**共识收敛故障**”：警惕毫无分歧的全票通过，大概率是**数据污染**。

架构选择矩阵：何时单兵？何时建军？

	适用单兵（Solo Agent）	适用舰队（Agent Fleet）
任务特征	一句话可回答的简单标准 概念学习 单一数据查询	多维度问题 涉及复杂假设判断 跨节点交接
容错成本	错了代价极低 （例如：快速粗筛 20 家公司）	错了代价极高 （例如：真实资金买入决策）
验证需求	不需要系统验证 （仅查阅标准概念）	需要严密的事实核查与红队对抗
执行耗时	几秒钟至几分钟	1 - 3 小时

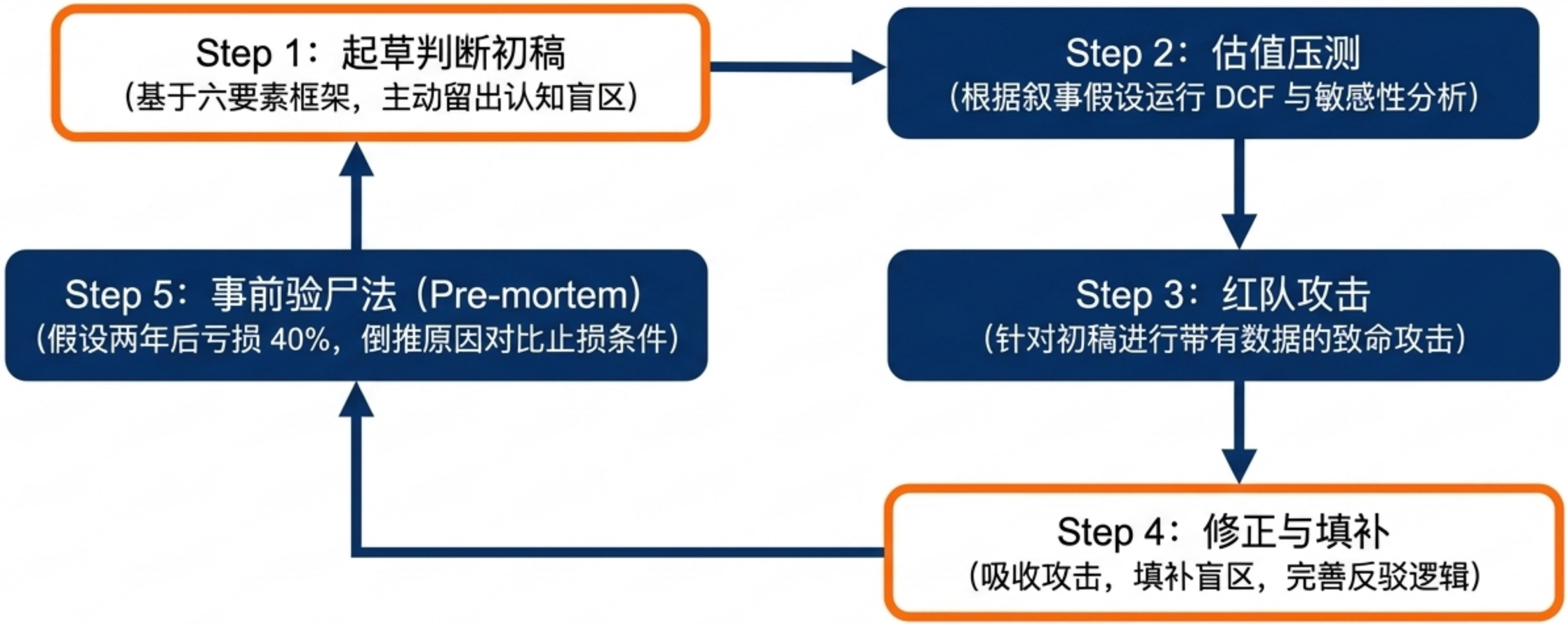
不要用设计复杂流程的快感，代替真正的投资判断。不到一小时的粗筛，不配用舰队。

现成架构一：研究管道 （1–1.5 小时）



最终交付物：一份经过人类审查、剔除虚假数据的公司全景概况。

现成架构二：判断压测流程（2-3 小时）

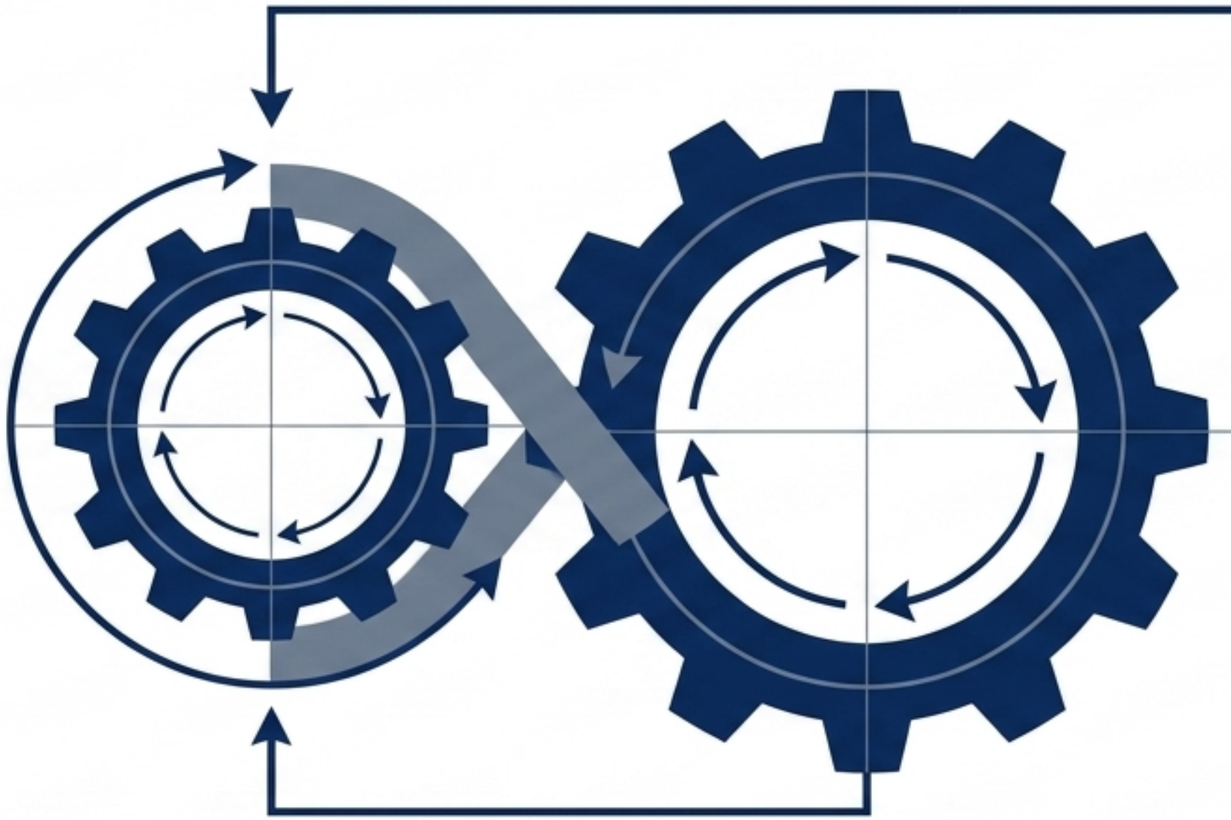


最终交付物：一份带有完整逻辑链与可观察止损条件的最终决策文档。

现成架构三：监控循环（让判断文档“活”起来）

Quarterly Loop (季度监控)

- 对 core hypothesis 标定状态（符合 / 超预期 / 不及预期）
- 检查止损条件是否触发
- 更新文档版本号，记录增减仓理由

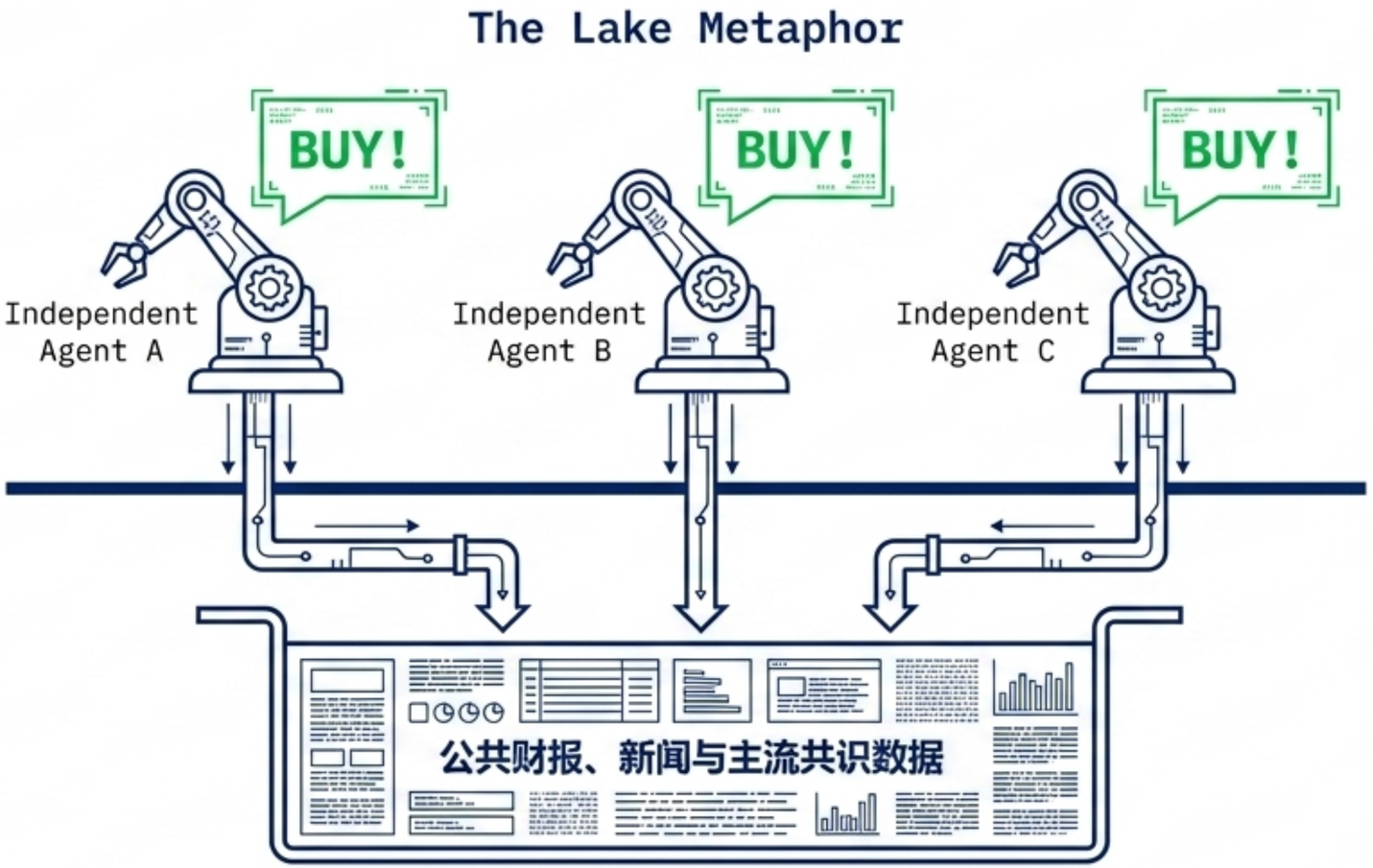


Annual Loop (年度复盘)

- 对比四个季度的轨迹，审视核心假设是否发生漂移
- 评估止损条件的调整：是因为捕捉到真实新信息，还是为了“合理化”亏损？

核心认知：判断文档不是写完就封存的静物，它必须随着新信息持续迭代更新。

虚假的共识：为什么“全票通过”是极其危险的警告信号?



底层逻辑：看似独立的分析其实是“同一片湖里钓出的三条鱼”。全票通过不代表正确概率高，只代表反映了市场共识。在投资里，共识本身没有超额收益，因为它已经被完全定价。

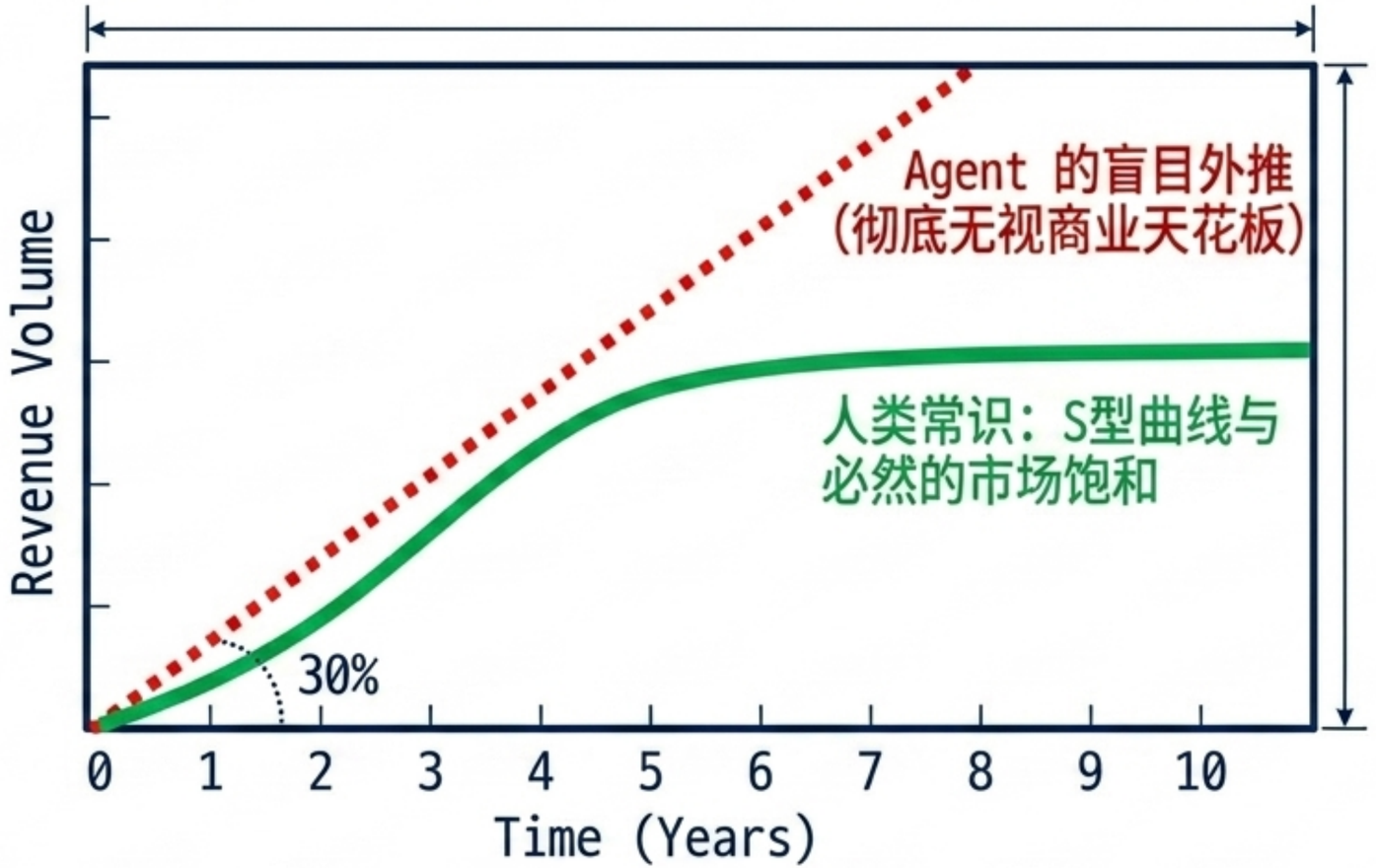
破局之道：人为制造摩擦与系统性分歧

破局法则：更有价值的信号是“部分一致”。在系统发生不一致的裂缝里，往往藏着你漏掉的风险或市场的盲区。

风险登记册：金融 Agent 的四大高频失败模式

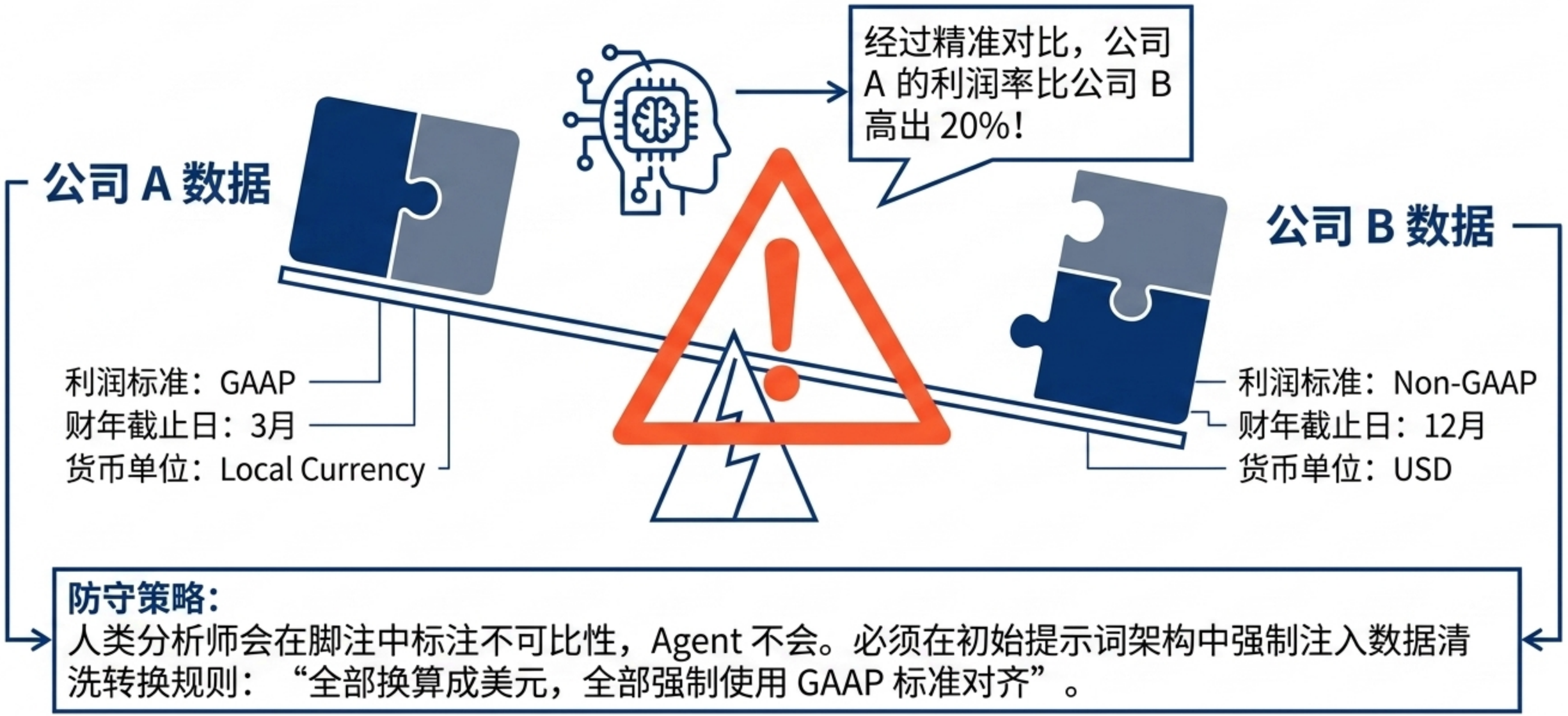
失败模式（Failure Mode）	临床症状（Symptoms）	核心防御手段（Defense）
数据凭空捏造 (Fabrication)	出现精确到小数点后两位的虚假预测数据。	必须追溯影响判断的 3–5 个核心数字的原始来源文件。
共识包装成洞察 (Disguised Consensus)	变换高大上话术，实则完全复述主流券商研报。	拷问系统：随便搜三篇文章结论一样吗？没有增量即刻抛弃。
单位和口径混乱 (Metric Confusion)	无缝且自信地将 GAAP 与 Non-GAAP 放在同一张图表对比。	指令中显式且强硬地要求统一口径（如：全换算成美元/滚动数据）。
增长线性外推 (Linear Extrapolation)	理所当然地假设 30% 的历史高增速能平稳维持十年。	常识反问：按此测算，十年后其收入是否超过全行业总规模？

失败模式深潜：增长线性外推陷阱（Linear Extrapolation）



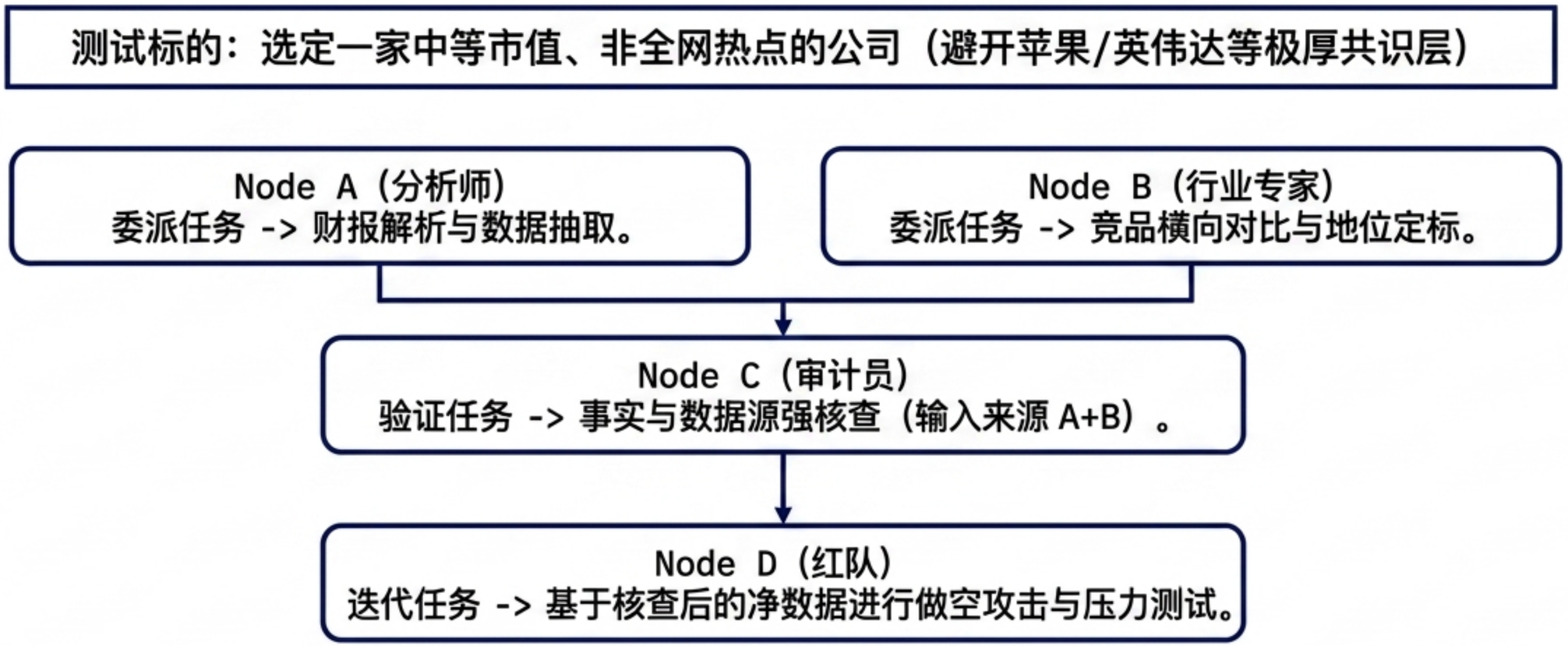
分析： Agent 不会主动质疑历史增速。30% 的增速作为初始起点看似合理，但作为十年的复利假设极其荒谬。人类架构师必须负责向系统引入“物理减速”与“行业天花板”的经济学常识。

失败模式深潜：单位与口径的无声混淆（Metric Confusion）



防守策略：
人类分析师会在脚注中标注不可比性，Agent 不会。必须在初始提示词架构中强制注入数据清洗转换规则：“全部换算成美元，全部强制使用 GAAP 标准对齐”。

Workshop：设计你的第一个四人舰队（v0.1）



实战复盘要求：跑完流程后写下这句话——『这次流程里，我手动做出的最关键判断是 _____。』
这个空白，就是硅基 Agent 永远填补不了的缝隙。

系统的终局：纪律归于硅基，灵感归于碳基

机器的底线（Agent Fleet）

提供系统化的防呆机制、交叉验证矩阵，无情拦截级联错误。

它是【不做蠢事的纪律】。

人类的巅峰（The Architect）

提供违背共识的罕见洞察、设定复杂的假设条件边界、拍板最终的风险承受度。

它是【极少数的好想法】。

“你不需要很多好想法，你需要不多的好想法加上不做蠢事的纪律。” —— Charlie Munger

模式只是手段，你无可替代的判断，才是系统运行的唯一目的。