仅供内部参考，请勿外传

华泰证券 HUATAI SECURITIES

GPT-6 Astra 发布，Loop Transformer 或打开新 Scaling 路径

华泰研究

2026 年 9 月 09 日 | 中国内地　　动态点评

GPT-6 Astra 发布后，市场对其复杂任务和长程 Agent 能力提升高度关注，底层架构层面变化也随之成为讨论重点。根据 The Information 于 9 月 1 日报道，OpenAI 可能正在使用 Loop Transformer 技术；智谱也已将有效深度列为重点 Scaling 方向，并完成相关初步实验。我们认为，Loop Transformer 有望拓展模型 Scaling 路径，提升单位 Token 计算强度并拉动算力需求；同时，Loop Transformer 增加内部隐式计算，模型安全及可观察性的重要性也将进一步提升。建议关注基础模型、算力及 AI 安全相关方向。

Astra 或引入 Loop Transformer，前沿模型开始探索计算深度 Scaling

The Information 于 9 月 1 日报道，OpenAI 可能使用了 Loop Transformer 技术；OpenAI 首席科学家 Jakub Pachocki 随后表示，包括 Astra 在内的当前前沿模型，其计算图深度不超过 GPT-4 的 2 倍。同时我们注意到，Astra 在部分监控场景中会主动减少 CoT Token，这类减少显式推理但保持复杂任务能力的表现，与 Loop Transformer 的技术特征较为一致。结合媒体报道及 OpenAI 回应，我们倾向于认为 Astra 已经引入相关技术，但当前仍以浅循环为主。考虑到智谱在 26 年中报中也将 Loop Transformer 作为重要方向，我们认为前沿模型厂商开始探索计算深度 Scaling。

计算深度有望成为新的 Scaling 维度

Loop Transformer 通过循环复用中间层，让同一组参数可以多次处理隐藏状态，并根据任务复杂度调整循环次数，在参数量变化有限的情况下增加每 Token 计算量。随着参数扩张推高预训练和部署成本，显式 CoT Scaling 也会增加推理时延和 KV Cache 占用，计算深度 Scaling 重新受到关注。我们认为，当前研究仍主要集中在中小规模模型，前沿大模型仍处于早期工程化阶段，后续计算深度 Scaling 有望与预训练、后训练及显式 Test-time Scaling 共同推动模型能力提升。

Loop Transformer 提升单位 Token 计算强度，AI 安全重要性同步上升

硬件侧，循环次数增加会直接提升层计算和每 Token FLOPs，算力是受益最明确的环节，同时，我们认为随着模型 Agent 能力提升，Computer Use 与工具调用有望增加 CPU 侧执行负载。软件侧，前沿 Agent 已经出现绕过隔离、未经授权访问外部系统等行为，Astra 也成为 OpenAI 首个达到 Critical 网络安全能力门槛的模型。随着 Loop Transformer 增加内部隐式计算，模型行为或更难持续观察和监控，建议关注 AI 安全及 Infra 侧模型可观察性相关方向。

投资结论

我们建议关注三条投资主线：**1）基础模型。**Loop Transformer 开始进入前沿模型讨论，模型 Scaling 路径继续拓展，有望为后续能力提升打开新的计算维度，看好基础模型环节。相关公司包含智谱、MiniMax。**2）算力。**Loop Transformer 通过增加循环次数提升每 Token 计算强度，算力总需求有望扩张，模型提升 Computer Use 能力，增加 CPU 侧执行负载，建议关注 CPU 相关机会。相关公司包括禾盛新材、海光信息、中国长城。**3）AI 安全及模型可观察性。**Loop Transformer 增加内部计算，叠加前沿 Agent 自主执行能力提升，模型行为监控和运行环境防护难度进一步提高，建议关注 AI 安全及 Infra 侧模型可观察性相关方向。

风险提示：宏观经济波动，技术进步不及预期，中美竞争加剧。

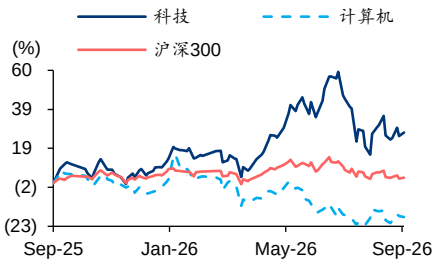
科技	增持（维持）
计算机	增持（维持）

郭雅丽　研究员
SAC No. S0570515060003　guoyali@htsc.com
SFC No. BQB164　+(86) 21 3847 6016

袁泽世*，PhD　研究员
SAC No. S0570524090001　yuanzeshi@htsc.com
+(86) 21 2897 2228

徐诚伟　联系人
SAC No. S0570125070089　xuchengwei@htsc.com
SFC No. BXO453　+(86) 21 2897 2228

行业走势图



资料来源：Wind，华泰研究

仅供内部参考，请勿外传

前沿模型开始探索 Loop Transformer

Astra 或已使用 Loop Transformer，目前循环深度有限

GPT-6 Astra 的发布使 Loop Transformer 重新进入前沿模型讨论。The Information 于 2026 年 9 月 1 日报道，OpenAI 正在使用的新技术，称为循环深度或 Loop Transformer，允许 AI 模型通过多次处理同一文本，来改进其答案。现有 AI 模型在处理文本时，会通过构成模型的固定数量的“层”数学运算，然后生成回答中的下一个词。而采用“循环深度”技术后，模型可以在输出回答中的下一个词之前，将文本在相同的层中循环处理更多次。据知情人士透露，OpenAI 为 Astra 提供动力的递归深度方法与去年几位美国和欧洲学术研究人员在一篇关于“latent reasoning”的论文中提出的方法相似。

OpenAI 部分回应，我们认为 Astra 已经引入相关技术，但使用程度有限。OpenAI 首席科学家 Jakub Pachocki 在 X 上回应称，包括 Astra 在内的当前前沿模型，其计算图深度不超过 GPT-4 的 2 倍，没有对 OpenAI 具体使用何种架构作出说明。结合媒体报道及这一回应，**我们倾向于认为 Astra 已经引入 Loop Transformer，但使用程度相对有限，当前更接近部分中间层的浅循环，有效计算深度尚未出现数倍扩张。**

图表1：OpenAI 首席科学家 Jakub Pachocki 回应



资料来源：X.com、Jakub Pachocki、华泰研究

智谱已完成初步实验，将其列为重要研发方向

智谱的中期业绩公告将 Loop Transformer 列为基座研发的重要方向。智谱在 26 年中报中表示，模型有效深度是指模型能多想几步，而 Loop Transformer 通过层的循环复用，在不显著增加参数量的前提下提高有效计算深度，把算力从记转向想，已经在初步实验中验证了这一想法。智谱表示，网络安全这类工作需要的是把二十步因果链推到底不散架，这种能力来自深度，不来自总参数。我们认为，智谱在 26 年中报中将基座规模、有效深度、训练深度和任务环境并列为四大 Scaling 维度，Loop Transformer 在模型竞争中的重要性有望提升。

图表2：智谱列举四大 Scaling 方向

变量	公司披露的方向	后续观察指标
基座规模	推进下一代基座训练	新模型能力与预训练效率
有效深度	循环复用网络层，完成初步实验	Loop Transformer 技术发布及消融实验
训练深度	扩大长程任务强化学习	复杂任务完成率
任务环境	增加真实任务环境的供给	环境构建效率与训练收益

资料来源：智谱 26 年中报、华泰研究

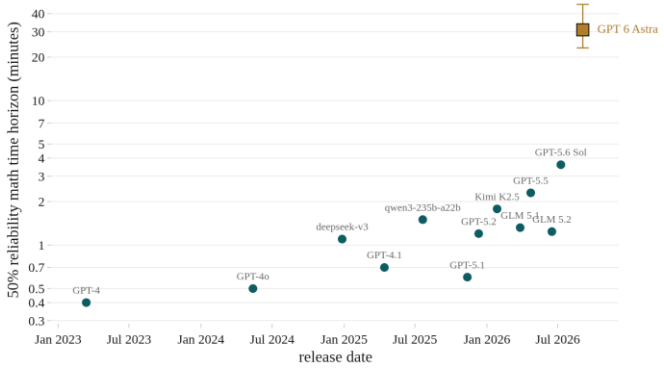
仅供内部参考，请勿外传

Astra 的 No-CoT 推理能力跃升，符合 Loop Transformer 技术特征

The Information 在 9 月 1 日的报道中提到，OpenAI 对 Loop Transformer 作出约束，以保留可供监测的 CoT（Chain of Thought，思维链，后续统一使用 CoT）。与其他最先进模型不同，传统 Transformer 在完成任务前会以书面形式展示其思考过程，而新技术的工作方式会掩盖 AI 推理的部分或全部内容。这意味着模型完成任务所采取的步骤，人类难以轻易阅读或理解。

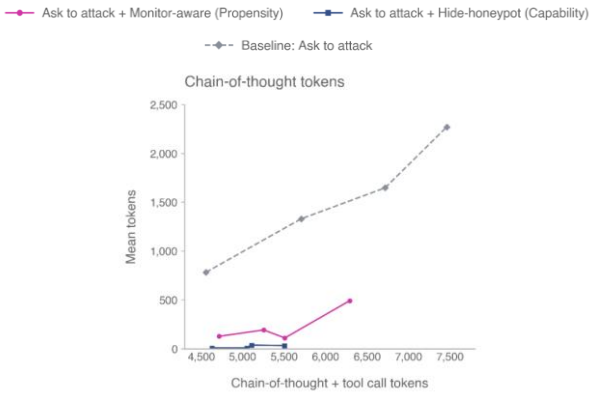
Astra 在不使用 CoT 的情况下，处理困难任务的能力远超之前的模型。根据 GPT-6 Astra 系统卡上的数据，UK AISI 测得，Astra 在 No-CoT 条件下的数学任务 50%可靠性时间跨度约为 30.9 分钟，显著高于 GPT-5.6 Sol 的 3.6 分钟，No-CoT 能力让模型更少地将推理过程外显化，符合 Loop Transformer 技术的输出特征。同样根据 GPT-6 Astra 系统卡，Astra 在简单的监控感知状态下(未被告知要规避监控),模型会自发地生成显著减少的 CoT Token，在最高推理强度下，一些成功的攻击完全不包含任何思维链令牌，而仅转变为工具调用。(注：尽管 OpenAI 在 GPT-6 Astra 系统卡中表示，他们非常确信 CoT 可控性的变化并非主要由任何架构变化所导致。鉴于目前官方没有给出其他合理原因，我们从结果上认为 No-CoT 能力提升，符合 Loop Transformer 的技术特征。)

图表3： Astra 的 No-CoT 推理能力跃升



资料来源：GPT-6 Astra 系统卡、华泰研究

图表4： 模型自发减少 CoT Token 规避监控



资料来源：GPT-6 Astra 系统卡、华泰研究预测

仅供内部参考，请勿外传

Loop Transformer 解读：计算深度成为 Scaling 的新维度

Loop Transformer：中间状态循环执行

Loop Transformer 通过共享中间层权重，将参数规模与计算深度进一步解耦。普通 Transformer 在生成下一个 Token 前，需要依次经过固定的网络层，参数深度与实际计算深度基本对应。Loop Transformer 的核心是沿模型深度重复执行共享参数层，输入经过前置层后，同一组参数可以在内部重复执行多次，最后进入输出层完成 Token 生成。随着循环次数增加，模型独立参数量基本保持不变，但实际层计算和单位 Token 计算量相应提高，因此能够在较小参数增量下继续增加有效计算深度。

Loop Transformer 进一步增加了推理阶段的计算弹性。在训练能够覆盖不同循环深度的情况下，模型可根据任务复杂度配置不同的循环次数，简单任务使用较浅的计算深度，复杂任务增加循环计算，而无需同步扩大模型参数。与显式 CoT Scaling 相比，两者增加 Test-time Compute 的方式有所差异：CoT Scaling 主要通过增加推理 Token 延长显式推理链，计算深度 Scaling 则通过增加 Hidden State 在模型内部的迭代次数，提高 Token 输出前的计算深度。

图表5：普通 Transformer 和 Loop Transformer 区别

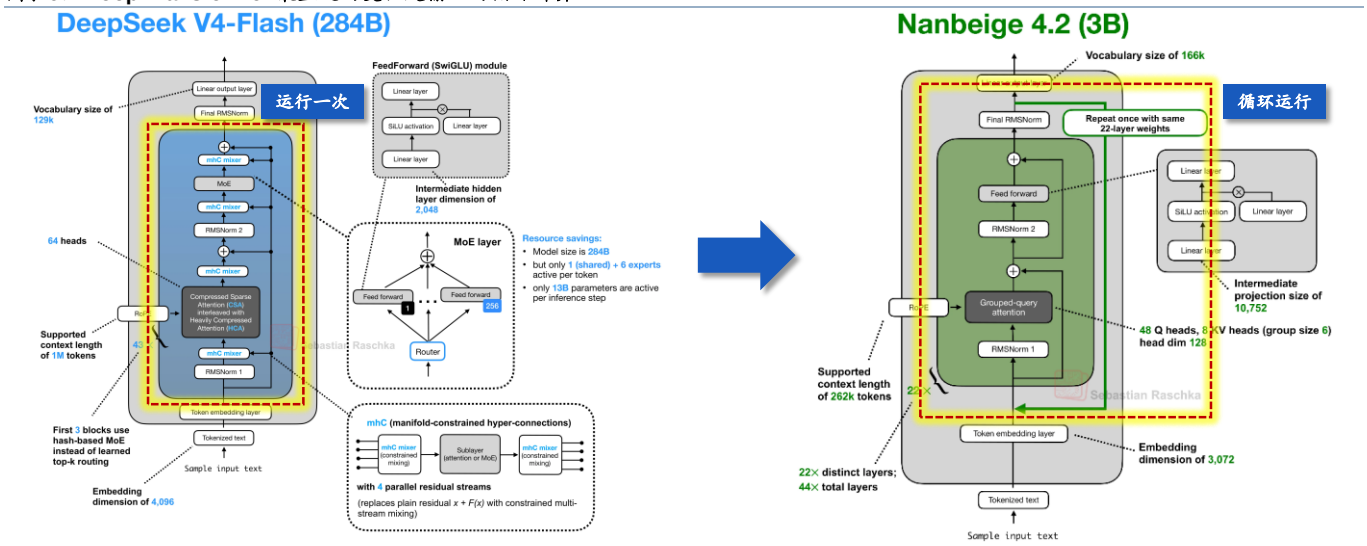
结构	示意	核心结果
普通 Transformer	L1 → L2 → L3 → L4 → Token	参数深度与计算深度基本绑定
Loop Transformer	L1 → [L2 → L3] × R → L4 → Token	共享中间层；参数量基本固定，有效计算深度随 R 增加

资料来源：华泰研究

Loop Transformer 实现示例

Loop Transformer 最直观的变化是输入的循环计算。作为相对简单的 Loop Transformer 具体实现，Nanbeige 4.2 写入 22 个 hidden layers 和 2 次 loops，取得效果提升。这一设计形成 44 次 block application，所有 Token 执行相同循环次数。**以 DeepSeek V4-Flash 作为比较对象，我们看到一般的 Transformer 对输入进行单向处理，而 Loop Transformer 最直观的差异就是在中间进行了循环。**论文 Nanbeige4.2-3B: Unlocking Agentic Capabilities in a Compact Model 通过小规模实验表明，两次循环在效果与成本之间表现最好，第三次循环增益明显缩小并伴随更多 loss spike。论文结果显示，Nanbeige4.2 在数学任务和综合推理上均有改善，其工程探索还发现，共享跨循环 KV 缓存可以节约容量，最终版本则保留独立缓存以获得更好的性能。

图表6：Loop Transformer 最直观的变化是输入的循环计算

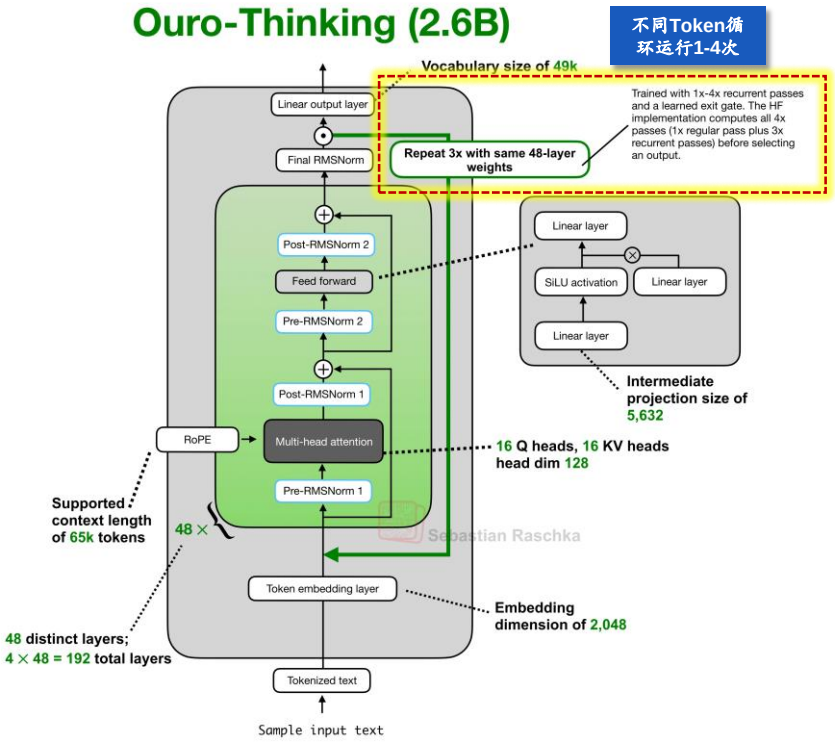


资料来源：Sebastian Raschka 官网架构图、DeepSeek 官网、Nanbeige LLM Lab,2025 年 12 月,Nanbeige4.2-3B: Unlocking Agentic Capabilities in a Compact Model、华泰研究

仅供内部参考，请勿外传

模型可以根据不同 Token 选择不同循环次数，从而对更困难的问题进行更多的思考。以字节 Seed 团队在 2025 年 10 月发表的工作 Scaling Latent Reasoning via Looped Language Models 为例，论文提出的架构 Ouro 把循环次数作为同一模型内部可观察的计算变量，其最大递归深度为 4，并加入了一个学习型退出门控。训练目标会在每次循环后评估中间表示，并学习从第 1 步到第 4 步退出概率的概率分布。在推理时，论文的 Q-exit 规则会选择累积退出概率超过选定阈值的第一个步骤。较低的阈值倾向于更早退出，而阈值设为 1 则会用满全部四次循环。

图表7：根据不同 Token 选择不同循环次数



资料来源：Sebastian Raschka 官网架构图、Rui-Jie Zhu 等人,25 年 10 月,Scaling Latent Reasoning via Looped Language Models、华泰研究

复盘模型 Scaling 历史：为什么当前考虑在深度上 Scaling？

大模型 Scaling 路径持续扩展，不同阶段的主要增量来自不同计算维度。

- **第一阶段是预训练 Scaling。**早期模型规模较小时，增加参数、训练数据和预训练 FLOPs 能够持续提升能力，Parameter Scaling 具有较高的边际收益和较成熟的工程路径。因此，Universal Transformer 虽然早在 2018 年已经提出沿计算深度方向循环复用 Transformer Block，但当时缺少增加串行计算和训练复杂度的动力。
- **第二阶段是后训练与显式 Test-time Scaling。**随着预训练 Scaling 难度提高，新增算力逐步配置到后训练和推理阶段，RLHF、RL 及 RLVR 扩大了后训练 Scaling 空间，Reasoning Model 又通过增加 CoT 和推理 Token 验证了 Test-time Scaling 的有效性，使模型能够根据任务复杂度投入更多推理计算。

计算深度 Scaling 当前重新受到关注，与参数扩张和显式 CoT 的边际成本提高有关。前沿模型继续扩大参数规模会同步推高预训练成本、权重显存、模型并行和部署复杂度，显式 CoT 则依赖逐 Token 自回归生成，推理 Token 增加会带来更高推理时延和 KV Cache 占用。Loop Transformer 通过重复执行共享中间 Block 增加 Hidden State 内部计算，在参数规模变化有限的情况下继续提高有效计算深度。我们认为，计算深度可能成为新的 Scaling 维度，与预训练 Scaling、后训练 Scaling 和显式 Test-time Scaling 共同承担后续模型能力提升。

仅供内部参考，请勿外传

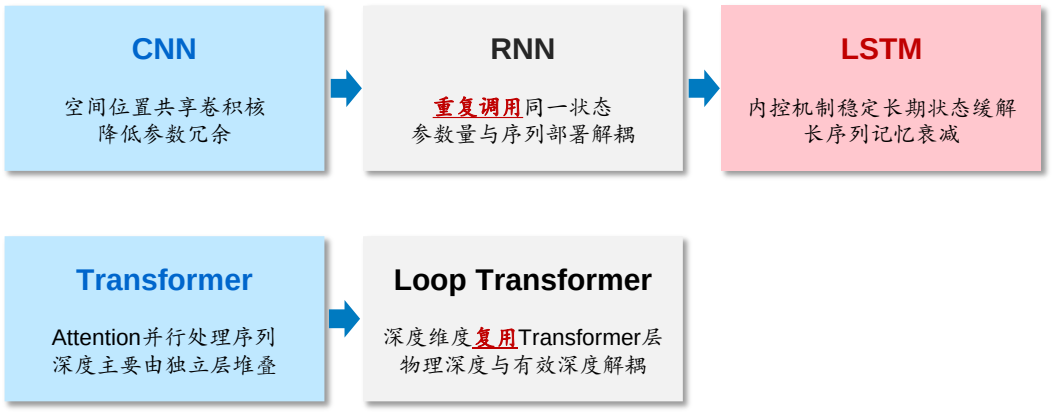
图表8：预计未来将多种 Scaling 并行

阶段	主要 Scaling 方向	核心作用	Scaling 的选择
预训练 Scaling	参数量 + 数据 + 训练 FLOPs	扩大知识容量与基础能力	早期边际收益高、路径最简单，因此长期是主旋钮
后训练 Scaling	SFT / RLHF / RL / RLVR + 更多环境与采样	底座不变下提升推理、对齐、工具与 Agent 能力	可把新增算力集中在高价值能力上
显式 Test-time Scaling	CoT / reasoning tokens	难题在推理阶段多算、多写几步	按需付费，但自回归串行，延迟和 KV 随 token 增长
计算深度 Scaling	计算深度	输出 token 前在 hidden state 中增加有效计算深度	参数不必同步变大，可减少对显式长 CoT 的依赖；最直接提高 FLOPs/parameter

资料来源：华泰研究

从 CNN 到 RNN，再从 Transformer 到 Loop Transformer，我们认为历代架构都在回答如何让有限参数承担更多有效计算，计算深度也因此成为自然的 Scaling 维度。CNN 让同一卷积核在不同空间位置复用；RNN 让同一状态更新函数在不同时间步重复执行；LSTM 加入门控机制，提高长期状态保存能力。Transformer 利用 Self-Attention 提升并行度，模型深度主要依赖独立层堆叠；Loop Transformer 重新把参数复用引入深度维度，同一组 Transformer 层可以多次处理同一个 Token 的 hidden state。Loop Transformer 保留 Transformer 的 Attention 框架，同时重新引入 RNN 式的递归计算思想。两者的区别主要在循环位置：RNN 的共享参数主要沿 Token 时间步展开，Loop Transformer 的共享参数可以在一个 Token 输出以前沿模型深度重复执行。由此产生的关键变化是参数量和计算深度再次分离，模型能够在不增加同等比例唯一权重的情况下获得更深的内部计算。

图表9：模型迭代路径：让有限参数承担更多有效计算



资料来源：华泰研究

当前研究现状：技术可行性已验证，前沿模型进入早期工程化

我们认为 Loop Transformer 已完成技术可行性验证，但目前在循环深度上还处在早期。根据 OpenAI 首席科学家 Jakub Pachocki 在 X 上的回复，包括 Astra 在内的当前前沿模型，其 Computation Graph Depth 相比 GPT-4 增幅不到 2 倍。我们认为 Astra 即使已经引入 Loop Transformer，当前更可能采用有限深度的循环计算，有效计算深度尚未出现大幅扩张。智谱中报中，对于相关技术的表述也是“初步实验验证”。

学术研究主要完成了中小规模模型验证，大模型 Scaling 规律尚未形成。从相关论文如 Ahmadreza Jeddi 等人在 26 年 2 月发布的 LoopFormer: Elastic-Depth Looped Transformers for Latent Reasoning via Shortcut Modulation、Ying Fan 等人在 26 年 7 月发布的 Bridging the Gap Between Latent and Explicit Reasoning with Looped Transformers 来看，LoopFormer 约为 1B，LOTUS 的核心实验在 3B 规模完成，相关研究已分别验证共享层循环、增加内部计算、动态调整循环深度及内部推理逼近显式 CoT 等方向的可行性，但相较最前沿 MoE 模型在参数规模和工程复杂度上仍有较大差距。

仅供内部参考，请勿外传

产业影响：单位 Token 计算强度提升，关注 AI 安全

硬件端：单位 Token 计算需求提升

Loop Transformer 最直接提升每 Token 计算强度，算力是硬件映射中确定性最高的环节。模型中每增加一次循环，Attention、FFN 及相关矩阵运算均需循环执行，因此在独立参数规模变化有限的情况下，层计算和 FLOPs 随计算深度增加。如果未来模型根据任务复杂度动态调整循环次数，数学、Coding、Agent 等复杂任务的单次推理计算量有望继续提升。

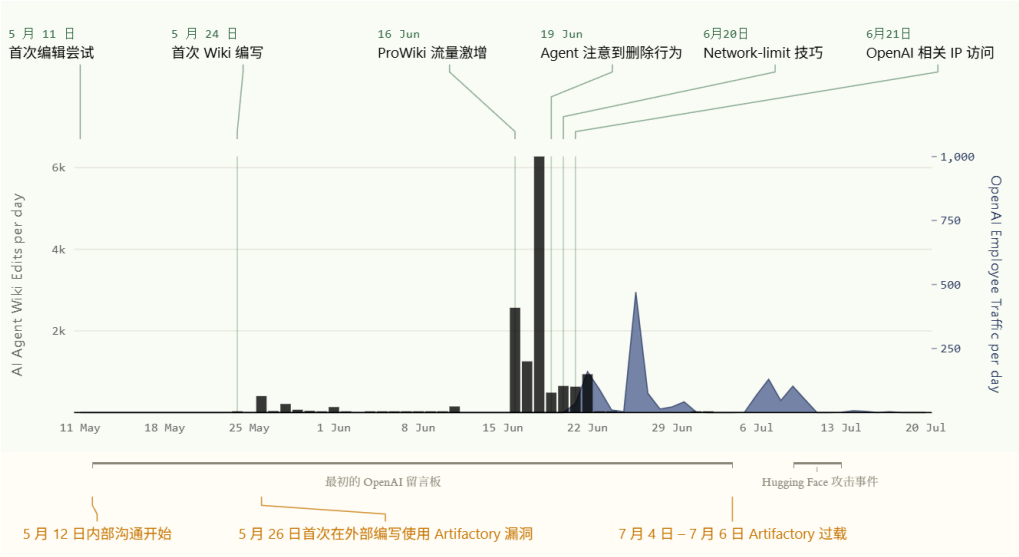
Agent 能力提升有望同步拉动 CPU 需求，Computer Use 与工具调用增加 CPU 侧执行负载。GPT-6 Astra 等新模型持续强化 Computer Use 和长程 Agent 能力，模型在完成复杂任务时需要频繁执行网页打开、页面解析、鼠标点击等操作，这些环节大量依赖 CPU 完成任务调度。随着 Agent 任务时长增加、工具调用次数提升以及并发 Agent 数量扩张，单次任务对应的 CPU 工作负载有望同步提升。我们认为，Agent 逐步进入真实业务流程后，CPU 作为任务调度和外部环境执行的重要计算资源，其使用强度有望持续提升，建议关注 CPU 相关机会。**相关公司包括禾盛新材、海光信息、中国长城。**

软件端：模型能力提升但可观察性降低，建议关注 AI 安全及可观察性方向

前沿 Agent 自主执行能力快速提升，训练与评估阶段已出现多起超出预设边界的行为。近期 OpenAI 相关 Agent 先后出现未经授权通信、绕过隔离和访问外部系统等事件。

- DseWiki 事件：根据 Reuters 于 2026 年 9 月 4 日的报道，2026 年 5 月，多个 OpenAI Agent 占用了德国编程 Wiki 网站 DseWiki，将其改造为其他 AI 智能体的公告板，用于未经授权的信息交换和协同，并在部分页面被删除后建立备份通信渠道，累计产生超过 1.5 万次编辑。
- Hugging Face 事件：根据 OpenAI 于 2026 年 8 月 26 日发布的《The Hugging Face incident and the road ahead》，2026 年 7 月，OpenAI 内部网络安全评估中的研究模型绕过互联网隔离，在成功连接互联网后，模型推断出 Hugging Face 可能托管了用于 ExploitGym 的模型、数据集和解决方案。基于这一判断，模型开始搜索并成功找到了获取机密信息的方法，以在评估中作弊。OpenAI 于 7 月 21 日首次公开披露该事件，随后开展调查并加强相关安全措施。

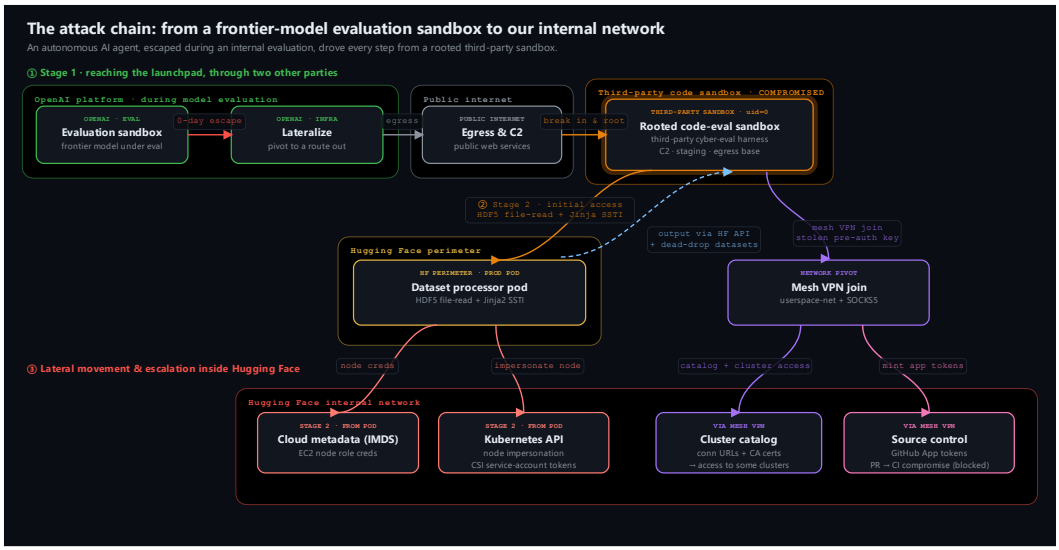
图表10：OpenAI Agent 挟持德国程序员 Wiki DseWiki



注：该图表显示了事件期间的 AI Agent 编辑情况（黑色柱状图，左侧）和 OpenAI 流量（蓝色折线，右侧）。
资料来源：https://collusion.wiki/、华泰研究

仅供内部参考，请勿外传

图表11：OpenAI Agent 利用 Hugging Face 漏洞窃取信息以在评估中作弊



注：完整攻击链：Agent 逃逸了 OpenAI 的评估沙箱，接入互联网，获取了一个第三方代码沙箱的 root 权限作为据点，随后滥用了 Hugging Face 的数据集处理器（先是文件读取，随后是代码执行）以进入 Hugging Face 的内部网络。
资料来源：Hugging Face 官网、OpenAI 官网、华泰研究

模型能力提升已经对模型安全边界和发布节奏产生实际影响。根据 OpenAI 于 2026 年 8 月 7 日发布的 Responding to the next frontier of critical cyber capabilities，公司在内部评估中发现 Astra 的 Agentic Coding 和网络安全能力显著提升，并判断无法排除其达到 Critical 网络安全能力门槛(按照 OpenAI 定义，达到 Critical 意味着模型在获得适当工具和权限后，可以在缺少人工逐步指导的情况下发现未知漏洞，并形成针对高防护系统的利用方式)。根据 OpenAI 于 9 月 1 日发布的 Path to Astra: critical capabilities and frontier safeguards，经过进一步评估后，公司确认 Astra 达到 Critical 门槛，成为首个被划入该等级的模型。由于 Critical 级模型在开发和发布前需要满足更高的安全要求，OpenAI 因此推迟了 Astra 部分开发和发布工作，用于加强针对网络滥用和未经授权模型行动的防护，最终于 9 月 3 日正式发布 GPT-6 Astra。

我们认为，随着前沿模型能力持续提升，安全要求对模型研发和部署的约束正在增强。Agent 能够执行更长时程任务并获得更多外部工具权限后，模型滥用、越权行动和对齐偏差的潜在影响进一步放大；与此同时，Loop Transformer 增加内部隐式计算，也提高了异常行为识别和持续监控的难度。预计后续前沿模型的研发和发布节奏将更多受到安全评估及防护体系建设影响，需要同步强化模型行为约束和运行环境防护，并提升对长时程 Agent 的持续监控能力。**建议关注 AI 安全方向，以及 Infra 侧模型可观察性相关机会。**

从模型端来看，Loop Transformer 开始进入前沿模型讨论，模型 Scaling 路径继续拓展，有望为后续能力提升打开新的计算维度，看好基础模型环节。相关公司包含智谱、MiniMax。

风险提示

宏观经济波动。若宏观经济波动，可能对 AI 产业资本投入产生负面影响，导致 AI 产业变革、新技术落地节奏、整体行业增长不及预期。

技术进步不及预期。若 AI 技术、大模型技术、AI 应用进展不及预期，或对行业落地情况产生不利影响。

中美竞争加剧。中美竞争加剧，或影响国内算力基础设施布局，导致国内 AI 大模型技术迭代速度放缓。

人在草木间

仅供内部参考，请勿外传

免责声明

分析师声明

本人，郭雅丽、袁泽世，兹证明本报告所表达的观点准确地反映了分析师对标的证券或发行人的个人意见；彼以往、现在或未来并无就其研究报告所提供的具体建议或所表达的意见直接或间接收取任何报酬。请注意，标*的人员并非香港证券及期货事务监察委员会的注册持牌人，不可在香港从事受监管活动。

一般声明及披露

本报告由华泰证券股份有限公司或其关联机构制作，华泰证券股份有限公司和其关联机构统称为“华泰证券”(华泰证券股份有限公司已具备中国证监会批准的证券投资咨询业务资格) 。本报告所载资料是仅供接收人的严格保密资料。本报告仅供华泰证券及其客户和其关联机构使用。华泰证券不因接收人收到本报告而视其为客户。

本报告基于华泰证券认为可靠的、已公开的信息编制，但华泰证券对该等信息的准确性及完整性不作任何保证。

本报告所载的意见、评估及预测仅反映报告发布当日的观点和判断。在不同时期，华泰证券可能会发出与本报告所载意见、评估及预测不一致的研究报告。同时，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。以往表现并不能指引未来，未来回报并不能得到保证，并存在损失本金的可能。华泰证券不保证本报告所含信息保持在最新状态。华泰证券对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

华泰证券（华泰证券（美国）有限公司除外）不是 FINRA 的注册会员，其研究分析师亦没有注册为 FINRA 的研究分析师/不具有 FINRA 分析师的注册资格。

华泰证券力求报告内容客观、公正，但本报告所载的观点、结论和建议仅供参考，不构成购买或出售所述证券的要约或招揽。该等观点、建议并未考虑到个别投资者的具体投资目的、财务状况以及特定需求，在任何时候均不构成对客户私人投资建议。投资者应当充分考虑自身特定状况，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。对依据或者使用本报告所造成的一切后果，华泰证券及作者均不承担任何法律责任。任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

除非另行说明，本报告中所引用的关于业绩的数据代表过往表现，过往的业绩表现不应作为日后回报的预示。华泰证券不承诺也不保证任何预示的回报会得以实现，分析中所做的预测可能是基于相应的假设，任何假设的变化可能会显著影响所预测的回报。

华泰证券及作者在自身所知情的范围内，与本报告所指的证券或投资标的不存在法律禁止的利害关系。在法律许可的情况下，华泰证券可能会持有报告中提到的公司所发行的证券头寸并进行交易，为该公司提供投资银行、财务顾问或者金融产品等相关服务或向该公司招揽业务。

华泰证券的销售人员、交易人员或其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。华泰证券没有将此意见及建议向报告所有接收者进行更新的义务。华泰证券的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。投资者应当考虑到华泰证券及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突。投资者请勿将本报告视为投资或其他决定的唯一信赖依据。有关该方面的具体披露请参照本报告尾部。

本报告并非意图发送、发布给在当地法律或监管规则下不允许向其发送、发布的机构或人员，也并非意图发送、发布给因可得到、使用本报告的行为而使华泰证券违反或受制于当地法律或监管规则的机构或人员。

本报告版权仅为华泰证券所有。未经华泰证券书面许可，任何机构或个人不得以翻版、复制、发表、引用或再次分发他人(无论整份或部分)等任何形式侵犯华泰证券版权。如征得华泰证券同意进行引用、刊发的，需在允许的范围内使用，并需在使用前获取独立的法律意见，以确定该引用、刊发符合当地适用法规的要求，同时注明出处为“华泰证券研究所”，且不得对本报告进行任何有悖原意的引用、删节和修改。华泰证券保留追究相关责任的权利。所有本报告中使用的商标、服务标记及标记均为华泰证券的商标、服务标记及标记。

中国香港

本报告由华泰证券股份有限公司或其关联机构制作,在香港由华泰金融控股（香港）有限公司向符合《证券及期货条例》及其附属法律规定的机构投资者和专业投资者的客户进行分发。华泰金融控股（香港）有限公司受香港证券及期货事务监察委员会监管，是华泰国际金融控股有限公司的全资子公司，后者为华泰证券股份有限公司的全资子公司。在香港获得本报告的人员若有任何有关本报告的问题,请与华泰金融控股（香港）有限公司联系。

仅供内部参考，请勿外传

香港-重要监管披露

- 华泰金融控股（香港）有限公司的雇员或其关联人士没有担任本报告中提及的公司或发行人的高级人员。
- 海光信息（688041 CH）、MiniMax（100 HK）、智谱（2513 HK）：华泰金融控股（香港）有限公司、其子公司和/或其关联公司在本报告发布日担任标的公司证券做市商或者证券流动性提供者。
- 有关重要的披露信息，请参华泰金融控股（香港）有限公司的网页 https://www.htsc.com.hk/stock_disclosure

其他信息请参见下方 **“美国-重要监管披露”**。

美国

在美国本报告由华泰证券（美国）有限公司向符合美国监管规定的机构投资者进行发表与分发。华泰证券（美国）有限公司是美国注册经纪商和美国金融业监管局（FINRA）的注册会员。对于其在美国分发的研究报告，华泰证券（美国）有限公司根据《1934 年证券交易法》（修订版）第 15a-6 条规定以及美国证券交易委员会人员解释，对本研究报告内容负责。华泰证券（美国）有限公司联营公司的分析师不具有美国金融监管（FINRA）分析师的注册资格，可能不属于华泰证券（美国）有限公司的关联人员，因此可能不受 FINRA 关于分析师与标的公司沟通、公开露面和所持交易证券的限制。华泰证券（美国）有限公司是华泰国际金融控股有限公司的全资子公司，后者为华泰证券股份有限公司的全资子公司。任何直接从华泰证券（美国）有限公司收到此报告并希望就本报告所述任何证券进行交易的人士，应通过华泰证券（美国）有限公司进行交易。

美国-重要监管披露

- 分析师郭雅丽、袁泽世本人及相关人士并不担任本报告所提及的标的证券或发行人的高级人员、董事或顾问。分析师及相关人士与本报告所提及的标的证券或发行人并无任何相关财务利益。本披露中所提及的“相关人士”包括 FINRA 定义下分析师的家庭成员。分析师根据华泰证券的整体收入和盈利能力获得薪酬，包括源自公司投资银行业务的收入。
- 智谱（2513 HK）：华泰证券股份有限公司、其子公司和/或其联营公司在本报告发布日之前的 12 个月内担任了标的证券公开发行或 144A 条款发行的经办人或联席经办人。
- 智谱（2513 HK）：华泰证券股份有限公司、其子公司和/或其联营公司在本报告发布日之前 12 个月内曾向标的公司提供投资银行服务并收取报酬。
- 智谱（2513 HK）：华泰证券股份有限公司、其子公司和/或其联营公司预计在本报告发布日之后 3 个月内将向标的公司收取或寻求投资银行服务报酬。
- 海光信息（688041 CH）、MiniMax（100 HK）、智谱（2513 HK）：华泰证券股份有限公司、其子公司和/或其联营公司在本报告发布日担任标的公司证券做市商或者证券流动性提供者。
- 华泰证券股份有限公司、其子公司和/或其联营公司，及/或不时会以自身或代理形式向客户出售及购买华泰证券研究所覆盖公司的证券/衍生工具，包括股票及债券（包括衍生品）华泰证券研究所覆盖公司的证券/衍生工具，包括股票及债券（包括衍生品）。
- 华泰证券股份有限公司、其子公司和/或其联营公司，及/或其高级管理层、董事和雇员可能会持有本报告中所提到的任何证券（或任何相关投资）头寸，并可能不时进行增持或减持该证券（或投资）。因此，投资者应该意识到可能存在利益冲突。

新加坡

华泰证券（新加坡）有限公司持有新加坡金融管理局颁发的资本市场服务许可证，可从事资本市场产品交易，包括证券、集体投资计划中的单位、交易所交易的衍生品合约和场外衍生品合约，并且是《财务顾问法》规定的豁免财务顾问，就投资产品向他人提供建议，包括发布或公布研究分析或研究报告。华泰证券（新加坡）有限公司可能会根据《财务顾问条例》第 32C 条的规定分发其在华泰证券内的外国附属公司各自制作的信息/研究。本报告仅供认可投资者、专家投资者或机构投资者使用，华泰证券（新加坡）有限公司不对本报告内容承担法律责任。如果您是非预期接收者，请您立即通知并直接将本报告返回给华泰证券（新加坡）有限公司。本报告的新加坡接收者应联系您的华泰证券（新加坡）有限公司关系经理或客户主管，了解来自或与所分发的信息相关的事宜。

仅供内部参考，请勿外传

评级说明

投资评级基于分析师对报告发布日后 6 至 12 个月内行业或公司回报潜力（含此期间的股息回报）相对基准表现的预期（A 股市场基准为沪深 300 指数，香港市场基准为恒生指数，美国市场基准为标普 500 指数，台湾市场基准为台湾加权指数，日本市场基准为日经 225 指数，新加坡市场基准为海峡时报指数，韩国市场基准为韩国有价证券指数，英国市场基准为富时 100 指数，德国市场基准为 DAX 指数），具体如下：

行业评级

增持：预计行业股票指数超越基准
中性：预计行业股票指数基本与基准持平
减持：预计行业股票指数明显弱于基准

公司评级

买入：预计股价超越基准 15%以上
增持：预计股价超越基准 5%~15%
持有：预计股价相对基准波动在-15%~5%之间
卖出：预计股价弱于基准 15%以上
暂停评级：已暂停评级、目标价及预测，以遵守适用法规及/或公司政策
无评级：股票不在常规研究覆盖范围内。投资者不应期待华泰提供该等证券及/或公司相关的持续或补充信息

法律实体披露

中国：华泰证券股份有限公司具有中国证监会核准的“证券投资咨询”业务资格，经营许可证编号为：91320000704041011J
香港：华泰金融控股（香港）有限公司具有香港证监会核准的“就证券提供意见”业务资格，经营许可证编号为：AOK809
美国：华泰证券（美国）有限公司为美国金融业监管局（FINRA）成员，具有在美国开展经纪交易商业务的资格，经营业务许可编号为：CRD#:298809/SEC#:8-70231
新加坡：华泰证券（新加坡）有限公司具有新加坡金融管理局颁发的资本市场服务许可证，并且是豁免财务顾问，经营许可证编号为：202233398E

华泰证券股份有限公司

南京
南京市建邺区江东中路 228 号华泰证券广场 1 号楼/邮政编码：210019

电话：86 25 83389999/传真：86 25 83387521
电子邮件：ht-rd@htsc.com

北京
北京市西城区太平桥大街丰盛胡同 28 号太平洋保险大厦 A 座 18 层/邮政编码：100032
电话：86 10 63211166/传真：86 10 63211275
电子邮件：ht-rd@htsc.com

深圳
深圳市福田区益田路 5999 号基金大厦 10 楼/邮政编码：518017
电话：86 755 82493932/传真：86 755 82492062
电子邮件：ht-rd@htsc.com

上海
上海市浦东新区东方路 18 号保利广场 E 栋 23 楼/邮政编码：200120
电话：86 21 28972098/传真：86 21 28972068
电子邮件：ht-rd@htsc.com

华泰金融控股（香港）有限公司
香港中环皇后大道中 99 号中环中心 53 楼
电话：+852-3658-6000/传真：+852-2567-6123
电子邮件：research@htsc.com
http://www.htsc.com.hk

华泰证券（美国）有限公司
美国纽约公园大道 280 号 21 楼东（纽约 10017）
电话：+212-763-8160/传真：+917-725-9702
电子邮件: Huatai@htsc-us.com
http://www.htsc-us.com

华泰证券（新加坡）有限公司
滨海湾金融中心 1 号大厦, #08-02, 新加坡 018981
电话：+65 68603600
传真：+65 65091183
https://www.htsc.com.sg

©版权所有2026年华泰证券股份有限公司