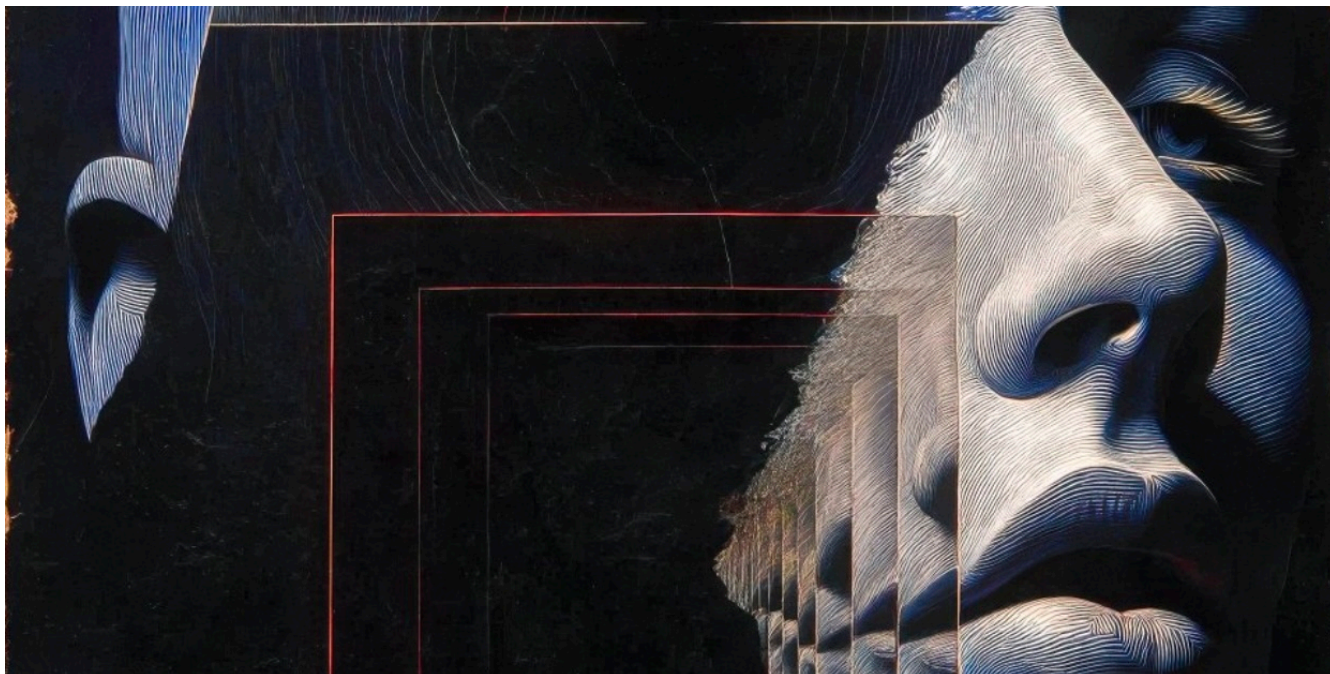


当 AI 无需我们也能学习

May 30, 2025

本文由 GPT-5.4 于 2026年3月10日 自动翻译。



图片来源: Substrata @blac_ai

传统上，一直是人类在训练人工智能。我们为数据打标签、设计奖励信号、手工设定目标。但一种微妙的范式转移正在发生：AI 系统开始利用来自其内部的信号来改进自己。

最新一代模型不再只依赖人类反馈或外部奖励，而是开始根据自身的置信度、一致性，甚至通过修改自己的代码库来提升自身能力。

这是一场安静的转变，但它将在未来几年改变 AI 的演化与学习方式。

从标签到自我反思

一种名为 Intuitor 的突破性方法，使模型能够把自己的“自我确信感”作为唯一需要的反馈信号。也就是说，模型会根据自己对答案的信心来学习；如果它足够自信，就会强化这种行为。不需要标签，不需要标准答案，也不需要地面真值，仅仅依靠 AI 从自身学习。

另一种方法则建立在“自治性”的思想之上。模型会用多种方式去解决同一个困难问题，只有当这些推理路径收敛到相同结论时，它才会信任这个结果。随着时间推移，这会不断增强模型对自己答案的信心，并使它能够在没有外部监督的情况下训练自己，甚至适用于数学这类复杂领域。

会为自己写升级版本的 AI

有些模型如今已经具备了设计自己更优版本的能力。有一个实验让 AI 的任务变成：生成新的算法来提升自身表现。它写代码、做测试、评估结果，再重复这一循环。最终，这个 AI 找到了连人类研究者都没有设计出来的优化策略，其效果超过了人类方案。

DeepMind 的 AlphaEvolve 则更进一步。它通过多轮演化改进代码，以优化基础设施并解决复杂科学任务。在一个案例中，它重新发现了一种更快的矩阵乘法方法，打破了一项自 1969 年以来一直保持的数学纪录。

代码的演化

也许最近最具戏剧性的例子，也是促使我写下这篇文章的原因，是 Darwin Godel Machine。它是一种通过重写自身代码来演化新智能体的系统。它会生成自己的多个版本，测试每个版本，并保留表现最好的那一个。经过许多轮迭代后，它会产出越来越复杂的编码智能体，而且每一代都优于上一代。这些智能体发展出更聪明的工作流、更好的调试策略，甚至还会形成内部机制来评估它们自己的修改。这是一种开放式的 AI 演化形式，在这里，智能增长并不是来自更多训练数据，而是来自递归式的自我改进。

我们正在见证一种转变：AI 正从“从数据中学习”，转向“从自身中学习”。这开启了新的机会，也带来了复杂的问题：

当 AI 以我们无法预测的方式进行创新时，会发生什么？

对于那些正在演化自身学习过程的系统，我们如何维持控制？

当这些系统深度嵌入经济体系时，我们又该如何维持治理与审计结构，以确保其行为安全、可靠且具韧性？

随着我们把更多自主权交给机器，我们做的已不仅仅是训练它们；我们也在赋权于它们。我们正在创造一种条件，使智能能够与我们共同演化。

近期论文：

- Learning to Reason without External Rewards arXiv 预印本 arXiv
- Can Large Reasoning Models Self-Train? arXiv 预印本 arXiv
- Can Large Language Models Invent Algorithms to Improve Themselves? arXiv 预印本 arXiv
- AlphaEvolve: A Gemini-Powered Coding Agent for Designing Advanced Algorithms Technical Report – DeepMind AlphaEvolve.pdf

- Darwin Godel Machine: Open-Ended Evolution of Self-Improving Agents arXiv 预印本 , Zhang, Hu, Lu, Lange, Clune , 2025 年 5 月 arXiv