

When AI Learns Without Us

May 30, 2025

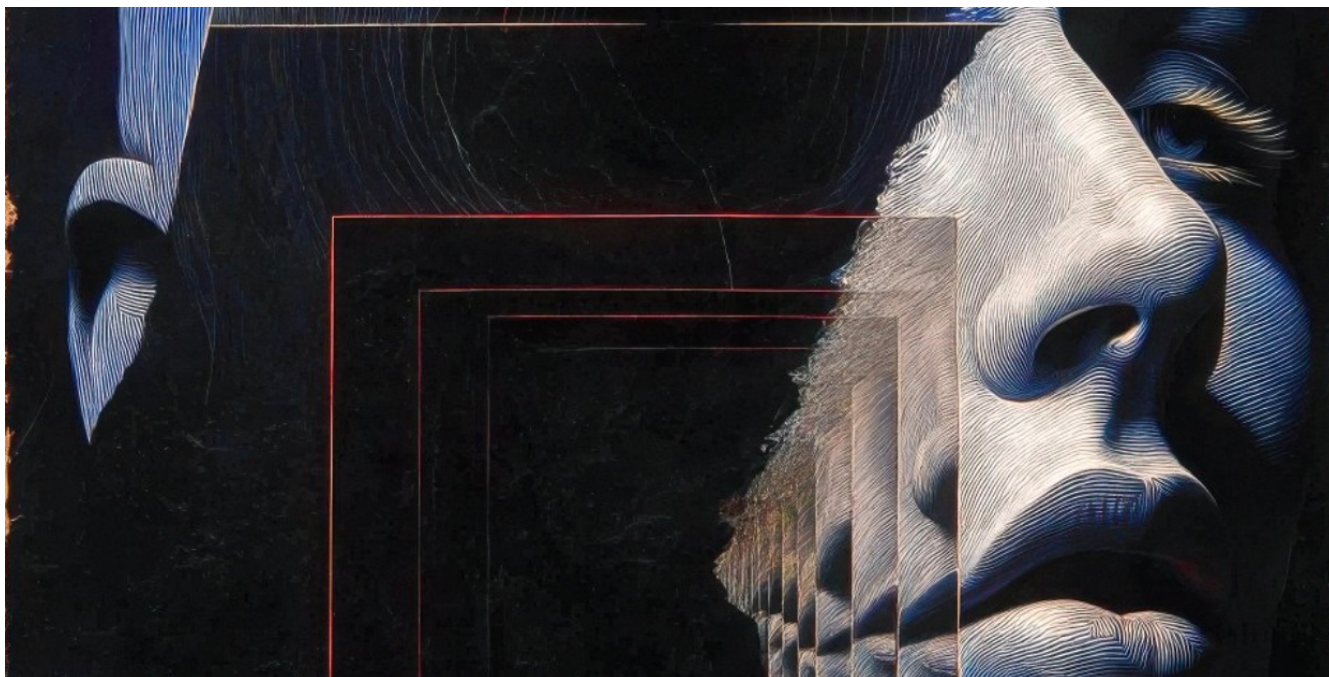


Image credit: Substrata @blac_ai

Humans have traditionally trained artificial intelligence – we label data, design reward signals, and hand-craft the objectives. But a subtle paradigm shift is underway: AI systems are beginning to improve themselves using signals from within.

Instead of relying solely on human feedback or external rewards, the newest AI models learn from their own confidence, consistency, and even improving themselves by adjusting their own codebase.

It's a quiet transformation that will change how AI evolves and learns in the coming years.

From Labels to Self-Reflection

One breakthrough method, called Intuitior, enables a model to use its own self-certainty — a sense of how confident it is in an answer — as the sole feedback it requires. If it's confident, it reinforces that behaviour. No labels. No gold-standard ground truth answers. Just the AI learning from itself.

Another approach builds on this idea of self-consistency. The model addresses a challenging problem in multiple ways and only trusts the results when those paths converge to the same conclusion. Over time, this builds confidence in its answers and allows it to train itself — even in complex domains like math, entirely without external supervision.

AI That Writes Its Own Upgrades

Some models are now capable of designing better versions of themselves. One experiment gave an AI the task of generating new algorithms to improve its own performance. It wrote code, tested it, evaluated the results — and repeated the loop. The AI ultimately discovered optimisation strategies that outperformed anything designed by human researchers.

DeepMind's AlphaEvolve takes this further. It evolves code over many iterations to improve infrastructure and solve complex scientific tasks. In one case, it rediscovered a faster method for multiplying matrices — breaking a mathematical record that had stood since 1969.

The Evolution of Code

Perhaps the most dramatic recent example (prompting me to write this!) is the Darwin Gödel Machine — a system that evolves new agents by rewriting its own code. It generates versions of itself, tests each one, and keeps the best performers. Over many iterations, it creates increasingly sophisticated coding agents, each better than the last. These agents develop smarter workflows, better debugging strategies, and even internal mechanisms for evaluating their own changes. It's a form of open-ended AI evolution, where intelligence grows not from training data — but from recursive self-improvement.

We're witnessing a shift from AI that learns from data to AI that learns from itself. This opens up new opportunities — and complex questions:

What happens when AI innovates in ways we can't predict?

How do we maintain control over systems that are evolving their own learning processes?

As these systems integrate deeply into the economy, how do we maintain structures of governance and audit to ensure safe, reliable, and resilient behaviour?

As we hand over more autonomy to machines, we're not just training them; we're also empowering them. We're creating the conditions for intelligence to co-evolve with us.

Recent papers:

- **Learning to Reason without External Rewards** [arXiv preprint](#) [arXiv](#)
- **Can Large Reasoning Models Self-Train?** [arXiv preprint](#) [arXiv](#)
- **Can Large Language Models Invent Algorithms to Improve Themselves?** [arXiv preprint](#) [arXiv](#)

- **AlphaEvolve: A Gemini-Powered Coding Agent for Designing Advanced Algorithms Technical Report** – DeepMind AlphaEvolve.pdf
- **Darwin Gödel Machine: Open-Ended Evolution of Self-Improving Agents** arXiv preprint Zhang, Hu, Lu, Lange, Clune. May 2025 arXiv