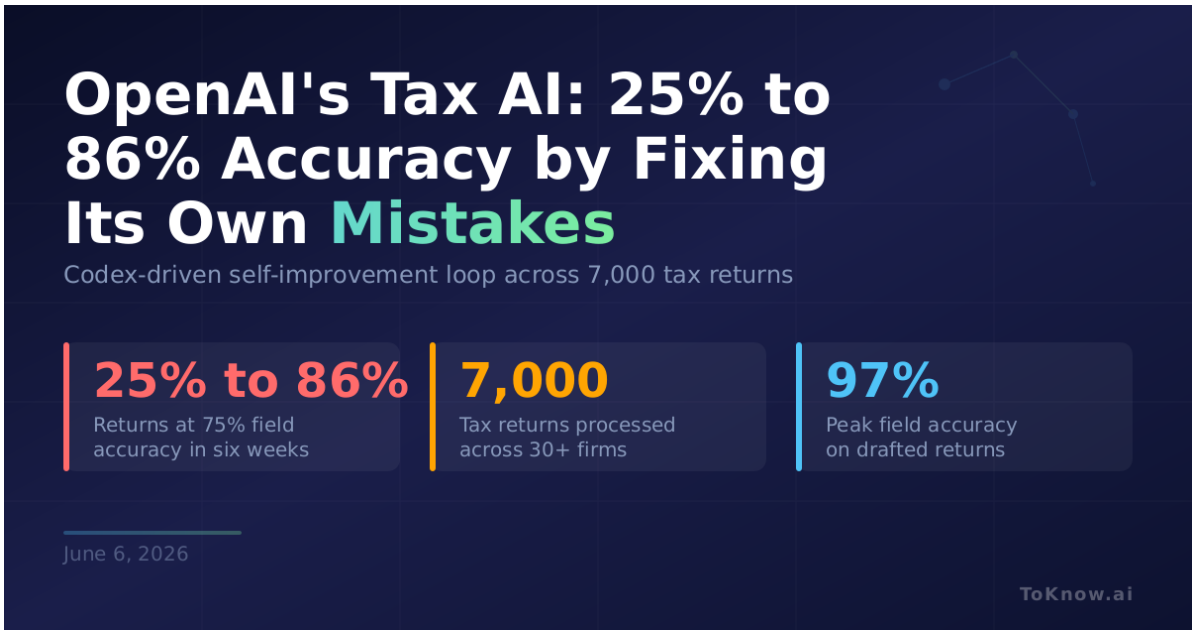
OpenAI's Tax AI Went from 25% to 86% Accuracy in Six Weeks by Fixing Its Own Mistakes

Kabui, Charles

2026-06-06

Read at ToKnow.ai



OpenAI and Thrive Holdings co-developed a tax preparation AI for Current (formerly Crete Professionals Alliance), a network of 30+ accounting firms. Practitioners upload source files, and Tax AI drafts 1040 and 1041 returns ready for review. It handled 7,000 returns this

season, saving practitioners about a third of their time and increasing throughput by roughly 50%. At launch, only 25% of returns hit 75% correct field completion. Six weeks later, 86% reached that mark, with peak accuracy reaching 97%. The improvement came not from manual engineering but from a three-part loop: practitioner corrections become structured evaluations, Codex investigates each failure's root cause, writes a fix, validates it against regression tests, and surfaces a pull request for review.

When a practitioner corrects a value, the system records exactly what Tax AI proposed versus what was filed. Repeated patterns, like consistently missing "fair rental days" on Schedule E, get grouped into eval targets. Codex inspects the full pipeline from document extraction through tax-engine mapping, diagnoses whether the issue is a schema gap or a missed extraction pattern, and writes a fix. One senior accountant who spent 180 hours on tax prep last year spent 15 this year, freeing her to call every client and take on new business.

Traditional ML improvement cycles take weeks: collect data, retrain, redeploy. Here, production errors become engineering tasks within hours, and the agent doing the engineering is itself an AI. Any expert-review workflow where corrections carry structured signal, from legal documents to insurance claims, could run the same loop.

Read More: Hermes Agent by NousResearch: Self-Improving AI Through Reflective Learning

Sources:

- Building Self-Improving Tax Agents with Codex (OpenAI)
- Harness Engineering: Leveraging Codex in an Agent-First World (OpenAI)
- Symphony: Open-Source Codex Orchestration (OpenAI)
- Current (formerly Crete Professionals Alliance)
