

Introducing the Agents API

PRIMARY

Publisher	OpenAI
Published	2026-09-10
Retrieved	2026-09-16
Cite key	openai2026introduci
URL	https://openai.com/index/introducing-the-agents-api/

Claims drawn from this source. CLAIM-2026-09-16-01, CLAIM-2026-09-16-02

Extracted passages (WebFetch, 2026-09-16)

The Agents API lets developers "create a production-ready agent in a single API call by specifying the task, model, tools, and environment."

Sessions and orchestration. The service provides a "durable session and orchestration layer" with "context management that helps agents carry relevant information across longer sessions."

Context compaction. "The Agents API automatically compacts earlier context as a session approaches its context limit, preserving information the agent needs to continue."

Subagents. With "multi-agent support, the Agents API can break complex tasks into independent pieces and delegate them to subagents that work in parallel."

Execution environments. Code executes in an OpenAI-managed sandbox, on self-hosted infrastructure, or in a sandbox partner environment. OpenAI is "partnering with ecosystem providers, including Blaxel, Cloudflare, Daytona, DigitalOcean, E2B, Modal, Oracle, Runloop, and Vercel, to provide first-class integrations for a range of needs." The OpenAI-hosted sandboxes use "the same sandboxing infrastructure that powers Codex and ChatGPT."

Tools. The platform "supports MCP, custom functions, and built-in tools like web search."

Availability and pricing. The Agents API is "available in public beta today to all developers. There are no additional fees for using the Agents API - you simply pay for the tokens and tools your agents use."

Absent. The announcement page carries no statement on data residency, regions, retention, logging or compliance. The model shown in the code examples is gpt-6-astra.