

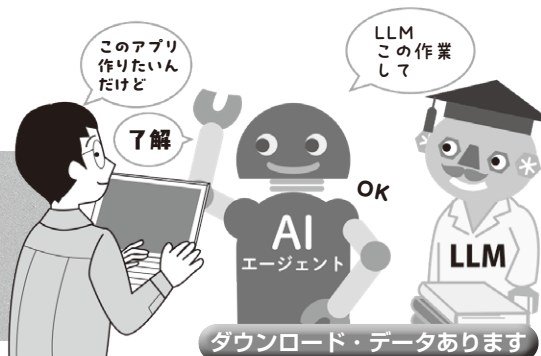
OpenAI Agents SDKで始める

自分専用

AIエージェント開発

新連載

第1回 OpenAI Agents SDKの紹介とインストール



ご購入はこちら

氏森 充

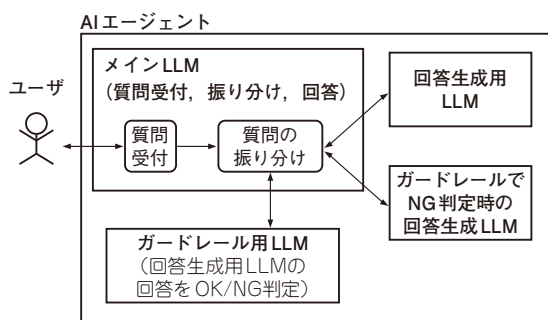


図1 本連載で作成するAIエージェントのイメージ

LLMの回答が適切か判定する、別のLLM（ガードレール用LLM）を組み込む

AIエージェントはClaude Codeなどのサービス経由で使うだけでなく、自分で設計、実装することも可能です。最近、AIエージェント構築に必要な機能をあらかじめ備えた専用のSDK（ソフトウェア開発キット）やフレームワークがそろってきました。

本連載では、筆者が特に注目しているOpenAI Agents SDKを使って、AIエージェントを開発していきます。AIエージェントはさまざまなことができますが、本連載では例として、プロンプトや生成された内容が不適切とされた場合、処理を中断させるものを中心に上げます（図1）。これを開発していく中で、その特徴や設計思想、実装スタイルについて詳しく解説します。

● 今回のテーマ

AIエージェントを作るためのツールとして、OpenAI Agents SDKの紹介とインストール手順を説明します。

AIエージェントを作るために
知っておきたいこと

● エージェントという概念は昔からあった

本稿で扱うエージェントとは、特定の目標に向かっ

て自律的に行動し、外部環境とやり取りを行うシステムを指します。概念自体は新しいものではなく、以前からルール・ベースのシステムや認知アーキテクチャ、強化学習などを用いたエージェントが長年にわたり研究、応用されてきました。こうした従来型のエージェントでは、タスクに応じたルール設計や状態管理、学習アルゴリズムの実装などが必要であり、構築や拡張には高度な専門知識と相応の労力が求められました。

しかし、最近の大規模言語モデル（LLM：Large Language Model）の登場によって、状況は大きく変わりつつあります。LLMは膨大なテキスト・データをもとに事前学習されたモデルを有しており、自然言語による柔軟な対話や推論、知識検索が可能です。これをエージェントの頭脳として活用することで、明示的なルールを設計せずとも、高度な言語理解と行動判断が可能なエージェントを、比較的容易に構築できるようになりました。こうしたLLMを中核としたエージェントは、一般にAIエージェント、あるいはLLMエージェントと呼ばれ、タスクの自動実行、外部ツールとの連携、柔軟なインタラクションなど、幅広い分野で活用が始まっています。

● ゼロから作るためには高度な技術がたくさん必要

AIエージェントは、自律的な処理や柔軟なタスク実行を可能にする非常に有用な仕組みです。それゆえ、エージェントには多岐にわたる高度な技術が必要であり、特に次のような技術的要素が複合的に関与しています。

- LLMの選定とプロンプト設計
- タスクの分解と処理フローの制御
- 外部ツールやAPIとのインターフェース連携
- 状態・メモリの管理と永続化
- 非同期処理やエラー対応などの実行制御
- 入力の検証や動作の制限により、信頼性と安全性を確保

開発者にとって、これら全てを自前で設計、実装する