

レベル3 AIエージェントにおける LLM層とアプリケーション層の役割

Genspark、Manus、OpenAI Operator、Deep Researchの事例分析

 LLM層（認知エンジン）

 アプリケーション層（実行環境）

レベル3 AIエージェントの定義

自律型 AI エージェント

- 🕒 長期間（数日）にわたる自律的タスク実行
- 🧠 人間介入なしの独立意思決定
- 🔄 変化する状況への適応能力
- 🎯 定義された目標の持続的 pursuit

AIエージェントの進化

レベル1
基本対話型AI



レベル2
推論型AI



レベル3
自律型AI

OpenAIフレームワークにおける位置づけ

「エージェント」（レベル3）は「リーズナー」（レベル2）の次段階として位置づけられる

基本アーキテクチャ：二層構造



LLM層

「脳/心」としての役割



自然言語理解 (NLU)



意思決定



ツール呼び出し



推論と問題分解



知識統合



スクリプト生成



アプリケーション層

「身体/手」としての役割



ツール実行インターフェース



状態管理と長期記憶



マルチエージェント連携



環境対話機能



タスク追跡・再開機能



API/ツール接続

LLM層の役割と機能

LLM層 - レベル3 AIエージェントの「知性」

「単なるプランナー」から「動的スクリプター」「継続的ディレクター」へ進化



自然言語理解 (NLU)

複雑な指示の解釈、意図理解、文脈の維持と追跡



推論と問題分解

複雑な目標を実行可能なサブタスクに分解し、論理的に計画



意思決定

不確実性の中での判断、優先順位付け、次のステップ選択



知識統合

事前学習知識と実行中に得た情報の統合、一貫性維持



ツール呼び出し

適切なツール選択、パラメータ設定、結果の解釈と統合



動的スクリプティング

環境変化に応じた実行コードの生成・修正、長期タスク管理

アプリケーション層の役割と機能

アプリケーション層 - レベル3 AIエージェントの「身体/手」

「単なる実行環境」から「エージェントOS（運用基盤）」へ進化



ツール実行と環境対話

Web API、データベース、コード実行など、外部環境との実際のインタラクション



状態管理と長期記憶

タスクの進捗状況追跡、履歴保持、MCP/LTM長期記憶管理



セッション永続性

複数セッションに渡る一貫した状態維持、再開機能、休止機能



エラー回復と頑健性

外部システム障害の処理、リトライメカニズム、例外ハンドリング



ユーザーインターフェース

人間との対話窓口、進捗状況の可視化、介入/フィードバック受付



マルチエージェント連携

複数エージェント間の調整、通信プロトコル、リソース管理

事例分析：4つのAIEージェント比較

Genspark

- 💡 LLM層
 - MoA（複数LLM連携）
 - タスク分解と動的スクリプト生成
 - 反復検証と再計画

- ⚙️ アプリケーション層
 - MCP（状態管理）
 - 多様なツール連携
 - API直接統合

Manus

- 💡 LLM層
 - プランナー・実行・検証サブエージェント
 - 階層的タスク分解
 - 状況適応型計画修正

- ⚙️ アプリケーション層
 - 長期記憶システム
 - クラウド永続運用基盤
 - 状態回復メカニズム

OpenAI Operator

- 💡 LLM層
 - CUA（視覚認識LLM）
 - 画面理解と操作判断
 - UI要素認識と意味解釈

- ⚙️ アプリケーション層
 - ブラウザGUI自動化
 - 人間介入メカニズム
 - スクリーンインタラクション

Deep Research

- 💡 LLM層
 - 強力単体LLM
 - 再帰的情報収集計画
 - 知識合成と評価

- ⚙️ アプリケーション層
 - ウェブ検索・閲覧
 - Pythonツール実行
 - 短期集中型バースト運用

比較分析表：4つのAIEージェント

評価項目	 Genspark	 Manus	 OpenAI Operator	 Deep Research
 LLM層の能力	MoA（複数LLM連携） 動的スクリプト生成	階層的タスク分解 複数サブエージェント	UI要素認識 CUA（視覚認識LLM）	強力単体LLM 再帰的情報収集
 アプリケーション層	MCP（状態管理） API直接統合	長期記憶システム 状態回復メカニズム	ブラウザGUI自動化 人間介入機能	ウェブ検索・閲覧 結果保存・出力
 主要ツール群	多様なツール連携 コード実行	クラウドストレージ メッセージング	スクリーン操作 ブラウザ操作	Pythonツール 検索エンジン
 長期タスク対応	反復検証 再計画能力	クラウド永続運用 状態復元	進捗監視 ヒューマンループ	短期集中型 バースト型実行

結論と今後の展望

「LLM層は知能の司令塔、アプリケーション層は現実との接地・運用基盤」



両輪型協調の必要性

LLM層とアプリケーション層の緊密な連携が、真のレベル3自律性実現には不可欠



役割明確化の重要性

LLM = 知性・計画・オーケストレーション、アプリケーション = 実行・永続性・運用



相互補完的な進化

片方だけの強化では本格的な「レベル3」の長期自律性は実現不可能



🚀 アプリケーション層の未来：エージェントOS

🔧 プロセス管理

📄 標準化メモリサービス

🛡️ スケジュール・リソース管理

🗨️ エージェント間通信

🔗 フレームワーク標準化

🌐 分散エコシステム