

Google I/O 2026: Gemini 3.5, Spark, Omni 評価報告

「単一モデルの性能」から「エージェント・フルスタックの実行力」への転換と実践的デプロイメント評価



Model Wars



I/O 2026の本質は単一の「神モデル」の発表ではない。モデル、サンドボックス、配布面を一気通貫で揃えた「実行 (Execution)」と「統治 (Governance)」のエコシステム始動である。



即時導入可能 (Deploy)

Gemini 3.5 Flash
- 1M Context
- Interactions API



検証・パイロット段階 (Pilot)

Managed Agents in Gemini
API
- Public Preview

Gemini Spark
- Trusted Testers / Beta

Gemini Enterprise Agent
Platform



評価未成熟 (Watch)

Gemini Omni Flash
- API未提供
- 独自ベンチマーク不足

Gemini 3.5 Pro
- API Docs未掲載



1M Context Window

Intelligence Index: 55

\$1.50 / \$9.00 (High Thinking Setting)

アーキテクチャ

Gemini 3系のツール群を継承。「Thinking levels」導入により品質・コスト・レイテンシを動的調整可能。

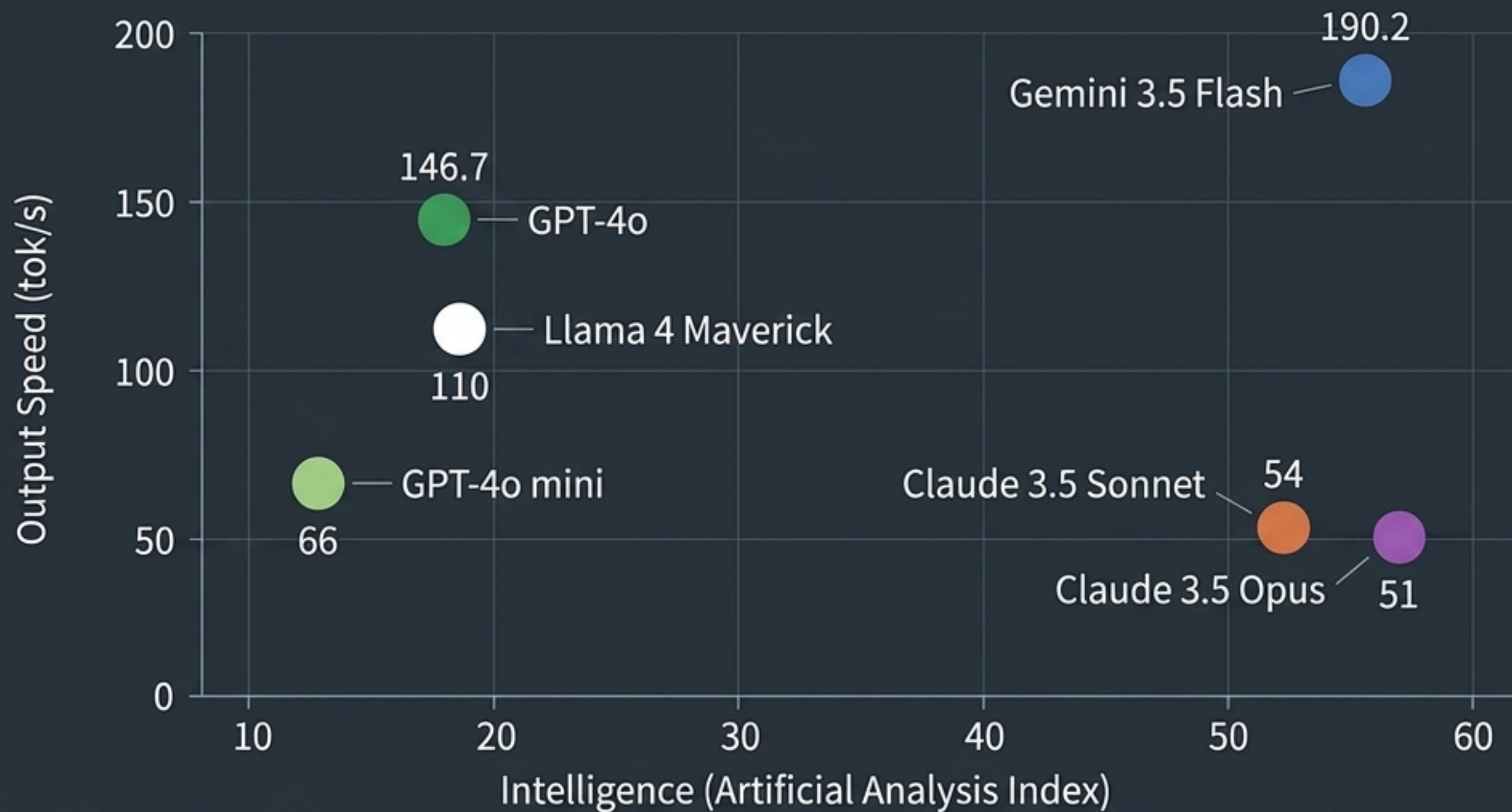
モダリティ

入力 (Text, Image, Audio, Video, PDF) -> 出力 (Text)

制約事項

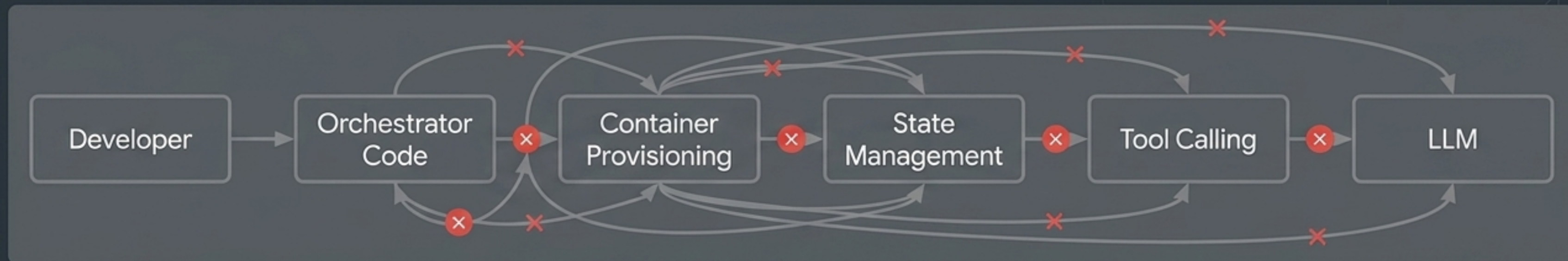
Computer Useは未対応。GUI自動操作ワークロードには不完全。

フラッシュ級の数度でありながら上位帯の推論能力を持つ、「エージェント・コーディング向け」の実用特化モデル。



3.5 FlashはGPT-4oを突き放し、Claude 3.5 Sonnetと同クラス以上の実力を示す。
絶対首位（Opus 3.5）には届かないが、「能力対速度」の最適解を提示。

Traditional Implementation

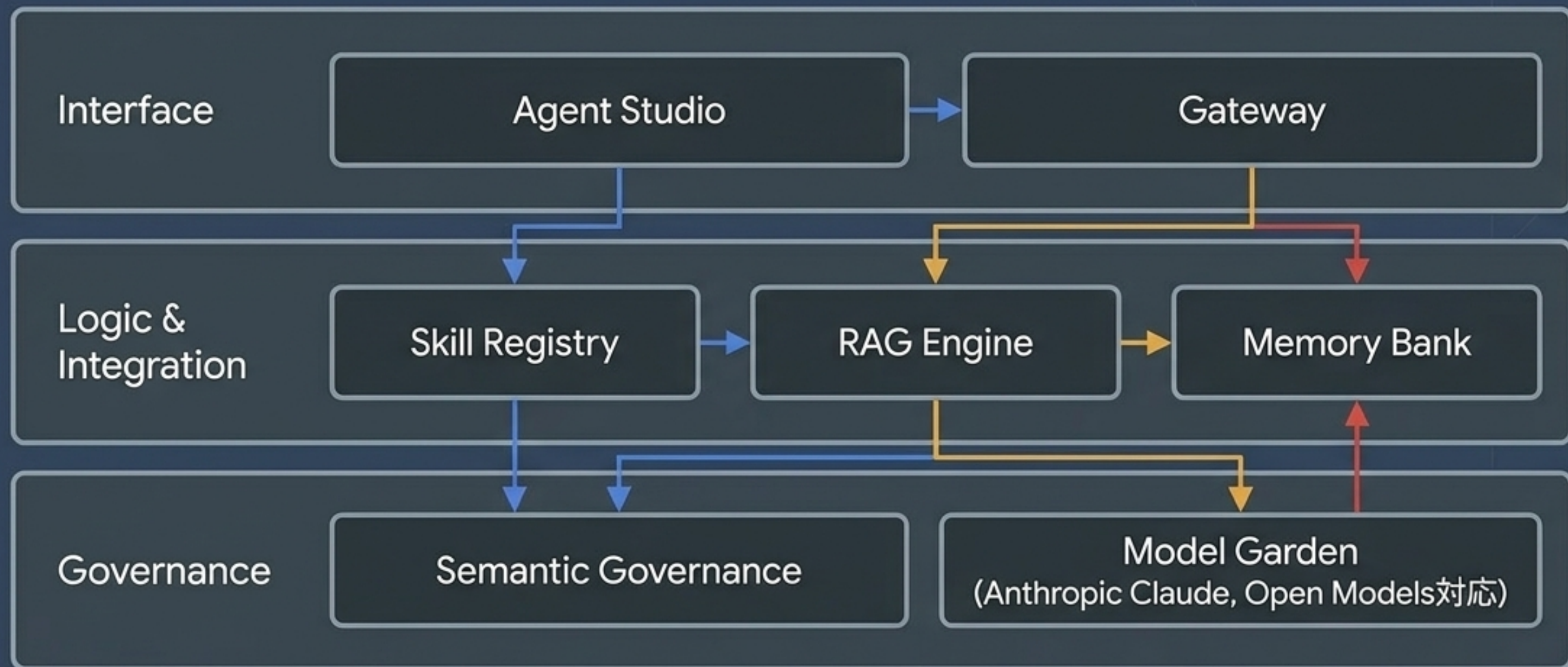


Managed Agents API



DXの真のブレークスルー。オーケストレーター構築、コンテナのライフサイクル管理、メモリ永続化の負担をGoogleが肩代わりする設計。

Gemini Enterprise Agent Platform



- 旧Vertex AIからの発展。単なるホスティングではなく、「マルチモデル・マルチエージェントの統制・評価基盤」としての再定義。
- 新規顧客向けに\$300クレジットを提供し、エンタープライズ導入を強力に推進。

理論上のDX - Theory

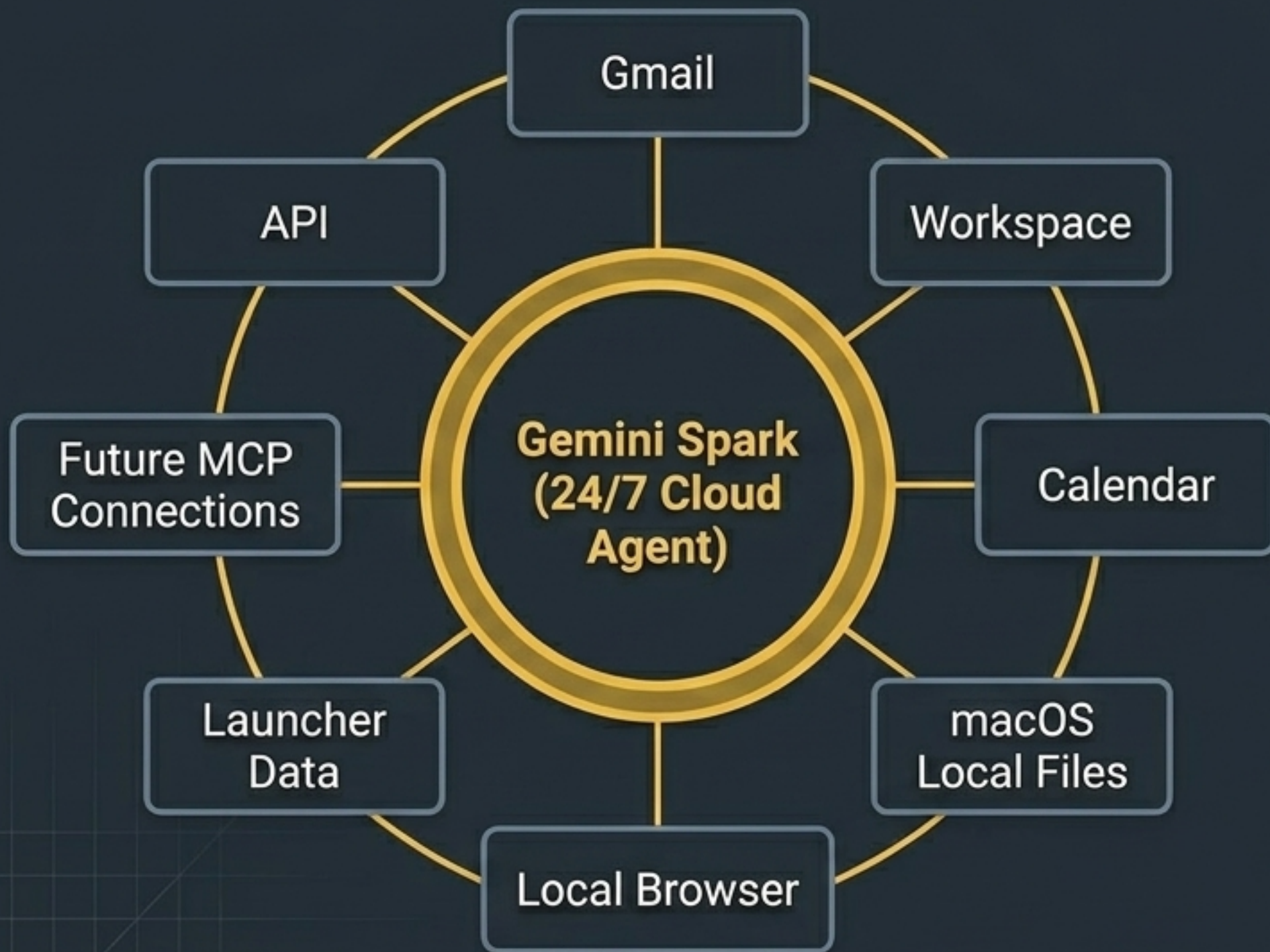
- Interactions APIと`google-genai` SDKによる洗練された実装。
- Search, Maps, Code, Filesのネイティブサポート。
- シームレスなAgent Studio統合。

初週の実運用摩擦 - Reality

- 厳しいQuota (割当) 制限 (Weekly / 24-hour rate limits)。
- 高トラフィックエラー (High traffic errors) と Infinite loading loops。
- 「Flash」のブランド期待値を超えるコスト (3 Flash Preview比で高額化)。

理論上のDXは大幅に向上したが、実運用品質は「即時の全社標準化」に耐えうるほど安定していない。段階的導入が必須。

Blast Radius (Risk Topology)



機能

クラウド常駐でバックグラウンド作業を横断実行。

リスクの性質変化

懸念点は「AIのハルシネーション」から「誤った自動化・過剰権限の付与」へ移行。

ガードレール

高リスク操作（送金・送信）前の確認、オプトイン方式、接続アプリの選択制。

従来のチャットボットとは比較にならないデータ集中と権限設計の課題を伴う。

	Gemini Omni Flash	Veo 3.1	Sora 2 (OpenAI)
モダリティ	Text/Image/Audio/Video -> Video+Audio	Video	Text/Image -> Video+Audio
API提供状況	未公表 (数週間内予定)	Gemini API/Vertex経由 で利用可	廃止予定 (2026-09-24)
価格	未公表	秒課金 (公開実績あり)	\$0.16/s, \$0.32/s 以上
独立評価	未成熟 (自己申告のみ)	VBench評価データあり	独自比較あり

物理的モーションや一貫性にまだ課題あり。「戦略的ポテンシャル」は極めて高いが、定量比較の準備は整っていない。現時点ではUI先行。

Model Safety (Gemini 3.5 Flash)	Authorization Risk (Managed Agents & Spark)	Generative Integrity (Gemini Omni)
<p>現状</p> <p>拒絶トーン (Refusal tone) の改善。 過剰な拒絶 (Unjustified refusals) の抑制。</p> <p>課題</p> <p>公開時点でのハルシネーション率 (Hallucination rate) の外部指標が未公開。</p>	<p>現状</p> <p>隔離Sandboxによる実行。 デフォルトで外部ネットワーク・認証情報へのアクセス「拒否 (Deny)」。</p> <p>課題</p> <p>Sparkにおける個人データとの広範な結合リスク。</p>	<p>現状</p> <p>SynthID電子透かし。 AI Content Detection APIの同時リリース。</p> <p>課題</p> <p>生成メディアの来歴確認が企業導入のボトルネックに。</p>

英語圏メディア (The Verge, WIRED)

「単なるチャットボットから全目的AIハブへの進化。しかし、常時稼働エージェントへ『委任してよいのか』という懸念が残る。」

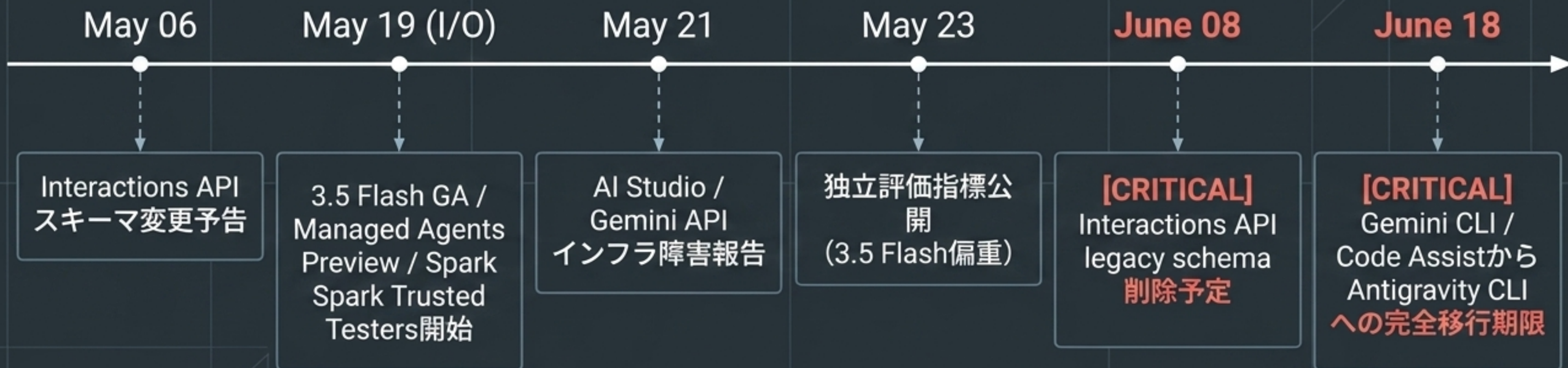
開発者コミュニティ (Reddit, Google AI Forum)

「能力と速度は実務で十分 (3.1 Pro以上)。しかし『Flash』の価格期待値から外れており、Quota制限と無限ループのバグが実運用を妨げている。」

現場の反応・不在のフラッグシップ

「最も期待されていた最上位モデル『Gemini 3.5 Pro』の即時公開が見送られたことへの不満の声 (Business Insider)。」

開発陣は能力を高く評価しつつも、運用環境の粗さと価格設定に不満を抱えている。



6月にかけて破壊的変更 (Breaking Changes) が連続する。早期導入組は直ちにCLIとスキーマの移行計画を策定すること。

今すぐ導入 (Deploy) - Gemini 3.5 Flash

高スループット、コーディング、エージェントのバックエンドとして即時適用。一般ベンチマークを鵜呑みにせず、自社ワークロードに基づくEvalルーティングを構築せよ。

慎重にパイロット検証 (Pilot) - Managed Agents / Spark

独自オーケストレータからの移行を検証開始。ただし、厳格なサンドボックス設計とHuman Oversight（人間の監視）を前提とした限定導入に留める。

動向注視 (Watch) - Gemini Omni / 3.5 Pro

APIの安定供給と独立ベンチマークが出揃うまで実稼働への組み込みは待機。生成コンテンツ検知API（Detection API）の検証を先行させる。

**Googleの新たな「堀」はLLM単体の知能ではなく、
配布経路とエージェント実行環境の統合にある。
このフルスタックに最適化できる企業が次の勝者となる。**