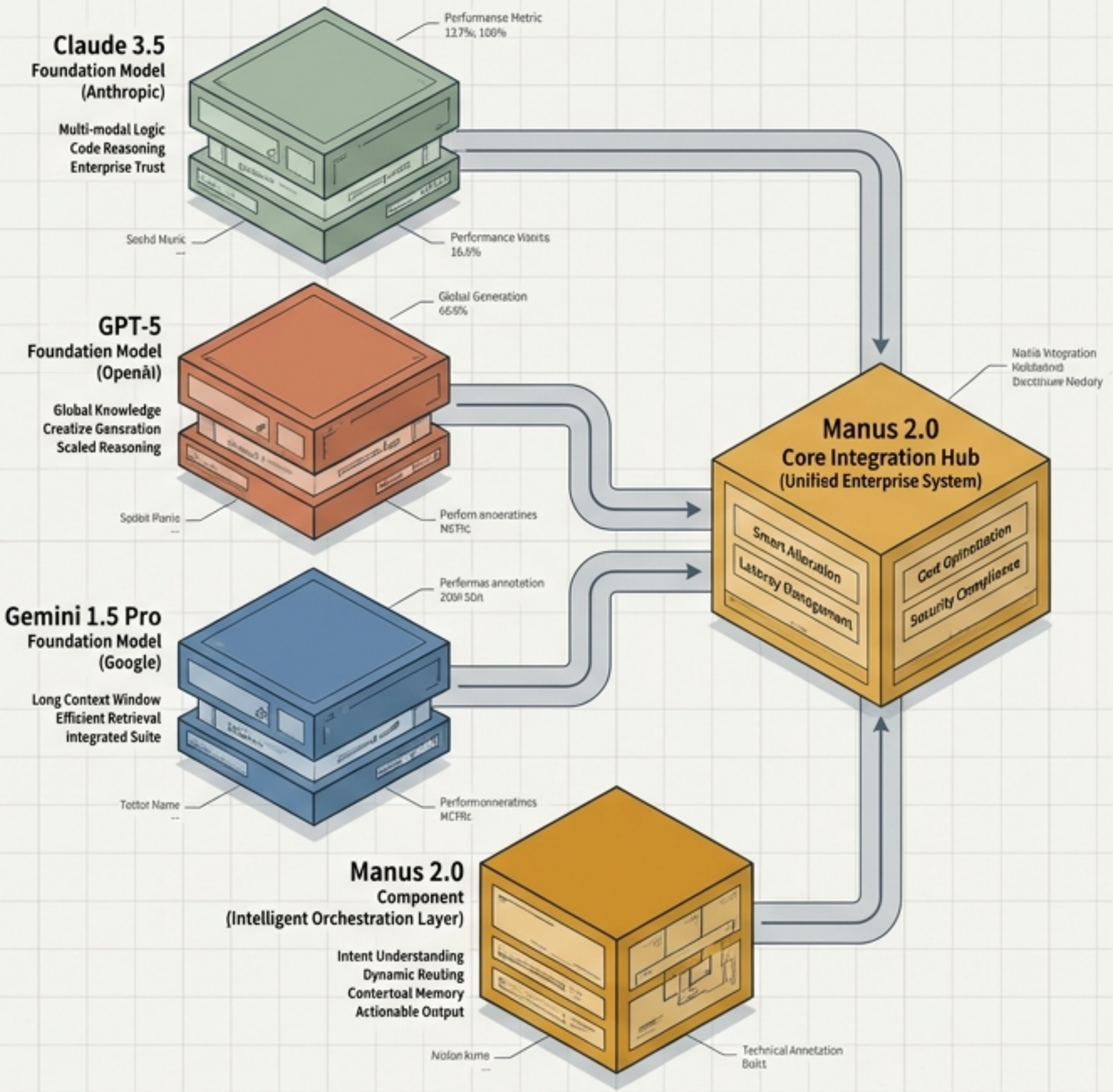
Manus 2.0 vs 3大基盤モデル（GPT / Claude / Gemini）
技術比較・導入戦略調査（2026年秋季版）

「最も賢いAI」を選ぶ時代から、「最適なシステム層」を選ぶ時代へ

本資料では、AI戦略を単一モデルの性能競争から、エンタープライズ環境における最適化されたシステム全体の構築へとシフトするためのアーキテクチャ的アプローチを詳述します。Manus 2.0が、最新の基盤モデルをどのように統合、調整、管理し、ビジネス価値を最大化するシステム層として機能するかを分析します。

エンタープライズ技術エグゼクティブ・
AIアーキテクト向け内部資料



Gemini Notebook

比較の前提となる3つのパラダイムシフト



前提の崩壊：レイヤーの不一致

Manus 2.0とGPT/Claude/Geminiは同列の製品ではない。前者は「エージェント・システム (Cascade) 」であり、後者は「基盤モデル (API) 」である。パラメータ数や単一コンテキストの直接比較は無意味。



指標の転換：単体知能からタスク完遂力へ

評価軸はベンチマークの正答率から、エージェント業務における「トークン削減率」「タスク完了時間」「最終コスト」といったTask-level efficiencyへと移行している。



第3の道：スタックの分離

Manus Flexの登場により、推論モデル（プロバイダー）とエージェント実行基盤を分離して選択することが可能になった。「どちらか」ではなくハイブリッドが合理的な解となる。

Gemini Notebook

2026年9月：チャットから「自律型エージェント」への一斉シフト

9/2: Gemini 3.8 Flash
(長文脈・高速化)

9/3: GPT-6 Astra
(推論の深化)

9/22: Claude Opus 5.5
(複雑な判断)

9/28: Claude Sonnet 5.5
(Agentic workloads)

基盤モデルアップデート（認知の高度化）

オーケストレーション・エージェント（行動の自動化）

9/28: Manus 2.0
(Cascade Agent Harness発表)

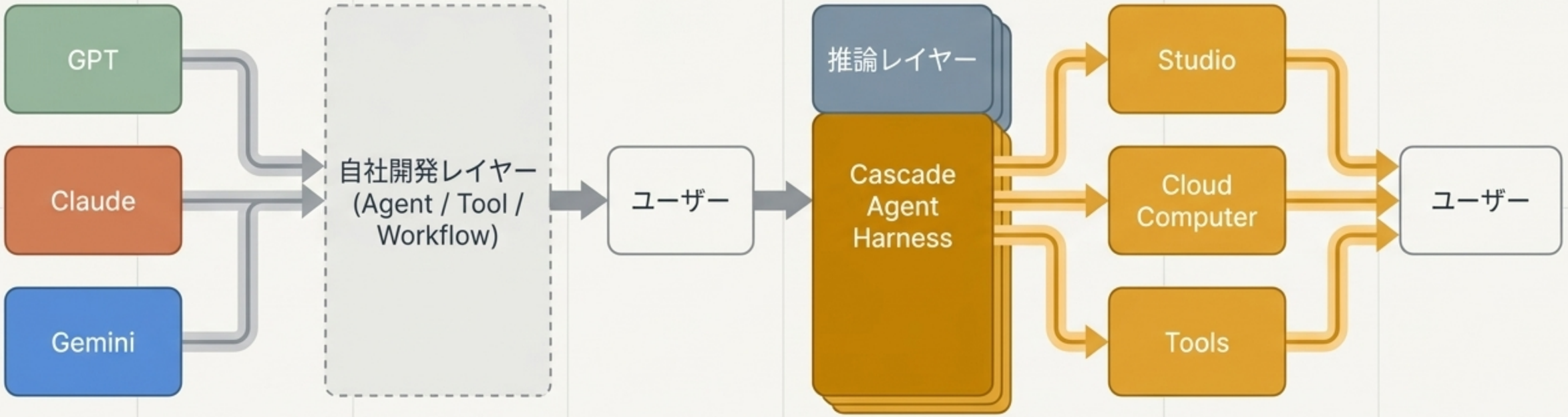
9/29: GPT-6.1 Sol
(Agentic coding / Computer use)

9/29: Manus Flex
(推論と実行の分離)

※Gemini 4 Argon (10/1日本発表) は現在Fairwind Program経由の限定展開中であり、本資料の実用API比較からは除外する。

Gemini Notebook

アーキテクチャの階層：自社で「どこまで構築するか」の違い



基盤モデルAPI

- **特徴:** 高い透明性と制御性。モデル単位でコンテキストやツールを厳密に定義可能。
- **課題:** Planning、コンテキスト管理、実行環境（Sandbox）は全て自社で設計・保守が必要。

Manus 2.0 / Flex

- **特徴:** 「仕事への変換」を丸ごと引き受ける。必要な能力とツールを必要な時点で動的に編成。
- **課題:** モデル単体の透明性は低下する（※Flexを利用することで推論プロバイダーの選択は可能）。

Gemini Notebook

「比較不能」なスペック：コンテキスト長とモーダルの誤解

Spec-sheet Myths

Architectural Realities

コンテキスト長 (Context Window)

GPT-6.1 (1.05M) / Claude (1M) / Gemini (1M)
vs **Manus** (5,000制限)

Manus API v2の約5,000トークンは「1回の直接プロンプトの上限」に過ぎない。**Manus**はFiles、Projects、Task referenceを別コンテキスト源として扱うシステム設計であり、基盤モデルの**総コンテキスト長とは別概念**ある (Cascadeの真のコンテキスト長は未公開)。

マルチモーダル (Multimodality)

「Manusは動画もゲームも作れるから、Geminiより高度なマルチモーダルである」

性質が異なる。Gemini 3.8 Flashは単一モデルが直接映像や音声を解釈する「Native Modality」。**Manus**はVideo Editor等のツール群を介して素材をオーケストレーションする「Application-level Orchestration」である。

Gemini Notebook

パフォーマンスの再定義：正答率から「実行の経済性」へ





消費トークン:
↓-23.2%

タスク完了時間:
↓-28.2%

実行コスト:
↓-32.0%

基盤モデルの戦い

各社が多様な外部ベンチマークで「1回の推論品質」を競う。しかしエージェント業務では「何回モデルを呼ぶか」「不要なコンテキストを省けるか」「ツール失敗をどうリカバーするか」が最終コストを決める。

Manusの最適化軸

Cascade導入によるシステム効率の改善（※特定テスト・旧システム比）。
インサイト: 「最も賢いモデル」ではなく、**「コンテキストと実行の経済性」**が実務の勝敗を分ける。

Gemini Notebook

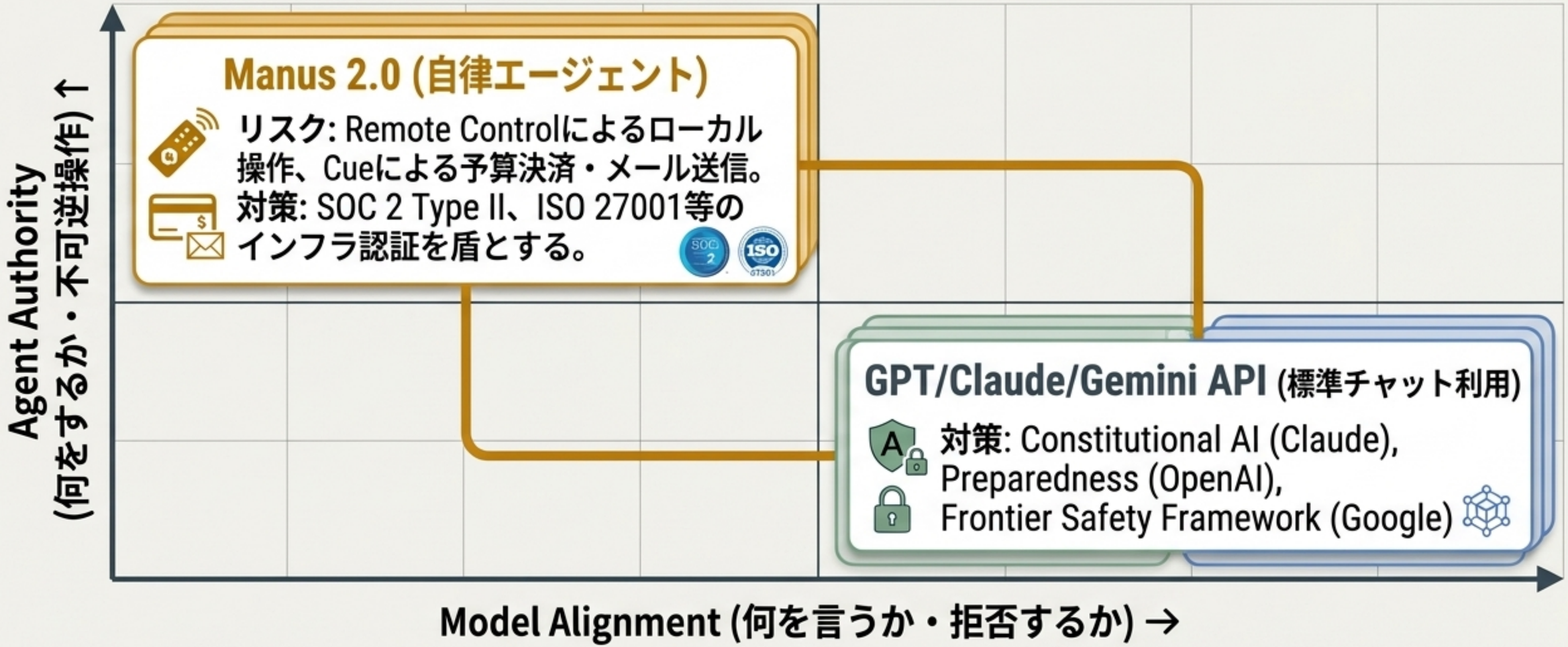
基盤モデルAPI 比較マトリックス（自社開発レイヤー向け）

GPT-6.1 Sol	Claude Sonnet 5.5	Gemini 3.8 Flash
位置付け: 知能・コストのバランス型 （Astraの技術を踏襲）	位置付け: 速度・知能バランス型 （Agentic/Tool workloads特化）	位置付け: 高速・大量処理・長文脈 （2026年末まで特別価格）
コンテキスト: 1.05M Input / 128K Output	コンテキスト: 1M Input / 128K Output	コンテキスト: 1M Input / 65K Output
API料金（1M）: Input $2 / Output $10	API料金（1M）: Input $2 / Output $10 （Cache read $0.20）	API料金（1M）: Input $0.75 / Output $3.75 （※2027年以降値上げ予定）
外部統合: Responses API中心。 Web search, File search, Computer use, MCP。	外部統合: Tool use, Computer use, Claude Code。	外部統合: Function calling, Code execution, Search Grounding標準搭載。

※各社APIレイヤーにおいてもエージェント機能（Tool/Computer use等）の標準サポートが進行している。

Gemini Notebook

安全性の2つのベクトル：アライメント vs. アクション権限



エージェント導入の最大リスクは「暴言」ではなく「不可逆操作（削除・決済・本番Deploy）の誤実行」である。Human-in-the-loopの設計が必須。

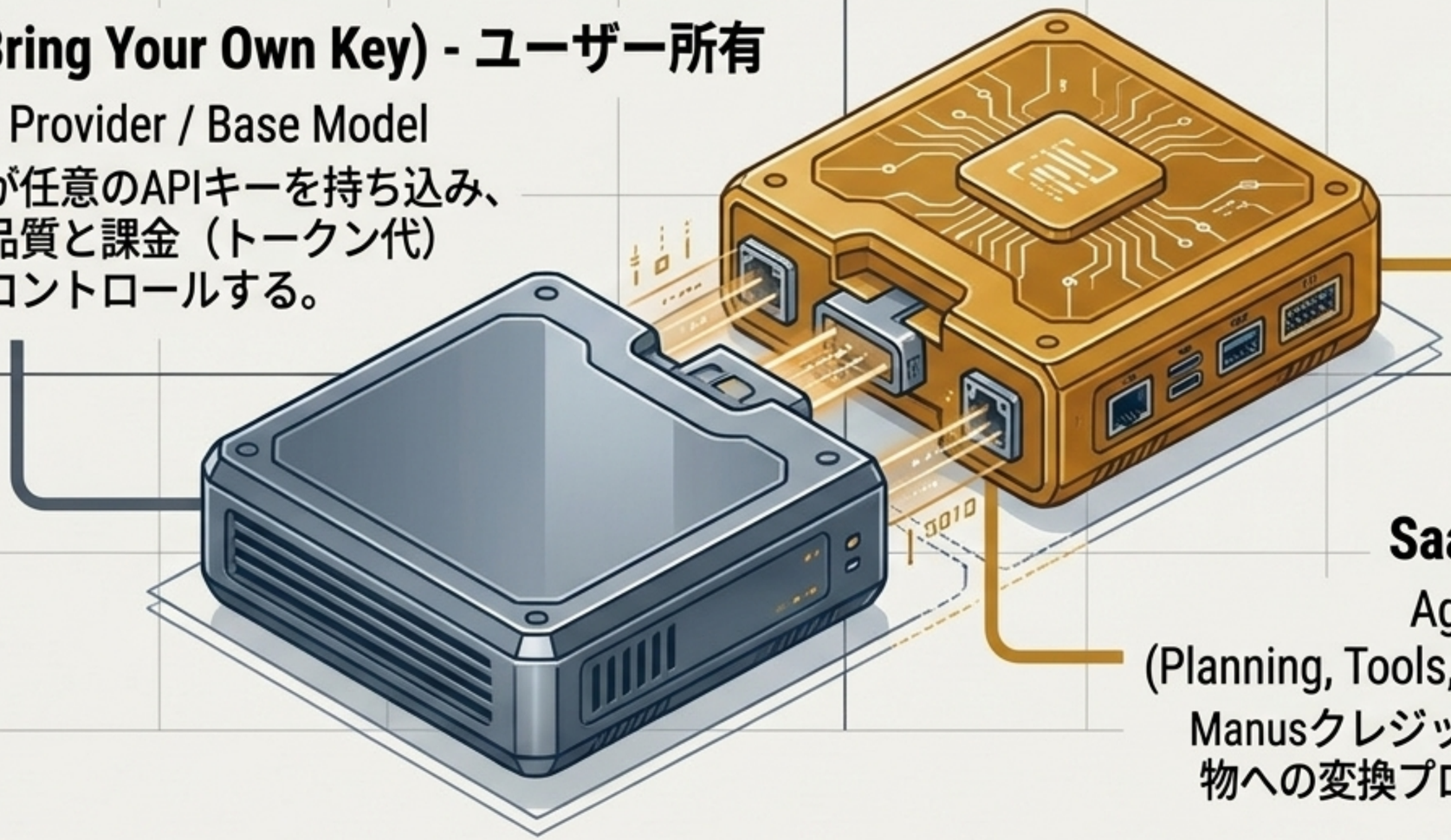
Gemini Notebook

第3の選択肢「Manus Flex」：推論とインフラのハイブリッド

BYOK (Bring Your Own Key) - ユーザー所有

Inference Provider / Base Model

ユーザーが任意のAPIキーを持ち込み、
モデルの品質と課金（トークン代）
を自社でコントロールする。



SaaS - Manus提供

Agent Infrastructure
(Planning, Tools, Cloud Computer)

Manusクレジットを消費し、成果
物への変換プロセスを委譲する。

**「モデル品質は自分で選びたいが、エージェント基盤をゼロから保守したくない」
というエンタープライズのジレンマを解決する構造的アプローチ。**

Gemini Notebook

エンタープライズ導入判断フローチャート

要件（Requirement）	推奨（Recommendation）＆理由（Reason）
「調査・操作・資料化」という業務プロセス全体を丸投げしたい。	**推奨: Manus 2.0** （Agent基盤の開発コストをゼロにする）
自社SaaSの内部に深くAIを組み込み、トークンやツールを厳格に制御したい。	**推奨: GPT / Claude API** （制御性と透明性を最優先）
莫大な長大文書処理、またはネイティブ動画の直接解析を低コストで行いたい。	**推奨: Gemini 3.8 Flash API** （明文化されたマルチモーダル仕様とコスト優位性）
基盤モデルは自社の基準で選びつつ、エージェント実行環境だけ利用したい。	**推奨: Manus Flex** （ハイブリッド構成）

Gemini Notebook

PoCの再構築：計測すべき真のKPI



API Integration (自社開発の場合)

- First-token latency / Output throughput
- Reasoning effortに応じた動的コスト変動率
- 長文脈コンテキストにおける情報検索精度 (Needle in a haystack)

Agent Delegation (Manus採用の場合)

- Task completion rate (タスクのEnd-to-End完遂率)
- Human intervention count (人間の介入・修正が必要だった回数)
- End-to-end task cost (インフラ・ツール実行費を含めた最終コスト)

Gemini Notebook

結論と提言（The Final Synthesis）

「AI技術の選択は、『誰が最も賢いか』という単一のレースから、AIスタックの『どの層を自社で所有するか』というアーキテクチャ設計へと移行した。」

AIに「回答」を求め、
自社システムへ組み込むなら
→ GPT / Claude / Gemini

AIに「仕事の完了」を求め、
基盤構築を省くなら
→ Manus 2.0

「制御性」と「利便性」を
両立させるなら
→ Manus Flex



【日本企業へのデータガバナンス注記】無料版と有料版（API/Enterprise）ではデータ学習利用の規約が明確に異なる（例：Gemini API Additional Terms等）。PoCおよび本番導入時には、一般消費者向け規約ではなく、必ず有償APIプランの当該時点のデータ処理条項を基準に選定すること。

Gemini Notebook