Gemini 4 Hydrogen 徹底解剖レポート

実務導入のための真実・独立評価・戦略的含意

TARGET:	gemini-4-hydrogen
SOURCE:	Claude Fable 5.2 / Independent Benchmarks
STATUS:	Verified Active (2026-11-13)

Gemini Notebook

STATUS WINDOW 01

MODEL: GEMINI 4H | STATUS: ACTIVE | DATE: 2026-11-12

1. 実在と世代交代

「Gemini 4 Hydrogen」は実在し、既に提供開始（2026年11月12日）。事実上「3.8 Flash」の完全な後継モデルであり、日常タスクの主役（Daily Driver）となる。

STATUS WINDOW 02

METRIC: SPEED X2.5 | COST X2.0 | CTX: 1M TOKENS

2. 速度とコストの等価交換

タスク完了速度は「2.5倍」に向上し、1Mトークンの長文脈を維持。　しかし、API単価は3.8 Flashの「2倍」に設定されている。

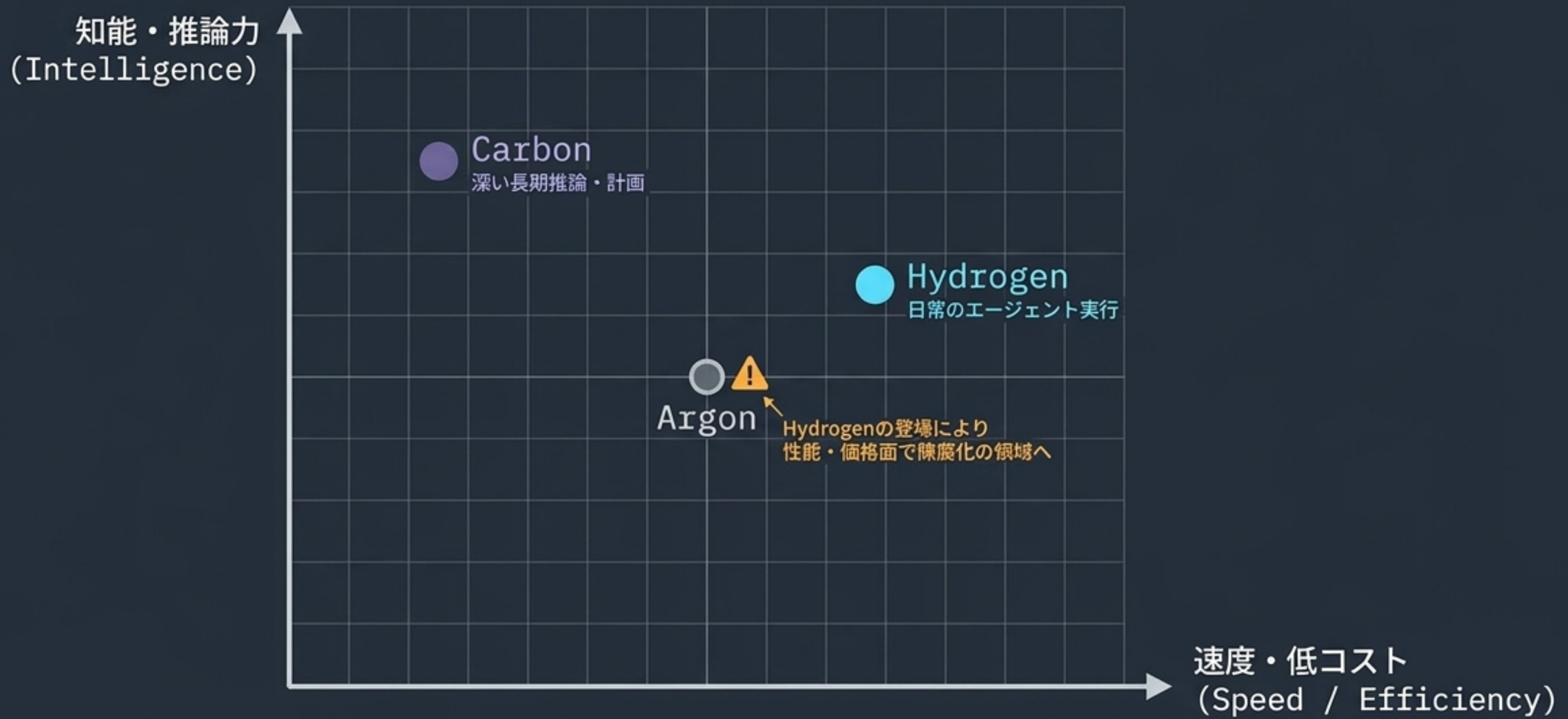
STATUS WINDOW 03

DEPLOY: AGENT SPLIT | OPTIMIZATION: COST/EFFICIENCY | ALERT: VERIFY LEGAL CLAIMS

3. 企業導入の最適解

最上位モデル（Carbon）を計画役とし、Hydrogenを実行役とする「分業型エージェント運用」が最も費用対効果が高い。 法務用途でのメーカー公式主張には独自の検証が必要。

Gemini Notebook



インサイト：Hydrogenは単なる軽量モデルではなく、中位クラスの知能を軽量級の速度で提供するゲームチェンジャーである。

Gemini Notebook

処理能力（Capacity）



- 最大1M（1,048,576）トークン コンテキスト
- 同期出力：最大64Kトークン
- 知識カットオフ：2026年3月

新機能（Thinking Levels）



4段階の思考レベルを搭載：

low
medium
high（既定）
xhigh（新設）

価格体系（Pricing - API標準）



- 入力: $1.50 / 100万トークン （キャッシュ入力: $0.075）
- 出力: $7.50 / 100万トークン

提供プラットフォーム（Availability）



以下で一般提供（GA）済み：

- Gemini API
- AI Studio
- Google Cloud
- Gemini for business 等

Gemini Notebook

Performance Gains（メリット）

- タスク完了速度：2.5倍高速化 ↑
- Terminal-Bench 4.1：18.2% から 57.0%へ急上昇 ↑
- 事実誤認（幻覚）：1/5未満に大幅削減 ↑
- 安全性：画像からテキストへの安全性が+7.00pp改善 ↑

Cost & Trade-offs（デメリット）

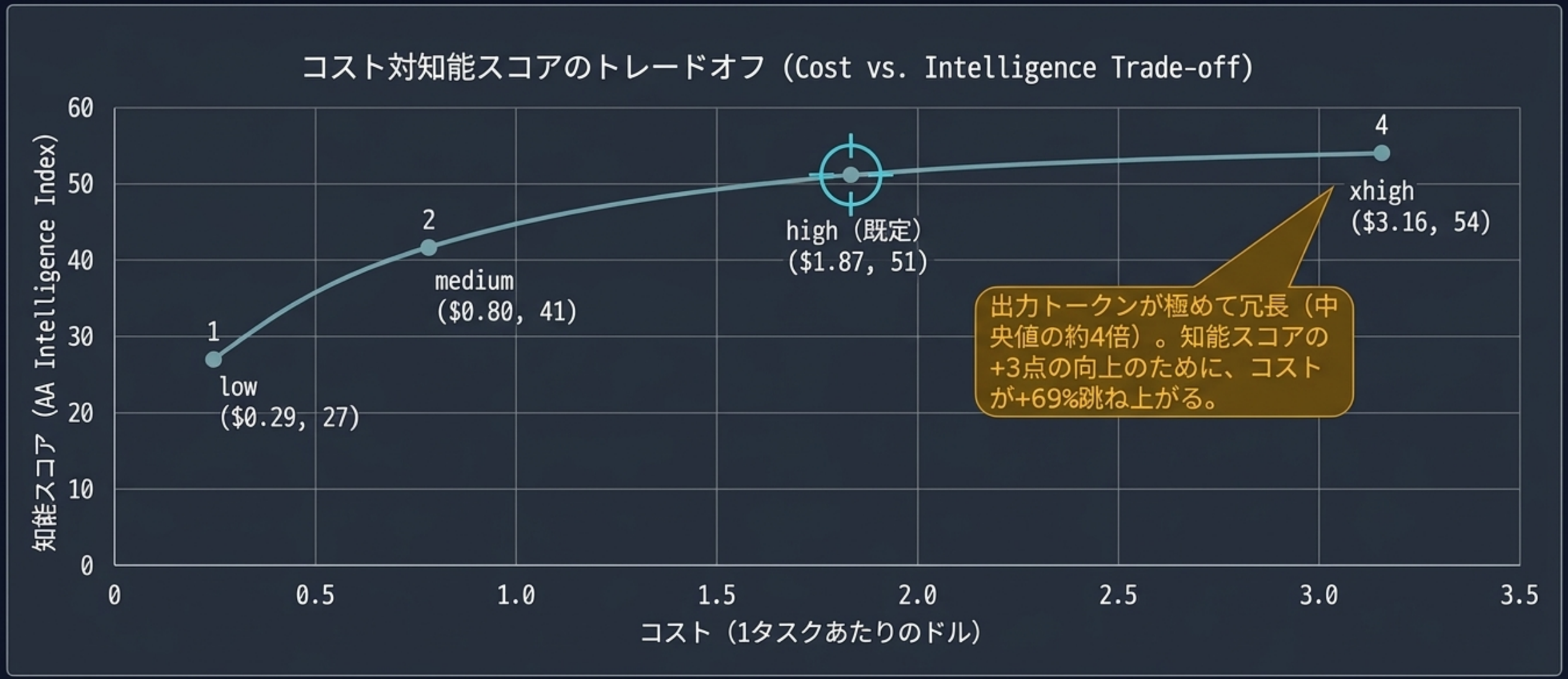
- ⚠ API単価：3.8 Flashの2倍（入力$1.50/出力$7.50） ⚠
- ⚠ 不当な拒否率：マルチモーダルで +5.57pp悪化 ⚠



診断結果：時間（速度と精度）を金（コスト）で買う設計。大量バッチ処理よりも、リアルタイム応答が求められるエージェント業務で真価を発揮する。

Gemini Notebook

Telemetry Dashboard



キーインサイト：パレート最適の限界点。
実務におけるスイートスポットは既定の「high」であり、xhighの常用は費用対効果が著しく悪い。

Gemini Notebook

Google公式の主張（ベストケース）

- 対Claude Sonnet 5.5で「7勝3敗」をアピール
- OSWorld-2.1: 77.6% (vs 83.9% Sonnet)
- Agents' Last Exam: 71.6% (vs 60.4% Sonnet)

独立評価の現実 (Arena / Artificial Analysis)

- Arena Agent: 総合5位。しかし上位10モデル中で「最速」（タスク中央値5.3分）
- Terminal-Bench 4.0: 54%。Sonnet 5.5の64%には明確に劣る
- AA-Omniscience: 作り話の率は約8%（幻覚の少なさは本物）

統合メッセージ：コーディング等の特定領域では最上位モデルに及ばないが、エージェント用途における「速度とコストのバランス」では現在市場トップクラスの実力。

Gemini Notebook

データドリブンなベンチマーク比較：注意報
(Data-Driven Benchmark Comparison: Alert)



- 上位6モデルは誤差2%以内にひしめいており、公式発表にはチェリーピッキング（都合の良いデータ抽出）の疑いがある。
- 特許・知財分野に特化した公開データは存在しない。法務用途でのメーカー公称値の鵜呑みは極めて危険。
- 対策：過去の先行技術調査やクレーム対比を用いた、自社データによる検証プロセスが必須。

Gemini Notebook

中位モデル（Mid-Tier）競合比較

	Gemini 4 Hydrogen (high)	Claude Sonnet 5.5	GPT-6.1 Terra
知能スコア	51	55	51
1タスクコスト	$1.60	$2.00+	$0.96
速度	最速（236 tokens/s）	中	中

用途別勝者（診断結果）

複雑なコーディング / UI操作 → Claude Sonnet 5.5	コスト最優先の大量バッチ処理 → GPT-6.1 Terra	超高速エージェント実行 / Workspace連携 → Gemini 4 Hydrogen

Gemini Notebook

動的モデルルーティング: 最適化パイプライン (Dynamic Model Routing: Optimization Pipeline)

Step 1: 複雑な計画 (The Brain)

【Gemini 4 Carbon】

深い推論を行い、タスクを分割・計画する。

Step 2: 動的ルーティング (The Switch)



【Orchestrator (Antigravity)】

セッション内で**モデルを最適に切り替え、クォータ消費を24%削減。**

Step 3: 高速実行・並列処理 (The Hands)

【Gemini 4 Hydrogen】

ファイル読み込みと検索を並列実行し、高速にタスクを完了させる。

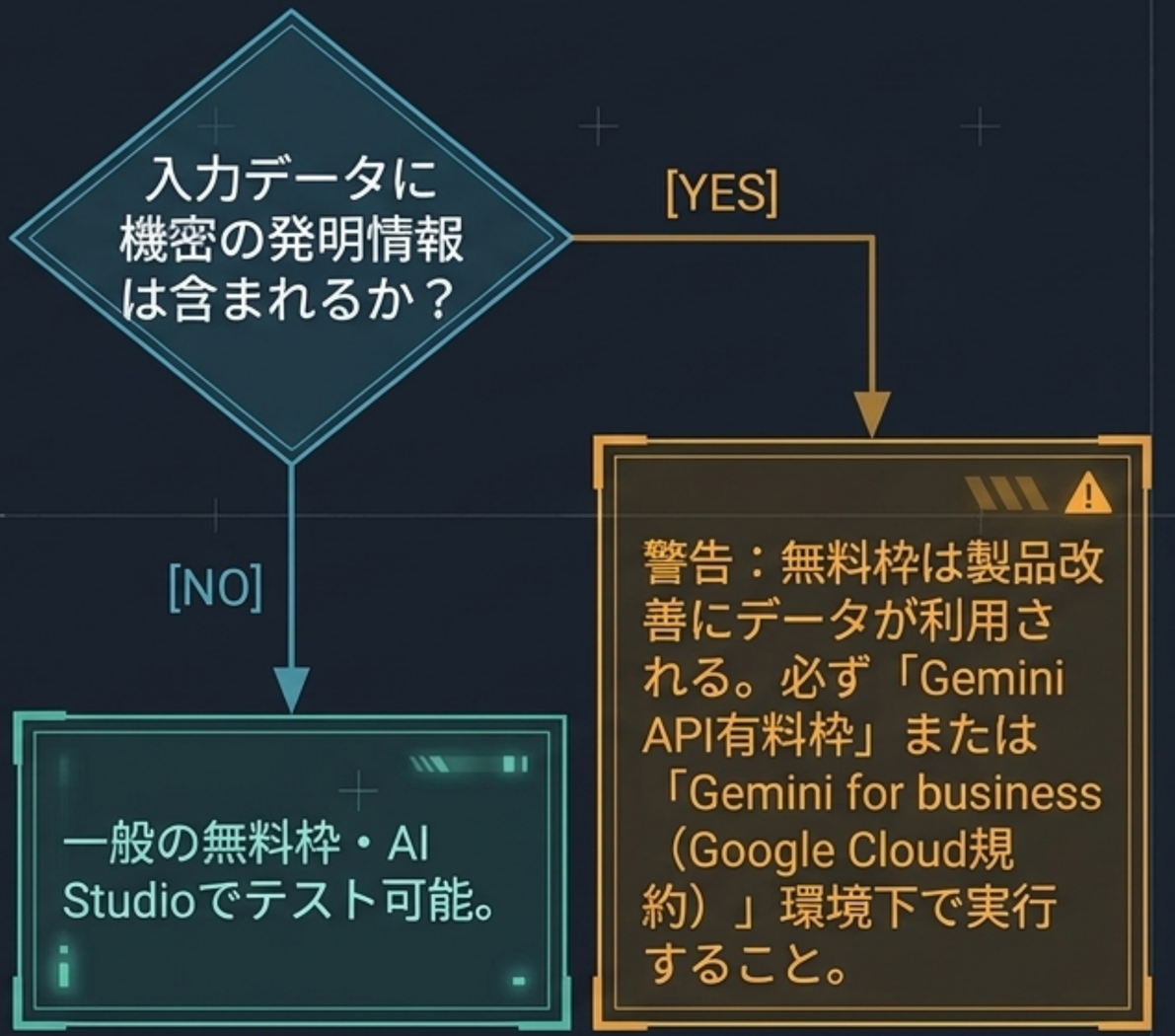
解説：単一モデルに全てを依存するのではなく、役割に応じてモデルを動的に切り替えることで、「Carbonの精度」と「Hydrogenの速度・低コスト」を両立させる。

Gemini Notebook

知財・法務での実践的ユースケース（Strengths）

1Mコンテキストを活かした先行技術文献の一括読解。

RAG（検索グラウンディング）を利用した最新公報の要約（※知識カットオフが26年3月のためRAG必須）。



Gemini Notebook

Timeline to Jan 2027



3.8 Flashからの移行チェックリスト（Migration Action Plan）

[ ] 思考レベルの再設定: 既定がmediumから high に変更されたことを確認。
[ ] エラー回避: thinking_budget と MINIMAL 指定を削除（HTTP 400エラーの原因）。
[ ] パラメータの無効化認識: temperature, top_p, top_k は仕様上無視される。
[ ] キャッシュ戦略: 暗黙キャッシュは4,096トークンから有効になることを組み込む。

Gemini Notebook

Action 1: 移行テストの即時開始

2027年1月のFlash値上げに備え、本番ワークロードでHydrogen（medium/high）のコストと時間を実測する。



Action 2: エージェント分業の設計

「難しい判断はCarbon、実行はHydrogen」というAntigravity（Orchestrator）前提のアーキテクチャへ移行する。



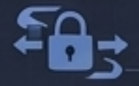
Action 3: コスト最適化の再評価

コスト最優先のバッチ処理については、同等知能で安価なGPT-6.1 Terra（$0.96/タスク）との比較検討を行う。



Action 4: 法務データの防衛と検証

Googleのベンチマークを鵜呑みにせず、自社の過去案件（正解データ）を用いて独自評価を実施する。機密データは有料枠に限定する。⚠



TELEMETRY END. AWAITING EXECUTIVE ACTION.（テレメトリー終了。経営判断を待機中。）

Gemini Notebook