

次世代画像AIの実践的評価： GPT Image 2 vs Nano Banana 2

単なるスペック比較を超えた、クリエイティブと実務のための選定・統合ガイド

2026.05 / AI Strategy & Creative Innovation Dept.

設計思想の明確な分岐：アトリエの「確実性」か、研究所の「高速展開」か

GPT Image 2

「完成度と論理の追求」

文字、レイアウト、編集の確実性を極める設計。最終的な誌面化や商用ポスターなど、細部の破綻が許されない「仕上げ」の領域で絶対的な優位性を持つ。

Nano Banana 2

「高効率と外部情報の統合」

低遅延、大量反復、そして「検索接地 (Search Grounding)」を前提とした設計。実在物の確認や、情報量の多い濃密な背景、4Kや極端なアスペクト比での高速試作に優れる。

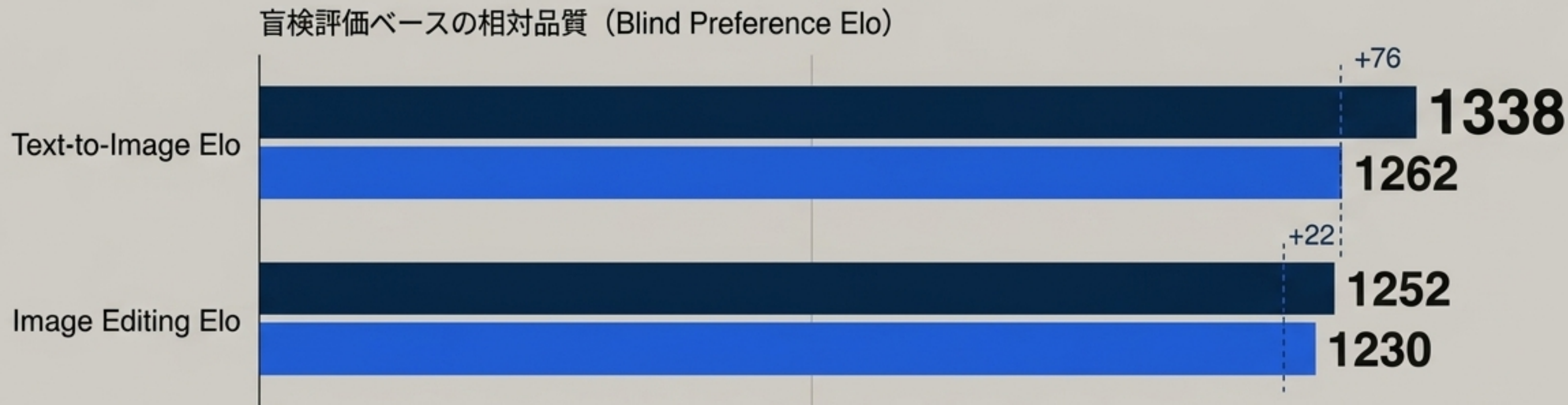
コア・インサイト

単一の勝者は存在しない。目的が「クライアント向け最終成果物」であればGPT、「内製の高速反復と情報探索」であればNanoが最適解となる。

俯瞰的アーキテクチャ比較： 実務導入を左右するコア・スペック

	GPT Image 2	Nano Banana 2
API / 提供チャンネル	ChatGPT, Image API, Responses API, Batch	Gemini, AI Studio, Gemini API, Vertex AI
解像度と比率の自由度	任意サイズ 最大エッジ3840px / 比率3:1まで	固定プリセット 0.5K, 2K, 4K / 比率1:4~8:1まで (16:9ネイティブ対応)
生成品質の方向性	総合選好・タイポグラフィ 複雑な視覚テキスト、高忠実度編集	検索接地・高密度指示 長文テキスト、Dense prompt alignment
基本速度のスタンス	Quality "low"で最速化 (複雑なプロンプトは最大2分)	"low latency / high-volume" を公式に強調 (数秒クラス)

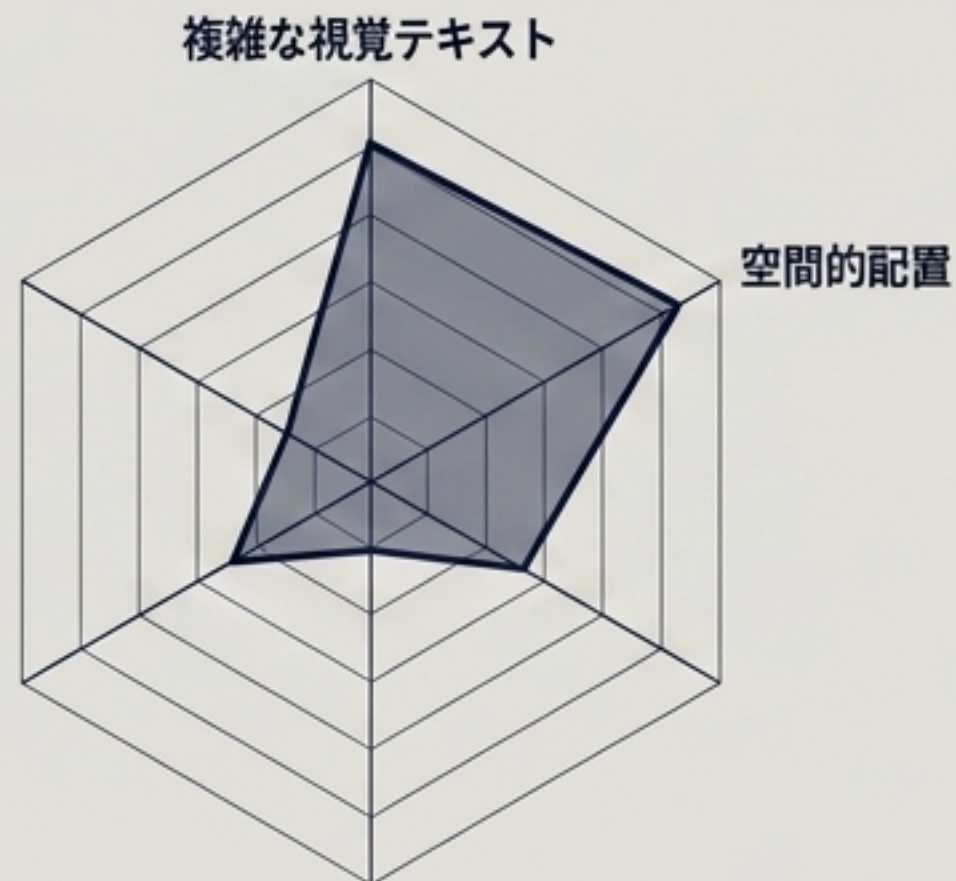
ユーザー選好における純粋な生成・編集力はGPTが先行する



ブラインドテスト（盲検評価）および「Text & Typography」リーダーボードにおいて、GPT Image 2は一貫して上位を獲得。新規生成の美しさだけでなく、既存画像に対する**「編集（Editing）の安定性」**において、実務上の大きなアドバンテージを持っている。

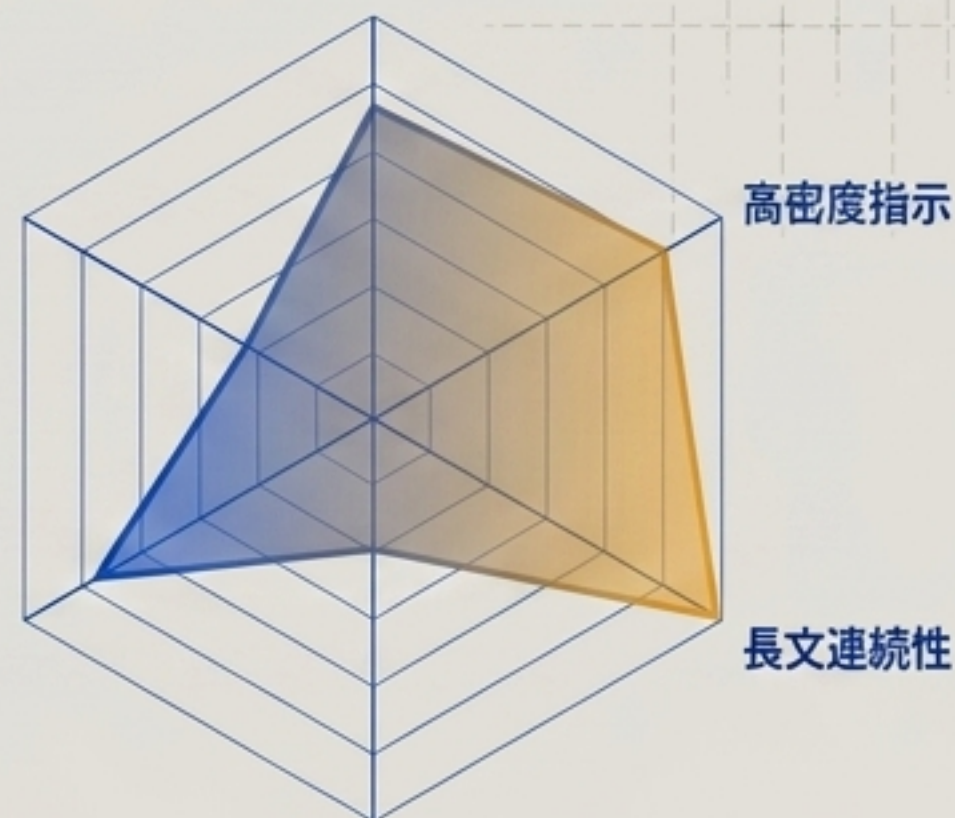
「論理的整合」と「絵としての密度」のトレードオフ

【GPT Image 2】 構造と文字の支配力



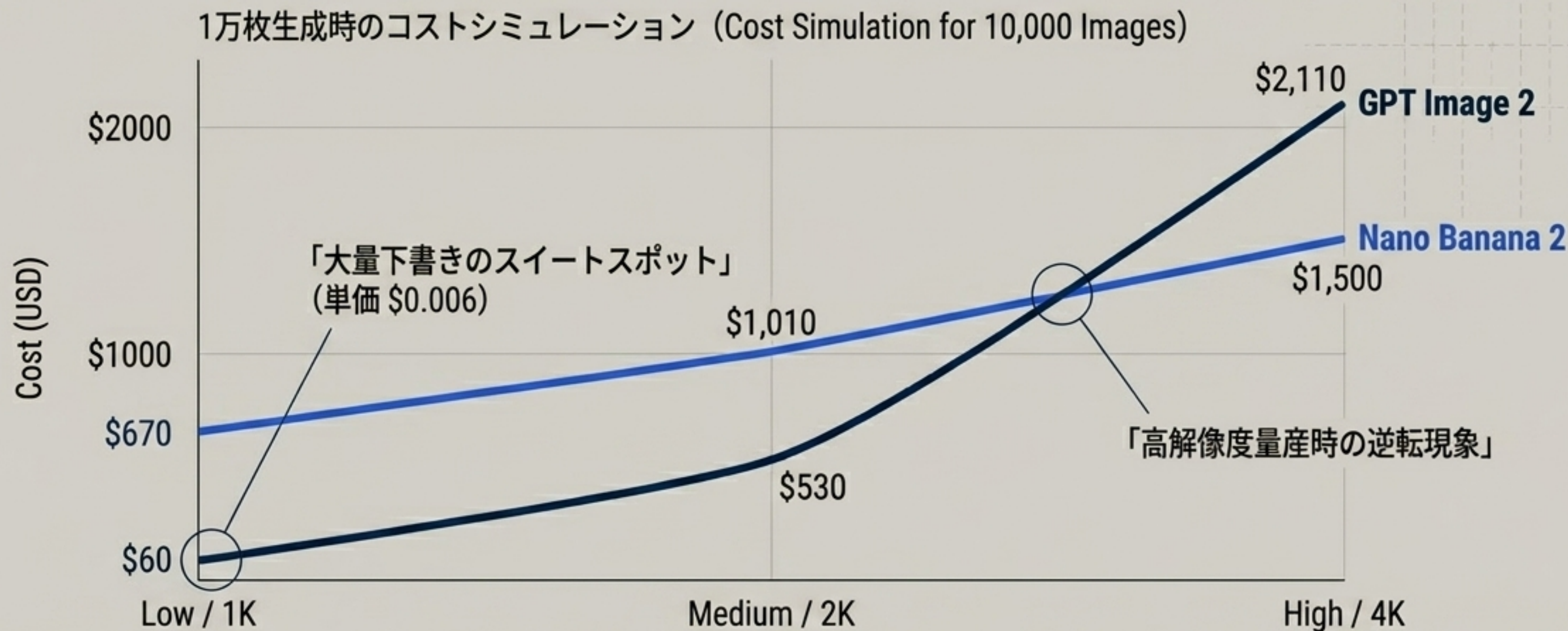
- **優位領域:** 複雑な視覚テキスト (CVTG: 0.9003), 空間的配置 (GenEval: 0.89), 広告・UIデザイン。
- **特徴:** 指示の論理的整合性 (Logic) が高く、指定した要素が正確なレイアウトとタイポグラフィで出力される。ポスターや図表に強い。

【Nano Banana 2】 情報量と連続性の支配力



- **優位領域:** 高密度な指示への追従 (DPG-Bench: 86.90), 長文テキストの連続レンダリング (LongText-Bench: 0.980)。
- **特徴:** 背景・前景の「絵としての情報量 (密度)」や物語性の表現に長ける。複雑で長い条件文を破綻なく一枚の絵に落とし込む能力が高い。

コスト・ダイナミクス：生成解像度によって生じる「価格の逆転」

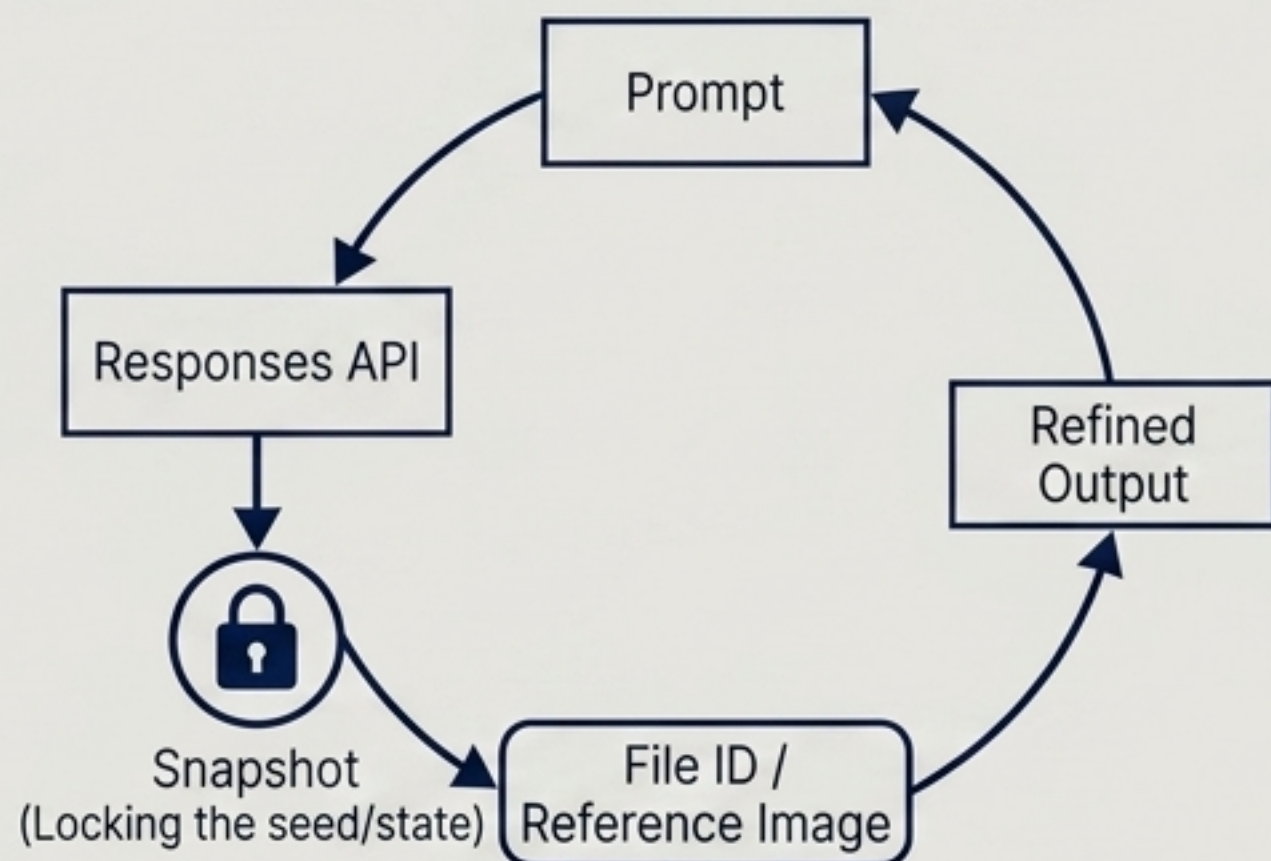


単純な単価比較では本質を見誤る。ラフ案を大量に生成する初期フェーズではGPTの「low」が圧倒的なコスト優位性を持つが、最終アセットレベルの高解像度（4K級）を量産するフェーズに入ると、Nano Banana 2の価格体系のほうが予測可能で低コストになる。

システム制御アプローチ： 再現性のAPIか、事実確認の検索拡張か

GPT Image 2

Snapshot & Multi-turn Architecture



Responses APIとSnapshotsの組み合わせにより、生成結果を固定しながら多段階の編集を行うことが可能。A/Bテストや研究、厳密な回帰テストに不可欠な「再現性」を担保する。

Nano Banana 2

Google Search Grounding Architecture



google_searchとのネイティブ連携により、外部の事実（風景、実在の物体）を接地させた画像生成が可能。インフォグラフィックや旅行・地域特化のクリエイティブにおいて、事実誤認（ハルシネーション）の摩擦を大幅に低減する。

ガバナンスとデータプライバシー： 企業導入におけるリスク評価

透明性と安全性 (Transparency & Safety)	Nano Banana 2
✓ 公表System Cardあり。C2PAおよび不可視可視ウォーターマーク実装。	⚠ 単体のSystem Cardは未確認。 SynthIDウォーターマーク実装。
データ利用・学習ポリシー (Data Usage Policy)	
✓ API入力は原則学習利用なし（オプトイン方式）。Abuse logは通常30日保持、ゼロデータ保持の承認ルートあり。	⚠ 無償ティアでは入力・出力が製品改善に利用される可能性あり。有償ティアは学習ティアは学習利用なし（No）。
出力の商用利用と所有権 (Commercial Ownership)	
利用者が Output を所有（非一意性には留意）。	Googleは所有権を主張しないが、同一出力の可能性と「法令遵守の責任」を利用者側に強く明示。

機密性の高い未公開クリエイティブを扱う場合、Nano Banana 2は必ず「有償ティア」を前提にアーキテクチャを設計する必要がある。

ユースケース別・最適アプローチのマッピング

商用広告・販促物

推奨: GPT Image 2

根拠: タイポグラフィと複雑な視覚テキストにおける優位性。クライアント提出前に文字崩れやレイアウト破綻を排除できる。

大量バリエーション生成

推奨: Nano Banana 2 (またはGPT low)

根拠: 高効率・低遅延設計により秒単位の反復が可能。ただし、品質を落としてもよい下書き用途ならGPTのLowも極めて低コスト。

アート制作・世界観探索

推奨: Nano Banana 2

根拠: 高密度なプロンプトへの追従性。背景の情報量や物語性の濃いアセットを4K解像度で一気に引き出すことができる。

研究・プロトタイピング

推奨: プロジェクト依存

根拠: 「実験の再現性 (Snapshots)」重視ならGPT。既存の「Google Stackとの連携」や検索拡張を重視するならNano。

究極の実戦解：「着想はNano、仕上げはGPT」のハイブリッド・ワークフロー

STEP 1: 情報探索と構図の量産 (Powered by Nano Banana 2)

- **Action:** 検索接地 (Search Grounding) を活用し、実在情報の裏付けを取りながら、極端なアスペクト比や4K解像度でアイデアのパリエーションを高速に大量出力する。
- **Goal:** 世界観と構図の確定。



STEP 2: キュレーションと抽出 (Human in the Loop)

- **Action:** 生成された大量のベース画像から、意図に最も近いレイアウトと情報量を持つものを参照画像 (Reference Image) としてピックアップ。

STEP 3: 最終レイアウトとタイポグラフィ (Powered by GPT Image 2)

- **Action:** ピックアップした画像をベースに、Responses APIを用いた多段階編集を実施。正確な文字組み、ロジカルなレイアウト補正、不要なアーティファクトの除去を行う。
- **Goal:** クライアント提出水準の完成品 (Production-ready Asset)。

意思決定の最終基準：「自社はどの失敗を最も嫌うか？」

「文字崩れ、誌面崩れ、手戻り」を嫌う

→ GPT Image 2 を主軸に据える

アクション: Image APIとResponses APIの使い分けを検証し、Snapshotsを用いた出力の安定化プロセスを社内に構築する。

「反復速度の遅さ、事実誤認、単調な背景」を嫌う

→ Nano Banana 2 を主軸に据える

アクション: Vertex AI等の有償環境を整備し、google_search連携を利用したファクトベースのプロトタイピングを試験導入する。

Next Action: 自社のプロンプト長、想定解像度、編集回数に基づき、両モデルのAPIコストを再試算し、POC（概念実証）のフェーズへと移行する。