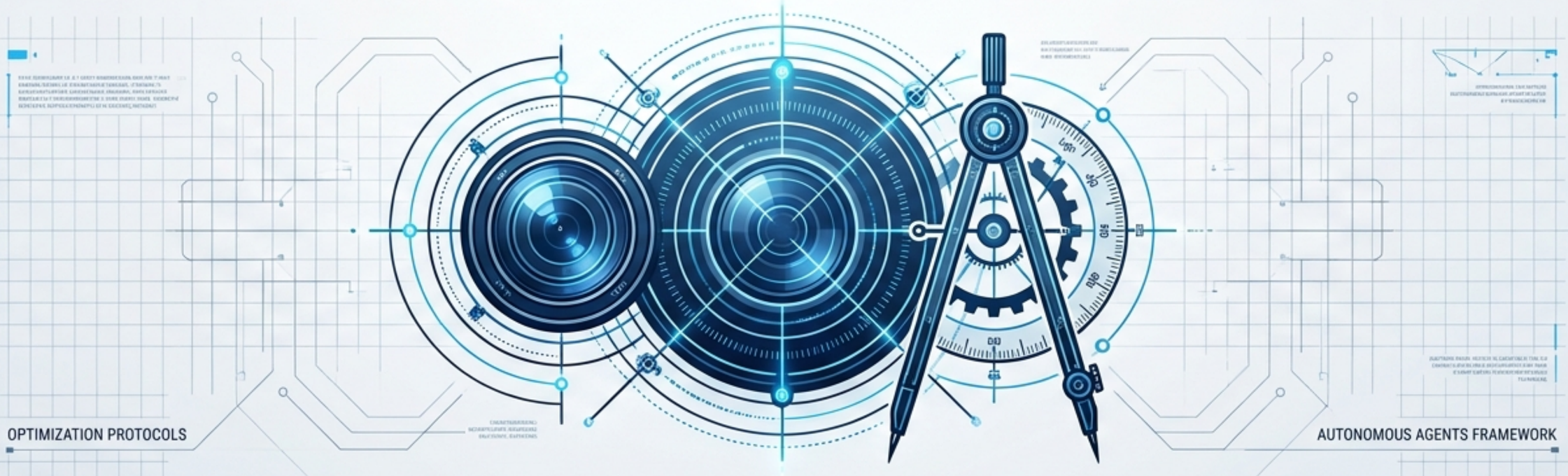
Gemini 3.7 Flash 徹底解剖：次世代ワークホースAIの産業的インパクトと実装戦略



OPTIMIZATION PROTOCOLS

AUTONOMOUS AGENTS FRAMEWORK

「チャット」から「自律型エージェント」へ。
TCO最適化とタスク完了率を極めるエンタープライズAIの最新見取り図。

基礎学習のやり直しを廃し、「アルゴリズム改修」で達成した最速の進化

従来モデルの限界

- 指示の曖昧さによる途絶
- 手戻りの多発

異例の3週間

- Gemini 3.6 Flashからわずか3週間
- コア推論基盤へのアルゴリズム改修
- 開発者フィードバックの直接適用

エージェント的ワークフロー

- 複数工程の自律遂行
- プロダクション環境での実行
  完了率獲得

最難関課題を解くための超大型モデルではなく、日常業務の低コスト・高精度な自動化を支配する「ワークホース（主力実務モデル）」への劇的なシフト。

日常業務の低コスト・高精度な自動化を支配する「ワークホース」の確立

戦略的意義

「単発の応答 (Chat)」から「プロダクション環境での実行完了率 (Execution)」へターゲットを完全シフト。

100万トークンを直接解釈し、即座にアクションへ変換するマルチモーダル処理基盤

アーキテクチャの核心：タスクの難易度に応じて「推論の深さ」を直接制御する

LOW

リアルタイム性重視。
即答タスク・単純な
テキスト抽出向け。

MEDIUM
(Default)

速度と推論の
バランス型。

HIGH

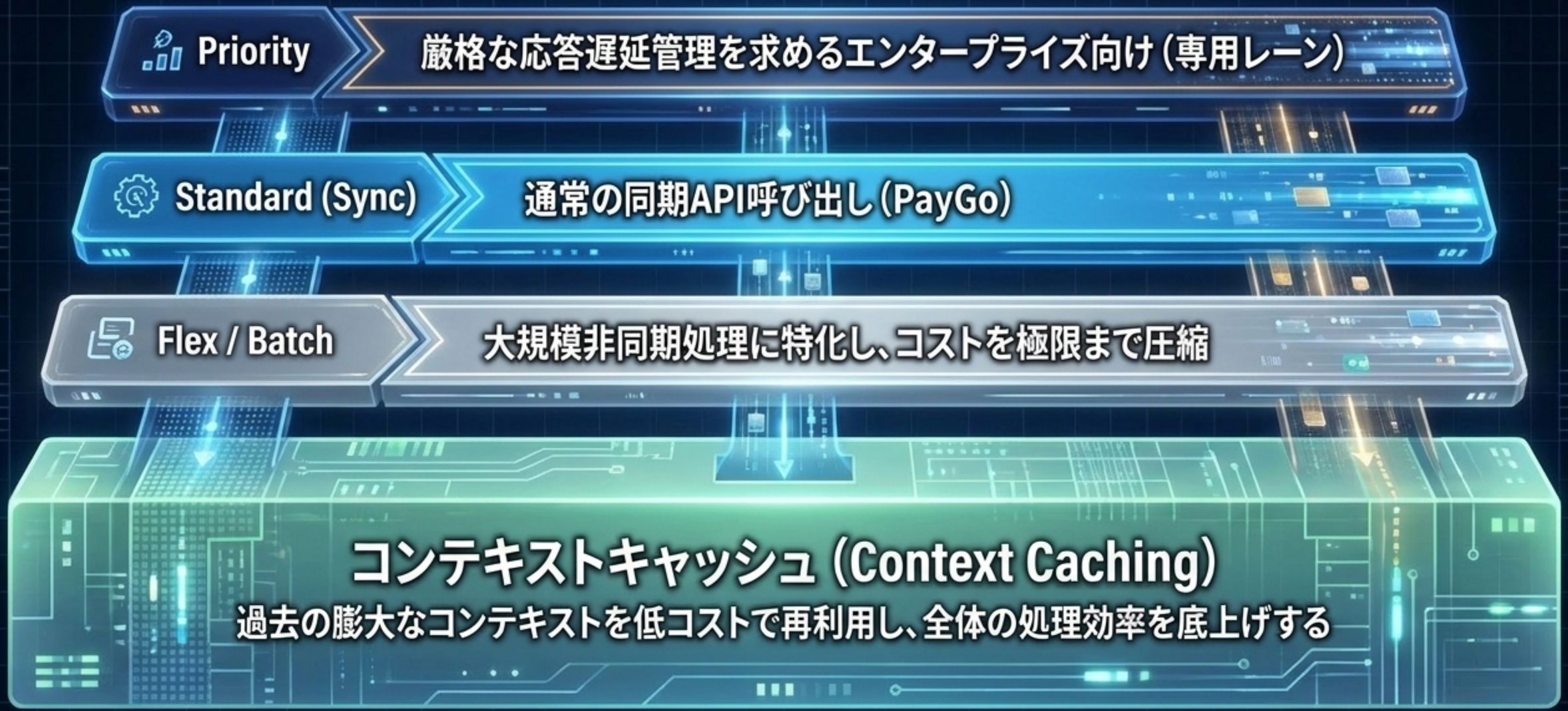
複雑なプログラミング、
多段階の自動計画
立案向け。

MINIMAL

非対応・エラー返却

モデルが自律的な安全評価や文脈補正
を行うための最低限の推論領域を強制
確保しているため、Minimal指定は不可。

用途と遅延許容度に応じた3つのAPI推論経路とコンテキスト再利用



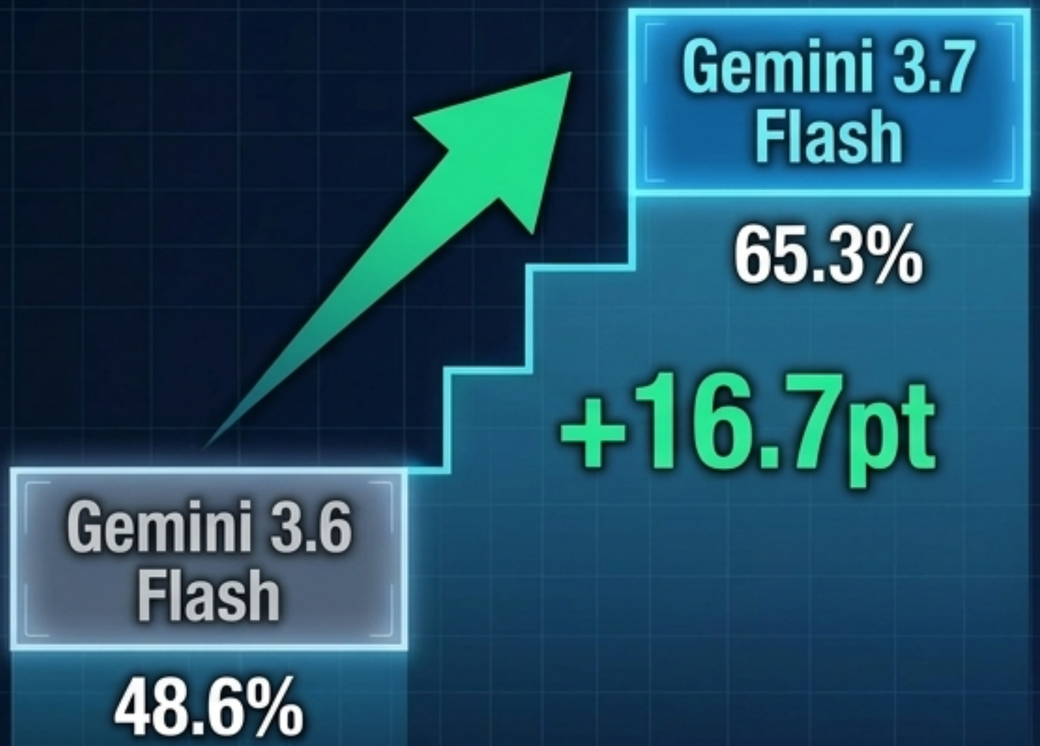
業界標準モデルを凌駕する「エージェント・自動化」領域の圧倒的パフォーマンス

評価指標	Gemini 3.7 Flash	Gemini 3.6 Flash	Claude Sonnet 5	GPT-5.6 Terra
【総合知能】				
AI Index	56	52	55	57
【コーディング】				
FrontierCode 1.1	43.6%	34.4%	42.7%	41.3%
DeepSWE v1.1	65.3%	48.6%	53.8%	69.6%
Code Arena	1,588	1,538	1,541	1,523
【エージェント】				
Terminal-bench 2.1	85.8%	78.0%	80.4%	87.4%
AutomationBench	30.4%	17.0%	10.7%	23.6%
OSWorld-2.0	47.9%	33.8%	—	50.2%
Harvey LAB-AA	90.7%	85.1%	90.1%	85.2%

特にDeepSWE（実用コード生成）とAutomationBench（複合業務自動化）において
旧バージョンおよびClaudeを劇的に引き離しており、「自律型実務モデル」としての地位を確立。

「動くコード」から「実用に耐えるコード」へ。初回精度（First-pass Accuracy）の劇的向上

DeepSWE v1.1 スコア

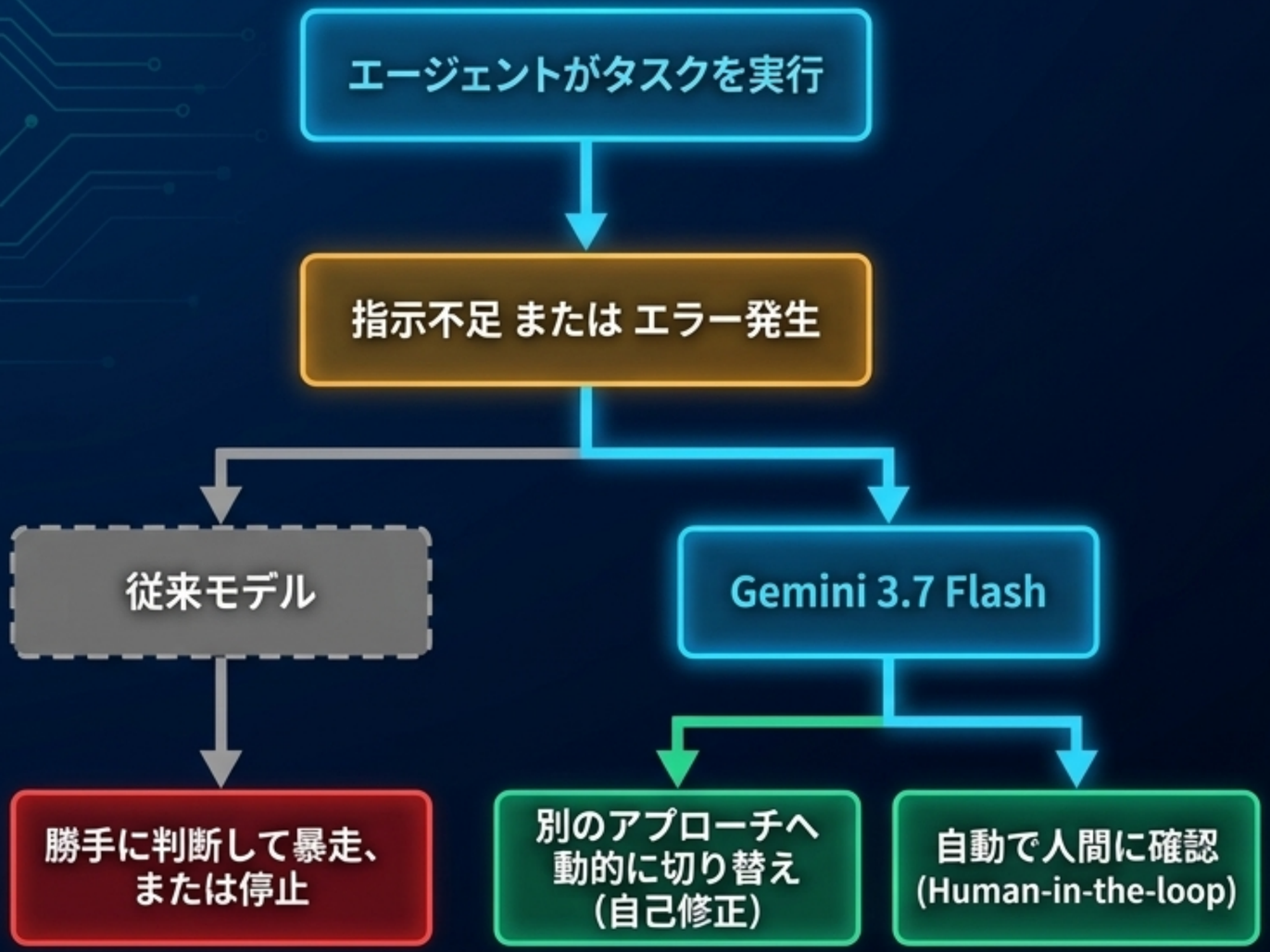


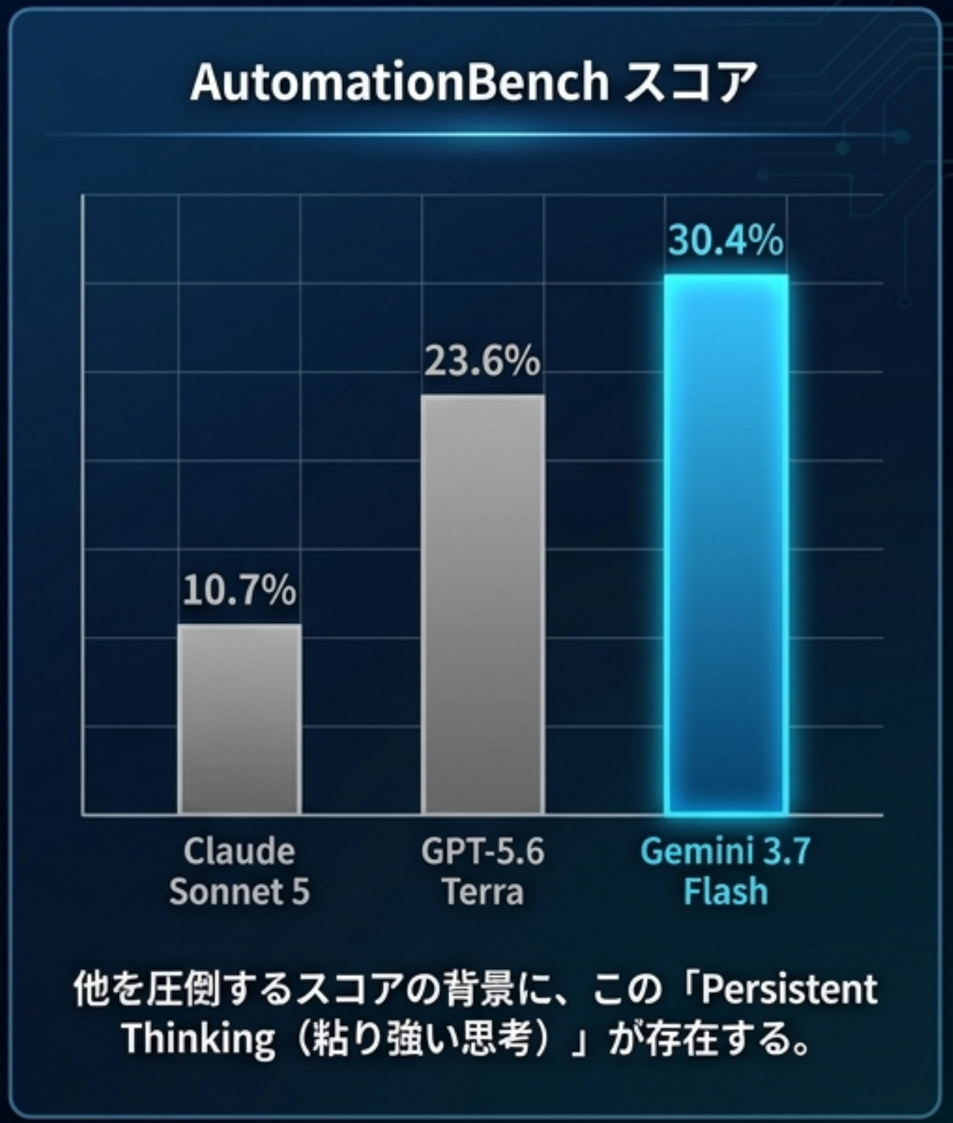
なぜ初回精度が向上したのか？

1. 複数ファイルにわたる依存関係の正確な把握
2. 既存のリポジトリ命名規則・レイアウトへの忠実な準拠
3. エラー処理パターンの網羅とレスポンス幅の遵守

人間（エンジニア）がコード修正に費やす「隠れた手戻りコスト」を根絶する。

エラーで暴走しない「粘り強い思考」が実現する複合業務の自律完遂





検証ハードルを下げるプロモーション価格と「2027年問題」への対応

TCOの罠を回避せよ：「思考トークン(Thinking Tokens)」という隠れた変数

[最終出力テキスト] + [思考トークン] = [実際の課金対象（出力料金）]

「High」レベルの乱用は実質的な請求金額の非線形な爆発を招く。タスク難易度に応じた思考レベルの緻密な制御が不可欠。

Premium Technical Blueprint

個人環境からエンタープライズ統合までを網羅するデプロイメント・エコシステム

開発者・インフラ

- Google AI Studio
- Gemini API
- Google Antigravity
- Android Studio

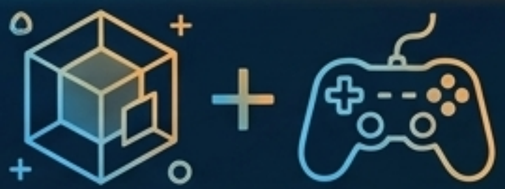
エンタープライズ

- Gemini Enterprise App
- Gemini Enterprise Agent Platform
- トグルスイッチ一つで組織全体へデプロイ可能

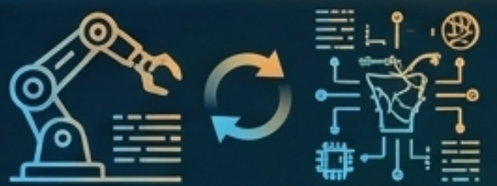
ユーザー (Consumer/Pro)

- Gemini Spark
  パーソナルAIエージェント、24時間横断作業

静的PDFから動的Web
ダッシュボードへの一括変換



Nano Banana連携による
3Dゲーム空間の動的生成



マルチモーダル視覚推論を組み込んだ
ロボティクス・フィードバックループ

高度化する自律能力と、徹底されたセーフガードの境界線 (Frontier Safety)

CBRN / サイバーセキュリティ

専門知識は警戒閾値に達するが、
独立完遂能力（T/CCL）には未到達で安全に制御。

有害なアライメント・影響力

高い状況把握能力を持つが、
不整列行動（Misalignment）は確認されず。

テキスト自動安全性評価

+1.17pp 改善

過剰拒絶（Unjustified-refusals）の抑制

+0.84pp

正当な応答の維持

戦略的青写真 1：移行コストを最小化する「モデル・アブストラクション・アーキテクチャ」

Layer 1: App Layer
ユーザーインターフェース / 業務アプリケーション

Layer 2: Abstraction Layer (必須)
プロンプト管理 / 権限 / キャッシュ / ログ収集

⚠ ここにモデルID (gemini-3.7-flash) を直書きしない！

Layer 3: Model Layer
Gemini 3.7 Flash, その他のモデル

モデルの更新サイクルが短期化（3週間）する環境下において、システム本体とモデルを分離・独立させ、新モデル登場時の検証リスクを極小化する設計。

戦略的青写真 2：「思考レベルの動的ルーティング」によるTCO最適化パイプライン

2026年中の特別価格期間にこの「動的ルーティングロジック」と「初回精度基準の社内ベンチマーク」を確立することが、2027年以降の標準料金移行後における、エンタープライズAI戦略最大の競争優位性となる。