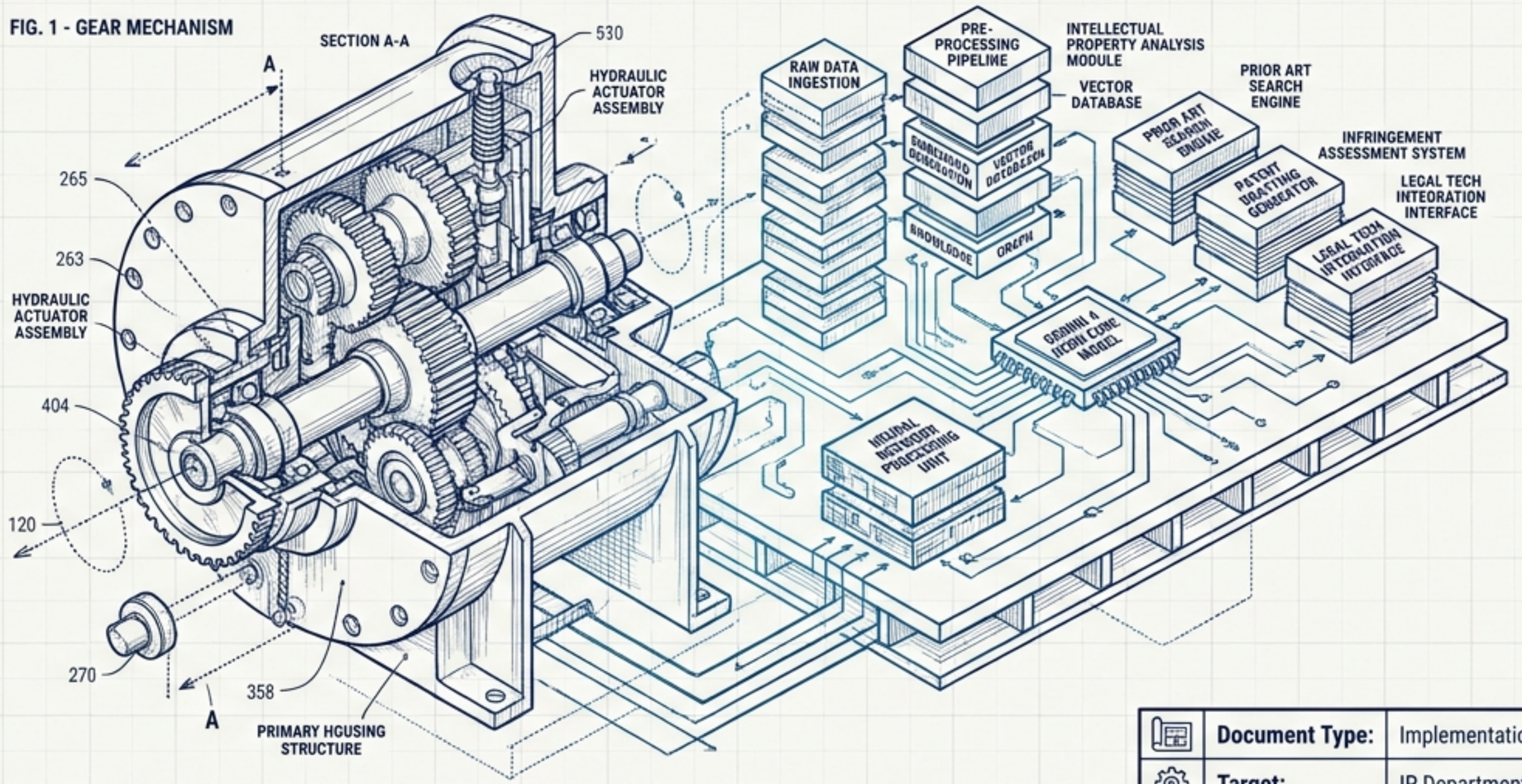
Gemini 4 Neon: 知財業務のパラダイムを革新する生成AI戦略

技術仕様・経済性検証に基づく、特許実務・システム実装への完全プレイブック



	Document Type:	Implementation Blueprint
	Target:	IP Departments / Legal Tech Architects
	Date:	2026.11

Gemini Notebook



知財実務における「文脈分断」「出力上限」「高コスト」という3大制約の同時解消。

Gemini Notebook

Gemini 4 モデルラインナップの再編と系譜

Gemini 4 Hydrogen (11/12発表)

日常作業・高速実行・サブエージェント

API Price (In/Out)
$1.50 / $7.50

Gemini 4 Neon (11/18発表)

Carbon派生の高効率推論・指揮役 (Orchestrator)

API Price (In/Out)
$3.00 / $15.00

中位Proプランから実務展開可能

Gemini 4 Carbon (11/10発表)

超長期・最高難度向フラッグシップ

API Price (In/Out)
$4.00 / $20.00
(12/1以降)

Gemini 4 Argon (10/13提供)

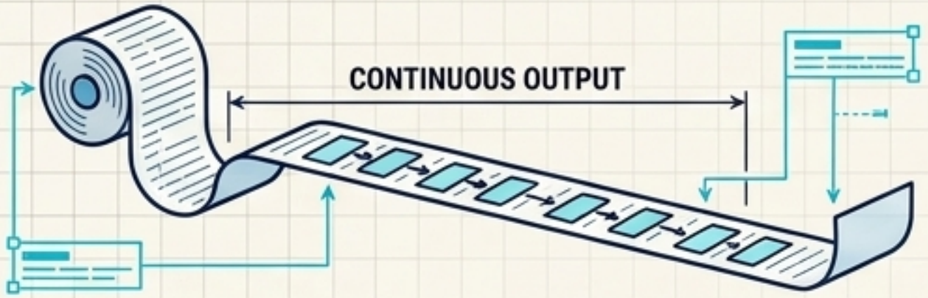
初期モデル
(Neonへの移行を公式案内)

Carbonから直接派生したNeonは、最高難度の研究タスク以外の大半の高度実務において、フラッグシップと同等の性能をより低コストで提供するパレートフロンティアを担う。

Gemini Notebook

実務設計を根底から変える4つの技術仕様

統合100万空間＆ロングデコード



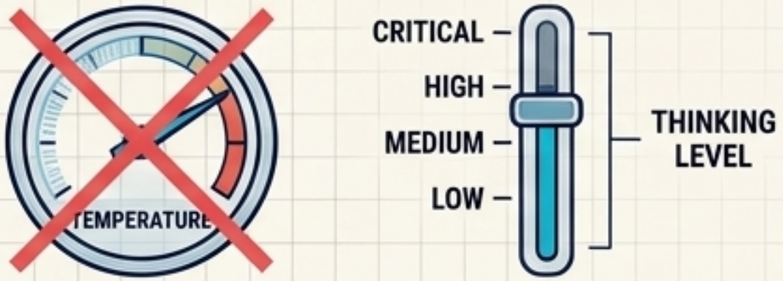
入力・思考・出力を合算した100万トークン上限 (Unified Token Limit)。
continuation_tokenによる最大1Mまでの継続生成
(途切れない超長文特許明細書・対比表の出力)。

ネイティブOffice入力



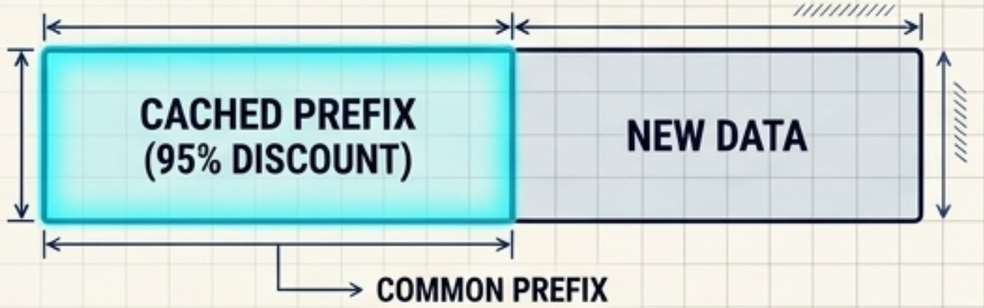
Cloud環境で .docx, .xlsx, .pptx を直接投入。
テキスト変換による構造崩れなしで、関数呼び出しの戻り値へ
直接格納可能。

サンプリング制御の厳格化



temperature等の従来パラメータは無視される。
推論の深さを4段階のthinking_levelで制御。出力形式はOpenAPI
準拠のresponseSchemaによる構造化拘束へ転換。

暗黙的キャッシュ機構 (Implicit Caching)

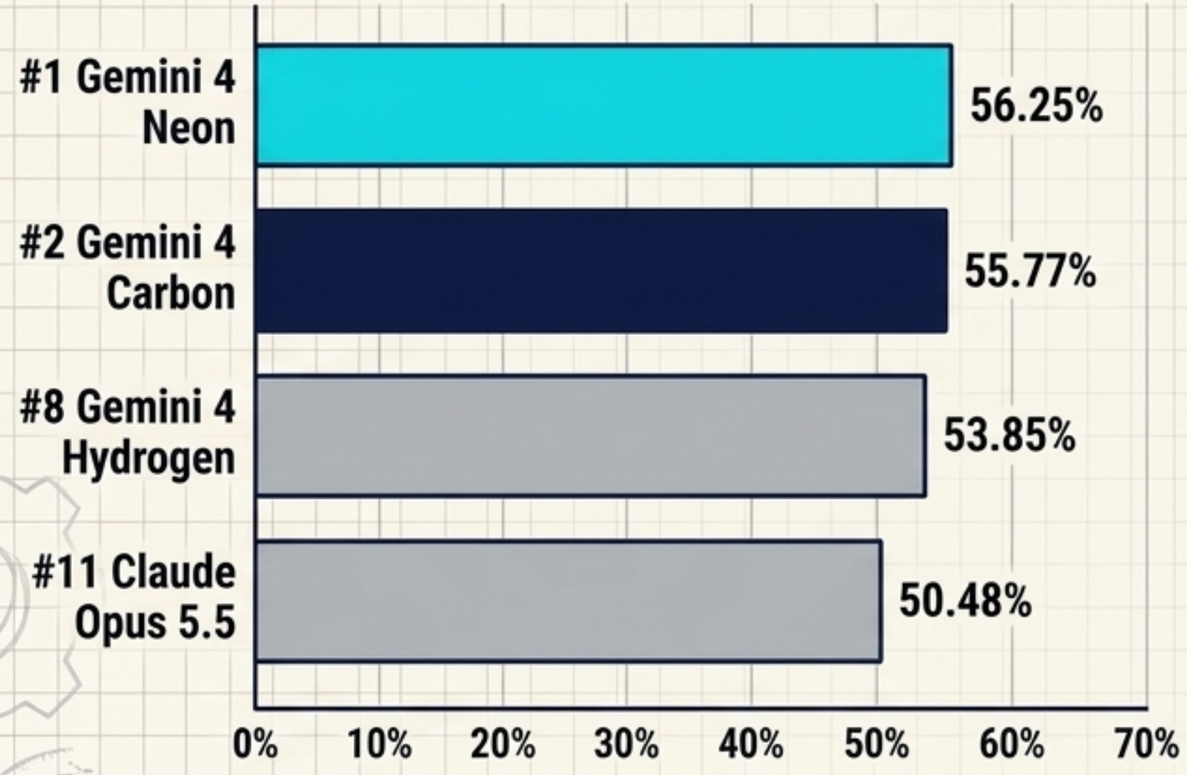


明示的キャッシュの廃止。4,096トークン以上の
共通プレフィックスに対し自動適用（95%割引）。

Gemini Notebook

ベンチマーク検証：法務推論と処理速度の圧倒的優位性

Vals AI Legal Research Bench (完全充足率)



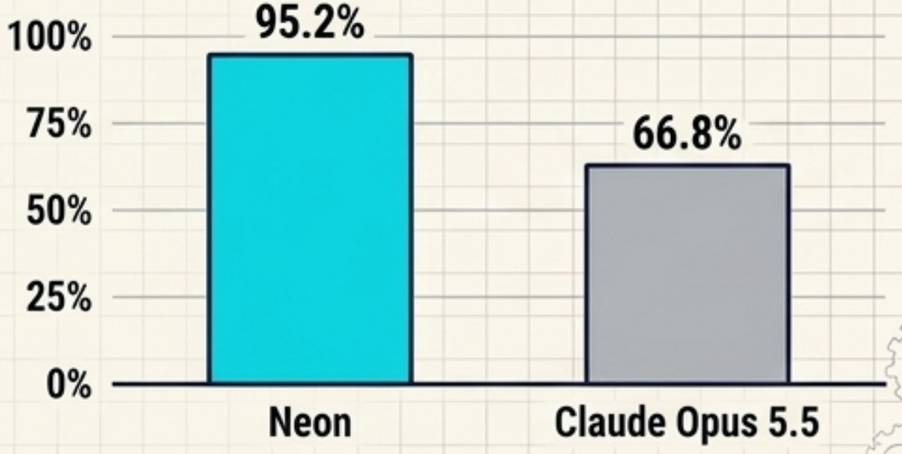
単なる知識テストではなく、現役弁護士が作成した
多段階の法的推論で最上位Carbonを凌駕。
コストはClaude Opus 5.5の1/3以下。

出力生成速度



Carbon(約57 tok/秒)の約2倍。
平均タスク時間も約40%短縮。

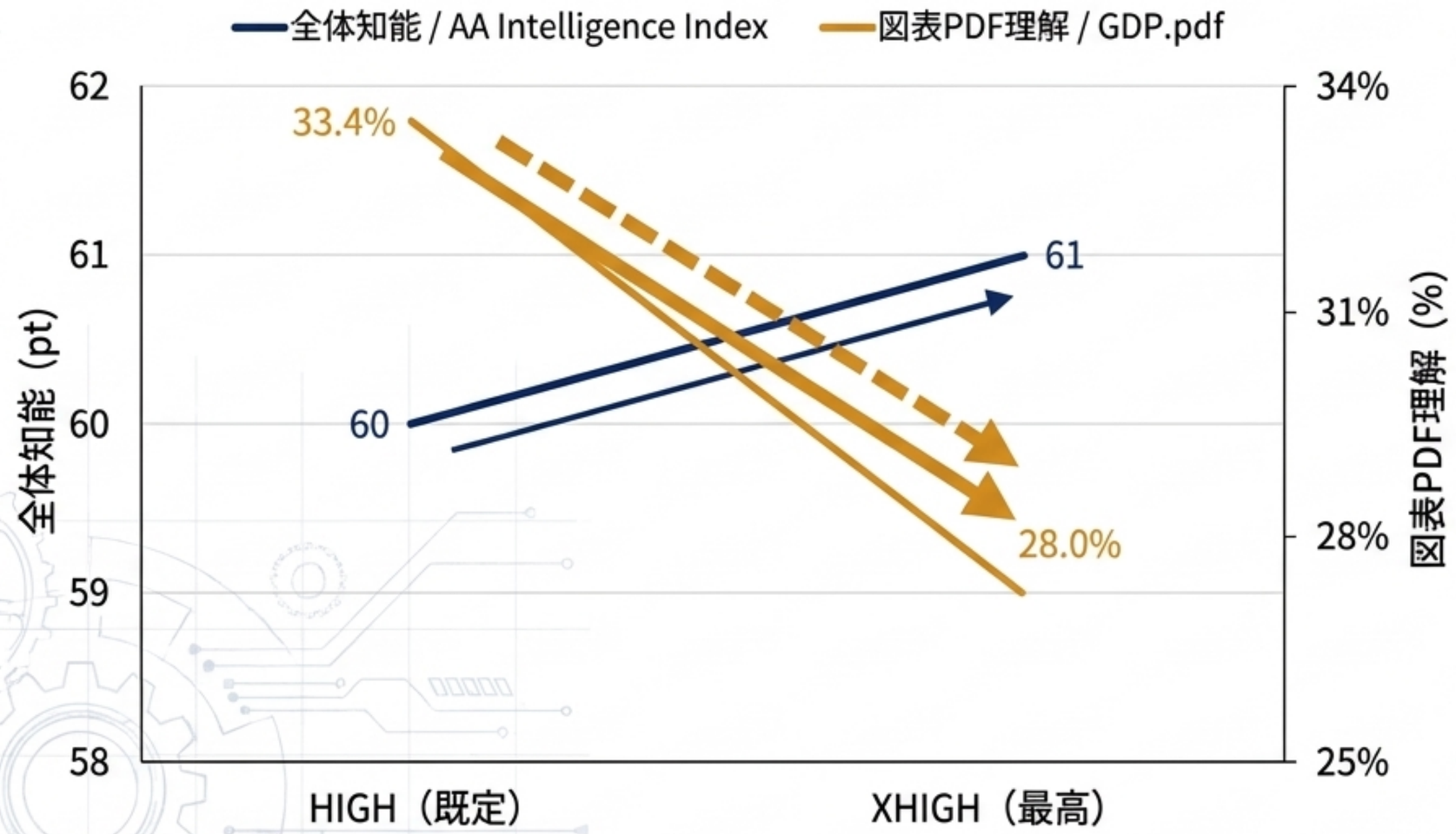
長文脈コンテキスト処理 (GraphWalks)



Gemini Notebook

【重要知見】視覚資料理解における「過剰推論の罠」

診断的ラインチャート: 過剰推論によるパフォーマンス低下の逆説



過剰推論 (Overthinking) のパラドックス

視覚情報や特許図面を含むPDF解析において、推論トークンを過剰に増やす（XHIGH設定）と逆に精度が低下する。

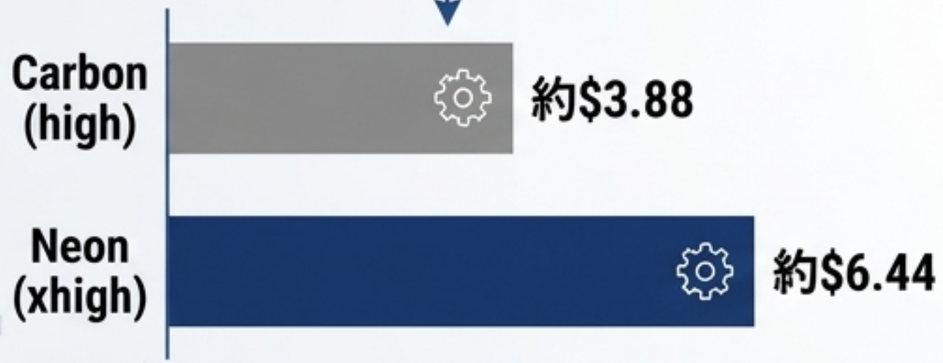
図面符号監査やグラフ読取では必ず HIGH を基本とし、複雑な文字ベースのクレーム論理検証でのみ XHIGH を適用するアーキテクチャ設計が必須。

Gemini Notebook

2026年12月1日：コスト優位性が発効する「価格逆転」

11月30日まで
Carbon発売記念半額期間

一時的にCarbonが安価



12月1日

12月1日以降
Carbon定価移行・Neonの本格優位

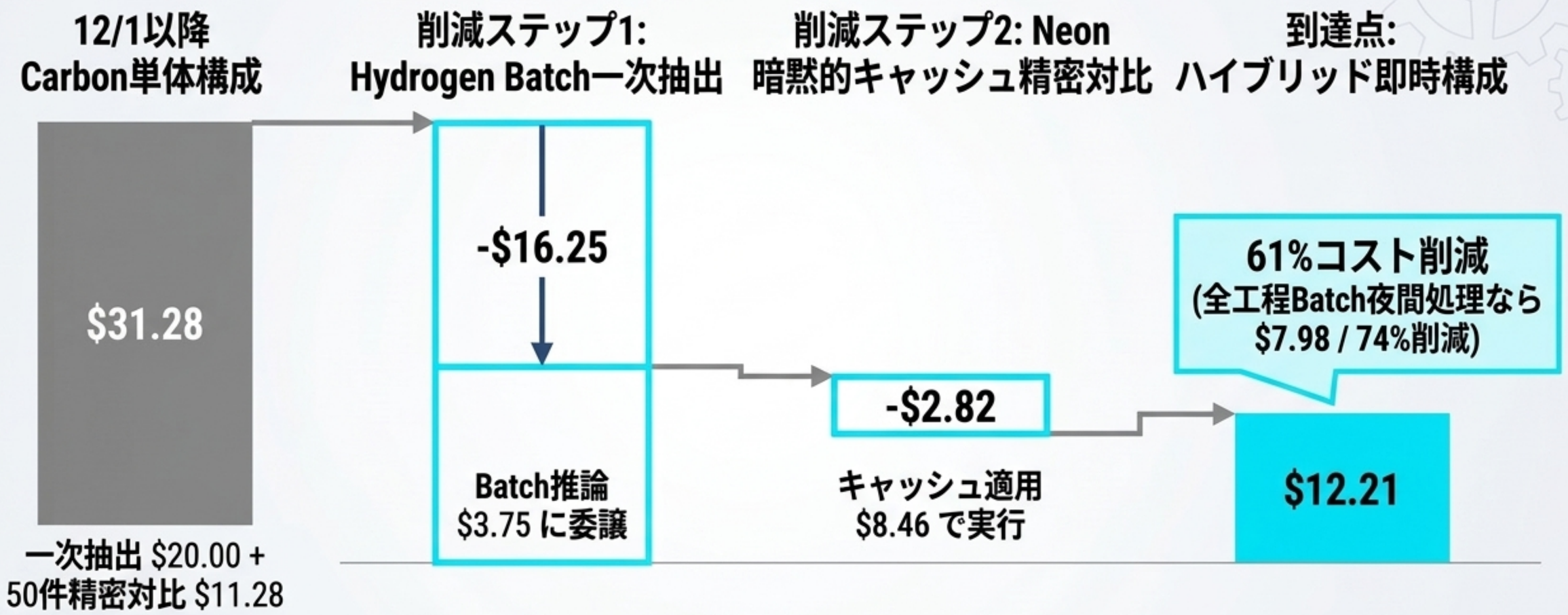
Carbon定価: 入力$4.00 / 出力$20.00
Neon単価: 入力$3.00 / 出力$15.00 (Carbon比 25%安価)



システム構築とテストは11月中にCarbon/Neon並行で実施し、本格稼働は
コストリーダーとなるNeonへ12月1日付で切り替える戦略的スケジューリングが最適。

Gemini Notebook

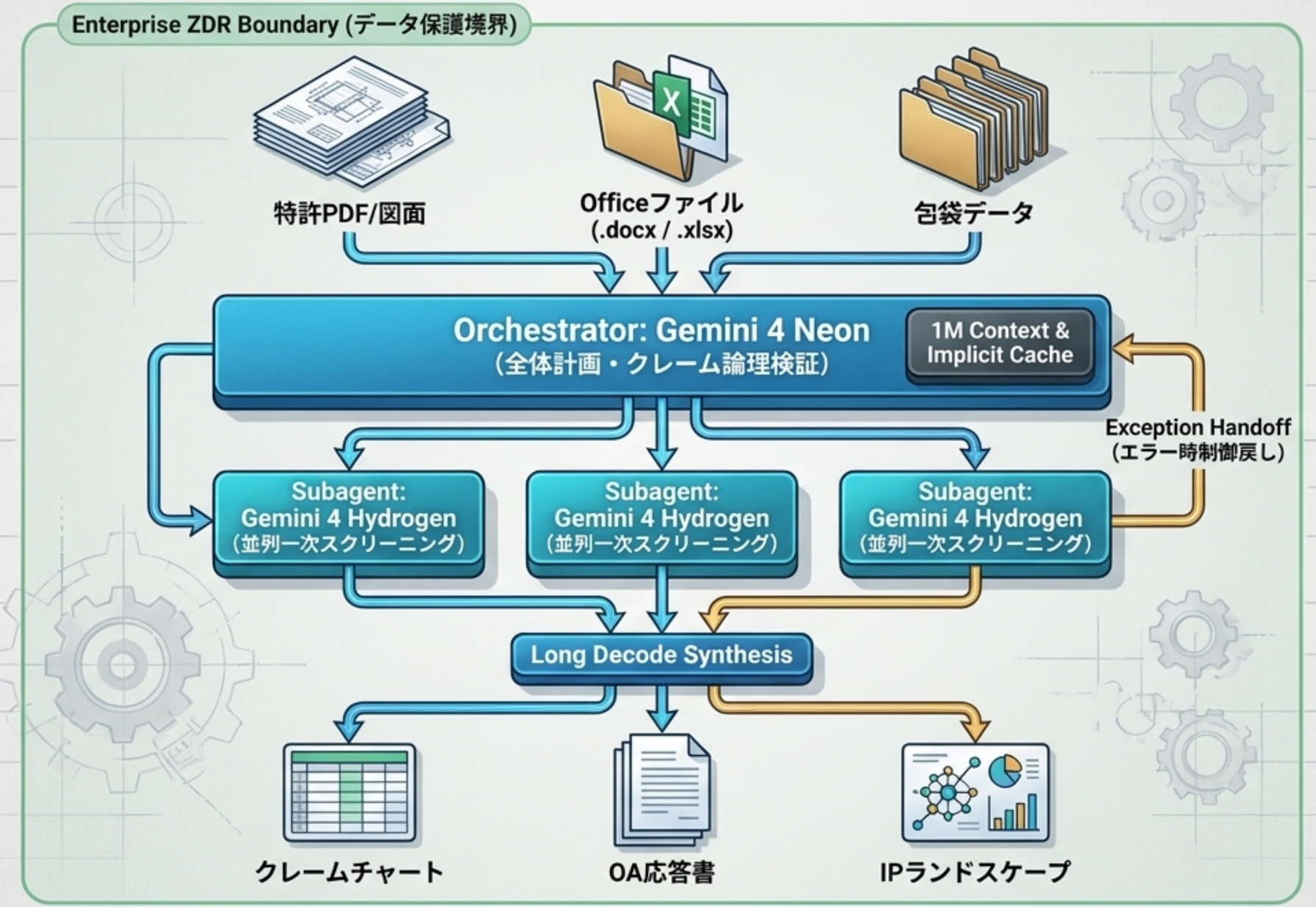
コストシミュレーション：1,000件規模のFTO調査



APIの名目単価ではなく、「HydrogenによるBatch一次抽出」と「Neonの暗黙的キャッシュによる精密対比」のハイブリッド設計こそが、FTO調査の巨大な経済効果を生み出す。

Gemini Notebook

マルチエージェント・アーキテクチャの基本設計図



Gemini Notebook

ユースケース① FTO・無効資料調査とクレームチャート

製品仕様のインプットと判定設計

対象のWord/Excel仕様書（約2万トークン）をNeonへ直接投入。

Keynote Box

Key: テキスト変換時の構造崩れを回避（ネイティブOffice入力）。

1,000件規模の並列スクリーニング

Hydrogen（Batch推論）が公報要約・請求項を並列で評価し、重要50件を抽出。

網羅的構成要件対比（ディープFTO）

製品仕様を先頭に固定（暗黙的キャッシュ：$0.15/1M）。抽出されたPDF/図面全文を一括投入。

Keynote Box

Key: 従来のRAGによる「参照構造の分断」を100万統合コンテキストで解決。

クレームチャート生成

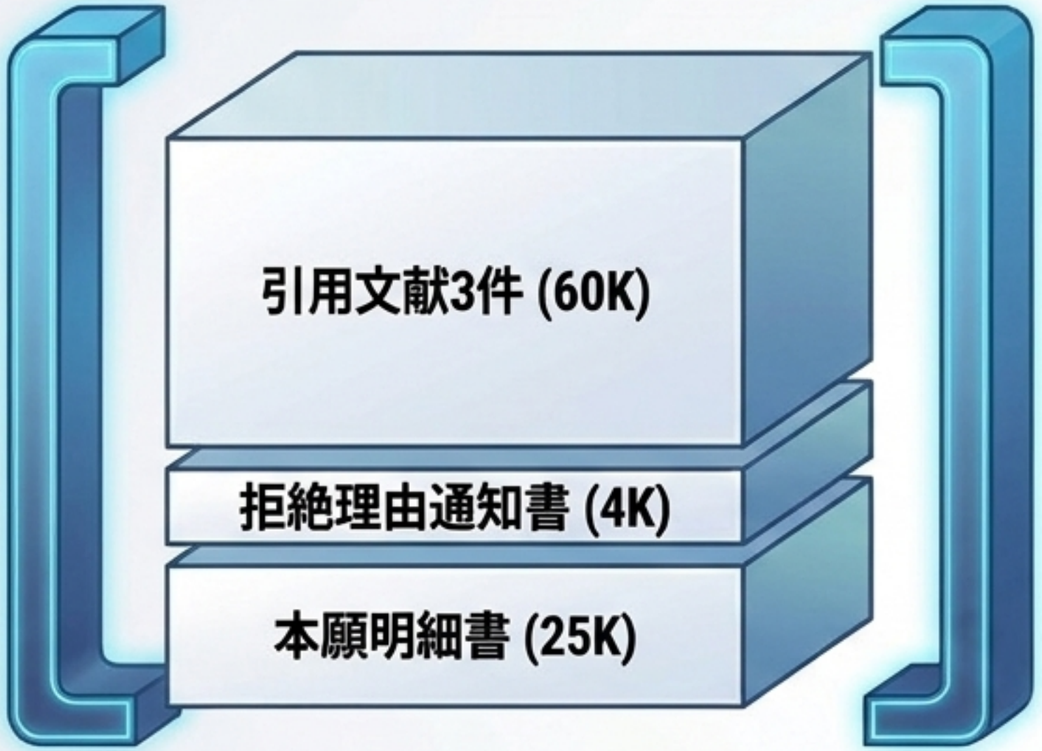
全請求項×全引用文献の長大な表を一括生成。

Keynote Box

Key: ロングデコード機能により、出力上限で途絶させることなく完遂。

Gemini Notebook

ユースケース② 拒絶理由通知（OA）対応における文脈保持



合計9万トークン：暗黙的キャッシュ
(Implicit Cache) に格納

弁理士の指示：「請求項3ではなく段落[0065]の数値範囲で補正案を再検討せよ」

Neonの処理：キャッシュを保持したまま低コスト（$0.15/1M）で再考。
Agent Arena修正対応力世界2位の追従性を発揮。

実務インサイト: 包袋一式を単一セッションに展開することで、主引例と副引例の論理的飛躍や、対応外国出願での主張と矛盾する補正（包袋禁反言リスク）を横断的に検証可能。

Gemini Notebook

ユースケース③ 明細書起案と図面・UI動画解析

明細書起案と「自己問答」制御

⚠ **課題:** モデルが勝手に確認質問を作り架空の実施形態を捏造する Self-Answering 挙動。

システムプロンプト指示:
開示資料に明記されていない構成・効果を推測補完せず、不明点は必ず [発明者確認事項] として列挙し出力を停止せよ。

responseSchemaによる構造的拘束でハルシネーションを防止。

図面符号・UI動画のマルチモーダル監査

図面と明細書の符号一致チェックや、ソフトウェア製品操作動画の侵害立証。



解決策: 思考レベルを必ず HIGH に設定。XHIGHによる過剰推論での精度低下を回避する。

複雑な三次元断面図の重なり等は人間が最終目視を行う。

Gemini Notebook

【統合】知財業務特化型アーキテクチャ・マトリクス

知財タスク	推奨モデル構成	推奨思考レベル	主要API機能	リスク管理・留意点
一次スクリーニング (SDI/ウォッチング)	Hydrogen(Batch) + Neon	Neon: HIGH / Hydrogen: HIGH	Batch推論	Free Tierの学習利用回避
FTO詳細調査 / クレームチャート	Neon単体	テキスト論理: XHIGH 図面: HIGH	暗黙的キャッシュ / ロングデコード	IPI（間接プロンプト / インジェクション）防御
拒絶理由通知 (OA) 対応	Neon (対話型推論)	HIGH ～ XHIGH	Interactions API	最新資料のコンテキスト手動注入
特許明細書起案 / 多言語翻訳	Neon (起案・用語統一)	HIGH	Officeファイル直接入力	Self-Answering挙動のシステム制御
IPランドスケープ / OSS監査	Orchestrator(Neon) + Subagents	HIGH ～ XHIGH	Web Grounding for Enterprise	通常検索のクエリ保存回避

Gemini Notebook

法務・倫理的要請と潜む罠

弁理士法第30条（守秘義務）と新規性喪失

罠：「Hydrogen」等のAPI無料枠（Free Tier）は入力データがGoogleの学習に利用される。

対策：未公開発明を扱う場合、学習非利用保証が適用される有料枠・Enterprise環境への完全統一が必須。クライアントへの事前合意取得。

知識カットオフと善管注意義務

罠：学習データの知識カットオフは「2026年3月」。2026年4月以降のAI関連発明の制度見直しや最新判例は欠落している。

対策: RAGやEnterprise Groundingによる最新資料のコンテキスト注入と、専門家による一次情報のファクトチェック。

著作権法と他社権利への配慮

罠: 第三者の著作物 (非特許文献など) を入力する際の適法性と、類似表現出力による侵害リスク。

対策: 著作権法第30条の4の範囲内での情報解析に限定し、元著作物の表現を享受する目的での生成を抑止する。

Gemini Notebook

データ保持ゼロ（ZDR）実装 4大チェックリスト

Google Cloud環境で機密性を担保するための厳格な設定ガイドライン

Interactions API の状態保存オフ

設定: store = false

デフォルトの true ではプロンプトや会話状態が保存されるため、明示的に無効化する。

ON

永続キャッシュ (Durable Caching) の排除

設定: retentionType: 'EPHEMERAL' 限定化

ストレージへの保持を避け、インメモリ限定の揮発キャッシュを選択。または完全無効化 (disableCache: true) を検討。

ON

Enterprise Web Groundingの選択

設定: 通常のGoogle検索グラウンディングの回避

通常版はデバッグ目的でクエリが最大3日間保存されるため、機密業務ではEnterprise版の採用が必須。

ON

不正監視ログの確認とIPI防御

設定: Frontier Model監視からの除外申請

ログ保存を無効化する申請手続きと、外部PDF取り込み時の間接プロンプトインジェクション(IPI)防御を実装。

ON

The architecture is ready.

この設計図に基づき、セキュアで高収益な知財AIシステムの実装へ。

Gemini Notebook