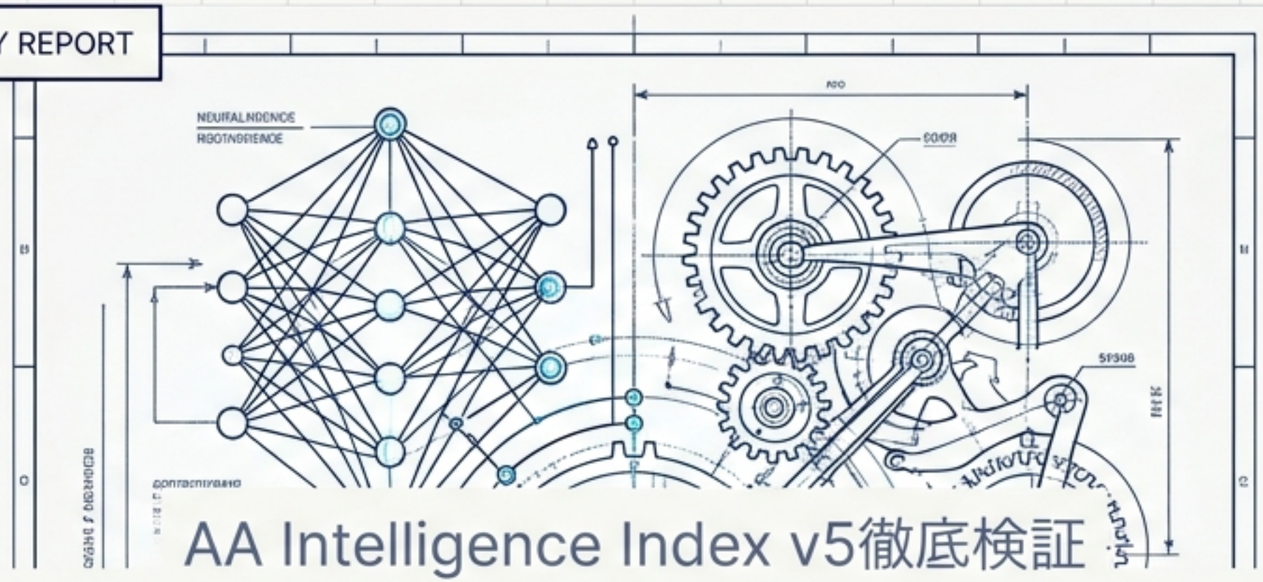
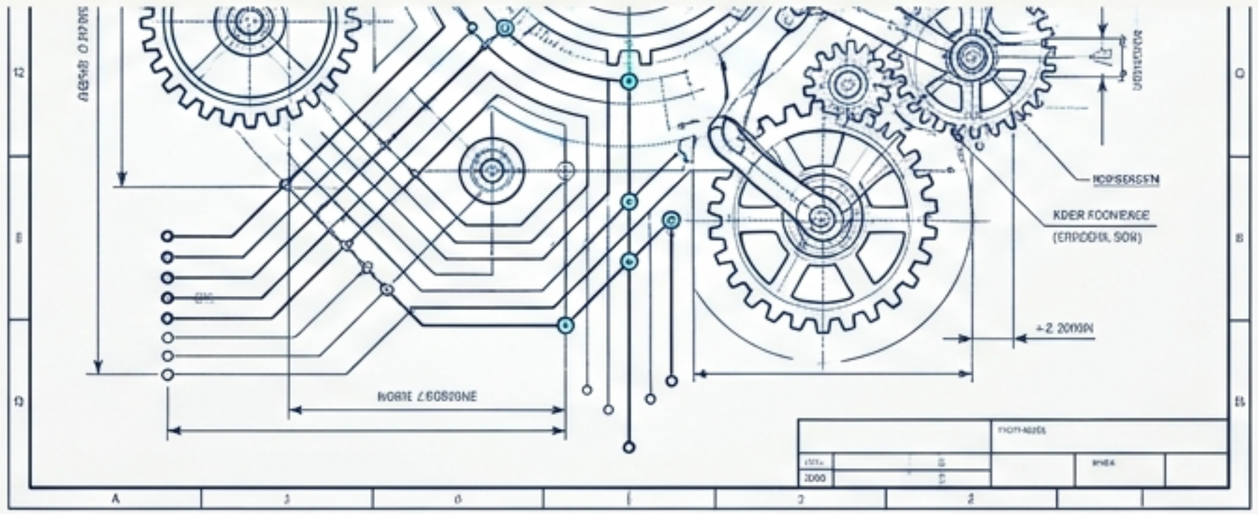
CONFIDENTIAL / INTERNAL STRATEGY REPORT



AA Intelligence Index v5徹底検証

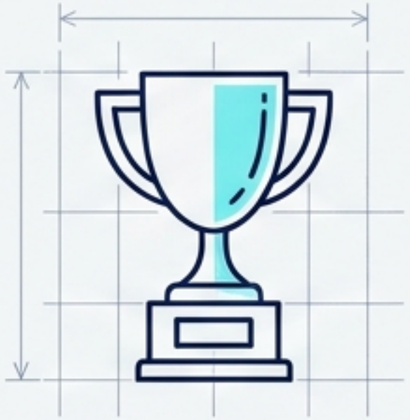
# Gemini 4 Carbonの現在地と知財実務の生存戦略



2026年10月29日

Gemini Notebook

# エグゼクティブ・サマリー：事実、影響、そして行動



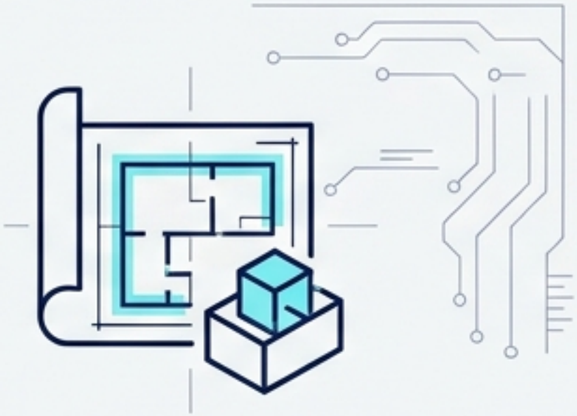
## 事実：覇権奪還と限定的アクセス

- AA Index v5で59点を獲得し首位
- サイバー防御者向けに限定提供中
- 一般提供日・詳細スペックは非公開



## 影響：長文脈の恩恵とハルシネーションの罠

- 100万トークンが特許読解プロセスを革命的に効率化
- 一方、ハルシネーション率は4%から11%へ悪化
- AI同士の相互チェックは法的に不十分



## 行動：自社独自の評価・ガバナンス構築

- 現状はOpus/Astra等の既存モデルを維持
- 20〜50件の自社固有「正解データ」を至急構築
- ゼロデータ保持環境での利用ルールを徹底

Gemini Notebook

# 華々しいベンチマーク首位の裏にある「非公開」の現実



**59 Points (1位)**

## 見出し（Headline）

- AA Intelligence Index v5で首位
- Claude Opus 5.5やGPT-6 Astraを凌駕
- 4つの評価（科学、コーディング等）で1位



**限定提供**

## 現実（Reality）

- Fairwind Program経由の一部ユーザーのみ
- 公開API、技術レポート、パラメータ数は一切非公開
- コストは$7.76/タスク（Astraの約1.2倍）

Gemini Notebook

# 主要フロンティアモデルの徹底比較（2026年10月時点）

| モデル | AA v5 スコア | タスク単価 | ハルシネーション率 | 提供状況 |
|---|---|---|---|---|
| **Gemini 4 Carbon (High)** | **59** | **$7.76** | **11%** | **限定提供** |
| Claude Opus 5.5 (Max) | 58 | $10.04 | 50% | 一般提供 |
| GPT-6 Astra (Max) | 57 | $6.25 | 50% | 一般提供 |
| GPT-6.1 Sol (Max) | 55 | $1.27 | 未検証 | 一般提供 |

**総合力は最高峰だが、知財実務における「実質的な費用対効果」では依然として他モデルにも優位性がある**

Gemini Notebook

# 長文脈エージェントがもたらす破壊的効率化




入力：1Mトークン
数十〜数百件の特許公報

**GraphWalks評価: 96.2%**
(Astra: 71.8% / Opus: 66.8%)

出力：自動化された対比表
IPランドスケープ（IPL）

インサイト：検索式の作成から精読、絞り込みまでのパイプラインが自動化され、
知財実務者の役割は「作業者」から「レビュアー」へ完全移行する。

Gemini Notebook

# 知能向上と引き換えに悪化したハルシネーション率




**推論スコア向上**

Argon(50) → Carbon(59)




**ハルシネーション率**

Argon(4%) → Carbon(11%)

**11%の捏造リスク。知財実務において「もっともらしい嘘」は、単なる誤答よりも致命的なリスク（偽陰性・権利喪失）をもたらす。**

Gemini Notebook

# In re Mitchell事件の教訓：AI同士の検証は免罪符にならない

AIがクレームチャートを作成（捏造を含む）

↓

別のAIがレビューを実施（エラーをすり抜け）

↓

代理人がUSPTOへ提出

↓

**結果：懲戒処分**
（架空引用47件、審査経過への混入26件）

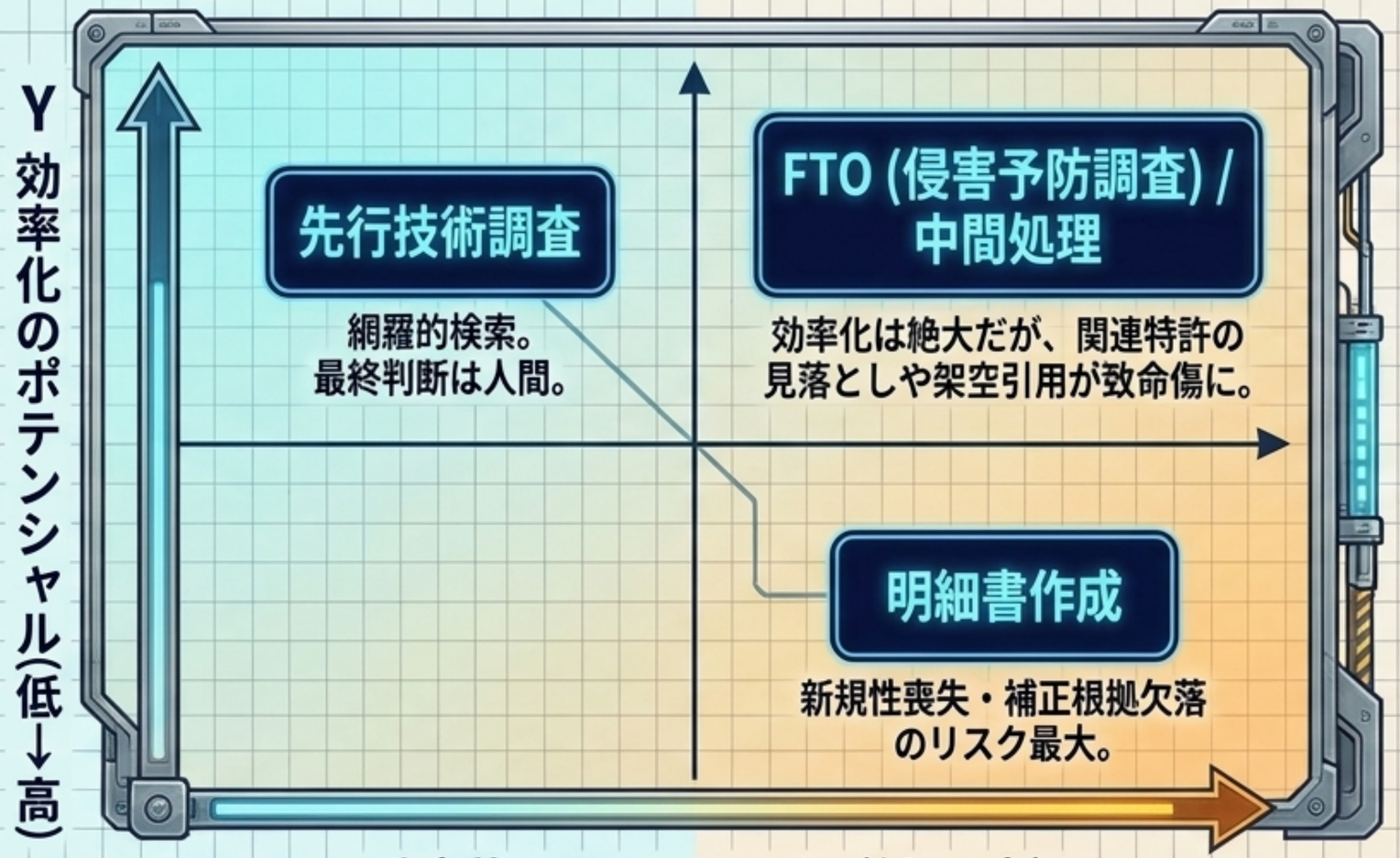
**人間による原典照合**

絶対的ルール：
「AIによる相互チェックは法廷での検証義務を果たしたとみなされない」

段落番号・図番号・頁行の原典照合は必ず人間が行うインフラが不可欠。

Gemini Notebook

# 業務別インパクト・マトリクス：どこからAIを適用すべきか




結論：偽陰性（関連特許の見落とし）のリスクをどう設計するかが、AI適用の分岐点となる。

Gemini Notebook

# 越えてはならないガバナンスの壁

## 知財実務のAI活用

### 日本弁理士会（JPAA）

秘密情報入力による守秘義務違反・新規規性喪失リスク。厳格な同意取得と人間によるファクトチェックが必須。

### 米国特許商標庁（USPTO）

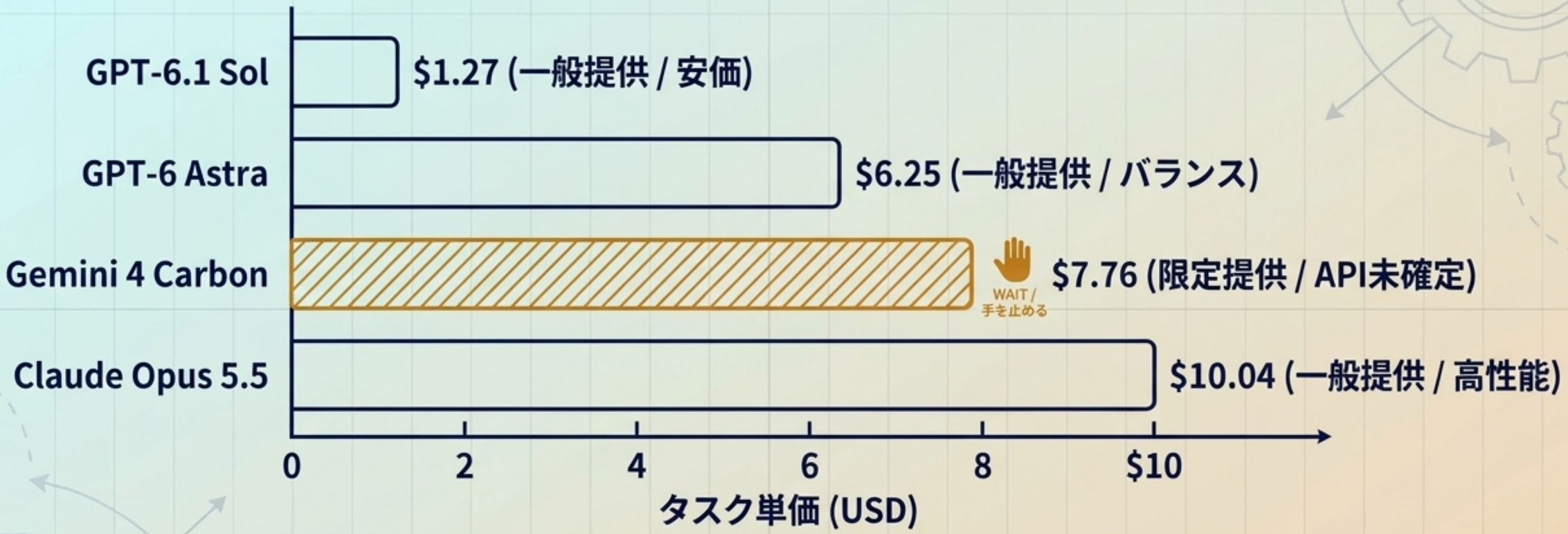
AIパイロット終了。AI支援の有無に関わらず、提出書類の最終責任は実務家が負うとする厳格な指針。

### 欧州特許庁（EPO）

2026年指針にて、2026年指針にて、書面の責任は当事者にあり「AIがやった」という抗弁を明確に無効化。

Gemini Notebook

# アクション１：現状は「一般提供済モデル」を維持する合理性




API仕様もデータ保持条件も未確定な現段階での乗り換えはリスクが高い。まずはOpus 5.5やGPT-6 Astra等の既存モデルで実務パイプラインを回し、一般提供開始を待つべき。

Gemini Notebook

# アクション２：自社固有の「正解データ」構築を急ぐ

## Step 1: 抽出 (Extraction)

過去案件から20〜50件の成功事例（無効化成功、複雑な拒絶応答、中日翻訳対比など）を抽出する。

## Step 2: 基準設定 (Metrics)

一般的なIQスコアは無視。「主要文献の再現率」「架空引用の数」「1件あたりの処理時間」を独自の評価軸とする。

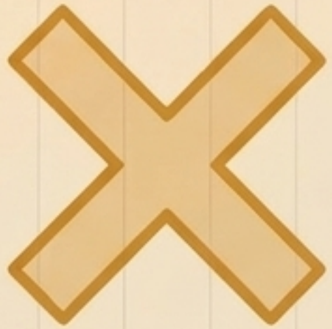
## Step 3: ブラインド評価 (Blind Eval)

Carbon一般公開時、この自社データセットを用いて既存モデル（Opus/Astra）と精度を直接比較する。

**外部の「59点」を鵜呑みにせず、自社特有の図面符号読解やクレーム解釈における精度を測る独自の物差しが、今後の最大の競争優位となる。**

Gemini Notebook

# アクション３：エンタープライズ・ガバナンスの徹底



## コンシューマー向け API（標準Gemini）

- プロンプトが監視目的で保持されるリスク。
- 弁理士会ガイドライン違反の可能性。
- 未公開発明の入力による新規性喪失の危険。



## エンタープライズ環境（Vertex AI 等）

- 「ゼロデータ保持（Zero Data Retention）」の契約的担保。
- 学習データへの利用を完全にブロック。
- クライアント同意の下、安全な処理が可能。

**絶対的ルール：「出願前の発明情報は、ゼロデータ保持が契約上担保された環境以外には絶対に入力しない」**

Gemini Notebook

# 結論と2026年末に向けたロードマップ

11月4日: Kimi K3.1 Pro 重み公開

未定: Gemini 4 Carbon 一般提供開始 / モデルカード公開

年末: Argonローンチ割引終了 / 競合の対抗モデル発表

2026年末

- ☐ 所内AI利用ガイドラインの策定（JPAA/USPTO準拠）
- ☐ 引用文献の「人間による原典照合」の義務化（Mitchell対策）
- ☐ 50件の自社用「正解評価データセット」の構築開始

**最強のAIを選ぶことではなく、最強の「検証インフラ」を持つ組織だけが生き残る。**

Gemini Notebook

# 出典およびリファレンス

## ベンチマーク・技術仕様

- Artificial Analysis (Intelligence Index v5)，2026年10月
- Google The Keyword (Koray Kavukcuoglu)，Gemini 4 Carbon: our next frontier model
- Gemini API Documentation (Zero Data Retention / Vertex AI)

## 法的判例・ガイドライン

- USPTO（In re Mitchell: D2026-16)，2026年8月
- 日本弁理士会（JPAA）「弁理士業務AI利活用ガイドライン」，令和7年4月
- 欧州特許庁（EPO）2026年版審査ガイドライン

※本レポートは2026年10月29日時点で利用可能な情報に基づいています。AIモデルの性能、アクセス条件、および価格は急速に変更される可能性があります。

Gemini Notebook