上位モデルと中位モデルの知財業務における使い分け

Claude・GPT・Gemini最新6モデルの性能評価と、実務適用のための5つの戦略的フレームワーク

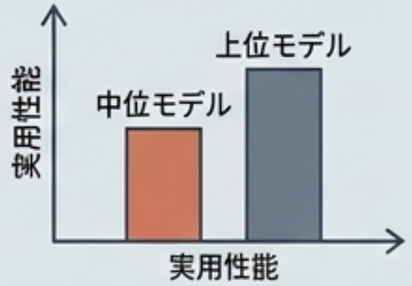
2026年11月19日（最新公開データに基づく）

Gemini Notebook

知財業務におけるAIモデル選定のパラダイムシフト

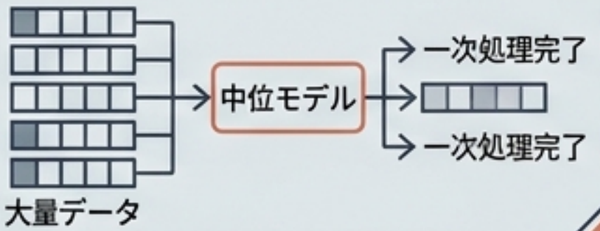
階層と性能は一致しない

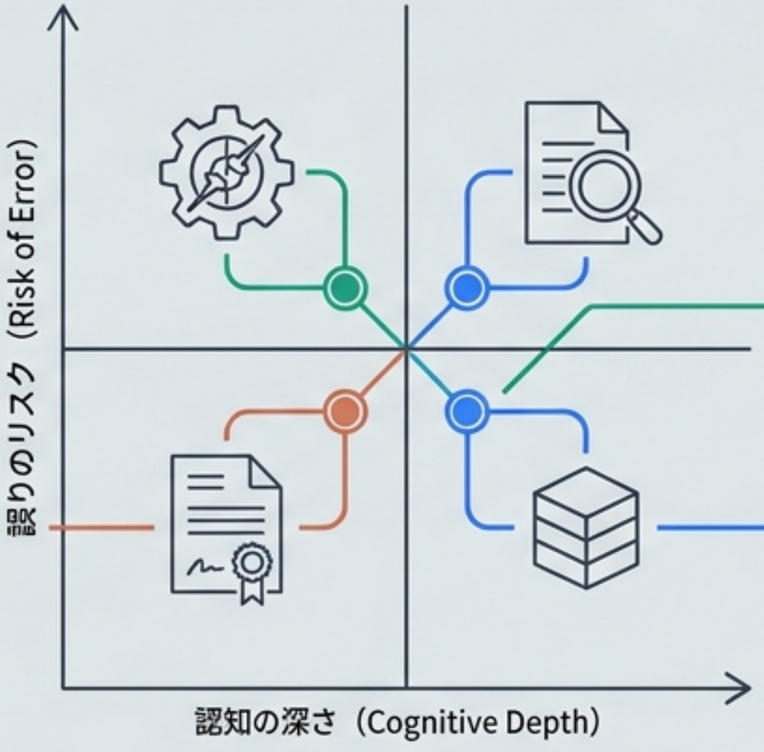
中位モデル（Neon等）が上位モデルを凌駕する逆転現象が発生。



日常業務は中位モデルで完結

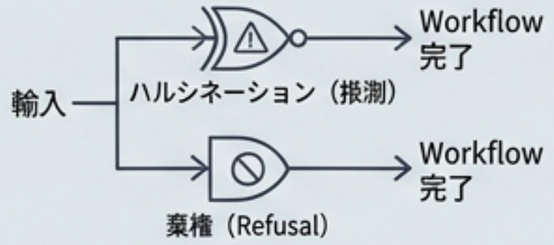
圧倒的なコストパフォーマンスで、大量の一次処理（バルク処理）をカバー。





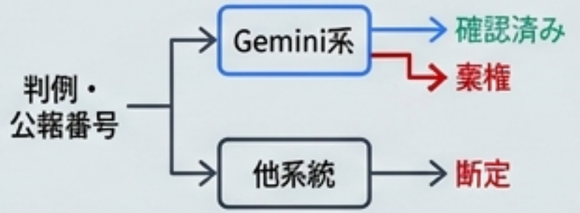
「正答率」より「誤りの出方」

ハルシネーション（推測）か、棄権（Refusal）か。系統の癖が業務の適性を決める。



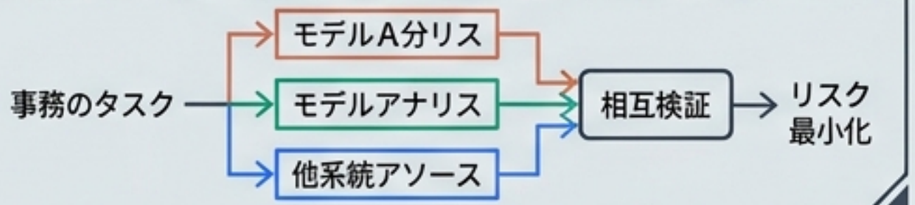
事実検証におけるGeminiの優位性

判例や公報番号の確認には、棄権率が高く断定を避けるGemini系を配置。



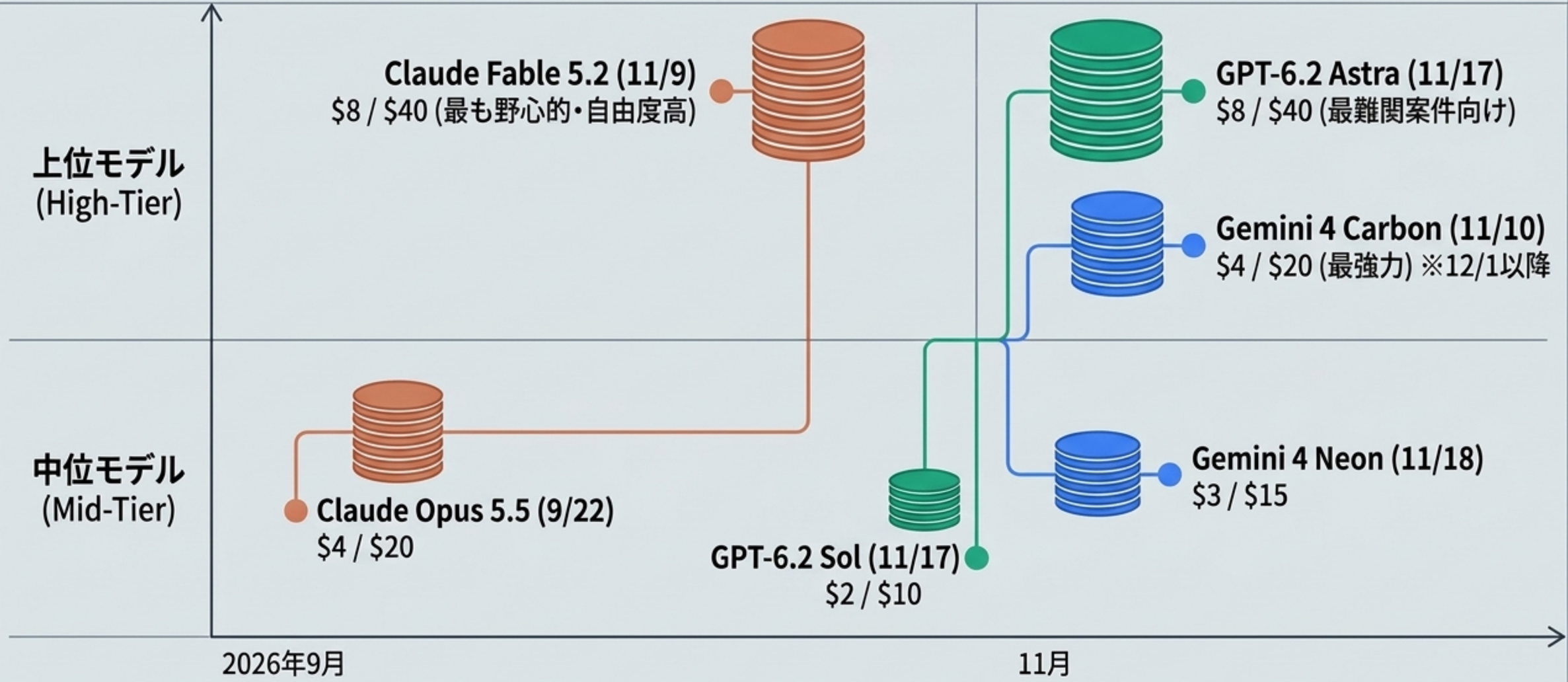
分業（カスケード）と相互検証

複数モデルを組み合わせ、知財特有の「見落としリスク」を最小化する設計へ。



Gemini Notebook

最新6モデルのランドスケープと圧倒的なコスト差

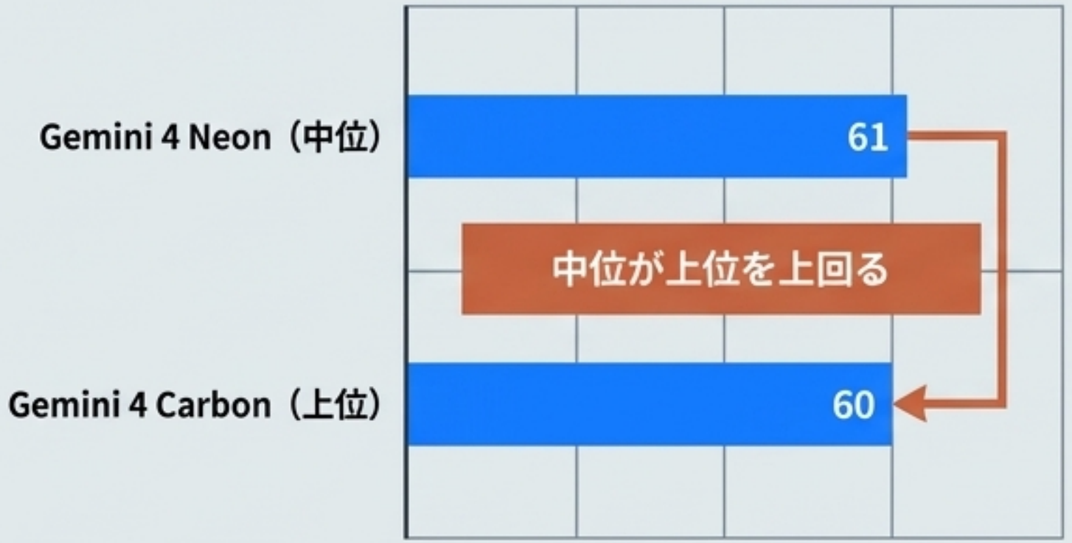


上位と中位でAPI単価に2倍〜4倍の差が存在。しかし、提供元は共通して「まず中位モデルから」を推奨している。

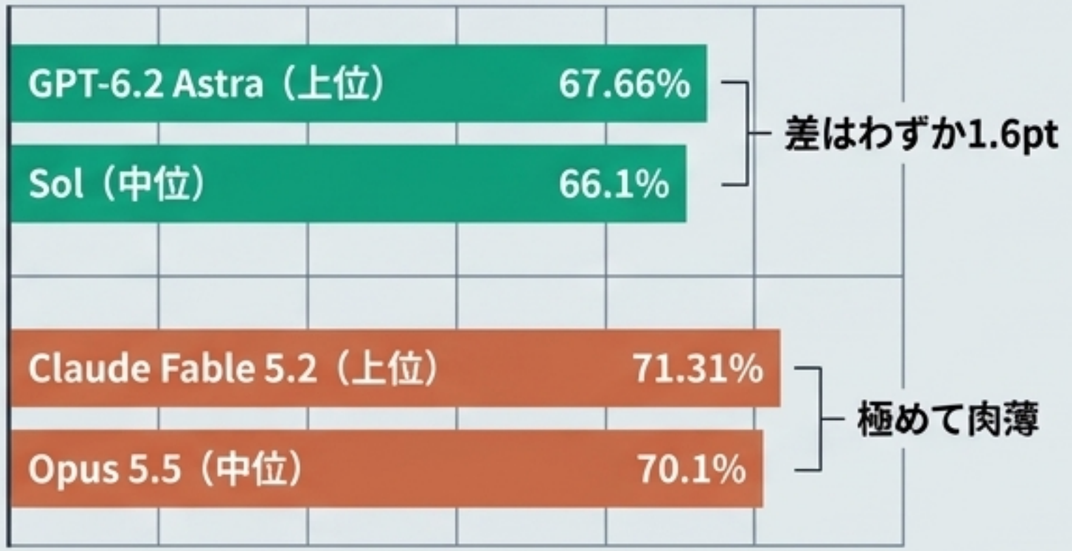
Gemini Notebook

「高いモデル＝賢い」は過去のもの（The Performance Paradox）

総合指数の逆転現象（AA 総合指数）



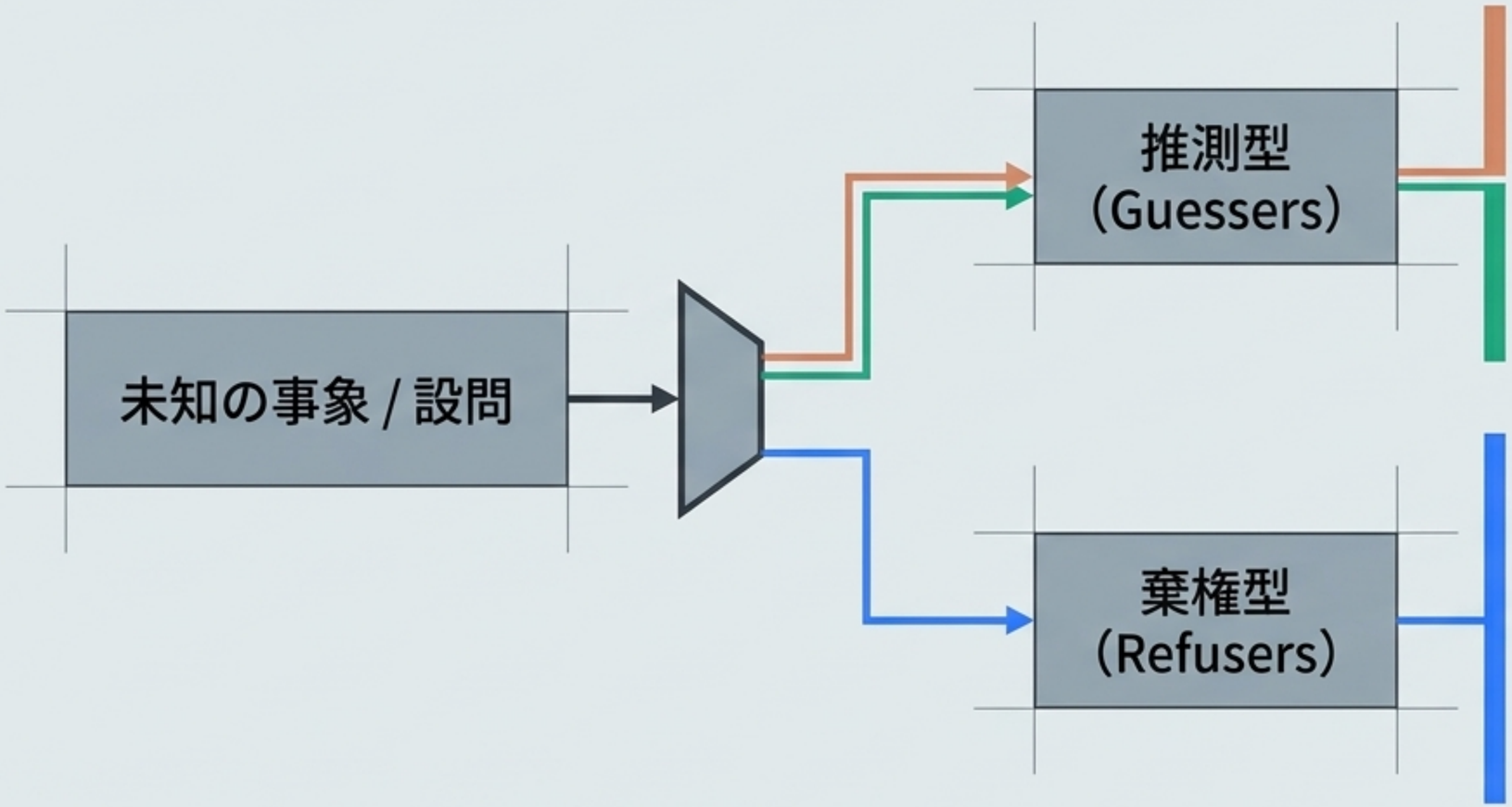
僅差のパフォーマンス（Vals AI 総合指数）



1タスクあたりのコストが数倍異なるにもかかわらず、総合性能はほぼ同等。
単なる「正答率」による選定は、投資対効果を著しく損なう。

Gemini Notebook

致命的なのは「正答率」よりも「未知の事象への振る舞い」



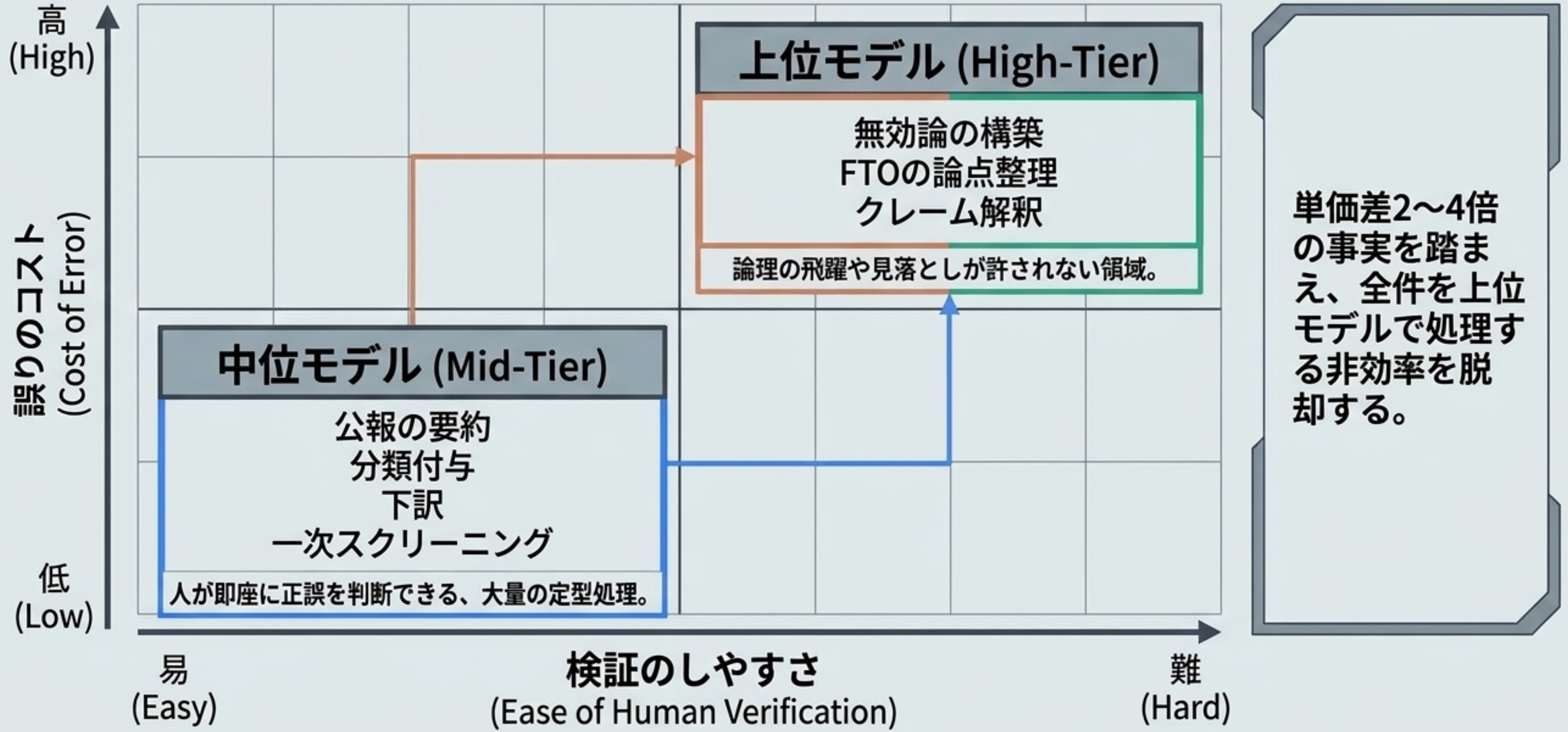
Claude & GPT: 正答率は高いが、正答できない設問の50〜60%で「もっともらしい誤答」を返す。
全設問に対する誤答率（試算）：13%〜21%。

Gemini:
正答率で劣る場面もあるが、知識がない場合は回答を「棄権」する傾向が強い。
全設問に対する誤答率（試算）：わずか5%〜6%。

知財実務においては、誤った断定（存在しない判例や公報番号の提示）が致命的となる。モデルの「系統（Lineage）」による誤答の癖を理解することが不可欠。

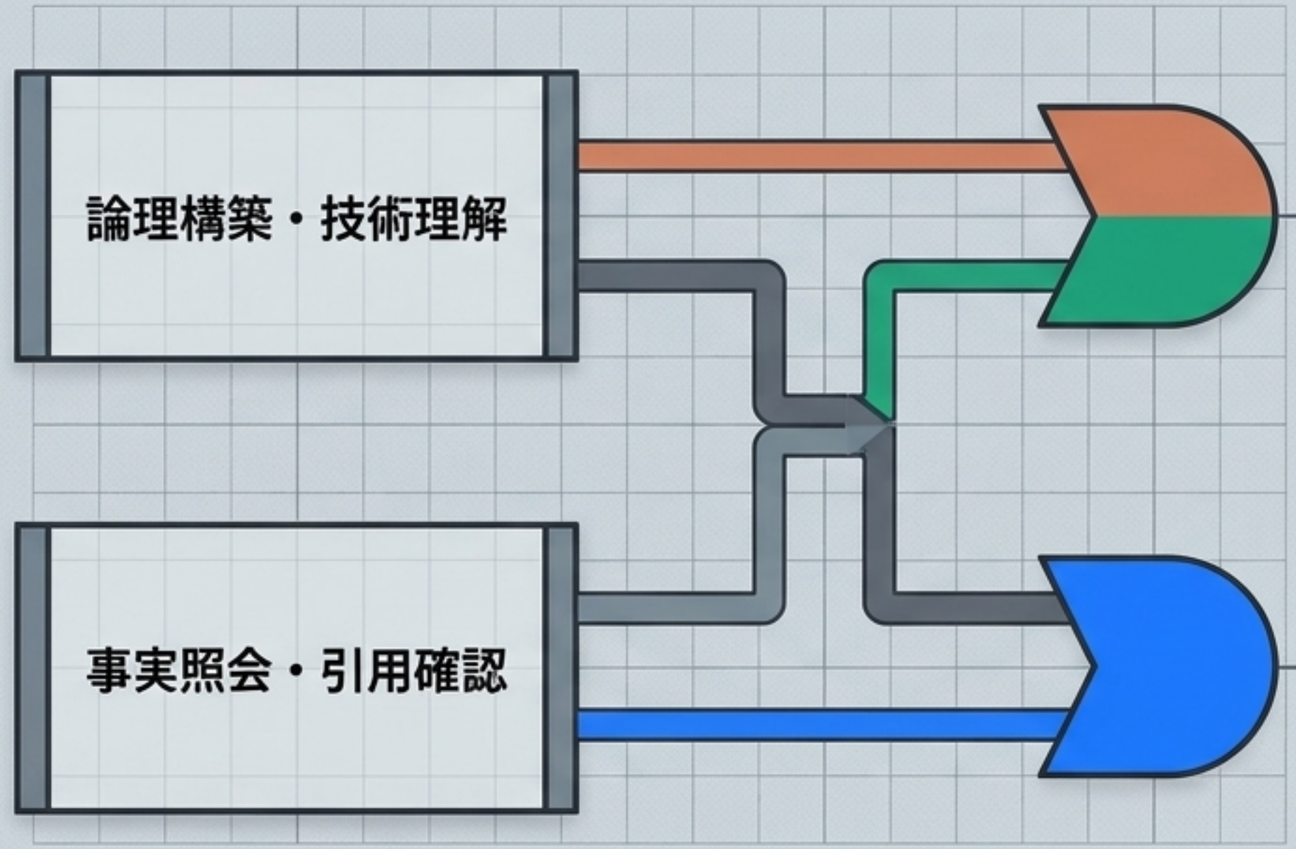
Gemini Notebook

Strategy 1 - リスクベースの階層ルーティング



Gemini Notebook

Strategy 2 - 「誤り方の系統」に基づく適性マッチング



適合モデル: Claude Fable 5.2 / GPT-6.2 Astra
適したタスク: 反論の起案、技術の深掘り、無効理由の論理組み立て。
強み: 知識と推論の深さ。

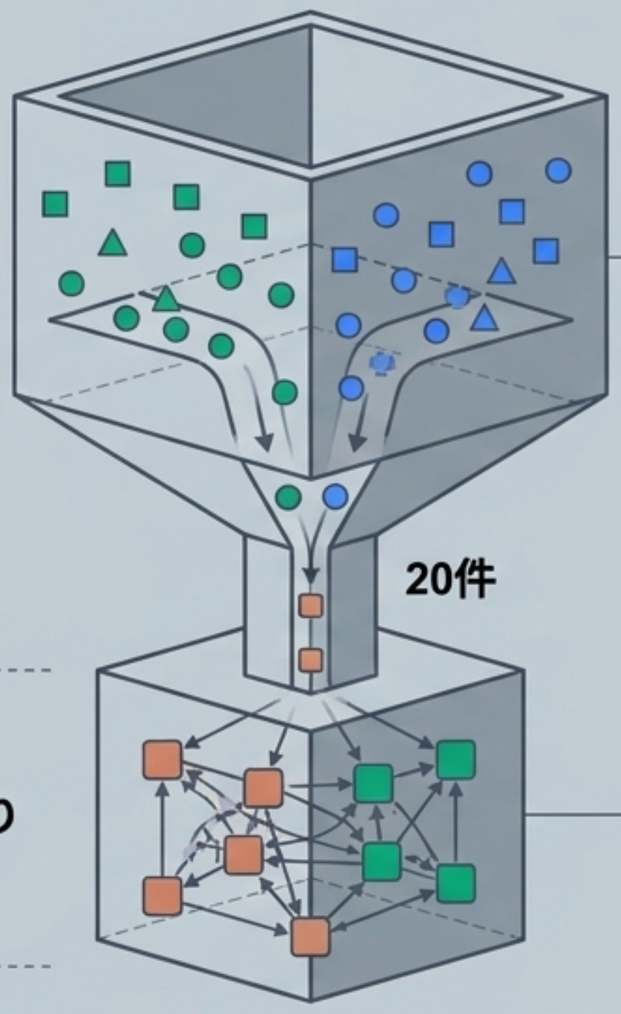
適合モデル: Gemini 4 Neon / Carbon
適したタスク: 判例番号、条文、審査基準の該当箇所、公報番号、日付の照会。
強み: 未知の事項に対する高い棄権率（誤った断定のリスク最小化）。

【Rule】どのモデルを使用する場合でも、必ず「原文の添付」と「検索による根拠づけ」を必須とする。

Gemini Notebook

Strategy 3 - AIカスケード（オーケストレーターと実行者の分業）

1,000件の先行文献



20件

Stage 1:
大量処理
(Mid-tier)

モデル: GPT-6.2 Sol または Gemini 4 Neon

処理: 先行文献1,000件の一次判定。広く浅くスクリーニング。

Stage 2:
難所の深掘り
(High-tier)

モデル: Claude Fable 5.2 または GPT-6.2 Astra

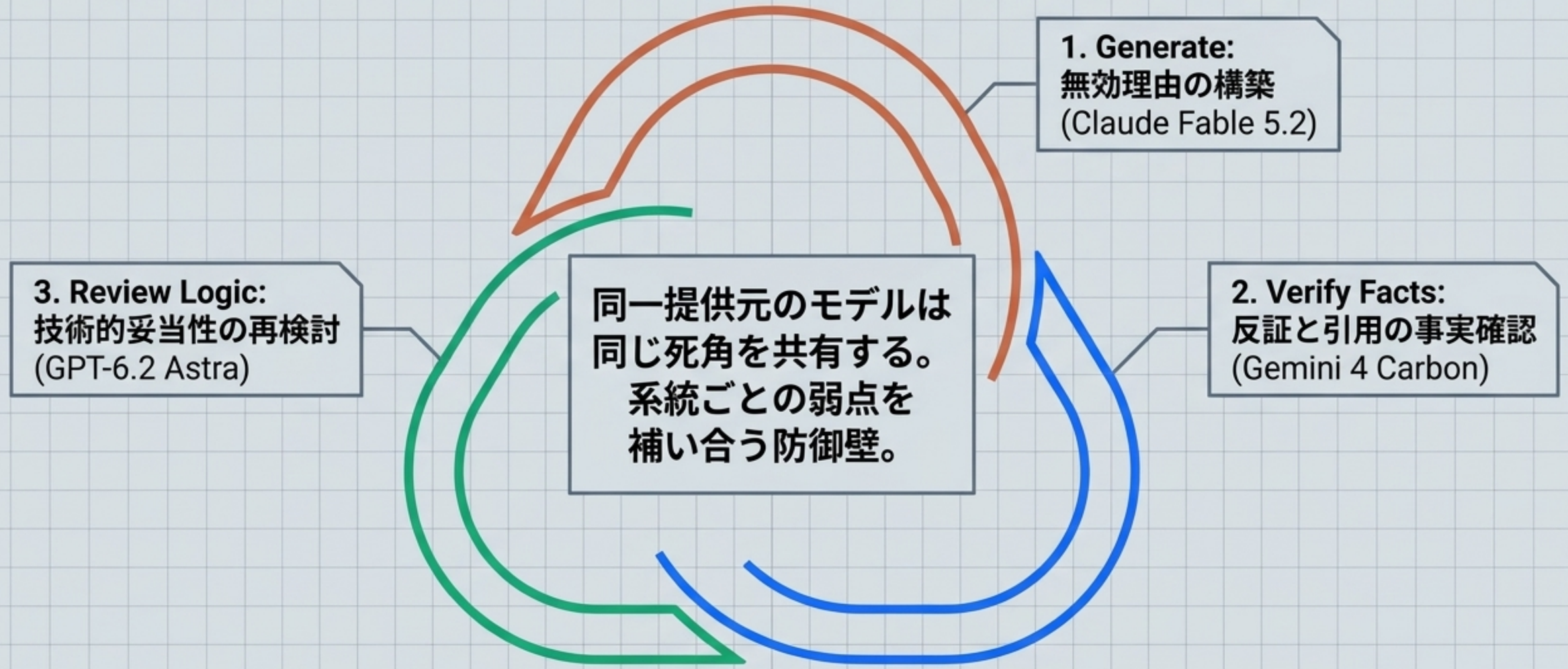
処理: 絞り込んだ20件の厳密な対比と、進歩性ロジックの構築。

【コスト最適化の裏技】

バッチ実行を活用すれば、上位のFable 5.2を中位Opus 5.5の通常料金と同等で使用可能。推論努力（effort）設定の調整で、処理時間・コストを半減しつつ同等精度を維持。

Gemini Notebook

Strategy 4 – 異系統モデルによる相互検証アーキテクチャ



GPTの省略癖、Claudeの推測癖、Geminiの棄権癖。
異なる性質を意図的に組み合わせることで、知財特有の「見落としリスク」を相殺する。

Gemini Notebook

Strategy 5 – 専門分野と入力フォーマットの適合性



化学・材料・物理

GPT-6.2 Astra / Claude Fable 5.2（科学系評価でAstraが首位）



法的リサーチ

Gemini 4 Neon / Carbon（法務評価で首位。※ただし米国法評価に基づく）



証拠資料・マルチモーダル

Gemini 4 Carbon（動画・音声・PDFに一貫対応。Astraは動画未対応）



【警告】バイオ・医薬分野

Claude Fable 5.2は、安全対策により生物分野の研究タスクの大部をブロック（旧世代Opus 5が代替応答する仕様）。GPT-6.2にも制限あり。実機での挙動確認が必須。

Gemini Notebook

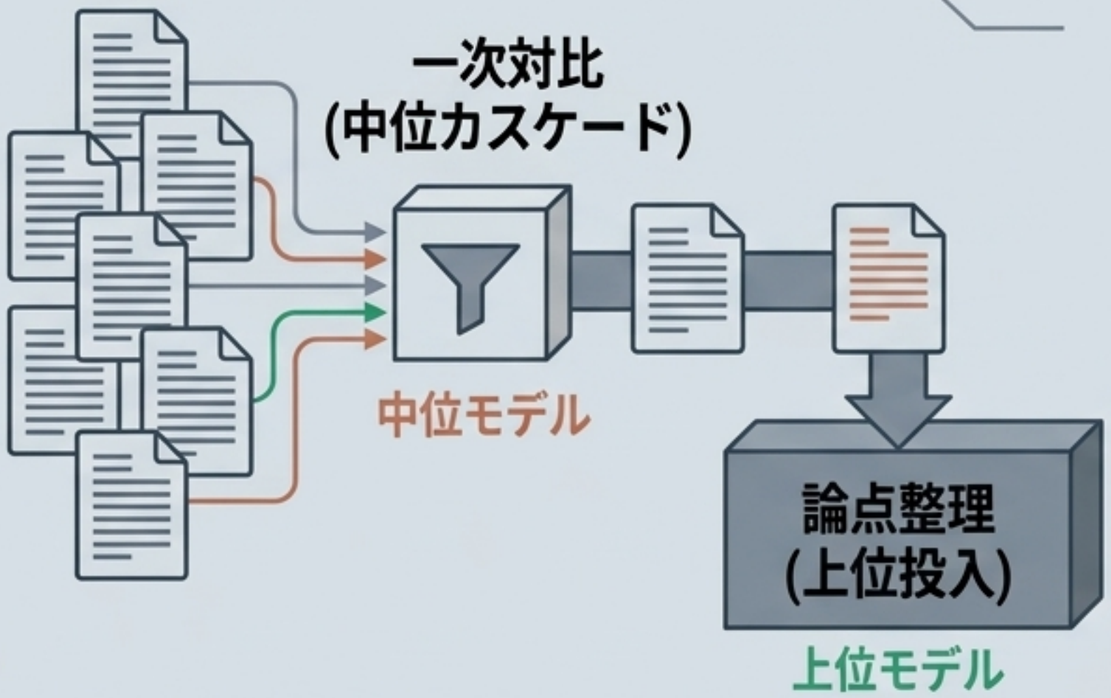
The IP Task Master Matrix: 業務別推奨モデル一覧

知財タスク (IP Task)	第一候補 (Primary Model)	検証・補完 (Validation)	理由・注意 (Rationale)
公報要約・一次スクリーニング	● Sol / Neon (中位)	抜き取りで上位モデル	量が多く、人が検証しやすい。Solは省略に注意
発明の技術的理解（化学・物理）	● Astra / Fable 5.2	発明者による確認	科学系評価でAstraが首位
バイオ・医薬明細書の読解	● Opus 5.5 / Gemini	専門家による確認	Fable 5.2のブロックに注意
無効・理由・進歩性ロジック	● Fable 5.2 / Astra	● Carbon / Neonで引用確認	誤りのコストが大きく、検証が困難
判例・条文の調査と引用	● Neon / Carbon	必須：原文との照合	誤った断定が致命的。日本法は実機確認
FTO・クレームチャート	一次対比:中位 / 論点整理:上位	別系統で再検討	見落としと誤認の双方が高リスク

Gemini Notebook

実務ディープダイブ 1: 先行技術調査とFTO

ワークフローの最適化



見落としと誤認の双方が高リスクとなる領域。一次対比は中位モデルでカスケード処理し、論点整理にのみ上位モデルを投入。

【警戒事項】GPT-6.2系の「省略の傾向」

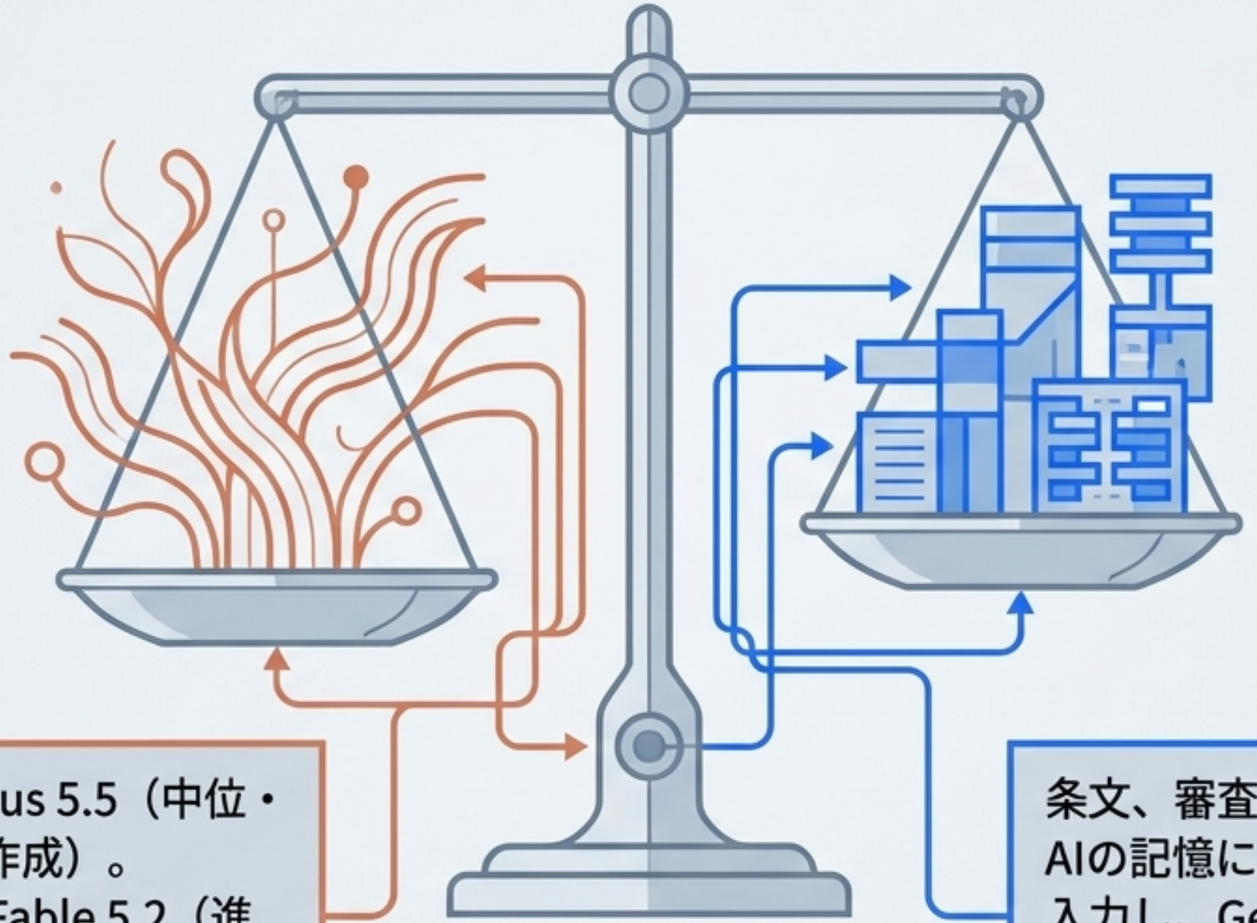
最新のGPT-6.2 Sol/Astraは、前世代の2倍のツール呼び出しを行う一方、回答が短くなり「重要事実を落とす」または「ファイルの欠落等の障害に言及しない」事例が報告されている。

対策：GPT系を使用する際は、「検討した論点と除外した理由の列挙」「入手できなかった資料の明示」をプロンプト　で強制する。

Gemini Notebook

実務ディープダイブ 2: 拒絶理由対応（意見書・補正案）

起案から論理構築までの一気通貫



事実の厳格な固定

ベースの起案: Claude Opus 5.5（中位・低コストで全体の構成を作成）。
難所の論理強化: Claude Fable 5.2（進歩性の反論など、技術的に高度な部分のみを上位に引き上げ）。

条文、審査基準、公報の引用テキストはAIの記憶に頼らず、必ず原文テキストを入力し、Gemini（Carbon/Neon）で一言一句の相違がないか照合・ロックする。

Gemini Notebook

運用上の制約とセキュリティガイドライン

未公開発明の入力リスク（データ保持）

Claude Fable 5.2には、ゼロデータ保持 (ZDR) 適格企業を除き、安全監視のための「30日間のデータ保持」が適用される。

OpenAI・Googleを含め、入力データの学習利用・保持条件の事前確認が必須。



保護措置による業務ブロック

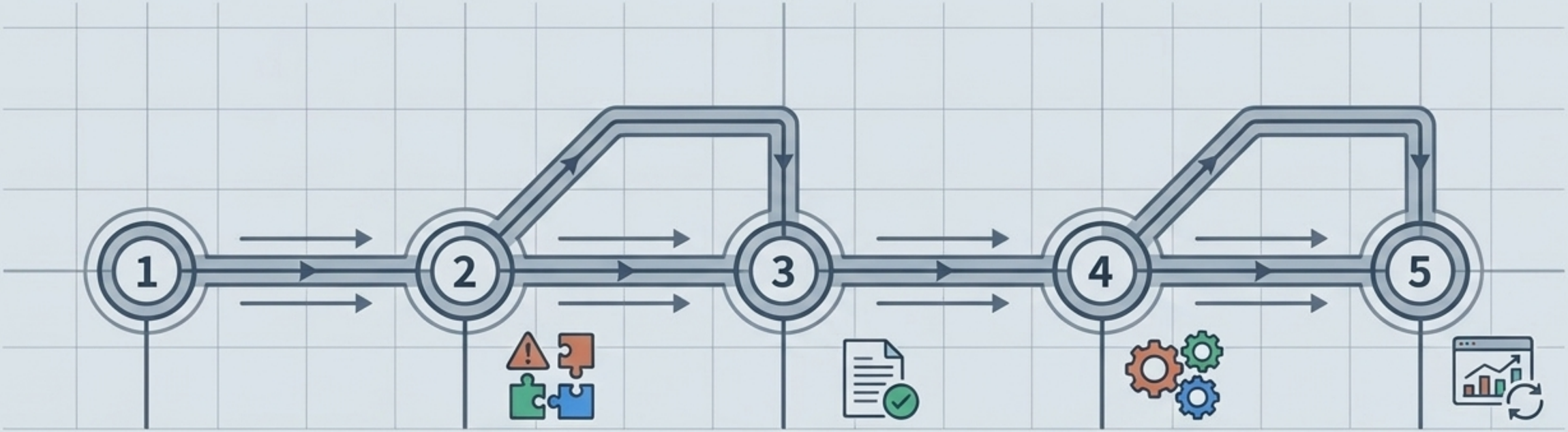
Fable 5.2の生物分野ブロックに加え、GPT-6.2にも生物・化学分野に対する新たな保護措置と制限が追加されている。

API価格の変動リスク

Gemini 4 Carbonの単価優遇は短期間で終了（12/1に2倍へ変動）。継続的なコスト監視が必要。

Gemini Notebook

The Starting Blueprint: 明日から実行すべき5つのステップ



Step 1. **既定モデルを「中位」に設定**：日常業務の出発点をOpus 5.5またはNeonに固定し、コストを最適化する。

Step 2. **「上位へのエスカレーション条件」の定義**：「誤りのりのコスト大」「検証困難」「中位で行き詰まった」の3条件でのみ上位モデルを許可。

Step 3. **「原文照合」のプロセス義務化**：事実照会（判例・条文等）には必ず原文を添付し、可能であればGemini系でクロスチェック。

Step 4. **異系統モデルによる「相互検証」導入**：VIP案件・高リスク案件では、出力と検証で意図的に別系統のAIを組み合わせる。

Step 5. **自社定型タスクでの「パイロット運用」**：自社の過去案件件10〜20件を用い、モデル更新のたびに適性を実測・再評価する。

常に最新の公開ベンチマークと実務データを注視し、ルーティングを更新

Gemini Notebook