

# DeepSeek V4 実務適用戦略ダッシュボード

知財・法務領域における「圧倒的コスト優位性」と  
「ガバナンス・リスク」の完全解剖

THE STRATEGIC BLUEPRINT FOR ENTERPRISE IP & LEGAL



**V4**  
Preview



**MIT**  
License



**1M**  
Context Length

# 結論：知財実務での導入検討に値するが、単独での信用は不可

## Key Takeaway

前工程（検索初動、構造化、候補抽出）においては、他社旗艦モデルを凌駕するコストパフォーマンスを発揮。

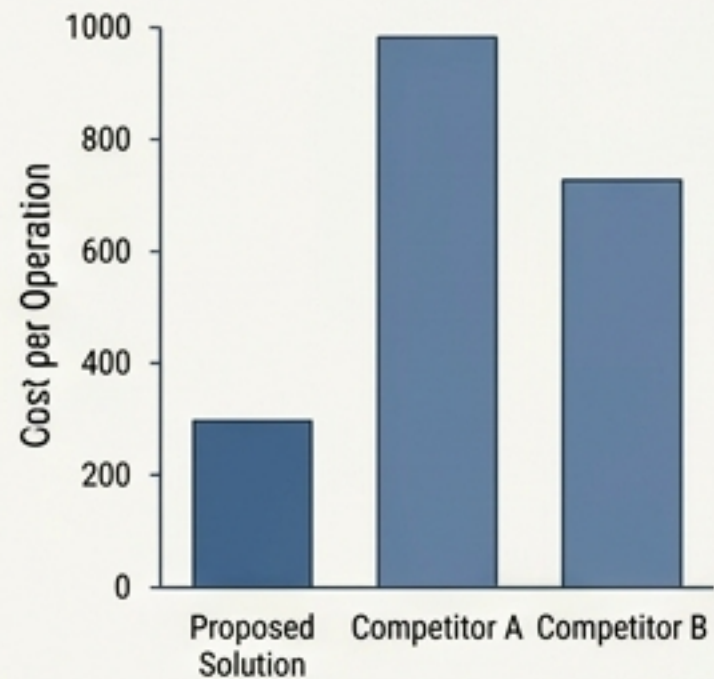
Cost-Performance Index:  
**8.5/10**

Initial Search Efficiency:  
**+40% vs Benchmark**

Structure Automation Rate:  
**75%**

## Opportunity Assessment

前工程の検索者、講義語に対してモデルを構造し、候補抽出においては、他社旗艦モデルを凌駕するコストパフォーマンスを発揮する。



Reject / High Risk



Hybrid Adoption

# 機密性・法的影響が高い工程における絶対条件

## Key Takeaway

セルフホスト化、RAG接続、出典強制、別モデルまたは人間による二重レビューが運用上の前提となる。

Confidentiality Risk Level:  
**HIGH**

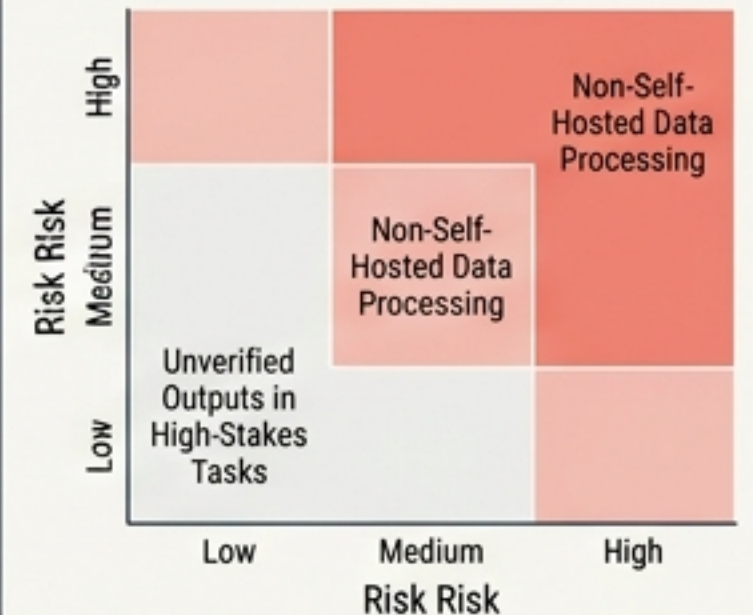
Legal Impact  
**CRITICAL REQ.**

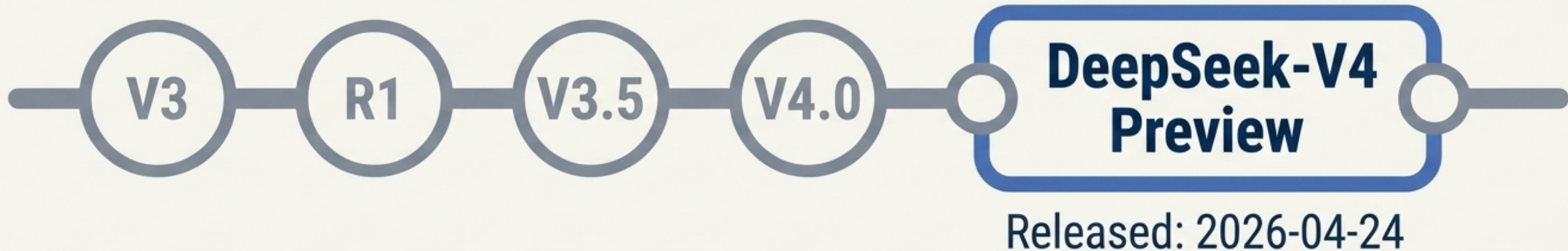
Self-Hosting Mandatory:  
**YES**

## Absolute Indicator

### Absolute Requirements:

- セルフホスト化
- RAG接続
- 出典強制
- 二重レビュー (Dual Review: Human/Alt-Model)





## APIエンドポイント廃止警告

旧APIエンドポイント（deepseek-chat / deepseek-reasoner）は2026年7月24日に完全廃止予定。新規PoCは必ずV4系の明示名で構築すること。

## Previewステータスの運用解釈

現在は「Preview」表記。ベンダー側で仕様や価格の変動余地がある状態として扱うべきであり、固定化された本環境への全面依存は控える。

## DeepSeek-V4-Pro (Think Mode Capable)

総パラメータ	1.6T
推論時アクティブ	149B
事前学習	10T トークン
コンテキスト長	1M Tokens

強み：高度な比較表作成・長文推論

## DeepSeek-V4-Flash (Fast Extraction)

総パラメータ	284B
推論時アクティブ	27B
事前学習	10T トークン
コンテキスト長	1M Tokens

強み：大量文書の構造化・正規化

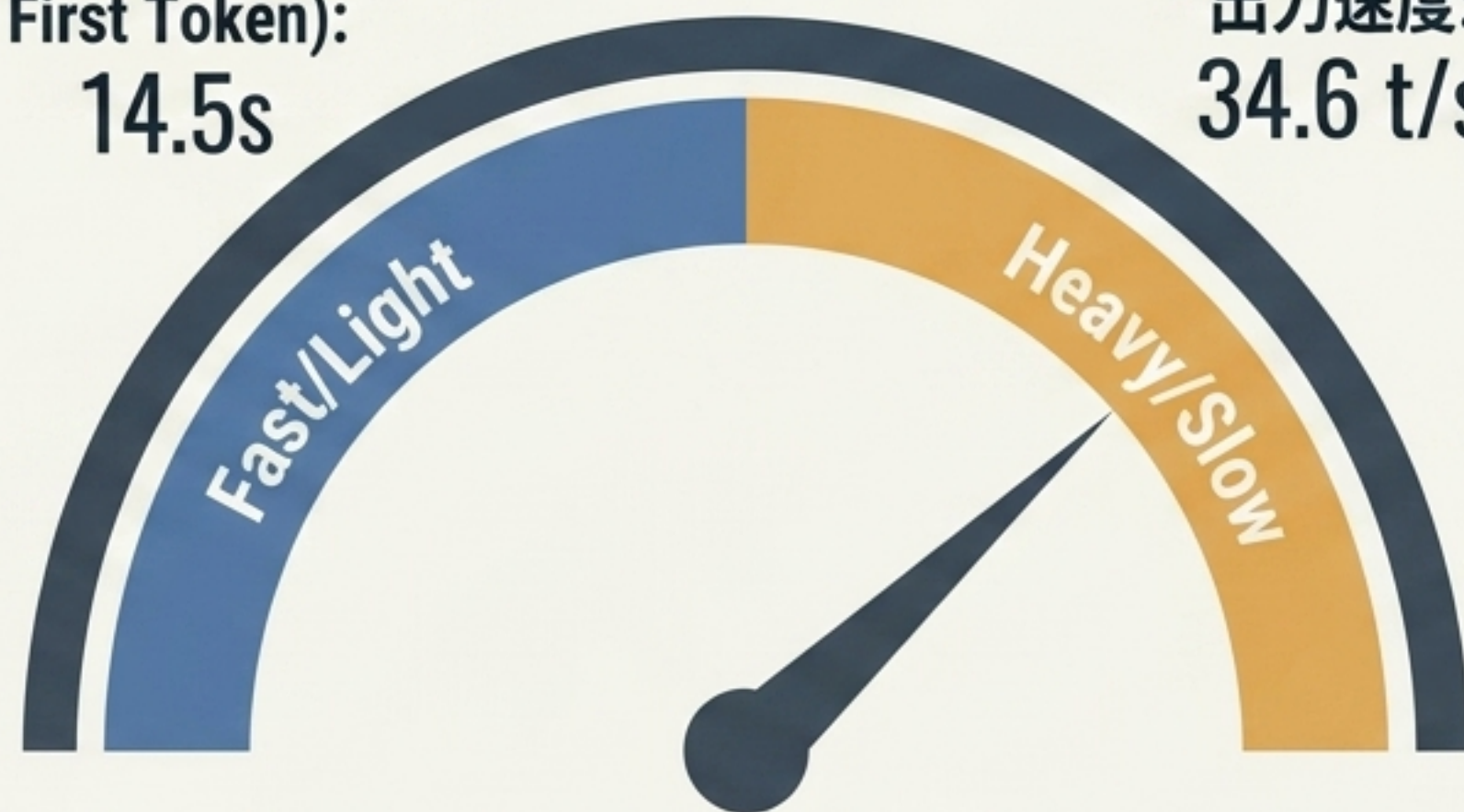
共通技術基盤: Text-only modality | OpenAI/Anthropic SDK互換 | MIT License (商用可)

Time to First Answer Token:

128.46s

TTFT (Time to  
First Token):  
14.5s

出力速度:  
34.6 t/s



思考モード (Think Mode)  
トレードオフメーター

### 向いている業務 (Batch Processing)

- バッチ型レビュー
- 長文の一括読解
- 複雑なクレームチャート下書き

### 向かない業務 (Real-time Interaction)

- クライアント対話
- 即応型チャットアシスタント
- 単発の短いQ&A

(API基盤・プロバイダ変更で短縮可能だが公式は長考型)

## 自社報告：コード・長文・検索に強み



## 独立評価：法務・知財領域の実力



59位 / 118モデル中 (Vals公開スコア)

評価の乖離：法務・知財領域においては「法務特化で飛び抜けている」わけではなく、NIST評価でも「フロンティア最先端より約8か月遅れ」。最終判断を完全に任せる水準にはない。

## TCO Iceberg

API入力コスト

V4-Pro Input Cost: \$0.14/M

(GPT-4oの1/10以下、Claude Opusの1/30レベル)

冗長な出力によるトークン消費増 (112M output tokens生成傾向)

セルフホスト環境構築・維持費

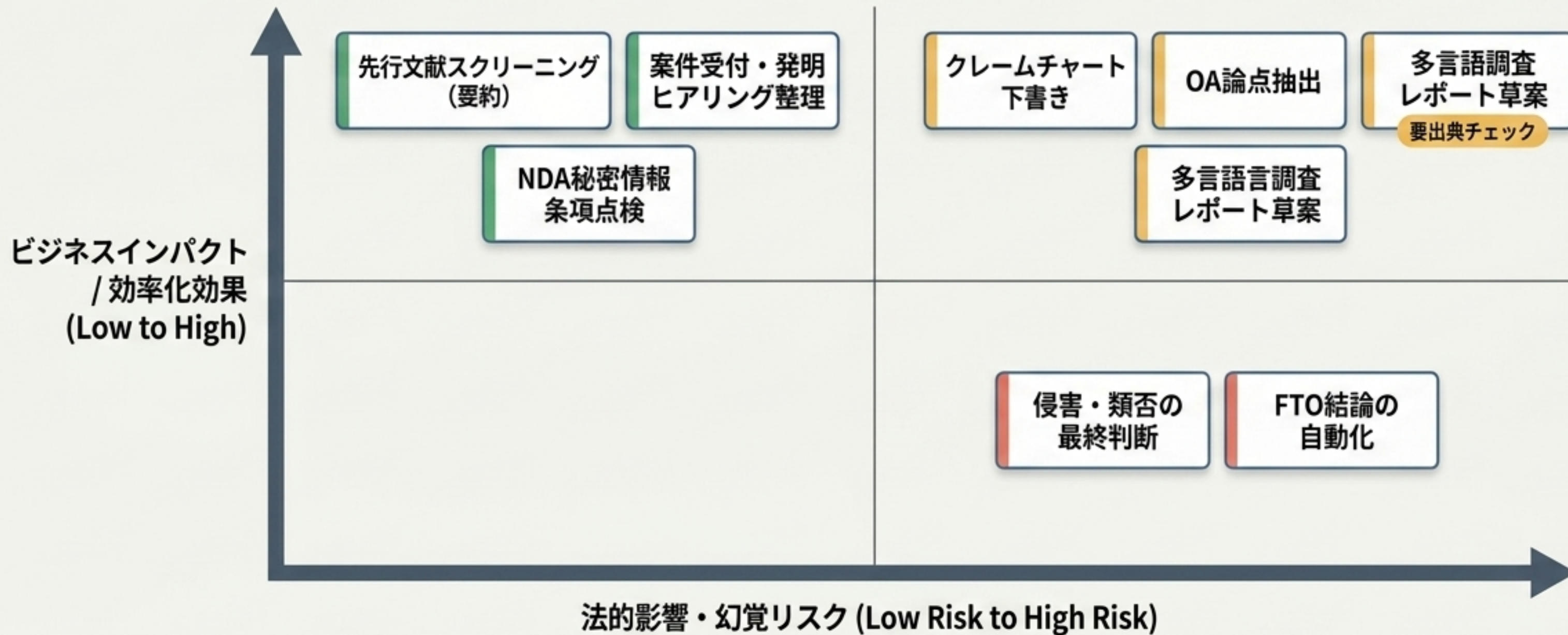
人間・別モデルによる「二重レビュー」の追加プロセス費用

単なるAPI単価比較は危険。出力の要約長制御（ポリシー設定）と、浮いたコストの「検証・ガバナンス」への再投資が不可欠。

# フロンティア AI 対競者判断鋭表

	DeepSeek V4-Pro	OpenAI GPT-4o	Anthropic Claude 3.5	Google Gemini 1.5 Pro
モダリティ	テキスト中心	マルチモーダル / ツール豊富		
強み	圧倒的低価格・セルフホスト可	複雑調査・ツール連動	エンタープライズ監査性・長文読解	画像・証拠統合
弱み・懸念	中国保管・SaaS準拠法リスク	セルフホスト不可	高価格帯	Google Grounding時のデータ保存
知財推奨ポジション	低～中リスクの下書き・整理・候補抽出	複雑な調査・最終クロスチェック	高信頼の契約レビュー・証拠読解	商標画像・多媒体証拠の統合評価

# Risk/Value Positioning Map



**結論の自動化には使用不可。コスト優位性を活かし、「候補抽出」「比較観点整理」「下書き」に特化させること。**

プライバシー：個人データは**中国で処理・保管**。API専用の詳細TTLは現行ドキュメントで明示確認できず。

データプライバシー

地政学：準拠法は**中国本土法**（杭州法院管轄）。海外での規制動向・他社からの**蒸留疑義**リスク。

地政学・  
規制

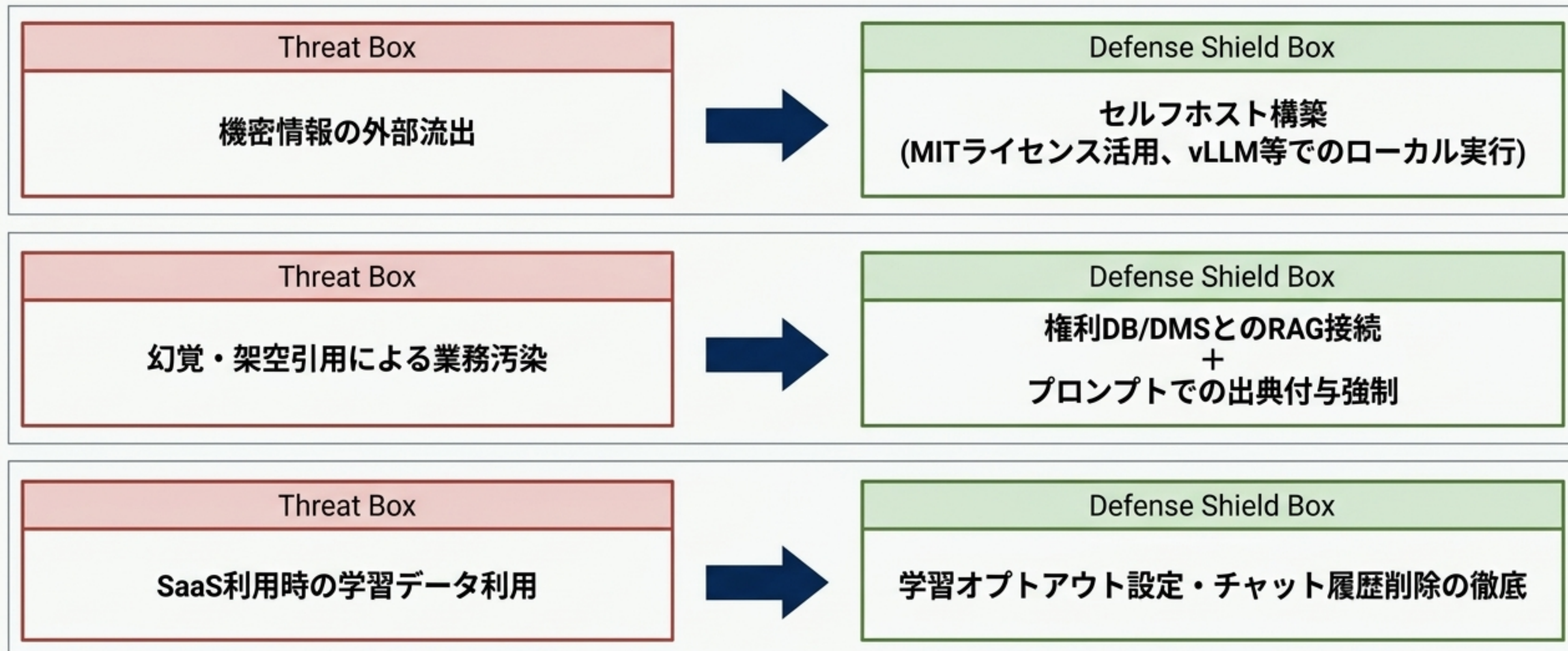
著作権・  
知財権

知財権：出力の**独自性**は保証されず、他ユーザーへ**類似出力**が生成される可能性あり。

幻覚（Hallucination）

幻覚：架空の判例、存在しない**文献引用**の危険性（法的推論での**事実幻覚率**に課題）。

# Defense Matrix: Threats and Mitigation Solutions



鉄則：「削除できること」より「最初から外に出さないこと」。特許出願前情報や侵害判断メモの公式Web/Appへの投入は厳禁。

# 「適材適所のAIパイプライン設計」

DeepSeek単独の検索に期待せず、権威ソース（J-PlatPat, DMS等）を先に整備し、その読解にモデルを使う構成が必須。

## 副査・下書き担当



DeepSeek V4 (Flash / Pro)

大量文書の抽出、分類、候補生成、第一ドラフトの高速作成。圧倒的コスト優位性を活かした「力技の整理」。



## 主査・最終判断担当



人間（弁理士/法務担当） +  
権威モデル（Claude 3.5 Opus等）

事実確認、法的推論の検証、最終的なリスク判断と出力承認。

# 複数のAIモデルと人間が直列に連なる「ハイブリッド知財ワークフロー」の完成形

1.  
案件受付 /  
文書収集  
と機微情  
報分類  
(権利DB /  
判例DB /  
契約台帳 /  
DMS検索)

2.  
DeepSeek  
Flash  
で抽出・  
正規化・  
候補生成

3.  
DeepSeek  
Pro  
High/Max  
で比較  
分析・  
ドラフト

4.  
出典  
リンク  
付与と  
妥当性  
チェック

5.  
弁理士・  
弁護士・  
担当者  
レビュー

6.  
採否ログ  
保存 /  
評価デー  
タ更新と  
再学習用  
フィード  
バック

# PoC Roadmap (導入ロードマップ)

## 1. Targeting

「会話満足度」ではなく「事故なき再現性」で検証。検索・抽出系から着手。

## 2. Data Prep

50~100件の「採否判断済み過去成果物（ゴールドセット）」を用意。機密データはセルフホスト環境でのみテスト。

## 3. Evaluation

自動完了率ではなく「保守的な要確認率」を測定。

# Metrics Dashboard (検証KPI)

## 正確性

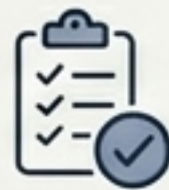


Citation Precision  
/ Recall@k



Precision: 92% / Recall@k: 88%

## 適合性

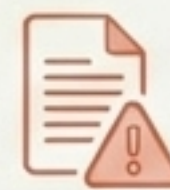


Issue Spotting  
Recall

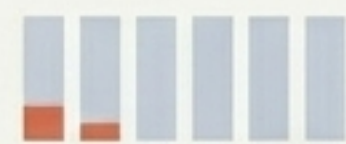


94%  
Recall

## 安全性



Unsupported  
Assertion Rate  
(架空引用率)



5.2%  
Rate

## 運用性



1件あたり処理時間・  
トークンコスト



Avg Time: 1.2s / Avg Cost: \$0.05

# 最終推奨方針 (Executive Directive: June 2026)

DeepSeek V4は、知財部門における「第一候補の本番判断モデル」ではなく、「超高コスパの実務補助・下書きモデル」として導入せよ。

## Actionable Next Steps



FTO結論や侵害認定の自動化には絶対に使用しない。



情報整理・翻訳・抽出タスクに限定し、大幅なコスト削減を即時実現する。



削減したコストを、セルフホスト基盤の構築と二重レビュー体制の強化に再投資する。



Architecture  
Approved